Logistics for healthy healthcare

Reducing Waiting Lists in Healthcare by Coordinated Placement of Clients

The Dutch healthcare system is facing increasing strain due to population aging

Rebekka Arntzen
TimeLab

Rob van der Mei
Centrum Wiskunde & Informatica, Stochastics Group,
Vrije Universiteit Amsterdam, Department of Mathematics

Summary

The Dutch healthcare system is facing increasing strain due to population aging, persistent financial constraints, and a growing shortage of qualified staff. One of the most visible consequences is the emergence of excessive waiting lists for access to healthcare domains such as elderly care, youth care, disability care and mental healthcare. These challenges are exacerbated by fragmented regional cooperation and the lack of reliable, real-time information on available capacity and waiting times.

This article presents how mathematical modeling, specifically queueing models and assignment optimization techniques, can support more efficient admission processes in healthcare. By combining economies of scale with client preferences, the proposed approach has the potential to reduce waiting times dramatically and improve system-wide flow. We illustrate the benefits with empirical data and demonstrate that both waiting times and logistical performance can be greatly improved without compromising client choice.

Introduction

The healthcare sector in the Netherlands is under increasing pressure due to structural staff shortages, a growing and more complex care demand, and persistent financial constraints. The current shortfall of roughly 70,000 healthcare employees is expected to rise to 270,000 over the next decade (FNV Zorg & Welzijn , 2024). This system-wide pressure translates directly into long waiting lists, particularly for patients with complex care needs or specific preference profiles. While these clients may wait months for appropriate placements, available capacity in other facilities remain unfilled, resulting from regional variation in information exchange, fragmented governance, and the lack of reliable, real-time data on both capacity and waiting times (NOS Nieuws , 2024).

Throughout this article, we focus on the placement of older adults in nursing homes, one of the care domains that face excessive waiting times. The health condition of older adults deteriorates when they remain on waiting lists for prolonged periods and thus do not receive the care they need. Such deterioration increases subsequent care demand, leading to higher pressure on an already strained system. In addition, waiting lists for nursing homes are often managed *de-centrally* rather than in a coordinated regional framework. Clients frequently register with one or two specific, and often popular, preferred nursing homes. As a result, economies of scale are not exploited, and the overall waiting time increases, representing a missed opportunity from both a societal and logistical perspective.

Although this article focuses on elderly care, it is important to stress that similar logistical challenges arise in other domains such as mental healthcare and disability care. In these sectors, waiting lists are frequently long for clients with severe psychiatric disorders or multiple disabilities, who depend on intensive, lifelong support. Meanwhile, regional disparities in capacity utilization persist; some facilities have long queues, while others operate below full capacity due to inadequate information exchange and fragmented referral pathways. The Dutch Healthcare Authority has therefore highlighted the need for better regional coordination and more efficient use of scarce resources across the healthcare system (Skipr , 2025).

Principle of Economies of Scale in Queueing

To illustrate the missed opportunity if the waiting lists are decentralized, we discuss the *principle of economies of scale* in the queueing context. First, we have the following extremes for the organization of waiting lists: (1) every patient takes place in the queue of *only one* healthcare organization or health home (de-centralized) and (2) the allocation to the different organizations or health home is organized by a regional admission portal (centralized). See Figure 1 for an illustration.

Figure 1 Organization of waiting lists (de-centralized vs centralized)

Now, we consider the classical Erlang-C model, with Poisson arrivals (rate λ), exponentially distributed service times (mean β), and c parallel servers, where customers are served in order of arrival. For this system, we define

$$\rho := \frac{\lambda\beta}{c}$$

as the *load per server*. Note that for the stability of the system we must have $\rho < 1$. Figure 2 shows the mean waiting time $E[W]$ as a function of ρ for different values of c, and for β = 1. The plot illustrates that as the number of servers increases, the mean waiting time decreases for a given load *per server*.

In the context of nursing homes, this means that if multiple nursing homes operate a joint waiting list system, the mean waiting time decreases for a given load per bed. In other words, for a given customer demand λβ and total number of beds in a region c, regional cooperation can simultaneously achieve high bed utilization (ρ), needed as a revenue source for nursing homes, and low waiting times. Thus, waiting times in nursing homes can be greatly reduced by using a centralized and coordinated admission system that exploits economies of scale.

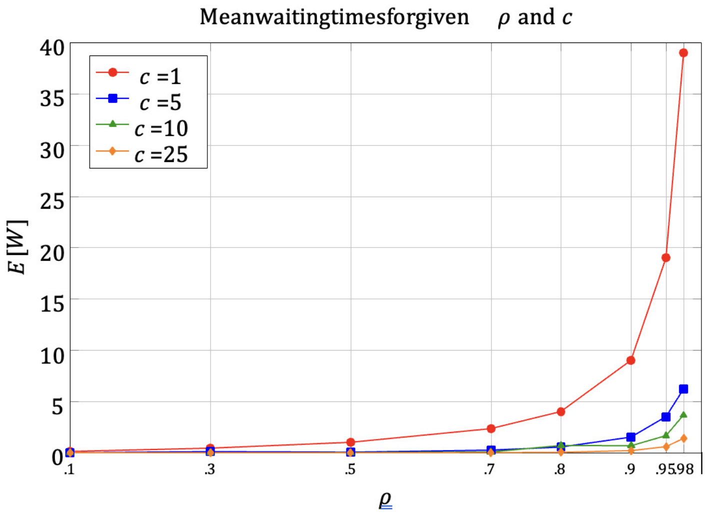

Figure 2 Mean waiting time (MWT) for different values of load per server (ρ) and c parallel servers (with mean service time β = 1).

Assignment of Older Adults to Nursing Homes

Knowing that working together can reduce waiting times, we looked for a way to combine two goals that are often seen as conflicting: making sure that older adults can go to a place that fits their individual preferences, while at the same time using the available beds in a region as efficiently as possible. To address this, we developed an assignment model that helps decide which older adult should be placed in which nursing home, and when (Arntzen e.a., 2024).

Loosely speaking, the model works as follows. At any given moment, there are older adults waiting for placement. Some live at home, others may be in temporary facilities. At the same time, there are nursing homes with different numbers of available beds. Each older adult has individual preferences: they may prefer certain nursing homes, want to stay close to family, or avoid long travel distances. The model takes these preferences into account.

To compare different placement options, the model uses a score (also called "utility") that reflects how well a particular nursing home matches a particular older adult. A higher score means a better match. This score can depend on several factors, such as:

- whether the nursing home is the person's preferred location,
- whether the person is currently waiting at home or in another facility,
- how long the person has already been waiting,
- and any special care needs that must be matched with the right facility.

The goal of the model is to find the combination of placements that gives the highest total score in the system while respecting two important rules:

- No nursing home receives more patients than it has beds available.
- Every older adult can be assigned to at most one nursing home at a time.

In other words, the model looks for the "best possible assignment" given the preferences, available capacity, and current waiting situation.

An important strength of this approach is that it can handle different types of patients and regional situations. For example, one can create different "preference profiles". Some older adults may want to be placed urgently and accept several locations, while others prefer to wait longer until a bed becomes available in their first-choice nursing home (Arntzen e.a., 2022). The model can incorporate these different behaviors without difficulty.

It is also worth noting that although this model is based on mathematical methods, it remains straightforward and scalable in practice. It can make decisions for dozens, or even hundreds, of clients and multiple nursing homes in a fraction of a second. While such assignments could still be made manually in small situations, the model becomes especially useful in larger regions with many waiting clients and many facilities, where manual coordination is difficult and time consuming.

Overall, the assignment model provides a systematic and transparent way to allocate scarce nursing home beds while respecting patients' personal preferences and reducing waiting times.

Waiting-time performance for the Assignment Model

In this section we investigate the waiting-time performance for difference bed allocation policies.

Toy example

First, we show by a toy example the working of our assignment model. In Figure 3, we compare our model to the extreme situations: a common queue (without preferences) which is very efficient and *separate queues* (where each patient takes place on the queue of their preferred nursing home). We considered an example of four nursing homes, with twenty beds each; we refer to Arntzen e.a. (2024) (Table 3) for a complete specification of the parameter values. Figure 3 shows the expected waiting times as a function of the led utilization.

The results in Figure 3 confirm that the assignment model does what it is constructed for: (1) the assignment algorithm leads to a strong reduction in waiting times, and (2) the inclusion of individual preferences increases the waiting times only mildly.

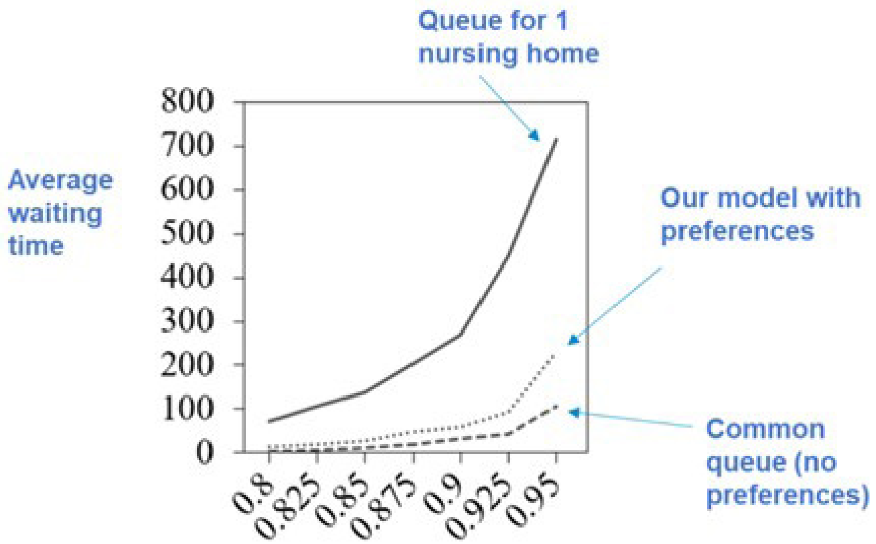

Figure 3 Results of the allocation model: MWT as a function of the bed utilization, for a small-scale example with four nursing homes with twenty beds each.

Usecase: placement at nursing homes in region of Amsterdam

To estimate how well the assignment model works in practice, and what reductions in waiting time can be achieved by applying the model, we analyzed data from recent registrations and placements of clients at a number of nursing homes in the region of Amsterdam. We compared (1) the actually realized waiting times under the current method of assigning clients to nursing homes with (2) the waiting times that *would have been achieved* if the assignment algorithm had been used.

Current practice of placement

In current practice, patients are allowed to sign up for a single preferred nursing home and wait at home until placement. However, if waiting times become excessive (typically fifteen months), patients may opt for immediate placement due to impatience. Since bed availability is not centrally managed and information is limited, a regional office manager contacts nursing homes individually to inquire about temporary residency options, ceasing calls once a suitable bed is found. Under this policy, patients are initially placed on a waiting list for their preferred nursing home and later added to a secondary waiting list shared among all regional nursing homes. Throughout, this placement policy is referred to as the current practice (CP) model.

Classification of patients

For the allocation model (AM), we distinguish between two classes of clients: fast placement (FP) and preferred placement (PP). FP patients prefer quick placement in a (temporary) nursing home, which PP patients opt to wait for availability at their preferred nursing home.

Specification of the parameters

For the case study, we focus on a specifc patient group with severe somatic conditions, categorized as ZZP6 and ZZP8 in the Dutch healthcare system. These patients exhibit physical symptoms and require intensive care. In Amsterdam, there are 39 nursing homes with dedicated somatic departments accommodating these patients. Utilizing non-public micro-data from Statistics Netherlands, we derive essential parameters. The arrival rate (1.25 patients per day) is determined from nursing allowance requests, while the capacity (780 beds) reflects the maximum number is simultaneous residents. The average length of stay is (666 days) is estimated from patient declaration data. Also, we assume that 50% of the client population prefers fast placement, and the other 50% preferred placement. We refer to Arntzen e.a. (2024) for more details on the specifics parameter choices.

Comparison of policies

Table 1 shows the results for different bed-allocation policies, for the case where each client has a single preferred nursing home. Here, ”Full sharing” refers to the case where patients are placed on a first-come-first-serviced basis, regardless of their preferences. ”Allocation model” refers to the combined performance of FP and PP clients under the allocation model. Separate queues” refers to the policy in which clients only subscribe to a single nursing home.

Table 1 Comparison of waiting times and bed occupancy for different bed allocation policies for teh Amsterdam region (for the case of a single preference).

Policy	MWT till placement (days)	MWT till preferred (days)	bed occupancy (%)
Full sharing	40	40	100
Allocation model	47	191	99
Current practice	234	257	72
Separate queues	256	256	67

The results in Table 1 show that by using the allocation model instead of current practcie the MWT till placement is reduced from 234 days to 47 days, and the MWT till placement at the preferred nursing home is reduced from 257 to 191 days. Also, the bed occupancy is shown to increase from 72% to 99% by using the allocation model, which is an important benefits for the revenue of the nursing homes.

Next, we investigate the the performance improvements that can be obtained when patients have more multiple preferred nursing home. Table 2 show the results.

The results in Table 2 show that that when patients indicate more than one preference, the MWTs decrease dramatically compared to the current practice.

Table 2 Comparison of waiting times and bed occupancy for the allocation model depending in the number of preferences for the Amsterdam region. Note that the values are based on simulation results with a confidence bound, which makes that the results might slightly fluctuate.

Policy	MWT till placement	MWT till preferred	bed occupancy
	days	days	%
AM with 1 preference	47	191	99
AM with 2 preferences	40	98	99
AM with 3 preferences	43	61	99
Current practice	234	257	72

Conclusions & Recommendations

Our analyses lead to the following conclusions (see Arntzen et al. (2024, 2022) for more details):

1. **Waiting-time reduction. By applying the assignment algorithm, the waiting times can be strongly reduced (typically by a factor 2 to 4) compared to the current practice.**

2. **Inclusion of preferences. Taking individual client preferences into account increases the waiting times only mildly.**

It is important to stress that the assignment model finds the right balance between, on the one hand, the economies of scale that regional cooperation offers, and, on the other hand, the individual preferences of the clients. It is also interesting, and good news, that the waiting times when using the assignment model are almost equal to the waiting times that would be achieved if the economies of scale were fully exploited at the expense of the preferences of the individual clients.

This work shows that mathematical tools can effectively reduce waiting times in elderly care while preserving client preferences. Empirical results indicate that the proposed allocation method substantially outperforms current practice without lowering regional capacity utilization or limiting client choice. More broadly, waiting lists in elderly care are not only medical or social issues but also logistical ones. Inefficiencies arise from decentralized decision-making, limited information sharing, and the absence of automated allocation tools. Addressing these logistical gaps can therefore complement traditional policy or staffing interventions.

Specifically, the study leads to the following practical recommendations:

1. **Implement regional coordination mechanisms.** Centralized or federated admission systems can unlock economies of scale and contribute to more balanced utilization of beds across facilities.

2. **Adopt data-driven allocation models.** Utility-based assignment algorithms allow for systematic and transparent matching of clients to nursing homes while respecting individual preferences.

3. **Invest in real-time information infrastructures.** Accurate and up-to-date insight into waiting times, available beds, and client profiles is essential for effective coordination.

Evaluate applicability beyond elderly care. Given the general nature of the model, similar approaches may benefit other domains with structural waiting list challenges, such as youth care, mental healthcare, and disability care.

In conclusion, mathematical modeling and operations research can meaningfully improve access to care. Integrating data, preferences, and capacity constraints into allocation decisions leads to more transparent, equitable, and efficient healthcare systems.

References

Arntzen, R. J., Bekker, R., and van der Mei, R. D. (2024). Preference-based allocation of patients to nursing homes. *Operations Research for Health Care*, 42, 100442.

Arntzen, R. J., Bekker, R., Smeekes, O. S., Buurman, B. M., Willems, H. C., Bhulai, S., and van der Mei, R. D. (2022). Reduced waiting times by preference-based allocation of patients to nursing homes. *Journal of the American Medical Directors Association*, 23(12), 2010–2014.

FNV Zorg & Welzijn. (2024). Personeelstekort in de Zorg & Welzijn verviervoudigt. (In Dutch)

NOS Nieuws. (October 2024). Wachttijden ggz niet goed vastgelegd, zeggen onderzoekers. (In Dutch)

Skipr. (2025). NZa: zorgkantoren moeten meer doen om wachtlijsten terug te dringen in de langdurige zorg. (In Dutch)