Appointment Optimisation for Urgent in Time Model of Care

Li-Hsuan Chung

Report for

3 November 2025

Table of contents

Abstract

Specialist clinics are under pressure: demand has grown, urgent patients wait too long, and session time cannot be expanded quickly. This report builds a simple and repeatable way to adjust appointment templates so that existing session capacity can be used more effectively. The pipeline is three parts: first we screen all 26 clinics with five signals to find where a change would help the most; next we estimate short-term demand and compare it with typical weekly supply; finally we test a few realistic template settings and show what they mean for waitlists.

Using de-identified data, we reconstruct weekly counts by priority and visit type, then run scenario tests that lift access for urgent new patients while still protecting reviews. For clinic 2078B we evaluate three options: keep urgent new at seventeen per week, lift it to nineteen, or lift it to twenty-four. We simulate backlog paths for three hundred weeks starting from the current waitlist. The results are clear: the first option holds steady but does not reduce the urgent list, the second clears it in roughly four years with moderate pressure on semi-urgent and routine, and the third clears it in about fifteen months but needs extra care to avoid spill-over. The outcome is a set of decisions the clinic can actually use, plus a simple way to monitor progress and adjust settings over time.

Background & Motivation

Motivation Public specialist clinics operate with persistent demand that exceeds short-term capacity. This issue has become even more pronounced in the post-COVID period, when patient flow and referral patterns changed significantly. Session time is largely fixed by staffing, rosters, room availability, and funding cycles. As a result, throughput cannot be lifted quickly by simply adding more clinics. Meanwhile, historical booking practices did not enforce clear rules for the mix of new and review visits across priority tiers. This looseness has produced a gradual mismatch between capacity and need, most evident in prolonged waits for urgent new referrals. The practical question, therefore, is not how to create more time in the short term, but how to use the existing time more effectively.

Research question

How can appointment templates be re-balanced under fixed weekly supply so that urgent waiting times fall, while continuity of care is preserved for semi-urgent and routine cohorts? The study seeks a policy that is operationally feasible, auditable, and explainable to clinical leads and schedulers.

Context and rationale

Referral volumes and the priority mix have shifted in the post-pandemic period. Legacy templates that once worked now under-serve urgent new patients. Clinics also face external performance

expectations around timely access and equity across cohorts. A structured method is needed to identify where reallocation will help most, to quantify near-term pressure on each class, and to test realistic template settings before any change is rolled out.

1 Introduction

1.1 Aim and contribution

The aim is to provide a transparent, reproducible analytics pathway that decision makers can trust. The contributions are fourfold:

A screening framework that ranks twenty-six clinics using five objective indicators: priority mix, new-to-review balance, demand volume, exposure to urgent waiting-time risk, and recent throughput stability.

Short-term demand estimates that place arrivals next to session-based supply, highlighting actionable gaps by class.

A simple optimisation that maximises weighted service subject to real operational rules, including fixed capacity and floors for reviews to protect continuity.

Scenario tests that translate model outputs into options, with clear timelines for urgent clearance and a view of spill-over to semi-urgent and routine.

1.2 Scope and data

The analysis focuses on weekly templates at the clinic level rather than individual clinicians or sessions. Results are intended to support operational planning conversations. The outputs describe what can be achieved with current session supply and how different settings trade urgent gains against pressure on other cohorts. The scenarios are designed to be implemented in the booking system and monitored over time.

1.3 Approach overview

The workflow proceeds in three stages. First, screening identifies clinics where a change in the template mix is most likely to yield measurable improvement. Second, near-term arrivals are reconstructed by priority and visit type and are compared with typical weekly supply to quantify the gap. Third, three template settings are tested by simulation over three hundred weeks starting from the current waitlist. The study reports urgent clearance times, the direction of change for semi-urgent and routine, and a plain recommendation that links model evidence to a policy choice.

1.4 Contributions

Built three policy scenarios based on the Consulting Team's requirements and constraints, then used the current waitlist as Week 0 to simulate 300 weeks and compare outcomes.

For each scenario, reported time to clear Urgent New (UN) and quantified the spill-over on Semi-urgent and Routine so trade-offs are explicit and decision-ready.

Provided a clear, auditable pathway from data to template settings, with recommendations that can be implemented and monitored for ongoing refinement.

2 Methods

2.1 Cross-clinic screening (26 clinics)

Because there are 26 clinics within Urology, we begin with a portfolio-level screening to identify where a change in the appointment template would produce the largest impact. The screening compares clinics on a common set of signals that capture both case mix and timeliness risk. Specifically, we examine the priority mix across urgent, semi-urgent, and routine cohorts; the balance between new and review visits as a proxy for continuity of care; and timeliness indicators that reflect urgent waiting exposure and progress against KPIs. Looking across clinics on the same scale allows us to separate structural differences from noise and to focus attention on clinics where demand pressure is both large and persistent. The outcome is a short list of clinics that merit detailed forecasting and scenario testing, ensuring that subsequent modelling effort is directed to the parts of the system where reallocation is most likely to improve access.

Attended by priority

What the figure shows Figure Figure 1 reports FY2025 attended volume by clinic with the distribution across urgent, semi-urgent, and routine. Clinic 2078B exhibits the largest throughput and an unusually high urgent share, signalling sustained urgent demand rather than a transient surge.

Interpretation The joint pattern of high scale and high urgent mix is a known risk for access: if the template does not reserve sufficient first-visit capacity for urgent patients, queuing and prolonged waits are the predictable result.

Action Prioritise 2078B for template reallocation. Increase protected capacity for urgent new patients to relieve immediate timeliness pressure, while calibrating safeguards so that semi-urgent and routine follow-ups remain adequately served.
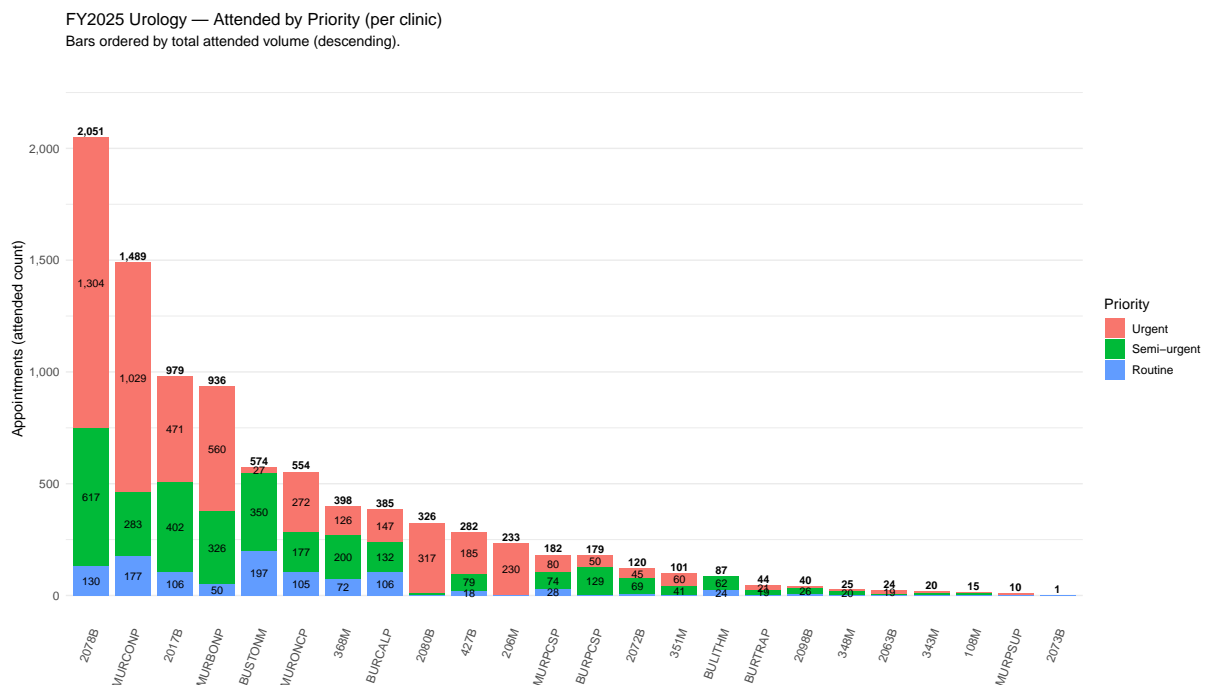
Figure 1: *FY2025 Urology --- Attended by Priority (per clinic)*

New vs Review

Why this matters   Figure 2 shows attended activity by clinic, split into New and Review. When New is persistently under-allocated in high-volume clinics, urgent new referrals accumulate and first-appointment waits lengthen. The aim is to detect this structural imbalance early and target clinics for template reallocation.

How to read it   Use two signals: total bar height for throughput, and the New segment for first-visit capacity. Large totals with a small New segment indicate scope to shift capacity toward New while protecting continuity of care for follow-ups.

FY2025 evidence and action   Clinic 2078B combines high throughput with a below-average New share in FY2025. This pattern aligns with pressure on urgent first-visit timeliness. Prioritise 2078B for a controlled shift toward urgent new appointments, with Review reductions kept within the agreed cap.
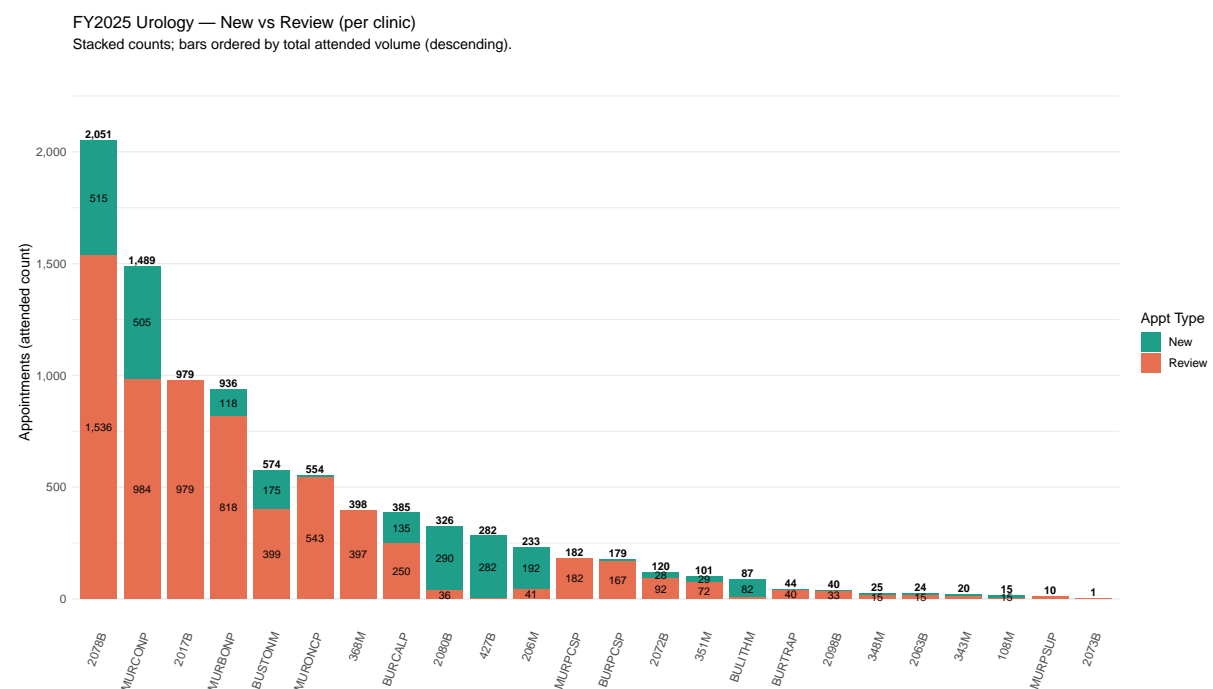
Figure 2: *FY2025 Urology --- New vs Review (per clinic)*

Urgent first-appointment average wait

What the figure shows Figure Figure 3 compares clinics by the average days from referral to first attended appointment for urgent patients. Higher values signal timeliness risk and indicate insufficient first-visit capacity relative to urgent demand.

How to use it with the other figures Read Figure 3 together with Figure 1 and Figure 2. Prioritise clinics that show three conditions at once: long urgent waits, a high urgent share of activity, and a small New segment within total throughput. This combination points to a structural imbalance rather than a one-off surge.

Action Shortlist clinics meeting the above profile for template reallocation. Increase protected capacity for urgent new patients, while setting safeguards to maintain adequate access for semi-urgent and routine follow-ups.
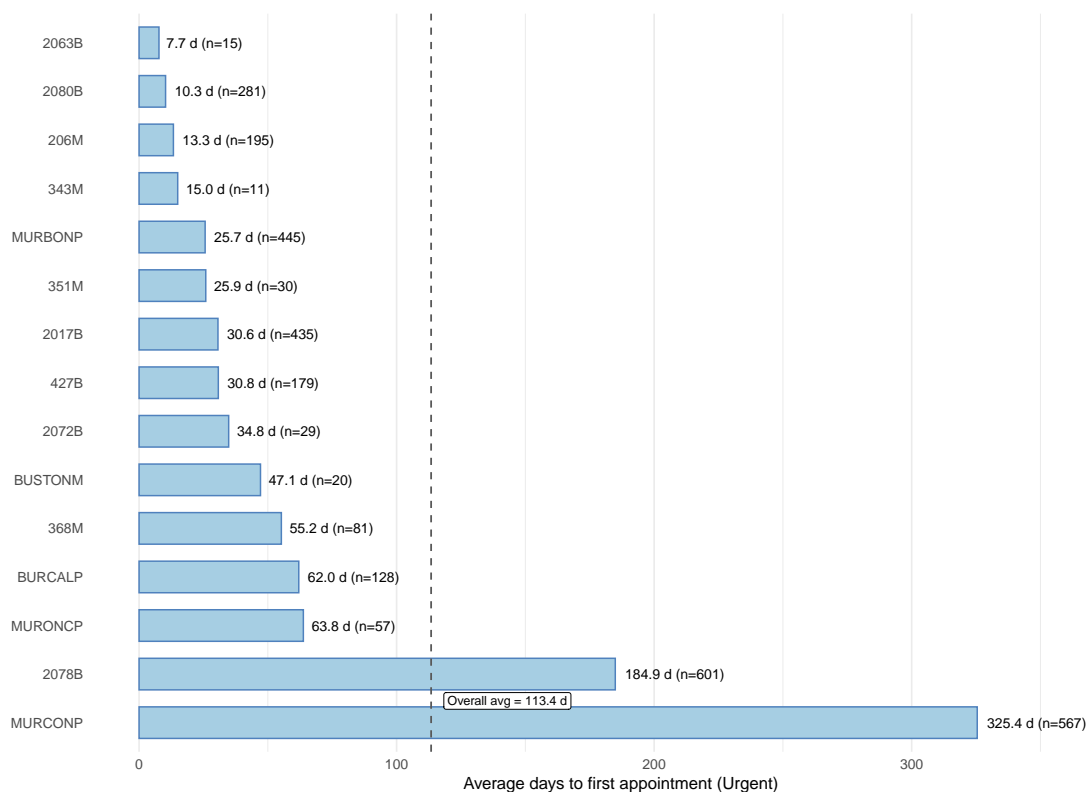
Figure 3: *FY2025 Urology --- Urgent: average wait to first appointment*

KPI

Building on the earlier comparisons of priority mix, new--review balance, and attended volume, we identified a subset of clinics most likely to require adjustment. Figure Figure 4 presents their Urgent and Routine KPIs, defined respectively as "seen within 30 days" and "seen within 365 days."

Among these higher-risk clinics, both 2078B and MURCONP---our largest by throughput---fall noticeably below the service-wide average on both measures. This pattern suggests systemic access constraints rather than random variation and reinforces the earlier screening results, confirming 2078B as the leading candidate for targeted template reallocation and process improvement.
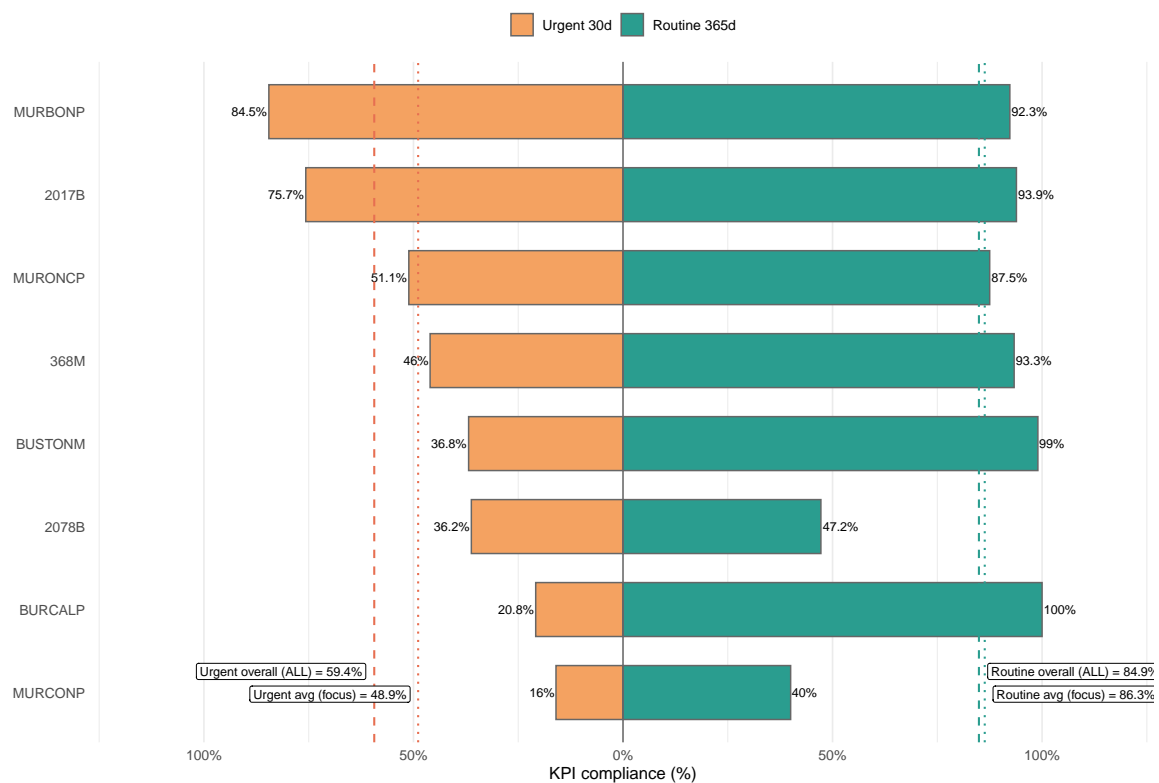
Figure 4: *FY2025 Urology --- KPI: Urgent 30d vs Routine 365d (focus clinics)*

2.2   Forecasting — Demand for 2078B

What we mean by demand

In this analysis, demand represents the total appointment capacity required for the clinic to meet all patient needs within a given period. It includes both external and internal components.

The external component arises from new referrals entering the system---patients who have not yet been seen and require an initial consultation.

The internal component stems from those same patients once they enter the care pathway: each new patient generates a sequence of review appointments for ongoing management, monitoring, or treatment follow-up.

Together, these two streams define the clinic's total workload. Forecasting demand therefore means estimating not only how many new patients will arrive, but also how many review visits those patients will subsequently generate.

Planning identity

Based on this definition, total clinic demand can be expressed as a simple planning identity:

Total demand = External (new referrals) + Internal (reviews)
= Total new accepted referrals + (Total new accepted referrals * Review rate * Avg review visits per patient).

This formulation links activity planning directly to referral inflows and review behaviour, allowing future capacity requirements to be forecast under different referral or follow-up scenarios.

2.3   Forecasting — New referrals for 2078B

Why we forecast

While the planning identity defines total demand in principle, no dataset directly specifies how many new referrals will arrive in the upcoming financial year. To estimate the external component of demand for FY2026, we therefore forecast the volume of new referrals using historical referral trends. This provides a data-driven basis for projecting overall workload and determining whether current capacity settings will be sufficient to meet expected demand.

Aim

The objective of this section is to forecast the monthly number of new referrals for clinic 2078B, thereby quantifying the external component of total demand. This forecast forms the foundation for subsequent calculations of overall appointment requirements, as the inflow of new patients determines both immediate first-visit workload and the future stream of review appointments generated downstream.

Data & split

The modelling uses monthly referral data from July 2023 to September 2025. Data were extracted from the internal referral database, cleaned to remove duplicates and cancelled cases. To ensure an unbiased evaluation of forecast performance, the data are divided into two parts:

- **Training period:** July 2023 -- June 2025, used to estimate model parameters and capture recent seasonal and trend behaviour.

- **Test period:** July 2025 -- September 2025, reserved for validation. This structure allows for a realistic assessment of each model's ability to predict unseen months immediately following the training window.

Models

Four standard univariate time series forecasting models are compared:

ARIMA -- captures both autoregressive and moving-average dynamics, adjusting for trend and auto-correlation.

ETS -- exponential smoothing with automatic selection of additive or multiplicative components for level, trend, and seasonality.

TSLM -- a time series linear model incorporating deterministic trend and seasonal regressors.

SNAIVE -- a seasonal naïve benchmark that repeats the last observed value from the same month in the previous year.

Together these models provide a balance between statistical sophistication and interpretability, ensuring that forecast differences can be attributed to identifiable model structure rather than arbitrary complexity.

Evaluation

Each model is trained on the July 2023 -- June 2025 window and used to predict the July -- September 2025 test months. Forecast accuracy is evaluated primarily using the Mean Absolute Percentage Error (MAPE), which expresses the average deviation as a percentage of the actual values. For completeness, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are also reported to capture absolute and squared error magnitudes.

A visual overlay of the observed series and the test-period predictions is shown in Figure 5, allowing direct comparison of each model's short-term forecast behaviour. The quantitative accuracy measures are summarised in Table 1, providing an empirical basis for selecting the most reliable model to project new referrals into FY2026.
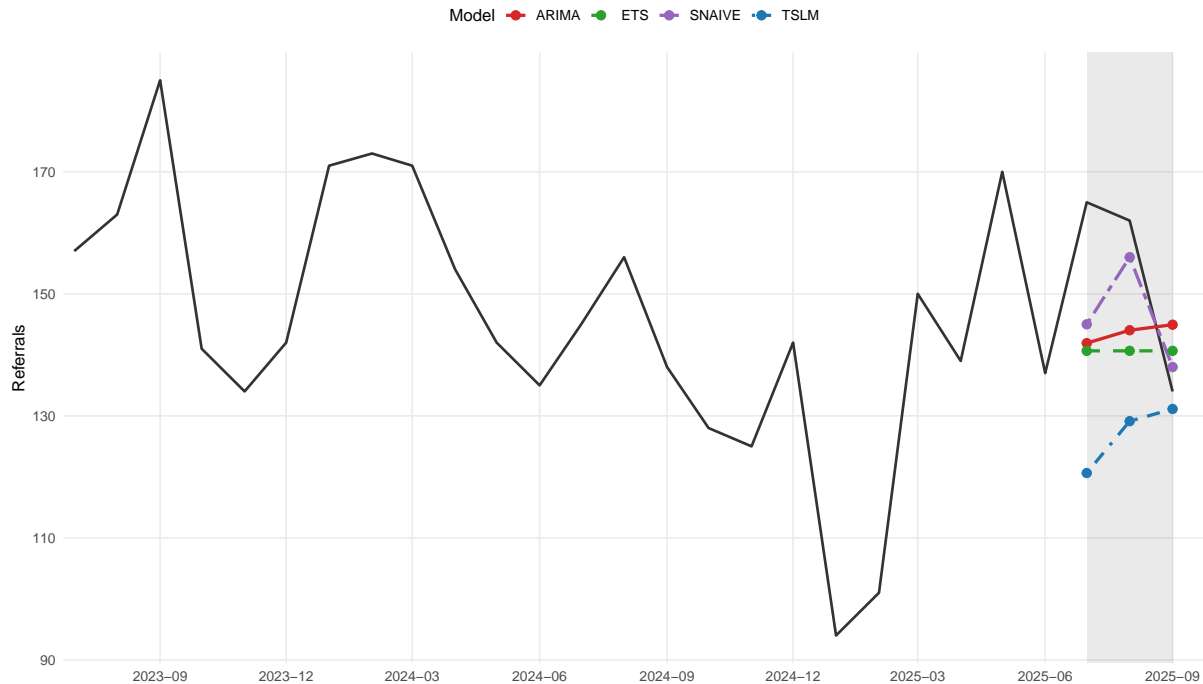
Figure 5: *2078B --- New referrals: Actual vs model predictions (Train: 2023-07..2025-06; Test: 2025-07..09). Shaded band marks the hold-out.*

Table 1: *Hold-out errors for 2078B (Jul--Sep 2025). Lowest MAPE highlighted.*

Model	RMSE	MAE	MAPE
SNAIVE	12.275	10.000	6.27%
ETS	19.085	17.449	10.97%
ARIMA	18.033	17.332	11.08%
TSLM	31.928	26.708	16.44%

From Table 1, the SNAIVE model records the lowest MAPE on the Jul--Sep 2025 hold-out, indicating the best short-term predictive accuracy among the tested methods. Its strong performance suggests that recent referral patterns are highly seasonal and stable over time, making a simple seasonal repetition more reliable than complex models. Consequently, SNAIVE is adopted as the short-horizon model for forecasting the external (new referral) component of demand for clinic 2078B.

Chosen model: SNAIVE
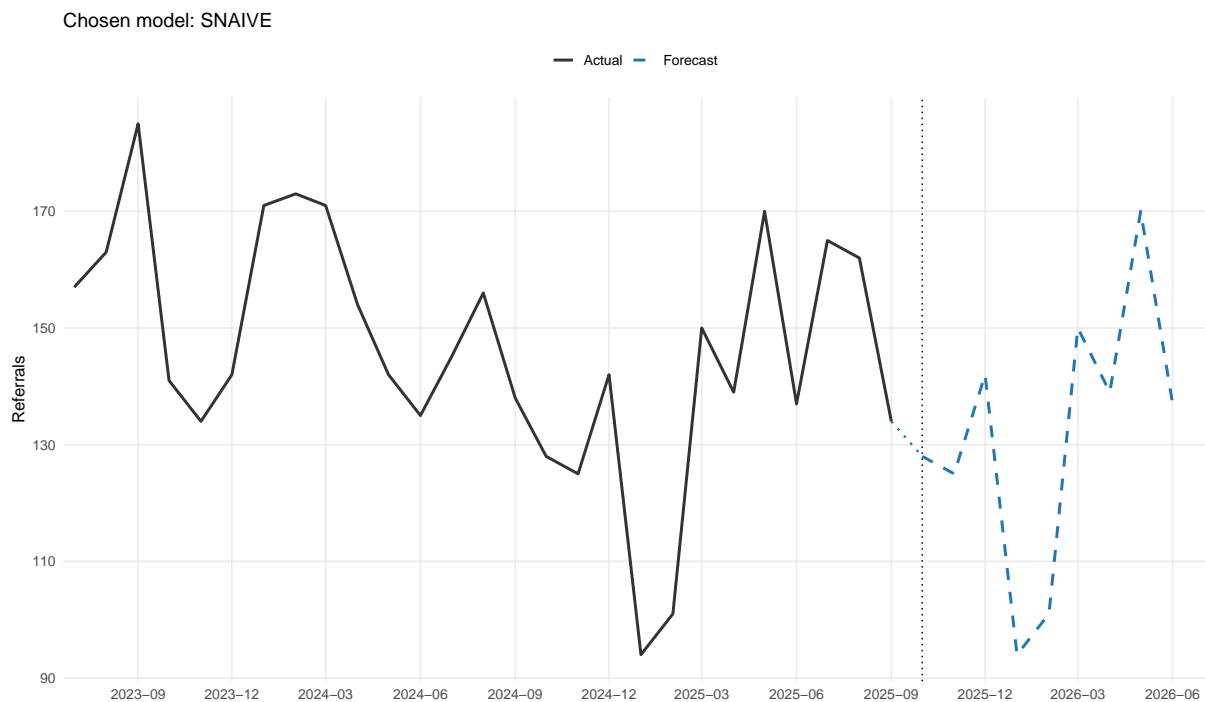
Figure 6: *2078B --- Actual through Sep-2025 and forecast (chosen model) to Jun-2026.*

For clinic 2078B, the external component of demand for FY2026 is estimated by combining both observed and forecast new referrals across the full financial year. We start by calculating the number of accepted new referrals already recorded during the first three months of FY2026, from July to September 2025, which amounts to 461 cases. These represent the actual external inflow received so far.

To complete the year, we use the forecasting model that achieved the lowest error in the previous evaluation to estimate referrals for the remaining nine months (October 2025 to June 2026). The model predicts approximately 1,186 new accepted referrals over this period.

By summing the observed and forecast components, we obtain a projected total of 1,647 new referrals for FY2026. This value represents the expected external demand and will be used as the baseline input for estimating total appointment needs, including the associated review workload generated by these new patients.

Table 2: *2078B --- FY2026 external (new) demand: Jul--Sep 2025 actual + Oct--Jun 2026 forecast (single row).*

ClinicCode	Actual JulSep 2025	Forecast Oct 2025Jun 2026	FY2026 Total
2078B	461	1186	1647

This arithmetic is consistent with the planning identity used throughout the report and ensures that near-term realities are respected while still providing a full-year view.

Translating external referrals to total demand

Using historical data from FY2023 to FY2025, we can estimate how new referrals translate into follow-up activity. Over this three-year period, approximately 35% of patients who attended a first appointment required at least one review visit, and those who returned for reviews attended an average of 1.54 review visits (excluding the first). These parameters allow us to approximate the internal workload that will be generated by new referrals in FY2026.

We begin with the forecasted number of new referrals for clinic 2078B, which is 1,647 for FY2026. This represents the external demand component. The internal demand---that is, the number of follow-up review appointments generated by these new patients---can be estimated as:

- `Internal demand = 1647 * 0.35 * 1.54 = 887.73`

Adding both components gives the total demand:

- `Total Demand = 1647 + 887.73 = 2534.73`

  2,535 appointment slots

Rounded to the nearest whole number, this corresponds to approximately 2,535 appointment slots required for FY2026. In summary, clinic 2078B will need capacity for around 1,647 new appointments and 888 review appointments to meet expected demand. This total serves as the baseline for subsequent capacity planning and template allocation scenarios.

Using FY2023--FY2025 priority mix for 2078B: **Urgent 54.2%**, **Semi-urgent 34.1%**, **Routine 11.7%**. With **Total Demand = 2,535 slots** in FY2026, that is about **49 slots/week** overall.

Combined weekly by priority **Urgent 26/wk**, **Semi-urgent 17/wk**, **Routine 6/wk**.

Broken down into **New** vs **Review**:

	Urgent (54.2%)	Semi-urgent (34.1%)	Routine (11.7%)
New	17 / wk	11 / wk	4 / wk
Review	9 / wk	6 / wk	2 / wk

2.4   Analysis — Average supply

To establish a realistic baseline for modelling, we first examine the actual attended appointments for clinic 2078B during FY2025. These completed activities are categorised into six operational classes, defined by the combination of visit type (New or Review) and urgency level (Urgent, Semi-urgent, or Routine). This classification reflects how appointment templates are structured in practice and allows direct linkage between demand categories and available supply.

Using the full-year totals of attended appointments in each class, we then calculate the average number of appointments per week by dividing each annual total by 52 weeks. This converts yearly activity into a measure of steady-state weekly throughput, assuming a consistent level of clinician session time and roster configuration throughout the year.

Type	Priority	FY2025 Count	Avg / week
New	Urgent	439	8.4
New	Semi-urgent	71	1.4
New	Routine	5	0.1
Review	Urgent	865	16.6
Review	Semi-urgent	546	10.5
Review	Routine	125	2.4

The resulting figures represent the clinic's baseline supply capacity, that is, the volume of appointments that can be realistically delivered per week under current operational settings. In the absence of major staffing or sessional changes, we assume that FY2026 supply will remain equal to FY2025 levels. This assumption allows scenario testing to focus on how template reallocation---rather than total capacity expansion---can improve alignment between supply and demand, especially for urgent new appointments.

Demand–Supply Gap and the Need for Optimisation

Comparing the estimated weekly demand for clinic 2078B with the average weekly supply observed in FY2025 reveals that demand persistently exceeds available capacity. This sustained shortfall explains the extended waiting times faced by urgent patients, despite stable session availability. The underlying problem is not inadequate clinical effort, but an imbalance in how appointment types are distributed within the existing template. Currently, too few slots are reserved for Urgent--New patients, while a disproportionate share of capacity is occupied by follow-up reviews.

Because clinician session time is effectively fixed in the short term, expanding total capacity is not immediately feasible. A more practical approach is to reconfigure the existing template, redistributing

appointment types to better align with the actual pattern of need. The aim is to increase capacity for Urgent--New patients while maintaining adequate access for Semi-urgent and Routine cohorts. In this way, the clinic can improve timeliness for the most critical patients without compromising continuity of care for those requiring ongoing management.

2.5 Reallocation Model

Literature basis

To address the demand--supply imbalance, I apply a linear programming (LP) reallocation model informed by three established streams in the outpatient scheduling literature. First, Patrick, Puterman, and Queyranne (2008) highlight the importance of protecting higher-priority patients through thresholds or quotas, ensuring that urgent cohorts receive guaranteed access before lower-priority ones. Second, Green and Savin (2008) support an urgent carve-out policy, coupled with the operational practice of releasing unused urgent slots on the day of service---an approach that improves timeliness without permanently locking capacity. Third, Gupta and Denton (2008) demonstrate how fairness and KPI targets (such as service levels by priority or visit type) can be encoded as linear constraints, allowing managers to balance urgency protection with continuity of care.

Together, these principles inform the LP formulation used here: to maximise weighted service value, enforce floors and caps that safeguard Urgent--New access, and preserve minimum service levels for semi-urgent, routine, and review patients. The model yields weekly class quotas and tracks unmet demand transparently for subsequent backlog simulation.

Objective

The objective function is defined as

$$\max \sum_{k \in \{\mathrm{UN,SN,RN,URev,SRev,RRev}\}} w_k x_k$$

subject to capacity and fairness constraints.

Symbols

- x_k: weekly slots allocated to class k.
- Class set k $\{\mathrm{UN, SN, RN, URev, SRev, RRev}\}$, where

    UN = Urgent--New

    SN = Semi-urgent--New

    RN = Routine--New

    URev / SRev / RRev = Urgent / Semi-urgent / Routine **Review**

- w_k: class weight capturing policy value/risk for serving one slot in class k, set to reflect urgency > visit type (New > Review), decreasing monotonically: $w_{UN} = 6$, $w_{URev} = 5$, $w_{SN} = 4$, $w_{SRev} = 3$, $w_{RN} = 2$, $w_{RRev} = 1$.

Constraints

While maximising the weekly *service value*, we enforce the following requirements requested by the Consulting Team to ensure feasibility and fairness.

1) Weekly capacity

This rule says that the total number of appointment slots given to all categories cannot go over 39 per week. In other words, we assume the clinic's total staffing and session time are fixed, so we cannot add extra sessions. The model must work within the existing weekly capacity.

$$\sum_{k \in \{UN, SN, RN, URev, SRev, RRev\}} x_k \leq 39$$

2) Scenario cap --- Urgent--New

We test three different policy settings that decide how many Urgent--New (UN) patients can be guaranteed each week: Scenario A = 17/week, Scenario B = 19/week, and Scenario C = 24/week. This cap allows us to see what happens to waiting times and backlog if we give more protected space to urgent new patients.

$$x_{UN} \leq q_s, \quad q_s \in \{17, 19, 24\}$$

3) New-share upper bound

This rule limits how much of the total weekly capacity can go to new patients. We set a maximum of 50% (so that New appointments can never take more than half of the schedule). This helps maintain a fair balance between bringing in new patients and providing follow-up care for existing ones. In Scenario C, this rule is relaxed to see the "best case" where the clinic gives more space to new patients.

$$x_{New} = x_{UN} + x_{SN} + x_{RN}$$

$$x_{New} \leq \alpha S, \quad \alpha = 0.5$$

4) Review floors (protect continuity of care)

These are the minimum guaranteed numbers of review appointments each week:

$$x_{\text{URev}} \geq 5, \qquad x_{\text{SRev}} \geq 3, \qquad x_{\text{RRev}} \geq 1$$

They make sure that every urgency group still has enough follow-up visits so that ongoing patients are not delayed or forgotten when we shift capacity toward new patients.

5) Non-negativity and integrality

This rule just means we can't assign a negative or fractional number of appointments. Each x_k must be a whole, non-negative number, since we can only book complete appointment slots in real life.

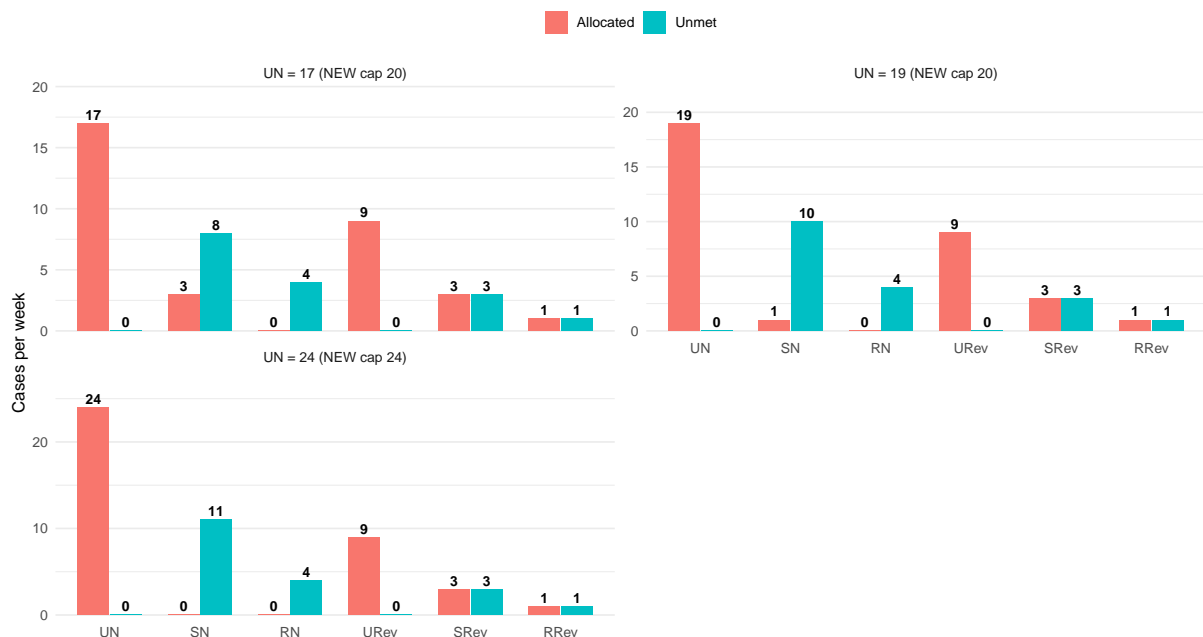$$x_k \geq 0, \quad x_k \in \mathbb{Z}_{\geq 0}, \quad \forall k \in K$$

3 Results

3.1 Reallocate Result

From Figure 7, the red bars show how many appointment slots the model allocates each week to the six categories after solving the optimisation. The blue bars show the number of patients arriving that same week who cannot be booked under the given capacity and fairness rules --- in other words, the unmet demand.

- Across all three policy settings where Urgent--New (UN) capacity is fixed at 17, 19, and 24 slots per week, the results follow a clear pattern:

- Urgent--New demand is fully met in every scenario. The model always protects this highest-priority group first.

- All review categories (Urgent Review, Semi-urgent Review, and Routine Review) also reach their minimum weekly floors, meaning ongoing patients are being followed up as planned.

Figure 7: *Allocated vs Unmet per week by class --- Clinic 2078B (scenarios: UN=17, 19, 24).*

When we increase the weekly target for Urgent--New from 17 to 24 slots, the model directs more capacity to urgent patients, which is exactly what we expect. However, because total weekly supply stays fixed, this shift slightly reduces the space available for other new patients. As a result, the blue bars for SN and RN indicate persistent unmet demand.

This means that while raising Urgent--New capacity improves access for the most critical cases, it does not eliminate waiting for Semi-urgent and Routine patients. To understand how these weekly shortfalls accumulate over time, we next look at the waitlist by class. This helps us see how the backlog evolves under each policy and how much timeliness improvement we actually gain for Urgent--New, compared with the impact on other groups.

3.2 Backlog influence

300-week simulation setup

At the start of the simulation (Week 0), we use the current data extract to set the baseline backlog: Urgent--New (UN) = 453, Semi-urgent--New (SN) = 1,092, and Routine--New (RN) = 440 patients waiting.

From Figure 8, we can see how each group's backlog changes over the next 300 weeks under the three capacity settings.
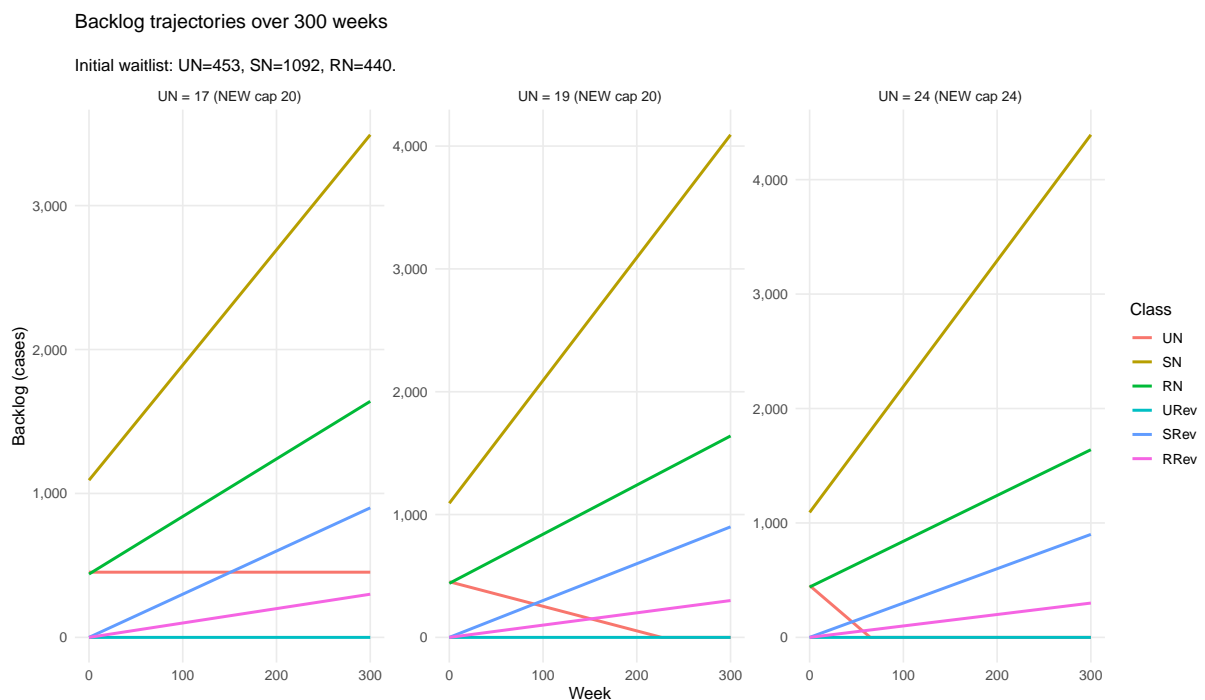
Figure 8: *Backlog trajectories over 300 weeks --- Clinic 2078B.*

For Semi-urgent and Routine, the backlog keeps growing in all scenarios. This happens because their weekly service remains smaller than their inflow --- meaning more patients are joining the queue each week than are being seen.

For Urgent--New, the pattern is quite different:

In Scenario A, capacity just matches weekly inflow. The backlog stops growing but never clears, staying flat over time.

In Scenario B, capacity slightly exceeds inflow, so the backlog slowly declines and fully clears after about 227 weeks (around 4 years).

In Scenario C, extra capacity accelerates the reduction --- the backlog clears in roughly 65 weeks (about 1.25 years).

Overall, the simulation shows that small increases in Urgent--New capacity lead to large improvements in clearance time, while Semi-urgent and Routine groups remain constrained by the fixed total supply.

3.3   Recommendations

Together, these three scenarios give clinic managers a clear set of policy options:

Scenario A --- Maintain stability without clearance In this setting, the clinic allocates 17 slots per week to Urgent--New patients. This level exactly matches the weekly inflow, meaning the clinic can keep up with new urgent demand but cannot make progress on reducing the existing waitlist. The backlog stays flat over time --- it does not grow further, but it also never clears. Scenario A is therefore only suitable if the clinic's goal is to stabilise the waiting list rather than reduce it. It provides short-term control without additional strain on other categories, but does not achieve long-term improvement in timeliness.

Scenario B --- Balanced improvement over the medium term Here, the clinic increases Urgent--New capacity to 19 slots per week, slightly above the current inflow rate. This small adjustment allows the backlog to start decreasing, though the clearance is gradual --- it would take about 227 weeks to fully remove the Urgent--New waitlist. Scenario B represents a balanced option, improving timeliness for the most critical group while keeping Semi-urgent and Routine pressures manageable. It is a practical choice for clinics that want visible progress but need to protect review capacity and avoid large disruptions to other cohorts.

Scenario C --- Rapid clearance with trade-offs This scenario prioritises urgent access more aggressively, allocating 24 slots per week to Urgent--New. With this much capacity, the backlog clears much faster --- in roughly 65 weeks. However, because total weekly capacity is fixed, this setting leaves less room for Semi-urgent and Routine patients, whose unmet demand will continue to grow unless additional sessions or efficiency improvements are introduced. Scenario C is therefore a high-intensity option, suitable if leadership wants to clear the Urgent--New list quickly, accepting that other categories may temporarily experience longer waits.

These results can be presented to the Urology leadership team so they can decide which option best fits their service goals and risk tolerance --- whether to prioritise speed, balance, or stability.

4 Conclusion

This project demonstrates that, even within fixed session capacity, a data-driven optimisation framework can meaningfully enhance access equity across patient cohorts.

The analysis was implemented in two main phases.

Phase 1 --- Screening of clinics: We evaluated 26 Urology clinics using five key indicators: priority mix, New-to-Review ratio, overall demand volume, Urgent--New (UN) waiting time, and KPI attainment. The aim was to identify where a change in appointment template would have the greatest operational and patient impact. This comparative screening highlighted clinic 2078B as the site with the strongest case for intervention, given its high throughput, large urgent share, and persistent timeliness pressure.

Phase 2 -- Capacity optimisation and simulation: We then developed a weekly allocation model that maximises weighted service value under fixed total capacity. The model incorporates fairness and feasibility constraints, including minimum review floors and an upper bound on the total New share, ensuring that changes remain realistic and clinically balanced. Three scenarios were tested, corresponding to Urgent--New quotas of 17, 19, and 24 slots per week, and each configuration was simulated over 300 weeks starting from the current waitlist position.

5   Limitations and Future Work

5.1   Limitations

With fixed session capacity, any scenario that protects Urgent-New will push pressure onto SN/RN to some extent. The best real-world outcome is likely a combo of modest capacity uplift plus reallocation, so UN improves without unacceptable spillover to SN/RN.

5.2   Future Work

Augment the objective with an unmet-demand penalty so the model learns more calibrated service weights and targets the right trade-offs. Optionally, design a phased policy after UN clears (automatic spillover to SN) and test sensitivity to the penalty strength.

Operationalisation with schedulers.   Co-design per-week constraints for all six classes (UN/SN/RN/URev/SRev/RRev) with clinic schedulers so the final allocation can be implemented directly in the hospital booking system.

6 References

Green, L. V., & Savin, S. (2008). Reducing delays for medical appointments: A queueing approach. *Operations Research, 56*(6), 1526--1538. https://doi.org/10.1287/opre.1080.0575

Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions, 40*(9), 800--819. https://doi.org/10.1080/07408170802165880

Patrick, J., Puterman, M. L., & Queyranne, M. (2008). Dynamic multipriority patient scheduling for a diagnostic resource. *Operations Research, 56*(6), 1507--1525. https://doi.org/10.1287/opre.1080.0590

R Core Team. (2024). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Monash Health. (2023). *Specialist Clinics Reporting Business Rules (Version 12/04/2023)*. Monash Health Business Intelligence. [Internal document].

GitHub Repository:

https://github.com/lihsuanchung/ETC5543-lihsuan-project