Simulating waiting list management

John Bowers

Abstract

Patients' experiences of waiting for treatment have changed dramatically in recent years in the United Kingdom's National Health Service. There has been a substantial reduction in the mean wait but the characteristics of the distributions have also changed significantly, implying a change in priorities and waiting list management. Simulations are often used to assess proposals for investment and reorganisation that might affect waiting time. However, a more realistic model incorporating waiting list management behaviour is needed since empirical distributions of waiting times indicate that a first-in-first-out model is not valid. This study develops a model of waiting list management that includes an explicit measure of priority associated with the patient's wait compared to a specified target. The model can generate a range of behaviours, from the near negative exponential distributions associated with some historic examples of waiting list management, through to the well defined modes of a rigid first-in-first-out system. In one illustration of the use of the waiting list management model, the impact of the seasonal variations in demand and supply was explored. Simulation experiments demonstrated the consequent seasonality in waiting times, implying a need for care when monitoring progress towards any targets.

Keywords Waiting time · Queuing · Simulation · Capacity planning

1 Waiting list targets

The National Health Service (NHS) has adopted a number of key targets to help direct a programme of improvements in healthcare in the United Kingdom. While there is much debate about the value of performance management and the effect of adopting such simple measures [1-2], these targets have been the focus of much activity in the NHS in recent years. Even though the emphasis can vary with a changing political agenda [3], simple measures of performance continue to have a high profile [4]: one such measure is the waiting time for elective care. While this has special prominence in the United Kingdom waiting time is also a concern in other countries, for example Finland [5], Australia [6], Canada [7] and the USA [8].

This study was undertaken in a Scottish NHS Board and the relevant targets are noted in Table 1. The targets distinguish two waiting periods: the outpatient wait being the time between referral, typically from a General Practitioner or Accident and Emergency, and a first outpatient appointment; the inpatient/ daycase wait between a decision to treat and the patient's surgical procedure. Additional investment and a focus on these targets has produced a significant reduction in waiting times in recent years [9-10]. At the time of the study, NHS management in Scotland were working towards the target of 18 weeks for the total wait during the whole process, from referral for the initial outpatient consultation to treatment [11].

Patient group	Maximum wait	Target date
Inpatients and daycases	12 months	1 April 1997
Inpatients and daycases	9 months	31 December 2003
Outpatients and inpatients	26 weeks*	31 December 2005
Inpatients and daycases	18 weeks	31 December 2006
Outpatients	18 weeks	31 December 2007
Total wait (outpatient, diagnostics and surgery)	18 weeks	31 December 2011

Table 1 NHS Scotland waiting list guarantees (* for either component)

2 Capacity planning and models of waiting

The initial requirement in meeting waiting list targets is an adequate sustained or "recurrent" capacity throughout the patient journey (outpatients, diagnostics and surgery) to meet demand. Capacity may be enhanced through a combination of investment and more efficient use of existing resources. Some temporary, non-recurrent capacity may also be necessary to eliminate excessive existing waiting lists; this is often provided as a series of "waiting list initiatives" such as additional weekend sessions. Matching supply and demand is an

essential step in achieving the waiting list targets but not sufficient. Demand may have to be managed, monitoring referral practice so that reductions in waiting times do not stimulate unjustified additional demand. These actions can all help reduce the mean waiting times substantially but the nature of the waiting list management, and the priorities accorded to different patients, may produce a substantial distribution in waiting times with many patients still experiencing excessive waits. Capacity planning models are valuable in managing this process and assessing proposals for achieving the targets but they have to include realistic models of waiting. Other studies have examined the variation in waiting between hospitals [12-13]; the current model aims to capture the variation in patients' experiences of waiting.

A simple first-in-first-out (FIFO) waiting list model is often adequate for capacity planning and has provided the basis for many useful studies [14-15]. Modelling the patients as indistinguishable entities simplifies the analysis and enables a relatively rapid assessment of resource requirements [16-17]. However, more substantive modelling of health services requires a consideration of the variation in patients' characteristics and the attitudes of the staff. Distinguishing major categories, such as urgent and non-urgent patients, explains some of the variation in waiting [6, 18] but this does not model the true range in waiting times. The NHS targets place an emphasis on the tail of the waiting time distribution and there is a need to understand the variation of patients' experiences. Some simulation models have incorporated empirical distributions of waiting times and the implied priorities in order to replicate the likely range in waiting associated with different proposals [19]. The current study attempts to develop a more fundamental model of waiting list management that replicates the key characteristics observed in empirical analyses. The objective was to provide a more realistic forecast of the future waiting lists and to develop a better understanding of the causes of the variation in the patients' waiting times. Reducing the variation, when clinically acceptable, should produce a fairer allocation of care, and help NHS management achieve their targets.

The FIFO model may be acceptable for many service industries but in healthcare clinical need is usually the dominant factor. At the very least the patients designated as urgent should receive care before those categorised as "routine". However, more subtle forms of clinical priority may be used; in some specialties these may be systematised [20] but often the priority will be determined by individual clinical judgement. Reallocating resources and scheduling appointments is not always deemed acceptable [21]; some suggest that healthcare resources should be allocated purely on the basis of clinical need with no consideration of waiting time, except where it affects the health of the patient [22]. A waiting list model is required that can reflect this range of attitudes towards patient priorities.

3 Waiting list data

The study was based on empirical data describing waiting times for elective surgical procedures in a number of specialties using data for the period April 2004 – March 2007 from one Scottish NHS Board. This paper describes the analyses of three specialties, Orthopaedics, General Surgery and Ear, Nose and Throat (ENT), illustrating the different behaviours found in their waiting lists and management. The patient records were categorised by the month of the "decision to treat", i.e. the month that the patient was added to the waiting list. Care is required in ensuring the use of a consistent set of definitions when identifying the appropriate waiting list data: these analyses employed the total wait to treatment for each patient added to the list in each specified month, as recommended [23], rather than the elapsed wait of those still on list. In addition to assimilating data routinely collected for management purposes, a number of interviews were undertaken with relevant staff from each specialty, including nurses, consultants and management. These interviews helped develop an understanding of the factors behind the relationships suggested by the waiting list data and the practices adopted in selecting patients for treatment.

4 Empirical distributions of waiting times

Figures 1-3 depict the changing characteristics of waiting from 2004 to 2007 as the three specialties responded to the targets of Table 1. A further example is provided by the distribution of Figure 4, adapted from a study of the waiting times for trauma and orthopaedics for England and Wales in 2001/2 [21]; this pattern of waiting is not unique to the NHS and the United Kingdom and it has been noted in other countries where waiting list data are collected, such as Finland [5]. The apparent negative exponential characteristics are confirmed in hazard function plots, at least for waits of 15-50 weeks. Such plots are valuable in providing more subtle insights, for example they suggest that there may be four phases to the distribution of Figure 4. Indeed the distributions of Figures 1-4 have many similarities to various phase-type distributions [24], including the hyperexponential, and this could well deserve further study. However, no satisfactory generic analytic distribution was identified that could adequately describe the set of empirical distributions of Figures 1-4. Some of the negative exponential characteristics of Figure 4 can also be seen in the waiting times for General Surgery in Figure 1 and, to a lesser extent, in 2004-6 in ENT. This distribution suggests a waiting list management system in which appointments are made on the basis of individual judgements, with no systematic method of determining priorities: a multitude of factors may influence the allocation but no simple variables can explain the scheduling of patients' appointments.

With the impending waiting time target in 2006/7, the distributions of waiting time in ENT and orthopaedics changed, with a clear priority emerging to avoid patients breaching the target time. However, the characteristics of waiting in General Surgery (Figure 1) did not alter: the mean wait fell from 10.5 to 8.9 weeks and this was sufficient to ensure that relatively few patients' wait exceeded the target. In General Surgery the target was met with all patient groups benefiting from a universal reduction in waiting times. Such an approach would not have been sufficient in ENT and a more focussed attitude was needed. This is reflected in the two distinct modes that appeared in the distribution of 2006/7 waiting times as illustrated in Figure 3. Some patients were designated as "urgent" and every effort was made to provide treatment quickly, typically within 3-5 weeks; most "routine" patients had to wait 15-19 weeks. The emergence of the well defined second mode at 17 weeks suggests that "routine" ENT patients whose wait might otherwise have exceeded the 18 weeks target were now given a high priority, relative to other "routine" patients. Comparing the distributions of Figure 2 suggests that there was a systematic change in waiting list management from 2004/5 to 2006/7, as illustrated by the development of a mode at 17 weeks. This mode is less well defined than that of ENT in Figure 3: Orthopaedics provides an intermediate example of a specialty adjusting its attitude towards waiting time without having to adopt such a focussed approach as ENT.
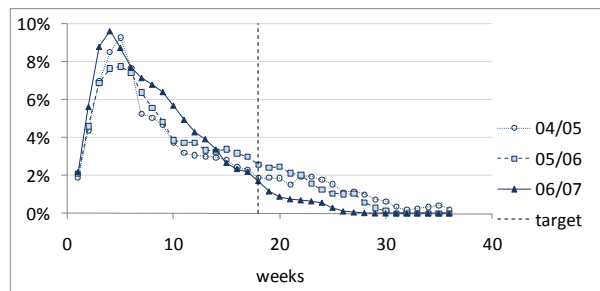

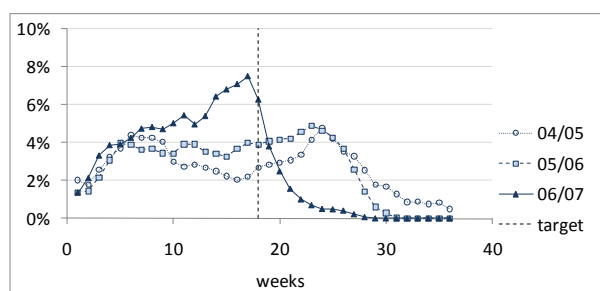Figure 1 Observed waiting times in General Surgery

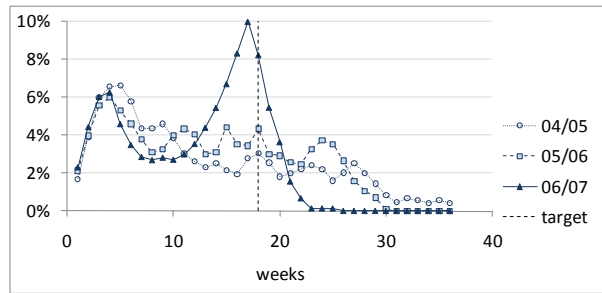Figure 2 Observed waiting times in Orthopaedics
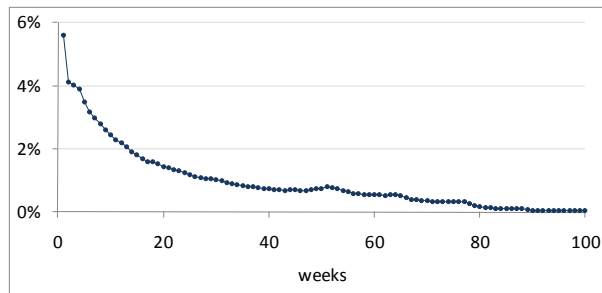
3

Figure 3 Observed waiting times in ENT

Figure 4 Waiting time for trauma and orthopaedics 2001/02

5 Developing a model of waiting list management

The objective was to develop a simple model that could capture the shift in priorities suggested by the interpretation of Figures 1-3, replicating the empirical distributions with sufficient accuracy for assessing future proposals. A simple simulation was constructed using Simul8 to model the effects of different possible responses to targets; Figure 5 illustrates the major elements of the model. The referrals are generated assuming a Poisson model with a weekly mean rate which varies according to a specified seasonal pattern reflecting the typical holiday periods. A similar seasonal activity pattern was applied to model the treatment capacity. The model generates a mix of "urgent" and "routine" patients reflecting the historic proportions. Although all patients may follow the same set of activities, their priority in the waiting list will vary. Indeed, the key element of the model is the ordering of the waiting list according to a specified priority for each patient. While a range of waiting list management policies can be modelled, typically each patient's individual priority is updated weekly such that the probability of being allocated an appointment increases over time. In the basic simulation experiments the annual referral rate is assumed to match the annual treatment capacity: the mean queue length remains approximately constant though the seasonal variations in activity and the stochastic nature of the individual referrals can produce substantial fluctuations in patients' waiting times. Other simulation experiments can explore non-steady state systems, with unequal or changing patterns of demand and supply; this can result from a variety of causes including: changes in referral behaviour, a loss of capacity, increased investment or new follow-up practices in outpatients clinics.
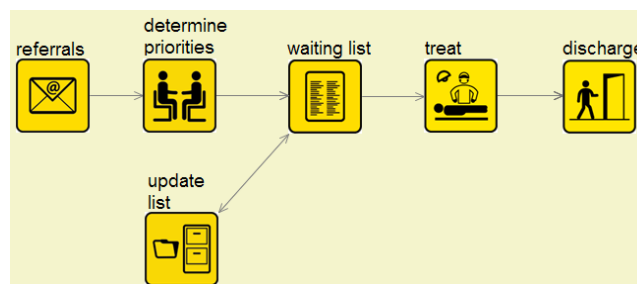

Figure 5 Basic simulation logic of the waiting list model

The priority accorded to each patient consists of three components:

- a clinical priority; although a more sophisticated model might be adopted if the data were available, the current study simply distinguishes urgent and routine patients;
- time spent waiting compared to a target, with different targets for urgent and routine patients;
- a random component modelling the aggregation of factors that are difficult to quantify but may influence the decision to allocate a particular appointment time.

While a variety of priority measures might be used, a simple measure, updated weekly in the simulation, is:

$$P_i = a\frac{w_i}{t} + (1-a)r_i$$

with different values of a and t for urgent and routine patients

Where:

P_i = individual priority accorded to patient i, used to determine his/her position in the waiting list
$0 \le a \le 1$ specifies the priority accorded to waiting time relative to a target wait
a_u = relative priority for urgent patients
a_r = relative priority for routine patients
u = proportion of urgent patients
m = mean wait of all patients
w_i = wait of patient i = current simulation time – time of referral of patient i
t_u = the specified target wait for urgent patients
$t_r = (m - u\, t_u)/(1-u)$ = target wait for routine patients
$0 \le r_i \le 1$ = a random variable that is re-sampled for each patient when the waiting list is updated

The basic behaviour is illustrated in Figure 6 which summarises the waiting time distributions generated from simulation experiments of 20 trials of 5 years for three hypothetical specialties. To aid comparison, patients with a common routine clinical need are considered, i.e. $u = 0$, and the mean wait is identical, $m = 11.4$ weeks, for each specialty. Three values of a are specified reflecting different relative priorities for waiting time in each specialty. When no regard is paid to waiting time, $a = 0.0$, the distribution is similar to a negative exponential, though it is slightly modified by the component seasonal models of referrals and treatment. If a FIFO approach is adopted to the waiting list management, waiting time alone is the criterion for determining priority and $a = 1.0$; in this example FIFO produces a well defined mode around 11 weeks. The experiment suggests that even in the extreme case of focussing solely on waiting time there is a significant spread in the waiting times due to both the seasonal variation in activity over the year and the stochastic nature of referrals. Intermediate values of a, e.g. $a = 0.15$, can produce a variety of intermediate distributions.
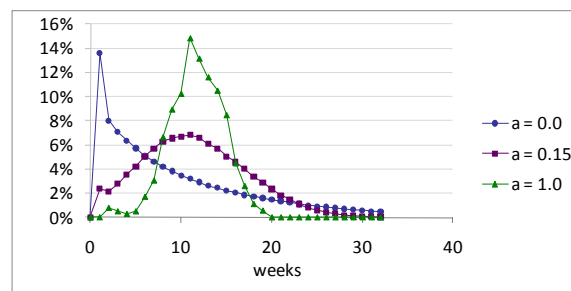

Figure 6 Family of basic distributions generated with the waiting list management model

6 Comparing the model generated waiting times with the empirical distributions

Varying the value of a_u and a_r a variety of distributions of waiting times may be generated and compared with those observed in Figures 1-4. In order to ensure a common basis for comparison a number of parameters were specified reflecting typical values for the three specialties as in Table 2. Various values of a_u and a_r were explored but those noted in Table provided a reasonable fit to the empirical distributions, as illustrated in Figure 7.

Waiting list model parameter	General surgery	Orthopaedics	ENT
proportion of urgent patients, u	0.20	0.20	0.20
initial mean wait (weeks), m	12	12	12
target wait for urgent patients (weeks), t_u	3	3	3
relative waiting time priority (urgent) a_u	0.30	0.20	0.50
relative waiting time priority (routine) a_r	0.10	0.40	0.90

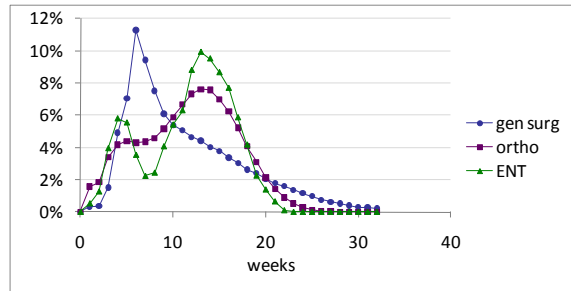Table 2 Waiting list model parameters selected to model the three specialties

Figure 7 Replicating the empirical waiting experience

A formal goodness-of-fit test highlights the significant discrepancies between the waiting list model and the empirical data, reflecting the scope for further work to identify better waiting list management models. However, the current model offers a simple approach, readily incorporated in simulations of proposals for future healthcare systems. It provides a significant improvement compared to the standard FIFO model ($a =1.0$) with more realistic estimates of the range of waiting times experienced by patients.

7 Assessing specific proposals: investing in waiting list initiatives

The waiting list model was designed for inclusion in larger simulations to assess specific local proposals for actions to reduce waiting times. A simple illustration of this role is offered by an assessment of a proposal for a substantial one-off waiting list initiative. Such action may be appropriate where the supply matches demand but the backlog implies that many routine patients continue to experience long waits. It was proposed that a specialty should receive investment to schedule additional weekend theatre sessions. This should enable the waiting list to be reduced with the mean wait falling from 11.4 to 7.5 weeks. The existing approach to waiting list management was similar to that of General Surgery, see Figure 7, with many patients having an excessive wait as noted in Table 3. Using the waiting list simulation, it was estimated that the proposed waiting list initiative should ensure that 97% would be treated within 18 weeks of the decision to treat. However the revised target is 18 weeks for whole care process, see Table 1; staff suggested that the surgical component for of the patient's total wait should be no more than 12 weeks. Combining the additional weekend sessions with a greater focus on waiting time, using a similar approach to that of ENT, simulation experiments suggested that 92% of patients would be treated within the component 12 week target. Waiting list management cannot solve any fundamental mismatch of capacity and demand but it can make useful marginal improvements in reducing the range of patients' waiting experiences.

waiting list management	mean (weeks)	<12 weeks	<18 weeks
"General surgery"	11.4	64%	85%
"General surgery"	7.5	86%	97%
"Orthopaedics"	7.5	89%	100%
"ENT"	7.5	92%	100%

Table 3 Waiting list management and investment in a waiting list initiative

8 Exploring generic behaviour: appreciating the seasonal variation

A further application of the waiting list model is illustrated by an analysis of the seasonal variations in activity and their effect on waiting times. While the magnitude of the effect depends on local conditions, the analysis offers a generic insight into the issue of seasonality and queues for healthcare. Capacity typically varies over the year, reflecting the popular holiday periods for staff. This affects both the demand, generated by the decisions to treat, and the supply of theatres and beds. Using typical seasonal patterns of staff availability, as illustrated in Figure 8, the three approaches to waiting list management represented by General Surgery, orthopaedics and ENT were examined in a series of simulation experiments. The demand, capacity and patterns of availability were specified as being identical in each specialty to facilitate comparison: just the key parameters describing the waiting list management policy were adjusted in this set of simulation experiments. The mean wait of patients being treated in each month varies by 2.7 weeks over the year as illustrated in Figure 9; the 95% confidence intervals are noted, based on simulation experiments of 20 trials of 5 years of activity. The minimum mean wait is experienced by patients treated in June-July and also November-December while those treated in February have the longest waits. Given the basis for this set of experiments, the monthly mean waits do not vary significantly between specialties but the 95-percentile waits, also depicted in Figure 9, vary substantially reflecting the different waiting list management policies. The "General Surgery" 95-percentile wait is of particular concern, varying between 21.6 weeks in June and 24.9 weeks in January-March. The peak in waiting for patients treated in January-March is caused by the backlog of demand reflecting the reduced treatment activity over Christmas/ New Year: the patients who might otherwise have been treated during this period have to wait an extra few weeks.

These experiments demonstrate the need for care when monitoring patient waits and interpreting performance. Waiting times often appear to fall in October-December and this is sometimes interpreted as a trend whereas the reality is it is a seasonal variation and the patient wait may increase to unacceptable lengths by February. This need for care in interpreting the effects of relative changes in demand and supply has been noted in other studies, such as that of the "winter bed crisis" [25]. The common assumption is that such crises arise from an increase in admissions associated with the weather and influenza epidemics whereas the problem appears to be largely due to a reduction in capacity arising from NHS staff leave and holidays in social services departments.
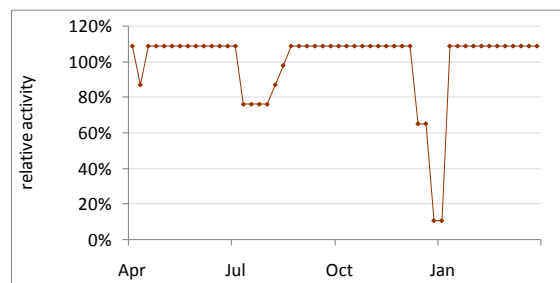

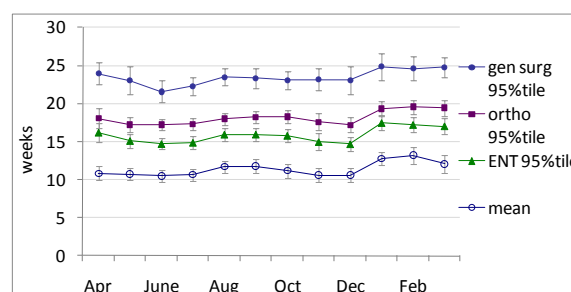Figure 8 Typical seasonal variation in activity

Figure 9 Seasonal variation in mean and 95-percentile waits

9 Conclusions

The empirical distributions of waiting times highlight the inappropriate nature of a simple FIFO model of waiting list management. A range of waiting list characteristics may be observed, changing over the years in response to targets, and also varying across specialties. However, a relatively simple model of waiting list management offers a useful approximation to the observed behaviour. This model provides an insight into the changing characteristics of waiting, indicating that the evolving distributions may be explained in terms of a shift in relative priorities. The model also provides a basis for exploring the effects of specific proposed investments in capacity and it offers a better appreciation of the variations in waiting, such as those arising from the seasonal variations in activity.

References

1. Bevan G, Hood C (2006) Have targets improved performance in the English NHS? BMJ 332: 419-422.
2. Carvel J (2007) Public sector targets to be scrapped. The Guardian Jul 18 2007
3. Department of Health (2010) Revision to the Operating Framework for the NHS in England 2010/11, p.7 available at
   http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_116860.pdf accessed 27 May 2011
4. Timmins N (2011) NHS breaches waiting time targets. The Financial Times Apr 14 2011 available at http://www.ft.com/cms/s/0/170a0340-66cb-11e0-8d88-00144feab49a.html#axzz1NYiiDMuE accessed 27 May 2011
5. Torkki M, Linna M, Seitsalo S, Paavolainen P (2002) How to report and monitor the performance of waiting list management. Int J Tech Assess Health Care 18(3): 611-618
6. Everett JE (2002) A Decision Support Simulation Model for the Management of an Elective Surgery Waiting System. Health Care Manag Sci 5(2): 89-95
7. O'Grady M, Calder D (2005) Waitlist Strategies The Effect of Management, Lessons Learned, Saskatchewan Surgical Care Network
8. Institute for Healthcare Improvement (2003) The Breakthrough Series, IHI's Collaborative Model for Achieving Breakthrough Improvement, Innovation Series 2003, p. 11
9. Audit Scotland (2006) Tackling waiting times in the NHS in Scotland, available at http://www.audit-scotland.gov.uk/docs/health/2005/nr_060216_waiting_times_km.pdf, accessed 26 May 2010
10. ISD Scotland (2010) Waiting time statistics available at http://www.isdscotland.org/isd/5557.html, accessed 28 May 2010
11. NHS Scotland (2008) 18 Weeks: Scotland's Referral to Treatment Standard. http://www.18weeks.scot.nhs.uk/ accessed 26 May 2010
12. Martin S, Rice N, Jacobs R, Smith P (2007) The market for elective surgery: Joint estimation of supply and demand. J Health Econ 26(2): 263-285
13. Tuft S, Gallivan S (2001). Computer modelling of a cataract waiting list. Brit J Opthal 85: 582-585.
14. Koizumi N, Kuno E, Smith T (2005) Modeling Patient Flows Using a Queuing Network with Blocking. Health Care Manag Sci 8(1): 49-60
15. Worthington DJ (1987) Queuing Models for Hospital Waiting Lists. J Opl Res Soc 38(5): 413-422.
16. Jun JB, Jacobson SH and Swisher JR (1999) Application of discrete-event simulation in health care clinics: a survey. J Opl Res Soc 50(2): 109-123
17. Proudlove NC, Black S, Fletcher A (2007) OR and the challenge to improve the NHS: modelling for insight and improvement in in-patient flows. J Opl Res Soc 58(2): 145-158
18. Vasilakis C, Sobolev BG, Kuramoto L, Levy AR (2007) A simulation study of scheduling clinic appointments in surgical care: individual surgeon versus pooled lists. J Opl Res Soc 58(2): 202-211
19. Bowers J (2009) Waiting list behaviour and the consequences for NHS targets. J Opl Res Soc 61(2): 246-254
20. Mullen P (2003) Prioritising Waiting Lists: How and Why? Eur J Oper Res 150(1): 32-45
21. Appleby J, Boyle S, Devlin N, Harley M, Harrison A, Locock L et al. (2005) Sustaining reductions in waiting times: identifying successful strategies, Final Report to the Department of Health

22. Culyer AJ and Cullis JG (1976) Some economics of hospital waiting lists in the NHS. J Soc Pol 5: 239–264
23. Dixon H, Siciliani L (2009) Waiting-time targets in the healthcare sector: How long are we waiting? J Health Econ 28(6): 1081-1098
24. Fackrell, M. 2009. Modelling healthcare systems with phase-type distributions. Health Care Manage Sci 12(1): 11-26.
25. Vasilakis C and El-Darzi E (2001) A simulation study of the winter bed crisis. Health Care Manag Sci 4(1): 31-36.