Improving emergency department performance by revising the patient-physician assignment process

Peer-reviewed author version

VANBRABANT, Lien; BRAEKERS, Kris & RAMAEKERS, Katrien (2021) Improving emergency department performance by revising the patient-physician assignment process. In: Flexible services and manufacturing journal (Print), 33(3), p. 783-845.

DOI: 10.1007/s10696-020-09388-2
Handle: http://hdl.handle.net/1942/31721

Flexible Services and Manufacturing Journal

Improving emergency department performance by revising the patient-physician assignment process

--Manuscript Draft--

Manuscript Number:	FLEX-D-19-00186R1
Full Title:	Improving emergency department performance by revising the patient-physician assignment process
Article Type:	Manuscript
Keywords:	discrete-event simulation; Emergency department; Patient-physician assignment; Real-life case study; Healthcare operations
Corresponding Author:	Lien Vanbrabant Universiteit Hasselt Diepenbeek, BELGIUM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universiteit Hasselt
Corresponding Author's Secondary Institution:	
First Author:	Lien Vanbrabant
First Author Secondary Information:	
Order of Authors:	Lien Vanbrabant
	Kris Braekers
	Katrien Ramaekers
Order of Authors Secondary Information:	
Funding Information:	Fonds Wetenschappelijk Onderzoek (1173019N) — Ms Lien Vanbrabant
	Fonds Wetenschappelijk Onderzoek (S007318N) — Prof. dr. Kris Braekers

Response to reviewers' comments

We would like to thank the reviewers and the editors for thoroughly reading the paper and making useful comments. Their remarks have contributed considerably to the quality of our work. The text provided below between quotation marks is incorporated in the revised manuscript (revisions are shown in italics, removed parts are crossed out).

Editor in Chief

I agree with the review team on asking a clearer positioning of the paper in comparison to the current literature.

In order to make the contributions of the paper more clear, to better position the paper within the current literature, and especially to clarify the differences with the paper of Campello et al. 2017, some adjustments have been made to the paper. An overview of the main contributions is provided in the introduction (Section 1). All three main contributions differentiate this paper from existing literature on the case manager approach. Section 2.3 is added to the paper and provides a detailed discussion of the three main contributions. After providing an overview of the existing literature in Sections 2.1 and 2.2, it is indicated in Section 2.3 how this paper differs from current literature and why this paper significantly contributes to the existing body of literature for all three contributions. Special attention is given to the difference between this paper and the paper of Campello et al. 2017, as the reviewers indicated that this was not clear.

> " 1. Introduction
> […] The main goal of this study is to analyse the overall effect of a case manager approach with limited caseloads on the performance of a complex service system, such as an ED. **The paper has three main contributions. First of all**, this study is the first to show the benefits of a case manager approach with limited caseloads in a complex and realistic ED setting by use of discrete-event simulation. The simulation model is based on the ED of a Western European university hospital. Electronic health record (EHR) data of the ED under study is used as input to the simulation model. In addition to evaluating the case manager approach in a more realistic setting, the **second contribution** of this paper involves that the paper is the first to examine the impact of different caseload limits and queueing disciplines on the outcomes of introducing a case manager approach. Queueing disciplines are important from both an operational and patient safety perspective, given the limit on the number of patients per physician. An experimental design is conducted in order to determine the optimal case manager setting. As **third contribution**, the benefits of introducing a case manager approach with limited caseloads to enhance ED performance are shown by use of a real-life case study. It is concluded that the introduction of a caseload limit may significantly improve both length-of-stay (i.e. the time between patient arrival and discharge in the ED (LOS)) and door-to-doctor time (i.e. the time between patient arrival and the first consultation with a physician (DTDT)) of patients.
> The remainder of this paper is organised as follows. Section 2 provides an overview of the related literature, *positions this paper within the literature and outlines the main contributions of this study with regard to existing literature.* […]"
>
> "2. Related literature
> […] The discussion of the literature on this topic is structured as follows. Section 2.1 outlines the problem of physician multitasking, while Section 2.2 provides an overview of current literature on the case manager approach. *In addition, the contributions of this paper compared to the existing literature are outlined in Section 2.3.* […]
>
> *2.3. Contribution and relation to existing literature*
> The potential of a case manager approach to improve ED performance highly depends on the caseload limit and queueing disciplines. Nevertheless, a customised and optimised case manager system may result in increased physician throughput, better physician utilisation, lower service times and higher quality of care (Campello et al., 2017; Dobson et al., 2013; Kc, 2014). This study contributes to the current body of literature by investigating the potential of introducing a case manager system in an ED as a way to improve physician productivity, and consequently ED performance. A discrete-event

simulation model, based on a real-life case study, is used to examine the case manager system and to extend previous findings in a more complex and realistic setting.

In the study of Campello et al. (2017), queueing theory is used to investigate the case manager approach. Queueing theory is a popular method for modelling and analysing patient flow, particularly because of its simplicity and efficiency. However, as such analytical models mostly rely on closed-form mathematical formulations, they are not suitable to model the complex, stochastic and dynamic nature of healthcare systems unless introducing simplifying assumptions (Bhattacharjee and Ray, 2014; Saghafian et al., 2015). Therefore, when patient flows are highly complex, which is the case in EDs, simulation may be a suitable alternative. The main advantage of simulation over analytical modelling techniques such as queueing theory is that a great level of detail can be taken into account. In addition, stochastic and time-dependent characteristics of an ED can be included in the model. These characteristics of simulation enable to model the ED without extensively simplifying real-life operations by use of assumptions (Vanbrabant et al., 2019a). Consequently, our simulation model does not require many of the assumptions made in the queueing model of Campello et al. (2017) to simplify actual ED operations. As a result, our study provides a more realistic evaluation of ED performance under the case manager approach. In Table 1, a comparison of our study with the queueing model of Campello et al. (2017) is provided regarding the most important assumptions.

In addition to avoiding the use of simplifying assumptions by using simulation techniques, this paper contributes to the current body of literature on two other aspects. Firstly, all three system parameters of the case manager approach are investigated simultaneously by means of an experimental design. Campello et al. (2017) only look at the optimal caseload limit for given queuing disciplines, while Dobson et al. (2013) focus on the optimal queueing disciplines without determining the optimal caseload limit. As all three system parameters may have an impact on performance under the case manager approach, and the parameters may be interdependent, analysing them simultaneously gives a more complete view of the potential benefits of introducing a case manager approach. Secondly, the potential performance improvements that can be obtained by introducing a case manager approach are shown by use of a real-life case study, while previous studies focus on a theoretical setting only."

Table 1: Comparison of assumptions in queueing theory model of Campello et al. (2017) and our simulation model

	Campello et al. (2017)	Our study
Arrival rate	Stationary	Non-stationary, depending on the hour of the day and the day of the week
Patient types	Homogeneous	Heterogeneous
Patient priority	FIFO	FIFO and based on triage code
Physician consultation duration	Single distribution	Distribution dependent on patient type and stage of a patient in ED patient flow
External delay probability	Single probability	Probability dependent on patient type
External delay duration	Single distribution for total external delay duration	All external delay processes (i.e. examinations, waiting for results) are modelled in detail
Physician capacity	Constant capacity of three physicians	According to a realistic shift schedule
Model parameters	Based on Graff et al. (1993)	Based on analysis of real EHR-data

I think the work as a potential, but it suffers from lack of clarity, repetitions of sentences and "excess" of word that creates more confusion than clarity. This probably does not help the reader to fully receive the main messages of this work. I would consider shortening some parts (for example, figure 1 in introduction could perhaps be avoided, presentation of results in section 5 is too dispersive and does not help to fully catch the most relevant results) and applying more Occam's razor principles though all the text.

The paper has been rewritten and restructured on multiple places in order to enhance the clarity. In addition, several paragraphs are shortened or removed from the paper to avoid repetition and to mainly focus on the main contributions and insights of the paper. The revised version of the paper has the same number of pages than the original version, but the dimensions of the graphs are increased, two graphs are added and an additional paragraph on the contributions of the paper and the positioning of the paper within current literature has been added. Consequently, the original paper has been shortened significantly through the avoidance of repetition, removal of unnecessary paragraphs and restructuring of the paper. The most important adjustments are described below:

1. As requested, Figure 1 containing a comparison of physician tasks with and without multitasking, and the definition of multitasking, have been removed from the introduction. This information was not necessary to be able to understand the content of the paper.
2. The detailed description of patient flow under the case manager approach, and the example used to clarify this patient flow, are removed from the introduction to avoid repetition.
    - Previously, patient flow under the case manager approach was explained in the introduction (Section 1), the literature review on the case manager approach (Section 2.2) and the description of the simulation model under the case manager approach (Section 4.1.1). The general description of patient flow under the case manager approach is now provided in Section 2.2.
    - The example of patient flow under the case manager approach is placed in Section 4.1.1, as we already referred to the example in this section. Section 4.1.1 provides a description of the adjustments that should be made to the simulation model in order to include the case manager approach. By giving an example, the explanation of the difference in patient flow between the situation with and without case manager approach is more clear. As the example provides a description of patient flow under the case manager approach, the general description is removed from Section 4.1.1.
    - Figure 5 (in the original version of the paper) is removed from Section 4.1.1 and the figure belonging to the example is added to this section. Figure 5 was a general figure that schematically illustrated patient flow under the case manager approach. Section 4.1.1 describes our simulation model and not the general situation, so the general figure was confusing. The figure belonging to the example schematically represents patient flow as included in our simulation model. As this figure is now included in Section 4.1.1, the general figure became unnecessary.
3. Section 2.1 has been restructured and shortened in response to the comments of reviewer 1. A detailed description of the adjustments to this section can be found in our response to the comments of Reviewer 1.
4. The results section (Section 5) has been rewritten, restructured and shortened in order to better focus on the main findings.
    - Firstly, the discussion of the scenario without multitasking is removed from Section 5.2. In Section 5.1, it was already concluded that the findings for both multitasking scenarios are comparable (only the effect sizes differ). In addition, the goal of Section 5.1 is to compare the scenario with and without multitasking, so the next sections should not necessarily focus on this difference anymore. As a description of both the scenario with and without multitasking effect resulted in a lot of repetition in Section 5.2, the scenario without multitasking effect is not discussed anymore and the graphs are placed in Appendix D.
    - Secondly, Section 5.3 has been rewritten and shortened. In addition, through adding more structure to this section, the focus is only on the main findings and the description of the main findings is more clear.

In addition to the comments below, I would like the authors to motivate better their design of experiments. The use of common random numbers (CRN) is for reducing the variance with the same experimental effort, but one hundred replications should be a large number, so I don't see large benefits of using CRN. However, CRN might complicate statistical analysis because experiments are not independent anymore and I wonder if the Fischer Test is affected by that. Apparently, the variance between groups should be underestimated biasing F statistics (will it be underestimated ?). If this was true, the test would be more robust but the significant factors (or interactions of) could be not significant anymore in reality. Please, try to check this technical issue so that we are safe from this potential risk.

In order to check the potential problems with the reliability of the statistical tests through the use of common random numbers, multiple books on statistics and discrete-event simulation have been consulted (e.g. Field 2013, Kelton et al. 2015, Law 2007). The first concern deals with the use of both common random numbers and multiple replications. Common random numbers are used to ensure that a statistical comparison of multiple experimental design settings (i.e. factor combinations) based on the simulation model is reliable. Through the use of common random numbers, the differences in simulation output between the different factor combinations are not the result of a difference in random numbers. Without common random numbers, the random numbers that are used to generate patients, service times, etc. can differ between the simulation model used for testing factor combination 1 and the simulation model used for testing factor combination 2. Through using common random numbers, all random numbers that are used in the simulation models for testing factor combinations 1 and 2 are the same. Consequently, if simulation output differs between the two factor combinations, the difference in output is not caused by different random numbers but by the different factor combination.

Multiple replications, on the other hand, are used to ensure that the simulation model output of a single factor combination is reliable. Because of the use of random numbers, two replications of the same factor combination can generate different results. As a result, no conclusion can be made based on a single replication because of the randomness. Therefore, multiple replications are executed such that the mean over all replications is a reliable estimate of actual performance for a specific factor combination (i.e. the sample mean is a good estimate of the population mean). In our simulation model, we used the method of Law (2007) to determine the appropriate number of replications such that the mean over all replications is a reliable estimate of actual performance, which resulted in 100 replications. In summary, common random numbers are used to enhance the reliability of a comparison of multiple factor combinations, while multiple replications are used to get a reliable estimate of performance measures for a single factor combination.

The second concern deals with the reliability of the statistical tests when using common random numbers, because the experiments are not independent from each other as every factor combination is tested in the same ED setting. A repeated-measures full factorial ANOVA is used, as this test is adjusted to the fact that every factor combination (i.e. full factorial) is evaluated in the same ED setting (i.e. repeated-measures) through the use of common random numbers. In addition, all assumptions of the repeated-measures full factorial ANOVA are tested to ensure that the F-statistics are reliable, and the assumptions are fulfilled. Besides a repeated-measures full factorial ANOVA, a univariate ANOVA is executed for the caseload limit factor. The use of a balanced design (every factor combination is tested in the same ED setting for the same number of replications) improves the reliability of this test (Field, 2013). In addition, the univariate ANOVA consists of multiple paired t-tests, and Kelton et al. 2012 state that the reliability of a paired t-test is not impacted by the use of common random numbers. As a univariate ANOVA executes multiple t-tests based on the same set of experiments, the Bonferonni correction is applied to further ensure reliability of the results.

Regarding the Fisher test, we are not sure what the link between this test and our statistical analyses is. According to us, the Fisher test is used for categorical data, namely to determine whether the frequency distribution of one categorical variable is different depending on the value of another categorical variable. We use categorical variables in our experimental design, but the performance measures for which the different experimental design settings are compared by means of the ANOVAs are continuous variables and no frequency distributions. Therefore, we don't think the Fisher test is suitable for our experiments.

In the paper, we adjusted the explanation of the statistical analysis in Section 4.4 slightly to make the choice of the statistical tests and the reliability of those tests more clear.

> "4.4 Statistical analysis
> […] The experimental design consists of a full factorial design with only repeated-measures factors, corresponding to priority Qp, priority Qi, priority stage and caseload. All 300 factor combinations of the experimental design are evaluated by use of the simulation model. Each factor combination is tested in the same ED setting, as common random numbers are used. As a result, a repeated-measures full factorial analysis of variance (ANOVA) is executed on the simulation results. An ANOVA is performed to test whether or not a relationship between one of the factors and the KPIs is statistically significant. In addition, interactions between experimental design factors can be identified. *A repeated-measures full factorial ANOVA is a specific version of the ANOVA-test which is adjusted to the fact that all factor combinations of the experimental design are tested in the same setting, for example by use of common random numbers (Field, 2013).* A total of eight repeated-measure full factorial ANOVAs, one for each KPI, is executed on the simulation results.
>
> A first important assumption of a repeated-measures full factorial ANOVA in order to ensure accuracy of the F-statistic, is sphericity (i.e. equality of variances of the differences in output between factor combinations for a single ED setting). The assumption of sphericity is tested with Mauchly's test. A violation of the sphericity assumption results in an increased Type I error rate in the statistical analysis. This involves that the probability of finding a significant effect in the ANOVA while there is no effect in

reality, increases. *Therefore, in case the sphericity assumption is violated based on the results of Mauchly's test, the conservative Greenhouse-Geisser (G-G) estimate of the F-statistic is used. The G-G estimate adjusts the degrees of freedom to compensate for the violation of the sphericity assumption (i.e. increased Type I error rate) (Field, 2013). Otherwise, the sphericity assumed estimate of the F-statistic can be used.* A second assumption is normality of the dependent variable (i.e. KPI). When the degrees of freedom are sufficiently large (at least 20) and group sizes are equal, the F-statistic controls the Type I error rate well under conditions of non-normality (Field, 2013). As the simulation is replicated 100 times for each factor combination, the degrees of freedom are sufficiently high for the normality condition to be fulfilled. In addition, all factor combinations are tested on the same set of patients by use of common random numbers in the simulation model, *which makes the experimental design balanced. This implies that the F-statistic is a reliable estimate in our setting, making robust checks (e.g. bootstrapping) unnecessary.* […]"

Other minors:

- **Many graphs (e.g. Figures 8 and 9), shows only the average, but authors say that some treatments are really different, maybe boxplots (or some other graphical tool) could represent better**

  The graphs showing only averages (Figures 8 and 9 in the original version of the paper) have been removed from the paper and are replaced by two graphs that show the minimum, maximum, median and mean DTDT and LOS, respectively, over all priority settings and patient types (Figures 8 and 9 in the revised version of the paper).

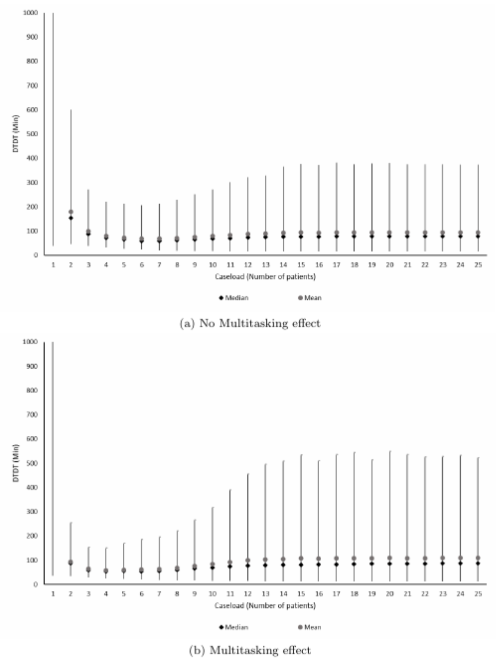


Fig. 8: Minimum, median, mean and maximum DTDT over all priority settings and patient types as a function of caseload.

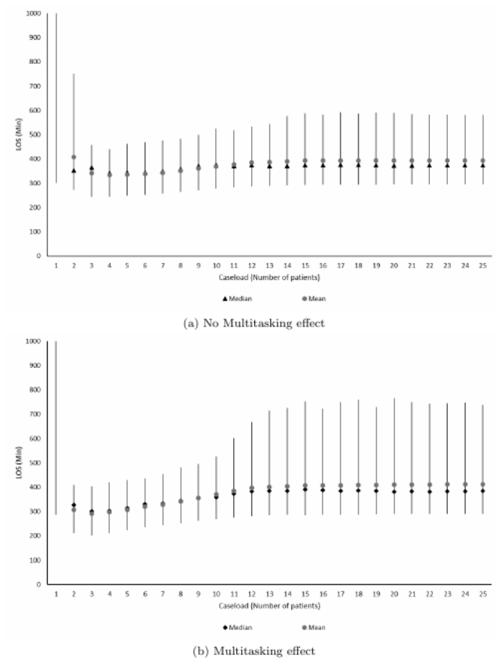


Fig. 9: Minimum, median, mean and maximum LOS over all priority settings and patient types as a function of caseload.

- **Graph dimensions are too small**

  The dimensions of the graphs have been increased where possible, and the newly added graphs are also shown in these larger dimensions. Only the size of Figures 12 and 13 (in the revised version of the paper) has not been adapted, as these seem readable and this would otherwise result in these graphs being spread over 4 pages, which would increase the length of the paper very much. However, we can also increase the size of these graphs if this is preferred.

- **Anova table should be in the main text in my opinion**

  We added all tables regarding the statistical analyses as appendix to the paper (Appendix C), instead of as online appendix.

Area Editor 1

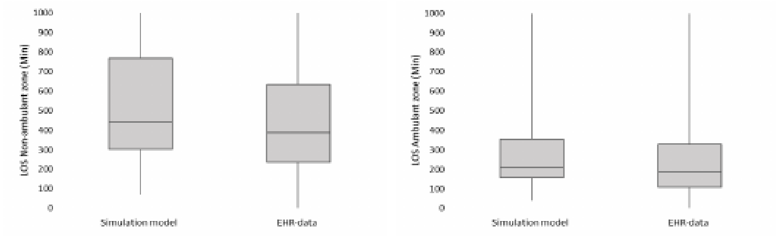
All reviewers and AE have concerns about the contribution. Even though using simulation, it is expected that there exist some novel modeling approaches or specific insights can be obtained through this study. It is important that the simulation needs to be validated by real data. Only verification by expertise is not enough, since that is typically confirmed qualitatively but no quantitative check. In addition, if the expertise can predict or the model is only corrected in those parts, then why do we need a simulation? Thus, it is important to validate the model using emergency department data.

As already indicated in the answer to the first comment of the Editor in Chief, the contributions of this paper and the positioning of this paper within existing literature are made more clear throughout the paper. This also involves a discussion of the modelling approach we use, and the advantages and differences compared to existing studies. A detailed overview of the adjustments that are made to clarify the contributions of the paper, can be found in our answer to the first comment of the Editor in Chief.

In this paper, a simulation model is developed based on real data extracted from the electronic health records (EHRs) of the ED under study. An overview of the available EHR-data is added to the paper in Appendix A. The simulation model has been validated by use of EHR-data from the ED under study. In order to clarify this in the paper, Section 3.2 has been rewritten and extended, and some examples of quantitative validation are presented in Appendix B. Nevertheless, as not every aspect of patient flow through the ED is included in the EHR-data, the quantitative validation is extended with a validation process that consisted of meetings with ED staff and operational management. This way, we could ensure that processes not included in the EHR-data are also realistically represented and that performance measures which cannot be retrieved from the EHR-data are a realistic representation of the actual performance of the ED. The fact that ED staff and operational management is consulted in the validation phase, does not mean that they can predict outcomes from the simulation model. They can evaluate whether the simulation model provides a good representation of current ED operations, but they cannot evaluate the impact of adjustments to current ED operations on ED performance. The validated simulation model of current ED operations is used in order to investigate the impact of the case manager approach and the experimental design factors on ED performance, which cannot be predicted beforehand.

> "3.2 Verification and validation
> The simulation model is verified and validated in several ways to ensure that current ED operations are correctly represented. *Firstly,* verification is used to evaluate the patient flow logic in the simulation model. The process involved checking the internal logic of the individual modules and the relationship between the modules, by use of the visual animation tool in Arena and by going through the model with ED staff and operational management
>
> *Secondly,* validation is executed to ensure that the model is a close representation of the real system. *Validation involved a comparison of model output against actual operations and performance by use of the available data. Both ED characteristics (e.g. patient arrival patterns, patient characteristics such as triage codes, service times and boarding time) and key performance indicators (KPIs) are compared. The comparison is executed based on mean values, distributions or boxplots, depending on the ED characteristic or performance measure to compare. For example, the length of stay of patients in the ambulant and non-ambulant zone is compared by means of boxplots and the hourly arrival patterns are compared graphically (see Appendix B). A high amount of overlap exists, so the simulation model is assumed to give a reliable representation of actual ED performance. In addition to a comparison of model output against actual ED operations by means of the available EHR-data, the validation phase also consisted of meetings with the operational management and ED staff. This is necessary, as not all ED processes and KPIs are included in the EHRs (e.g. no start times of activities, physician consultations not registered)."*

“Appendix B: Validation



(a) Non-ambulant zone (b) Ambulant zone

Fig. 1: Validation boxplot for the LOS of patients in the non-ambulant and ambulant zone

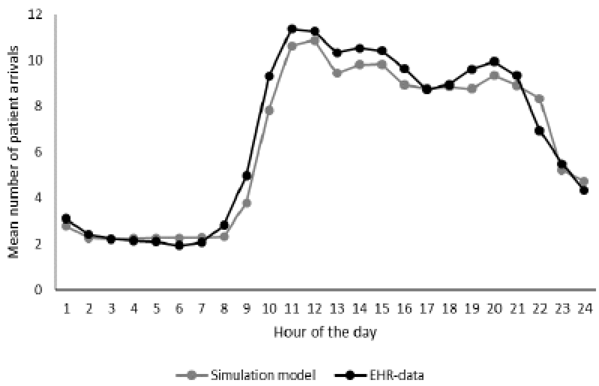


Fig. 2: Graphical validation of hourly number of patient arrivals ”

Area Editor 2

Both reviewers questioned about the contribution of this paper. Please clarify the contribution in the new version. Please clarify the relationship between this paper and Campello et al. (2017).

The contributions of this paper are clarified in Section 2.1, with a special focus on the differences with the paper of Campello et al. (2017). A detailed comparison between the queueing model of Campello et al. (2017) and our simulation model is provided. An overview of the adjustments we made to the paper, can be found in the answer to the first comment of the Editor in Chief.

Reviewer 1

General evaluation

In this paper, the authors introduce a case manager approach with limited caseloads to improve the emergency department (ED) performance. To verify the effectiveness of this approach, this paper built a discrete-event simulation model in a complex and realistic ED setting. The case manager approach is characterized by three parameters, i.e., caseloads limit, pre-assignment queueing discipline and internal queueing discipline. The impact of these parameters on ED performances in terms of length-of-stay (LOS) and door-to-doctor time (DTDT) is evaluated. Experimental results show that both LOS and DTDT can be improved significantly after the implementation of a case manager approach.

Overall assessment

1) The manuscript is well written and the results seem interesting. Both the ideas and the insights are presented clearly in the manuscript.

We thank the reviewer for his positive comments on the content and relevance of the paper.

2) The whole paper is based on a discrete-time simulation model built in Arena. Although both the idea and the conclusions are interesting, I am doubtful about the contribution of this paper for this journal.

Because of the practical relevance of improving ED operations, the innovative nature of the case manager topic within healthcare research, and the use of simulation as an operations research method to investigate the idea of a case manager approach in a realistic setting by use of a real-life case study, we are convinced that our paper fits within the scope of Flexible Services and Manufacturing Journal. In addition, emergency department simulation studies have been published in this journal in the past (e.g. Kuo, Y.-H., Rado, O., Lupia, B., Leung, J.M.Y., Graham, C.A., 2016. Improving the efficiency of a hospital emergency department: a simulation study with indirectly imputed service-time distributions. Flexible Services and Manufacturing Journal 28, 120–147). Finally, the editor in chief, the area editors and the other reviewer did not mention the fact that the paper does not fit within this journal.

3) Literature review, especially the subsection 2.1 should be reorganized.

The original introduction of the literature review has been removed and replaced by a new introduction in order to better introduce the content of Sections 2.1, 2.2 and 2.3. Sections 2.1 and 2.2 have been rewritten and restructured to improve the clarity. In addition, Section 2.3 has been added, as explained in our response to the first two comments of the Editor in Chief.

“2. Related literature
An interesting and obvious way to improve ED operations is by focusing on staffing solutions (Gul and Guneri, 2015; Vanbrabant et al., 2019a). Because of the strict healthcare budgets, human resources should be used as efficiently as possible (Duma and Aringhieri, 2018). Current staffing solutions merely concern personnel capacity and shift schedules. The reorganisation of nurse- and physician-related processes in order to improve their productivity is barely investigated in ED simulation research. An innovative way to improve the efficiency of physician-related processes is the application of a case manager approach with limited caseloads as a way to reduce the negative effects of excessive physician multitasking (Campello et al., 2017). The discussion of the literature on this topic is structured as follows. Section 2.1 outlines the problem of physician multitasking, while Section 2.2 provides an overview of current literature on the case manager approach. In addition, the contributions of this paper compared to the existing literature are outlined in Section 2.3.”

“2.1 Multitasking
Among ED staff, physicians are the most costly resource and one of the main bottlenecks. In order to sustain acceptable performance levels, it is crucial to make optimal use of the available physician capacity. In this regard, physician multitasking is essential, which involves that a physician is responsible for a set of patients at a single moment in time. Multitasking limits the amount of physician idle time that would otherwise be caused by the external delays (e.g. laboratory or radiological examinations) between patient consultations (Kc, 2014; Gunal and Pidd, 2006).

Nevertheless, as a result of crowding, the number of patients simultaneously assigned to a single physician may reach very high levels. *This may negatively impact ED performance, as a too high workload may offset the productivity gains obtained through a reduced physician idle time (Levin et al., 2007; Kc, 2014). Multitasking has both advantages and disadvantages, and the overall effect of multitasking on ED performance and physician productivity depends on multiple factors. Delasay et al. (2015) developed the 'load effect on service times' (LEST) framework to predict the impact of system load on server productivity. The number of patients simultaneously managed by a physician (i.e. the level of multitasking) is directly related to physician workload, as a higher number of patients results in more care-related activities to be performed (Kc and Terwiesch, 2009). Therefore, the LEST framework can be used to predict the effect of physician multitasking and workload on physician productivity in the ED, as explained in the next two paragraphs.*

On the one hand, multitasking can have a positive effect on physician productivity. Firstly, as already indicated, multitasking limits the amount of idle time. Secondly, multitasking results in a high (perceived) workload among physicians, which can induce rushing (i.e. decrease in service times) and task reduction (i.e. decrease in the number of executed tasks such as examinations, or early discharges) (Batt and Terwiesch, 2016; Delasay et al., 2015; Forster, 2003; Kc and Terwiesch, 2009). In addition, Delasay et al. (2015) mention a reduction of service times in overloaded systems when server performance is visible to others waiting in the queue because of social pressure (Kc and Terwiesch, 2009; Tan and Netessine, 2014). In an ED, this may be the case when physicians are observable by patients in the waiting room. *An undesirable side effect of this second category of advantages, especially in a healthcare context, is a reduced quality of the provided services (Kc and Terwiesch, 2009). In addition to health risks, a deteriorated quality of care may result in patient revisits.* These revisits create additional workload in the future (e.g. the next day) and may be avoided. Because of the negative side effects, the presence of service time reductions is questionable in an ED setting, especially for physician consultations, where quality of care is a major concern (Batt and Terwiesch, 2012).

On the other hand, excessive multitasking may negatively impact physician productivity. Firstly, multitasking incurs a setup time when switching between tasks, or patients. This setup time may increase with the number of patients per physician because of cognitive limitations (Aral et al., 2012; Batt and Terwiesch, 2016; Delasay et al., 2015; Kc, 2014). The setup time before each consultation consists of a fixed component which is caused by multitasking but independent of the number of patients, and a variable component which depends on the number of patients. The fixed component is a result of forgetting, which implies the loss of patient information from immediate
memory when switching between patients. The variable component is called cognitive sharing, and represents the fact that reviewing a patient file before a consultation takes longer when the number of patients per physician increases (Delasay et al., 2015; Kc, 2014). *Secondly, the service time per patient may increase with the number of patients per physician, because of a higher number of task interruptions (e.g. by nurses, specialists, patients or family members) (Chisholm et al., 2000; Dobson et al., 2013; Gunal and Pidd, 2006; Kc, 2014) and the occurrence of fatigue after an extended period with high workloads (Batt and Terwiesch, 2012; Delasay et al., 2015; Kc and Terwiesch, 2009). Besides a decreased physician productivity, fatigue also has a negative impact on quality of care (Batt and Terwiesch, 2012; Delasay et al., 2015; Kc and Terwiesch, 2009).*

The overall effect of multitasking on physician productivity is the result of a complex interplay between all the different mechanisms (Batt and Terwiesch, 2016). Aral et al. (2012) find that both the advantages and disadvantages of multitasking are likely to have non-linear effects on productivity. In case the workload becomes too high, the negative effects possibly offset the positive effects. Kc (2014) concludes that a reduced idle time and an increased setup time between consultations are the two predominant mechanisms in an ED setting. Kc (2014) finds a concave relationship between the level of multitasking and productivity, with the level of multitasking measured as the number of patients per physician. Similar results are found in the setting of a recruiter firm by Aral et al. (2012)."

4) Please recheck the whole manuscript and correct the mistakes in sentences.

We checked the whole manuscript and corrected all the mistakes.

I have some comments that I am sure the authors can fix or offer an explanation. For these reasons, I recommend minor revision for this manuscript.

<u>Detailed comments</u>
Comment #1: Related literature

1) **Page 7, line 15: “Firstly, sharing physician capacity among multiple patients increases the service time per patient because of time sharing mechanisms”, which contradicts the saying on page 6 that “the positive effect of multitasking on physician efficiency may be amplified as high workloads induce rushing (i.e., decrease in service time)”. These two sayings are reasonable under different situations, but you should better reorganize this part to explain the impact of multitasking clearer.**
2) **Page 7, 2nd paragraph: The authors list some literatures to show the negative impact caused by multitasking, i.e., the increasing in service time and setup time. The first, third and fourth points in this paragraph are all related to the “increasing in service time”. I suggest that these three reasons should merge into one.**

Both comments deal with Section 2.1. This section has been rewritten and restructured by taking both comments into account. In addition, as suggested by the reviewer, the three reasons that are provided for an increasing service time are merged into one. The revised version of this section is provided in the answer to the third ‘overall assessment’-comment above.

Comment #2: Case study and simulation model description

1) **Page 12, Figure 3: According to this figure, there are at most two physician consultations for each patient. However, the saying on page 2 states that “physicians have multiple consultations with each patient, which are interspersed by external delays due to radiological and laboratory examinations”. The authors should give more explanations regarding Figure 3.**

The saying on page 2 is linked to the general description of ED patient flow and the case manager approach. In general, patients in an ED have multiple consultations with a physician, and how many consultations this are depends on the specific ED and the characteristics of the patient. In our study, we use the ambulant zone of a Western-European ED as basis for simulation model construction. In this zone, the majority of patients only has two consultations with a physician. Figure 3 (Figure 1 in the revised manuscript) provides an overview of ED patient flow in the ambulant zone of the ED under study. As we assume that all patients have two consultations with an ED physician, after which they can have an additional consultation with a specialist, this figure provides a correct representation of patient flow in our simulation model. In order to clarify this, we added the assumption explicitly in the description of the figure.

“3.2.1 Current patient flow through the ED
[…] The results of the laboratory and/or radiological examinations are analysed by the responsible physician, after which a second physician consultation takes place. *It is assumed that all patients have two consultations with an ED physician, as this is the case for the large majority of patient in the ambulant zone.* Additional examinations or advice from a medical specialist may be requested before a diagnosis can be made. In case specialist advice is needed, the patient enters the consultation-examination cycle again. Based on the diagnosis, a treatment can be prescribed and/or initiated within the ED (e.g. medication, plaster, etc.). […]”

2) **Page 14: I am not sure the subsection 3.2.2 is necessary in this paper. The authors can either verify and validate the simulation model with more experimental data or just put these words right at the end of subsection 3.2.1.**

The area editor requested to clarify the validation process, and especially the part of the validation process in which the simulation model is validated by use of real emergency department data. Therefore, this section is extended with a more comprehensive description of the validation process. In addition, an example of the validation by use of real data is provided in Appendix B. The revised version of Section 3.2.2 and Appendix B can be found in the answer to the comment of Area editor 1.

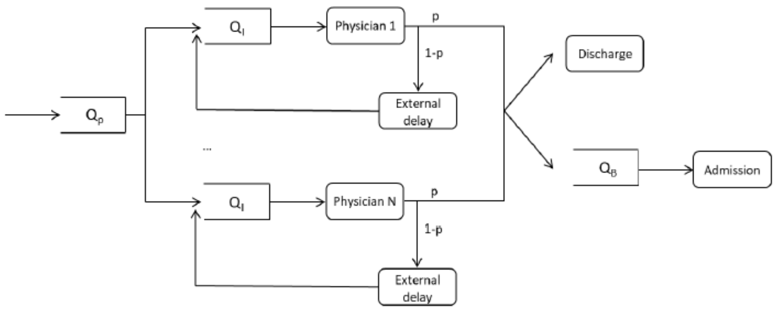
Comment #3: Methodology

1) **Page 15, line 48: the defined QB (boarding queue) here has never been used in the following part. My advice is to delete QB to avoid unnecessary confusion for readers.**
2) **Page 16, Figure 5: according to Figure 5, it seems that the number of consultations for each patient subjects to a geometric distribution with parameter p. However, the authors “assume that all patients have two consultations with a physician, as this is true for the large majority of patients in the**

ambulant zone", which contradicts the information in figure 5. Given the assumption that at most two consultations are considered, the saying that "after each physician consultation, a patient has a certain probability p to be medically finished" is incorrect.

We agree with the reviewer that Figure 5 was misleading. In answer to both comments, and because the Editor in Chief indicated that repetition should be avoided in the paper, Figure 5 and the accompanying description of patient flow based on this figure have been removed from the paper. Figure 5 provided a general description of patient flow under the case manager approach, not adjusted to our case study setting. Therefore, the figure contained a boarding queue and a probability to be medically finished. Because this resulted in some confusion, Section 4.1.1 has been rewritten such that patient flow under the case manager approach is explained by means of an example and a figure which are based on the patient flow in our case study.

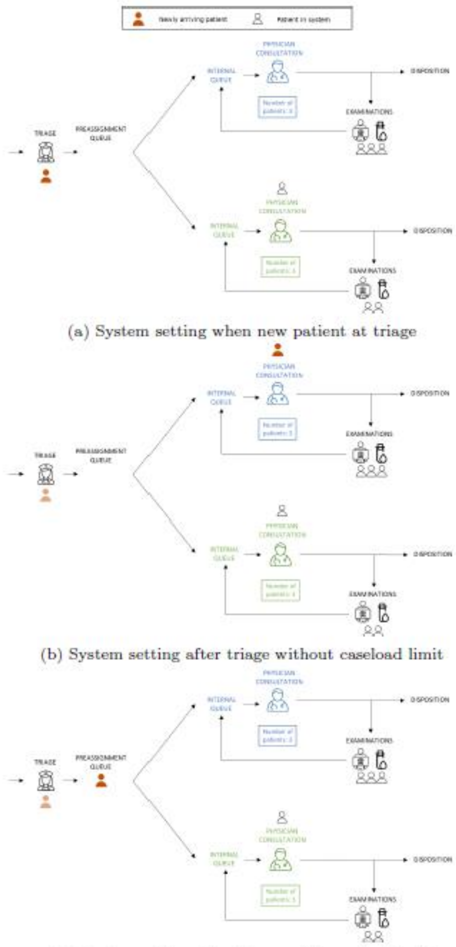
"4.1.1 Patient flow under the case manager approach
~~Figure 2 already illustrated ED patient flow under the case manager approach with an example. Figure 5 provides a more general representation of ED patient flow in the ambulant zone under the case manager approach. The figure focuses on the processes in which a physician is involved, as only these processes are organised differently. After triage, ambulant patients have to wait in the pre-assignment queue (QP) until capacity becomes available with one of the physicians (i.e. case managers). A physician may be working below the maximum caseload, in which case a newly arriving patient can be assigned to a physician immediately. If multiple physicians are working below capacity, a patient is assigned to the physician with the highest remaining capacity. If no physician is working below capacity, the patient has to wait in the pre-assignment queue until another patient is medically finished (i.e. a physician finished the treatment of a patient in the ED and gives approval to the patient to leave the ED). Once assigned to a physician, the patient enters the internal queue (QI) of that specific physician and stays with the same physician during the complete stay in the ED. The internal queue consists of all patients waiting for the services of a single physician. The number of internal queues is equal to the number of physicians in the ED, N.~~



~~Fig. 5: ED patient flow under case manager approach~~

The difference in patient flow between the situation with and without a case manager approach is illustrated by use of an example in Figure 3. We suppose that seven patients and two physicians are present in the ED, and the caseload limit is set at three patients. Patient flow consists of the same processes in both systems, but the difference lies in the assignment of patients to physicians. After triage, each patient should be assigned to a physician for a first consultation. In Figure 3a, the seventh (orange) patient is undergoing triage and should still be assigned to a physician. In the situation without caseload limit (Figure 3b), physician assignment takes place whenever a physician becomes available to have a first consultation with a newly arriving patient. There is no limit on the number of patients per physician. Since the blue physician is idle, the patient at triage can be assigned to that physician, resulting in an empty pre-assignment queue. With the presence of a caseload limit (Figure 3c), physician assignment can only take place when a physician is working below the caseload limit of three patients. In the example, three patients are assigned to each physician, so the seventh patient has to wait in the pre-assignment queue until another patient is medically finished and disposed from the ED. When a place becomes available at a physician, the patient is assigned to that physician and joins the internal queue. Once assigned to a physician, a patient stays with the same physician during the complete stay in the

ED. The internal queue of a physician consists of all patients waiting for the services of a single physician. The number of internal queues is equal to the number of physicians in the ED. […]"



(a) System setting when new patient at triage

(b) System setting after triage without caseload limit

(c) System setting after triage with caseload limit

Fig. 3: Comparison of ED patient flow with and without caseload limit

Comment #4: Results and discussion

1) **Page 23, Figure 8: what is the priority setting (i.e. a possible combination of the priority QP, priority QI, and priority stage factor levels) in this experiment? If adopt stage 1 first policy, this result shown in Figure 8 is obviously wrong. Details of the parameter setting in this experiment should be presented clearly.**

   In Section 5.1, the focus is on the difference between the scenario with and without multitasking effect. Therefore, the figures in this section show the average LOS and DTDT over all patient types and priority settings. So these figures are not related to a specific priority setting, but show the average results over all 12 priority settings. It is correct that the graphs would have been wrong if they were related to a specific priority setting in which stage 1 patients have priority.

   As there might be large differences in the effect of a caseload limit on LOS and DTDT depending on the specific priority setting, Figures 8 and 9 of the original manuscript are replaced by two new graphs (Figures 6 and 7 in the revised manuscript) which show the minimum, maximum, mean and median LOS and DTDT over all priority settings. The minimum, maximum, mean and median LOS and DTDT are calculated based on the results of all simulation runs for all priority settings. Based on insights from these graphs, it is concluded that a case manager approach impacts ED performance in both multitasking scenarios, and as the impact is higher for the multitasking scenario we decide to only discuss this scenario in the remainder of the results section.

2) Page 24-25: mistake in sentence: "The results of the scenario without multitasking effect are included in A These results..."

This sentence is adjusted: *"*The results of the scenario without multitasking effect are included in *Appendix A. These results...".*

Reviewer 2

This paper focus on the case manager approach in emergency departments using a discrete-event simulation model based on a real-life case study. The impact of three main parameters, i.e. caseload limit, pre-assignment queuing discipline and internal queuing discipline, on the ED performance in terms of length-of-stay and door-to-doctor time is well investigated based on numerical experiments. The work also carries out a case study to confirm the application of the proposed approach in a complex and realistic settings, and provides managerial insights to the management ED. The problem is of good significance and the paper is well written. Here are some questions and comments for authors:

We thank the reviewer for his positive comments on the content and relevance of the paper.

1. **As is stated, Campello et al. (2017) introduced and investigated the case manager approach with limited caseloads in different service system using queuing theory with restricting assumptions and this paper applies a discrete-event simulation to study the case manager problem in a more complex and realistic environment. The authors should clarify the relation between these two papers clearer. What assumptions make it difficult to solve the ED problem by theoretical model and what is the theoretical contribution of this paper?**

   The main contributions of the paper are summarised and clarified in Section 2.3. In addition, a comparison of the modelling approach used in Campello et al. (2017) and the modelling approach used in our paper is provided in this section. The advantages of discrete-event simulation compared to queueing theory are discussed, and the assumptions made in the model of Campello et al. (2017) which are modelled more realistically in our model are indicated. The content of Section 2.3 can be found in our answer to the first comment of the Editor in Chief, as this comment deals with the same concern.

2. **The simulation model is validated by expertise of ED staff due to the lack of some data in the EHRs. Can the model be partly validated, for example, the service time at one physician?**

   Although this may not have been clear from the original manuscript, the simulation model has also been quantitatively validated by means of electronic health record (EHR) data of the emergency department under study. A comparison of service times against actual data was not possible, as start timestamps of all activities are lacking in the EHR-data. Therefore, these are validated by expertise of ED staff and operational management. However, simulation model output regarding several key performance indicators and ED characteristics is validated by means of real data. The description of the validation process in Section 3.2.2 is extended and adjusted to clarify the validation process. In addition, an example of the quantitative validation of simulation model output is provided in Appendix B. The specific adjustments made to Section 3.2.2 and Appendix B can be found in our response to the comment of Area editor 1.

3. **Do all patients with different TDs share the same medical finishing probability p?**

   In Figure 5 (removed in the revised manuscript), a general representation of patient flow under the case manager approach was provided. This figure contained a medical finishing probability after each physician consultation. However, in the ambulant zone of the ED that we are modelling, all patients have two consultations with a physician (i.e. we do not need this probability p in our simulation model). In other EDs the number of consultations can depend on the triage code of a patient. Therefore, the medical finishing probability is not very relevant in our paper. We removed the figure from the paper, as reviewer 1 also indicated that the figure was confusing.

4. **Authors should briefly introduce the data settings and confidence of the simulation results in the main body rather than the technical report.**

   A more detailed description of the EHR-data used as basis for simulation model construction is included in Section 3.2. In addition, an overview of all data captured by the EHRs is provided as an appendix (Appendix A). The explanation of the validation of the simulation results by means of real data is also extended in the paper, as explained in the answer to comment 2. We hope this suffices to clarify the use of real-life data in our simulation model construction. Because the Editor in Chief indicated that we should try to shorten the paper, we did not extend the simulation model description in Section 3.2.1 by including the accompanying input data analysis for every simulation model component. This information is provided in the technical report.

"3.2. Simulation model description

Among the simulation techniques, discrete-event simulation is the most appropriate technique to model ED operations. Discrete-event simulation enables to capture randomness and complexity of ED patient flow at the level of an individual patient (Brailsford and Hilton, 2001; Maidstone, 2012; Mohiuddin et al., 2017). The Arena simulation software is used for simulation model construction. Several complex processes and some patient flow logic are modelled by writing custom code in Visual Basic for Applications, which is embedded in Arena. In addition, key performance indicators not included in the standard Arena output are defined and calculated by use of Visual Basic for Applications.

The reliability of a discrete-event simulation model highly depends on the availability of input data on the real system. In our study, data extracted from the electronic health records of the hospital under study is used as basis for simulation model construction. The dataset consists of anonymised patient records containing personal, medical and patient flow information of all patients that visited the ED in 2016. This information contains, amongst others, symptoms, diagnosis, type of inflow, timestamps of the different processes a patient undergoes in the ED, types of examinations, outflow destination, etc. *Tables 1 to 4 in Appendix A provide an overview of the data fields in the extracted dataset.* The quality of the dataset has been analysed based on the framework and assessment techniques described in Vanbrabant et al. (2019b). *Multiple quality problems were identified in the dataset, such as missing values and violations of the logical order of activities. These quality issues are taken into account when using the input data as a basis for simulation model construction, for example by removing incorrect or missing values.* The use of a cleaned, high quality dataset enabled to model the complete patient flow in a detailed and realistic way, for instance by including time- and day dependent patient arrival rates, individualised patient pathways and stochastic service times. *In order to obtain this information, the EHR data is analysed by use of Excel, R and the Arena input analyser. The information from the input data analysis is supplemented with observations, interviews and surveys, as the EHR data does not (correctly) capture every aspect of patient flow through the ED. An example of information that is obtained by use of surveys are process times, as not all activities are recorded in the EHRs and start timestamps of activities are lacking. The surveys provided a minimum, most likely, and maximum duration for all ED processes, which are used to estimate a triangular distribution for each service time.* A detailed description of the content of the EHR dataset and how the dataset is analysed and used as input to the simulation model can be found in a technical report.

The remainder of this section contains a general description of the most important aspects of the patient flow through the ambulant zone as included in the simulation model. Furthermore, *the verification and validation of the simulation model are briefly discussed*. A more detailed description of the simulation model regarding the ambulant zone, and a description of the non-ambulant zone, can be found in the technical report."

5. Why DTDT and LOS of multitasking scenario are better than that of scenario without multitasking when the caseload limit is small, while worse when the caseload limit is large?

The difference can be explained by the fact that when the caseload limit is small, the mean consultation setup time is smaller in the scenario with multitasking than in the scenario without multitasking. When the caseload limit is large, the mean consultation setup time is higher in the scenario with multitasking than in the scenario without multitasking. As the impact of physician multitasking (i.e. caseload limit) on ED performance is mainly determined by the reduced idle time and the increased consultation setup time, the difference in consultation setup time between the two scenarios has an impact on the results.

6. Authors should pay more attentions to the statement of the whole paper for the syntax errors. For example, the quotes in Line 19 and Line 46 on Page 2.

This issue has been resolved, and the complete paper is checked and corrected with regard to syntax errors.

Noname manuscript No.
(will be inserted by the editor)

Improving emergency department performance by revising the patient-physician assignment process

Received: date / Accepted: date

Abstract Emergency departments (EDs) are continuously exploring opportunities to improve their efficiency. A new opportunity lies in revising the patient-physician assignment process by limiting the number of patients simultaneously assigned to a single physician, which is defined as the application of a case manager approach with limited caseloads. The potential of introducing a case manager approach with limited caseloads as a way to improve physician productivity, and consequently ED performance, is investigated by use of a discrete-event simulation model based on a real-life case study. In addition, as the case manager system is characterised by three parameters that can be customised and optimised (i.e. caseload limit, pre-assignment queueing discipline and internal queueing discipline), the impact of these parameters on the effectiveness to improve ED performance in terms of length-of-stay and door-to-doctor time is evaluated. To the best of our knowledge, this paper is the first to examine the potential of a case manager system with limited caseloads in a complex service system like a real-life ED, and to investigate the impact of the three system parameters on the results. The outcomes of the study show that performance can be improved significantly by introducing a case manager system, and that the system parameters have an impact on the effect size.

Keywords Discrete-event simulation · Emergency department · Case managers · Real-life case study · Healthcare operations

1 Introduction

Nowadays, emergency departments (EDs) worldwide are inevitably confronted with crowding. The negative consequences of crowding on patients, caregivers and the hospital are countless: high waiting times, deteriorated quality of care, high stress-levels of caregivers, financial losses, etc. (Hoot and Aronsky, 2008;

Address(es) of author(s) should be given

Paul and Lin, 2012; Pines et al., 2011). In order to overcome these negative consequences, hospital managers are constantly looking for opportunities to improve ED performance. Because of the strict budgets, the main focus is on improving the efficiency of ED operations while preserving a high quality of care (Abo-Hamad and Arisha, 2013; Saghafian et al., 2015; Vanbrabant et al., 2019a).

An interesting and straightforward way to improve ED operations is by focusing on staffing solutions (Gul and Guneri, 2015; Vanbrabant et al., 2019a). Because of the strict healthcare budgets, human resources should be used as efficiently as possible (Duma and Aringhieri, 2018). Among ED staff, physicians are the most costly resource and one of the main bottlenecks. In most EDs, each patient is assigned to a single dedicated physician. Physicians have multiple consultations with each patient, which are interspersed by external delays due to radiological and laboratory examinations, and waiting for the results of these examinations. Because of these external delays, a physician is responsible for a set of patients at a single moment in time to avoid unproductive idle time. Being responsible for a set of patient necessitates multitasking and results in productivity gains by reducing the idle time of a physician caused by the external delays. However, cognitive limitations of physicians, and the resulting patient switching costs, cause productivity losses if the number of patients assigned to a single physician increases. In addition, an excessively high workload may negatively impact the quality of care (Kc, 2014). Therefore, an interesting opportunity to improve such a system entails the revision of the patient-physician assignment process by limiting the number of patients simultaneously assigned to a single physician. This approach to patient-physician assignment is defined as a case manager approach with limited caseloads, and is proven in this paper to be an effective way to enhance performance in a realistic ED setting.

The main goal of this study is to analyse the overall effect of a case manager approach with limited caseloads on the performance of a complex service system, such as an ED. The paper has three main contributions. First of all, this study is the first to show the benefits of a case manager approach with limited caseloads in a complex and realistic ED setting by use of discrete-event simulation. The simulation model is based on the ED of a Western European university hospital. Electronic health record (EHR) data of the ED under study is used as input to the simulation model. In addition to evaluating the case manager approach in a realistic setting, the second contribution of this paper involves that the paper is the first to examine the impact of different caseload limits and queueing disciplines on the outcomes of introducing a case manager approach. Queueing disciplines are important from both an operational and patient safety perspective, given the limit on the number of patients per physician. An experimental design is conducted in order to determine the optimal case manager setting. As third contribution, the benefits of introducing a case manager approach with limited caseloads to enhance ED performance are shown by use of a real-life case study. It is concluded that the introduction of a caseload limit may significantly improve both length-of-stay (i.e. the time between patient arrival and discharge in the ED (LOS)) and door-to-doctor time (i.e. the time between patient arrival and the first consultation with a physician (DTDT)) of patients.

The remainder of this paper is organised as follows. Section 2 provides an overview of the related literature, positions this paper within the literature and outlines the main contributions of this study with regard to existing literature. In section 3, the ED under study is described in detail and the simulation model construction based on real-life input data is discussed. Section 4 describes the methodology used to investigate ED performance under the case manager approach. The section contains a description of the incorporation of the case manager approach within the simulation model. Furthermore, the experimental design, operational measures and statistical analyses used to determine the optimal case manager setting in terms of caseload limit and queueing disciplines are discussed. Section 5 provides a detailed discussion of the main results. The paper ends with some practical considerations in Section 6, and conclusions and possibilities for future research in Section 7.

2 Related literature

An interesting and obvious way to improve ED operations is by focusing on staffing solutions (Gul and Guneri, 2015; Vanbrabant et al., 2019a). Current staffing solutions merely concern personnel capacity and shift schedules. The reorganisation of nurse- and physician-related processes in order to improve their productivity is barely investigated in ED simulation research. An innovative way to improve the efficiency of physician-related processes is the application of a case manager approach with limited caseloads as a way to reduce the negative effects of excessive physician multitasking (Campello et al., 2017). The discussion of the literature on this topic is structured as follows. Section 2.1 outlines the problem of physician multitasking, while Section 2.2 provides an overview of current literature on the case manager approach. In addition, the contributions of this paper compared to the existing literature are outlined in Section 2.3.

2.1 Multitasking

Among ED staff, physicians are the most costly resource and one of the main bottlenecks. In order to sustain acceptable performance levels, it is crucial to make optimal use of the available physician capacity. In this regard, physician multitasking is essential, which involves that a physician is responsible for a set of patients at a single moment in time. Multitasking limits the amount of physician idle time that would otherwise be caused by the external delays (e.g. laboratory or radiological examinations) between patient consultations (Kc, 2014; Gunal and Pidd, 2006). Nevertheless, as a result of crowding, the number of patients simultaneously assigned to a single physician may reach very high levels. This may negatively impact ED performance, as a too high workload may offset the productivity gains obtained through a reduced physician idle time (Levin et al., 2007; Kc, 2014).

Multitasking has both advantages and disadvantages, and the overall effect of multitasking on ED performance and physician productivity depends on multiple factors. Delasay et al. (2015) developed the 'load effect on service times' (LEST) framework to predict the impact of system load on server productivity. The number of patients simultaneously managed by a physician (i.e. the level of multitasking) is directly related to physician workload, as a higher number of patients results in more care-related activities to be performed (Kc and Terwiesch, 2009). Therefore, the LEST framework can be used to predict the effect of physician multitasking and workload on physician productivity in the ED, as explained in the next two paragraphs.

On the one hand, multitasking can have a positive effect on physician productivity. Firstly, as already indicated, multitasking limits the amount of idle time. Secondly, multitasking results in a high (perceived) workload among physicians, which can induce rushing (i.e. decrease in service times) and task reduction (i.e. decrease in the number of executed tasks such as examinations, or early discharges) (Batt and Terwiesch, 2016; Delasay et al., 2015; Forster, 2003; Kc and Terwiesch, 2009). In addition, Delasay et al. (2015) mention a reduction of service times in overloaded systems when server performance is visible to others waiting in the queue because of social pressure (Kc and Terwiesch, 2009; Tan and Netessine, 2014). In an ED, this may be the case when physicians are observable by patients in the waiting room. An undesirable side effect of this second category of advantages, especially in a healthcare context, is a reduced quality of the provided services (Kc and Terwiesch, 2009). In addition to health risks, a deteriorated quality of care may result in patient revisits. These revisits create additional workload in the future (e.g. the next day) and may be avoided. Because of the negative side effects, the presence of service time reductions is questionable in an ED setting, especially for physician consultations, where quality of care is a major concern (Batt and Terwiesch, 2012).

On the other hand, excessive multitasking may negatively impact physician productivity. Firstly, multitasking incurs a setup time when switching between tasks, or patients. This setup time may increase with the number of patients per physician because of cognitive limitations (Aral et al., 2012; Batt and Terwiesch, 2016; Delasay et al., 2015; Kc, 2014). The setup time before each consultation consists of a fixed component which is caused by multitasking but independent of the number of patients, and a variable component which depends on the number of patients. The fixed component is a result of forgetting, which implies the loss of patient information from immediate memory when switching between patients. The variable component is called cognitive sharing, and represents the fact that reviewing a patient file before a consultation takes longer when the number of patients per physician increases (Delasay et al., 2015; Kc, 2014). Secondly, the service time per patient may increase with the number of patients per physician, because of a higher number of task interruptions (e.g. by nurses, specialists, patients or family members) (Chisholm et al., 2000; Dobson et al., 2013; Gunal and Pidd, 2006; Kc, 2014) and the occurrence of fatigue after an extended period with high workloads (Batt and Terwiesch, 2012; Delasay et al., 2015; Kc and Terwiesch, 2009). Besides a decreased physician productivity, fatigue also has a negative impact on quality of care (Batt and Terwiesch, 2012; Delasay et al., 2015; Kc and Terwiesch, 2009).

The overall effect of multitasking on physician productivity is the result of a complex interplay between all the different mechanisms (Batt and Terwiesch, 2016). Aral et al. (2012) find that both the advantages and disadvantages of multitasking are likely to have non-linear effects on productivity. In case the workload becomes too high, the negative effects possibly offset the positive effects. Kc (2014) concludes that a reduced idle time and an increased setup time between consultations are the two predominant mechanisms in an ED setting. Kc (2014) finds a concave relationship between the level of multitasking and productivity, with the level of multitasking measured as the number of patients per physician. Similar results are found in the setting of a recruiter firm by Aral et al. (2012).

2.2 Case manager approach

The concave relationship between the level of multitasking and productivity implies that an optimal level of multitasking may exist (Kc, 2014). The case manager system with limited caseloads[1] is proposed by Campello et al. (2017) as a way to optimise performance of a service system by limiting excessive multitasking.

Campello et al. (2017) define a case manager as “a server who is assigned multiple customers and repeatedly interacts with those customers”. In an ED context, the server is a physician, and the customers are the patients in the ED. During their stay in the ED, each patient is assigned to a single dedicated physician. In addition to the use of dedicated physicians, an upper limit (i.e. caseload limit) may be imposed on the number of patients simultaneously assigned to a single physician. The introduction of a caseload limit entails the formation of two separate physician queues in the ED. In the pre-assignment queue, patients are awaiting physician assignment. As most EDs are crowded, the patient census in the ED may exceed the sum of the maximum caseloads of all physicians, resulting in a significant pre-assignment queue. Once patients are assigned to a physician, they join the internal queue of that specific physician to await their consultation. Most patients are repeatedly placed in the internal queue as they need multiple interactions with a physician, which are interspersed by laboratory or radiological examinations, before disposition from the ED (Campello et al., 2017)

Dobson et al. (2013) state that the prioritisation rules in both queues have a major impact on physician throughput and ED performance under a case man-

[1] Shortened as case manager approach in the remainder of this paper.

ager approach. This results in the case manager system being characterised by three parameters that should be customised and optimised. Besides a decision on the caseload limit, the queueing disciplines for the pre-assignment and internal queue should be defined (Campello et al., 2017; Dobson et al., 2013; Kc, 2014). Classical queueing disciplines include First In First Out (FIFO), Last In First Out (LIFO) and Random Order of Service (ROS) (Li et al., 2019; Tan et al., 2012). As these classical queueing disciplines fail to take differences in the urgency of patients into account, static priority queueing disciplines can be used as an alternative. In these queueing disciplines, the priority is determined based on some patient attributes. Triage codes are frequently used as they reflect patient acuity levels. Patients with the same triage code are prioritised by use of FIFO. In ED simulation literature, FIFO and a static priority based on triage codes are the most commonly used queueing disciplines and best reflect reality (Batt and Terwiesch, 2016; Ferrand et al., 2018; Li and Stanford, 2016; Saghafian et al., 2015; Tan et al., 2012).

As most ED patients encounter multiple consultations with a physician, choosing the optimal queueing discipline in the internal queue is complicated by an additional decision. In addition to selecting the most appropriate queueing discipline for patients waiting in the same internal queue (e.g. FIFO or triage code), prioritisation rules based on the stage of patients within their ED stay can be determined. Newly arriving patients or patients near the end of their ED stay may be prioritised, or an equal priority can be applied for all patient waiting for the same resource regardless of their stage (Dobson et al., 2013). Patients are first prioritised based on their stage, after which the selected queueing discipline is applied to prioritise patients within a single stage. Ferrand et al. (2018) conclude that in order to minimise LOS, newly arriving patients should not be prioritised over patients waiting for a follow-up consultation. Cildoz et al. (2018) on the other hand prioritise newly arriving patients over patients waiting for a second consultation. This may be explained by the inclusion of DTDT in their objective function. Dobson et al. (2013) find that the optimal decision to maximise throughput depends on the capacity of the ED and the presence of interruptions. Tan et al. (2012) investigate the potential of prioritising patients based on the shortest remaining consultation time. However, this is difficult to implement in practice and benefits are questionable.

2.3 Contribution and relation to existing literature

The potential of a case manager approach to improve ED performance highly depends on the caseload limit and queueing disciplines. Nevertheless, a customised and optimised case manager system may result in increased physician throughput, better physician utilisation, lower service times and higher quality of care (Campello et al., 2017; Dobson et al., 2013; Kc, 2014). This study contributes to the current body of literature by investigating the potential of introducing a case manager system in an ED as a way to improve physician productivity, and consequently ED performance. A discrete-event simulation model, based on a real-life case study, is used to examine the case manager system and to extend previous findings in a more complex and realistic setting.

In the study of Campello et al. (2017), queueing theory is used to investigate the case manager approach. Queueing theory is a popular method for modelling and analysing patient flow, particularly because of its simplicity and efficiency. However, as such analytical models mostly rely on closed-form mathematical formulations, they are not suitable to model the complex, stochastic and dynamic nature of healthcare systems unless introducing simplifying assumptions (Bhattacharjee and Ray, 2014; Saghafian et al., 2015). Therefore, when patient flows are highly complex, which is the case in EDs, simulation may be a suitable alternative. The main advantage of simulation over analytical modelling techniques such as queueing theory is that a great level of detail can be taken into account. In addition, stochastic and time-dependent characteristics of an ED can be included in the model. These characteristics of simulation enable to model the ED without extensively simplifying real-life operations by use of assumptions (Vanbrabant et al., 2019a). Consequently, our simulation model does not require many of the assumptions made in the queueing model of Campello et al. (2017) to simplify actual ED operations. As a result, our study provides a more realistic evaluation of ED performance under the case manager approach. In Table 1, a comparison of our study with the queueing model of Campello et al. (2017) is provided regarding the most important assumptions.

Table 1: Comparison of assumptions in queueing theory model of Campello et al. (2017) and our simulation model

	Campello et al. (2017)	**Our study**
Arrival rate	Stationary	Non-stationary, depending on the hour of the day and the day of the week
Patient types	Homogeneous	Heterogeneous
Patient priority	FIFO	FIFO and based on triage code
Physician consultation duration	Single distribution	Distribution dependent on patient type and stage of a patient in ED patient flow
External delay probability	Single probability	Probability dependent on patient type
External delay duration	Single distribution for total external delay duration	All external delay processes (i.e. examinations, waiting for results) are modelled in detail
Physician capacity	Constant capacity of three physicians	According to a realistic shift schedule
Model parameters	Based on Graff et al. (1993)	Based on analysis of real EHR-data

In addition to avoiding the use of simplifying assumptions by using simulation techniques, this paper contributes to the current body of literature on two other aspects. Firstly, all three system parameters of the case manager approach are investigated simultaneously by means of an experimental design. Campello et al. (2017) only look at the optimal caseload limit for given queueing disciplines, while Dobson et al. (2013) focus on the optimal queueing disciplines without determining the optimal caseload limit. As all three system parameters may have an impact on performance under the case manager approach, and the parameters may be interdependent, analysing them simultaneously gives a more complete view of the potential benefits of introducing a case manager approach. Secondly, the potential performance improvements that can be obtained by introducing a case manager approach are shown by use of a real-life case study, while previous studies focus on a theoretical setting only.

3 Case study and simulation model description

The emergency department of a Western European university hospital is used as basis for simulation model construction. The operations of the ED under study are comparable to many other EDs. Patient flow (e.g. Duguay and Chetouane (2007); McKay et al. (2013); Saghafian et al. (2012); Vanbrabant et al. (2019a); Yang et al. (2016)), patient arrival patterns (e.g. Carmen et al. (2015); Kuo et al. (2016); Ghanes et al. (2015)) and the type of distributions used to model interarrival times (i.e. exponential) and service times (i.e. triangular) (e.g. Duguay and Chetouane (2007); Kang et al. (2014); Zeinali et al. (2015)) are similar to other ED simulation studies. Consequently, the results of our study are not limited to this single case study and can be extended to be generally applicable to other EDs. The emergency department used as real-life case in this study is described in Section 3.1. The simulation modelling approach is presented in Section 3.2.

3.1 Emergency department description

The ED under study is confronted with crowding. The yearly number of patient visits recognised an increase of 15% throughout the last 5 years. In 2017, over 60.000 patients visited the ED. This number is expected to increase even further during the coming years. The ED consists of four zones, namely the paediatric, psychiatric, ambulant and non-ambulant zone. The paediatric and psychiatric zone are excluded from our study, as these units work completely independent from the rest of the ED, they have their own resources, only a minority of patients is assigned to these zones and they are less affected by the problem of crowding.

The ambulant and non-ambulant zone are physically separated and have their own dedicated resources. The main difference in patient flow between the ambulant and non-ambulant zone concerns the bed assignment. Patients in the non-ambulant zone occupy a bed during their complete stay in the ED. Contrarily, patients in the ambulant zone stay in the waiting room and are only assigned to a bed during physician or nurse consultations, and for laboratory examinations. After each consultation, they are rerouted to the waiting room.

Apart from bed assignment, both zones differ in terms of capacity, types of resources and patient characteristics (e.g. triage code and medical condition).

The main focus of this study is on the ambulant zone for several reasons. First of all, the maximum possible caseload is not restricted by the number of available beds. Secondly, as critical patients need immediate attention and cannot wait until a place becomes available with a physician, the patient set in the ambulant zone is more suitable to introduce this approach. The patient set in the ambulant zone consists mainly of low urgency patients, with only 5% triage code 2 patients and no triage code 1 patients (according to the ESI-triage system). Thirdly, all ambulant patients are initially treated by a trauma physician. This makes it straightforward to apply a case manager approach, as all patients are treated by the same discipline, and a single shift schedule is applicable. In contrast, the patient diagnoses in the non-ambulant zone are very diverse. As a result, internal medicine, urgency, neurology and cardiology physicians are present in this zone. In addition, four different shift schedules are applicable, with some disciplines only working at daytime.

The ambulant zone is modelled in great detail as to obtain reliable results on the impact of applying the case manager approach with limited caseloads on ED performance. Nevertheless, all non-ambulant patients and processes are included in the model as they may impact patient flow of ambulant patients. For example, these patients use the same radiology resources and observation beds. Also, physicians and nurses of the ambulant zone may sporadically assist in the non-ambulant zone, which impacts the operations of the ambulant zone.

3.2 Simulation model description

Among the simulation techniques, discrete-event simulation is the most appropriate technique to model ED operations. Discrete-event simulation enables to capture randomness and complexity of ED patient flow at the level of an individual patient (Brailsford and Hilton, 2001; Maidstone, 2012; Mohiuddin et al., 2017). The Arena simulation software is used for simulation model construction. Several complex processes and some patient flow logic are modelled by writing custom code in Visual Basic for Applications, which is embedded in Arena. In addition, key performance indicators not included in the standard Arena output are defined and calculated by use of Visual Basic for Applications.

The reliability of a discrete-event simulation model highly depends on the availability of input data on the real system. In our study, data extracted from the electronic health records of the hospital under study is used as basis for simulation model construction. The dataset consists of anonymised patient records containing personal, medical and patient flow information of all patients that visited the ED in 2016. This information contains, amongst others, symptoms, diagnosis, type of inflow, timestamps of the different processes a patient undergoes in the ED, types of examinations, outflow destination, etc.

Tables 1 to 4 in Appendix A provide an overview of the data fields in the extracted dataset.

The quality of the dataset has been analysed based on the framework and assessment techniques described in Vanbrabant et al. (2019b). Multiple quality problems were identified in the dataset, such as missing values and violations of the logical order of activities. These quality issues are taken into account when using the input data as a basis for simulation model construction, for example by removing incorrect or missing values. The use of a cleaned, high quality dataset enabled to model the complete patient flow in a detailed and realistic way, for instance by including time- and day dependent patient arrival rates, individualised patient pathways and stochastic service times. In order to obtain this information, the EHR data is analysed by use of Excel, R and the Arena input analyser. The information from the input data analysis is supplemented with observations, interviews and surveys, as the EHR data does not (correctly) capture every aspect of patient flow through the ED. An example of information that is obtained by use of surveys are process times, as not all activities are recorded in the EHRs and start timestamps of activities are lacking. The surveys provided a minimum, most likely, and maximum duration for all ED processes, which are used to estimate a triangular distribution for each service time. A detailed description of the content of the EHR dataset and how the dataset is analysed and used as input to the simulation model can be found in a technical report[2].

The remainder of this section contains a general description of the most important aspects of the patient flow through the ambulant zone as included in the simulation model. Furthermore, the verification and validation of the simulation model are briefly discussed. A more detailed description of the simulation model regarding the ambulant zone, and a description of the non-ambulant zone, can be found in the technical report.

3.2.1 Current patient flow through the ED

The developed discrete-event simulation model forms a realistic representation of the patient flow and current operations in the ED under study. The general patient flow through the ED, with a focus on the ambulant zone, is shown in Figure 1. Patients arrive in the ED by ambulance or walk-in. The arrival pattern is highly hour and day dependent. Four exponential distributions for interarrival times are discerned from the EHR data. The details of the arrival data analysis are presented in the technical report. Based on the day of the week and hour of the day, the most appropriate distribution is used to generate patient arrivals in the simulation model.

Registration and triage are executed for all arriving patients. Patient priorities for these two processes are determined based on a First-In First-Out (FIFO) principle. During the triage process, a triage nurse assesses the severity

[2] The technical report can be obtained upon request and will be made available online after acceptance of the paper for publication in order to ensure a double-blind review process.

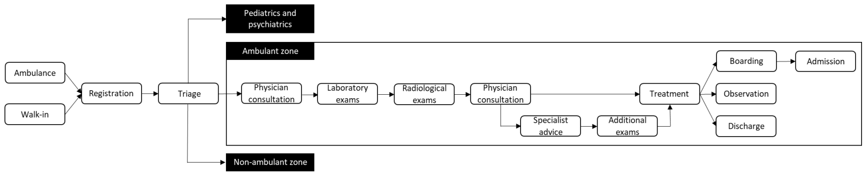


Fig. 1: Schematic overview of the general patient flow in the ambulant zone of the ED under study

of a patient's condition and assigns a triage code based on the 5-scale Emergency Severity Index (ESI). In this index, patients are divided into five groups with a triage code from 1 (most urgent) to 5 (least urgent) based on their acuity level. In addition to acuity, the triage code also determines the priority of a patient in the next ED processes.

After triage, each patient is allocated to the most appropriate zone in the ED by the triage nurse. Of all ED patient visits in 2016, 60% were assigned to the non-ambulant zone and 25% to the ambulant zone. The remaining patients are either paediatric or psychiatric patients, and are excluded from the simulation model after triage. For patients assigned to the ambulant or non-ambulant zone, a clear relationship exists between zone assignment and triage code. The higher the triage code (i.e. lower urgency), the higher the probability that the patient is treated in the ambulant zone. In the ambulant zone, 85% of the patients has triage code 4 or 5.

Once assigned to the most appropriate zone, patients follow an individualised trajectory based on their condition and individual needs. The main difference between individual patient pathways lies in the required examinations and treatments, and in the outflow process. In the ambulant zone, patients wait in the waiting room until a physician becomes available for consultation, and only occupy a box during consultations or examinations. A total of eight boxes are available in the ambulant zone. In case the needs of a patient change during their ED stay, the patient can be replaced to one of the non-ambulant beds for further treatment or observation. Patients that are referred to the ED by a general practitioner are immediately assigned to the responsible medical specialist, while patients attending the ED on their own initiative are assigned to an ED physician. A single physician type is responsible for patient consultations in the ambulant zone, namely a trauma physician. The shift schedule of the trauma physicians is presented in Figure 2.

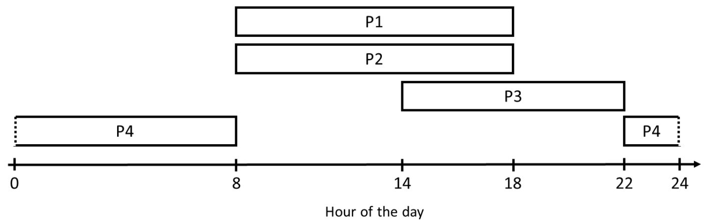


Fig. 2: Shift schedule of trauma physicians (P = physician)

After a first physician consultation, laboratory and radiological examinations may be ordered and executed, or an ECG test may be taken. In the simulation model, a specific combination of laboratory and radiological examinations is assigned to each patient. The proportion of patients that undergoes a specific examination is determined based on the EHR data. Laboratory examinations and an ECG are executed by an ED nurse, while radiological examinations are the responsibility of the in-house radiology department. Laboratory examinations mostly consist of blood and urine tests. The set of possible radiological examinations is diverse (e.g. RX, CT or Ultrasound), and multiple of them may be required for a single patient.

The results of the laboratory and/or radiological examinations are analysed by the responsible physician, after which a second physician consultation takes place. It is assumed that all patients have two consultations with an ED physician, as this is the case for the large majority of patient in the ambulant zone. Additional examinations or advice from a medical specialist may be requested before a diagnosis can be made. In case specialist advice is needed, the patient enters the consultation-examination cycle again. Based on the diagnosis, a treatment can be prescribed and/or initiated within the ED (e.g. medication, plaster, etc.).

Finally, a disposition decision is made and the patient enters the outflow process. At this moment, a patient is medically finished in the ED. Patients may be either admitted to an inpatient unit of the hospital, or discharged home. The disposition type is related to the triage code of a patient, as this reflects the acuity level and, consequently, the possibility of a patient to leave the hospital without a health risk. The higher the triage code (i.e. lower urgency), the higher the probability that a patient is discharged home and vice-versa. A triage code 1 patient has a 90% probability to be admitted to the hospital, while a triage code 5 patient is only admitted in 9% of the cases.

The outflow process depends on the disposition decision. For discharged patients, only administrative tasks should be executed before a patient can leave the ED. Admitted patients have to wait in the ED until a bed becomes available on an inpatient unit. This waiting time is defined as boarding time, and is mainly determined by the inpatient units. The inpatient units are not included in the simulation model, but based on the data analysis a clear link between the boarding time and the hour of the day at which a patient starts boarding is detected. This can be explained, for example, by the fact that most discharges on inpatient units take place around noon, or by the avoidance of new admissions on inpatient units during the night. As a result, the impact inpatient units have on boarding time is indirectly included in the simulation model by modelling the boarding time as a waiting time which depends on the time of the day.

3.2.2 Verification and validation

The simulation model is verified and validated in several ways to ensure that current ED operations are correctly represented. Firstly, verification is used to evaluate the patient flow logic in the simulation model. The process involved checking the internal logic of the individual modules and the relationship between the modules, by use of the visual animation tool in Arena and by going through the model with ED staff and operational management.

Secondly, validation is executed to ensure that the model is a close representation of the real system. Validation involved a comparison of model output against actual operations and performance by use of the available data. Both ED characteristics (e.g. patient arrival patterns, patient characteristics such as triage codes, service times and boarding time) and key performance indicators (KPIs) are compared. The comparison is executed based on mean values, distributions or boxplots, depending on the ED characteristic or performance measure to compare. For example, the length of stay of patients in the ambulant and non-ambulant zone is compared by means of boxplots and the hourly arrival patterns are compared graphically (see Appendix B). A high amount of overlap exists, so the simulation model is assumed to give a reliable representation of actual ED performance. In addition to a comparison of model output against actual ED operations by means of the available EHR-data, the validation phase also consisted of meetings with the operational management and ED staff. This is necessary, as not all ED processes and KPIs are included in the EHRs (e.g. no start times of activities, physician consultations not registered).

4 Methodology

This section outlines the research methodology used to investigate the effect of introducing a case manager approach on ED performance. Section 4.1 explains the ED operations under the case manager approach and how the simulation model of current ED operations is adjusted accordingly. The experimental design is outlined in Section 4.2. Section 4.3 describes the operational measures used to evaluate ED performance. Finally, Section 4.4 provides an overview of the statistical analyses used to evaluate the impact of the experimental design factors on ED performance under the case manager approach, and to investigate the significance of performance improvements.

4.1 Case manager approach

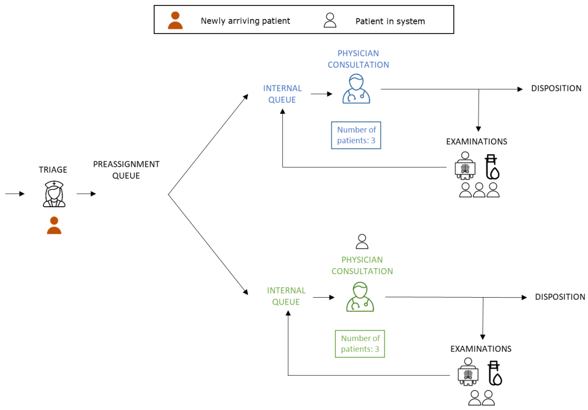
The main goal of this study is to investigate the effect of the introduction of a case manager approach on ED performance. The implementation of a case manager approach has an impact on patient flow and service times in the ambulant zone, and consequently on simulation model construction.

4.1.1 Patient flow under the case manager approach

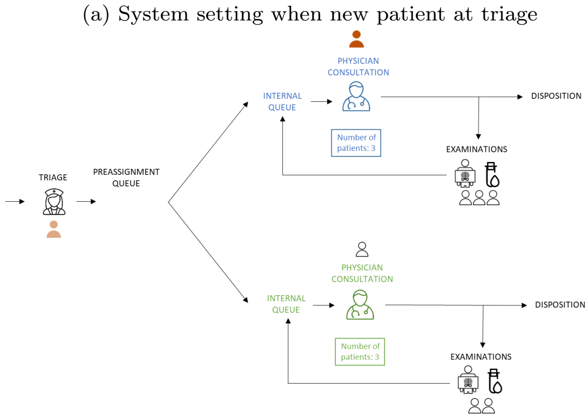
The difference in patient flow between the situation with and without a case manager approach is illustrated by use of an example in Figure 3. We suppose that seven patients and two physicians are present in the ED, and the caseload limit is set at three patients. Patient flow consists of the same processes in both systems, but the difference lies in the assignment of patients to physicians. After triage, each patient should be assigned to a physician for a first consultation. In Figure 3a, the seventh (orange) patient is undergoing triage and should still be assigned to a physician. In the situation without caseload limit (Figure 3b), physician assignment takes place whenever a physician becomes available to have a first consultation with a newly arriving patient. There is no limit on the number of patients per physician. Since the blue physician is idle, the patient at triage can be assigned to that physician, resulting in an empty pre-assignment queue. With the presence of a caseload limit (Figure 3c), physician assignment can only take place when a physician is working below the caseload limit of three patients. In the example, three patients are assigned to each physician, so the seventh patient has to wait in the pre-assignment queue until another patient is medically finished and disposed from the ED. When a place becomes available at a physician, the patient is assigned to that physician and joins the internal queue. Once assigned to a physician, a patient stays with the same physician during the complete stay in the ED. The internal queue of a physician consists of all patients waiting for the services of a single physician. The number of internal queues is equal to the number of physicians in the ED.

A first consultation with a physician is mostly followed by examinations, after which the patient again joins the internal queue. The examinations, and waiting for the results of these examinations, cause an external delay between consecutive physician consultations. In case of a second consultation, this is preceded by a setup time which is modelled as indicated in Section 4.1.2. In our simulation model, we assume that all patients have two consultations with a physician, as this is true for the large majority of patients in the ambulant zone. The duration of the physician consultations and external delays, as well as the proportion of patients undergoing specific examinations, and the proportion of patients per disposition type, are determined based on our input data analysis.

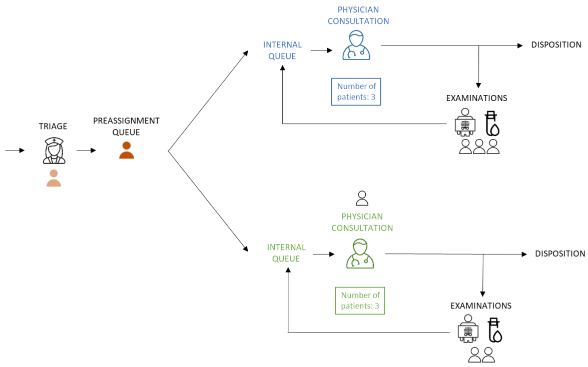
The case manager approach is incorporated in the simulation model such that the current ED setting (referred to as base case setting) is a specific parameter setting of the case manager approach. The number of physicians is determined by the current shift schedule (Figure 2). Each physician has a



(a) System setting when new patient at triage



(b) System setting after triage without caseload limit



(c) System setting after triage with caseload limit

Fig. 3: Comparison of ED patient flow with and without caseload limit

capacity equal to the caseload. In the base case, a caseload of 25 patients is used as the model output in this situation corresponds with the absence of a caseload limit (i.e. no pre-assignment queue).

4.1.2 Multitasking effect on physician service times

The introduction of a case manager approach affects the level of physician multitasking. The literature review provided evidence for a relationship between the level of multitasking and productivity. In our ED setting this implies that physician productivity is related to the number of patients simultaneously assigned to a single physician, which is limited by the caseload. Multiple factors contribute to this relationship, both in a positive and negative way. The two most prominent determinants of physician productivity in case of multitasking are a reduction in idle time and an increased setup time before each consultation as a result of patient switching (Kc, 2014). Other factors are not included in our simulation model as we assume that quality of care is preserved in the ED at all times.

The reduced idle time is not explicitly modelled in our simulation model, as it automatically results from increasing caseload limits in the simulation model. However, the effect of caseload on consultation setup times should be determined in advance, as service times are a required input to the simulation model. In the simulation model, the actual number of patients per physician at the time the consultation setup takes place is used as a proxy for the level of multitasking. This number may differ from the caseload limit, which indicates the maximum number of patients per physician.

Except from the probable non-linear nature, the specific relationship between the level of multitasking and task setup times is not clearly defined in literature (Delasay et al., 2015). In our study, a sigmoid relationship is assumed as we expect the consultation setup time to increase slowly at lower numbers of patients, since cognitive limitations have no great impact in case a physician is responsible for two patients instead of one patient, for example. From a certain number of patients onwards, the consultation setup time increases sharply until the cognitive limits of a physician are reached and consultation setup implies reading through the complete patient files again.

The sigmoid relationship can be represented by the function $S(X) = \frac{c}{d+e^{-fX+b}} + a$, with $S(X)$ the expected consultation setup time as a function of the actual number of patients per physician, X. The values of the parameters of the $S(X)$ function are uncertain. Figure 4 represents four potential sigmoid functions, obtained with different parameter settings for the function $S(X)$. The minimum and maximum value of the consultation setup time are held constant throughout the different parameter settings, but the evolution of the curve is changed. The minimum value for the mean consultation setup time appears for an actual caseload of only one patient, and is set equal to the minimum consultation setup time indicated by ED staff. The maximum value is determined in a similar way, but the number of patients for which the maximum value is reached depends on the parameters of the sigmoid function.

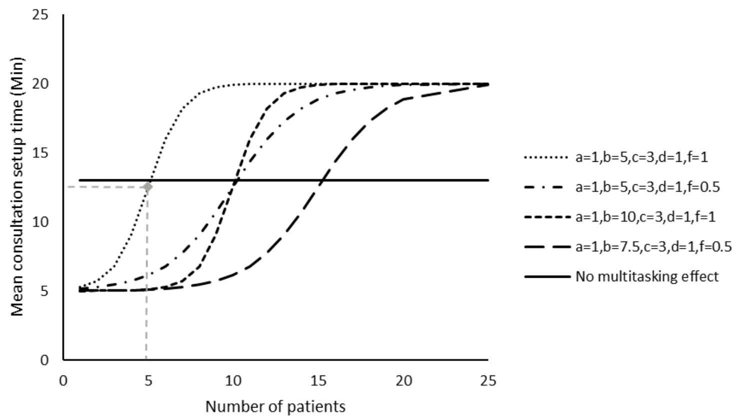


Fig. 4: Examples of sigmoid curves representing the multitasking effect on mean consultation setup time

In the simulation model, the consultation setup time for an actual number of patients X is determined based on the sigmoid function $S(X)$. This value is used as most likely value in a triangular distribution. The width of the interval between the minimum and maximum possible value in the triangular distribution is assumed to be independent from the actual number of patients per physician, and set equal to 4 minutes. As a result, the consultation setup time has a distribution TRIA($S(X)-2$,$S(X)$,$S(X)+2$) in the simulation model. Each time a patient enters the setup process, the value of $S(X)$ is recalculated in the simulation model. For example, when looking at the first curve on the graph, in case a physician is responsible for five patients the most likely setup time according to the sigmoid function is 12.5 minutes. This implies that the triangular distribution which is used to determine the setup time in the simulation model is TRIA(10.5,12.5,14.5).

The results of introducing a case manager approach may be impacted by the relationship between multitasking and consultation setup times. As a result, the experimental design is executed for two different scenarios with regard to the sigmoid function. The first scenario represents the absence of a multitasking effect on consultation setup times. The consultation setup time is independent from the number of patients per physician. In the second scenario, the consultation setup time is determined by the number of patients per physician according to the formula: $S(X) = \frac{15}{1+e^{-X+5}} + 5$ (i.e. the first curve on Fig 4 with as parameters a = 1, b = 5, c = 3, d = 1 and f = 1). This parameter setting of the sigmoid curve is used because of the early and sharp increase, which is expected to have the highest impact on ED performance of the curves in Figure 4. The two scenarios represent the two most extreme cases of a multitasking effect. The scenarios will determine a minimum and maximum effect size of introducing a case manager approach and form the boundaries of an interval that contains the actual effect size. Figure 5 plots the relationship between the number of patients per physician and the consultation setup time for the two scenarios. The consultation setup time in case of no multitasking effect is distributed as TRIA(11,13,15) minutes. This distribution is determined based on an estimate of the mean consultation setup time by ED staff. In the multitasking case the distribution is TRIA($\frac{15}{1+e^{-X+5}}+3, \frac{15}{1+e^{-X+5}}+5, \frac{15}{1+e^{-X+5}}+7$).

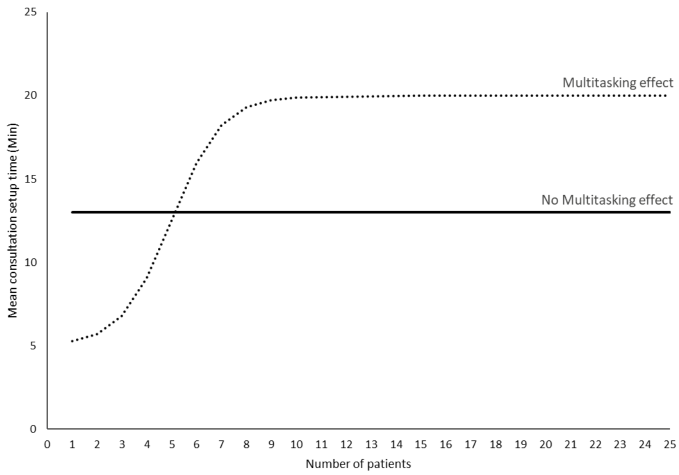


Fig. 5: Sigmoid functions of the two multitasking scenarios

4.2 Experimental design

The case manager approach is characterised by three parameters: caseload limit, pre-assignment queueing discipline and internal queueing discipline. In addition to the caseload limit, queueing priorities may have a major impact on ED performance. Under current ED operations, physicians prioritise patients based on their triage code. However, for low caseload limits this may result in very long waiting times for low urgency patients. In order to evaluate the effect of these parameters on ED performance, an experimental design has been developed. Table 2 provides an overview of the factors and accompanying factor levels.

The first factor, **priority** Q_P, determines the priority in the pre-assignment queue. The two factor levels are FIFO and priority based on triage codes (TC). The second factor, **priority** Q_I, represents the queueing discipline for patients in the internal queue of a single physician, with FIFO and TC as the two factor levels. An additional factor to determine priority in the internal queue, is **priority stage**. This factor is used to differentiate between patients in the

Table 2: Experimental design

Factor	Factor levels
Priority Q_P	(1) FIFO
	(2) TC
Priority Q_I	(1) FIFO
	(2) TC
Priority Stage	(1) Equal
	(2) Stage 1
	(3) Stage 2
Caseload limit	1, 2, ..., 25

same internal queue based on the stage in their ED stay. The factor has three levels, enabling to give priority to patients waiting for their first physician consultation, patients waiting for their second consultation, or equal priority for both consultations. Patients in the same stage are prioritised based on the value of the factor priority Q_I. The fourth factor is the **caseload limit**, which determines the maximum number of patients per physician at a single moment in time. The caseload limit ranges from 1 to 25. The value of 25 represents the situation in which each patient is immediately assigned to a physician, so it corresponds to the absence of a binding restriction on the number of patients per physician (i.e. current situation), as it is very unlikely that more than 25 patients per physician are present simultaneously in the ED.

4.3 Operational measures

The performance measures that are used to evaluate the case manager approach are DTDT and LOS. DTDT is relevant because the time until the first consultation with a physician is the most critical period of a patient's stay in the ED, and will be directly impacted by the introduction of a caseload limit. LOS includes the total time a patient is assigned to a physician, which consists of all waiting times for a physician, consultation times, consultation setup time, and the time of examinations. LOS is included as a KPI because the case manager approach may have a contradictory effect on DTDT and LOS. Only focusing on DTDT may lead to a high LOS, as the second consultation with a physician will be neglected by always prioritising the first consultation.

As the patient mix in the ED is not evenly distributed over all different triage codes, and queueing disciplines may be based on triage codes, the effect of a case manager approach on DTDT and LOS may differ for the different patient types. In addition, the importance that patients attach to either DTDT or LOS may differ based on triage codes. Because of this, both KPIs are measured separately for each patient type (i.e. triage code), leading to eight KPIs in total. This may provide relevant information about when a case manager approach is beneficial (i.e. for which patient types).

4.4 Statistical analysis

The simulation model should be initialised in order to obtain statistically reliable results from our simulation analyses for all eight KPIs. Each simulation run consists of a warm-up period of 7 days, followed by an actual run length of 28 days for which output statistics are collected. A total of 100 replications is executed for each factor combination in the experimental design. More information on the specification of the run parameters can be found in the technical report. In addition to specifying appropriate run parameters, common random numbers are applied in the simulation model to improve the validity and reliability of results. Common random numbers are used such that all variability in simulation output between different experimental design settings can be attributed to changes in the system setting (i.e. experimental design factors) instead of randomness (Kelton et al., 2015).

The experimental design consists of a full factorial design with only repeated-measures factors, corresponding to priority Q_P, priority Q_I, priority stage and caseload. All 300 factor combinations of the experimental design are evaluated by use of the simulation model. Each factor combination is tested in the same ED setting, as common random numbers are used. As a result, a repeated-measures full factorial analysis of variance (ANOVA) is executed on the simulation results. An ANOVA is performed to test whether or not a relationship between one of the factors and the KPIs is statistically significant. In addition, interactions between experimental design factors can be identified. A repeated-measures full factorial ANOVA is a specific version of the ANOVA-test which is adjusted to the fact that all factor combinations of the experimental design are tested in the same setting, for example by use of common random numbers (Field, 2013). A total of eight repeated-measure full factorial ANOVAs, one for each KPI, is executed on the simulation results.

A first important assumption of a repeated-measures full factorial ANOVA in order to ensure accuracy of the F-statistic, is sphericity (i.e. equality of variances of the differences in output between factor combinations for a single ED setting). The assumption of sphericity is tested with Mauchly's test. A violation of the sphericity assumption results in an increased Type I error rate in the statistical analysis. This involves that the probability of finding a significant effect in the ANOVA while there is no effect in reality, increases. Therefore, in case the sphericity assumption is violated based on the results of Mauchly's test, the conservative Greenhouse-Geisser (G-G) estimate of the F-statistic is used. The G-G estimate adjusts the degrees of freedom to compensate for the violation of the sphericity assumption (i.e. increased Type I error rate) (Field, 2013). Otherwise, the sphericity assumed estimate of the F-statistic can be used. A second assumption is normality of the dependent variable (i.e. KPI). When the degrees of freedom are sufficiently large (at least 20) and group sizes are equal, the F-statistic controls the Type I error rate well under conditions of non-normality (Field, 2013). As the simulation is replicated 100 times for each factor combination, the degrees of freedom are sufficiently high for the normality condition to be fulfilled. In addition, all factor combinations are tested on the same set of patients by use of common random numbers in the simulation model, which makes the experimental design balanced. This implies that the F-statistic is a reliable estimate in our setting, making robust checks (e.g. bootstrapping) unnecessary.

The repeated-measures full factorial ANOVA is used to determine whether a factor, or an interaction between factors, has a significant effect on the outcome of the case manager approach. In case a factor has a significant impact, post hoc tests may provide insights in the specific relationship between a factor and the different KPIs. As the focus is on investigating the impact of a caseload limit on ED performance, given a specific priority setting (i.e. a possible combination of the priority Q_P, priority Q_I, and priority stage factor levels), post hoc tests are performed for the caseload factor. More specifically, a univariate ANOVA is executed to investigate the caseload factor in more detail. A total of 12 priority settings exist, each with a different combination of the factor levels of the three priority factors. For each of these priority settings, a pairwise comparison of all 25 factor levels of the caseload factor is executed with regard to their impact on the KPI under investigation. This way, caseload limits which result in a significant improvement in ED performance compared to the current ED operations (i.e. caseload 25) can be determined for each priority setting. The analysis is executed for all eight KPIs. The Bonferroni correction of the significance level, and consequently the p-value, is used in the univariate ANOVA to ensure the overall Type I error rate across all comparisons remains at 0.05. When evaluating multiple hypotheses based on a single dataset of experiments, the Bonferroni correction approach makes the results robust in terms of power and control of the Type I error rate (Field, 2013).

The results of the statistical analyses (i.e. Mauchly's test and the repeated-measures full factorial ANOVA) are provided in Appendix C.

5 Results and discussion

The results show that the case manager approach may improve ED performance in terms of both DTDT and LOS. However, the effect size depends on the multitasking scenario, patient type (i.e. triage code) and experimental design factors. In Section 5.1 the impact of the multitasking scenario on the effectiveness of a case manager approach is investigated. The difference in effect size based on the KPI and patient type under investigation is discussed in Section 5.2. Finally, the effect of the experimental design factors on the outcome of implementing a case manager approach is described and statistically analysed in Section 5.3. The results of applying the case manager approach in the case study setting are presented in Section 5.4.

5.1 Multitasking scenario

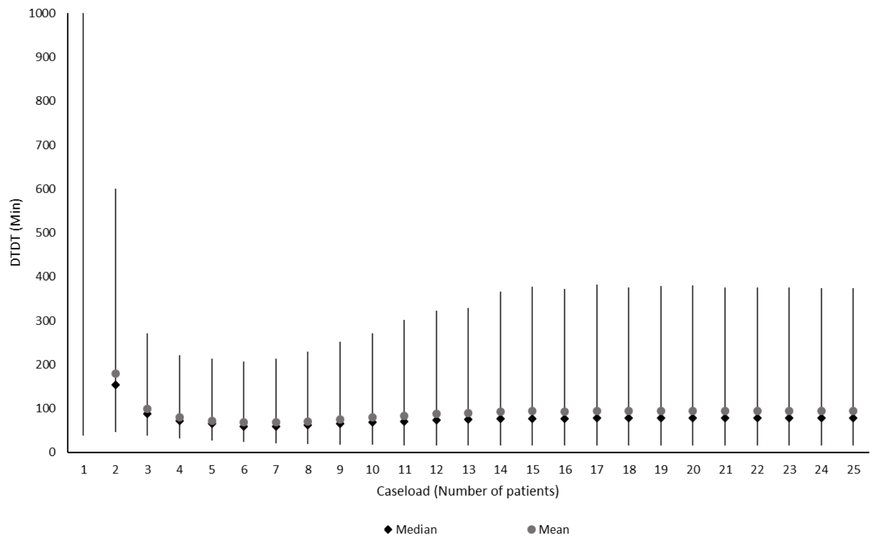
The effect of introducing a case manager approach is evaluated in a scenario with and without the presence of a multitasking effect on consultation setup times (see Section 4.1.2). Figures 6 and 7 show the minimum, maximum, median and mean DTDT and LOS, respectively, as a function of caseload for both multitasking scenarios. The minimum, maximum, median and mean are calculated over all patient types and priority settings. From these graphs, it can be concluded that the introduction of a caseload limit, on average, has a positive effect on both DTDT and LOS. In both scenarios, the mean DTDT and LOS can be reduced compared to a caseload limit of 25 (i.e. absence of a caseload limit). This positive effect can be seen on the graphs as the reduction in DTDT (Figure 6) or LOS (Figure 7) of introducing a strict caseload limit compared to a caseload of 25. The extent of the effect depends on the multitasking scenario. In the scenario with multitasking effect, the mean DTDT can be reduced by 34.87%, while LOS may be reduced by 44.96%. In the scenario without multitasking effect, mean DTDT and LOS can be reduced by 19.01% and 26.89%, respectively.

The results indicate that a caseload limit may positively impact ED patient flow, even without the presence of a multitasking effect. This implies that the positive effect is not only related to the reduced consultation setup time due to physician multitasking and the accompanying assumptions we make with regard to the relationship between multitasking and consultation setup times. There are several possible explanations for this finding. Firstly, a caseload limit not only reduces the number of patients per physician, but also the number of patients in other ED processes. In addition to a reduced waiting time for physicians, this may result in shorter waiting times for ED processes such as bed assignment and radiological or laboratory examinations. Secondly, workload may be divided more evenly over the different physicians by introducing a caseload limit. Not every patient entails the same workload. By assigning all arriving patients immediately to a physician, it might happen that one physician gets idle while there is a significant waiting time for another physician as a result of a difference in workload between their set of patients. As a result, patients queue for a physician consultation while there is capacity available at another physician. A case manager approach overcomes this situation with the formation of a pre-assignment queue. Once a patient is medically finished, a new patient can be assigned to a physician.

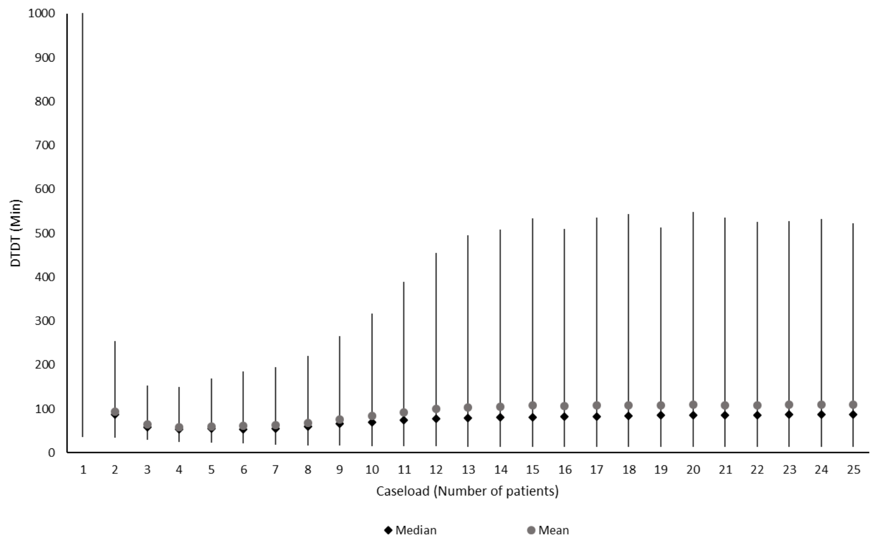
In reality, we expect a multitasking effect to be present, but the shape of the sigmoid curve is unknown. Given the higher impact of a case manager approach in case of a multitasking effect, and as this scenario seems more realistic, the next subsections provide a detailed discussion of the results of the scenario with multitasking effect. The results of the scenario without multitasking effect are included in Appendix D. These results mainly show the same pattern with respect to a positive or negative effect, but the effect size is less pronounced.

5.2 Patient type

Previous findings indicate that the DTDT and LOS over all patient types are positively impacted by the introduction of a caseload limit. However, ED

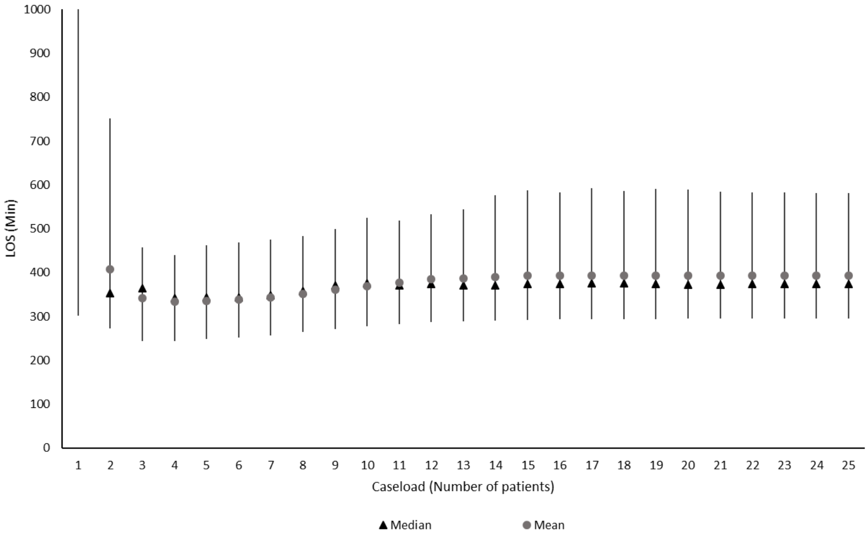


(a) No Multitasking effect

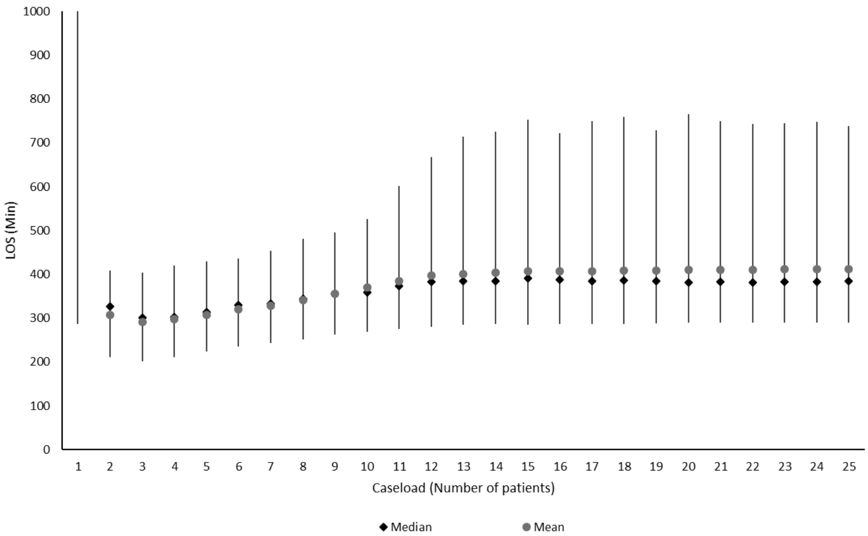


(b) Multitasking effect

Fig. 6: Minimum, median, mean and maximum DTDT over all priority settings and patient types as a function of caseload.



(a) No Multitasking effect



(b) Multitasking effect

Fig. 7: Minimum, median, mean and maximum LOS over all priority settings and patient types as a function of caseload.

patients can be differentiated based on their triage code. Figures 8 and 9 present the mean DTDT and LOS, respectively, per patient type as a function of the caseload limit. The mean values in the graphs are calculated over all priority settings. All patient types benefit from introducing a caseload limit, in terms of both DTDT and LOS. Nevertheless, the effect of a caseload limit on ED performance differs based on the patient type. The potential reduction in DTDT is similar for triage code 2, 3 and 4 patients with an average reduction around 40%. For triage code 5 patients, the average reduction lies even 10% higher. With regard to LOS, a 40% reduction can be obtained for triage code 4 and 5 patients, which is twice the effect size of triage code 2 and 3 patients.

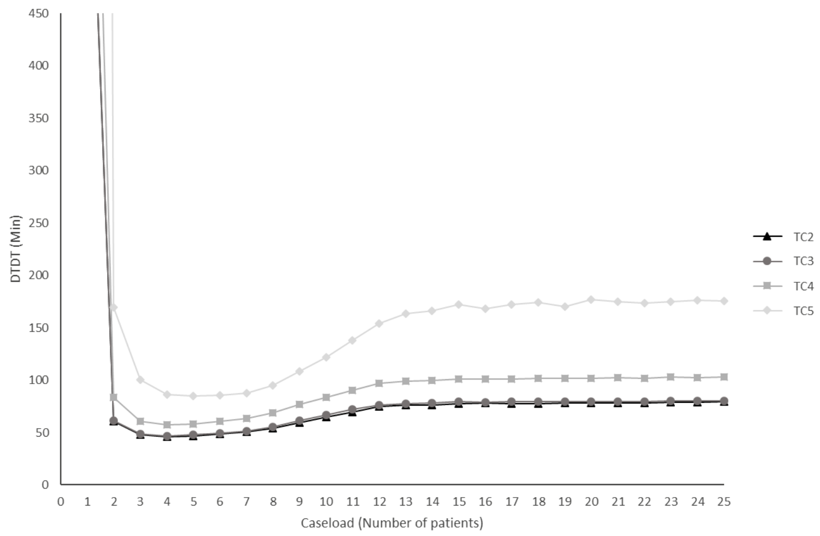


Fig. 8: Mean DTDT per TC as a function of caseload. Note: DTDT at caseload 1 is very high because of the large amount of physician idle time, and will never be used in practice. These values are not presented in the figures for clarity purposes. The DTDT at caseload 1 equals (in minutes): TC2: 707.70, TC3: 712.62, TC4: 983.85, TC5: 6185.81.

5.3 Experimental design factors

The experimental design consists of four factors, namely priority Q_P, priority Q_I, priority stage and caseload limit. An analysis of variance (ANOVA) is performed for all eight KPIs in order to test whether or not a statistically significant relationship exists between the four experimental design factors and the eight KPIs. The appropriate F-statistics of the repeated-measures full

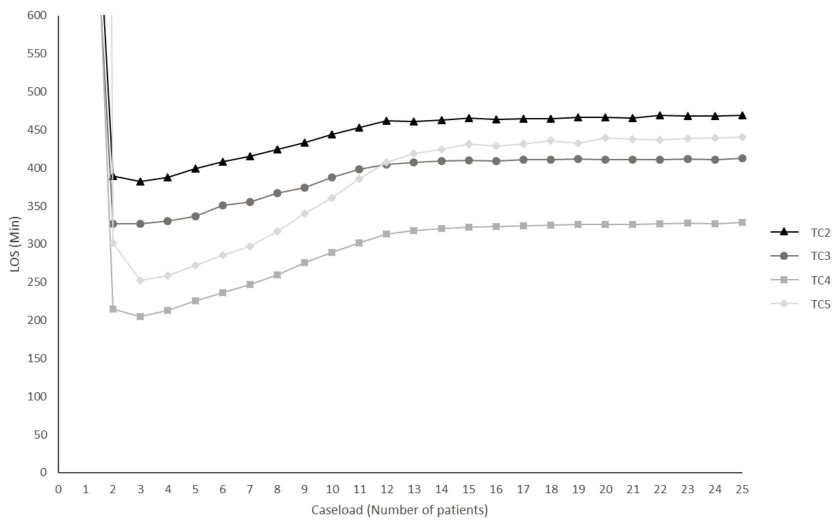


Fig. 9: Mean LOS per TC as a function of caseload. Note: LOS at caseload 1 is very high because of the large amount of physician idle time, and will never be used in practice. These values are not presented in the figures for clarity purposes. The LOS at caseload 1 equals (in minutes): TC2: 1032.14, TC3: 962.66, TC4: 1102.88, TC5: 6311.80.

factorial ANOVA are provided in Appendix C. The main focus of this paper is on the effect of a caseload limit on DTDT and LOS. For all eight KPIs, the repeated-measures full factorial ANOVA shows that the caseload factor and all 2-way interactions including the caseload factor are significant. As a result, the effect of introducing a caseload limit on ED performance is impacted by the queuing disciplines that are applicable in the ED. The three priority factors have an impact on the pattern of the relationship between the caseload limit and both DTDT and LOS, the effect size (i.e. potential reduction in the KPI), and the optimal caseload limit.

In order to determine if the introduction of a caseload limit is beneficial for an ED as a way to improve performance, the 12 individual priority settings are investigated. The combination of a priority setting and a specific caseload limit can either significantly increase, significantly decrease or not significantly impact ED performance for a specific KPI. For each priority setting, the caseload ranges for which each KPI is significantly increased or decreased at the 5% significance level compared to the situation without a caseload limit (i.e. caseload = 25), are determined by use of the post hoc tests in the univariate ANOVA (see Section 4.4). Figure 10 shows these caseload ranges for the DTDT of all four patient types. Figure 11 presents the same information with regard to LOS. The dark grey bars indicate for each priority setting the caseload ranges that result in a significant increase in DTDT or LOS. The caseload ranges for which a KPI is significantly decreased are indicated by the light grey bars. In case no significant increase or decrease in DTDT or LOS is found for a specific priority setting and caseload limit, this implies that ED performance is not significantly changed compared to the situation without a caseload limit. This is the case for all case manager settings for which no light or dark grey bar is present on Figures 10 and 11.

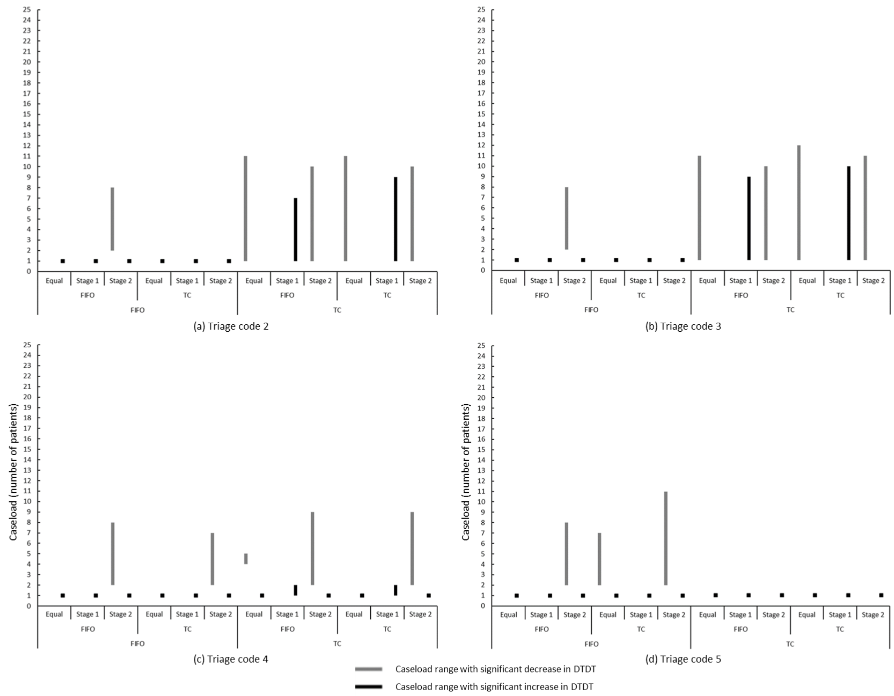


Fig. 10: Caseload range resulting in significant DTDT improvement in comparison with no caseload limit for each priority factor combination (per triage code)

The results indicate that not all KPIs can be improved significantly for all priority settings by introducing a caseload limit. Nevertheless, an improvement of ED performance is a more frequent outcome of introducing a case manager approach than a significant deterioration of ED performance. The priority settings for which a caseload limit significantly improves ED performance differ based on the KPI under investigation. The most important findings regarding the impact of the priority factors on the results of introducing a caseload limit are described below.

A first main finding deals with the priority stage factor. When patients waiting for the first physician consultation get priority (i.e. stage 1), DTDT

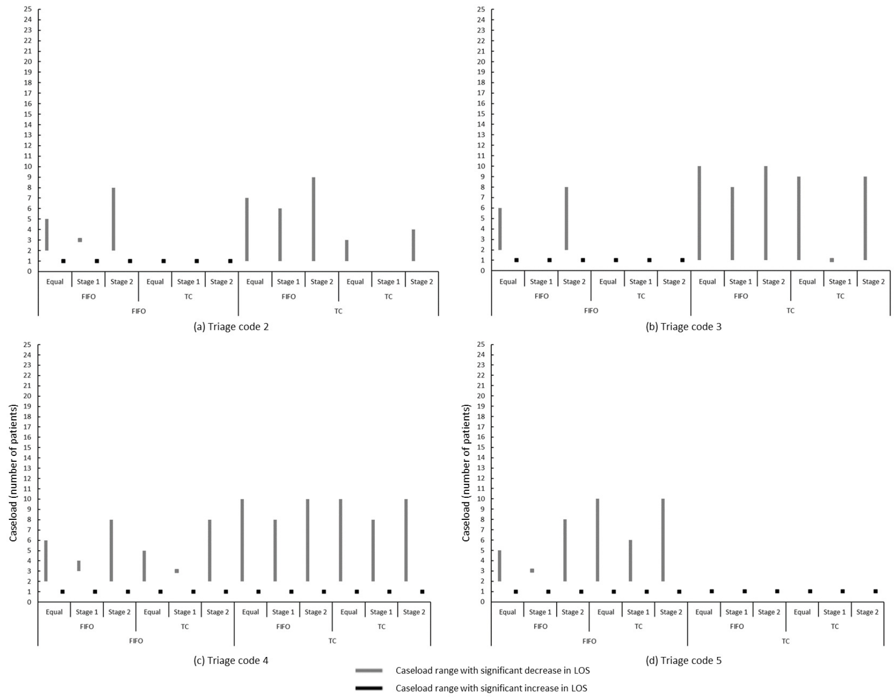


Fig. 11: Caseload range resulting in significant LOS improvement in comparison with no caseload limit for each priority factor combination (per triage code)

cannot be improved by introducing a caseload limit, independent of the patient type. As the absence of a caseload limit involves that all newly arriving patients are immediately assigned to a physician, and these patients get priority over patients further in their ED stay, DTDT cannot be improved by limiting the number of patients per physician. However, LOS can be optimised by introducing a caseload limit, even in situations where a stage 1 priority is applicable. This is explained by the fact that a stage 1 priority without caseload limit may result in a high LOS as patients have to wait for their second consultation with a physician until no newly arriving patients are awaiting their first consultation.

A second finding concerns the pre-assignment queue priority. The DTDT of high urgency patients (i.e. low triage codes) can be mainly reduced by introducing a caseload limit when the pre-assignment queue priority is based on TC. Lower urgency patients, on the other hand, benefit from a caseload limit when the pre-assignment queue priority is FIFO. This can be explained by the fact that a caseload limit is especially advantageous for patients with a high priority in the pre-assignment queue, so high urgency patients in case of a TC queueing discipline, as they get priority once physician capacity becomes available. For low urgency patients, which are overtaken by newly arriving patients with a higher urgency, a pre-assignment queue priority based on TC in combination with a caseload limit may extend waiting times in the pre-assignment queue. As DTDT is the time until the first consultation, this is not beneficial for the DTDT of these patients. The previous findings with regard to DTDT are also applicable for LOS, but the introduction of a caseload limit can significantly reduce LOS in more priority settings. LOS consists of more time intervals than DTDT. By taking the waiting (and process) times of patients after their first physician consultation into account, which can be reduced by limiting the number of patients per physician, LOS can be reduced for priority settings for which DTDT is not significantly improved by introducing a caseload limit.

A third finding deals with the caseload ranges that significantly deteriorate ED performance. The most important conclusion is that a caseload limit of one patient is detrimental for ED performance in almost all priority settings. Furthermore, we find that in most situations where a caseload range exists that significantly increases DTDT or LOS, this range only contains a caseload limit of one patient. Only when a pre-assignment queue priority based on TC is combined with a stage 1 priority, the DTDT of triage code 2, 3 and 4 patient is significantly deteriorated for a larger caseload range. An explanation is that this priority setting benefits patients with a higher urgency that are waiting for a first physician consultation. Without a caseload limit all arriving patients are immediately assigned to a physician and patients waiting for their first consultation get priority over patients further in their ED stay. With a caseload limit, there might be a high pre-assignment wait as patients have to wait until a place becomes available with a physician, which increases DTDT. As a result, DTDT of high urgency patients is optimised without a caseload limit and significantly deteriorates when a low caseload limit is introduced for this priority setting.

A final interesting finding is found when combining the insights of all eight graphs in Figures 10 and 11. The only priority setting for which all eight KPIs can be significantly improved by introducing a caseload limit, is FIFO - FIFO - Stage 2. Nevertheless, this does not imply that the case manager approach is always worse than a situation without caseload limit for all other priority settings. A case manager setting (i.e. combination of a priority setting and caseload limit) without a significant increase or decrease for a specific KPI implies that ED performance remains unchanged with regard to that KPI. As a result, multiple priority settings exist for which a specific caseload limit may result in a significant improvement of some KPIs, and an unchanged value for all other KPIs in comparison with the situation without caseload limit. Depending on the importance attached to each KPI by the decision maker, situations in which an important KPI is highly improved, while a less important KPI stays unchanged, might be an interesting option for implementation in an ED. The only caseload limit that will always be detrimental for at least one KPI in each priority setting, is one patient. As a result, this caseload limit is never optimal when trying to improve overall ED performance.

5.4 Case study setting

The priority setting has a major impact on ED performance, and on the potential benefits of introducing a case manager approach. In order to determine if the introduction of a caseload limit is beneficial in the ED under study, the minimum obtainable DTDT and LOS for the current priority setting in the ED (i.e. TC, TC, equal) are investigated. Table 3 provides an overview of the minimum obtainable value for each KPI, and the corresponding caseload for which this minimum value is realised. The DTDT and LOS of triage code 2 and 3 patients, and the LOS of triage code 4 patients can be significantly improved by introducing a caseload limit. For the other KPIs, no significant improvement can be obtained by introducing a caseload limit. The DTDT of triage code 4 and 5 patients, and the LOS of triage code 5 patients, are significantly worse than the current value when the caseload limit is set at one patient.

Table 3: Significant potential KPI improvements under the current queueing disciplines (TC-TC-Equal) with corresponding caseload limit - Multitasking effect

KPI	Current value (Min)	Minimum value (Min)	Caseload limit (Min)	% improvement
DTDT - TC2	63.94	34.37	2	46.25%
DTDT - TC3	64.69	35.99	2	44.37%
DTDT - TC4	-	-	-	-
DTDT - TC5	-	-	-	-
LOS - TC2	428.71	350.28	1	18.29%
LOS - TC3	383.81	285.97	1	25.49%
LOS - TC4	315.57	204.04	3	35.34%
LOS - TC5	-	-	-	-

As indicated in the previous section, for each priority setting there might exist a caseload limit for which some KPIs are significantly improved, other KPIs remain unchanged and no KPIs are significantly deteriorated. These caseload limits are interesting to implement in practice. In the case study setting, a caseload limit that lies between 2 and 12 patients per physician always significantly improves at least one KPI while the other KPIs remain unchanged. When looking at the number of KPIs that significantly improves and the size of these improvements, a caseload limit of three patients is the most interesting. A caseload of three patients leads to a significant improvement in the DTDT of triage code 2 and 3 patients, and in the LOS of triage code 2, 3 and 4 patients. These significant improvements are presented in Table 4. The size of the improvements ranges from 15% up to even 42%. The other KPIs improved on average, but the improvements are not significant on the 5% significance level.

Table 4: Significant KPI improvements under the current queueing disciplines (TC-TC-Equal) for caseload limit 3 - Multitasking effect

KPI	Current value (Min)	Value at caseload 3 (Min)	% improvement
DTDT - TC2	63.94	36.99	42.15%
DTDT - TC3	64.69	38.72	40.15%
DTDT - TC4	-	-	-
DTDT - TC5	-	-	-
LOS - TC2	428.71	362.33	15.48%
LOS - TC3	383.81	311.12	18.94%
LOS - TC4	315.57	204.04	35.34%
LOS - TC5	-	-	-

The results show that a case manager approach is beneficial in a complex service system such as an ED. The results are based on a simulation study that takes real-life characteristics of the ED into account. In the ED under study, improvements of up to 40% can be obtained for some KPIs by only introducing a caseload limit without changing the current priority setting. These results are found in case a multitasking effect on consultation setup times is present. Nevertheless, when looking at the scenario without multitasking effect, a caseload limit between 2 and 8 patients can significantly improve ED performance in the case study for at least one KPI while not significantly impacting the other KPIs. Table 2 in Appendix D shows the significant performance improvements when a caseload limit of two patients is introduced in the scenario without multitasking effect. This is the caseload limit with the highest number of significant KPI improvements. Only three KPIs can be improved simultaneously and the size of the improvements is lower, but still reaches up to 17%.

6 Practical considerations

This is a first study to investigate the potential of a case manager approach to improve ED performance in a complex and realistic setting by use of discrete-event simulation. The results indicate that the introduction of a caseload limit may significantly improve ED performance. As the results are promising, future research may focus on how to implement the case manager approach in practice. In order to fully exploit the benefits, some practical considerations for both physicians and patients should be taken into account when implementing the approach in an ED.

The implementation of a case manager approach has a direct impact on the way ED physicians work. From an operational point of view, their work will be organised more efficiently. However, two psychological aspects should be taken into account. Firstly, physicians are used to working at utilisation rates near 100 %, and are frequently paid in proportion to the number of patients treated. Nevertheless, utilisation rates near 100% are not desirable from an operational point of view as they undermine the ability to provide quick and qualitative care. Placing a limit on the number of patients per physician in order to limit the amount of multitasking may imply a reduction of their utilisation, but improves ED performance. However, it requires a change in the physician mindset. Secondly, not all patients involve an equal workload. When assigning patients to physicians, this can be taken into account to ensure that all physicians have a comparable patient mix and workload. The optimal number of patients per physician may depend on the workload of the assigned patients. Besides, physicians may prefer certain types of patients over others in terms of complexity, medical condition, etc. As a result, appropriate patient assignment rules should be defined in order to ensure fairness.

Patients are also impacted by the introduction of a case manager approach. As the results show, their stay in the ED can be shortened significantly. However, as with physicians, two psychological aspects should be taken into account. First of all, patients may get the idea that physicians are not working at their full capacity. Nevertheless, minimising DTDT and LOS requires physicians to work at high utilisation rates, so we expect this problem to be minimal. Also, the observed service rates are higher as physician productivity is improved. In addition, quality of care may be improved by limiting physician workload to manageable levels. A second, and probably more important, consideration is the patient perception of waiting. Patients tend to attach more importance to the waiting time before a first consultation with a physician. After a first consultation with a physician, they feel to be taken care of and are more tolerant towards waiting times. This implies that a reduction in the pre-assignment waiting time increases patient satisfaction more than a reduction in the internal waiting time, independent of the total waiting time.

Besides psychological aspects related to physicians and patients, some practical implications on ED operations of implementing the case manager approach should be investigated. Introducing case managers in the context of an ED implies the alignment of this approach with the complex ED setting. First, as shift work is the primary employment model of ED physicians, patient handovers are necessary at shift changes. Furthermore, physician capacity may differ between shifts. Consequently, a handover policy should be determined, especially when physician capacity decreases (e.g. a systematic decrease in caseload when approaching shift changes, or a temporary increase in caseload limits after a shift change). In our simulation model, physicians are supposed to complete their current task at the moment of a shift change. All patients assigned to the physician whose shift ends are assigned to a new physician. In case of a reduction in the number of physicians, patients that were assigned to a physician get priority over newly arriving patients once a place becomes available. Second, as a high urgency patient may enter the ED at all times, a policy to deal with these patients should be foreseen. This may imply the introduction of spare (or reserved) capacity or the allowance of a temporary exceedance of the caseload limit. As the case manager approach is applied to the ambulant zone of the ED under study, triage code 1 patients are not present and the amount of triage code 2 patients is limited. As a result, the problem of high urgency patients that have to wait while needing immediate care is minimal in the current setting. Nevertheless, when optimising the case manager system, and especially the caseload limit, a limit on the maximum DTDT for these high urgency patients can be included in the constraints to enforce these patients to be treated immediately.

7 Conclusion and future research

Emergency departments are confronted with the problem of crowding, and physicians are one of the main bottlenecks in an ED. The introduction of a case manager approach with limited caseloads, which entails the introduction of an upper limit on the number of patients simultaneously assigned to a single physician, may improve physician productivity and ED performance. Physician multitasking results in a reduced idle time and increased patient switching costs (i.e. consultation setup times). The positive effect of a reduced idle time reduces with an increasing level of multitasking, while the negative effect of increased patient switching costs increases. As a result, an optimal number of patients per physician exists that optimises ED performance. This paper is the first to empirically investigate the potential of introducing a case manager approach with limited caseloads in a complex and realistic ED setting by use of discrete-event simulation. The simulation model is based on a real-life case study of the ED of a Western European university hospital.

Results show that the introduction of a case manager approach with limited caseloads may improve ED performance significantly and that the effect size depends on the specific case manager setting. The case manager approach is characterised by three parameters: the pre-assignment and internal queue priority, and the caseload limit. These parameters are interrelated, and the combination of the three priority parameters determines the potential performance improvement of introducing a caseload limit. Besides the system parameters, the effect size is also related to the KPI under investigation. Both DTDT and LOS can be improved for all patient types, but the extent of the possible improvement differs based on the patient type, queueing disciplines and caseload limit. In the ED of the case study, improvements of up to 40% can be reached by introducing a caseload limit, even without changing the current queueing disciplines. These insights can help ED management to improve ED performance without significant financial investments by more efficiently using the scarce and expensive ED physician capacity.

Several opportunities exist for future research. Firstly, future research may focus on determining the optimal case manager setting with regard to the priority rules and caseload limit by use of a multi-objective approach taking both KPIs for all patient types into account. Secondly, the queueing disciplines that are taken into account may be extended by investigating the accumulating priority queue. FIFO does not take the patient's condition into account, and TC systematically disadvantages low urgency patients. The accumulating priority queue is a relatively new approach, in which patients accumulate priority based on their waiting time in the ED. The rate at which their priority increases depends on their triage code. Determining optimal accumulation rates is a complex optimisation problem in itself, little research exists on this topic, and the implementation might be difficult in practice. As a result, this queueing discipline is not included in the experimental design of this initial study, which has as main goal to investigate the potential of a case manager approach with limited caseloads. Thirdly, the effect of multitasking on consultation setup times is found to have a major impact on the effectiveness and potential performance improvement of the case manager approach. As the exact relationship between the number of patients and the consultation setup times is unknown, future research may focus on determining the actual curve. This requires numerous on-field observations to determine the actual consultation setup time for all possible physician caseloads, which is time-consuming and out of the scope of this paper. Fourthly, caseload is currently measured as the number of patients per physician. Alternatively, workload can be used as not every patient involves the same workload. This is expected to enhance the results, since the assignment of patients to physicians is more balanced. Also, physicians may more easily accept a limit on the workload than on the number of patients. However, this is only possible when patient workload can be measured based on patient characteristics that are known before the first consultation with a physician (e.g. at triage). Fifthly, the case manager approach can be evaluated in other ED settings, for example in terms of bed capacity, shift types or arrival patterns. The external validity of this case study is limited as the empirical study only consists of a single ED. The insights can be tested in another ED setting to improve the generalisability. Finally, the effect of a case manager approach on other ED processes can be investigated, as interactions and trade-offs between consecutive processes may exist. An increased physician throughput may, for example, lead to an increase in the number of boarding patients, which can result in an increased length of stay for admitted patients. Based on these findings, improvement options in other ED processes that reinforce the case manager approach may be identified.

A Electronic health record data

Table 1: Timestamp attributes in extracted input file of EHRs (T = time)

Attribute	Explanation
Start date (only date)	The date a patient arrives at the ED and is first registered in the system
T Arrival	Timestamp expressing patient registration in the system
T First triage	Timestamp representing the completion of triage (the point at which a triage code is entered in the system)
T First physician	Timestamp at which the doctor starts writing a report after a first consultation with the patient
T First physical location other than waiting room	Timestamp when the patient was moved out of the waiting room to another physical location (box) for the first time

Attribute	Explanation
T Start observation	Timestamp at which the doctor decides that the patient needs to be placed in observation
T Medically finished	Timestamp at which the doctor “signs off” the patient (all medical actions are completed from the perspective of the ED)
T Mutation request	Timestamp when a bed in the hospital was requested for the patient
T Mutation plan	Timestamp when a bed in the hospital is assigned to the patient
T Departure	Timestamp when patient left the ED
T Rx request	Timestamp of the first request for a radiological examination (entered by the physician)
T Rx start execution	Timestamp when the radiological examinations are executed
T Rx first report	Timestamp of the first finished report of the radiological examinations
T Rx last report	Timestamp of the last finished report of the radiological examinations
T Lab request	Timestamp of the first request for a lab test (blood, urine, etc.)
T Lab first sample received	Timestamp when the first sample is taken for a lab test
T Lab first report	Timestamp when the first finished report was written of the lab results
T Lab last report	Timestamp when the last finished report was written of the lab results
T Pharmacy first use	Timestamp when something was taken from the electronic medicine cabinet (eg. medication, band aid, etc.)
T Last triage	Timestamp when the final triage code was given

Table 2: Numerical attributes in extracted input file of EHRs

Attribute	Explanation
Patient number	Unique number assigned to every patient, used for identification purposes
File number	Unique number for every file available for a patient e.g. every time a patient visits the hospital, a new file is opened
Age	The age of the patient
Discharged outside the hospital	Dummy variable indicating if a patient is discharged to a place outside the hospital e.g. home, other hospital, nursing home, etc.

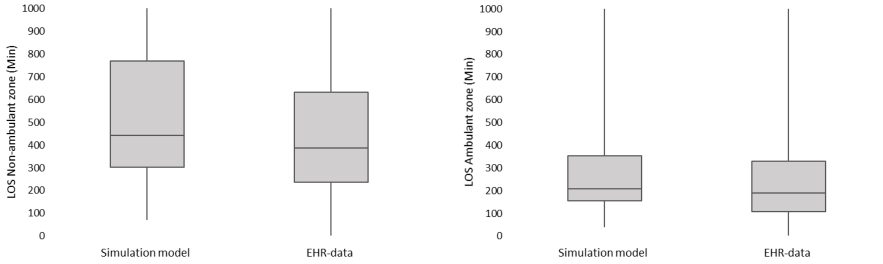
Table 3: Categorical attributes in extracted input file of EHRs

Attribute	Explanation
First triage code	The first triage code assigned to a patient (ESI-triage, code between 1-5)
Last triage code	The last triage code assigned to a patient
Mutation unit	The inpatient unit an admitted patient is assigned to
Brought by	Indicates if a patient came to the hospital by ambulance, police, walk-in, transfer or internal transport
Referral type	Indicates whether the patient is referred to the ED or came to the ED on own initiative
Destination after ED	Indicates the destination of the patient after the ED, for example home, inpatient unit, nursing home, other hospital, passed away, etc.
Discharge type	A patient can be discharged on medical advice, admitted, left without being seen, left against medical advice or passed away.

Table 4: Free text attributes in extracted input file of EHRs

Attribute	Explanation
Main complaint	Most important symptoms of a patient when arriving in the ED
Diagnosis	The final diagnosis made by a physician, registered at the time of departure. This should be a categorical attribute, but the ICD-9 coding is not consistently followed.

B Validation



(a) Non-ambulant zone (b) Ambulant zone

Fig. 1: Validation boxplot for the LOS of patients in the non-ambulant and ambulant zone

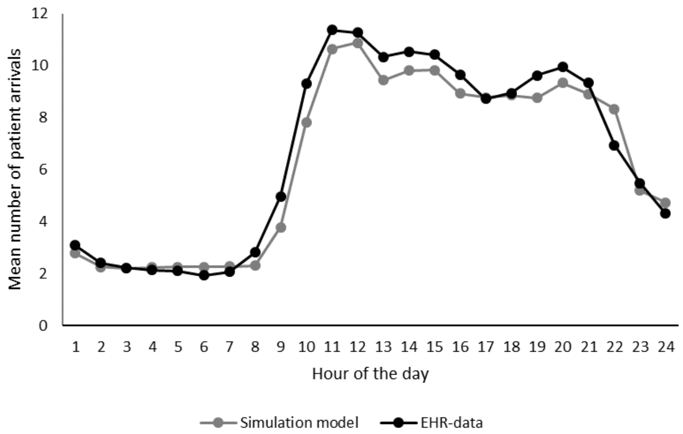


Fig. 2: Graphical validation of hourly number of patient arrivals

C Statistical analysis

This online appendix provides the results of Mauchly's test of sphericity and the repeated-measures full factorial ANOVA. For all main effects and 2-way interactions in the ANOVA, the most appropriate F-statistic is determined by the results of Mauchly's test of sphericity Tables 1 and 10. In case the results of Mauchly's test provide evidence for the violation of the sphericity assumption at the 5% significance level (p-value < 0.05), the G-G estimate of the F-statistic is used in the ANOVA. Otherwise, the sphericity assumed estimate of the F-statistic is used.

C.1 Scenario without multitasking effect

Table 1: Mauchly's test results on sphericity - No multitasking effect

Factor	Mauchly's W	χ^2	df	p-value
DTDT TC2				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.348	103.527	2.000	0.000
Caseload	< 0.001	15395.554	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.700	34.994	2.000	0.000
Priority Q_I x Priority stage	0.245	137.959	2.000	0.000
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	3721.754	299.000	0.000
Priority stage x Caseload	0.000		1175.000	
DTDT TC3				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.270	128.407	2.000	0.000
Caseload	< 0.001	16642.258	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.715	32.853	2.000	0.000
Priority Q_I x Priority stage	0.162	178.502	2.000	0.000
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	4198.760	299.000	0.000
Priority stage x Caseload	0.000		1175.000	
DTDT TC4				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.116	210.945	2.000	0.000
Caseload	< 0.001	19567.202	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.875	13.066	2.000	0.001
Priority Q_I x Priority stage	0.248	136.720	2.000	0.000
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	3385.762	299.000	0.000
Priority stage x Caseload	0.000		1175.000	
DTDT TC5				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.076	79.804	2.000	0.000
Caseload	0.000		299.000	
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.646	13.554	2.000	0.001
Priority Q_I x Priority stage	0.080	78.373	2.000	0.000
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	0.000		299.000	
Priority stage x Caseload	0.000		1175.000	
LOS TC2				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.952	4.771	2.000	0.092
Caseload	< 0.001	7328.667	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.977	2.315	2.000	0.314
Priority Q_I x Priority stage	0.989	1.098	2.000	0.578
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	2902.390	299.000	0.000
Priority stage x Caseload	0.000		1175.000	

Factor	Mauchly's W	χ^2	df	p-value
LOS TC3				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.965	3.492	2.000	0.174
Caseload	< 0.001	9359.974	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.991	0.929	2.000	0.629
Priority Q_I x Priority stage	0.978	2.174	2.000	0.337
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	2749.743	299.000	0.000
Priority stage x Caseload	0.000		1175.000	
LOS TC4				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.478	72.418	2.000	0.000
Caseload	< 0.001	17332.699	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.979	2.040	2.000	0.361
Priority Q_I x Priority stage	0.872	13.406	2.000	0.001
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	< 0.001	2596.656	299.000	0.000
Priority stage x Caseload	0.000		1175.000	
LOS TC5				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.338	33.592	2.000	0.000
Caseload	0.000		299.000	
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.897	3.380	2.000	0.185
Priority Q_I x Priority stage	0.466	23.679	2.000	0.000
Priority Q_P x Caseload	0.000		299.000	
Priority Q_I x Caseload	0.000		299.000	
Priority stage x Caseload	0.000		1175.000	

Table 2: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC2 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	174826095.522	1.000	174826095.522	1618.626	0.000
Priority Q_I	9276360.038	1.000	9276360.038	1206.297	0.000
Priority stage	32680114.757	1.210	26998583.800	1240.468	0.000
Caseload	3684538337.863	1.007	3660722170.411	1577.006	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	398.478	1.000	398.478	22.751	0.000
Priority Q_P x Priority stage	1098.855	1.538	714.412	27.856	0.000
Priority Q_I x Priority stage	7823019.253	1.139	6865897.207	907.767	0.000
Priority Q_P x Caseload	3793908437.425	1.005	3776670580.209	1591.267	0.000
Priority Q_I x Caseload	1747627.838	2.334	748871.846	426.924	0.000
Priority stage x Caseload	6749948.150	2.304	2929423.723	528.630	0.000
Residuals					
Between subjects	67990.740	99	686.775		
Within Priority Q_P	10692883.249	99.000	108008.922		
Within Priority Q_I	761304.916	99.000	7689.949		
Within Priority stage	2608153.539	119.833	21764.835		
Within Caseload	231304996.513	99.644	2321311.962		
Within Priority Q_P x Priority Q_I	1733.934	99.000	17.514		
Within Priority Q_P x Priority stage	3905.346	152.274	25.647		
Within Priority Q_I x Priority stage	853169.350	112.801	7563.502		
Within Priority Q_P x Caseload	236036405.280	99.452	2373373.326		
Within Priority Q_I x Caseload	405260.013	231.034	1754.111		
Within Priority stage x Caseload	1264106.096	228.115	5541.535		

Table 3: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC3 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	175255305.864	1.000	175255305.864	1491.560	0.000
Priority Q_I	8331433.270	1.000	8331433.270	1633.802	0.000
Priority stage	33687669.056	1.156	29144129.247	1718.160	0.000
Caseload	3727366913.863	1.004	3711560207.348	1448.223	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	431.151	1.000	431.151	38.548	0.000
Priority Q_Pp x Priority stage	1110.685	1.557	713.520	39.040	0.000
Priority Q_I x Priority stage	7413058.966	1.088	6813378.017	1172.993	0.000
Priority Q_P x Caseload	3809706138.099	1.003	3797543927.197	1456.443	0.000
Priority Q_I x Caseload	1602369.750	2.112	758601.646	606.120	0.000
Priority stage x Caseload	6994468.812	2.112	3311583.553	778.915	0.000
Residuals					
Between subjects	69117.122	99	698.153		
Within Priority Q_P	11632302.759	99.000	117498.008		
Within Priority Q_I	504842.080	99.000	5099.415		
Within Priority stage	1941075.933	114.434	16962.406		
Within Caseload	254801506.000	99.422	2562838.031		
Within Priority Q_P x Priority Q_I	1107.288	99.000	11.185		
Within Priority Q_P x Priority stage	2816.532	154.106	18.277		
Within Priority Q_I x Priority stage	625658.091	107.714	5808.539		
Within Priority Q_P x Caseload	258960339.739	99.317	2607410.380		
Within Priority Q_I x Caseload	261721.261	209.115	1251.569		
Within Priority stage x Caseload	888995.974	209.100	4251.533		

Table 4: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC4 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	32845398.230	1.000	32845398.230	1181.353	0.000
Priority Q_I	56377.377	1.000	56377.377	57.923	0.000
Priority stage	66804560.540	1.062	62923421.146	2137.068	0.000
Caseload	9074824512.084	1.002	9058626383.465	1002.688	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	58.041	1.000	58.041	13.333	0.000
Priority Q_P x Priority stage	35.138	1.778	19.762	3.867	0.027
Priority Q_I x Priority stage	67810.462	1.141	59408.461	44.163	0.000
Priority Q_P x Caseload	755783855.354	1.004	752813967.333	1156.362	0.000
Priority Q_I x Caseload	30484.922	3.552	8581.331	23.488	0.000
Priority stage x Caseload	13720729.928	1.587	8643219.099	1068.417	0.000
Residuals					
Between subjects	200954.753	99	2029.846		
Within Priority Q_P	2752517.690	99.000	27803.209		
Within Priority Q_I	96358.617	99.000	973.319		
Within Priority stage	3094730.937	105.106	29443.804		
Within Caseload	895999575.087	99.177	9034346.047		
Within Priority Q_P x Priority Q_I	430.977	99.000	4.353		
Within Priority Q_P x Priority stage	899.555	176.028	5.110		
Within Priority Q_I x Priority stage	152011.176	113.001	1345.216		
Within Priority Q_P x Caseload	64705151.187	99.391	651019.084		
Within Priority Q_I x Caseload	128489.745	351.695	365.345		
Within Priority stage x Caseload	1271368.515	157.158	8089.740		

Table 5: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC5 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	517466281.210	1.000	517466281.210	59.335	0.000
Priority Q_I	8511433.401	1.000	8511433.401	282.449	0.000
Priority stage	49445792.053	1.040	47561774.363	308.142	0.000
Caseload	25691260941.281	1.002	25649693348.073	120.770	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	737.498	1.000	737.498	19.389	0.000
Priority Q_P x Priority stage	83.927	1.477	56.826	1.066	0.334
Priority Q_I x Priority stage	6510028.614	1.042	6250261.419	120.826	0.000
Priority Q_P x Caseload	11192915957.653	1.001	11181191862.334	52.094	0.000
Priority Q_I x Caseload	1686064.713	1.748	964346.211	74.211	0.000
Priority stage x Caseload	9871270.185	1.433	6889277.981	114.506	0.000
Residuals					
Between subjects	899717.075	32	28116.159		
Within Priority Q_P	279076548.606	32.000	8721142.144		
Within Priority Q_I	964301.020	32.000	30134.407		
Within Priority stage	5134859.054	33.268	154350.222		
Within Caseload	6807306891.816	32.052	212384153.093		
Within Priority Q_P x Priority Q_I	1217.202	32.000	38.038		
Within Priority Q_P x Priority stage	2518.995	47.261	53.299		
Within Priority Q_I x Priority stage	1724143.540	33.330	51729.553		
Within Priority Q_P x Caseload	6875532305.154	32.034	214635327.577		
Within Priority Q_I x Caseload	727032.369	55.949	12994.588		
Within Priority stage x Caseload	2758646.774	45.851	60165.397		

Table 6: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC2 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	177710959.677	1.000	177710959.677	1574.276	0.000
Priority Q_I	62571453.342	1.000	62571453.342	528.537	0.000
Priority stage	3443945.113	2	1721972.557	22.568	0.000
Caseload	3489232948.950	1.180	2956269227.826	1362.922	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	618.785	1.000	618.785	0.339	0.562
Priority Q_P x Priority stage	7074.354	2	3537.177	2.062	0.130
Priority Q_I x Priority stage	2346739.774	2	1173369.887	16.933	0.000
Priority Q_P x Caseload	3838743633.124	1.066	3600932768.079	1536.250	0.000
Priority Q_I x Caseload	11390707.422	5.269	2161775.519	63.521	0.000
Priority stage x Caseload	1580324.231	12.461	126822.118	4.724	0.000
Residuals					
Between subjects	255286.496	99	2578.651		
Within Priority Q_P	11175539.447	99.000	112884.237		
Within Priority Q_I	11720218.484	99.000	118386.045		
Within Priority stage	15107824.282	198	76302.143		
Within Caseload	253451058.924	116.848	2169066.832		
Within Priority Q_P x Priority Q_I	180671.696	99.000	1824.967		
Within Priority Q_P x Priority stage	339587.355	198	1715.088		
Within Priority Q_I x Priority stage	13720367.814	198	69294.787		
Within Priority Q_P x Caseload	247378720.911	105.538	2343975.396		
Within Priority Q_I x Caseload	17752811.102	521.645	34032.342		
Within Priority stage x Caseload	33121476.809	1233.634	26848.702		

Table 7: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC3 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	176361924.330	1.000	176361924.330	1479.202	0.000
Priority Q_I	44314027.494	1.000	44314027.494	1074.134	0.000
Priority stage	4244904.330	2	2122452.165	51.464	0.000
Caseload	3504112565.700	1.061	3301169691.634	1333.899	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	2457.700	1.000	2457.700	3.635	0.059
Priority Q_P x Priority stage	9325.879	2	4662.939	6.138	0.003
Priority Q_I x Priority stage	3202872.750	2	1601436.375	52.700	0.000
Priority Q_P x Caseload	3850335901.552	1.025	3756236648.438	1420.146	0.000
Priority Q_I x Caseload	7530260.253	5.609	1342624.205	114.097	0.000
Priority stage x Caseload	1617776.280	11.013	146895.879	11.727	0.000
Residuals					
Between subjects	133680.413	99	1350.307		
Within Priority Q_P	11803545.115	99.000	119227.728		
Within Priority Q_I	4084301.993	99.000	41255.576		
Within Priority stage	8165741.377	198	41241.118		
Within Caseload	260069991.413	105.086	2474827.020		
Within Priority Q_P x Priority Q_I	66939.081	99.000	676.152		
Within Priority Q_P x Priority stage	150405.422	198	759.623		
Within Priority Q_I x Priority stage	6016813.631	198	30387.948		
Within Priority Q_P x Caseload	268411280.456	101.480	2644964.796		
Within Priority Q_I x Caseload	6533869.279	555.253	11767.379		
Within Priority stage x Caseload	13657566.095	1090.295	12526.486		

Table 8: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC4 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	32779748.359	1.000	32779748.359	1169.184	0.000
Priority Q_I	140992.909	1.000	140992.909	49.383	0.000
Priority stage	7514246.542	1.314	5719812.835	307.050	0.000
Caseload	8582494020.454	1.004	8551548840.241	959.162	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	128.724	1.000	128.724	5.984	0.016
Priority Q_P x Priority stage	61.830	2.000	30.915	1.239	0.292
Priority Q_I x Priority stage	80289.618	1.773	45277.332	22.059	0.000
Priority Q_P x Caseload	755599844.839	1.006	750888567.905	1152.761	0.000
Priority Q_I x Caseload	82701.764	4.579	18060.907	23.549	0.000
Priority stage x Caseload	2926463.247	2.750	1064359.539	177.277	0.000
Residuals					
Between subjects	277255.652	99	2800.562		
Within Priority Q_P	2775606.552	99.000	28036.430		
Within Priority Q_I	282655.175	99.000	2855.103		
Within Priority stage	2422764.640	130.059	18628.265		
Within Caseload	885842850.812	99.358	8915644.863		
Within Priority Q_P x Priority Q_I	2129.744	99.000	21.513		
Within Priority Q_P x Priority stage	4941.884	198.000	24.959		
Within Priority Q_I x Priority stage	360343.555	175.555	2052.594		
Within Priority Q_P x Caseload	64891488.630	99.621	651382.632		
Within Priority Q_I x Caseload	347683.476	453.326	766.962		
Within Priority stage x Caseload	1634275.125	272.201	6003.925		

Table 9: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC5 - No Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	527456185.837	1.000	527456185.837	62.177	0.000
Priority Q_I	35424976.532	1.000	35424976.532	497.865	0.000
Priority stage	8501778.587	1.204	7063410.150	50.957	0.000
Caseload	25389193137.393	1.002	25338920878.095	122.493	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	2389.595	1.000	2389.595	12.554	0.001
Priority Q_P x Priority stage	484.319	2.000	242.159	1.012	0.369
Priority Q_I x Priority stage	1924019.161	1.304	1475846.203	26.040	0.000
Priority Q_P x Caseload	11437034800.451	1.001	11424748900.816	54.812	0.000
Priority Q_I x Caseload	5566720.802	2.056	2707309.619	117.843	0.000
Priority stage x Caseload	3274981.281	2.160	1516329.065	26.757	0.000
Residuals					
Between subjects	840213.903	32	26256.684		
Within Priority Q_P	271459045.304	32.000	8483095.166		
Within Priority Q_I	2276918.827	32.000	71153.713		
Within Priority stage	5338945.436	38.516	138614.971		
Within Caseload	6632659446.088	32.063	206860198.361		
Within Priority Q_P x Priority Q_I	6091.170	32.000	190.349		
Within Priority Q_P x Priority stage	15308.976	64.000	239.203		
Within Priority Q_I x Priority stage	2364416.067	41.717	56676.841		
Within Priority Q_P x Caseload	6677139083.526	32.034	208436448.826		
Within Priority Q_I x Caseload	1511631.543	65.798	22973.886		
Within Priority stage x Caseload	3916673.899	69.114	56669.851		

C.2 Scenario with multitasking effect

Table 10: Mauchly's test results on sphericity - Multitasking effect

Factor	Mauchly's W	χ^2	df	p-value
DTDT TC2				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.623	46.359	2.000	0.000
Caseload	< 0.001	12851.271	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.387	92.987	2.000	0.000
Priority Q_I x Priority stage	0.315	113.241	2.000	0.000
Priority Q_P x Caseload	< 0.001	18158.741	299.000	0.000
Priority Q_I x Caseload	< 0.001	3783.549	299.000	0.000
Priority stage x Caseload	< 0.001	11554.329	1175.000	0.000
DTDT TC3				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.498	68.225	2.000	0.000
Caseload	< 0.001	13576.657	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.249	136.313	2.000	0.000
Priority Q_I x Priority stage	0.289	121.724	2.000	0.000
Priority Q_P x Caseload	< 0.001	18596.334	299.000	0.000
Priority Q_I x Caseload	< 0.001	4045.429	299.000	0.000
Priority stage x Caseload	< 0.001	11711.719	1175.000	0.000
DTDT TC4				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.312	114.197	2.000	0.000
Caseload	< 0.001	14821.397	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.179	168.491	2.000	0.000
Priority Q_I x Priority stage	0.390	92.306	2.000	0.000
Priority Q_P x Caseload	< 0.001	18308.105	299.000	0.000
Priority Q_I x Caseload	< 0.001	3691.397	299.000	0.000
Priority stage x Caseload	< 0.001	14444.411	1175.000	0.000
DTDT TC5				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.155	179.131	2.000	0.000
Caseload	< 0.001	17949.574	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.187	161.195	2.000	0.000
Priority Q_I x Priority stage	0.107	214.813	2.000	0.000
Priority Q_P x Caseload	< 0.001	24975.713	299.000	0.000
Priority Q_I x Caseload	< 0.001	6003.041	299.000	0.000
Priority stage x Caseload	< 0.001	14591.138	1175.000	0.000
LOS TC2				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.968	3.146	2.000	0.207
Caseload	< 0.001	5665.878	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.995	0.522	2.000	0.770

Factor	Mauchly's W	χ^2	df	p-value
Priority Q_I x Priority stage	0.995	0.443	2.000	0.801
Priority Q_P x Caseload	< 0.001	10530.176	299.000	0.000
Priority Q_I x Caseload	< 0.001	2384.495	299.000	0.000
Priority stage x Caseload	< 0.001	4964.262	1175.000	0.000
LOS TC3				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.988	1.192	2.000	0.551
Caseload	< 0.001	7292.645	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.979	2.085	2.000	0.353
Priority Q_I x Priority stage	0.970	2.945	2.000	0.229
Priority Q_P x Caseload	< 0.001	12726.955	299.000	0.000
Priority Q_I x Caseload	< 0.001	2292.634	299.000	0.000
Priority stage x Caseload	< 0.001	4811.898	1175.000	0.000
LOS TC4				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.837	17.417	2.000	0.000
Caseload	< 0.001	12558.760	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.743	29.110	2.000	0.000
Priority Q_I x Priority stage	0.765	26.285	2.000	0.000
Priority Q_P x Caseload	< 0.001	15479.481	299.000	0.000
Priority Q_I x Caseload	< 0.001	2530.526	299.000	0.000
Priority stage x Caseload	< 0.001	10900.716	1175.000	0.000
LOS TC5				
Priority Q_P	1.000	0.000	0.000	
Priority Q_I	1.000	0.000	0.000	
Priority stage	0.374	94.483	2.000	0.000
Caseload	< 0.001	16481.216	299.000	0.000
Priority Q_P x Priority Q_I	1.000	0.000	0.000	
Priority Q_P x Priority stage	0.651	41.275	2.000	0.000
Priority Q_I x Priority stage	0.198	155.564	2.000	0.000
Priority Q_P x Caseload	< 0.001	22468.975	299.000	0.000
Priority Q_I x Caseload	< 0.001	4079.721	299.000	0.000
Priority stage x Caseload	< 0.001	10376.624	1175.000	0.000

Table 11: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC2 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	25207366.349	1.000	25207366.349	395.117	0.000
Priority Q_I	11479205.463	1.000	11479205.463	1007.812	0.000
Priority stage	40606396.276	1.453	27955458.623	937.873	0.000
Caseload	475580209.935	1.021	465621993.345	341.318	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	1637.929	1.000	1637.929	33.993	0.000
Priority Q_P x Priority stage	1677.753	1.240	1352.952	25.091	0.000
Priority Q_I x Priority stage	10343614.942	1.187	8715049.406	728.637	0.000
Priority Q_P x Caseload	518064172.901	1.003	516664803.272	362.164	0.000
Priority Q_I x Caseload	2671863.855	2.863	933198.279	352.933	0.000
Priority stage x Caseload	10497778.821	3.249	3231229.902	330.183	0.000
Residuals					
Between subjects	65092.363	99	657.499		
Within Priority Q_P	6315931.458	99.000	63797.287		
Within Priority Q_I	1127632.815	99.000	11390.230		
Within Priority stage	4286328.736	143.801	29807.287		
Within Caseload	137943136.629	101.117	1364189.222		
Within Priority Q_P x Priority Q_I	4770.291	99.000	48.185		
Within Priority Q_P x Priority stage	6619.879	122.767	53.922		
Within Priority Q_I x Priority stage	1405387.645	117.500	11960.751		
Within Priority Q_P x Caseload	141616506.570	99.268	1426605.848		
Within Priority Q_I x Caseload	749475.944	283.449	2644.126		
Within Priority stage x Caseload	3147591.264	321.636	9786.189		

Table 12: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC3 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	25091378.944	1.000	25091378.944	393.184	0.000
Priority Q_I	10302888.039	1.000	10302888.039	1160.734	0.000
Priority stage	43117002.441	1.332	32370377.410	1162.371	0.000
Caseload	480997638.252	1.014	474514889.664	341.085	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	1200.748	1.000	1200.748	36.017	0.000
Priority Q_P x Priority stage	1449.122	1.142	1268.823	21.354	0.000
Priority Q_I x Priority stage	9562705.526	1.169	8181930.700	905.249	0.000
Priority Q_P x Caseload	515876702.750	1.002	515023629.237	355.032	0.000
Priority Q_I x Caseload	2445925.610	2.791	876273.032	435.067	0.000
Priority stage x Caseload	11319431.402	2.962	3820941.573	460.376	0.000
Residuals					
Between subjects	61418.214	99	620.386		
Within Priority Q_P	6317775.482	99.000	63815.914		
Within Priority Q_I	878741.936	99.000	8876.181		
Within Priority stage	3672308.210	131.867	27848.585		
Within Caseload	139609656.623	100.353	1391192.299		
Within Priority Q_P x Priority Q_I	3300.525	99.000	33.339		
Within Priority Q_P x Priority stage	6718.220	113.068	59.418		
Within Priority Q_I x Priority stage	1045798.286	115.707	9038.321		
Within Priority Q_P x Caseload	143851362.724	99.164	1450641.258		
Within Priority Q_I x Caseload	556573.160	276.337	2014.110		
Within Priority stage x Caseload	2434150.552	293.285	8299.617		

Table 13: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC4 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	7938763.154	1.000	7938763.154	242.887	0.000
Priority Q_I	176023.994	1.000	176023.994	80.228	0.000
Priority stage	83765862.935	1.185	70705250.170	1385.940	0.000
Caseload	932194805.491	1.011	922153778.715	359.368	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	254.044	1.000	254.044	17.538	0.000
Priority Q_P x Priority stage	79.264	1.098	72.162	2.331	0.127
Priority Q_I x Priority stage	231609.438	1.242	186458.880	59.634	0.000
Priority Q_P x Caseload	186659482.415	1.001	186423993.333	239.264	0.000
Priority Q_I x Caseload	92262.486	5.144	17935.382	23.212	0.000
Priority stage x Caseload	21714542.109	2.532	8576219.890	612.085	0.000
Residuals					
Between subjects	111813.015	99	1129.424		
Within Priority Q_P	3235811.973	99.000	32684.969		
Within Priority Q_I	217211.573	99.000	2194.056		
Within Priority stage	5983535.557	117.287	51016.103		
Within Caseload	256804331.315	100.078	2566042.360		
Within Priority Q_P x Priority Q_I	1434.070	99.000	14.486		
Within Priority Q_P x Priority stage	3366.657	108.743	30.960		
Within Priority Q_I x Priority stage	384501.256	122.973	3126.723		
Within Priority Q_P x Caseload	77233770.466	99.125	779154.874		
Within Priority Q_I x Caseload	393509.909	509.272	772.691		
Within Priority stage x Caseload	3512161.554	250.663	14011.496		

Table 14: 2x2x3x25 full factorial repeated-measures ANOVA on DTDT TC5 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	1167222127.006	1.000	1167222127.006	167.435	0.000
Priority Q_I	55986384.460	1.000	55986384.460	244.721	0.000
Priority stage	274018178.505	1.084	252816161.552	494.756	0.000
Caseload	40953406362.434	1.002	40865486186.277	193.406	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	777.805	1.000	777.805	5.215	0.025
Priority Q_P x Priority stage	3554.483	1.103	3222.957	7.218	0.007
Priority Q_I x Priority stage	50511587.794	1.056	47816539.619	163.207	0.000
Priority Q_P x Caseload	26238213483.745	1.000	26231181621.584	157.567	0.000
Priority Q_I x Caseload	18612043.858	2.019	9217406.844	90.348	0.000
Priority stage x Caseload	82191126.749	2.609	31500821.456	183.071	0.000
Residuals					
Between subjects	3430533.076	97	35366.320		
Within Priority Q_P	676207289.348	97.000	6971209.169		
Within Priority Q_I	22191304.970	97.000	228776.340		
Within Priority stage	53722944.303	105.135	510991.328		
Within Caseload	20539559610.133	97.209	211293449.275		
Within Priority Q_P x Priority Q_I	14468.124	97.000	149.156		
Within Priority Q_P x Priority stage	47766.896	106.978	446.512		
Within Priority Q_I x Priority stage	30020929.424	102.467	292981.044		
Within Priority Q_P x Caseload	16152570604.413	97.026	166476718.583		
Within Priority Q_I x Caseload	19982410.454	195.865	102021.295		
Within Priority stage x Caseload	43548844.527	253.090	172068.688		

Table 15: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC2 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	25256172.713	1.000	25256172.713	363.764	0.000
Priority Q_I	60677067.740	1.000	60677067.740	369.129	0.000
Priority stage	8635307.338	2	4317653.669	43.092	0.000
Caseload	424126981.304	1.368	310131312.660	255.311	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	78.551	1.000	78.551	0.047	0.828
Priority Q_P x Priority stage	3012.090	2	1506.045	0.816	0.444
Priority Q_I x Priority stage	3239676.778	2	1619838.389	21.442	0.000
Priority Q_P x Caseload	533309165.320	1.132	471292370.960	333.878	0.000
Priority Q_I x Caseload	12606464.568	4.447	2834812.785	54.778	0.000
Priority stage x Caseload	4678625.608	10.905	429017.038	11.971	0.000
Residuals					
Between subjects	342860.120	99	3463.234		
Within Priority Q_P	6873572.531	99.000	69430.026		
Within Priority Q_I	16273545.835	99.000	164379.251		
Within Priority stage	19838643.470	198	100195.169		
Within Caseload	164460308.715	135.390	1214718.435		
Within Priority Q_P x Priority Q_I	163908.213	99.000	1655.639		
Within Priority Q_P x Priority stage	365379.492	198	1845.351		
Within Priority Q_I x Priority stage	14957820.343	198	75544.547		
Within Priority Q_P x Caseload	158134641.940	112.027	1411572.494		
Within Priority Q_I x Caseload	22783464.451	440.255	51750.630		
Within Priority stage x Caseload	38692937.751	1079.640	35838.742		

Table 16: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC3 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	25772455.540	1.000	25772455.540	377.828	0.000
Priority Q_I	54141790.037	1.000	54141790.037	900.136	0.000
Priority stage	10364023.706	2	5182011.853	128.603	0.000
Caseload	409494981.257	1.147	356999200.384	268.528	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	145.357	1.000	145.357	0.219	0.641
Priority Q_P x Priority stage	2519.918	2	1259.959	1.472	0.232
Priority Q_I x Priority stage	3747332.501	2	1873666.250	60.316	0.000
Priority Q_P x Caseload	524656542.230	1.041	503857870.559	356.548	0.000
Priority Q_I x Caseload	11461705.739	5.236	2188992.386	143.111	0.000
Priority stage x Caseload	4486110.710	9.167	489360.568	26.527	0.000
Residuals					
Between subjects	153040.912	99	1545.868		
Within Priority Q_P	6752998.064	99.000	68212.102		
Within Priority Q_I	5954699.908	99.000	60148.484		
Within Priority stage	7978363.308	198	40294.764		
Within Caseload	150971206.219	113.558	1329467.087		
Within Priority Q_P x Priority Q_I	65820.925	99.000	664.858		
Within Priority Q_P x Priority stage	169421.194	198	855.663		
Within Priority Q_I x Priority stage	6150680.870	198	31064.045		
Within Priority Q_P x Caseload	145677384.715	103.087	1413155.313		
Within Priority Q_I x Caseload	7928855.221	518.370	15295.733		
Within Priority stage x Caseload	16742073.136	907.562	18447.309		

Table 17: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC4 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	7923267.392	1.000	7923267.392	242.550	0.000
Priority Q_I	338441.922	1.000	338441.922	70.468	0.000
Priority stage	15866799.463	1.720	9225175.517	241.184	0.000
Caseload	812449974.538	1.040	781284781.322	315.969	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	570.887	1.000	570.887	17.004	0.000
Priority Q_P x Priority stage	83.015	1.591	52.175	0.923	0.380
Priority Q_I x Priority stage	196333.553	1.619	121260.746	23.212	0.000
Priority Q_P x Caseload	186964464.147	1.004	186232854.329	236.943	0.000
Priority Q_I x Caseload	191706.505	5.627	34068.499	19.761	0.000
Priority stage x Caseload	7274784.408	2.941	2473476.220	108.616	0.000
Residuals					
Between subjects	209537.751	99	2116.543		
Within Priority Q_P	3233990.728	99.000	32666.573		
Within Priority Q_I	475472.681	99.000	4802.754		
Within Priority stage	6512916.609	170.275	38249.488		
Within Caseload	254558115.894	102.949	2472660.486		
Within Priority Q_P x Priority Q_I	3323.866	99.000	33.574		
Within Priority Q_P x Priority stage	8899.744	157.519	56.499		
Within Priority Q_I x Priority stage	837365.789	160.291	5224.031		
Within Priority Q_P x Caseload	78117942.677	99.389	785982.421		
Within Priority Q_I x Caseload	960400.619	557.082	1723.985		
Within Priority stage x Caseload	6630732.729	291.171	22772.669		

Table 18: 2x2x3x25 full factorial repeated-measures ANOVA on LOS TC5 - Multitasking effect

	Sum of squares	df	Mean square	F	p-value
Main effects					
Priority Q_P	1164087352.986	1.000	1164087352.986	167.199	0.000
Priority Q_I	155749349.652	1.000	155749349.652	489.482	0.000
Priority stage	93467084.896	1.230	76000953.920	171.331	0.000
Caseload	39579804809.290	1.004	39439217175.185	187.849	0.000
Two-way interactions					
Priority Q_P x Priority Q_I	259.658	1.000	259.658	0.774	0.381
Priority Q_P x Priority stage	2398.783	1.482	1618.524	2.515	0.100
Priority Q_I x Priority stage	26124479.042	1.110	23540663.243	80.632	0.000
Priority Q_P x Caseload	26254450172.621	1.000	26244945714.781	157.858	0.000
Priority Q_I x Caseload	37686267.061	2.129	17705427.540	143.525	0.000
Priority stage x Caseload	45827123.773	3.400	13477832.969	83.261	0.000
Residuals					
Between subjects	3822628.551	97	39408.542		
Within Priority Q_P	675341491.261	97.000	6962283.415		
Within Priority Q_I	30864650.269	97.000	318192.271		
Within Priority stage	52916874.256	119.292	443591.069		
Within Caseload	20437877978.781	97.346	209951366.492		
Within Priority Q_P x Priority Q_I	32538.241	97.000	335.446		
Within Priority Q_P x Priority stage	92533.560	143.762	643.659		
Within Priority Q_I x Priority stage	31427653.348	107.647	291951.880		
Within Priority Q_P x Caseload	16132700723.776	97.035	166256293.536		
Within Priority Q_I x Caseload	25469936.742	206.466	123361.440		
Within Priority stage x Caseload	53389133.425	329.818	161874.562		

D Results scenario without multitasking effect

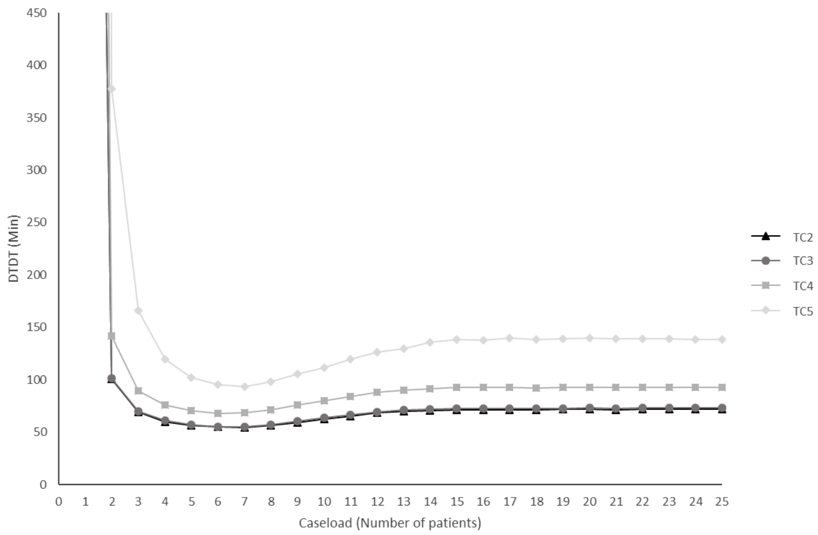


Fig. 1: Mean DTDT per TC as a function of caseload. Note: DTDT at caseload 1 is very high because of the large amount of physician idle time, and will never be used in practice. These values are not presented in the figures for clarity purposes. The DTDT at caseload 1 equals (in minutes): TC2: 1855.96, TC3: 1867.53, TC4: 2893.89, TC5: 8738.38.

Table 1: Significant potential KPI improvements under the current queueing disciplines (TC-TC-Equal) with corresponding caseload limit - No Multitasking effect

KPI	Current value (Min)	Minimum value (Min)	Caseload limit (Min)	% improvement
DTDT - TC2	55.44	38.33	1	30.86%
DTDT - TC3	58.47	46.13	1	21.10%
DTDT - TC4	-	-	-	-
DTDT - TC5	-	-	-	-
LOS - TC2	409.68	355.54	1	13.22%
LOS - TC3	373.22	302.24	1	19.02%
LOS - TC4	-	-	-	-
LOS - TC5	-	-	-	-

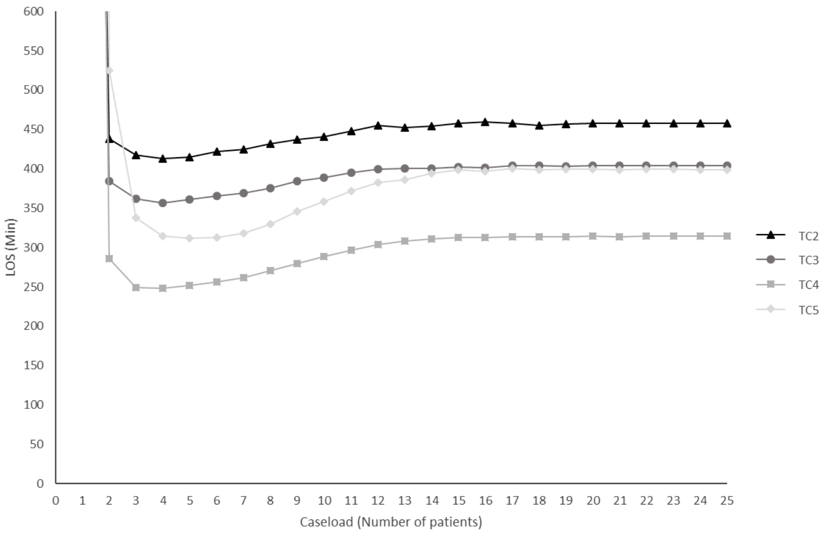


Fig. 2: Mean LOS per TC as a function of caseload. Note: LOS at caseload 1 is very high because of the large amount of physician idle time, and will never be used in practice. These values are not presented in the figures for clarity purposes. The LOS at caseload 1 equals (in minutes): TC2: 2138.78, TC3: 2133.13, TC4: 3021.15, TC5: 8925.27.

Table 2: Significant KPI improvements under the current queueing disciplines (TC-TC-Equal) for caseload limit 2 - No multitasking effect

KPI	Current value (Min)	Value at caseload 2 (Min)	% improvement
DTDT - TC2	55.44	46.02	17.01%
DTDT - TC3	58.47	48.50	17.06%
DTDT - TC4	-	-	-
DTDT - TC5	-	-	-
LOS - TC2	-	-	-
LOS - TC3	373.22	329.18	11.80%
LOS - TC4	-	-	-
LOS - TC5	-	-	-

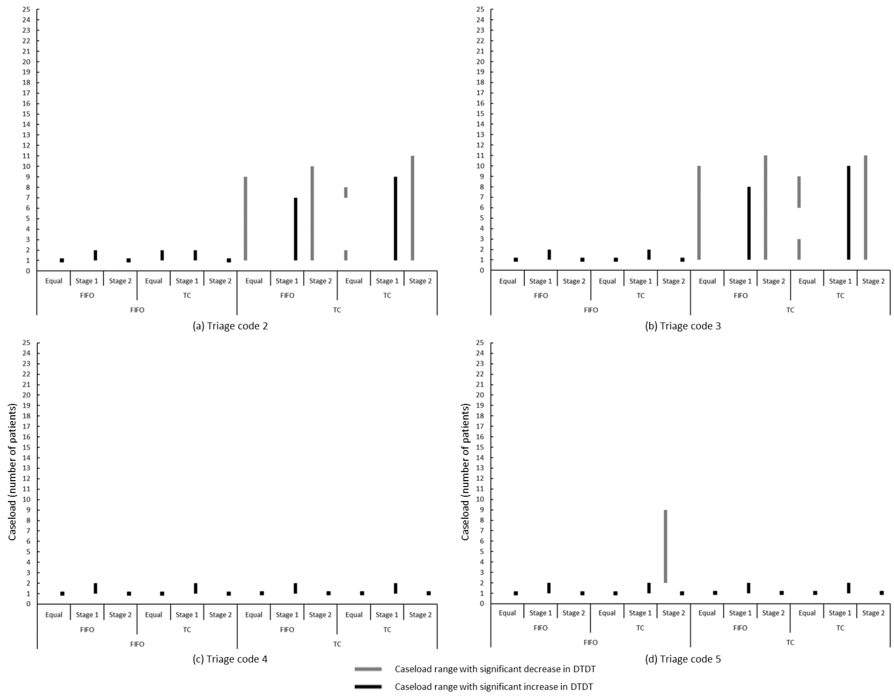


Fig. 3: Caseload range resulting in significant DTDT improvement in comparison with no caseload limit for each priority factor combination (per triage code) - No multitasking effect

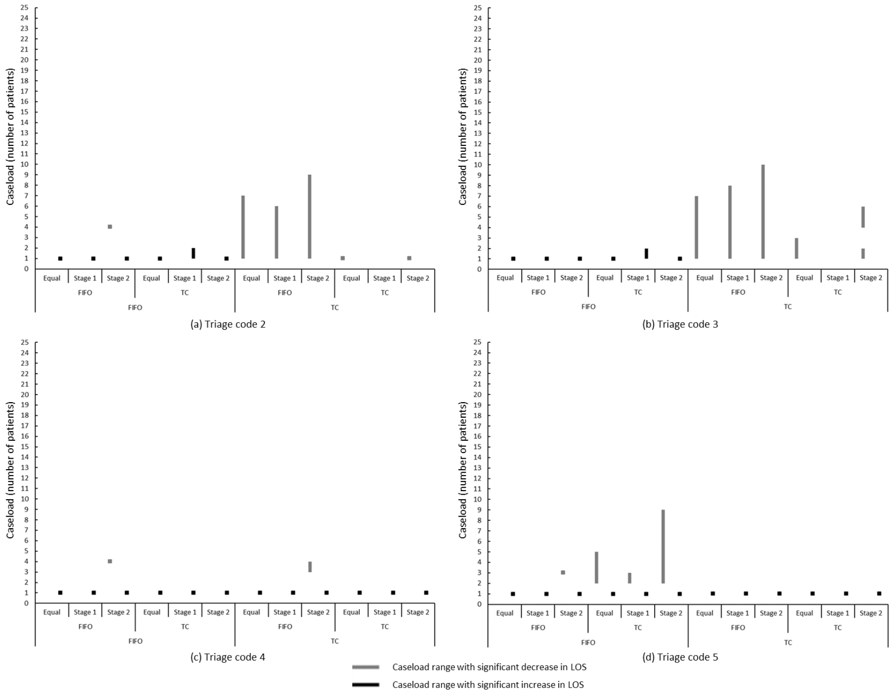


Fig. 4: Caseload range resulting in significant LOS improvement in comparison with no caseload limit for each priority factor combination (per triage code) - No multitasking effect

Acknowledgements This work is supported by the Strategic Basic Research project Data-driven logistics (S007318N), funded by the Research Foundation Flanders (FWO).

References

Abo-Hamad W, Arisha A (2013) Simulation-based framework to improve patient experience in an emergency department. European Journal of Operational Research 224(1):154–166, DOI 10.1016/j.ejor.2012.07.028

Aral S, Brynjolfsson E, Van Alstyne M (2012) Information, Technology, and Information Worker Productivity. Information Systems Research 23(3-part-2):849–867, DOI 10.1287/isre.1110.0408

Batt RJ, Terwiesch C (2012) Doctors under load: An empirical study of state-dependent service times in emergency care. The Wharton School, the Univ ersity of Pennsylvania, Philadelphia, PA 19104

Batt RJ, Terwiesch C (2016) Early Task Initiation and Other Load-Adaptive Mechanisms in the Emergency Department. Management Science 63(11):3531–3551, DOI 10.1287/mnsc.2016.2516

Bhattacharjee P, Ray PK (2014) Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: A review and reflections. Computers & Industrial Engineering 78:299–312, DOI 10.1016/j.cie.2014.04.016

Brailsford SC, Hilton NA (2001) A comparison of discrete event simulation and system dynamics for modelling health care systems. In: Riley J (ed) Planning for the Future: Health Service Quality and Emergency Accessibility, Glasgow Caledonian University

Campello F, Ingolfsson A, Shumsky RA (2017) Queueing Models of Case Managers. Management Science 63(3):882–900, DOI 10.1287/mnsc.2015.2368

Carmen R, Defraeye M, Van Nieuwenhuyse I (2015) A Decision Support System for Capacity Planning in Emergency Departments. International Journal of Simulation Modelling 14(2):299–312, DOI 10.2507/IJSIMM14(2)10.308

Chisholm CD, Collison EK, Nelson DR, Cordell WH (2000) Emergency Department Workplace Interruptions Are Emergency Physicians ”Interrupt-driven” and ”Multitasking”? Academic Emergency Medicine 7(11):1239–1243, DOI 10.1111/j.1553-2712.2000.tb00469.x

Cildoz M, Mallor F, Ibarra A (2018) Analysing the ED Patient Flow Management Problem by using Accumulating Priority Queues and Simulation-based Optimization. In: 2018 Winter Simulation Conference (WSC), IEEE, Gothenburg, Sweden, pp 2107–2118

Delasay M, Ingolfsson A, Kolfal B, Schultz KL (2015) Load Effect on Service Times. SSRN Electronic Journal DOI 10.2139/ssrn.2647201

Dobson G, Tezcan T, Tilson V (2013) Optimal Workflow Decisions for Investigators in Systems with Interruptions. Management Science 59(5):1125–1141, DOI 10.1287/mnsc.1120.1632

Duguay C, Chetouane F (2007) Modeling and Improving Emergency Department Systems using Discrete Event Simulation. SIMULATION 83(4):311–320, DOI 10.1177/0037549707083111

Duma D, Aringhieri R (2018) An ad hoc process mining approach to discover patient paths of an emergency department. Flexible Services and Manufacturing Journal pp 1–29

Ferrand YB, Magazine MJ, Rao US, Glass TF (2018) Managing responsiveness in the emergency department: Comparing dynamic priority queue with fast track. Journal of Operations Management 58-59(1):15–26, DOI 10.1016/j.jom.2018.03.001

Field A (2013) Discovering Statistics Using IBM SPSS Statistics. SAGE, google-Books-ID: c0Wk9IuBmAoC

Forster AJ (2003) The Effect of Hospital Occupancy on Emergency Department Length of Stay and Patient Disposition. Academic Emergency Medicine 10(2):127–133, DOI 10.1197/aemj.10.2.127

Ghanes K, Wargon M, Jouini O, Jemai Z, Diakogiannis A, Hellmann R, Thomas V, Koole G (2015) Simulation-based optimization of staffing levels in an emergency department. SIMULATION 91(10):942–953, DOI 10.1177/0037549715606808

Graff LG, Wolf S, Dinwoodie R, Buono D, Mucci D (1993) Emergency physician workload: A time study. Annals of Emergency Medicine 22(7):1156–1163, DOI 10.1016/S0196-0644(05)80982-5

Gul M, Guneri AF (2015) A comprehensive review of emergency department simulation applications for normal and disaster conditions. Computers & Industrial Engineering 83:327–344, DOI 10.1016/j.cie.2015.02.018

Gunal MM, Pidd M (2006) Understanding accident and emergency department performance using simulation. In: Proceedings of the 38th conference on Winter simulation, Winter Simulation Conference, pp 446–452

Hoot NR, Aronsky D (2008) Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions. Annals of Emergency Medicine 52(2):126–136.e1, DOI 10.1016/j.annemergmed.2008.03.014

Kang H, Nembhard HB, Rafferty C, DeFlitch CJ (2014) Patient Flow in the Emergency Department: A Classification and Analysis of Admission Process Policies. Annals of Emergency Medicine 64(4):335–342.e8, DOI 10.1016/j.annemergmed.2014.04.011

Kc DS (2014) Does Multitasking Improve Performance? Evidence from the Emergency Department. Manufacturing & Service Operations Management 16(2):168–183, DOI 10.1287/msom.2013.0464, URL http://pubsonline.informs.org/doi/abs/10.1287/msom.2013.0464

Kc DS, Terwiesch C (2009) Impact of Workload on Service Time and Patient Safety: An Econometric Analysis of Hospital Operations. Management Science 55(9):1486–1498, DOI 10.1287/mnsc.1090.1037

Kelton WD, Sadowski RP, Zupick NB (2015) Simulation with Arena, 6th edn. McGraw-Hill Education, New York, NY

Kuo YH, Rado O, Lupia B, Leung JMY, Graham CA (2016) Improving the efficiency of a hospital emergency department: a simulation study with indirectly imputed service-time distributions. Flexible Services and Manufacturing Journal 28(1-2):120–147, DOI 10.1007/s10696-014-9198-7

Levin S, Aronsky D, Hemphill R, Han J, Slagle J, France DJ (2007) Shifting Toward Balance: Measuring the Distribution of Workload Among Emergency Physician Teams. Annals of Emergency Medicine 50(4):419–423, DOI 10.1016/j.annemergmed.2007.04.007

Li N, Stanford DA (2016) Multi-server accumulating priority queues with heterogeneous servers. European Journal of Operational Research 252(3):866–878, DOI 10.1016/j.ejor.2016.02.010

Li N, Stanford DA, Sharif AB, Caron RJ, Pardhan A (2019) Optimising key performance indicator adherence with application to emergency department congestion. European Journal of Operational Research 272(1):313–323, DOI 10.1016/j.ejor.2018.06.048

Maidstone R (2012) Discrete event simulation, system dynamics and agent based simulation: Discussion and comparison. System 1(6)

McKay KN, Engels JE, Jain S, Chudleigh L, Shilton D, Sharma A (2013) Emergency departments:"repairs while you wait, no appointment necessary". In: Handbook of Healthcare Operations Management, Springer, pp 349–385

Mohiuddin S, Busby J, Savović J, Richards A, Northstone K, Hollingworth W, Donovan JL, Vasilakis C (2017) Patient flow within UK emergency departments: a systematic review of the use of computer simulation modelling methods. BMJ Open 7(5):e015007, DOI 10.1136/bmjopen-2016-015007

Paul JA, Lin L (2012) Models for Improving Patient Throughput and Waiting at Hospital Emergency Departments. The Journal of Emergency Medicine 43(6):1119–1126, DOI 10.1016/j.jemermed.2012.01.063

Pines JM, Hilton JA, Weber EJ, Alkemade AJ, Al Shabanah H, Anderson PD, Bernhard M, Bertini A, Gries A, Ferrandiz S, Kumar VA, Harjola VP, Hogan B, Madsen B, Mason S, Öhlén G, Rainer T, Rathlev N, Revue E, Richardson D, Sattarian M, Schull MJ (2011) International Perspectives on Emergency Department Crowding. Academic Emergency Medicine 18(12):1358–1370, DOI 10.1111/j.1553-2712.2011.01235.x

Saghafian S, Hopp WJ, Van Oyen MP, Desmond JS, Kronick SL (2012) Patient Streaming as a Mechanism for Improving Responsiveness in Emergency Departments. Operations Research 60(5):1080–1097, DOI 10.1287/opre.1120.1096

Saghafian S, Austin G, Traub SJ (2015) Operations research/management contributions to emergency department patient flow optimization: Review and research prospects. IIE Transactions on Healthcare Systems Engineering 5(2):101–123, DOI 10.1080/19488300.2015.1017676

Tan KW, Wang C, Lau HC (2012) Improving patient flow in emergency department through dynamic priority queue. In: 2012 IEEE International Conference on Automation Science and Engineering (CASE), IEEE, Seoul, Korea (South), pp 125–130, DOI 10.1109/CoASE.2012.6386409

Tan TF, Netessine S (2014) When Does the Devil Make Work? An Empirical Study of the Impact of Workload on Worker Productivity. Management Science 60(6):1574–1593, DOI 10.1287/mnsc.2014.1950

Vanbrabant L, Braekers K, Ramaekers K, Nieuwenhuyse IV (2019a) Simulation of emergency department operations: a comprehensive review of KPIs and operational improvements. Computers & Industrial Engineering DOI 10.1016/j.cie.2019.03.025

Vanbrabant L, Martin N, Ramaekers K, Braekers K (2019b) Quality of input data in emergency department simulations: Framework and assessment techniques. Simulation Modelling Practice and Theory 91:83–101, DOI 10.1016/j.simpat.2018.12.002

Yang KK, Lam SSW, Low JM, Ong MEH (2016) Managing emergency department crowding through improved triaging and resource allocation. Operations Research for Health Care 10:13–22, DOI 10.1016/j.orhc.2016.05.001

Zeinali F, Mahootchi M, Sepehri MM (2015) Resource planning in the emergency departments: A simulation-based metamodeling approach. Simulation Modelling Practice and Theory 53:123–138, DOI 10.1016/j.simpat.2015.02.002

Title Page

Noname manuscript No.
(will be inserted by the editor)

Improving emergency department performance by revising the patient-physician assignment process

Lien Vanbrabant · Kris Braekers · Katrien Ramaekers

Received: date / Accepted: date

Abstract Emergency departments (EDs) are continuously exploring opportunities to improve their efficiency. A new opportunity lies in revising the patient-physician assignment process by limiting the number of patients simultaneously assigned to a single physician, which is defined as the application of a case manager approach with limited caseloads. The potential of introducing a case manager approach with limited caseloads as a way to improve physician productivity, and consequently ED performance, is investigated by use of a discrete-event simulation model based on a real-life case study. In addition, as the case manager system is characterised by three parameters that can be customised and optimised (i.e. caseload limit, pre-assignment queueing discipline and internal queueing discipline), the impact of these parameters on the effectiveness to improve ED performance in terms of length-of-stay and door-to-doctor time is evaluated. To the best of our knowledge, this paper is the first to examine the potential of a case manager system with limited caseloads in a complex service system like a real-life ED, and to investigate the impact of the three system parameters on the results. The outcomes of the study show that performance can be improved significantly by introducing a case manager system, and that the system parameters have an impact on the effect size.

Keywords Discrete-event simulation · Emergency department · Patient-physician assignment · Real-life case study · Healthcare operations

Lien Vanbrabant
UHasselt, Research Group Logistics, Agoralaan, 3590 Diepenbeek, Belgium
Research Foundation Flanders (FWO), Egmontstraat 5, 1000 Brussel, Belgium
E-mail: lien.vanbrabant@uhasselt.be

Kris Braekers, Katrien Ramaekers
UHasselt, Research Group Logistics, Agoralaan, 3590 Diepenbeek, Belgium
E-mail: kris.braekers@uhasselt.be, katrien.ramaekers@uhasselt.be