5th International Industrial simulation Conference ISC'2007 June 11-13 Delft, The Netherlands
Eurosis-ETI. ISBN 978-77381-34-2

SIMULATION OF A HOLDING UNIT BETWEEN FIRST AID CENTER AND SPECIALTY WARDS: A CASE STUDY

A.C. van Bruggen, H.P.M. Veeke, G. Lodewijks
Delft University of Technology, Faculty of 3mE
F.K. van Lambalgen
Leiden University Medical Center
e-mail: *H.P.M.Veeke@tudelft.nl*

KEYWORDS
Health care, Resource management, Discrete simulation

ABSTRACT

The Leiden University Medical Center wants to increase its productivity. One suggestion is a temporary separation of acute and planned patients by means of a holding unit between the First Aid Center and the specialty wards. Acute patients can wait in the holding unit for a free bed in the wards and some can already be released from it after treatment. A simulation model shows that a holding unit indeed increases the productivity due to a scale advantage in the holding unit and increased bed occupancies in the wards. Moreover, it also leads to an improved quality of care and an improved controllability of the system

INTRODUCTION

Health institutes in the Netherlands have been cut back retrospectively on their government budget in order to stimulate an efficiency increase. Furthermore, in the near future there will be increasing market forces in the health care sector and also an increase in the quest for health care while the available personnel is decreasing. These developments force a productivity increase at the Leiden University Medical Center (LUMC). One of the hypotheses is that a further separation between elective (planned) and acute care will lead to a higher utilization of the available bed capacity. More specifically, it is suggested to have a temporary separation between acute and elective admission patients. Instead of going directly to the ward after visiting the First Aid Center, acute patients will then stay in a holding unit for a while. In this holding unit the diagnosis for the patient should be determined and the treatment should also already be started.

Several other hospitals make use of similar units but the objectives are most often different from the objective in this case. Some objectives that are often mentioned are to relieve the overcrowded First Aid Center or to make sure acute patients are diagnosed within a certain amount of time [e.g. 1]. However, the transfer of acute patients to the specialty wards is found to be a bottleneck and a good transfer system is thus seen as one of the key prerequisites for a good functioning of a holding unit [2-6]. In this research project the focus is especially on the interaction between the holding unit and the wards because the main objective of the holding unit at the LUMC is to have a productivity increase in the clinical capacity of the hospital. Furthermore, there are also large differences between the hospitals regarding the maximum length of stay of patients in the unit. This question is also addressed in this research project.

APPROACH

In order to investigate the hypothesis first the current processes are analyzed with the help of a systems approach [7, 8]. From this analysis it is determined what the shortcomings are of the current situation. Then it is investigated if a temporary separation between acute and elective patients by means of a holding unit can help to overcome these shortcomings and especially if it will lead to a productivity increase. A computer simulation is made with the simulation package TOMAS [10] based on Delphi to give a quantitative judgment of the advantages of a holding unit compared to the current situation.

ANALYSIS CURRENT SITUATION

In the current situation the acute patients first visit the First Aid Center of the LUMC. There it is determined whether an admission in the hospital is needed. If this is the case the patient will be transferred to the ward of the

Eurosis-ETI. ISBN 978-77381-34-2

concerning specialty or the patient will be transferred to another nearby hospital. On the specialty wards the acute and elective patients come together.
For every acute admission patient it is tried to find an admission spot in the LUMC but this is not always possible. Therefore, part of the acute admission patients has to be transferred to another hospital. Because the LUMC is an academic hospital it should at least be able to admit the acute patients who have a complex need of care. However, it also occurs that acute patients who actually need an admission in an academic hospital have to be transferred. These are called forced transfers. Therefore, more choices should be made regarding which patients should be admitted to the LUMC and which patients could just as well be transferred to a nearby non-academic hospital.

To find out how at the wards the inflow of acute patients is dealt with, the processes of several specialties have been further analyzed. Four specialties were chosen for this analysis, which all have a high number of acute patients. It happens to be that there are large differences between these specialties in the way they handle the inflow of acute patients. For example, one specialty uses all kinds of flexibility to be able to create an admission spot for an acute patient, among other things by placing a patient at the ward of another specialty or by calling off an elective patient. Another example is a specialty that does not have many possibilities for flexibility and therefore uses some overcapacity to be able to cope with the peaks in the inflow of acute patients.
However, overcapacity leads to a decrease in productivity while a high amount of flexibility in the current situation leads to a decrease in the overall quality of care as well as in the controllability of the system. So both an overcapacity and a high flexibility lead to unwanted side effects.
The question is now, can a holding unit lead to a productivity increase while at the same time contributing to overcome the unwanted side effects in the current situation? The proposed function structure in the LUMC with a holding unit is presented in Figure 1.

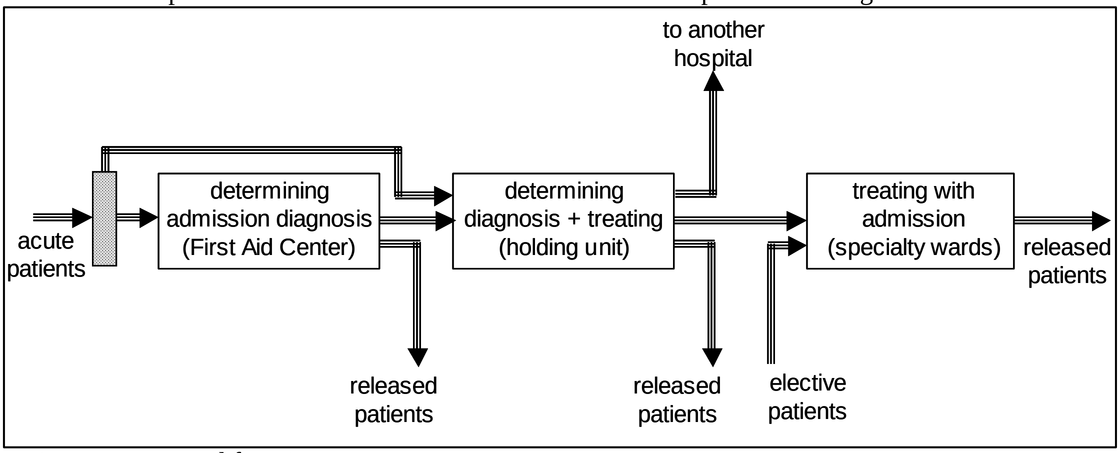

Figure 1: proposed function structure

As explained before most acute patients in the current situation first visit the First Aid Center after which they are transferred to a specialty ward if necessary. With the holding unit there will be one step in between. Upon arrival, part of the acute patients can be sent immediately to the holding unit if it is clear that the patient needs an admission. This will not be clear for each patient at the start so other patients first go to the First Aid Center and from there possibly to the holding unit and to the ward.

The holding unit would have two main functions as seen from the perspective of the wards, firstly filtering and secondly buffering. Filtering means that acute patients with a short length of stay can be diagnosed and treated in the holding unit and be released after treatment. These patients are thus kept away from the wards. For example, approximately one quarter of the acute patients is released on the same day as they were admitted. Buffering means that acute patients with a higher length of stay can stay in the holding unit until a bed is available at the ward.

A productivity increase with a holding unit is expected because of two reasons. Firstly, with a holding unit the number of acute patients at the wards will decrease and the length of stay of the acute patients at the wards will also decrease. Therefore, the bed capacity at the

Eurosis-ETI. ISBN 978-77381-34-2

wards can be decreased and this capacity can be used in the holding unit. Because of a scale advantage relatively less capacity is then needed in the holding unit. Secondly, the bed occupancy at the wards is expected to increase with a holding unit because acute patients can wait in the holding unit for a bed at the ward.

The non-occupied beds at the wards can then easily be filled with these acute patients. With a holding unit the acute admission patients actually become semi-elective patients because there is some time available to plan a bed for the acute patient at the ward. A holding unit in combination with a variable planning horizon for elective patients is therefore expected to lead to an even higher productivity. Instead of planning all the elective patients one week in advance some beds are kept free in the bed planning for possible acute patients. If it turns out that acute patients will not be occupying these beds it is then possible to call up some elective patients one day in advance. This system would then lead to less forced transfers of acute patients in combination with a high bed occupancy at the wards. One of the specialties already has positive experiences with such a variable planning horizon.

OBJECTIVE OF SIMULATION

To find out if the above-mentioned presumptions about the holding unit are true a computer simulation is made with the simulation package TOMAS based on Delphi. Two main questions will have to be answered with the simulation. The first question is whether the holding unit will indeed lead to a productivity increase, with and without a variable planning horizon. If this turns out to be true the second question is which holding time will be most advantageous. The holding time is the maximum length of stay of acute patients in the holding unit. Similar units in other hospitals mostly have a maximum length of stay varying from a few hours to two days. Therefore it will be useful to see if the simulation results can lead to a justified choice for the holding time.

OUTCOMES AND KEY PERFORMANCE INDICATORS

For several holding times it is determined with the simulation program if a productivity increase can be accomplished. That is, having the same amount of admissions as in the current situation only with less means. Therefore, the following outcomes are needed for this:

- The number of admissions.
- The number of means. The number of beds is used as a standard for the number of means. Both the beds at the wards and the beds in the holding unit get a weighting factor that indicates the number of nursing capacity needed per bed. For a holding bed this is two times as much as for a bed at the ward because of the higher level of care that is needed in the holding unit.

Besides the productivity increase, the holding unit is also expected to lead to an increase in the quality of care. The following outcomes are relevant for measuring the quality of care:

- The number of forced transfers. These are the patients for whom no bed is available in the LUMC and who have to be transferred to another hospital in the region. In the current situation the number of forced transfers as a percentage of the total number of acute admission patients is on average 15%. In the simulation this percentage will be set at a maximum of 3%.
- The number of elective patients who have to be rescheduled because there is no bed available for them at their admission date as a result of disturbances in the bed planning. This number as a percentage of the total number of elective patients is also set at a maximum of 3%.

The above two Key Performance Indicators are used to determine the number of beds needed at the wards. To determine the number of beds that is needed in the holding unit the following Key Performance Indicator is needed:

- The average waiting time of acute patients for a bed in the holding unit. This average waiting time is set at a maximum of half an hour.

The variables in the simulation are the number of beds at the wards, the number of beds in the holding unit and the holding time. The way each simulation run is executed will be explained later.

PROCESS DESCRIPTION

Figure 2 displays the patient flows in the simulation. In this figure only one ward is displayed but in the simulation for each of the four considered specialties one ward is used.

Eurosis-ETI. ISBN 978-77381-34-2

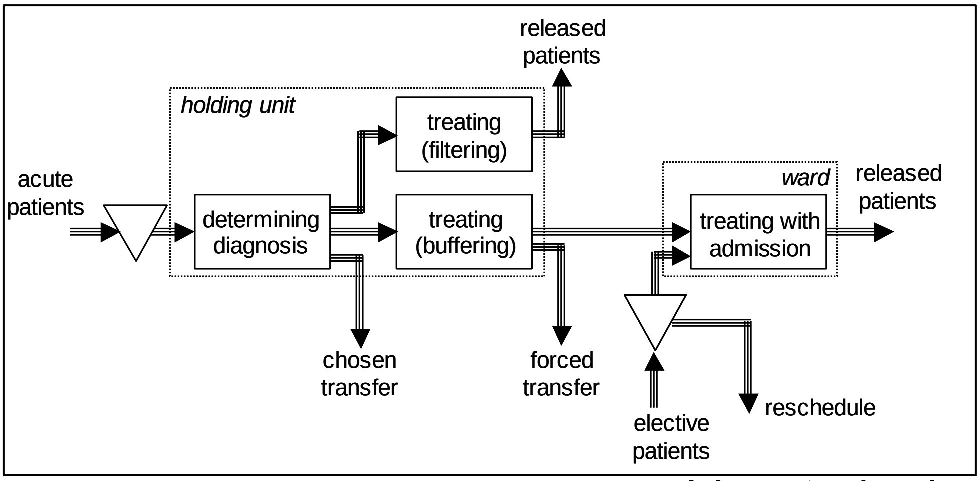

Figure 2: simulation description

The flowing elements in the simulation are the acute patients and the elective patients. All patients get an estimated length of stay on the basis of a distribution. They also get a real length of stay that is determined by adding or subtracting a correction to the estimated length of stay. This correction is also based on a distribution. The elements that are executing a process are as follows:

- Holding unit:
  Acute patients enter the holding unit according to an arrival pattern and they possibly have to wait until a bed is available in the holding unit. After determining the diagnosis, which takes a certain amount of time, for some of the patients it is chosen to transfer them to another hospital. For the other patients it is determined whether the estimated length of stay of the patient is smaller than the holding time. In that case the patient can be treated in the holding unit and be released from it (filtering). The other patients have to wait in the holding unit for a free bed at the ward while they are already treated (buffering). The Bed planning of the ward is informed that a bed has to be found for the acute patient. If no bed becomes available before the patient exceeds its maximum length of stay in the holding unit the patient has to be transferred to another hospital. This patient then has a forced transfer.
- Intake planning for elective patients (one for each ward):
  Every day a number of elective patients is planned and their admission date is set seven days later.
- Bed planning (one for each ward):
  For elective patients the bed at the ward is searched for that will become available as late as possible before the admission time of the patient. For acute patients it is first determined if there is already a bed available at the ward and if not the bed is searched for that will become available as soon as possible. If a bed is found it is reserved for the patient.
- Ward coordination (one for each ward)
  This process will release patients from the wards and admit acute or elective patients at the ward. Moreover, there may also be disturbances in the planning when the real length of stay of a patient is longer than its estimated length of stay. In that case it may be that other patients have to be rescheduled because their reserved bed will become available too late.

VALIDATION

Validation of the simulation model was done by a conceptual model validation, data validation and an operational validation [9]. The most important assumptions and simplifications in the simulation model are as follows:

- One ward per specialty. In reality most specialties have more wards.
- Constant number of beds at the wards. In reality this number may vary for example in holiday periods or because of ill personnel.
- The absence of flexibility. In reality patients can for example be released earlier or acute patients are placed temporary on another ward. In the simulation model this is not possible.

- A five-day working week. The influence of weekends is not accounted for.
- For the operational validation the holding time is set at 3 hours, which simulates the current situation. The percentage of forced transfers is taken as the indicator for the correctness of the model. With the current number of beds and the current number of patients this percentage would have to be equal to the real percentage of forced transfers. The operational validation is based on historical figures of both 2003 and 2004. The outcomes of this validation are shown in Table 1.

1. Make a simulation run with a high number of beds for both the holding unit and the wards.
2. Decrease the number of beds in the holding unit step by step until the average waiting time for acute patients reaches the maximum value of half an hour.
3. Decrease the number of beds at the wards step by step until one of the Key Performance Indicators reaches its maximum value.
4. If needed iterate step 2 and 3 until both the holding unit and the wards reach one of its maximum values for the Key Performance Indicators.

	Specialty 1 (23 beds)		Specialty 2 (21 beds)		Specialty 3 (30 beds)		Specialty 4 (45 beds)	
	2003	2004	2003	2004	2003	2004	2003	2004
Real percentage	8%	7%	26%	23%	0%	0%	15%	9%
Simulated percentage	8%	6%	27%	23%	2%	7%	10%	9%

Table 1: operational validation

The validation results show that the simulation model is performing well for specialty 1 and 2. Specialty 3 has a slight deviation from reality. In reality specialty 3 uses multiple wards, a variable planning horizon and there are special beds for acute patients. A similar validation for specialty 4 was not possible because specialty 4 often uses beds from other specialties to admit acute patients. Therefore, the current number of beds cannot be determined. It is only known that specialty 4 normally has 38 beds. The validation for this specialty is therefore done the other way around. It is thus determined what number of beds is needed in order to have the real percentage of forced transfers. Considering that specialty 4 often uses beds from other specialties the outcome of 45 beds seems acceptable as compared to the 38 beds that specialty 4 normally has of its own.

SIMULATION PLAN

For several holding times it is determined what number of means is needed and the number of beds is used as a standard for this. For each holding time the following steps should be taken:

First a reference run was made with the simulation model to determine the number of means in the current situation, which was found to be equal to 91. Although there is no holding unit in the current situation, the current situation can still be simulated with the model by setting the holding time at a maximum of only 3 hours as was just explained with the operational validation.
The main results of the simulation are shown in Figure 3. In each simulation run the number of elective and acute patients is the same as in the current situation.

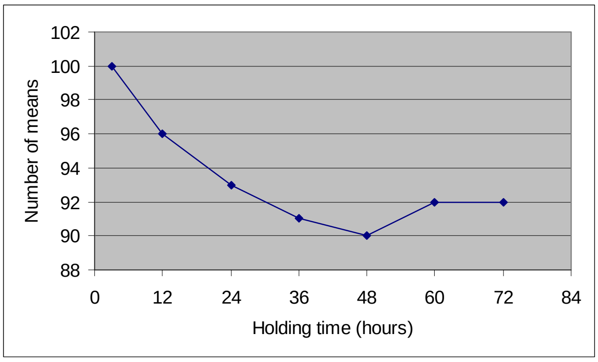

Figure 3: simulation results

The higher the holding time the more acute patients can be treated and be released from the holding unit. The scale advantage in the holding unit will therefore increase with increasing holding times. When the holding time increases it will also be easier to find a bed at the ward for an acute patient because the patient can stay longer in the holding unit. Therefore the bed occupancy at the wards can increase. Until a holding time of 48 hours those two effects lead to a decrease in the total number of means that is needed. For higher holding times there is a slight increase in the number of means. One of the reasons for this is that the bed occupancy at the wards is optimal for a holding time of 48 hours and decreases again with higher holding times. This decrease in bed occupancy is caused by the fact that the number of acute patients that has to go to the ward decreases so therefore, there are fewer possibilities to fill up the free beds at the wards.

As can be seen from Figure 3 the minimum number of means is with a holding time of maximum 48 hours, where the number is equal to 90. Compared with the reference run, where the number of means was 91, it seems like there is no productivity increase with the holding unit. However, the reference run concerns the current situation with approximately 15% forced transfers of acute patients to other hospitals while in the simulation model this percentage is reduced to 3%. This reduction was introduced in the simulation model because less forced transfers will lead to an increase in the overall quality of care. More acute patients are thus admitted with almost the same amount of means. This is a productivity increase but not as is wanted by the LUMC because they want to have the same number of admissions as in the current situation only with less means. In the analysis of the current situation it was mentioned that more choices should be made regarding which patients to admit in the LUMC and which patients to transfer to another hospital. Therefore a scenario is considered with the simulation model with a higher number of these chosen transfers. The results are shown in Figure 4 together with the first results.

This figure shows that with less means it is indeed possible to have the same number of admissions as in the current situation. Several sensitivity analyses with the simulation model have further shown that these results are reliable.

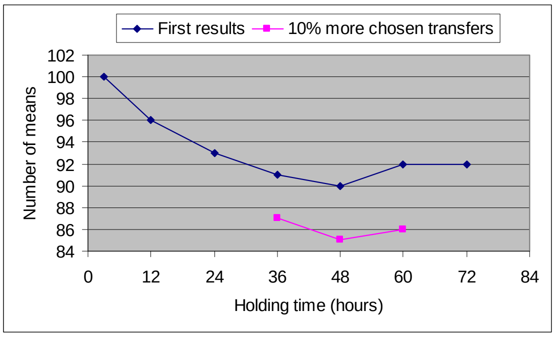

Figure 4: simulation results with more chosen transfers

The results in Figure 3 and Figure 4 are based on the historical figures for the year 2004. It follows from these results that the optimum holding time is 48 hours. To see if this is a reliable outcome the simulation results for 2003 are also determined. These are presented in Figure 5.

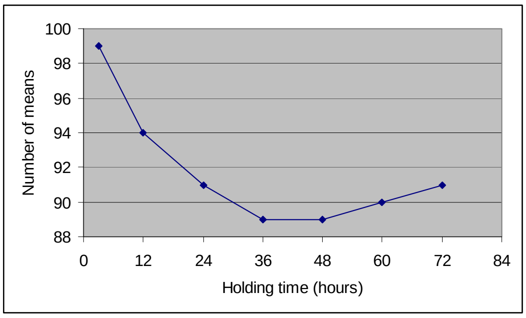

Figure 5: simulation results for the year 2003

The simulation results for the year 2003 are similar with the results for 2004. The optimum holding time is also around 48 hours.
It was suggested before to set up the holding unit in combination with a variable planning horizon for elective patients. Depending on the number of available beds extra elective patients can be called up for the next day. The expected productivity increase with this variable planning horizon is also investigated with the simulation model.
The results are shown in Figure 6 for the optimum holding time of 48 hours. The number of patients shown in the legend of the figure indicates how many patients at the most are called up for the next day per specialty.

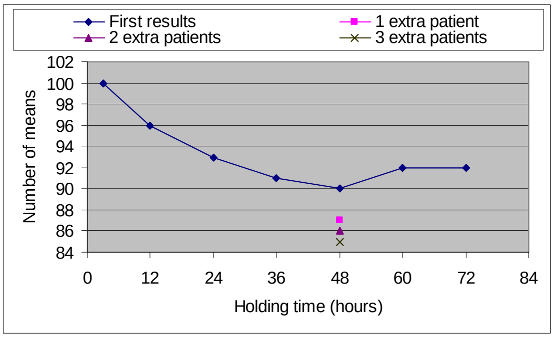

Figure 6: simulation results with a variable planning horizon

Figure 6 shows that a variable planning horizon can indeed further increase the productivity. This productivity increase is caused by higher bed occupancies at the wards.

CONCLUSIONS

It can be concluded that a further separation between elective and acute care will lead to a higher utilization of the available clinical capacity. The simulation has shown that with a separation of maximum 48 hours of acute and elective patients by means of a holding unit less clinical capacity is needed to admit the current number of patients in the LUMC. This is firstly because of a scale advantage in the holding unit and secondly because the bed occupancy at the wards can be increased. A variable planning horizon might further contribute to a productivity increase.
Moreover, a holding unit will lead to an increase in the quality of care because of less forced transfers to other hospitals, less elective patients to be called off and less transfers of acute patients between the wards. Finally, the overall controllability of the system will increase compared to the current situation.
On the basis of this research it was therefore recommended to the LUMC to set up a holding unit with a maximum length of stay of two days. Furthermore, it was recommended to have directives for each specialty indicating which acute patients to admit in the LUMC and which patients to transfer to a regional non-academic hospital. Ideally, these directives should also be part of an agreement system between the hospitals in the region. Regarding the variable planning horizon it was recommended to do some further research into the possibilities for an introduction of this measure. This will depend mainly on the willingness of the patients to be called up one day in advance and on the possibility to plan other processes in the same time.

References

1. Davies, M., 2003, Proposal to Establish a new Medical Assessment Unit at the Royal United Hospital Bath
2. Alberti, G., 2003, Emergency Assessment Units – a Checklist, Department of Health UK
3. Brennan, R., 2004, Medical Assessment & Medical Admission Units, The Irish College of General Practitioners
4. Clinical Epidemiology & Health Service Evaluation Unit, 2004, Models of Care to Optimise Acute Length of Stay, Melbourne Health
5. Huang, X.-M., 1998, Decision making support in reshaping hospital medical services, Health Care Management Science 1, 165-173
6. Levy, J., 2001, Medical Assessment & Planning Unit (MAPU) – Outcomes, Royal Brisbane Hospital, Proceedings of the National Demonstration Hospitals Program Phase 3 Conference
7. Bikker, H., 2001, Analysis and Design of Production Organizations, Lecture Notes, Delft University of Technology
8. Veld, J. in 't, 2002, Analyse van Organisatieproblemen, een Toepassing van denken in Systemen en Processen, 8th edition, Stenfert Kroese, ISBN 90-207-3065-7
9. Sargent, R.G. 2000. Verification, Validation, and Accreditation of Simulation Models, Proceedings of the 2000 Winter Simulation Conference, 50-59.
10. Veeke, H.P.M., Ottjes, J.A., 2000, "TOMAS: Tool for Object oriented Modeling And Simulation",Proceedings of Advanced Simulation Technology ASTC 2000, pp. 76-81, WashingtonDC, SCS International