# Keeping pace with the ebbs and flows in daily nursing home operations

René Bekker
Vrije Universiteit Amsterdam

Dennis Moeke
HAN University of Applied Sciences
KennisDC Logistiek Gelderland
dennis.moeke@han.nl

Bas Schmidt
Deloitte, Amsterdam

1st March 2018

**Abstract**

Nursing homes are challenged to develop staffing strategies that enable them to efficiently meet the healthcare demand of their residents. In this study, we investigate how demand for care and support fluctuates over time and during the course of a day, using demand data from three independent nursing home departments of a single Dutch nursing home. This demand data is used as input for an optimization model that provides optimal staffing patterns across the day. For the optimization we use a Lindley-type equation and techniques from stochastic optimization to formulate a Mixed-Integer Linear Programming (MILP) model. The impact of both the current and proposed staffing patterns, in terms of waiting time and service level, are investigated. The results show substantial improvements for all three departments both in terms of average waiting time as well as in 15 minutes service level. Especially waiting during rush hours is significantly reduced, whereas there is only a slight increase in waiting time during non-rush hours.

# 1 Introduction

Like many Western countries, the Netherlands is plagued by increasing financial pressure on its Long-Term Care (LTC) system. Between 2009 and 2012, the Dutch public spending on LTC increased by more than 20% and on average accounted for more than 40% of the total public healthcare spending [12]. With 4.3% of GDP in 2014, the Dutch public spending on LTC was the highest among the OECD countries [26]. In addition, recent projection scenarios show a non-negligible increase in public spending on LTC over the forthcoming decades [21][30][17].

In an attempt to reduce the growth rate of the expenditures on LTC, without compromising on quality, radical reforms were introduced by the Dutch government at the start of 2015 [38][23]. Under these reforms, the responsibility for most formal long-term care services has been decentralized and taken over by local authorities. New clients requiring lighter types of care no longer receive an indication for admitted care. Instead, they now receive care at home. As a result, the care needs of nursing home clients are becoming more severe as healthier people are drawn away from nursing home facilities.

Together with ongoing budget cuts, these reforms have put serious pressure on Dutch nursing homes, which face the challenge of providing high-quality, cost-efficient care in a rapidly changing healthcare landscape. Consequently, most nursing homes are searching for ways to further streamline their care and support activities with the purpose of lowering costs while maintaining an appropriate quality level of care [36]. The relevance of Quality of Care (QoC) in a nursing home differs from many other service settings. Often, nursing home residents are in need of ongoing assistance with basic activities of daily living due to physical or psychological disabilities. Therefore, in order to make it possible for nursing home residents to live their lives according to their own daily routines, the necessary care and support should be delivered as close as possible to the time preferences of the residents.

Capacity planning plays a key role in ensuring that an organization has the capability to respond sufficiently to the level of demand experienced, see e.g. [19]. For nursing homes, care workers are by far the most important resource. This is due to the fact that care workers are responsible for the daily care and supervision of the residents and their labor costs account for a significant proportion of the total healthcare expenditure [11]. As such, capacity planning in a nursing home setting consists for an important part of workforce planning. A prominent question that arises in this context is: how to create a staffing plan that meets the demand of the nursing home residents as closely as possible, without overstretching the available staffing hours?

In daily nursing home practice, staffing decisions are often made without a sound rational basis. In our view, nursing homes could greatly benefit from a more rational staffing approach. Hence, our preliminary hypothesis is that a mathematically-based staffing method leads to beter performance in terms of meeting the time preferences of residents for care and support activities.

Effective workforce planning starts with insight in and understanding of demand patterns [15]. Regarding the healthcare demand in a nursing home, a distinction should be made between ‘scheduled care' (i.e., ‘care by appointment') [20][25] and ‘unscheduled care' (i.e., ‘care on demand') [36]. This study focuses primarily on scheduled care, for which it is possible, based on the needs and preferences of the residents, to make a fairly detailed planning in advance. In the sequel, the predefined activities of scheduled care with a preferred starting time are referred to as *Preferred Activity Times* (PATs). In this study, we investigate how PATs fluctuate over time and during the course of a day, using data from three independent nursing home departments of a single Dutch nursing home. The PAT-data is the input for an optimization model that provides optimal staffing patterns across the day. The impact of both the current and proposed staffing patterns, in terms of waiting time and service level, are investigated.

## 1.1 Available literature

There is a growing body of Operations Research (OR) literature on capacity planning in healthcare. However, the vast majority of the ‘OR in healthcare' literature is on capacity decisions in hospitals [27] and to date the area of nursing home care has received hardly any attention. This finding is in line with [18], which provides an overview of studies in the field of OR and Management Science (MS) in healthcare and found that the body of OR/MS literature directed to residential care services is limited. To the best of our knowledge, [20][25][24][36] are the only studies that examined capacity planning issues in a nursing home context using OR-methods. The results of these studies show that, by using a data-driven approach, extensive gains in performance can be achieved. The authors of [24] show that a more flexible allocation of care workers has a substantial positive effect on the performance in terms of meeting the time preferences of the nursing home residents. However, we could find no study in which a more advanced algorithm is used to determine the optimal staffing pattern in a nursing home setting. We think that an important reason for this is the lack of (reliable) data. According to [34], most of the available data is qualitative and largely unstructured and often not used for decision making.

The allocation of care workers in a home care setting has been studied more extensively, see e.g. [13][22][28]. When it comes to home care, care workers are assigned to care and support activities, which should be carried out within a client-specific time window. As these activities are performed at the clients' homes, the spatial component is a crucial element. This scheduling problem is related to the vehicle routing problem with time windows, which is known to be NP-hard, see e.g. [8]. Since home care organizations usually have to schedule tens or hundreds of activities, approximation methods (or heuristics) have been proposed to give satisfactory outcomes within an acceptable time frame.

The primary objective of this study is to determine the appropriate capacity level across the day, such that waiting times of clients are avoided as much as possible. This differs from the area of nurse scheduling in hospitals, where the goal is to construct actual workforce plans, see e.g. [5] for a comprehensive review. The papers that are most closely related are [37] and [39]. These two papers use Markovian models to describe nursing workload at an inpatient hospital department. Thereby, they take admissions and discharges into account and the fluctuations in demand while a patient occupies a bed. The assumptions of their models are crucially different. Both papers assume stationary demand process, whereas there are large peaks in demand corresponding to natural moments of activities in daily living. In addition, the Markovian assumptions being used in [37] [39] can be questioned in this setting. Moreover, the authors do not consider differences in capacity across the day. Finally, we like to note that the impact of waiting in an inpatient hospital setting is considerably different from nursing home situtions. For inpatients the length of stay is in the order of days, causing some inconveniences if they have to wait (for example for help with washing). In nursing homes, waiting times directly affect the daily routine, while clients are long-term dependent on care being provided.

There are other application areas, like call centers, that have a much longer tradition in applying (advanced) OR-methods to workforce planning, see e.g. [2][7][6][10]. Like in a nursing home setting, the capacity in a call center cannot be inventoried and salaries typically account for a large share of the total operating costs. In order to meet the demand in a cost-efficient manner, call center managers vary the number of agents to track the predictable variability in the number of calls. Workforce planning in a call center is typically solved using a four-step hierarchical framework: (1) workload prediction, (2) determining staffing levels, (3) shift scheduling, (4) rostering, It should be noted that shift scheduling refers to generating shifts such that the required staffing levels are met, whereas rostering refers to the pairing of shifts into rosters and the assignment of employees to the rosters; see [6] or [14] for additional background. In call centers, step (1) is already advanced, see e.g. [3][32][35] for forecasting only the arrival pattern. Also, all four steps are traditionally executed separately, where multi-server queues play a prominent role for determining the required staffing capacity in step (2). For our nursing home, we essentially carry out steps (1) - (3), albeit step (1) now involves basic data analysis. As there is no obvious queueing model available, we combine steps (2) and (3).

Outside the field of OR there are numerous studies that address the relationship between staffing levels and QoC in a nursing home setting. The literature study of [33] provides a systematic overview. The authors conclude that existing studies mainly focus on clinical outcomes as a measure of QoC and it is difficult to draw conclusions and offer recommendations based on existing studies. In the more recent study of [4], which also uses clinical outcomes as measure, no consistent evidence was found for a positive relationship between staffing levels and QoC.

## 1.2 Contribution and outline

The study of capacity planning in nursing home care is still in its infancy, whereas there is a growing interest among nursing home managers and policymakers for this topic. The main contribution of this paper is that we provide a mathematical optimization approach for staffing, resulting in more timely care delivery and a better balanced workload for the care workers. For the optimization we use real-life PAT-data, which is scarce. As such, the data analysis is of interest on its own. Moreover, for the optimization we apply a Lindley-type equation and techniques from stochastic optimization to formulate a Mixed-Integer Linear Programming (MILP) model. This formulation differs from the standard call center shift-scheduling literature.

The remainder of this paper is structured as follows. A brief description

Table 1: Type of care per department

| Department | Type of care |
|---|---|
| C | Somatic care |
| D | Somatic care |
| E | Psychogeriatric care |

of the empirical context is provided in Section 2. Section 3 investigates how the PATs and workload fluctuate over time and during the course of a day. Next, in Section 4, we propose a MILP model for determining the optimal staffing patterns and provide an overview of the simulation approach used in this study to explore the impact of the staffing patterns in terms of waiting times and service level. In Section 5, we present the numerical results of the current and proposed staffing levels. Section 6 concludes and points out possible directions for further research.

## 2 Empirical context

The nursing home departments under study provide accommodation for people who require assistance with daily activities such as washing, dressing, eating, drinking and taking medication. Furthermore, medical attention is given as required. In this study, the following definition of nursing home applies: "a facility with a domestic-styled environment that provides 24-hour functional support and care for persons who require assistance with activities of daily living and who often have complex health needs and increased vulnerability" [29, p. 183]. Table 1 shows that the type of care differs per department.

In order to make it possible for nursing home residents to live their lives according to their own daily routines, all departments aim to deliver the necessary care as close as possible to the time preferences of the residents. The time preferences are inventoried on a regular basis, using a standardized, systematic method.

To adequately meet the demand of the residents, it is crucial to maintain an appropriate number and mix of care workers during the course of a day. The care worker-to-resident ratios currently applied by departments C, D and E, which have been provided to us by the nursing home manager, are

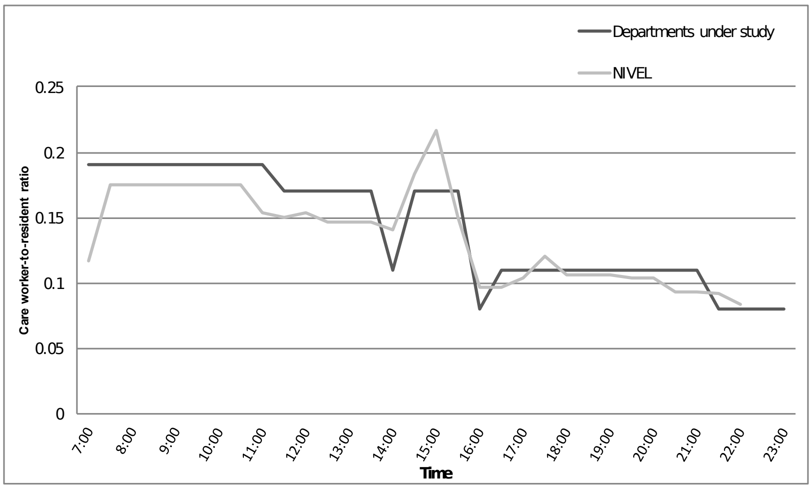


Figure 1: Care worker-to-resident ratios

shown in Figure 1. This figure also shows worker-to-resident ratios that are based on a study of The Netherlands Institute for Health Services Research (NIVEL) [16] among a large sample of Dutch nursing homes. There seems to be a reasonable fit between the ratios used by the nursing homes under study and the ratios presented by NIVEL. We note that the peak around 15:00 hours presented in the NIVEL study is due to information transfer between shifts.

For the delivery of care and support, the departments under study make use of the so-called differentiated practice. Based on their education and expertise, care workers are hierarchically divided into four distinct qualification levels (QLs). Depending on the required education and complexity of care, healthcare tasks are assigned to a healthcare worker with that specific qualification level. As the aim of this study is to provide insight on an aggregate level, task requirements in terms of qualification levels are not taken into account.

# 3 Demand and workload analysis

Three separate datasets were used (one for each department). Each dataset contains the PAT data of 91 days, from January 1 until April 4, 2014, and consists of the following variables:

Table 2: Number of residents and size of dataset

| Dep. | Number of residents | Average rows of data per day | Average rows of data per day / Number of residents |
|---|---|---|---|
| C | 34 | 175 | 5.1 |
| D | 28 | 257 | 9.2 |
| E | 33 | 251 | 7.6 |
| Total | 95 | 683 | 7.2 |

- *Resident ID* – the ID of a specific resident.
- *Preferred Activity Time (PAT)* – the preferred starting time of the healthcare activity.
- *Date* – the date of the PAT.
- *Task description* – a brief description of the activity (i.e. healthcare task) entered as free text.
- *Expected service time* – expected duration of the activity in minutes.

Table 2 shows the size of the datasets and the number of residents per department. The final column gives the average number of care activities per resident.

The analysis is structured as follows. First, for each department, the demand patterns (Subsection 3.1) and care duration (Subsection 3.2) are analyzed. In Subsection 3.3, the average workload over the course of a day is examined and essentially follows from the combination of demand patterns and care durations. Finally, Subsection 3.4 elaborates on the assumptions regarding the demand for unscheduled care.

## 3.1 PAT-analysis

To get an impression of the demand patterns, we divide the days into time periods (time buckets) of 5 (or 30) minutes and calculate the number of PATs within each time bucket over the course of a day. Observe that this only involves the prefered starting time of activities. We let $T$ denote the number of time intervals; for time buckets of 5 (or 30) minutes, $T$ equals 288 (or 48). Let $X_{t,d}(y)$ be the number of PATs in interval $t$ at day $d \in \{1, \ldots, 91\}$,

Table 3: Results of the Friedman test

| Department | 5 minute intervals $p$-value | 30 minute intervals $p$-value |
|---|---|---|
| C | 0.9063 | 0.999 |
| D | <2.2e-11 | <2.2e-11 |
| E | 1 | 1 |

for department $y$, where $y \in \{\mathrm{C}, \mathrm{D}, \mathrm{E}\}$. For department $y$ this results in a matrix of the form:

$$\begin{pmatrix} X_{1,1}(y) & X_{1,2}(y) & \cdots & X_{1,91}(y) \\ X_{2,1}(y) & X_{2,2}(y) & \cdots & X_{2,91}(y) \\ \vdots & \vdots & \ddots & \vdots \\ X_{T,1}(y) & X_{T,2}(y) & \cdots & X_{T,91}(y) \end{pmatrix}$$

The first step is to determine to what extent the PATs vary between days within each department. Hence, the following assumption is tested: For $y \in \{\mathrm{C}, \mathrm{D}, \mathrm{E}\}$,

$$X_{t,1}(y) = X_{t,2}(y) = X_{t,3}(y) = \ldots = X_{t,91}(y), \qquad t = 1, 2, \ldots, T. \tag{1}$$

Because the distribution for the number of PATs in a time interval is unknown, we use the non-parametric Friedman test. Table 3 shows the results of the Friedman test for both 5 and 30 minute intervals. The test results show that only department D has a $p$-value $< 0.05$. This allows a tentative conclusion to be drawn, indicating that only for department D the PATs vary between days. An interaction plot was used to analyze further the outcome of department D (see Appendix A). The interaction plot shows a different pattern for the first 10 days compared to the final 81 days. In fact, after day 10 the interaction plot indicates that there are little differences between the days.

Next, we investigate the daily PAT patterns in more detail using the column vector $\overline{X}(y) = (\overline{X}_1(y), \overline{X}_2(y), ..., \overline{X}_T(y))'$, where $\overline{X}_t(y) = \frac{1}{91}\sum_{d=1}^{91} X_{t,d}(y)$ represents the average number of PATs at time $t$ for department $y$ (here $'$ denotes transpose). To compare the demand patterns of the departments we plotted the relative PAT-values during intervals of 30 minutes in relation to the total number of residents on the vertical axis (see Figures 2, 3 and 4). To smooth out the inconsistencies in the first 10 days of department D, a 20% trimmed mean is used to determine the PAT-values per interval.

The plots show there is considerable fluctuation in PATs over the course of a day. Most of the fluctuation can be explained by the major activities of daily living. Between 0:00 and 6:00 hours there are hardly any activities (some residents have medication though). During the early morning, most residents need help with getting out of bed, washing, and/or dressing. Around lunch and dinner time, there is need for assistance with feeding. Finally, at the end of the day, some of the residents need assistance with getting to bed. Furthermore, it can be observed that during some intervals the peaks exceed 100%. This is because some residents have more than one PAT during some time intervals of 30 minutes. Departments C and E have clearly fewer PATs, but the peaks are comparable to department D. Although peaks are dominated by the activities in daily living, we see that there are some differences in the exact times that peaks occur.

## 3.2 Duration of scheduled care

In addition to the number of PATs, the workload is determined by the duration of care delivery. The average 'expected' care delivery times for C, D and E are 13.2, 10.0 and 12.4 minutes, respectively, with a standard deviation of 11.2, 10.3 and 8.9 minutes. Most care delivery times are short and take 0–5 (or 5–10) minutes. However, there is considerable variation in care delivery times with activities that may take well over half an hour. We refer to Appendix B for figures of the 'expected' care delivery times. There is no data available on realized care durations. For the simulations, we assume that the care durations follow a lognormal distribution, as such a family of distributions is used more often for durations; see e.g. [9] for a call center example.

## 3.3 Workload analysis scheduled care

Based on the PATs and care delivery durations, this subsection examines the workload and current staffing levels. The workload provides the aggregate demand for care. That is, the workload at time $t$ is the number of residents who need care at time $t$ ignoring capacity constraints. As such, it prescribes the required number of care workers at any time if demand would have been met directly (no waiting is allowed). In addition, the workload (i.e. the

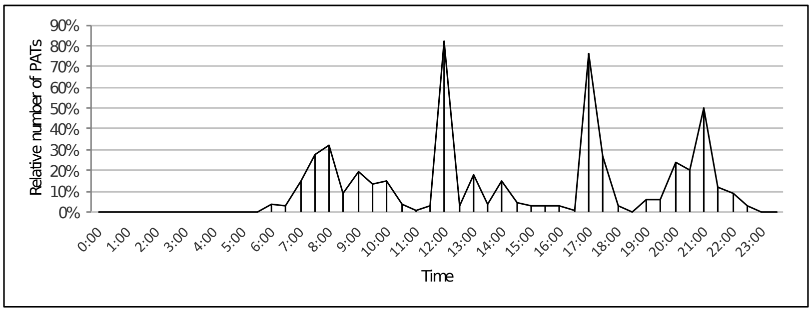


Figure 2: Relative number of PATs for department C

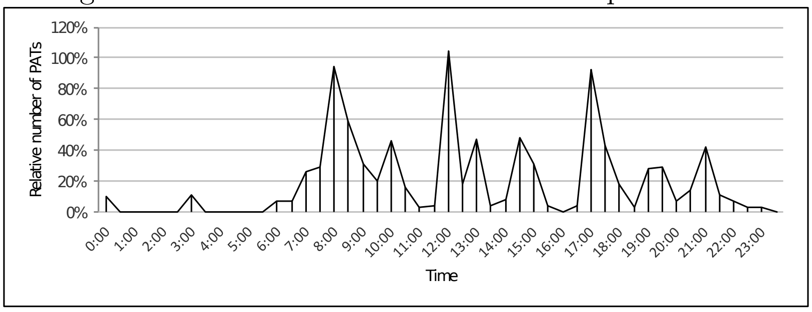


Figure 3: Relative number of PATs for department D

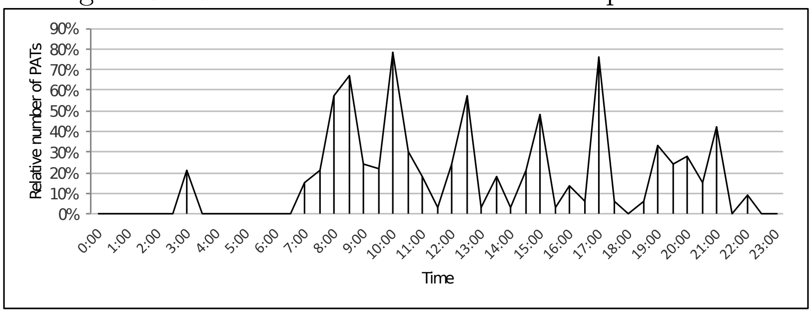


Figure 4: Relative number of PATs for department E

required number of care workers) is compared with the real-life staffing levels. As there is hardly any scheduled care during the night, we consider the time frame 7:00–23:00 hours. Now, we indicate how the workload and staffing levels are determined and then compare them to assess whether there is a mismatch. In contrast to Subsection 3.1, the workload and staffing levels are here sampled at multiples of 5 minutes, i.e. 7:00, 7:05, ....

**Current staffing levels** The staffing levels are determined by the actual worker-to-resident ratios, see Figure 1. Let $\overline{R}_t$ be the care worker-to-resident ratio in interval $t$ and $I(y)$ the number of residents of department $y$. The current available capacity in interval $t$ for department $y$ is then $\overline{C}_t(y) = \overline{R}_t I(y)$, rounded to the nearest integer.

**Workload** Let $L_{t,d}(y)$ be the workload at day $d$ at instant $t$, for department $y$, where $t = 1$ denotes the first epoch of 07:00 hours and $y \in \{\mathrm{C}, \mathrm{D}, \mathrm{E}\}$. Put differently, $L_{t,d}(y)$ represents all running activities plus starting activities at time $t$, day $y$, and department $y$. For department $y$, $S_{t,d}(y)$ denotes the cumulative number of start times at day $d$ until time $t$, i.e. the number of PATs between 7:00 hours and $t$, and $E_{t,d}(y)$ denotes the cumulative number of end times at day $d$ *before* time $t$. Then, $L_{t,d}(y)$ is determined by

$$L_{t,d}(y) = S_{t,d}(y) - E_{t,d}(y), \quad t = 1, 2, \ldots, T, \quad d = 1, \ldots, 91, \quad y \in \{\mathrm{C}, \mathrm{D}, \mathrm{E}\}.$$

We then consider the average workload $\overline{L}(y) = (\overline{L}_1(y), \overline{L}_2(y), \ldots, \overline{L}_T(y))'$, where $\overline{L}_t(y) = \frac{1}{91} \sum_{d=1}^{91} L_{t,d}(y)$ is the averaged workload at time $t$ for department $y$.

The workload (i.e. the number of residents in need of care) and the available capacity are visualized together in Figures 5, 6 and 7. The figures show that, at some moments during the day, the available capacity is insufficient to meet the time preferences of the residents. Hence, residents sometimes have to wait until a care worker is available to provide the necessary care and/or support. Especially during morning care (7:00–9:00 hours) the workload is high. This finding is in line with the findings of [25] and [31]. Furthermore, analysis of the task descriptions corresponding to the (high) peaks in demand shows that some peaks are caused by ‘serving coffee and tea’ and ‘giving medicines’. Observe that we lose some information regarding the workload process by sampling the workload at instants that are multiples of 5 minutes. It should be noted that all PAT’s are multiples of 5 minutes

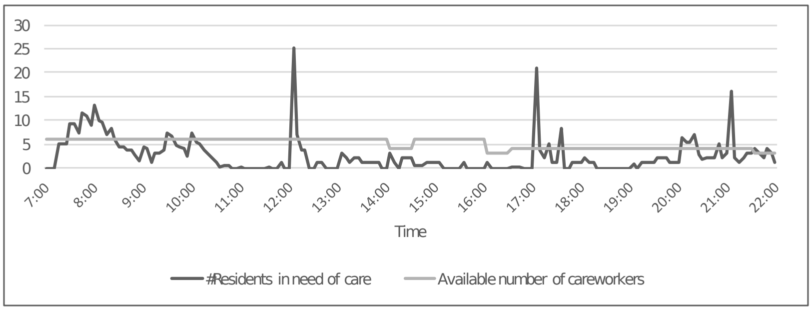


Figure 5: Workload for department C

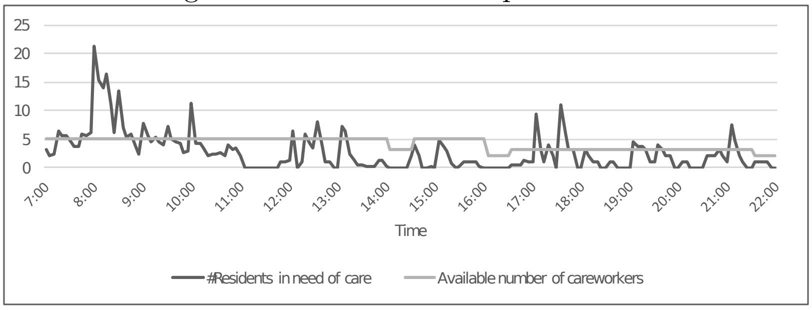


Figure 6: Workload for department D

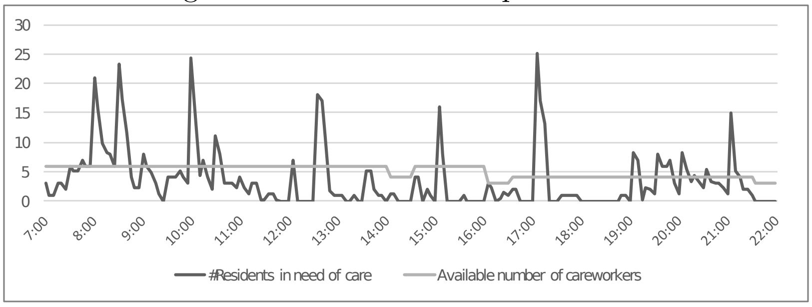


Figure 7: Workload for department E

and the majority of the care durations are also multiples of 5 minutes. As a consequence, the loss of information is relatively small and the workload is largest at moments that are multiples of 5 minutes.

The workload can also be analyzed for random care durations. Let $T_i$ be the starting time of PAT $i$, and let $S_i$ be the random variable denoting the duration of PAT $i$. At time $t \geq T_i$, the probability that PAT $i$ has not yet been finished is a Bernoulli random variable with probability $\mathbb{P}(S_i \geq t - T_i)$. Consequently, we obtain the mean and variance of the workload at time $t = 1, 2, \ldots, T$ as

$$\mathbb{E}L_t = \sum_{i:T_i \leq t} \mathbb{P}(S_i \geq t - T_i), \quad \text{and} \quad \mathbb{V}arL_t = \sum_{i:T_i \leq t} \mathbb{P}(S_i \geq t - T_i)(1 - \mathbb{P}(S_i \geq t - T_i)). \tag{2}$$

A similar argument can be used to analyze the case in which $T_i$ is also a random variable.

### 3.4 Unscheduled care

To make assumptions about the demand and service times of unscheduled healthcare tasks, we use the assumptions and results as presented in [24]. This study shows that it is reasonable to assume Poisson arrivals. In addition, regarding the delivery times of unscheduled healthcare tasks, the analysis presented in [36] indicates that the hyperexponential distribution gives the best fit with the data. In this case we evidently have that the workload due to unscheduled care follows a Poisson random variable at any moment.

## 4 Shift scheduling and waiting times

### 4.1 Shift scheduling

In this section, we schedule shifts that lead to minimal waiting times for scheduled care activities. In the call center domain, determining staffing levels and shift scheduling are often done as two consecutive and separate steps. This is often a reasonable approximation, as call lengths are in the order of minutes such that stationary models per interval are reasonable approximations of system performance. In the nursing home context, there are no obvious queueing models available for the performance analysis of waiting times; this is due to the deterministic and time-dependent arrival pattern of PATs in combination with random activity times and the multi-server nature of the staff. Below, we develop a MILP algorithm for the shift scheduling problem that uses slightly simplified queueing dynamics that are appropriate at a tactical level. Assigning care workers to tasks (operational decision) is outside the scope of this study.

Next, we first outline the performance of a given combination of shifts and then describe the MILP for the optimal shift combination. Let $K$ be the number of possible shift types and let $x_k$ be the number of care workers that are scheduled for shift type $k$. As before, we discretize time such that staffing levels and care activities may only start at epochs $t = 1, 2, \ldots, T$. We denote $a_{tk} = 1$ if shift type $k$ works during interval $t$, and $a_{tk} = 0$ otherwise. Hence, with the matrix $(a_{tk})_{t,k}$ the user defines which types of shifts are possible. The staffing level during interval $t$ is then $c_t = \sum_{k=1}^{K} x_k a_{tk}$, for $t = 1, 2, \ldots, T$. The main assumption that we are going to make is that a team of $c$ care workers work as a single care worker at speed $c$. This assumption is valid when there are sufficient residents (at least $c$) waiting for assistance, and is consistent with the level of aggregation for tactical capacity decisions. The key element is that this assumption avoids that we have to keep track of the individual status of each of the care workers.

The performance measure we are going to consider for the optimization is the backlog at the start of interval $t$, denoted as $Q_t$. During interval $t$, the new amount of work that has to be completed is given by $L_t$. Formally, for department $y$, it holds that $L_t = \int_t^{t+1} \overline{L}_s(y)\mathrm{d}s$, which can easily be calculated from the data. As the workload is more or less constant during 5-minute intervals, $L_t$ for department $y$ can also be approximated by $\overline{L}_t(y)$ as displayed in Figures 5–7. The available capacity is $c_t$, which is determined by $x_k$, $k = 1, \ldots, K$. This provides the following Lindley-type of recusion relation:

$$Q_{t+1} = \max\{Q_t + L_t - c_t, 0\}.$$

For a given realization of $L_u$, $u = 1, \ldots, T$, there is an almost linear relation between $Q_{t+1}$ and $Q_t$ (upto the $\max\{\cdot, 0\}$ operator); this is exploited in the MILP formulation. We are now going to capture the randomness in the workload by introducing scenarios. Specifically, we consider $S$ realizations of the workload pattern $L_t^{(s)}$, for $s = 1, \ldots, S$ and $t = 1, \ldots, T$. These scenarios are generated by simulating the workload process using random care durations and unscheduled care activities. Then, we have the following deterministic recursion for the backlog in scenario $s$:

$$q_{t+1}^{(s)} = \max\{q_t^{(s)} + L_t^{(s)} - c_t, 0\}.$$

As $c_t$ is a linear function of the decision variables $x_k$, we may use standard tricks to incorporate the backlog in an MILP.

| Sets | |
|---|---|
| $S$ | Number of scenarios. |
| $T$ | Number of time epochs. |
| $K$ | Total number of shift types. |
| Input parameters | |
| $a_{tk}$ | 1 if shift type $k$ works interval $t$, 0 otherwise. |
| $L_t^{(s)}$ | Workload for interval $t$ for scenario $s$. |
| $\overline{C}$ | Maximum number of staffing hours. |
| $c_{\min}$ | Minimum required staffing level at any moment. |
| $M$ | Large constant (or weight). |
| Decision variables | |
| $x_k$ | Number of care workers for shift type $k$. |
| $q_t^{(s)}$ | Backlog at instant $t$ for scenario $s$. |
| $c_t$ | Capacity at time $t$, being fully determined by $x_k$. |

Table 4: Notation used for the shift scheduling problem.

We specify the basic shift scheduling problem as follows, with the notation given in Table 4:

$$\text{Minimise} \quad M\frac{1}{S}\sum_{s=1}^{S}\sum_{t=1}^{T} q_t^{(s)} + \sum_{k=1}^{K} x_k \tag{3}$$

$$\text{subject to} \quad c_t = \sum_{k=1}^{K} x_k a_{tk}, \qquad t = 1,\dots,T \tag{4}$$

$$q_{t+1}^{(s)} \geq q_t^{(s)} + L_t^{(s)} - c_t, \qquad t = 1,\dots,T-1, \quad s = 1,\dots,S \tag{5}$$

$$\sum_{t=1}^{T} c_t \leq \overline{C} \tag{6}$$

$$c_t \geq c_{\min}, \qquad t = 1,\dots,T \tag{7}$$

$$q_t^{(s)} \geq 0, \qquad t = 1,\dots,T, \quad s = 1,\dots,S \tag{8}$$

$$x_k \in \mathbb{N}_0, \qquad k = 1,\dots,K \tag{9}$$

Here, $M$ is a positive weight such that the weighted combination of the total average backlog per scenario and the number of shifts is minimized in (3). In our experiments, we let $M$ be large such that the backlog is minimized first and then, if possible, the least number of shifts to obtain this is taken (giving a preference for longer shifts). Equation (4) gives the staffing capacity at time $t$ in terms of the shifts. Equation (5), jointly with the minimization, provides Lindley's recursion for each scenario. Equation (6) provides that the total number of staffing hours per day (excluding the night shifts) does not exceed the budget $\overline{C}$, whereas (7) quarantees that at any moment at least $c_{\min}$ care wokers are available. Finally, (8) provides that backlogs are non-negative and (9) makes sure that an integer number of each shift type is scheduled (including 0).

This basic shift scheduling problem works well for the three departments under study. In some cases, it might be worthwile to modify the formulation, see Remark 1.

*Remark 1* Depending on the practical situation, it is possible to choose alternative formulations. Two key elements are the backlog at the end of the day and the control of shift lengths. First, we note that the former is captured as $q_T^{(s)}$ is taken into account; if the transfer of work to the night shift should be avoided a (soft) constraint on $q_T^{(s)}$ can be included. As the total amount of capacity is sufficient in our case, backlogs are small, also at the end of the day. Secondly, the shift lengths are now controlled by choosing appropriate shift types $x_k$ and in the objective function (3). We note that there are many other choices, e.g. by using a staffing budget and differentiating the costs of certain shift types. An alternative view is to use the current formulation to determine the staffing capacity across the day, and then apply a standard shift scheduling approach using the just obtained staffing capacity per interval.

## 4.2 Waiting time analysis

To gain insight into the performance of the nursing home departments, in terms of meeting the time preferences of their residents, a Python-based simulation model is used. We simulate the operation during daytime, i.e., 7:00-23:00, using a discrete-event simulation approach. For each simulation round, the department under study is 'filled' with new residents. Together with their PATs for scheduled care and corresponding expected service times, each resident is randomly chosen from the empirical dataset of the corresponding department. In addition, demand for unscheduled healthcare tasks is generated according to a Poisson process, with hyperexponential durations (see Subsection 3.4. Furthermore, it is assumed that the provision of care and support is on a FCFS basis. In this fully reactive approach, no distinction is made between scheduled and unscheduled care. The simulation procedure can be described as follows:

- Each day is divided into time buckets of 1 minute ($t = 1, 2, \ldots, 960$).
- Starting at time 7:00 ($t = 1$), for every time step, healthcare tasks based on PATs are assigned to the first available care workers.
- When a resident has a care request during the handling of an earlier PAT (i.e., an additional PAT), this additional request will be adhered to by the same care worker. The waiting time for this additional PAT is set to 0.
- When all care workers are busy, a virtual queue is filled with the remaining activities.
- PATs are served on a FCFS basis.
- When a care worker is providing care during change of shifts, he or she will first complete the task.

As explained above, when all care workers are busy, a virtual queue is filled with the remaining activities. Based on this approach, we calculate the following waiting time measures:

- *Average waiting time at time $t$* – The average time a task has spent in the queue (for all tasks in the queue) at time $t$.
- *Total average waiting time* – The average waiting time over all $t$.
- *Overall 15 minute service level* – The % of care requests for which a care worker was present with the resident within 15 minutes[1].

The performance measures are averaged over all simulation runs.

# 5 Numerical insights

This section presents the optimal staffing patterns for the departments under study, using the MILP algorithm as described in Subsection 4.1. In addition, we explore the impact of the proposed staffing levels in terms of waiting times and service level, whereby the total available staffing hours are kept the same as in the current situation. The total available number of staffing hours per day for departments C, D and E are respectively 82, 64 and 82. For the current staffing pattern we refer to Figure 1.

## 5.1 Experimental design

For our numerical experiments, we made the following choices:

- The ratio of demand for scheduled care tasks to unscheduled tasks is 80-20%. This assumption is based on an estimation made by the managers of the departments under study.
- The demand for unscheduled care follows a Poisson process with rates 3, 4, 4 per hour for departments C, D, E, respectively.
- The care durations for unscheduled care are hyperexponentially distributed with $\hat{p}_1 = 0.10$, $\hat{p}_2 = 0.90$, $\hat{\mu}_1 = 0.11$, $\hat{\mu}_2 = 0.56$ (see [36][24]). These parameters can be interpreted as follows: with probability $\hat{p}_2 =$

[1] In the Netherland, a 15 minute target is regularly used [12, 25].

0.90, the client has a minor request taken on average $1/\hat{\mu}_2 = 1.79$ minutes, whereas with probability $\hat{p}_1 = 0.10$ the client has a large request taking $1/\hat{\mu}_1 = 9.28$ minutes on average.

- For safety reasons, the minimum number of care workers at every moment is 2 (per department), i.e. $c_{\min} = 2$.
- There are 8hr and 4hr shifts, with a preference for 8hr shifts. Breaks are not taken into account.
- Shifts can start at full hours only.
- The care durations of scheduled care follow a lognormal distribution, where the mean corresponds to the estimated activity duration and the standard deviation is taken to be 10 minutes.

## 5.2 Staffing optimization

For department C we experimented with the number of simulated scenarios, see Table 5. A scenario corresponds to a realization of the workload process across a single day. Such a realization is obtained by simulation using the design described above in Subsection 5.1. It can be observed that the number of scenarios has little influence on the results. This can be directly explained by the fact that the variability in the workload is relatively small compared to the mean, which can be analyzed using Equation (2) and Subsection 3.4. In particular, the largest value of the standard deviation of the workload is 1.96 for department C (at 21:05), and only 1% of the instants the standard deviation of the workload exceeds 1.5, whereas the expected workload is considerably larger (see Figure 5). Similar conclusions can be drawn for departments D and E. To be on the safe side, we used 100 scenarios throughout the rest of the experiments. All computations were performed on an Intel i7 CPU with 3.1 GHz and 16 GB RAM. The MILP was solved using the lpSolveAPI package in R.

The results of the MILP approach as presented in Subsection 4.1 can be found in Figure 8. It can be observed that the optimal staffing levels during the day differ substantially from the current levels as presented in Figure 1. Specifically, we see that the staffing during the morning and evening based on the MILP is larger than the current staffing, at the expense of staffing during the late morning and afternoon. Hence, the MILP-based staffing

Table 5: Influence of number of simulated scenarios for department C

| Number of sims | Average waiting time | 15 min. SL | CPU time |
|---|---|---|---|
| 1 | 3.1 | 95.2% | 39 s |
| 5 | 3.0 | 96.2% | 228 s |
| 10 | 3.0 | 95.6% | 468 s |
| 20 | 3.1 | 95.8% | 1192 s |
| 50 | 3.1 | 95.4% | 3181 s |
| 100 | 2.9 | 95.9% | 5946 s |

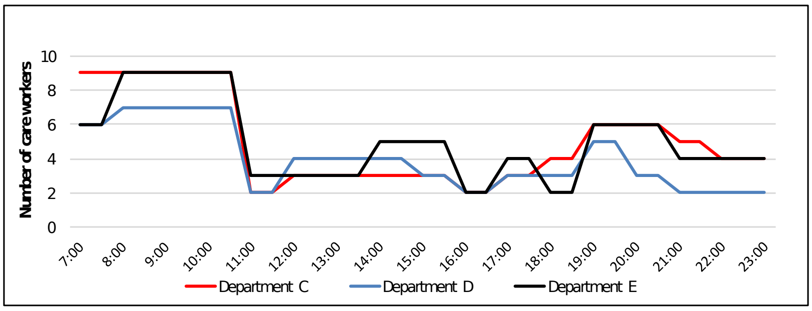


Figure 8: Optimal staffing pattern for departments C, D and E

better follows the demand pattern across the day. However, from Table 7 is can be seen that, in order to adequately follow the demand pattern, a relatively large number of shorter shifts (i.e., 4hr shifts) is required.

## 5.3 Waiting time and service level

The performance of the proposed staffing patterns from Figure 4.1 are compared with the current staffing levels. To do so, we used the simulation from Subsection 4.2. Using a 99% confidence level, and a maximum acceptable confidence interval width of 2 minutes, it was found that a minimum number of 2000 simulation runs is required.

Table 6 shows the main numerical results. We see a substantial improvement for all three departments both in terms of average waiting time as well as in 15 minutes service level. For instance, for department C we see a reduction of the average waiting time of about 70% (from 9.59 to 2.89 minutes) when changing from the current to the MILP-proposed staffing pattern. Also,

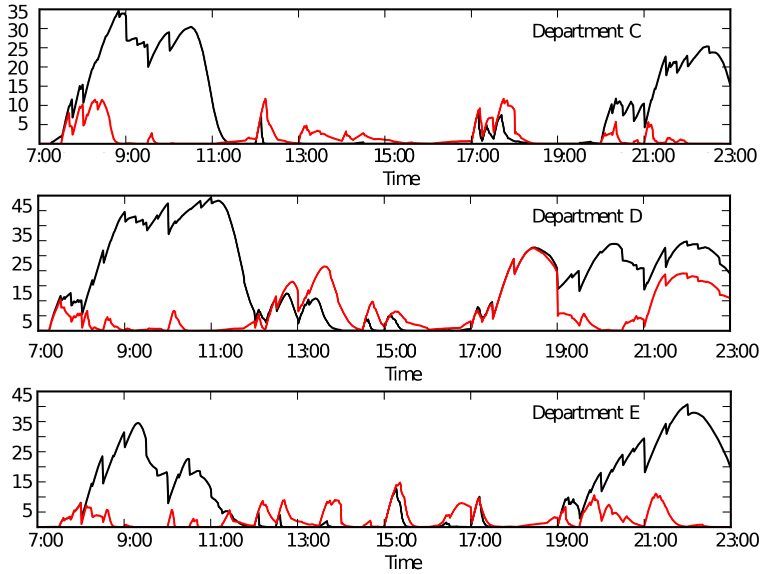


Figure 9: Waiting time (in minutes) during the day: current (black) versus optimal (red) for departments C, D and E

the service level can be increased by more than 12 percentage points for all departments, reaching a 15 minutes service level of at least 95% for departments C and E. The waiting time across the day is visualized in Figure 9. As expected, waiting reduces significantly during rush hours, whereas there is only a slight increase in waiting time during non-rush hours. For example, during the morning rush hour the maximum average waiting times for departments C, D and E drop from 35, 45 and 35 minutes to around 12, 15 and 8 minutes, respectively.

# 6 Conclusions & Discussion

The key message of this paper is that nursing homes could greatly benefit from a more mathematically driven staffing approach. The results show substantial improvements for all three departments both in terms of average waiting time as well as in 15 minutes service level. Especially waiting during rush hours is significantly reduced, whereas there is only a slight increase in waiting time during non-rush hours. Furthermore, as the proposed staffing

Table 6: Overview main results

| Current | Dep. C | Dep. D | Dep. E |
|---|---|---|---|
| Av. waiting time | 9.59 minutes | 14.49 minutes | 10.34 minutes |
| 15 min. SL | 78.5% | 70.9% | 77.4% |
| Optimal | Dep. C | Dep. D | Dep. E |
| Av. waiting time | 2.89 minutes | 7.20 minutes | 3.37 minutes |
| 15 min. SL | 96.0% | 83.5% | 95.0% |

pattern is more balanced, it allows for a more evenly spread workload for the care workers. The workload pattern is obtained using data about residents' prefered activity times. Then, the shift scheduling algorithm is formulated as an MILP relying on elements that finds its roots in stochastic optimization and the celebrated Lindley recursion.

The findings presented in this study are of great societal value as it shows that mathematical driven staffing approaches can greatly support nursing homes in their search for ways to further reduce their costs while maintaining an appropriate quality level of care. In our opinion, an important first step would be to introduce objective performance measures regarding timely delivery of care and support (e.g., mean waiting time and service levels). Nursing homes should have a sufficient information system to make it possible to work with those types performance measures. However, in practice, there is a lack of reliable and valid data. Fortunately, due to developments in technology (e.g., ICT support for domestic tasks, robotics and registration systems), data generation is likely to increase rapidly in the near future. As such, an important future challenge will be to transform these data into tools that support decision making. This will be a challenging task as nursing home processes have many complex characteristics and research on nursing home operations from an applied mathematical perspective is still in its infancy.

Furthermore, the results show that, in order better follow the demand pattern across the day, a relatively large number of shorter shifts (i.e., 4hr shifts) is required. Consequently, in order to implement a mathematically driven staffing approach, sufficient staffing flexibility must be ensured. In this case, we should distinguish between numerical and functional flexibility. Numerical flexibility can be defined as the ability of terms to adjust the number of workers, or the level of worked hours, in line with changes in the level of demand for them [1]. Numerical flexibility could, for example, be achieved by creating a flex pool. A flex pool consists of care workers who are 'on call' and available for work as and when required. Supplementing a core team of full-time care workers with flex pool workers allows nursing home managers to balance their staffing levels better over the course of a day. Functional flexibility can be defined as internal flexibility and refers to the ability of care workers to perform a broader range of tasks, which makes it possible to assign them to different tasks and jobs [1]. In practice, nursing homes often consists of multiple departments. As such, within a full-shift, a care worker could be assigned to multiple departments. In addition, long-term care organization often provide both nursing home care (i.e., intramural care) and home care (i.e., extramural care). It would therefore be possible to divide some of the full-shifts into two shorter intra- and extramural shifts.

This study is limited in scope due to some simplifying assumptions. To model capacity at a tactical level, we neglected some decision at the operational level. For instance, staff utilization and residents' waiting times depend on the different qualification level of the care workers as well as the assignment of care workers to activities (typically leading to NP-hard problems). Also, the results for the current shift scheduling algorithm does not take breaks into account (although that would be an easy adjustment). We envisage that such breaks are typically taken when time permits. In fact, we see in practice that health organizations utilize the flexibility of human capacity to some extent. Finally, we note that the data set is restricted to three departments of a single nursing home. Although the timing of activities of daily living seems rather universal, data from other sources would be valuable.

*Acknowledgements:* The authors would like to thank Ruben van de Geer for support with the modification of the simulation model.

# A Interaction plot department D

The plot below displays PAT levels, which imply that, on some day $d$, a level $x$ is attained when at least one of the $X_{t,d}(y)$ is equal to $x$. Each dot in the interaction plot represents a PAT-level.

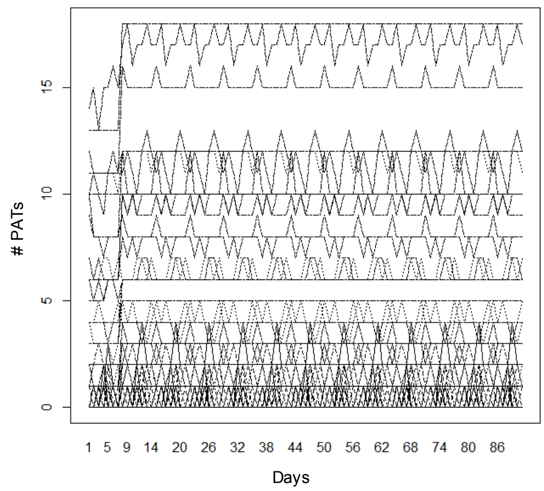


Figure 10: Interaction plot for department D (5 minute intervals)

# B Care delivery durations

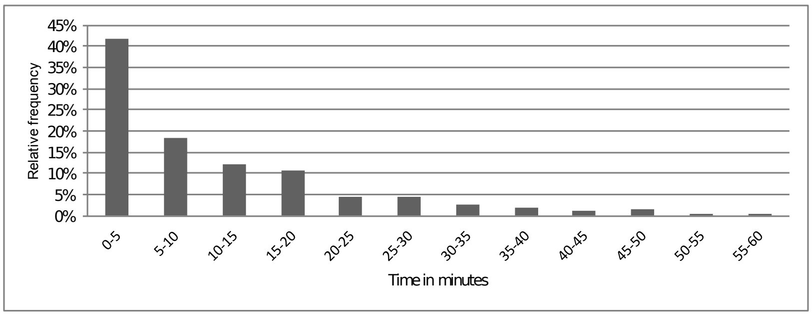


Figure 11: Distribution of care delivery times for department C

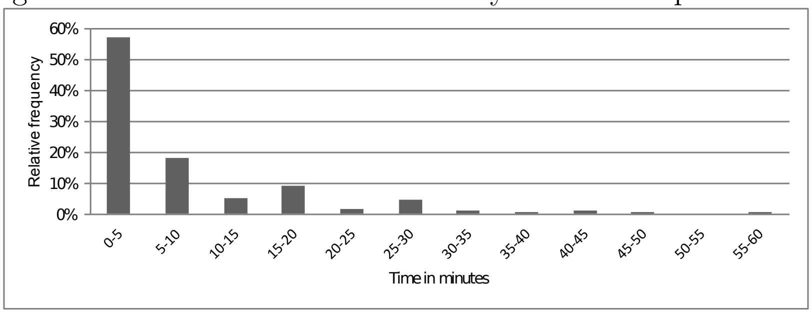


Figure 12: Distribution of care delivery times for department D

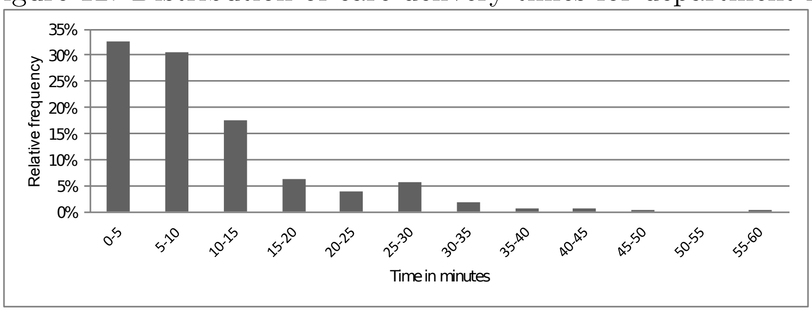


Figure 13: Distribution of care delivery times for department E

Table 7: Required number of shifts

| Dep. | Number of 4hr shifts | Number of 8hr shifts |
|---|---|---|
| C | 12 | 4 |
| D | 8 | 4 |
| E | 14 | 3 |
| Total | 34 | 11 |

# C Number of shifts

Table 7 shows the required number of 4hr and 8hr shifts in the optimal situation.

# References


[1] Atkinson, J.: Flexibility or fragmentation? The United Kingdom labour market in the eighties. Labour and Society **12**(1), 87–105 (1987)

[2] Atlason, J., Epelman, M., Henderson, S.: Optimizing call center staffing using simulation and analytic center cutting-plane methods. Management Science **54**(2), 295–309 (2008)

[3] Avramidis, A., Deslauriers, A., L'Ecuyer, P.: Modeling daily arrivals to a telephone call center. Management Science **50**(7), 896–908 (2004)

[4] Backhaus, R., Verbeek, H., van Rossum, E., Capezuti, E., Hamers, J.: Nurse staffing impact on quality of care in nursing homes: a systematic review of longitudinal studies. Journal of the American Medical Directors Association **15**(6), 383–393 (2014)

[5] Van den Bergh, J., Beliën, J., De Bruecker, P., Demeulemeester, E., De Boeck, L.: Personnel scheduling: A literature review. European Journal of Operational Research **226**(3), 367–385 (2013)

[6] Bhulai, S., Koole, G., Pot, A.: Simple methods for shift scheduling in multiskill call centers. Manufacturing & Service Operations Management **10**(3), 411–420 (2008)

[7] Borst, S., Mandelbaum, A., Reiman, M.: Dimensioning large call centers. Operations research **52**(1), 17–34 (2004)

[8] Bredström, D., Rönnqvist, M.: Combined vehicle routing and scheduling with temporal precedence and synchronization constraints. European Journal of Operational Research **191**(1), 19–31 (2008)

[9] Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., Zhao, L.: Statistical analysis of a telephone call center: A queueing-science perspective. Journal of the American statistical association **100**(469), 36–50 (2005)

[10] Cezik, M., L'Ecuyer, P.: Staffing multiskill call centers via linear programming and simulation. Management Science **54**(2), 310–323 (2008)

[11] Di Giorgio, L., Filippini, M., Masiero, G.: Implications of global budget payment system on nursing home costs. Health policy **115**(2), 237–248 (2014)

[12] Dutch Healthcare Authority: Annual statement of the Dutch healthcare system 2013, from fact to focus. Dutch Healthcare Authority (NZa), Utrecht (2013)

[13] Eveborn, P., Flisberg, P., Rönnqvist, M.: Laps Care–an operational system for staff planning of home care. European Journal of Operational Research **171**(3), 962–976 (2006)

[14] Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: Tutorial, review, and research prospects. Manufacturing & Service Operations Management **5**(2), 79–141 (2003)

[15] Higginson, I., Whyatt, J., Silvester, K.: Demand and capacity planning in the emergency department: how to do it. Emergency Medicine Journal **28**(2), 128–135 (2011)

[16] Hingstman, T., Langelaan, M., Wagner, C.: De dagelijkse bezetting en kwaliteit van zorg in instellingen voor langdurige zorg [Daily utilization and quality of care in long-term care institutions]. NIVEL, Utrecht (2012)

[17] Van der Horst, A., Erp, F.V., Jong, J.D.: Trends in gezondheid en zorg [Trends in health and care], CBS policy letter. CBS, The Hague (2011)

[18] Hulshof, P., Kortbeek, N., Boucherie, R., Hans, E., Bakker, P.: Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. Health systems **1**(2), 129–175 (2012)

[19] Jack, E., Powers, T.: A review and synthesis of demand management, capacity management and performance in health-care services. International Journal of Management Reviews **11**(2), 149–174 (2009)

[20] Lieder, A., Moeke, D., Koole, G., Stolletz, R.: Task scheduling in long-term care facilities: A client-centered approach. Operations Research for Health Care **6**, 11–17 (2015)

[21] Lipszyc, B., Sail, E., Xavier, A.: Long-term care: need, use and expenditure in the EU-27 (No. 469). Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission, Brussels (2012)

[22] Mankowska, D., Meisel, F., Bierwirth, C.: The home health care routing and scheduling problem with interdependent services. Health Care Management Science **17**(1), 15–30 (2014)

[23] Ministry of Economic Affairs: National Reform Programme 2014, The Netherlands. Ministry of Economic Affairs, The Hague (2014)

[24] Moeke, D., van de Geer, R., Koole, G., Bekker, R.: On the performance of small-scale living facilities in nursing homes: A simulation approach. Operations Research for Health Care **11**, 20–34 (2016)

[25] Moeke, D., Koole, G., Verkooijen, L.: Scale and skill-mix efficiencies in nursing home staffing: inside the black box. Health Systems **3**(1), 18–28 (2014)

[26] OECD: OECD Health Statistics 2016 (2017). http://www.oecd.org/els/health-systems/health-data.htm

[27] Rais, A., Viana, A.: Operations research in healthcare: a survey. International Transactions in Operational Research **18**(1), 1–31 (2011)

[28] Rasmussen, M., Justesen, T., Dohn, A., Larsen, J.: The home care crew scheduling problem: preference-based visit clustering and temporal dependencies. European Journal of Operational Research **219**(3), 598–610 (2012)

[29] Sanford, A., Orrell, M., Tolson, D., Abbatecola, A., Arai, H., Bauer, J., Cruz-Jentoft, A., Dong, B., Ga, H., Goel, A., et al.: An international definition for nursing home. Journal of the American Medical Directors Association **16**(3), 181–184 (2015)

[30] Schut, E., Sorbe, S., Hoj, J.: Health care reform and long-term care in the Netherlands, OECD Economics department working paper No. 1010. OECD Publishing, Paris (2013)

[31] Sloane, P., Miller, L., Mitchell, C., Rader, J., Swafford, K., Hiatt, S.: Provision of morning care to nursing home residents with dementia: opportunity for improvement? American Journal of Alzheimer's Disease and other Dementias **22**(5), 369–377 (2007)

[32] Soyer, R., Tarimcilar, M.: Modeling and analysis of call center arrival data: A bayesian approach. Management Science **54**(2), 266–278 (2008)

[33] Spilsbury, K., Hewitt, C., Stirk, L., Bowman, C.: The relationship between nurse staffing and quality of care in nursing homes: a systematic review. International Journal of Nursing Studies **48**(6), 732–750 (2011)

[34] Spruit, M., Vroon, R., Batenburg, R.: Towards healthcare business intelligence in long-term care: an explorative case study in the Netherlands. Computers in Human Behavior **30**, 698–707 (2014)

[35] Taylor, J.: Density forecasting of intraday call center arrivals using models based on exponential smoothing. Management Science **58**(3), 534–549 (2012)

[36] Van Eeden, K., Moeke, D., Bekker, R.: Care on demand in nursing homes: a queueing theoretic approach. Health care management science **19**(3), 227–240 (2016)

[37] Véricourt, F.d., Jennings, O.B.: Nurse staffing in medical units: A queueing perspective. Operations Research **59**(6), 1320–1331 (2011)

[38] VWS: Hervorming langdurige zorg: naar een waardevolle toekomst [Long-term care reforms: towards a valuable future] (Kamerbrief). Ministry of Health, Welfare and Sport (VWS), The Hague (2014)

[39] Yankovic, N., Green, L.V.: Identifying good nursing levels: A queuing approach. Operations research **59**(4), 942–955 (2011)