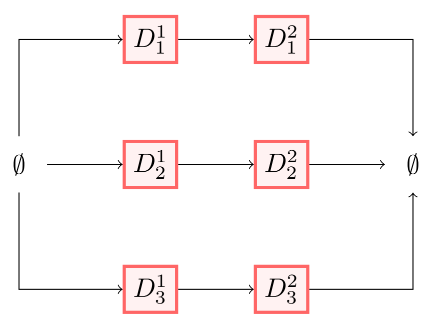
Optimal Resource Allocation in Healthcare

David Maxen

January 2024

1 Initial plan

The concept of optimal resource allocation in healthcare revolves around the strategic distribution of limited resources such as medical personnel, facilities, equipment, and financial funds to meet the demands of patient care, research, and administrative functions. Striking the right balance between supply and demand is essential not only for managing healthcare costs but also for enhancing patient outcomes and overall system performance. This project will focus on the optimal allocation of surgeries using a rule based model.



This diagram represents the flow of patients through a hospital. Here we consider 3 diseases, D_1, D_2, D_3, each with two stages of severity, e.g. D_1^1, D_1^2, giving us 6 types for our patient objects. Patients arrive in the first stage of a disease according to a Poisson process, after some time progress to the next stage, and then at some point are removed from the system. In addition to our patient object, we have another basic object, interventions. We model these as also arriving according to a Poisson process. There are two types of interventions, I^1, I^2, which represent treatments to patients in the first and second stage of illness respectively. We model the interaction of interventions and patients with the following counting rules.

$$D_j^1 + I^1 \to \emptyset \quad \forall j \in \{1,2,3\}$$
$$D_j^2 + I^2 \to D_j^1 \quad \forall j \in \{1,2,3\}$$

The project will begin by analysing this system using techniques from the course as well as queuing theory. We will explore how varying different rate parameters affects system performance. We will then explore how policies can be implemented, and how they affect performance. There are many interesting avenues to explore after this, for example when interventions can have different outcomes with different probabilities. Finally, we will apply the model to real world data, and compare our model to current state of the art resource allocation models.

2 First Implementation

When implementing the model, I first simplified to having only two different diseases, as this removes complexity while preserving the optimal resource allocation element. We can then write down our rule based model:

Variable	**Meaning**
D_1^1	disease 1, stage 1
D_1^2	disease 1, stage 2
D_2^1	disease 2, stage 1
D_2^2	disease 2, stage 2
I^1	stage 1 treatment
I^2	stage 2 treatment

$$\emptyset \rightarrow^{\lambda_1} D_1^1 \tag{1}$$
$$\emptyset \rightarrow^{\lambda_2} D_2^1 \tag{2}$$
$$D_1^1 \rightarrow^{r_1^1} D_1^2 \tag{3}$$
$$D_2^1 \rightarrow^{r_2^1} D_2^2 \tag{4}$$
$$D_1^2 \rightarrow^{r_1^2} \emptyset \tag{5}$$
$$D_2^2 \rightarrow^{r_2^2} \emptyset \tag{6}$$
$$\emptyset \rightarrow^{i^1} I^1 \tag{7}$$
$$\emptyset \rightarrow^{i^2} I^2 \tag{8}$$
$$D_1^1 + I^1 \rightarrow^{p_1^1} \emptyset \tag{9}$$
$$D_2^1 + I^1 \rightarrow^{p_2^1} \emptyset \tag{10}$$
$$D_1^2 + I^2 \rightarrow^{p_1^2} \emptyset \tag{11}$$
$$D_2^2 + I^2 \rightarrow^{p_2^2} \emptyset \tag{12}$$

We now discuss these rules and their drawbacks. Inherent in these stochastic rule-based systems are some key assumptions: that events are independent of each other and time.

Rules (1, 2) describe the Poisson processes, of rate λ_1, λ_2 respectively, that model the influx of patients with stage 1 diseases into the hospital. This assumption is fairly weak in the setting where you have a large population and a relatively small number of patients have the illness. Examples where this assumption could break down is if there is a strong seasonal aspect to the disease (could have λ_i as a function of time), or if it is an infectious disease, as this violates our independence assumption.

Rules (3, 4) describe how a disease develops from stage 1 to stage 2. This assumption can be fairly weak, for example for cancer development [Ta98]. However if patients develop to stage 2 one month after being in stage 1, this would be a very bad assumption. However, due in part to diagnosis delays, I believe this to be a reasonable assumption. The same is true for the assumptions behind rules (5, 6), which account for the rate that individuals with stage 2 diseases die.

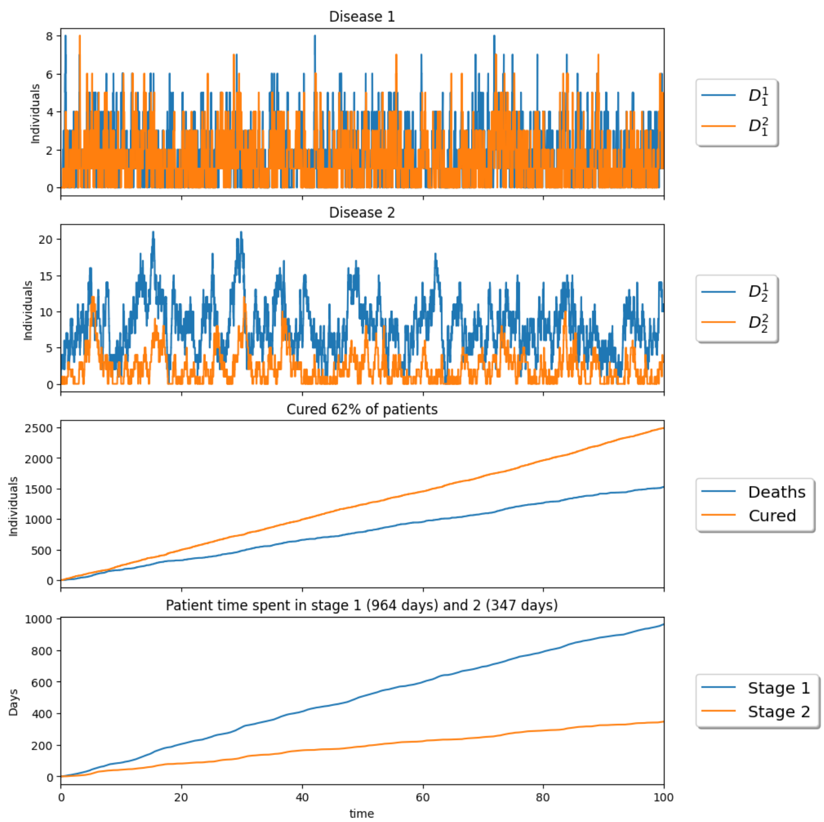
Rules (7, 8) impose that surgeries, or treatments, become available according to a Poisson process with rates i^1, i^2. This may go against the time independence assumption: if no opportunities have become available in the last week, this could be correlated with either a high (e.g. compensating) or low (e.g. machinery broken) number of opportunities the next week. Another reason this may be a bad assumption is if surgeries are carried out in batches, for example an entire day is allocated to one specific type of surgery.

Rules (9, 10, 11, 12) are govern the interactions between surgeries and diseases. Here we have that if a stage i treatment is used on a stage i patient then they are completely cured. This is an interesting place to add complexity; for example for the stage 2 treatment you could potentially have 4 outcomes: disease cured, disease regresses to stage 1, disease remains in stage 2 (no change), or death (e.g. risky operations). This would add very interesting dynamics that are relevant from a treatment allocation perspective.

The rates p_i^j are the product of two terms: time taken between a treatment becoming available and it taking place, and ratio of treatments allocated to disease j according to policy that we are analysing. For example, if the policy is to always treat patients with disease 1 first, then the ratio for disease 2 would be an indicator function. Alternatively, you could allocate surgeries in a ratio. As such, the rates p_i^j could be functions of the current number of patients.

3 Prioritise Urgency

The first question I asked was whether more urgent diseases should be prioritised, as is done in England's National Health Service (NHS) [Eng24]. This can be implemented in the model by setting the rate of progression of disease 1 to be far higher than that of disease 2. The following graph is an example realisation of this system using the Gillespie algorithm using parameters $\lambda_1 = \lambda_2 = 20$, $i^1 = 15$, $i^2 = 10$, $r_1^1 = r_1^2 = 10$ and $r_2^1 = r_2^2 = 1$. Here, the ratio is left as equal: the chance of a treatment being assigned to a patient with disease 1 is proportional to the total number of patients with disease 1.



The following test was ran to compare three ratios: for each ratio, run the Gillespie algorithm 50 times, and for each return the time spent with the disease and the percent of patients cured. This gives us the following data.

$D_1 : D_2$	Days in stage 1	Days in stage 2	Percent cured
1 : 9	800	402	62%
1 : 1	962	359	62%
9 : 1	1256	430	62%

The first interesting result is that the percentage of patients cured is always 62%. This is because of two features of this setting. Firstly, each disease ends in two possible scenarios, treatment or death. Secondly, the number of people with the diseases far outweighs the number of treatments available. In this case, the total number of people saved is simply equal to the number of treatments that were available over the 100 day period; it doesn't matter whether they were used on patients with disease 1 or 2.

However, there is a significant difference with regards to another metric, namely the total number of patient days spent with a disease in stage 1. Here, we can see that the chance of treating patients with disease 1 increases, the total patient time spent in stage 1 increases. One way of understanding this is the following situation: patient 1 has disease 1, patient 2 has disease 2, and there is one treatment available. If you treat the patient with disease 1, then the patient with disease 2 spends a long time in stage 1, then a long time in stage 2, before dying. On the other hand, if you treat the patient with disease 2, the patient with disease 1 quickly progresses to stage 2 then dies. That is to say, there is a trade-off between total patient time alive and total patient time with the disease. Prioritising urgency results in a higher total time alive, but there is an ethical dilemma to consider. If disease 2 is extremely painful, is it immoral to treat the patient with disease 1 and let them suffer for a prolonged period of time? This is equivalent to the ethics of euthanasia.

We now explore a new setting my adding the condition that there is a maximum number of treatments that you can carry over. This is true in many real world settings; if last week a specialist was available to perform a surgery but there were no patients requiring it, then if a new patient arrives this week, you can't still use that opportunity. The following table was created using following average results from 500 Gillespie runs at each of the three ratios, with arguments $\lambda_1 = \lambda_2 = 1$, $i^1 = 2$, $i^2 = 0$, $r_1^1 = r_1^2 = 10$, $r_2^1 = r_2^2 = 0.1$ and a maximum number of treatments of 1. In particular, treatments become available at the same rate as patients are diagnosed.

$D_1 : D_2$	Days in stage 1	Days in stage 2	Percent cured
$1 : 9$	60	53	66%
$1 : 1$	60	54	67%
$9 : 1$	68	62	68%

Notably, in this setting the percent cured increases as you prioritise the more urgent treatment. This is intuitive - if you expect more treatments to be available soon, then one should always treat the most urgent patient first. In conclusion, urgency is a key factor to consider when assigning care, however there are ethical considerations in some settings.

4 The Effect of Stochasticity

The Gillespie algorithm provides independent samples from the stochastic process defined by our system. However, when the number of patients and treatments tends to infinity, the computational burden increases. However, the laws of mass action also come into play, allowing us to model our system using a system of ordinary differential equations (ODE's). In this section we compare these two approaches. Firstly, we write down the ODE's corresponding to our system.

$$\frac{\partial}{\partial t}D_1^1(t) = \lambda_1 - r_1^1 D_1^2(t) - p_1^1 D_1^1(t) I^1(t) \tag{13}$$

$$\frac{\partial}{\partial t}D_2^1(t) = \lambda_2 - r_2^1 D_2^2(t) - p_2^1 D_2^1(t) I^1(t) \tag{14}$$

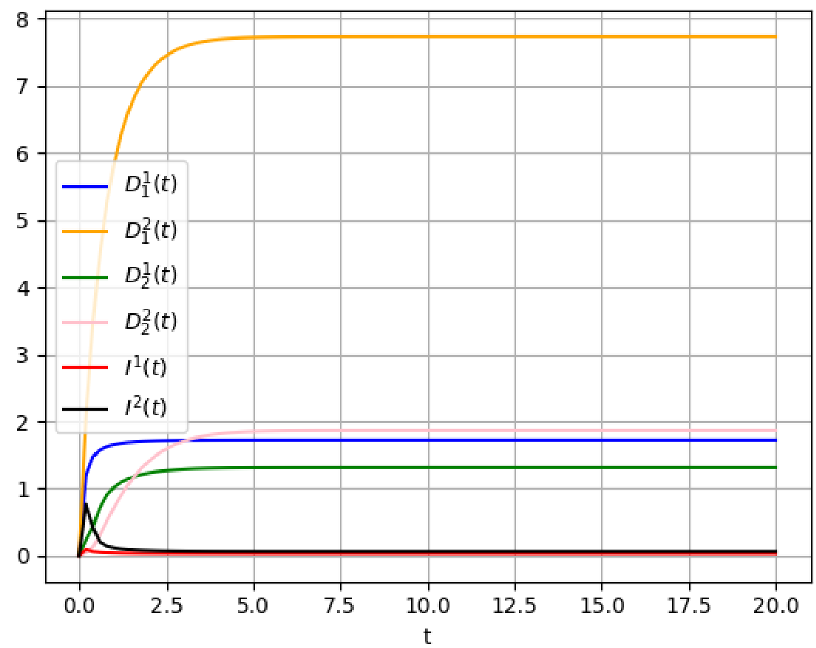
$$\frac{\partial}{\partial t}D_1^2(t) = r_1^1 D_1^1(t) - r_1^2 D_1^2(t) - p_1^2 D_1^2(t) I^2(t) \tag{15}$$

$$\frac{\partial}{\partial t}D_2^2(t) = r_2^1 D_2^1(t) - r_2^2 D_2^2(t) - p_2^2 D_2^2(t) I^2(t) \tag{16}$$

$$\frac{\partial}{\partial t}I^1(t) = i^1 - p_1^1 D_1^1(t) I^1(t) - p_2^1 D_2^1(t) I^1(t) \tag{17}$$

$$\frac{\partial}{\partial t}I^2(t) = i^2 - p_1^2 D_1^2(t) I^2(t) - p_2^2 D_2^2(t) I^2(t) \tag{18}$$

We can analyse this system numerically using an ODE solver.



Comparing our three ratios, again in the case that disease 1 progresses faster than disease 2, we get the following results. These suggest that in the mass action case, the best policy is also to prioritise urgency.

$D_1 : D_2$	Percent cured
1 : 9	63.18%
1 : 1	63.49%
9 : 1	64.21%

References

[Ta98] Prevost TC and Launoy G et al. "Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches". In: (1998). DOI: 10.1093/oxfordjournals.aje.a009687.

[Eng24] NHS England. *Frequently asked questions about surgical prioritisation.* 2024. URL: https://www.england.nhs.uk/coronavirus/secondary-care/other-resources/clinical-prioritisation-programme/clinical-prioritisation-programme-frequently-asked-questions/frequently-asked-questions-about-surgical-prioritisation/#how-many-priority-categories-are-there-for-surgical-validation (visited on 01/02/2024).