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

Citation for published version (APA):
The will to win means nothing without the will to prepare.
Juma Ikangaa
Summary

How to deal with fluctuations in hospital processes to improve accessibility?
S.1 Summary of General Introduction

In the near future, hospitals are expected to deliver more care because of the aging of the population. At the same time, they are being confronted with budget restrictions as well as a structural shortage of health care professionals due to this aging population. This situation requires hospitals to optimize their processes as much as possible and to use their resources as efficiently as possible. Moreover, hospitals have to improve their accessibility to be able to deal with the increasing competition from new initiatives such as private clinics and diagnostic centers. From the patient’s perspective, low access times are preferable so that the period of uncertainty and inconvenience during their illness is kept as short as possible. More importantly, a patient’s condition may deteriorate while spending time on a waiting list. For several types of cancer and cardiac disorders, the literature shows that if patients have to wait longer for treatment, their chance of full recovery decreases and their risk of death increases significantly.

The overall research objective of this thesis is 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 to the overall research objective, this thesis aims 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?

S.2 Summary of Research Studies

Chapters 2 through 8 are based on articles. Because this thesis should be seen as applied research, the order in which they are presented here is based primarily on the application area. Therefore, we have chosen to arrange the chapters according to the hospital departments rather than methodology.

Chapter 2, the first case study, aims to reduce the throughput times in the radiotherapy department of the Academic Medical Center (AMC) in Amsterdam, the Netherlands. We used a combination of queuing theory and computer simulation: queuing theory to provide insight into the effect of variability on the throughput times, and computer simulation to find the bottlenecks in the multi-step radiotherapy
process and to quantitatively compare the alternative solutions to reduce the throughput times efficiently. Despite investing in an additional linear accelerator to remove the presumed bottleneck, the throughput time targets were still not met. The simulation model indicated that the capacity of the outpatient department (OPD) had to be increased. A cost-effective alternative for increasing the capacity of this department was to reduce the substantial fluctuations in capacity. This resulted in both reduced access times and waiting times prior to the consecutive steps (the preparation phase and the actual treatment). This chapter concludes with practical suggestions on how to reduce fluctuations in capacity, and could be of interest to other radiotherapy departments or multi-step processes in hospitals.

Chapter 3 also describes a research study done in the AMC radiotherapy department. Another cost-effective alternative for increasing the capacity of the OPD was to pool or separate urgent and elective patients waiting for a consultation at a radiotherapy OPD. Queuing theory shows that pooling does not always shorten the waiting times of urgent patients. Computer simulation was used to investigate when to pool or separate urgent and elective patients in real hospital situations. The current situation at the AMC radiotherapy OPD requires the same number of consultations for both pooling and separating capacity. One slight advantage of separating capacity is that the service level improves somewhat. The more stringent the performance target for urgent patients, the more advantageous the separation of queues becomes. Jockeying patients from the elective to the urgent queue can further reduce the required capacity.

In Chapter 4, we also investigate whether to pool or separate urgent and elective patients, but this research study deals with the AMC endoscopy department. We used an iterative combination of computer simulation and integer linear programming; computer simulation was used to determine the minimum capacity per patient group to meet the corresponding access time target while keeping the percentage of double-bookings limited. One result was that the dedicated capacity for urgent patients could almost be halved without a large increase in double-bookings. Next, integer linear programming was used to reschedule the required capacity for all patient groups in an available procedure room with the required equipment and a specialized physician present. After we determined the minimum capacity per patient group, we used the simulation model to decide whether to pool or separate certain patient groups. First, the simulation model showed that separating the urgent general gastroenterological and colonoscopy patients into semi-urgent and urgent would use the total capacity more efficiently. Second, the simulation model showed 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.
Chapter 5 demonstrates that pooling various patient groups with the same urgency level can improve the efficiency of another central diagnostic facility, namely the magnetic resonance imaging scanner (MRI). In this case, we pooled elective patients from three patient groups and substantially reduced access times for MRI scans. To convince the radiologists to implement this solution, we used computer simulation to determine the order of magnitude of the reduction in access times: The reduction predicted with the simulation model was large enough to convince them, and the solution was successfully implemented in practice.

Chapter 6 describes how we also simulated the daily process to ensure that the reduction in access times for MRI scans would not lead to higher waiting times in the waiting room. In addition, we evaluate various alternative scenarios for reducing waiting times in the waiting room of the MRI department. One solution was to reduce 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, limiting such supervision would lead to an overall reduction in the variability of the process times. Subsequently, this would lead to lower waiting times in the waiting room. Another solution for reducing waiting times was to better estimate the scan duration. Consequently, we compared 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 contrast to Chapter 6, in Chapter 7 we actually demonstrate 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 not to reduce waiting times but to reduce the risk of overtime in the operating room and to reduce the percentage of cancellations due to overrun of previous surgeries. We concluded that applying econometric models can significantly improve the estimates of durations. From our investigation, it follows that applying the lognormal model (which is often used in the existing literature) is not the optimal one. Applying an objective method for estimating duration can prevent serious under- or overestimation of the duration. 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). Finally, the accurate estimates should be used to analyze a proposed operating room schedule objectively, and estimate the risk of overtime and the risk of cancellation in advance. Subsequently, if these risks exceed a certain threshold value, the proposed operating room schedule would have to be adapted until both performance measures are satisfactory. This way, management is able to improve these important performance measures proactively.
In Chapter 8, we describe an interactive tool that supports management with strategic patient-mix decisions and takes the key performance indicators (KPIs) into account. This tool enables management to alter the number of patients in various patient groups, and to see the consequences in terms of the KPIs of the OPD, operating room, and nursing ward. In this case study of the AMC ophthalmology department, we focused on the bottleneck, namely the operating room. To decide upon the appropriate level of detail, we started with a literature review to identify all factors that influence an operating room’s utilization rate. Next, we decided upon which factors were relevant to our study. For these factors, we used regression analysis and computer simulation to quantify the impact of an individual factor on the maximum allowed utilization rate. The average duration of an operation, the number of cancellations due to overrun of previous surgeries, and the waiting-time target for elective patients all turned out to have a significant impact. The interactive tool offers management quantitative decision support to enable them to be proactive towards expected alterations in the patient-mix. Hence, management can 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.

This research study also shows that the utilization rate of an operating room alone is not a good performance indicator for the operating room. To evaluate this performance correctly, management should also take the patient-mix and actual performance into account.

S.3 Summary of General Discussion

In Chapter 9, we present the main findings. Moreover, we describe the generalizability of the applied methodology and our results, the limitations of our research, and our recommendations for further research. We conclude Chapter 9 by listing the practical implications of this thesis.

S.3.1 Main findings

In terms of added value, the quantitative methods applied in our research studies contributed to the following:

1. Selecting an effective solution for a logistical problem by fact-based decision making with no hidden agenda (all chapters);
2. Providing an objective way of comparing alternative solutions before selecting the most cost-effective solution for meeting the performance target (all chapters);
3. Reducing the risk of undesirable outcomes for patients or personnel and unnecessary investments (Chapter 2);
SUMMARY

4. Quantifying the effect of variability reduction on demand, process times, and capacity (Chapters 2, 4, and 6);
5. Adapting historically grown capacity division of shared resources, and dividing the capacity among the subspecialties so that the overall performance is optimal (Chapter 4);
6. Deciding whether to pool patient groups that are different from a medical perspective but logistically similar (Chapters 3, 4, and 5); and
7. Enabling management to be proactive and anticipate the future situation to ensure that the preferred KPIs will be met structurally (Chapters 6, 7, and 8).

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.

Chapters 2, 4, and 6 demonstrate ways to decrease fluctuations in health care processes:
1. Decrease demand variability by stabilizing the capacity in the first step of a multi-step (radiotherapy) process;
2. Decrease variability in process times by increasing standardization of (MRI) protocols; and
3. Decrease capacity variability by introducing a backup system for physicians to lower the number of occasional closures of (endoscopic) procedure rooms and by preparing an overview of the total number of consultations for new (radiotherapy) patients per week for the coming quarter, and being proactive if the expected number is below a certain threshold value.
SUMMARY

In Chapters 3, 4, and 5, we explore the situations in which patient groups should be either pooled or separated to improve accessibility. We investigated two different reasons for this:

1. Pooling or separating patient groups based on urgency level: The radiotherapy study shows that the difference in urgency level will determine whether pooling or separating urgent and elective patients will be beneficial. 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 addition, the endoscopy study demonstrates that by separating the urgent patients into semi-urgent and urgent, the total capacity would be used more efficiently. As 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.

2. Pooling or separating capacity dedicated to specific patient groups that seem to be similar from a logistical point of view: In both the endoscopy study and the MRI study (Chapters 4 and 5), pooling elective patients in various patient groups will lead to reduced access times. Unfortunately, it was not possible to pool multiple patient groups in either study because personnel outside the (radiology) department needed to be present, or because different equipment and/or specialized physicians were required to perform the (endoscopic) procedures.

The research in Chapters 6, 7, and 8 was used to explore how to make hospital processes more predictable. By increasing standardization of (MRI) protocols, not only would the average duration decrease, but also the variability in process times. This variability reduction makes it easier to predict the actual duration more accurately. Although in Chapter 6 we showed only the effect of more accurate estimates of the duration on the relevant KPIs, in Chapter 7 we actually explain how to more accurately estimate the duration (in this case, of a surgical operation) by applying various econometric models. We conclude that applying such models can significantly improve these estimates. Finally, we show that more accurate predictions of the durations will result in less overtime and fewer cancellations due to overrun of previous surgeries.

In Chapter 8, we present an interactive decision-support tool we developed to enable the management of the AMC ophthalmology department to decide upon the future patient-mix to improve the KPIs. 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. Therefore, 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.

S.3.2 Generalizability of methodology and results
The queuing models that were used to provide insight into the effect of variability to the access times of the AMC radiotherapy OPD could also be used to provide such insights for other OPDs within the AMC and in other hospitals. Although the specific simulation models used in this thesis must be adapted to analyze similar problems in other hospital departments, the need for computer simulation remains. The various regression models used in Chapter 7 could be used to improve the accuracy of scheduling operations in other hospitals as well. Moreover, the applied regression models could also be used to improve the accuracy of scheduled consultations in an OPD, various procedures, and various scan durations in other hospitals. The two combined approaches - namely computer simulation and integer linear programming for the endoscopy study, and regression analysis and computer simulation to support strategic patient-mix decisions - are also applicable for other hospitals. In the latter case, computer simulation may have to be used for other included factors.

For the reasons mentioned earlier, other hospitals could also benefit from applying quantitative methods to support operational, tactical, and strategic decision making. All the examples of variability reduction presented in this thesis are useful for other hospitals as well. We cannot draw a general conclusion for the situations in which patient groups should be either pooled or separated to improve accessibility. In addition to the lognormal regression model, the alternative regression models described in Chapter 7 could also be applied in other hospitals. 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. 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.

S.3.3 Limitations
The first limitation of our research is that all studies concern unit logistics rather than chain or network logistics.

The second limitation of this thesis 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”. An obvious disadvantage of process improvements in a single unit is that the overall hospital performance might not improve.
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.

S.3.4 Further research
We suggest multiple directions for further research. 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.

Building on the research in this thesis, 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.

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.

Furthermore, additional research is necessary to develop scheduling guidelines that efficiently combine appointments within a single department or for multiple departments. Although an alternative way of reducing the number of hospital visits would be to introduce a partial walk-in system for central diagnostic facilities, research is necessary to investigate the situations in which the advantages counterbalance the disadvantage of increased fluctuations in demand during a given day.

S.3.5 Practical implications
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.

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. This thesis demonstrates multiple solutions for achieving cost reductions and improving quality and accessibility of hospital care.

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. Quantitative methods can assist in the difficult, often political, process of fairly redistributing the capacity 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. 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.

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. If these risks exceed a certain threshold value for a proposed operating room schedule, it should be adapted until both performance measures are satisfactory.

At tactical level. For multiple reasons, the capacity of processes performed by physicians in academic hospitals fluctuates substantially. By using tactical management to reduce this variability, accessibility can be improved significantly with no investments.

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
SUMMARY

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.

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 effectively with budget restrictions, increased competition among hospitals, and the consequences of the aging population for both demand and capacity.