How to deal with fluctuations in hospital processes to improve accessibility?
Joustra, P.E.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
“It is not the strongest of the species that survives, nor the most intelligent, but the one most responsive to change.”
Charles Darwin
CHAPTER 9

General Discussion
This chapter evaluates the research presented in Chapters 2 through 8, which are based upon articles. All chapters are used to elaborate on the overall research objective of this thesis: to show the added value of quantitative methods such as queuing theory, discrete event simulation, and regression analysis for generally applied process improvement methods in a complex hospital setting, in particular for decision support at strategic level.

In addition, the research in Chapters 2 through 8 was used to answer the following specific research questions:

1. How can waiting times, access times, and throughput times be reduced cost-effectively
   a. by decreasing the various fluctuations in health care processes and/or
   b. by pooling or separating various patient groups?
2. How can hospital processes be made more predictable so that management is able to anticipate future developments proactively, and structurally improve the hospital’s key performance indicators?

Multiple chapters are used to answer an individual research question (see Table 1.6). Some of the chapters are used to answer two specific research questions.

The outline of this chapter is as follows: The main findings are presented per research question. We start by describing the added value of quantitative methods to support decision making in a complex hospital setting. Subsequently, we answer each specific research question. After reporting our main findings, we clarify the extent to which the applied methodology and our results are also applicable in other hospitals. Next, we continue with the limitations of our research, followed by our recommendations for further research. We conclude this chapter by listing the practical implications of this thesis.

9.1 Main Findings

9.1.1 The added value of quantitative methods

Chapter 2 clearly shows the added value of quantitative methods such as queuing theory and discrete event simulation. This chapter describes a case study of the Academic Medical Center (AMC) radiotherapy department with the goal of reducing the throughput times for a multi-step process: (1) the outpatient department (OPD), (2) the preparation phase, and (3) the actual treatment using three linear accelerators. Before we started this study, management had already decided to invest in an additional linear accelerator to remove the presumed bottleneck. However, our simulation model indicated that while a fourth linear accelerator would indeed decrease throughput times, the decrease would not be enough to meet the
throughput time targets. These would only be met when the access times for the OPD were reduced as well. An alternative solution might have been to increase the capacity of the OPD and to reduce the preparation time for elective patients (10 days on average), which was relatively high compared to the throughput time target (21 days). Because the fourth linear accelerator had already been purchased, we did not investigate how large this reduction should have been and whether this was feasible in practice. Nevertheless, this study clearly demonstrates the added value of quantitative decision support.

The endoscopy case study described in Chapter 4 demonstrates that the existing subdivision of capacity among various patient groups was not efficient. By reducing the dedicated capacity for urgent patients by almost 50% and redistributing the released urgent capacity among other patient groups, all groups had sufficient capacity to meet their access time targets. An iterative combination of simulation and integer linear programming was used for quantitative decision support. The simulation model was used to determine the minimum urgent capacity keeping the percentage of double-bookings due to urgent patients under 3%. In addition, integer linear programming was used to redistribute the released urgent capacity among other patient groups, taking into account the limited availability and usability of both procedure rooms and physicians.

Chapter 5 also describes a case study for reducing the access times of the magnetic resonance imaging scanner (MRI). Although the solution of pooling various patient groups looked very promising in terms of logistics, it required more flexibility on the part of the radiologists. In the pooled situation, multiple subspecialties would use the same multipurpose capacity, and thus patients of one subspecialty could be scheduled within a much longer time period. Therefore, the specialized radiologists had to be available over a longer period, also to check the quality of the MRI scans. Hence, we developed a simulation model to determine the order of magnitude of the improvement to convince the radiologists to implement this solution. The reduction in access times predicted with the simulation model was large enough to convince them, and the solution was implemented in practice. In contrast to the endoscopy study, we did not need an integer linear programming model to reschedule the multipurpose MRI capacity. There are two reasons for this, namely, that none of the three subspecialties involved required a specific MRI scanner, and the presence of a specialized radiologist was not a limiting factor. Therefore, rescheduling the multipurpose capacity was straightforward and we did not need linear programming.

In the case studies described above, we could have tested the potential solutions during a pilot period. However, to test a single scenario to reduce access times and throughput times, the pilot period would have had to last a number of months. It might have even taken several years to quantitatively compare or possibly combine several alternative solutions. If a potential solution would then have turned out
to have had an adverse effect on access times or throughput times, the interests of patients would have been harmed and the motivation of personnel diminished. Furthermore, the experimental conditions would probably not have been the same during the entire pilot period, as confounding factors make it hard to isolate the effect of the intervention.

In the case study described in Chapter 7, developing and implementing the software to support more accurate scheduling of operations would be costly. Therefore, management wanted to know beforehand how large the positive impact would be so they could quantitatively support a business case. Using a regression model would be an objective way of predicting the duration.

Finally, Chapter 8 describes a study concerning a future situation, which is often the case with decision making at strategic level, because in practice trial and error is just not an option. Furthermore, this study shows that it is necessary to take fluctuations at operational and tactical levels into account for strategic-level decision making.

For these reasons, generally applied process improvement methods (such as Business Process Redesign, Total Quality Management, Theory of Constraints, Lean Management, and Six Sigma) need queuing theory and discrete event simulation to quantify the expected impact of a potential solution on waiting times, access times, or throughput times before implementing the solution in practice. Chapter 8 clearly demonstrates that this quantitative decision support is not only beneficial at operational and tactical levels, but also has an effect on strategic decision making as well.

Our conclusion that Lean Management could benefit from quantitative methods such as queuing theory and discrete event simulation might not be surprising. In contrast, our claim that this statement is also valid for Six Sigma may come as more of a surprise. Although Six Sigma is known as a data-driven, quantitative process improvement method, these quantitative methods are used mainly in the “Measure” and “Analyze” phases. Despite the fact that these methods can also be used in the “Improve” phase to analyze the expected impact of a potential solution on waiting, access, or throughput times, a pilot period is often required. For the reasons mentioned above, testing a potential solution during a pilot period is often not desirable or even feasible in a hospital setting.

9.1.2 How can the various fluctuations in health care processes be decreased?

To answer this specific research question, we investigated the effects of various types of variability on the accessibility of hospital care:
1. Demand variability,
2. Variability in process times, and
3. Capacity variability.
CHAPTER 9

Three studies in this thesis - Chapters 2, 4, and 6 - demonstrate the effect of variability reduction. First, Chapter 2 shows how a decrease in the fluctuations of capacity in an OPD leads to a reduction in both access times as well as waiting times prior to the consecutive steps of the radiotherapy process: the preparation phase and the actual treatment. Note that the reduction in waiting times prior to the preparation phase and the actual treatment results from lower demand variability prior to these steps. Because the average OPD capacity remains the same, this is a cost-effective solution for reducing throughput times.

Second, one of the solutions for reducing access times in the AMC endoscopy department (Chapter 4) was to introduce a backup system for physicians to decrease the number of occasional closures. Although physicians have to take on more shifts and therefore work more hours, the backup system will still save capacity because the weekly capacity will be more stable, which in itself will reduce access times.

Third, to reduce waiting times in the waiting room of the MRI department, we recommended reducing the supervision by radiologists through increased standardization of the MRI protocols. Since the variability of durations with supervision exceeded the variability of durations without supervision, this intervention would lead to an overall reduction in the variability of the process times. Subsequently, this reduction would lead to lower waiting times in the waiting room without investing in additional resources.

9.1.3 When should various patient groups be pooled or separated?

To explore the situations in which patient groups should by pooled or separated to improve accessibility, we investigated two different reasons for this:

1. Pooling or separating based on urgency level, and
2. Pooling or separating capacity dedicated to specific patient groups that seem to be similar from a logistical point of view.

In Chapter 3, the minimum required capacity for meeting the performance target was determined in the current situation (in which all patients use the same capacity) and in an alternative situation (in which urgent and elective patients both have dedicated capacity). This analysis clearly shows that the difference in urgency level determines whether pooling or separating urgent and elective radiotherapy patients will be beneficial. In the current situation, although the minimum required capacity is equal in both situations, the performance level in the separated case is higher. In addition, if elective patients are allowed to utilize free urgent timeslots (one day in advance), the capacity required to meet the access time targets would drop even further.

In Chapter 4, we also conclude that separating patients based upon urgency level improves efficiency. In contrast to the radiotherapy study, we do not separate urgent
and elective patients, but subdivide the urgent patients into semi-urgent and urgent patients. The endoscopy study demonstrates that separating the urgent general gastroenterological and colonoscopy patients into semi-urgent and urgent would use the total capacity more efficiently. Because the semi-urgent capacity can be used more efficiently than the urgent capacity, a smaller number of urgent patients will lead to a higher overall utilization. This effect offsets the negative effect of separating the capacity into semi-urgent and urgent timeslots.

In addition, the simulation model of the AMC endoscopy department shows that pooling various patient groups with the same urgency level (in this case, the elective general gastroenterological and colonoscopy procedures) can also lead to improved utilization. Unfortunately, because the semi-urgent procedures of other types of procedures required different equipment and/or specialized physicians, pooling was not an option. However, a larger reduction in required capacity was to be expected when pooling semi-urgent procedures than when pooling elective procedures.

Moreover, Chapter 5 illustrates a similar advantage of pooling various patient groups with the same urgency level. In this case, we pooled elective patients from three patient groups and substantially reduced the access times for MRI scans. Unfortunately, the cardiac MRI scans and MRI scans under sedation could not be pooled, because personnel outside the radiology department needed to be present (a cardiologist and an anesthesia team respectively). Nevertheless, because there was a large reduction in access times, the number of semi-urgent timeslots was reduced significantly, and only one specific capacity claim remained (concerning brain surgery). The other capacity claims of various clinical paths became redundant with an access time of less than two weeks. Obviously, this increased the flexibility of the MRI capacity even more.

9.1.4 How can hospital processes be made more predictable?

Chapter 6, 7, and 8 are used to answer this research question. In Chapter 6, the goal is to reduce the waiting times in the waiting room of the MRI department. One of the solutions for achieving this goal was to better estimate the scan duration. In contrast to Chapter 7, we did not use an econometric model to estimate the scan duration more accurately. Because the variance of the scan durations per scan type was limited, we knew that the scan type was an important explanatory variable for the average scan duration. Therefore, we compared only the scheduled duration and the actual duration per scan type. This comparison showed that for some types of scans the scheduled duration was overestimated, while for other types it was underestimated. By estimating the scan duration more accurately, waiting times could be reduced without decreasing the utilization rate of the MRI.

In Chapter 6, we showed only the effect of more accurate estimates of the duration on the relevant key performance indicators (KPIs). In Chapter 7, we actually explain
how to estimate the duration more accurately (in this case, of a surgical operation) by applying various econometric models. In Chapter 7, the goal was to reduce the risk of overtime in the operating room and to reduce the percentage of cancellations due to overrun of previous surgeries. In contrast to the MRI department, waiting times before an operation are not considered relevant in the AMC. We concluded that applying econometric models can significantly improve the estimates of durations. Applying an objective method for estimating the duration can prevent this. This does not mean the surgeon’s estimate is of no value, because it turned out to be an important explanatory variable for the econometric models. More specifically, the surgeon’s estimate was significant for all departments and all econometric models ($p < 0.01$).

Chapter 8 demonstrates how quantitative decision support at strategic level can contribute to improving the performance of the complete chain of the OPD, nursing ward, and operating room in the long term. If efficiency improvements are not sufficient to meet the performance targets and if it is not feasible or preferable to increase capacity, the only solution is to reduce demand. We developed an interactive decision-support tool to enable the management of the AMC ophthalmology department to decide upon the future patient-mix to improve the KPIs, especially the long waiting list for the operating room. With this interactive tool, management is able to alter the number of patients per patient group, and see the consequences in terms of maximum workload so that the preferred KPIs can be met. For the AMC ophthalmology department we focused on the main bottleneck, namely the operating room. We assessed that it is not enough to use last year’s utilization rate to calculate the future maximum workload for a different patient-mix and higher performance requirements. Therefore, we used a combination of regression analysis and simulation to determine the effect of patient-mix characteristics and preferred KPIs on the maximum workload. With this interactive tool, management is able to better anticipate the future situation, either to alter the expected patient-mix or to expand capacity to ensure that the KPIs will be met structurally in the future. Clearly, capacity issues cannot be solved without taking the preferred service levels into account as well.

This research also shows that assessing operating room usage based exclusively on the actual utilization rate is unfair (the patient-mix is not taken into account) and shortsighted (the other performance measures are not assessed). Furthermore, focusing on maximizing the utilization rate might even lead to more overtime and an increased number of cancellations due to overrun of previous surgery and prioritizing semi-urgent and urgent patients.
9.2 Generalizability of methodology and results

9.2.1 Generalizability of methodology
The queuing models in Chapters 2 and 3 were used to provide insights into the effect of variability on access times of the AMC radiotherapy OPD. These queuing models could also be used to provide insights for other OPDs within the AMC and in other hospitals.

Computer simulation was used in Chapters 2 through 6 to include the dynamic interaction between the different steps of a multi-step hospital process, to model the complex scheduling rules, to incorporate all relevant types of variability, and to supply specific performance indicators. Subsequently, these simulation models were applied to quantitatively evaluate and compare the alternative solutions for cost-effectively improving the accessibility of hospital care. Although the specific simulation models must be adapted to analyze similar problems in other hospital departments, the need for computer simulation remains. Note that the aim to keep the simulation model's scope and level of detail as limited as possible is in direct conflict with the generic application of the model.

In Chapter 4, we used a combination of computer simulation and integer linear programming (ILP) to reduce the access times for endoscopic procedures. The ILP model was necessary to reschedule the required number of timeslots for all procedure types in an available procedure room with the required equipment and a specialized physician present. It is relatively easy to adapt this ILP model for other endoscopy departments. It might even be useful for scheduling operating rooms, because these two scheduling problems are comparable. Note that we did not use the ILP model in Chapter 5 to reschedule the required timeslots for the various types of MRI scans: This scheduling problem was straightforward, and therefore could be solved without an ILP model.

As we already stated in the previous section, the various regression models used in Chapter 7 could be used to improve the accuracy of scheduling operations in other hospitals as well. These regression models could also be used to improve the accuracy of scheduled consultations in an OPD, various (endoscopic) procedures and (MRI) scan durations to increase the utilization rate given a maximum risk of overtime, a maximum percentage of cancellations due to overrun of previous operations/consultations/procedures/scans, and a waiting time target.

The combined approach of regression analysis and computer simulation we used in Chapter 8 is suitable for other hospitals as well. In our case study, we used only computer simulation to determine the effect of the waiting time target for elective patients on the maximum workload. In other case studies, computer simulation may have to be used for other factors included. This might be necessary if the historical data of an included factor cannot be retrieved, or if a specific factor was nearly
constant over the past several years. Moreover, if the preferred level of a KPI differs considerably from actual performance over the past several years, using regression analysis would require extreme extrapolation. Therefore, it might be better to model this KPI explicitly with computer simulation.

9.2.2 Generalizability of results

Other hospitals could also benefit from applying quantitative methods to support operational, tactical, and strategic decision making.

The positive effect of reduced demand variability and variability in process times on a hospital’s performance has been shown before (for example, [1] and [2] respectively), so they are clearly applicable to other hospitals. Capacity variability reduction could also be applied in other hospitals to reduce the access times of many OPDs. With respect to the applicability within central diagnostic departments, it depends on whether the available capacity is dependent on physicians. For example, the MRI capacity depends mainly on the availability of the MRI scanners and the presence of radiology assistants rather than on radiologists. Therefore, the capacity is not dependent on physicians, and the available capacity will be more stable. In contrast, an endoscopic procedure requires the presence of a specialized physician and the available capacity therefore depends mainly on these physicians, which often leads to higher fluctuations in capacity. Furthermore, less fluctuation in the capacity of a single step of a multi-step hospital process will probably also reduce the waiting times prior to the subsequent steps.

Although pooling different patient groups is often claimed to improve the efficiency of shared resources, this cannot be guaranteed. If the average duration per patient group differs substantially and if the changeover times per patient type are significant, efficiency might even decrease if the different groups are pooled. In addition, we cannot draw a general conclusion about when to pool or separate patient groups with different urgency levels. The radiotherapy case study presented here does show that this question of a trade-off depends mainly on the threshold value of the access time target of urgent patients.

Multiple studies (for example, [3] and [4]) have demonstrated that applying the lognormal regression model can improve the accuracy of scheduling operations. Clearly, the alternative regression models described in Chapter 7 could be applied in other hospitals as well.

Some modifications would have to be made to apply the interactive decision-support model described in Chapter 8 to other departments within the AMC or to other hospitals. First, to apply the current version of the interactive model, one has to analyze whether the operating room is the department’s main bottleneck. Next, one has to check whether the factors included are the same for the new department. If this is the case, the required data has to be gathered. Finally, one
has to decide whether to use regression analysis or computer simulation to quantify the impact of an individual factor on the maximum workload of the operating room. Now the interactive model is ready to be used by management to decide upon the future patient-mix, taking the preferred operating room performance into account. Obviously, the main message holds for other hospitals as well: Be proactive and anticipate the future situation, either by altering the expected patient-mix or by expanding capacity to ensure that the KPIs will be met structurally in the future.

9.3 Limitations

The first limitation of this thesis is that all studies concern unit logistics rather than chain or network logistics [5]. An obvious disadvantage of process improvements in a single unit is that the overall hospital performance might not improve. This might be especially true for the studies described in Chapter 4 and 5 about reducing access times in a diagnostic department. Although we succeeded in reducing access times for endoscopic procedures and MRI scans, we did not check whether the total throughput times of the corresponding patients actually decreased. We assumed that the high access times for endoscopic procedures and MRI scans were the main bottlenecks for the total throughput time of the patients, but maybe the follow-up consultations or other diagnostic tests could not be scheduled earlier and these were the real bottlenecks.

The second limitation of our research is that although we did succeed in meeting the access time targets of two weeks with the current capacity, the preferred access times seem to be even lower with the emergence of “one-stop shopping”. Moreover, a recent study concluded that total hospital costs are minimized when waiting times are less than ten days [6]. Capacity might have to be expanded to reduce access times even further, something that would require investments in additional personnel and/or equipment.

The third limitation is that if patients with different urgency levels are to be separated, the definition of an urgent patient should be unambiguous. Otherwise, a referring physician might use the urgent indication more often than actually necessary. This would increase the number of urgent patients and, consequently, the access times for urgent patients would increase and the overall efficiency would decrease. Pooling different patient groups also has a disadvantage. For example, one subspecialty within a diagnostic department might expand the indication area, which would increase the number of this subspecialty’s patients. If various subspecialties were to be pooled, this increase would probably lead to higher access times for all pooled subspecialties. This negative effect could be reduced by dedicating most of the capacity to a single subspecialty and making only a small part of the capacity
Another way to limit this risk is to maintain the dedicated capacity per subspecialty, and define conditional overflow rules that state when one subspecialty may use the dedicated capacity of another subspecialty.

### 9.4 Further research

Most articles about quantitative decision support in health care deal with a specific problem in a specific hospital setting. More research on how to implement best practices within other departments in the same hospital - and within the same departments in other hospitals - would be very valuable. This requires user-friendly, easily adaptable decision-support tools that clearly describe the goal, the assumptions, and the limitations.

In addition, it would be worthwhile to pay more attention to tactical-level decision support. In this thesis, we demonstrated that reducing fluctuations in OPD capacity through tactical management improves efficiency. Another application of tactical management is to adapt the operating schedule to better distribute the expected workload for nursing wards. Management could anticipate the remaining fluctuations in demand for nursing care by forming groups of multi-skilled nurses and a group of nurses with flexible working hours. Operations research techniques are required to decide upon the sizes of these groups, and to propose some rules of thumb about when to schedule certain types of nurses.

In this thesis, we investigated whether to pool logistically similar patient groups with different process times to improve efficiency. We also investigated whether to pool patient groups with different urgency levels. More research is necessary to explore the situations in which to either pool or separate patient groups that have both different process times and different urgency levels. It might even be possible to determine rules of thumb for when - and when not - to pool patient groups. An alternative to completely pooling or separating patient groups is to determine conditional overflow rules that set out the situations in which one patient group can use the dedicated capacity of another patient group. Chapter 3 illustrates that this alternative would probably be superior to completely pooling or separating patient groups. Another alternative for pooling entire patient groups is to limit the amount of multipurpose capacity, and to set out rules for the situations in which a patient is allowed to make use of this. Further research is needed to decide upon the amount of multipurpose capacity and to determine the set of rules. Although both alternatives seem promising, operational management will be required to correctly apply these overflow rules in practice. Note that the above suggestions for further research are applicable not only to OPDs and central diagnostic/therapeutic departments, but also to nursing wards.
Applying the interactive tool described in Chapter 8 to other hospital departments would require more detailed modeling of the OPD and nursing ward. Moreover, the financial perspective would need to be integrated to investigate all consequences of a future patient-mix. Finally, an aggregated model for the entire hospital would have to be developed to incorporate interrelated patient groups (for example, diabetes patients) and to determine the future capacity requirements for shared resources such as central diagnostic/therapeutic facilities and the operating room. With this aggregated model, hospital management would be able to be proactive and anticipate a future situation, either to alter the preferred patient-mix or to expand capacity to ensure that the KPIs will be met structurally in the future.

A limitation of our research was that all studies concern unit logistics. Once the accessibility of the OPDs and the central diagnostic facilities is satisfactory, the AMC’s next step will be to focus on network logistics. In contrast to chain logistics, network logistics does not take just one specific patient group into account, but also aims to avoid adverse consequences for other patient groups [5]. Increasing competition in the near future will probably lead to a demand for shorter throughput times for patients. One way to reduce throughput times for a large number of patients would be to combine appointments more often, though this might require complex scheduling rules to avoid a substantial drop in efficiency [7]. Further research is necessary to develop scheduling rules that efficiently combine appointments within one department or for multiple departments. If these scheduling rules could be integrated with a web-based application, patients would be able to schedule their own appointments. Subsequently, patients could decide for themselves if they want to combine multiple appointments on one day or schedule individual appointments on multiple days. An alternative way of reducing the throughput times is to offer a walk-in system for central diagnostic facilities. This will clearly be neither possible nor desirable for all patients, but will substantially reduce the throughput time for all patients using the walk-in facility. Although a walk-in facility is likely to increase fluctuations in demand during a working day and therefore might lead to higher waiting times, it also has some advantages. One advantage is that the no-show percentage will drop, because walk-in patients always show up. Another advantage is that the amount of slack capacity for limiting waiting times is not necessary if patients walk in. In addition, patients might be willing to wait longer if they are able to walk in instead of being forced to return to the hospital on another day. The situations in which these advantages counterbalance the disadvantage of increased fluctuations in demand during a day cannot be determined without quantitative decision support. A research project is ongoing in the AMC to quantitatively and qualitatively investigate the consequences of a partial walk-in system for CT scans.
9.5 Practical implications

9.5.1 The need for quantitative decision support

A hospital is a highly complex environment with numerous interactions between patient groups and hospital processes. Moreover, there are many types of variability that make it difficult to predict the effect of potential solutions for improving a hospital’s accessibility. This thesis has shown that quantitative methods such as queuing theory and computer simulation are very well-suited for quantifying the effect of a potential solution before implementing this solution in practice. Applying quantitative methods ensures that alternative solutions are compared objectively. These methods can be used to adapt historically grown capacity division of shared resources, and divide the capacity among the subspecialties so that the overall performance is optimal. Furthermore, these methods can be used to decide whether to pool patient groups that are different from a medical perspective but logistically similar. In short, queuing theory, computer simulation, and regression analysis support fact-based decision making with no hidden agenda. This reduces the risk of undesirable outcomes for patients or personnel, and also unnecessary investments. This thesis shows that the applicability of quantitative methods is not limited to operational- and tactical-level decision making, but that these methods are also valuable for strategic-level decision making.

Generally applied process improvement methods such as Lean (and) Six Sigma often require a pilot period to quantitatively predict a potential solution’s performance. Using a model instead of a pilot period has several advantages:

1. In a model, because one has control over the experimental conditions, the effect of a single intervention can be quantified.
2. Especially at tactical level, the experimentation period for comparing alternative scenarios will be substantially shorter, thus making it possible to compare more scenarios and probably find a better solution.
3. If the effect of the intervention turns out to be negative, no harm will have been done to the interests of either patients or personnel.
4. If a considerable investment is required, using a model reduces the risk of investing in additional resources that do not result in a substantial increase in the KPIs.

9.5.2 Costs versus quality

We expect that in the near future the goal of process improvements will shift increasingly from improving quality and accessibility of care towards reducing costs per patient. If costs are cut and resources are restricted without careful consideration, the quality and accessibility will probably decline (see Figure 1.2 in the General Introduction). The challenge will be to decrease costs and maintain the current level
of quality and accessibility. Preferably, the cost reductions will go hand-in-hand with further improvements in quality and accessibility. This thesis demonstrates multiple solutions for accomplishing this.

The capacity of shared resources such as those of an endoscopy department and a radiology department are often divided among many subspecialties and/or patient groups. Both the capacity division as well as the division of the patient population into many groups have often grown historically and are hard to change. This results in capacity that is fragmented and not flexible enough to respond to alterations (including temporary ones) in the number of patients per subspecialty and/or patient groups. Quantitative methods can assist in the difficult, often political, process of redistributing capacity fairly among the patient groups by providing objective decision support. Moreover, by pooling logistically similar patient groups, the scarce capacity becomes more flexible, which will result in greater efficiency.

To increase the flexibility of shared resources even more, capacity claims for clinical paths should be avoided as much as possible. We demonstrated that greater flexibility of shared resources will lead to lower access times. If access times are less than two weeks, many capacity claims for clinical paths become superfluous because the preferred timeslot will probably be available when the indication for a scan or examination is set two or three weeks in advance. If capacity claims for multiple clinical paths are still necessary with low access times, these claims should be pooled on one specific day and with the timeslots for semi-urgent or urgent patients to improve efficiency. Unused timeslots should be made available to other patient groups as soon as possible.

To further reduce the costs per patient, physicians will need to play an active role, because they are involved in a major part of the costs. An additional advantage of more active involvement by physicians and other medical personnel is that the results of the process improvement projects will be more sustainable. Moreover, by creating a continuous improvement culture among all medical personnel (for example, with Lean Management), awareness of quality and cost will become more their own responsibility. Quality improvement personnel should facilitate this continuous improvement culture, and support the medical personnel by step-by-step improvements.

An alternative way of reducing the costs per patient and improving the quality of care is to enlarge the economy of scale by increasing specialization among hospitals. First, separating highly complex care from less complex care would reduce the cost for the less complex care because the “focused factory” concept could be applied more often [8]. Second, by concentrating highly complex care more, accessibility could be improved efficiently and/or costs could be reduced. Moreover, existing literature shows that this will also lead to higher medical quality of care for complex cancer surgery [9-11] and high-risk cardiac surgery [12]. Our interactive tool supplies
quantitative decision support for these difficult strategic patient-mix decisions for both secondary and tertiary patient groups.

9.5.3 Reactive versus proactive

Finally, this thesis shows that being proactive rather than reactive is an efficient way of improving hospital performance:

- **At operational level.** By using accurate estimates of an operation’s duration, management can foresee the risk of overtime and the risk of a cancellation. Next, these accurate estimates should be used to objectively analyze a proposed operating room schedule and estimate the risk of overtime and the risk of cancellation in advance. Subsequently, if these risks exceed a certain threshold value, the operating room list should be adapted until both performance measures are satisfactory. In this way, management will be able to improve these important performance measures.

- **At tactical level.** Because many processes in hospitals are performed by physicians, the available capacity depends on the presence of these physicians. In academic hospitals, physicians are more frequently absent for different reasons. Moreover, due to the higher degree of subspecialty in academic hospitals, it is often more difficult to replace an absent physician. Also, due to the higher degree of subspecialty and other academic tasks (such as teaching medical students and doing research), patients can only be seen on a limited number of days during the week. Consequently, if one consultation session is cancelled, the access times of the patients increase more than with consultations that are held daily. Therefore, the irregular capacity of processes performed by physicians in academic hospitals fluctuates substantially. This artificial variability of capacity [13] reduces the efficiency or service level of the corresponding process. By using tactical management to reduce this variability, accessibility can be improved significantly with no investments. Furthermore, for a multi-step process, the waiting times prior to the subsequent steps will be reduced as well.

- **At strategic level.** By incorporating service levels into strategic planning, the additional capacity required could be calculated well in advance. Also, it would be possible to predict the effect of expected or preferred alterations in patient-mix on future service levels so that the necessary measures can be taken. By being proactive, high access times for consultations and high waiting times for surgery can be avoided. To make these patient-mix decisions, management should first decide upon the preferred levels of the KPIs, because the maximum workload depends on the preferred performance.

This conclusion also implies that an operating room’s utilization rate alone is not a good performance indicator for the operating room. To evaluate this performance correctly, management should also take into account the patient-mix and all KPIs
(such as the current waiting list for elective patients, the actual amount of overtime, and the actual number of cancellations due to overrun of previous surgery or due to prioritizing urgent patients).

This thesis shows how quantitative decision support, variability reduction of hospital processes, pooling or separating patient groups, and proactive behavior contribute to improved accessibility by more efficient use of resources. This will enable hospitals to face the challenges in the near future, and cope adequately with budget restrictions, increased competition among hospitals, and the consequences of the aging population for both demand and capacity.
References


