Scheduling of Nurses at a Community Health Care Centre

Alan Bi, Trevor Dallow, Samantha Zimmerman

Department of Mathematics, Simon Fraser University,
Surrey, BC, Canada

Abstract

This project looks at optimizing nurse schedules at a community health centre in Vancouver, Canada. The centre provides primary care to individuals who are unable to access conventional fee-for-service primary care. Nurses at the Centre have reported that their schedules do a poor job of matching patient arrival levels to the walk-in components of the clinic. This paper describes patient flow through specific components of the centre as a scheduling problem in the form of a Mixed Integer Linear Program (MILP). Having collaborated with the centre's management, we built diagrams outlining the flow and used electronic medical record (EMR) data and staff opinion to estimate the parameters needed in the model. The model is written in Python and solved using Gurobi to approximate the optimal schedule for nurses to maximize the amount of time spent with patients. The centre has expressed interest in implementing the generated schedules for nurses.

1 Introduction

1.1 Background

We collaborated with a local community health centre, which offers primary care services for people who cannot access traditional fee-for-service care, especially youth and patients undergoing Opioid Replacement Therapy (ORT). The centre has both capacity and accessibility issues, and nurses at the centre have reported that their original schedules are ineffective in meeting patterns of patient demand. Most Responsible Physicians (MRP's) at the centre maintain limited and fixed work schedules, with nurses meeting remaining demand. An effective schedule for nurses is not obvious, and must take into account both scheduled appointments and walk-in services. The aim of this project is to determine a nursing staff schedule for the centre that most effectively meets patient demand by maximizing the amount of time spent with patients.

1.2 Significance

Patients suffering from homelessness and complex conditions like mental illness and substance use disorder are frequently refused help from fee-for-service physicians and rely on clinics like the centre. With limited resources and staff to service their clients, the centre is facing pressure to meet growing demand while maintaining their level of service. Nurses at the centre have reported that their original schedules are ineffective in meeting patterns of patient demand. The centre differs from many other primary care service providers as they have a diverse offering of staff due to the complex needs of its patients. This includes medical professionals, social workers, and dietitians; who all need to be effectively scheduled. Nurse scheduling impacts the local health authority, the centre's staff, and the centre's patients. More efficiently scheduled staff would improve wait times for patients and make better use of employee time and skills, as well as funding. Staff would experience less idle time given the optimized schedule computed from our model.

1.3 Scope

The centre offers 32 different health services, however, our analysis focuses on the adult primary care core services offered by MRP's, nurses, and social workers (SW's) during morning booked appointments and afternoon walk-ins, as well as the services offered by the ORT team, consisting of 2 MRP's and 1 nurse, during afternoon walk-ins as the centre's management expressed particular interest in examining these services.

The purpose of the model discussed in this paper is to find weekly schedules for RN's and LPN's engaged in adult primary care and ORT services at the centre which best meet patterns of patient demand, while preserving current staffing levels. The schedule produced is not tailored to individual nurses, but rather is built for a group of indistinguishable staff; so it is left for the centre to determine which staff take which shifts. the centre adult primary care and ORT services employs 9.7 full time equivalent (FTE) of MRP's, 10 FTE of nurses, and 3 FTE of SW's. Taking this staff level into account, it changes which shifts are filled on each day of the week to generate a nursing staff schedule which maximizes the expected amount of time staff spend with patients. Patient demand levels can fluctuate on an hourly, daily, weekly, and seasonal basis. The model is built to address hourly and daily fluctuations, but not weekly or seasonal ones.

The model considers the patient flow process from when a patient finishes checking in for their booked or walk-in appointment, to when they finish their time with staff. EMR data was used to estimate the parameters of the underlying distributions for this process. Using this foundation we formed a MILP optimization model to determine optimal schedules.

2 Model and Data

2.1 Patient Flow Process

In collaboration with the centre's management, a set of process maps was developed that describe the flow of patients through a subset of the centre's services. The descriptions below will aid the reader in understanding the processes of a few of the services offered by the centre.

2.1.1 Booked Appointment Flow Process

Weekday morning operating hours are reserved for booked appointments. Patients can book appointments in advance to be seen by a health care provider, consisting of any one of MRP's, SW's, RN's, or LPN's. Patients currently registered with the centre are scheduled for appointments lasting 30 minutes, and unregistered patients are scheduled for one hour; using the additional time to register the patient into the centre's system. See appendix A for a visualization of this process.

2.1.2 Morning Walk-in Flow Process

Given that the centre dedicates its morning operating hours to booked appointments, staff will not see walk-in patients during these hours unless the patient is assessed to be in need of urgent care. For instances of walk-in patients during morning operating hours, a clinical assistant (CA) will make an initial decision as to whether a patient is urgent and if not, advises the patient to return during the afternoon for walk-in care. Otherwise, if the CA feels that the patient's needs are urgent, then the patient will be sent to a triage nurse who will again assess the urgency of the patient's state. Patients deemed urgent by the triage nurse will be have an appointment booked with the first available MRP, whereas non-urgent patients are requested to return during regular afternoon walk-in hours. At any given time, there is an RN or LPN tasked with triage duties in addition to their routine work, having the expectation of halting their current work in order to perform triage duties if the need arises. As such, it is important for our model to consider morning walk-ins as these instances are impacted by how nurses are scheduled at the centre. This process is not represented by the model as there was little data on urgent patients and so it could not be meaningfully interpreted. See appendix A for a visualization of this process.

2.1.3 Afternoon Walk-in Flow Process

The most complex scenario described by our model are afternoon walk-ins which span from 13:15 to 20:30 during weekdays and 10:00 to 18:00 for weekends. As with morning walk-ins, there exist distinctions between urgent and non-urgent patients; sharing the same assessment process as morning walk-ins. Patients deemed as urgent will be placed first in the queue to see an MRP, whereas non-urgent patients will be queued on a first come, first served basis. There are three initial treatment streams available to non-urgent patients: visiting an MRP, visiting a nurse, or undergoing ORT. Patients initially seeing MRP's may be referred to either a nurse or SW, otherwise they will leave the centre after seeing an MRP. For patients who see a nurse first, the nurse may consult an MRP, if needed, and return to treating the patient, or finish servicing the patient on their own. Finally, a patient may be in need of ORT, where MRP's, RN's and LPN's work as a unit to treat the patient, after which the patient will exit the centre. ORT services are provided daily from 13:15 to 16:30 and so interacts with regular walk-in services, which is why our model includes ORT services in afternoon walk-in. Given that the model does not consider the urgent patient distinction, the assessment process is not represented and all patients are prioritized on a first come, first serve basis. See appendix A for a visualization of this process.

2.2 Data Acquisition and Analysis

2.2.1 Acquisition

The EMR was the primary source of data for this project. Information about each appointment slot and staff schedules for the last week of October were obtained by a research assistant who manually entered data into excel spreadsheets. Any remaining data considered in this project was acquired through the centre signage and internal documentation of staff estimates, based off of their passive observations on the proceedings of the centre.

2.2.2 Composition

Each data set acquired spans the course of one week such that it contains the last Wednesday of the month and avoids holidays, taken from months for the year 2017. The reason for this choice is to maintain consistency between months in that they are all roughly the last week of the month. For the last weeks of November , August, and September (Nov. 26 - Dec. 2, Aug. 27 - Sept. 2, Sept. 24 - Sept. 30), only patient check-in times were collected so that a larger amount of observations could be analyzed to estimate arrival statistics.

For the week of October 22 to 28, a dataset for staff schedules and a dataset for patient appointments were obtained.

- Staff Schedules
  - Staff type a given schedule relates to
  - Start/end times of shifts and breaks
  - Whether staff member was designated as clinician of the day or triage nurse
  - Start/end times of instances where provider was working but unavailable to treat patients
- Patient Appointments
  - Expected start/end times of appointments
  - Whether appointment is booked or an instance of a walk-in
  - Patient check-in and seen times and time at which the provider closed the patients case (indicating that treatment is complete)

2.2.3 Analysis

2.2.3.1 Arrival Rates

One component that describes demand for service at the centre are walk-in arrival rates. In order to incorporate arrival rates in the model, the arrival rates were fit by a non-homogeneous Poisson process.

In order to fit the arrivals as a non-homogeneous poisson process, the NHPoisson package for R was used. The NHPoisson package [2] uses Maximum Likelihood Estimation (MLE) to compute parameters from the input data using the fitPP function. Only arrivals for the week of October (22-28) was considered as it could not be determined whether an arrival was a walk-in or not for the data from other months. The results from the fitPP function yield an intensity function for the arrival rates. Given this intensity function, the rate of arrival can be specified for a specific time. The intensity function was estimated using a cubic spline via the splinefun function in R. The resulting spline was integrated over 15 minute intervals to obtain the expected number of arrivals for each 15 minute interval.

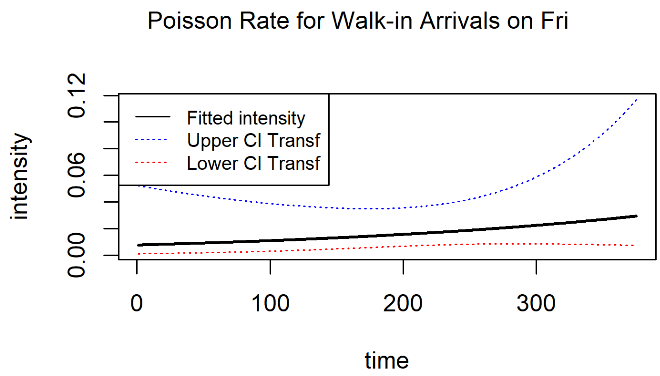


Figure 1: Plot of the intensity function for Friday (Oct. 27, 2017)

2.2.3.2 Booked Appointment Demand

We estimated demand patterns for booked appointments based on the booked appointments in the data. Demand was estimated by examining the count of morning check-ins, grouped into 15 minute intervals, for each day of the week. The number of appointments for each interval and weekday was averaged over 4 weeks for each interval-weekday grouping, (with a weighting depending on the total number of visits that week compared to the October week), to yield an estimated expected demand for booked appointments for each group. Only appointments during morning operating hours on weekdays were considered. It is assumed that morning operating hours are from 8:30 to 12:45 and that all check-ins during this time on weekdays are for booked appointments. Though some walk-in appointments occur during these hours, these walk-ins are typically urgent cases, happen infrequently, and are recorded as booked appointments. Note for the week in October, greater detail for booked appointments was included in the data and so the registered appointment start time was used instead of check-in times for this week.

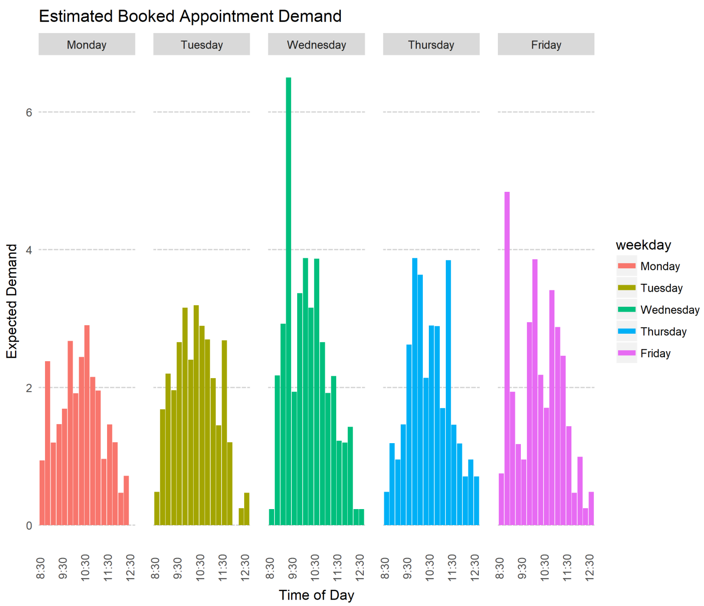


Figure 2: Estimated Booked Demand Levels per time interval

2.2.3.3 Lateness for Booked Appointments

Though not considered by the model, the lateness of patients for booked appointments for the week of October was analyzed. The data for booked appointments in the week of October contained 180 observation, two of which were filtered out given that they indicated the patient was late in excess of 90 minutes (the two removed observations were late by over 2.5 hours and 10 hours). The remaining 178 observations all fall within 90 minutes of the scheduled appointment start time. Normal and student-t distributions were considered when fitting the data using the MLE method with R's MASS package [4]. Quantile-Quantile (Q-Q) plots were used to compare the goodness of fit of each distribution, as well as examining the fitted density plots against the empirical density.

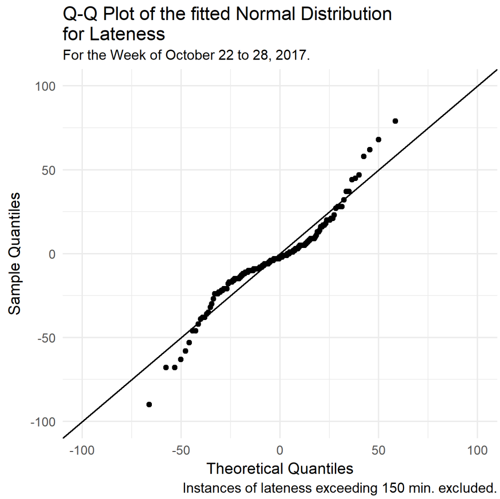


Figure 3: QQ-plot of normal distribution.

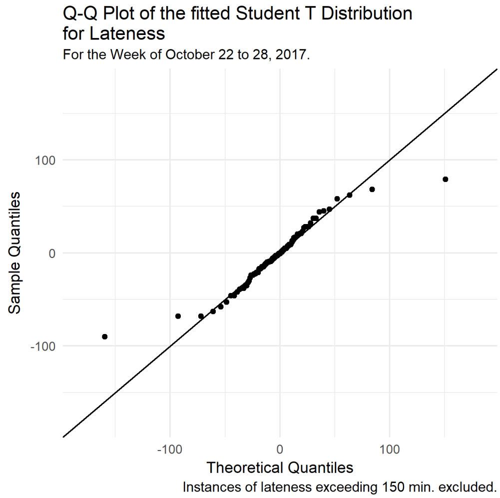


Figure 4: QQ-plot of student-t distribution.

By inspecting these plots, the student-t distribution appears to better fit the distribution of the actual data. It was found that on average, patients checked-in 3.9 minutes prior to their scheduled appointment start. Though a more useful metric would be to consider how many patients check-in 10 minutes late or earlier. The reason for this is that the centre will call the patient twice, at least 10 minutes apart to see whether the patient can still attend the appointment before canceling the appointment (though appointments are not always canceled after this point as is shown by the data). For this analysis, on-time will be considered as patients arriving at most 10 minutes late as this is assumed to be the earliest possible point that the centre will consider canceling the appointment, given they make the first call at the scheduled start time. It was found that 83.7% (n = 149) of patients (whose appointment were neither canceled nor were they a no-show) arrive at most 10 minutes late for the week examined in October.

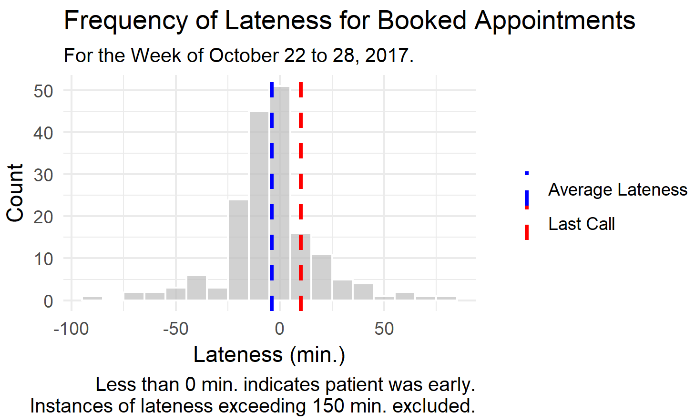


Figure 5: Plot of the frequency of lateness

2.2.3.4 Service Rates

Using the data, it was possible to estimate actual service times of patients for the week in October. The expected service times are used in the MILP. The MLE method was used from the MASS package in R to fit the data, where exponential and lognormal distributions were considered to estimate service times.

Service times from the data were computed by taking the minimum of the difference between the close and seen times of a patient with a provider and the difference between the seen time of one patient and the following patient, for the same provider.

Let i denote patient i, and ordered in terms of their seen time with provider, j denote provider j, $l_{i,j}$ be the time patient i is seen by provider j, $c_{i,j}$ be the time that provider j closes patient i's case and $v_{i,j}$ be the service time of provider j with patient i.

$$v_{i,j} = min\{c_{i,j} - l_{i,j}, l_{i+1,j} - l_{i,j}\}$$

Only comparing the difference between close and seen times of the same patient are inaccurate given the data set as the centre's leadership staff had indicated that it is possible for providers to close cases at the end of the day instead of right after finishing treatment with the patient. As a result, we considered the next seen time as an indicator of when the provider finished treating their last patient. Service times were filtered such that only service times greater than four minutes and less than 91 minutes were considered. Times over 90 minutes may be indicative of a provider who closed patients at the end of day, and was not able to see patients for a period of time (eg. going on break or other obligations), causing both service time computations to be large. Times below 5 minutes may be as a result of an error in recording a patient arrival and so when a provider closes a case, check-in and seen times take close to or the same time as the close time.

Data for appointment times that were scheduled for 30 minutes of both walk-in and booked is used for estimating service times in the model. There exist other scheduled lengths, most prominently 60 minutes (but still less so than 30), but recall 60 minute appointments are typically used for unregistered patients. Given that the data does not indicate whether a patient is registered, 60 minute appointments are possible for existing patients (though infrequent) and there were too few data points to conduct analysis on for scheduled lengths other than 30 minutes, this model assumes that all appointments are scheduled for 30 minutes. After filtering, MRP (n = 138, mean = 27.8 min, see B), nurse (n = 30, mean = 24.0 min, see B) and ORT (n = 33, mean = 25.9 min, see B) service types were considered, with all remaining services falling short of 30 observations and so not useful to fit these groups.

For all service types considered, each data set was fitted using exponential and lognormal distributions and compared using Q-Q plots as well as examining the empirical density with the fitted density plots. The lognormal yields a closer fit to the data for all groups over the exponential function, though the fit is still somewhat unconvincing. Ashton found that typical distributions were ineffective at fitting service times, but the end result from using any distribution was roughly the same [1].

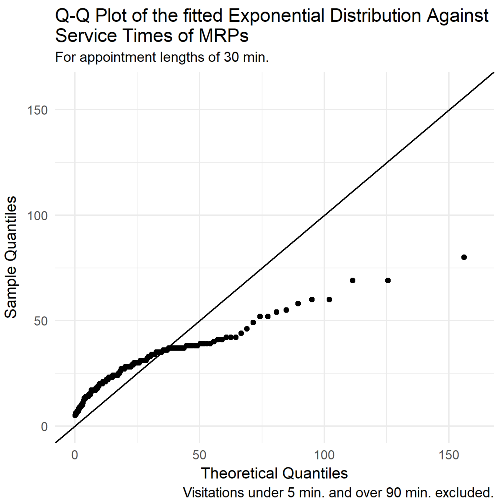


Figure 6: QQ-plot of exponential distribution.

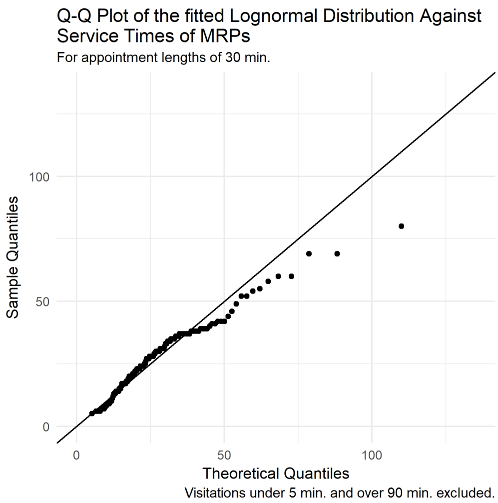


Figure 7: QQ-plot of lognormal distribution.

2.3 Simplified Model of Patient Flow

The model used for patient flow in the optimization part of this study is simplified from the process described above and built on a series of assumptions about how patients arrive and spend their time at the centre. The model doesn't explicitly take into account the involvement of multiple staff within one primary care visit, for example when a physician is consulted during a nursing visit. However, this is indirectly accounted for in our estimation of service times. We assumed that there was no difference in service time distribution for different staff members doing the same types of appointments. The majority of the time slots for patients in the data were 30 minutes long, which was also identified by the centre's management as a baseline for booked appointment lengths, so the model assumes that all time slots are 30 minutes in length. In our analysis of the estimated service times for 30 minute nursing slots, the lognormal distribution seemed to be the most appropriate distribution of the ones considered. Using the data to find the best lognormal parameters we determined a mean of 34.0 minutes and a standard deviation of 17.64 minutes.

With respect to patient walk-in arrivals we assumed that they followed a non-homogeneous Poisson process (NHPP) that had different rates for each time of day and day of the week. This modelling choice does rely on the general assumptions of a NHPP; namely that different time intervals have independent number of arrivals, that the probability of two arrivals in a short interval is negligible, and that the probability of a single arrival in a short increment is proportional to the length of the interval and the instantaneous intensity at the start of the interval [5]. While in theory patients could arrive as couples or families, the Poisson process assumptions seemed reasonable and have been made in many similar studies [1].

We assumed that the proportion of the total number of patients visiting the centre's adult primary care walk-in was relatively constant. Because of the additivity of NHPP's this means that the probability of an arriving patient requiring specific services is also these same probabilities. Using total appointment and walk-in counts we estimated 38.6% of walk-in's are for nurses, 20.4% of booked appointments are with RN's and 7.7% of booked appointments are with LPN's.

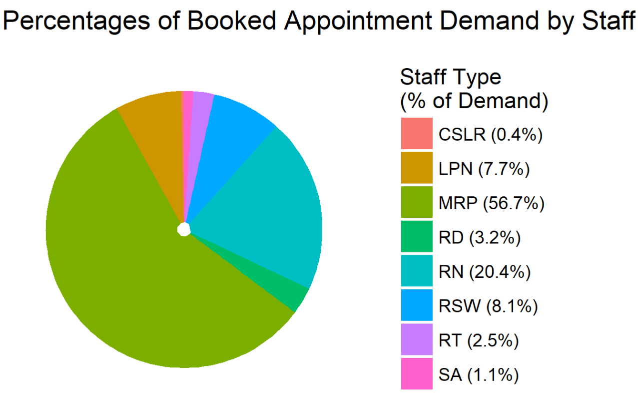


Figure 8: Proportion of types of patient demand for booked appointments.

In order to determine the expected demand for a given time interval, the intensity function of the non-homogeneous Poisson process was estimated using a cubic spline via the splinefun function in R. The resulting spline was then integrated over 15 minute intervals to estimate the expected number of arrivals for each 15 minute interval.

The following list details the modelling assumptions made.

- All appointment lengths are 30 minutes in length
- Time nurses spend with patients doesn't depend on type of nurse, time of day, or other factors
- Time nurses spend with patients follow a lognormal distribution
- The probability of a patient needing specific services is the same for a given walk-in arrival, it does not depend on the time of day
- The arrival rates for walk-in adult primary care follow NHPP's that varies by day of the week
  - Different time intervals have independent number of arrivals

  - The probability of two arrivals in a short interval is negligible
  - The probability of a single arrival in a short increment is proportional to the length of the interval and the instantaneous intensity at the start of the interval

- The demand patterns for booked appointments follow daily and weekly trends
- The proportion of demand for booked appointments which is for specific staff types is relatively constant
- There are no booked appointments on weekday afternoons (after 13:15)
- There are no walk-in designated slots available on weekday mornings
- During open hours there is a relatively constant amount of demand for a triage nurse except for during afternoon walk-in's when it is slightly higher
- During ORT walk-in hours (13:15 to 16:30) there is one nurse on walk-in duty
- Patients are treated individually
- Nurses only work 8 or 8.5 hour shifts (corresponding to 7.5 and 8 hours of paid work, respectively)
- The last half hour of a Nurse's shift is not available to be booked by patients
- Nurses who start before 10:00 attend a staff huddle from 13:00 to 13:30
- Nurses have one 30 minute break and two 15 minute breaks which they take at predictable times

2.4 Mixed Integer Linear Programming Formulation

2.4.1 Description

The goal of the optimization model is to maximize the expected amount of time RN's and LPN's spend with patients. Because it keeps the total staff hours and shift numbers constant, this also maximizes the proportion of these staff's time which is spent with patients. And because the distribution of time spent with each patient is kept the same, the number of patients on average should also be maximized. The model varies how many RN's and LPN's are on a given shift on a given day, constraining the total number of shifts by staff type and hour length to be the same as in the actual schedule used for the week given.

The Parameters, Variables, Objective Function, Constraint and Summary subsections below describe a generic model with parameters which could be changed to include more nurse types, demand levels, days and shifts considered, or different planning intervals; and also gives the values used here in our implementation of the model for the given situation and scope.

Note that the formulation takes into account two levels of demand; that which is specific to a particular type of nurse and that which can be met by any type of nurse. In our situation we segregated whether demand was specific or generic by time of day (with respect to walk-in hours), so our implementation was able to be simplified. However, the formulation given doesn't have to have this condition. Note also that if one were to increase the number of nurse types, the model given would only take into account demand for all nurses, or one specific nurse type. Such a formulation is possible, but the equations are not detailed here.

2.4.2 Parameters

Let K be the set of nurse types. Here $K = \{1, 2\}$ where 1 corresponds to RN and 1 corresponds to LPN.

Let S be the set of shifts in consideration, and σ be the number of shifts considered, so $S = \{1, 2, ..., \sigma\}$. Here the model considers ten types of shifts so $S = \{1, 2, ..., 10\}$, with shift numbers defined below.

Shift Number	Start Time	End Time	Break 1	Lunch & Admin	Break 2
1	8:30	17:00	10:00	12:00	14:00
2	8:30	4:30	10:15	12:00	14:15
3	9:00	17:00	10:30	12:00	14:30
4	9:00	5:30	10:45	12:00	14:45
5	9:30	18:00	11:00	12:00	15:00
6	10:00	18:00	12:15	13:00	16:00
7	10:00	6:30	12:45	13:00	16:15
8	12:00	20:00	14:15	16:30	19:00
9	12:15	20:45	14:30	16:30	19:15
10	12:15	20:45	14:45	16:30	19:30

Let $\nu_{k,8}$ be the total number of 8 paid hour shifts of nurse type k, and $\nu_{k,7.5}$ be the total number of 7.5 paid hour shifts of nurse type k, for all $k \in K$.

Let D be the set of days in consideration, and δ the total number of days, so $D = \{1, ..., \delta\}$, where day 1 is Sunday, and the rest follow. Here $\delta = 7$ and $D = \{1, 2, ..., 7\}$.

Let R be the set of shifts when different days are taken into account. So $R = \{1, 2, ..., \delta\sigma\}$, where element $r \in R$ corresponds to shift ceil(r/δ) on day r mod δ (where day 0 is day δ). Let R_8 be the index set of shifts with 8 hours of paid working time, and $R_{7.5}$ be the index set of shifts with 7.5 hours of paid working time.

Let T be the set of all time segments considered, and τ be the length of time segments considered in minutes. Let Ω be the number of minutes between the start of the first shift considered and the end of the last shift considered. Then $T = \{1, 2, ..., \delta\Omega/\tau\}$ and element $t \in T$ corresponds to the (t mod Ω/τ) $+1$ the shift on day ceil($t\tau/\Omega$). Here we consider shifts from 8:30 to 20:45, and use $\tau = 15$, so $T = \{1, ..., 343\}$.

Let W be the index set of time intervals which are considered walk-in periods at the centre.

Let $c_{t,r}$ be 1 when a nurse on shift r would be available to see patients at time interval t, for all $t \in T$ and $r \in R$. These values take into account start and end times, breaks, some staff meetings, and the last half hour of every shift being administrative time.

Let μ be the expected amount of time a patient spends with nurses for adult primary care services. Let $d_{t,k}$ be the expected number of patients requesting services specific to nurses of type k in time interval t, and d_t be the expected number of patients requesting generic nurse services (regardless of type) in time interval t, for all $t \in T$ and $k \in K$.

Let $b_{t,k}$ be a modifier for levels of nurse type k at time interval t to be added to scheduled staffing levels, for all $k \in K$ and $t \in T$. This allows the formulation to take into account nurses scheduled on duty but who are designated for ORT or triage duties. Note that one nurse during this week had a special shift that started and 6:30 and incorporated outreach and shelter work not considered in the scope of this project. The formulation does not include this shift as a variable, and has reduced total shift counts as appropriate. However, this special shift did have booked appointment components as well, so additional nurse hours are added during those times in the $b_{t,k}$ values.

Let $e_{t_1,t_2,k} = 1$ if a patient with demand for nurse type k, at time interval t_1, is willing to wait until time interval t_2 for service, and zero otherwise; for all $t_1, t_2 \in T$ and $k \in K$. Note that $e_{t_1,t_2,k}$ is always zero if $t_1 > t_2$, and always one for $t_1 = t_2$. We set the wait limit to be 90 minutes in our implementation. Let $f_{t_1,t_2,k}$ be the same in value as $e_{t_1,t_2,k}$ except for the case when $t_1 = t_2$, in which case it is zero. And let e_{t_1,t_2} and f_{t_1,t_2} be defined similarly for generic nurse demand.

2.4.3 Variables

Let $x_{r,k}$ be a non-negative integer variable representing the number of nurses of type k on shift r, for all $k \in K$ and $r \in R$. We denote X as the matrix containing these values.

Let $y_{t,k}$ be a non-negative continuous variable representing the expected amount of time spent with patients who require the specific services of nurse type k over all nurses of type k during time interval t, for all $t \in T$ and $k \in K$.

Let y_t be a non-negative continuous variable representing the expected amount of time spent with patients requiring generic nursing needs over all nurses of any type in K during time interval t, for all $t \in T$. We denote Y as the matrix containing the $y_{t,k}$ and y_t values.

2.4.4 Objective Function

The goal of the model is to maximize the expected amount of time nurses spend with patients. So our objective function, to be maximized, is

$$f(X,Y) = (\sum_{k \in K} \sum_{t \in T} y_{t,k}) + (\sum_{t \in T} y_t)$$

2.4.5 Constraints

1. The total number of scheduled nurses, by type of nurse and type of shift, must not be increased from the original schedule.

$$\sum_{r \in R_8} x_{r,k} \leq \nu_{k,8} \tag{1}$$

$$\sum_{r \in R_{7.5}} x_{r,k} \leq \nu_{k,7.5} \tag{2}$$

2. The expected amount of time spent with specific patients by nurses of any type k, during any time interval t, must be less than or equal to the amount of time nurses of that type are available during time interval t.

$$y_{t,k} \leq \tau c_{t,r} x_{r,k} + \tau b_{t,k} \qquad \forall t \in T, k \in K \tag{3}$$

3. The expected amount of time spent with general patients by all nurses with type in K, during any time interval t, must be less than or equal to the amount of time those are available during time interval t. The amount of time available is equal to the time provided by all nurse types subtract time spent on nurse type specific duties.

$$y_t \leq \tau \sum_{k \in K} c_{t,r} x_{r,k} + \sum_{k \in K} \tau b_{t,k} - \sum_{k \in K} y_{t,k} \qquad \forall t \in T \tag{4}$$

4. The expected amount of time spent with nurse specific patients by nurses of any type k, during any time interval t must also be less than or equal to the amount of nurse specific time demanded for that type and interval. The amount demanded at that time interval includes unmet demand from previous intervals within the range of the patient waiting

limit.

$$y_{t,k} \leq \mu \sum_{q \in T} e_{q,t,k} d_{t,k} - \sum_{q \in T} f_{q,t,k} y_{q,k} \qquad \forall t \in T, k \in K \tag{5}$$

5. The expected amount of time spent with patients requiring generic nursing, during time interval t, must be less than or equal to the amount of generic nurse time demanded for that interval. This demand includes unmet demand from earlier intervals within the patient wait time limit.

$$y_t \leq \mu \sum_{q \in T} e_{q,t} d_t - \sum_{q \in T} f_{q,t} y_q \qquad \forall t \in T \tag{6}$$

2.4.6 Summary

Altogether, the model formulation becomes:

$$\text{Maximize } f(X,Y) = (\sum_{k\in K}\sum_{t\in T} y_{t,k}) + (\sum_{t\in T} y_t)$$

Subject to:

$$\sum_{r\in R_8} x_{r,k} \leq \nu_{k,8}$$

$$\sum_{r\in R_{7.5}} x_{r,k} \leq \nu_{k,7.5}$$

$$y_{t,k} \leq \tau c_{t,r} x_{r,k} + \tau b_{t,k} \qquad \forall t \in T, k \in K$$

$$y_t \leq \tau \sum_{k\in K} c_{t,r} x_{r,k} + \sum_{k\in K} \tau b_{t,k} - \sum_{k\in K} y_{t,k} \qquad \forall t \in T$$

$$y_{t,k} \leq \mu \sum_{q\in T} e_{q,t,k} d_{t,k} - \sum_{q\in T} f_{q,t,k} y_{q,k} \qquad \forall t \in T, k \in K$$

$$y_t \leq \mu \sum_{q\in T} e_{q,t} d_t - \sum_{q\in T} f_{q,t} y_q \qquad \forall t \in T$$

$$x_{r,k} \geq 0, \qquad x_{r,k} \in \mathbb{Z} \qquad \forall k \in K,\ r \in R$$

$$y_{t,k} \geq 0, \qquad y_{t,k} \in \mathbb{R} \qquad \forall t \in T,\ k \in K$$

$$y_t \geq 0, \qquad y_t \in \mathbb{R} \qquad \forall t \in T$$

2.5 Simulation Model

For the purposes of validating the model output, AnyLogic PLE was used to build a simulation to compare the current nurse schedules and the model output. The average wait time of a patient in a queue is used for comparison between the centre's current staff schedule for nurses and the schedule obtained by optimizing our model. For the purpose of the simulation, we assume patients arrive according to a non-homogeneous Poisson process with rates determined by expected demand during each 15 minute interval, and only arrived for services during the hours of operations stated by the centre. Patients were assigned to a provider depending on the services available, and the expected proportion of patients seeing a provider-type. Nurses doing primary care work have a service rate according to a lognormal distributed service time in minutes with mean-log 3.37 and a standard-deviation-log of 0.62. Whereas nurses performing ORT duties have a lognormal distributed service time with mean-log 2.99 and standard-deviation-log 0.75.

3 Results

3.1 Optimization Results

We implemented the optimization model for our situation using the Gurobi optimizer accessed through python [3]. Our implementation has 428 continuous variables, 140 integer variables and 992 constraints. It runs in an average of 6.98 seconds with standard deviation 2.75 seconds. The original schedule is shown in Figure 8, and The resulting optimal schedule is shown in Figure 9. The expected amount of time spent with patients for the optimized schedule is 42.56 hours over the whole week. This corresponds to an adult primary care patient time to staff hours ratio of 15.9%. Note that this number does not include time spent with patients in triage or in ORT. The original nurse schedule was infeasible in our formulation. When taking out constraint components related to ORT and Triage duties, the original schedule yielded 40.45 hours of patient time and a 15.15% adult primary care patient time to staff hours ratio.

3.2 Simulation Results

Using the centre's current staff schedule, 100 trials were run; keeping track of the amount of time each patient spends waiting in a queue. To ensure independent patient wait times, one sample from each of the 100 trial runs was chosen according to a uniform distribution, for each of booked appointment, afternoon walk-in, and ORT to find the respective expected wait times. This process was then repeated using the MILP's outputted schedule. The MILP schedule results in lower waiting times for patients for all services considered. The following table provides a summary of the average wait times for each service is calculated in minutes.

Service	Current Schedule	MILP Schedule
Booked Appointments	1.27	0.49
Walk-ins (PC)	13.72	11.87
Walk-ins (ORT)	15.40	12.13

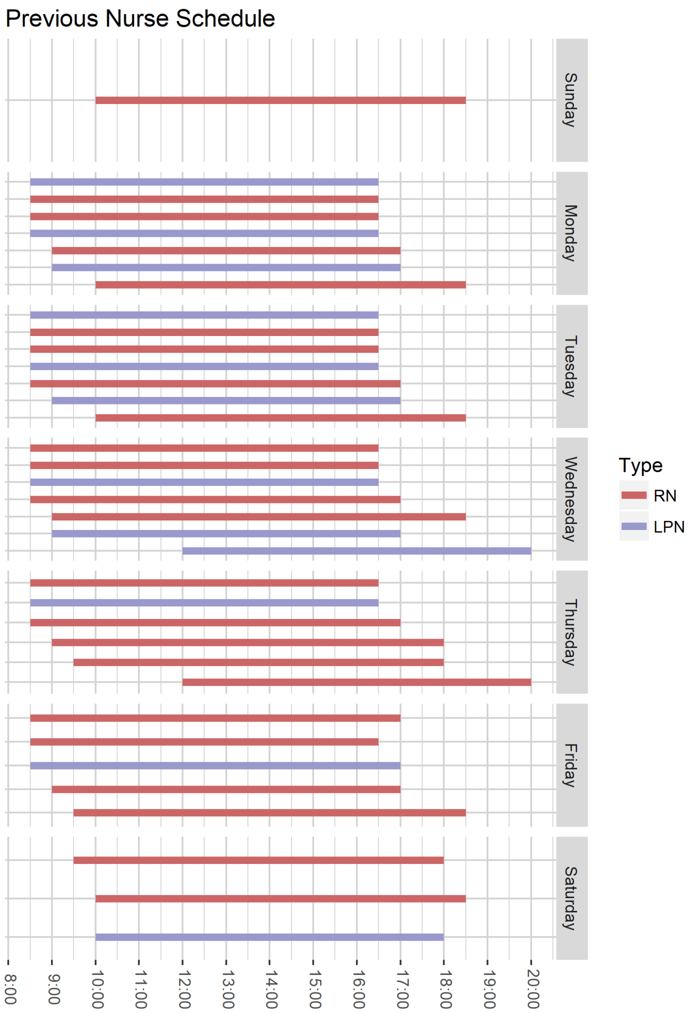


Figure 9: The original schedule.

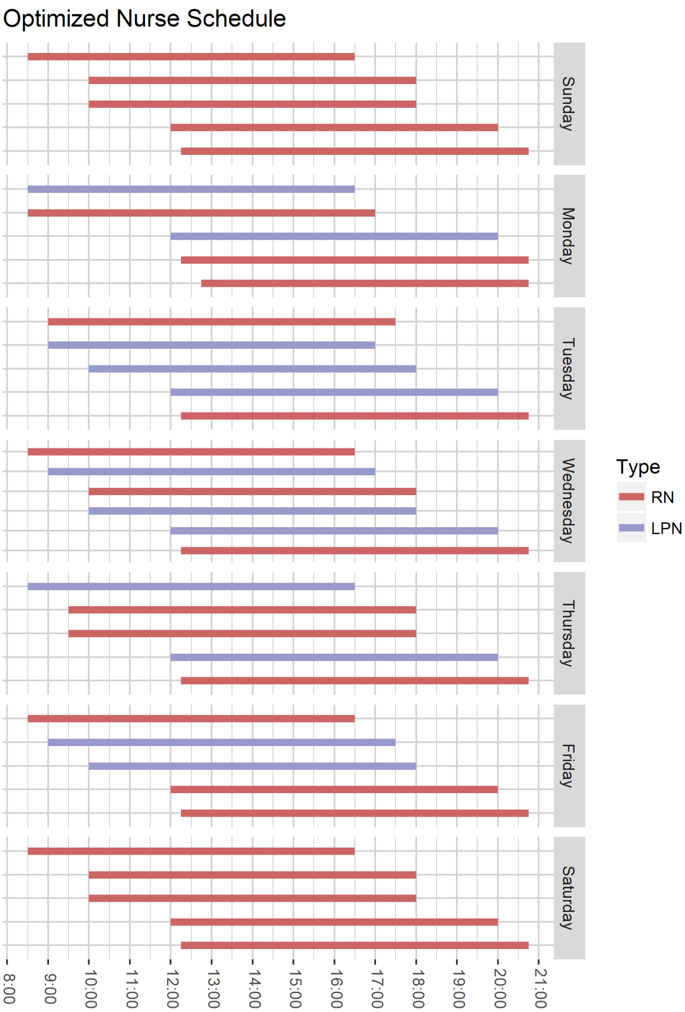


Figure 10: The resulting schedule from the optimization model.

3.3 Discussion and Conclusion

The original nurse schedules for the week considered at the centre struggle to match patient demand, made clear by its infeasible nature in the MILP formulation. Specifically, this infeasible nature indicates that the original schedule is not able to meet ORT and triage duties adequately at all times. The original schedule fared roughly the same as the optimized one when no reductions were made for ORT or triage duties. In other words, the original schedule has around the same expected patient time only when an extra 323 hours of availability are added. While available hours are not directly translated to time with patients (as that relies on patient arrivals), it is clear that the optimized schedule is more effective in meeting patient demand, as well as ORT and triage levels. Additionally, it is much more evenly spread out in comparison to the original with respect to how many staff are scheduled on each day.

This schedule is the most relevant to the last week in October of every year as that was where the majority of our data came from. However, if one considers this week as representative of the last weeks in other months, or even other weeks in general, it can be widely applied. Note that the schedule is only a number of generic shifts, and it is left to the centre to determine the allocation of nurses to shifts.

The model formulation is able to determine schedules in seconds based off of estimated parameters derived from the EMR that reflect demand, greatly shortening the time necessary to update nurse schedules and allowing for decisions surrounding schedules to be supported with by data. Changes to schedules can be easily explored and optimized without the need to real world trials.

3.4 Future Improvements

As mentioned previously, the scope of this model has been narrowed to focus on RN's and LPN's performing adult primary care and ORT. As such, a logical continuation of the model would be to include more of the components of the centre's services, such as their Youth Clinic, Immunization Clinic, and Trans Specialty Care, as well as include other staff types, such as social workers and clinical assistants. Furthermore, reflected in the EMR, the same nurse may be assigned to perform a variety of the centre's services on a single day. Thus, expanding the model to include the hours of operations and patient demand for all of the centre's different services, and scheduling staff with the flexibility of providing different services according to expected patient demand, may lead to a more suitable schedule for the centre's staff. Additionally, RN's, LPN's and MRP's have occasions where they attend to patients when it seems to be more appropriate for the patient to see a social worker or counselor. As such, a more efficient schedule for social workers and counselors may reduce the overflow social work and counseling patients seen by RN's, LPN's, and MRP's.

Although we are scheduling staff levels to meet expected demand, the reality of limited space to service patients has not been considered. Patients will occupy a room until their service is completed. In this case, the model may schedule more nurses than the maximum number of rooms to service patients. Furthermore, if the model were to include the centre's other services, the same rooms may need to be used for multiple purposes and thus require set up and take down time. The addition of a limited space and available room constraint may be considered for a more accurate model.

The model assigns shifts for RN's and LPN's from 10 possible shifts with predetermined start and end times, and breaks. The break times for nurses was taken from their original schedules. In reality, if the model were to include varying start times, end times, break schedules, union regulations would have to be considered . It is worth performing sensitivity analysis on break times, as well as shift start and end times. It may also be a point of consideration to have flexible break schedule be beneficial to staff working in places with high variability in demand and supply of services, like in the case of the centre.

The patients considered are homogeneous in their demand for services. Categorizing patients may help differentiate the patients in our model that demand certain services, thus providing a more reliable patient demand to use when scheduling staff. Expanding our data extraction and analysis to multiple weeks over multiple months would allow greater consideration of weekly and monthly variations in patient demand and arrival rates. We could also model triage and ORT assignment, as well as unavailable slots more dynamically to better reflect the nature of nurse hours provided. Furthermore, service rates are assumed homogeneous across both primary care nurses and ORT nurses. It may be worth sitting in the centre's waiting room tracking patient arrivals, wait times and service times of the providers.

References

[1] R Ashton, L Hague, M Brandreth, and D Worthington. A simulation-based study of a NHS Walk-in Centre. Journal of the Operational Research Society, 56(2):153–161, 2005.

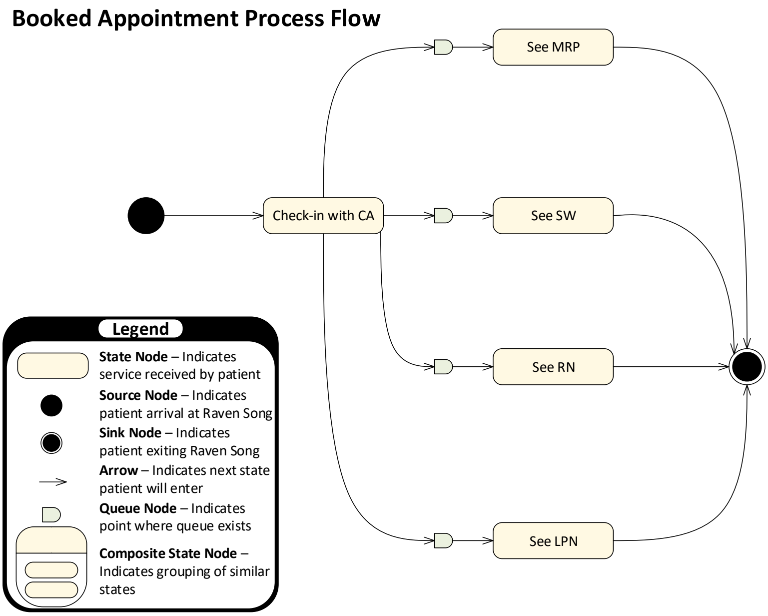
[2] Ana C Cebrian. Package 'NHPoisson'. https://cran.r-project.org/web/packages/MASS/MASS.pdf, 2015.

[3] Gurobi. Python Examples. http://www.gurobi.com/documentation/7.5/examples/python_examples.html.

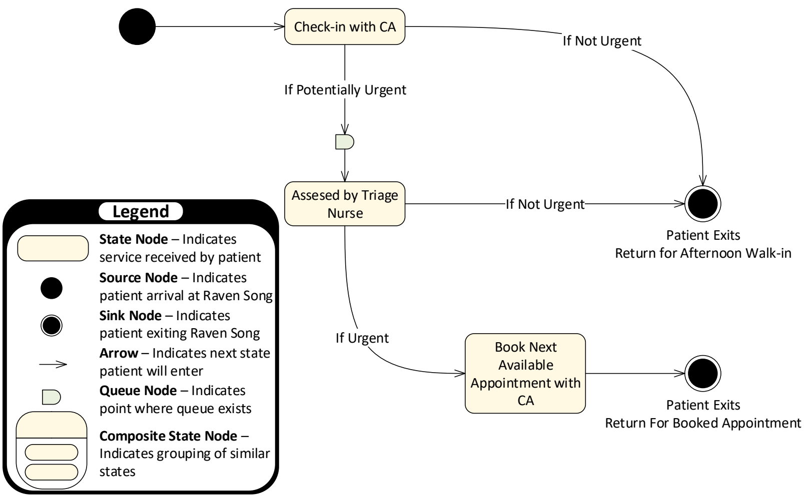
[4] Brian Ripley, Bill Venables, Doglas M Bates, Kurt Hornik, Albrecht Gebhardt, and David Firth. Package 'Mass'. https://cran.r-project.org/web/packages/MASS/MASS.pdf, 2018.

[5] Sheldon M Ross. Introduction to Probability Models. Cambridge, MA: Academic Press, eleventh edition, 2014.

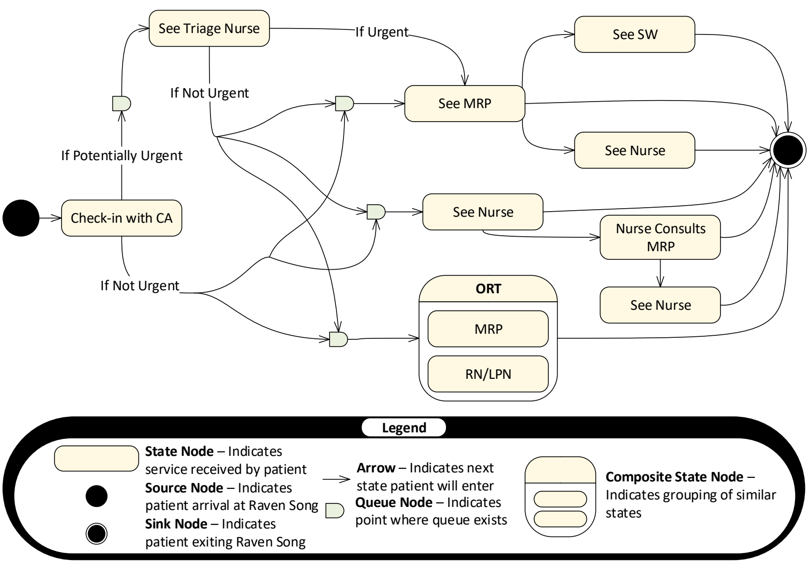
A Patient Flow Process Maps



Morning Walk-in Process Flow



Afternoon Walk-in Process Flow



B Distribution of Service Times