Operational improvement in a specialist out-patient clinic in Hong Kong

Key Messages

1. The maximum number of patients is an important measure for quantifying clinic congestion.
2. The mean patient waiting time is a useful measure for studying the operation of clinics.
3. An efficient smoothing-out of the number of patients booked in specialty-sessions can help to resolve common congestion and waiting time problems. This will lead to operational improvement of specialist out-patient clinics.

Introduction

Congestion is a major problem in specialist out-patient clinics administered by the Hong Kong Hospital Authority (HA). In some clinic sessions, the total number of patients and accompanying relatives/friends far exceeds the number of available seats in the waiting hall. Another problem is the long patient waiting time for a doctor consultation. In some instances, patients have to wait for more than 2 h, making the waiting hall even more crowded.

The main objective of this study was to quantify and address the problems of congestion and long patient waiting times in the clinic under study.

Methods

The study clinic was open 5.5 days per week and had six morning and five afternoon sessions. The clinic occupied two floors and multiple specialties were scheduled on each floor in each session. The clinic provided 10 specialised out-patient services: medicine, surgery, geriatrics, paediatrics, orthopaedics, gynaecology, antenatal, postnatal, ENT (ear, nose, throat), and dermatology. All patients attending this clinic were referred from either hospitals or other clinics. Both the registration time and the number of booked patients were set by hospital management staff.

Workload statistics and the schedule of clinic sessions during the study period from November 1995 to April 1996 were analysed. An illustration of a typical clinic schedule with three specialty-sessions is shown in Table 1.

Congestion model

In order to quantify the level of patient congestion in each specialty and clinic session, a modelling approach was adopted. This was in line with many overseas studies into the operational performance of medical facilities. Patient waiting time is a useful measure for studying clinic operation. Our previous experience indicated that the maximum number of patients in the waiting hall is also important when quantifying the level of congestion.

We adopted a congestion model to assess this problem. Input measures included the number of patients attending each session, case types, the mean doctor consultation time, and the number of available doctors. Output measures included the maximum number of waiting patients and the patient waiting time.

Model validation

Computer simulation was applied to model the flow of patients and used to validate the congestion model. We simulated a variety of patient arrival patterns through four different scenarios designed to highlight problems including long patient waiting time (>1 h), and overcrowding, with the maximum number of patients exceeding a ‘tolerable limit’.

Different values were set for the ‘tolerable limit’ of different specialties, taking into account the available space in the waiting hall of each floor. For example, the value of the ‘tolerable limit’ for the specialty medicine was much larger than that for orthopaedics, because the orthopaedic patients were more likely to be accompanied by relatives/friends and because some orthopaedic
Scenario A
Depicted the actual patient arrival pattern and was used as a control for comparison.

Scenario B
We hypothesised that a major cause of the overcrowding and long waiting times was the ‘time lag’ between the registration start-time and the consultation start-time. For some clinic sessions, the ‘time lag’ was nearly 2.5 h. Assuming that the doctor consultation start-time was fixed (due to the availability of doctors after ward rounds) and the ‘time lag’ was necessary for the movement of records, we proposed shortening the ‘time lag’ to 1 h. We set the registration start-time at 1 h before the consultation start-time for scenario B.

Scenario C
The original scenario was modified by manually evening out the number of booked patients between the two floors.

Scenario D
A smoothing technique (patient-peak smoothing) was built into the congestion model to help solve the congestion problem. The purpose was to smooth out the congested clinic sessions that had a large number of patients in the waiting hall. This was achieved by re-scheduling patients from the more-congested specialty-sessions to the less-congested specialty-sessions.

Results

Scenario A
Mean patient waiting time
It was found that all sessions on each floor had at least one specialty with a mean waiting time of >1 h. Of the 37 specialty-sessions, 31 had a mean waiting time of >1 h, 25 had >1.5 h, 4 had >2 h. ‘Time lag’ and overbooking for

maximum number of waiting patients
In six of the 37 specialty-sessions, the maximum number of waiting patients was over the ‘tolerable limit’. In two extremely congested specialty-sessions, there were more than 100 patients above the ‘tolerable limit’.

Scenario B
We performed two experiments under scenario B: (1) registration start-time was set at 1 h before consultation start-time; and (2) similar to (1) but the duration of registration was extended by 30 min. This aimed to spread the patient arrivals over a longer period.

Mean waiting time of patients
Of the 37 specialty-sessions, the number of sessions with a mean waiting time of >1 h, >1.5 h, and >2 h for scenarios A and B are illustrated in Table 2. Patients experienced less waiting time in scenario B(1) than scenario A. The registration start-time was set so that patients only came 1 h before the session start-time. Scenario B(2) was even better than scenario B(1). The extension of the registration start-time enabled patients to spread out their arrival pattern and resulted in an overall shortening of the waiting time per patient.

Maximum number of waiting patients
In scenario A we identified four specialty-sessions with a mean waiting time of >2 h and six specialty-sessions with the maximum number of patients over the ‘tolerable limit’. Since two of these specialty-sessions had both problems, there were eight problematic specialty-sessions (Table 3). Under scenario B(1), the maximum number of patients was reduced and brought below the ‘tolerable limit’ in five of the eight specialty-sessions (Table 3). Under scenario B(2), there was an even greater improvement over scenario A. The maximum number of patients were reduced and brought below the ‘tolerable limit’ for all eight problematic specialty-sessions (Table 3).

Scenario C
Different options were tried, especially for those specialty-

patients used wheelchairs, thus requiring more space.

Scenario A
Depicted the actual patient arrival pattern and was used as a control for comparison.

Scenario B
We hypothesised that a major cause of the overcrowding and long waiting times was the ‘time lag’ between the registration start-time and the consultation start-time. For some clinic sessions, the ‘time lag’ was nearly 2.5 h. Assuming that the doctor consultation start-time was fixed (due to the availability of doctors after ward rounds) and the ‘time lag’ was necessary for the movement of records, we proposed shortening the ‘time lag’ to 1 h. We set the registration start-time at 1 h before the consultation start-time for scenario B.

Scenario C
The original scenario was modified by manually evening out the number of booked patients between the two floors.

Scenario D
A smoothing technique (patient-peak smoothing) was built into the congestion model to help solve the congestion problem. The purpose was to smooth out the congested clinic sessions that had a large number of patients in the waiting hall. This was achieved by re-scheduling patients from the more-congested specialty-sessions to the less-congested specialty-sessions.

Results

Scenario A
Mean patient waiting time
It was found that all sessions on each floor had at least one specialty with a mean waiting time of >1 h. Of the 37 specialty-sessions, 31 had a mean waiting time of >1 h, 25 had >1.5 h, 4 had >2 h. ‘Time lag’ and overbooking for

maximum number of waiting patients
In six of the 37 specialty-sessions, the maximum number of waiting patients was over the ‘tolerable limit’. In two extremely congested specialty-sessions, there were more than 100 patients above the ‘tolerable limit’.

Scenario B
We performed two experiments under scenario B: (1) registration start-time was set at 1 h before consultation start-time; and (2) similar to (1) but the duration of registration was extended by 30 min. This aimed to spread the patient arrivals over a longer period.

Mean waiting time of patients
Of the 37 specialty-sessions, the number of sessions with a mean waiting time of >1 h, >1.5 h, and >2 h for scenarios A and B are illustrated in Table 2. Patients experienced less waiting time in scenario B(1) than scenario A. The registration start-time was set so that patients only came 1 h before the session start-time. Scenario B(2) was even better than scenario B(1). The extension of the registration start-time enabled patients to spread out their arrival pattern and resulted in an overall shortening of the waiting time per patient.

Maximum number of waiting patients
In scenario A we identified four specialty-sessions with a mean waiting time of >2 h and six specialty-sessions with the maximum number of patients over the ‘tolerable limit’. Since two of these specialty-sessions had both problems, there were eight problematic specialty-sessions (Table 3). Under scenario B(1), the maximum number of patients was reduced and brought below the ‘tolerable limit’ in five of the eight specialty-sessions (Table 3). Under scenario B(2), there was an even greater improvement over scenario A. The maximum number of patients were reduced and brought below the ‘tolerable limit’ for all eight problematic specialty-sessions (Table 3).

Scenario C
Different options were tried, especially for those specialty-
sessions where the maximum number of patients was over the 'tolerable limit'. However, most of the trial options were not operationally feasible due to the mix of specialties and the availability of manpower resources.

Only one option proved to be feasible and this involved dividing the 440 patients for the problematic specialty-session, Wednesday am (Medicine) [originally located on the second floor], evenly between the two floors. This option was feasible because there were no other specialties scheduled on the first floor. Although the mean waiting time per patient remained unchanged, the congestion problem was greatly eased with the reduction of the maximum number of patients from the original 374 on the second floor to 187 on each floor, which was below the 'tolerable limit'.

**Scenario D**

**Maximum number of waiting patients and mean patient waiting time**

A comparison of the performance before and after applying patient-peak smoothing is given in Figs 1 and 2. There was a significant improvement in the two most congested specialty-sessions, namely, sessions 11 and 22 in Fig 1. The maximum number of waiting patients was reduced by nearly 50%. However, the mean waiting time for these two specialty-sessions remained unchanged (Fig 2). Although specialty-sessions 7, 9, and 19 were less-congested performance worsened after smoothing. Smoothing increased both the maximum number of waiting patients and the mean waiting time for patients in these three specialty-sessions. However, the congestion was kept below the ‘tolerable limit’ and the waiting time was <2.5 h. This worsening effect in some specialty-sessions was unavoidable due to the shifting of patients from the most-congested sessions to less-congested sessions, while keeping the overall total number of patients unchanged. The overall performance in scenario D was better than that in scenario A.

**Discussion**

The original situation at the clinic had been quantified using modelling techniques. A study of the highlighted problematic specialty-sessions led to a better understanding of the major causes of the congestion and long patient waiting time problems. The time lag between the registration start-time and consultation start-time and the number of patients booked for each specialty-session were identified as the two main causes which could be modified for improvement.

Scenarios B, C, and D offer some solutions to address the problems in the clinic. Scenario B should be implemented with care as it involves the education of and acceptance by patients who have been attending the clinic for many years. Although the redistribution of patients (scenario C) between different floors involved fewer changes than scenario B, careful implementation with clear instructions to patients at the centralised registration counter will be necessary. Scenario D offered the most flexible option for application, though the usual limitations of manpower resources and operational feasibility should still be taken into consideration. In this scenario, we are able to postulate ‘what-if’ questions in a user-friendly decision support system for future planning of clinics addressing common problems related to physical capacity, availability of doctors, congestion, and long patient waiting time.

As our project finished, a new computerised appointment system for specialist out-patient clinics was introduced across all HA hospitals and helped to reduce some of the waiting time between registration and consultation. Regular audits were conducted into clinic performance, facilitating more punctual staff arrival time. Subspecialty clinic sessions were also introduced, so spare capacity in the less-congested sessions could be better utilised. As these few measures managed to reduce patient waiting time and congestion, the results of this study were not put
into practice. The clinic management acknowledged that the modelling framework of the project was ‘conceptually’ applied with the implementation of the measures mentioned above.

It was unfortunate that the results of the project were not properly evaluated at the time, as we were unable to verify the impact of the smoothing technique on patient congestion and waiting time. However, the problems of serious congestion and long patient waiting times have recently re-appeared due to the increasing workload and physical space constraints in out-patient clinics. The results of this project may be applicable in the current situation with minor modifications and these findings should be further evaluated.

Acknowledgements

This study was supported by the Health Services Research Fund (#512021). The authors thank Dr SS Lam for research assistance.

References

2. Franz LS, Miller JL. Scheduling medical residents to rotations: solving the large-scale multiperiod staff assignment problem. Oper Res 1993;41:269-79.