Paper 26 presented at the Health Economists' Study Group, July 2009, Sheffield

CAPACITY CONSTRAINED COST-EFFECTIVENESS ANALYSIS: A PILOT STUDY IN GLAUCOMA

Jon Karnon, School of Population Health and Clinical Practice, University of Adelaide
Adam Martin, School of Medicine, Health Policy and Practice, University of East Anglia
Robert Casson, South Australian Institute of Ophthalmology, University of Adelaide

Corresponding author:
Jon Karnon, University of Adelaide, Level 3, 122 Frome Street, SA 5005, Australia
Tel: 08 83033562
Email: jonathan.karnon@adelaide.edu.au

ABSTRACT

A review of applied cost-effectiveness analyses found that such analyses rarely explicitly account for the impact of the organisation and delivery of available resources on estimated costs and benefits. Likewise, no studies were identified that analysed the impact of alternative methods of service delivery with respect to patient-based outcomes. The aim of the reported study is to develop an evidence-based framework for the economic evaluation of service delivery and organisation (SDO) for specific patient groups at a hospital level.

A discrete event simulation model was developed to represent the services received by all patients with glaucoma attending the Royal Adelaide Hospital (RAH), whilst simultaneously modelling glaucoma progression. The model contains a set of individuals who are competing for access to hospital services. Individuals are assigned characteristics that determine their level of priority (e.g. their disease status). At each clinic visit, an individual's characteristics are updated, treatment decisions are implemented, and the next intended time of a clinic visit is defined. Across all individuals, these data inform the demand for services over time, which is combined with information on the supply of available resources within the system to facilitate analysis of alternative approaches to ordering access to services.

Preliminary results show that, across 7 alternative scenarios, the maximum difference in the mean QALY gain across the patient cohort (over a 3 year time horizon) is 22 QALYs. Paradoxically, a higher QALY gain is predicted when a smaller maximum clinic size is specified. This is because clinic size is only expanded if urgent patients are waiting beyond their intended clinic date. It is hypothesised that this result is due to limitations in the representation of the clinical decision making algorithm. This result also highlights that the optimal gap between clinics is not necessarily the shortest time, but rather a function of the probability of progression within different time periods.

Further developments to the glaucoma disease model could include both structural improvements and the identification and analysis of better data to inform the model's parameterisation. However, the presented results illustrate the potential for health gains from adjusting the process of health service delivery, in the absence of the introduction of new technologies.

INTRODUCTION

The provision of health services is subject to two broad constraints. Financial constraints limit the amount of resources that can be purchased, whilst physical constraints represent the volume of resources that are available to be purchased. Health technology assessment (HTA) generally compares alternative interventions for specific conditions without explicit consideration of the impact of physical resource constraints on the intervention being evaluated, or analysis of potential effects on the provision of interventions that share resources with the evaluated condition. Reimbursement decisions informed by cost-effectiveness analyses incorporate subjective thresholds for cost-effectiveness. It is then implicitly assumed that if a set of cost-effective interventions are specified, the health service will contrive to deliver a cost-effective health care system.

In practice, the organisation and availability of resources will affect the cost-effectiveness of supplied interventions. The impact of physical constraints is most obvious at the Specialty level within a hospital setting, where the introduction of a new technology requiring additional inputs (of labour and/or capital equipment) can mean resources are transferred within the Specialty and existing service provision is reduced. Moreover, there are potential opportunities to improve technical and allocative efficiency across patient groups within Specialties by considering the costs and benefits of alternative approaches to organising available resources.

A simple example of the decision problem involves a condition with three disease stages. The objective of treating disease stages 1 and 2 is to prevent progression, palliation is the aim in stage 3. Probabilities of preventing progression and of adequate palliation are increased with monitoring, and timely changes in treatment strategies. How should the scarce resources available for monitoring and treatment be allocated between patients who were at disease stages 1, 2 or 3 at the last point of contact? The standard approach involves a clinician making an independent assessment of patients' clinical needs, which may be qualitative (e.g. urgent or non-urgent) or quantitative (e.g. see patient again in 6 weeks). This information is then added to an existing dataset containing similar information for all patients waiting for a particular service (e.g. an outpatient appointment or an inpatient admission). The process of ordering patients is implemented by clerical staff, with reference to an implicit set of decision rules that have been developed on an ad hoc basis over time.

The aim of the reported study is to develop an evidence-based framework for the evaluation of SDO at a hospital level that identifies and corrects both technical and allocative inefficiencies (capacity constrained cost-effectiveness analysis). The first section reports the results of a literature review for

studies that have assessed the costs and benefits associated with SDO of health care resources, in order to identify existing methods for such analyses. The second section presents the methods and preliminary results from the first stage of development of a decision modelling framework to assess current and potential health system efficiency (with respect to health outcomes) at the clinical Specialty level. A discrete event simulation model was developed to represent the services received by all patients with glaucoma attending the Royal Adelaide Hospital (RAH), whilst simultaneously modelling the progression of glaucoma in each patient as a function of patients' frequency of contact with their consultant ophthalmologist.

Literature review of capacity constrained cost-effectiveness analyses

The objective of the literature review was to identify a comprehensive range of study types that have been applied to the SDO of health care resources. A search of Medline was undertaken that combined the following search terms: (cost* OR econ* /in Title) AND (queu* OR wait* /in Title or Abstract). This search strategy identified 348 titles and/or abstracts, which were reviewed. The findings from the review are presented below.

Koopmanschap et al [2005] considered the theoretical effect of ignoring the impact of waiting times on analyses of the costs and health benefits of interventions. They describe different scenarios with respect to deterioration in health status whilst waiting, and absolute levels of health benefits as a function of health status at the time of intervention. Koopmanschap et al hypothesise that costs and effects of health technologies are not fixed over time, and so the cost-effectiveness of extending the availability of an intervention will vary as a function of current waiting times for the intervention.

Gravelle & Siciliani [2008] developed a theoretical model to look at the allocation of health care, where care is rationed by waiting. They found that optimal waiting times are higher for treatments with demands more elastic to waiting time, higher costs, lower charges, smaller marginal welfare loss from waiting by treated patients, and smaller marginal welfare losses from under-consumption of care. Gravelle & Siciliani conclude that allocation rules based purely on cost effectiveness ratios are suboptimal because they assume that there is no rationing within treatments.

The applied studies can be categorised into five study types. A selection of relevant applied studies from each category were retrieved and reviewed to provide more details of the study objectives and examples of the applied approaches.

Type 1 studies are characterised as those that quantify the costs and health effects incurred by patients on waiting lists for various surgical procedures, including cholecystectomy, discectomy, hysterectomy, total knee and total hip replacements.[Núñez et al 2007, Fielden et al 2005, Quan et al 2002] A sample finding is that "longer waits for total hip arthroplasty incur greater economic costs and deterioration in physical function while waiting" [p990, Fielden et al, 2005]. Such data may be of direct interest, as well as informing the cost-effectiveness analyses proposed by Koopmanschap and colleagues.

The second study type looks at the cost-effectiveness of increasing capacity in order to reduce waiting times. In the case of cataracts surgery, Hopkins et al [2005] estimated the QALY gain as the difference in pre- and post-utility multiplied by the reduced waiting period. However, in other clinical areas it would be necessary to incorporate the relationship between waiting time and health status and treatment effect, as postulated by Koopmanschap et al.

This study type raises an interesting question around the appropriate threshold value to use when deciding whether the health gains are sufficient to justify the additional costs. A high proportion of the resources required to provide cataract surgery is the time of specialized health professionals. In the absence of significant new ophthalmic human resources, it is likely that these resources would be transferred from the provision of other ophthalmic interventions. Thus, the relevant threshold value is the incremental cost per unit of effect delivered by the services from which such resources would be transferred.

The third study type looks at alternative prioritisation strategies for one-off inpatient admissions, such as cataract surgery [Tuft & Gallivan 2001] or liver transplants [Ratcliffe et al 2001]. Tuft & Gallivan compared a triage system and a priority based waiting list system with a first come first served admission policy. Their objective was to minimize duration of priority weighted delay, i.e. where delay for higher priority patients was weighted more highly. It was not a great surprise, therefore, that the priority based waiting list system performed best. The benefits of this type of study are better illustrated by Ratcliffe et al, who modeled the survival benefits of alternative waiting list systems, which informs an explicit assessment of the trade-off between efficiency (maximizing survival) and equity. Both horizontal and vertical equity can be assessed by comparing waiting times for patients with similar and divergent needs (priorities). The study by Ratcliffe et al combines service delivery and HTA, but in a specific setting where capacity is assumed to be constrained only by the availability of organs.

The next study type is a mainstream operational research study that provides a detailed representation of service delivery, but focuses on maximising throughput as it's objective. Stahl et al [2005] used discrete event simulation to model the delivery of anesthesia care for laparoscopic cholecystectomy. The model covers the patient journey from admission through discharge, describing individual patient interactions with different health professionals. As the objective was to maximize the number of patients treated per day, the model did not consider the organization of waiting lists, or the impact of the increased throughput on health outcomes.

The final study type models capacity constraints for specific treatments to inform waiting times for patients, which then impact on the estimated long term costs and benefits. Only one such study was identified,[Cooper et al 2008] which modelled pathways for coronary heart disease (CHD) linked to available numbers of angiograms, angioplasties, and bypass grafts. The published application of the model evaluated prevention strategies at a national level, though the model could be adapted to evaluate the costs and health benefits of alternative prioritisation mechanisms for these big ticket interventions at a local level.

Cooper et al integrate physical capacity constraints and HTA, though the broad nature of the CHD model appears to preclude the more detailed analysis of local constraints, such as waiting times for smaller items, including outpatient visits and other diagnostic or prognostic tests.

Thus, the literature review identified no studies that combined operational research analyses of SDO with HTA at a local (hospital) level to analyse the economic impact of alternative methods of service delivery and/or patient prioritisation with respect to patient outcomes.

METHODS

A model-based framework for capacity constrained cost-effectiveness analysis

The framework combines analyses of SDO with HTA to assess the efficiency of hospital services. A prototype model has been developed around glaucoma services at the Royal Adelaide Hospital (RAH). Simulation techniques were used to develop models of disease progression that are linked to the delivery of services for specific conditions. The progression of disease is described in dynamic cohorts of patients in active contact with a hospital, where the speed of disease progression (or regression) is influenced by the frequency and content of hospital visits (outpatient and/or inpatient visits).

A generalized description of the model is that it contains a set of individuals who are competing for access to a system (in this case, hospital services). The individuals are assigned characteristics that determine their level of priority (e.g. their disease status). At each point of contact with the system (e.g. hospital attendance), an individual's characteristics are updated, actions are implemented (e.g. treatment decisions), and the next intended time and point of contact with the system is defined. Across all individuals, these data inform the demand for services over time, which is combined with information on the supply of available resources within the system to facilitate analysis of alternative approaches to ordering access to services.

The following sections describe in more detail the structure and parameterisation of the underlying glaucoma progression model, and the overlying SDO model.

Glaucoma model structure

In a healthy eye, the optic nerve consists of approximately 1 million nerve fibres. These fibres carry information about the visual environment from the retina to the brain. "Glaucoma" refers to family of disease states where there is intraocular-pressure (IOP) sensitive damage to the optic nerve fibres over time. The damage is thought to be initiated at the point where the optic nerve fibres exit the eye (the optic nerve head). The optic nerve head has a donut-like appearance and is referred to clinically as the "disc". There is a central portion to this disc, known as the "cup", which is naturally devoid of nerve fibres and which in glaucomatous eyes typically increases in size as nerve fibres are lost. The functional correlate of this structural damage is loss of the visual field, usually starting in the periphery, but if relentless, progresses to blindness. The visual field is routinely mapped using a computerized test which provides information about the overall depression in the sensitivity of the nerve fibres (the mean deviation; MD) and about localized defects in the visual field (the pattern standard deviation; PSD). Although reduction of the IOP is the sole treatment strategy for all forms of glaucoma, many individuals have glaucoma without the IOP ever exceeding the 97.5^{th} percentile (so called "normal tension glaucoma"). Conversely, many individuals have an IOP which routinely exceeds the 97.5^{th} percentile, but do not develop glaucoma (so called "ocular hypertension"). However, they remain at a greater risk of developing glaucoma. Furthermore, while the definition of glaucoma always includes characteristic, structural optic nerve changes, the exact processes and causes of glaucoma remain largely unknown.[Salmon, 2008]

The most common form of glaucoma (primary open-angle glaucoma) is a relatively slowly progressing condition. Treatment is targeted at reducing IOP, primarily through medical management; laser

treatment and surgery are options if IOP is not controlled. Many other types of glaucoma exist and are seen in tertiary practice and the model was developed with all the types of glaucoma that are encountered in mind.

The glaucoma disease progression model was developed by an iterative process of literature review and discussions with expert clinicians. The structure was chosen to be consistent with current understanding of the natural history of glaucoma, and to reflect the clinical decision-making process.

The resulting model describes disease status as a function of three parameters. IOP is clinically measured in units of pressure (mm of mercury; mmHg), and optic nerve head damage is measured as the ratio of the cup-to-disc (CDR). The MD and PSD are the principle measures of visual field progression in glaucoma patients. MD was used as the measure of visual field progression as a recently published papers have provided algorithms for mapping utility values to measures of MD.[Rein et al, 2007; Burr et al, 2007] Depending on the reason for the clinical appointment, these parameters may or may not be recorded; however, together they provide a widely employed strategy of disease surveillance.

The assumptions underlying the defined structure are that optic nerve head damage and visual field progression are functions of IOP, and the extent to which IOP is controlled. CDR and MD can progress in the presence of well controlled IOP, but the probability of progression will be lower than if IOP is not controlled. The variables were modelled as continuous variables, which provides more specific measures of disease progression, but also provides greater flexibility with respect to modelling progression over time.

Estimation of the glaucoma parameters

Disease progression was modelled as a two step process. Logistic regression models were estimated to inform probabilities of disease progression for each glaucoma variable (IOP, CDR, and MD). If progression occurred, ordered probit regression models informed the magnitude of the deterioration. Explanatory variables included patient, disease and treatment characteristics, as listed in Table 1. The analyses assumed that disease progression is independent of fellow eye status.

Table 1 Variables informing estimates of disease progression

Patient	Disease	Treatment
Age, sex	IOP, CDR, and MD at current and previous clinic visit	Time since last clinic visit, probability of treatment change at last clinic visit

Progression was not described as a function of specific treatments received because the prototype model focused on the impact of frequency of contact. The impact of treatment was incorporated as probabilities of patients changing treatment at each clinic visit, for which a look-up table was developed in consultation with glaucoma specialists. The probabilities were estimated as a function of changes in IOP and CDR, as presented in Table 2.

A review of the literature identified a range of data, including papers reporting rates of CDR progression as a function of change in IOP,[Rath et al, 1996; Kwon et al, 2003] and a review paper of algorithms describing the relationship between IOP and visual field progression.[Schmier et al, 2006] However, the main data source was a dataset comprising measurements of IOP, CDR and MD extracted from the records of glaucoma patients attending the RAH. Data from 47 different patients were extracted, which included observations from 775 clinic appointments [7.7% (60) were 'did not attends' (DNAs)]. Data capture was variable. IOP was recorded in 80% (570) of clinic visits, though CDR was only recorded in 29% (209) and MD in only 13% (90) of attendances.

The data extracted from patient notes was formatted to generate separate observations that combined data from consecutive visits at which visual fields were measured. Each observation in the dataset contained observed information on current and previous IOP, CD ratio, and MD, as well as the time gap between clinic visits, age, and sex. Estimates of the probability of treatment change were then added to each observation.

A wide range of regression models were tested for each of the six specified regression models, with each potential model being tested for specification error using the Ramsey RESET test. In addition, the intervals used in the ordered probit analyses were varied to inform the most robust regression models. Table 3 describes the final models specified.

Table 2 'Probability of treatment change' and 'Intended gap to next clinic' look-up table

IOP		% of patients changing treatment		Intended gap to next clinic (days)		
current	previous	If ΔCD<=0	If ΔCD>0	If ΔCD<=0	If ΔCD>0	
					A	B
30	40	50	60	60	30	60
30	35	55	70	60	30	60
30	30	60	80	60	30	30
30	<30	60	95	30	7	14
28	30	50	70	90	30	60
28	28	50	80	90	30	30
28	26	50	85	90	30	30
28	<26	50	95	90	7	14
26	30	30	65	120	60	90
26	28	30	70	120	60	90
26	26	30	75	180	60	90
26	<26	30	95	180	7	14
24	30	20	60	180	30	30
24	26	20	70	180	30	30
24	24	20	75	180	30	30
24	<24	50	90	180	30	30
22	28	0	65	180	60	90
22	25	0	70	180	60	90
22	22	0	75	180	60	90
22	<22	50	85	180	60	90
20	26	0	60	180	60	90
20	22	0	70	180	60	90
20	20	0	75	180	60	90
20	<20	50	80	180	60	90
18	26	0	60	180	60	90
18	22	0	70	180	60	90
18	20	0	75	180	60	90
18	<20	25	80	180	60	90
16	24	0	60	180	60	90
16	20	0	70	180	60	90
16	18	0	75	180	60	90
16	<18	0	80	180	60	90
<16	<18	0	0	180	60	90

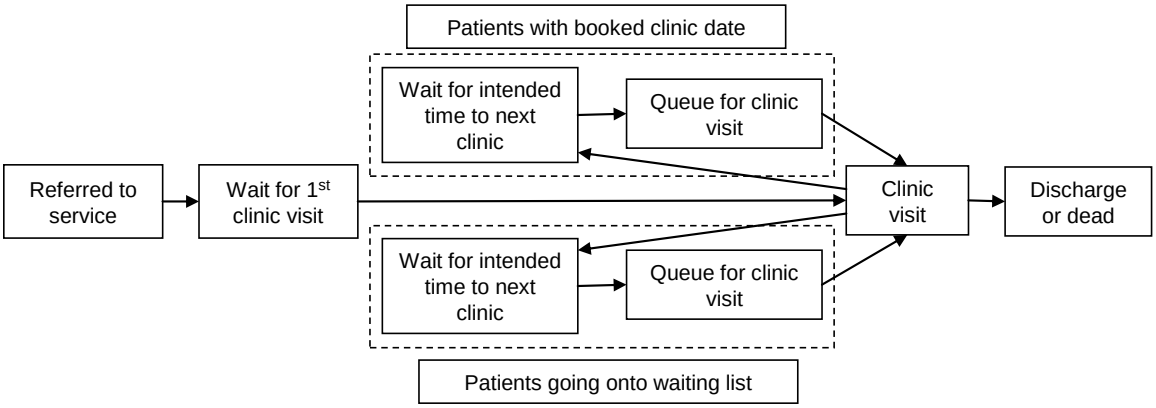
Table 3 Glaucoma parameter models

Logit model for probability of worsening IOP									
IOP_last_visit	age_last_visit	time_gap	rx_change	rx_chg x iop_last_visit	_cons				
-0.262	-0.025	0.001	-0.021	0.001	5.821				
Ordered probit model for patients with improving IOP									
IOP_last_visit	time_gap	/cut1	/cut2	/cut3	/cut4	/cut5			
0.196	0.000	3.773	4.475	5.024	5.636	6.135			
IOP_change categories	0	-2	-4	-6	-8	-10	-15		
Probability thresholds		0.40905	0.67724	0.84375	0.94738	0.983	1		
IOP increase | worsening IOP									
1	2	3	4	5	6	7	8	9	10
0.243	0.427	0.524	0.612	0.728	0.806	0.854	0.893	0.913	1
Logit model for probability of worsening CDR									
dummy_iop_worsened	CDR_last_visit	_cons							
0.448	-3.346	0.731							
Ordered probit model for patients with worsening CDR									
CD_last_visit	time_gap	/cut1	/cut2	/cut3					
-4.057	0.001	-1.757	-0.975	-0.352					
CDR_change categories	0	0.1	0.2	0.3	0.5				
Probability thresholds		0.61026	0.85543	0.95352	1				
Logit model for probability of worsening MD									
current_CDR	CDR_change	_cons							
-6.987	-5.994	4.222							
Ordered probit model for patients with worsening MD									
age	timegap	/cut1	/cut2	/cut3					
-0.026	-0.001	-1.901	-1.429	-0.789					
MD_change categories	0	-3	-6	-10	-20				
Probability thresholds		0.03005	0.07927	0.22065	1				

Service delivery and organization model structure

The SDO section of the prototype model represents the pathway of attending glaucoma patients between clinics over time. The model structure was developed in consultation with clinical and administrative staff at the RAH, and is presented in Figure 1.

Figure 1 Model structure



Moving through the model in patient chronological order, individual patients are referred to a glaucoma clinic by a GP, an optometrist, or a general ophthalmologist. Each new referral is assigned a set of patient and disease characteristics, and an urgency rating based on the referral information provided. Urgency ratings inform the intended time to clinic appointment. Patients then enter a queue, where patients with the earliest intended clinic date are prioritised, even if their intended date has passed.

At each clinic visit, the model determines whether a patient died between referral and clinic appointment (based on Australian life tables). The glaucoma variables of surviving patients are then updated, as informed by the regression models described above. The model then determines whether the patients attend the current clinic as a function of their current disease characteristics and their previous attendance record. Patients not attending three sequential clinics are assumed to be discharged from the waiting list. Other non-attendees are assumed to receive the same intended gap to next clinic as they were assigned at their last clinic attendance.

At the end of a 1st clinic visit, an intended time to next clinic is specified, based on patients' IOP and CDR. At repeat clinic visits, the intended time to next clinic is based on current IOP and CDR, and the change in these variables since the patient's previous visit.

The clinic booking system is open three months in advance, so patients with a next intended clinic date ≤3 months go to a separate queue for future clinics than patients intended to be seen >3 months. This allows the model to define the number of clinic spaces open to patients with and without a booking date provided immediately after their last clinic visit.

After a clinic appointment, patients enter a waiting state for the duration of their intended time to next clinic. They then enter a queue for their next clinic in which patients are prioritised according to the date of their intended next clinic. If no queue exists, patients are offered a clinic appointment on the date intended by their consultant.

Service delivery and organization parameterization

Inputs to the SDO model include numbers of referred patients and their glaucoma status at point of referral. In the prototype model, an informal audit of triaged referrals informed the numbers of patients. Disease status at referral was back extrapolated from disease status at 1st clinic attendance.

The capacity constraint was informed by the number and size of clinics. The RAH holds two glaucoma clinics each week with an intended 15 patients seen per clinic, though more patients are sometimes accommodated depending on urgency.

Model calibration

A calibration reference case was defined that replicated the currently applied decision rules regarding clinic capacity as closely as possible. Specified algorithms for intended time to next clinic (given presenting disease parameters) were varied until the model predicted feasible increases in patient numbers on the waiting list for the glaucoma clinics over a time period of one year.

Model analysis

The Glaucoma model facilitates two broad forms of analysis: clinical and organisational. Clinical variables include the effects of alternative algorithms for defining the 'probability of a treatment change' and the 'intended time to next clinic visit', which were defined as functions of current and previous IOP, and whether or not patients' cup disc ratio worsened between visits. Organisational variables include clinic capacity, the maximum number of new referrals seen in each clinic, and the numbers of clinic spaces that are booked in advance and that are assigned to patients on the waiting list.

Exploratory analyses were undertaken in which clinic capacity and prioritisation algorithms, and algorithms for defining the intended time to next clinic were varied. The efficiency of alternative clinical

decision making algorithms and/or approaches to SDO was measured with respect to the numbers of quality adjusted life years (QALYs) accruing to the cohort of patients attending glaucoma services at the RAH, over a 3 year time period. For each scenario, 200 2nd order model runs were undertaken that used alternative sets of input parameter values. For each 2nd order analysis, 50 1st order model runs simulated the RAH patient cohort over a three year period.

The first three models compared booking algorithms in which clinic capacity was fixed at 15 patients. Within the allowed capacity, either 8 or 10 places were held for more urgent patients who book a date for their next clinic visit at the time of their current visit, as opposed to joining the waiting list. Remaining clinic places were taken up by new referrals and patients on the waiting list.

Urgent (booked) patients are generally those whose cup disc ratio worsens between clinic visits. The impact of increasing the time gap between clinic visits for patients whose CDR worsens (i.e. reducing their urgency) was also tested. In Table 2, the reference case 'intended gaps to next clinic' for these patients are presented in column A, with the extended gaps presented in column B.

The final four model specifications assign unlimited priority to booked patients, such that new referrals and waiting list patients are only seen if there are clinic spaces not taken by booked patients. In models 4 and 6, maximum clinic capacity is also extended to 20, but only if required for booked patients, i.e. if more than 15 booked patients have waited beyond their intended clinic date, capacity is increased. Models 4 and 5 assign lower urgency (i.e. longer intended gaps) to patients showing a change in their CDR than in models 6 and 7.

RESULTS

Table 4 reports the total number of QALYs gained by patients attending the RAH for glaucoma services over a three year period, under each modelled scenario for controlling access to clinics and prioritising patients. The results show that the QALY gains estimated for models 1 to 3 are in-between the QALY gains demonstrated by models 4 to 7. These latter models, which demonstrate the largest differences in QALY gains, prioritise urgent cases absolutely, bounded only by the maximum clinic size.

The maximum difference in the mean QALY gain across the patient cohort (over a 3 year time horizon) is 22 QALYs, between models 5 and 7. This is an initially surprising result as one would expect a model with an expanded capacity (i.e. model 4 or 6) to produce the most QALYs, but instead it is model 7, which assigns high urgency to progressing patients with a clinic capacity of 15 patients. The reason for this result is demonstrated by the data presented in Figures 2a and 2b, which shows the distribution of the

numbers of urgent (booked) and less urgent (waiting list) patients attending clinics over the modelled three year period. The data are illustrative as they represent the distribution for one of the model's runs.

Table 4 Glaucoma model results

Model number	1	2	3	4	5	6	7
Maximum clinic capacity	15	15	15	20	15	20	15
Places held for booked patients*	10	8	10	20	15	20	15
Urgency for patients with ΔCDR†	high	high	low	low	low	high	high
Total QALYs:							
Mean	1962	1961	1960	1958	1943	1960	1965
lower 95% CI	1685	1687	1684	1674	1625	1675	1686
upper 95% CI	2189	2192	2191	2190	2183	2189	2192
Difference from modelled scenario with highest mean QALY gain:							
Mean	-3	-4	-5	-7	-22	-5	0
lower 95% CI	-10.43	-14	-13	-14	-43	-12.5	0
upper 95% CI	5.583	5.5	5.48	0.52	-7.99	1.667	0
Probability modelled scenario has highest QALY gain:							
Mean	0.125	0.08	0.06	0.01	0	0.055	0.685
lower 95% CI‡	0.085	0.05	0.03	0	0	0.03	0.64
upper 95% CI‡	0.18	0.11	0.09	0.02	0.005	0.095	0.73

The figures show that the extended clinic size was only used to reduce a backlog of urgent patients, but that this early increased circulation of urgent patients had knock on effects over the three year period. 51 additional waiting list patients attended a clinic in model 7, where clinic capacity was constrained to 15 places. Conversely, 95 additional booked (urgent) patients attended a clinic when maximum capacity was extended to 20 places. The lower QALY gain in the extended clinic scenario appears to arise because the earlier clinic dates for urgent patients means that they are less likely to have progressed by the time of the visit. A hypothesised scenario is described below:

1. Urgent patients are seen earlier (reduced gap between visits)
2. Urgent patients are less likely to have progressed, so classed as non-urgent and treatment maintained
3. In the absence of updated treatment, re-classified non-urgent patients are more likely to progress whilst waiting for their non-urgent appointment

4. At next visit, non-urgent patients are classified as more urgent, requiring smaller gap between visits

5. Urgent patients seen more often, and non-urgent patients have to wait longer and thus, have more time to progress.

Figure 2a Distributions of the numbers of urgent patients attending glaucoma clinics

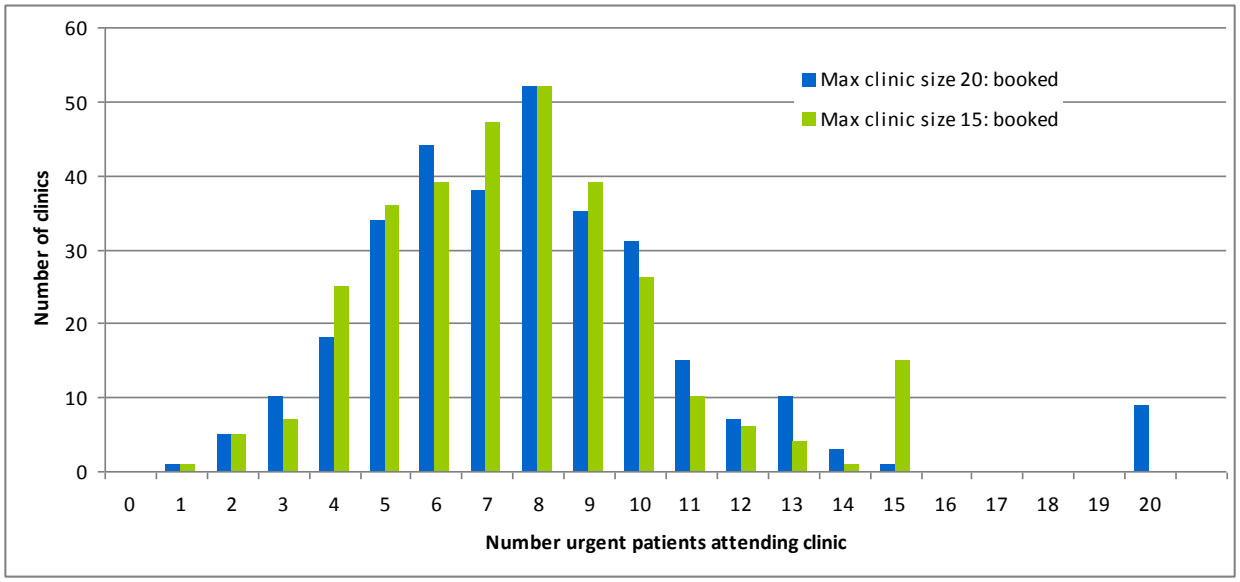
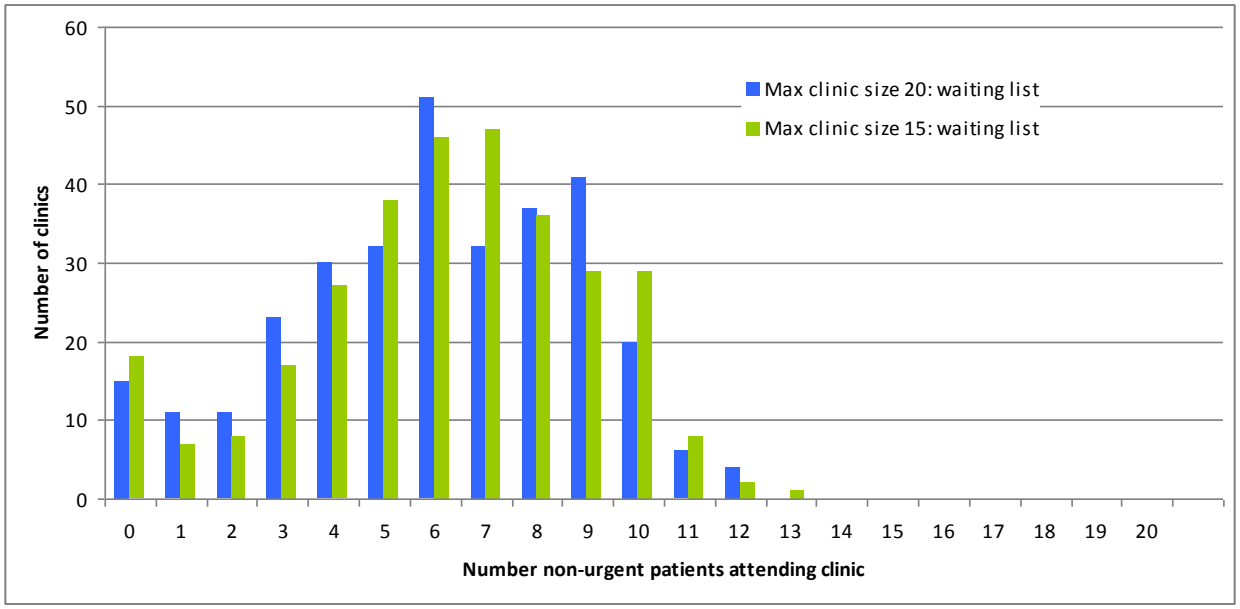


Figure 2b Distributions of the numbers of non-urgent patients attending glaucoma clinics



This hypothesis identifies the potential importance of incorporating disease history into the 'gap to next clinic' algorithm, which is something that the model currently does not do, but that clinicians do incorporate in their decision process. However, it also highlights that the optimal gap between clinics is not necessarily the shortest time, but rather a function of the probability of progression within different time periods. The relative priority given to urgent and non-urgent patients should also account for the likelihood that non-urgent patients will become urgent, whilst the rate of disease progression in urgent cases may decrease.

The opposite result is observed when patients with worsening CDR are assigned less urgency, i.e. more QALYs are gained when clinic capacity is expanded. This results is in line with the above hypothesis as it suggests that varying the intended gap between clinic visits, particularly for the more urgent patients, has more of an impact on aggregate health outcomes than increasing clinic capacity for urgent patients.

The other results presented in Table 4 illustrate the uncertainty around the QALY gains. The presented confidence intervals for the QALY gains should be interpreted with caution. The intervals for each modelled scenario are not independent as each model was evaluated using the same 200 sets of input parameter values. More meaningful are the intervals for the QALY differences from each modelled scenario to the scenario with the highest mean QALY gain (scenario 7). These results show that the upper confidence interval for 5 of the 6 other modelled scenarios is positive, which indicates a statistically significant likelihood that each scenario has higher QALY gains. This is further illustrated by the final rows that describe the probability that each scenario is the optimal scenario. Model 7 has the highest probability of 68.5%, with an upper estimate of 73%. The next most likely scenario is Model 1, with a probability of 12.5% (upper 95% CI 18%).

DISCUSSION

The presented modelled analysis of the assignment of clinical urgency, and the organisation of outpatient services for glaucoma patients illustrated the potential impact of alternative scenarios on health outcomes, in the form of total QALYs gained by the patient cohort receiving glaucoma services.

The maximum mean difference between scenarios was 22 QALYs over a three year period. This QALY gain may not appear to be of a significant magnitude. However, when we consider that bodies such as the Pharmaceuticals Benefits Advisory Committee in Australia use an implicit threshold value of around $50,000 per QALY gained, the implied financial value of the difference between the alternative scenarios is over $1 million in one hospital.

The scale of the potential benefits is much larger as a finalised model could be easily adjusted, and applied to any and all interested hospitals. The process of adjustment applies only to the service delivery and organisation aspects of the model as the natural history components should be applicable to all hospitals, particularly Australian hospitals as most of the input data are Australian.

As implied, the presented results should be treated as illustrative and preliminary. The model requires further development and validation before it can be properly applied to inform clinical and organisation policy within hospitals. Further development of the glaucoma disease model includes both structural improvements and the identification and analysis of better data to inform the model's parameterisation. The regression models fitted to the data extracted from patient records could be improved with the extraction of additional data. Based on the existing regression models, extraction of data from an additional 200 sets of patient notes would provide an adequate sample size to inform the required models. This would provide around 250 observations of visual field progression (the least frequently recorded glaucoma variable) at previous and current clinic visits.

A preliminary literature review identified a range of relevant secondary data, including papers reporting rates of CDR progression as a function of change in IOP,[Rath et al, 1996; Kwon et al, 2003] and a review paper of algorithms describing the relationship between IOP and visual field progression.[Schmier et al, 2006] Primary data analysis of two prominent treatment trials for ocular hypertension could usefully inform the natural history parameters as both included a placebo arm.[Kass et al, 2002; Heijl et al, 2002]

These literature-based data could be used to calibrate the regression models fitted to the primary data. Probabilistic calibration could be used to inform this process.[Karnon et al, 2008 & 2009; Carlton et al, 2008] An example of an application of probabilistic calibration involves sampling regression model coefficients from a multivariate normal distribution, and comparing the predictions to external estimates of the model's outputs. Probability weights are then applied to each set of regression coefficients that reflect the accuracy of their predictions.

The developed algorithm for probabilities of treatment change require further development to represent patient characteristics in more detail (e.g. adverse events and compliance, co-morbidities, etc.) and specific treatment options (e.g. 1^{st}, 2^{nd} and 3^{rd} line treatments). The primary goal of treatment is control of IOP, and the effect of specific treatments on IOP can be derived from reported clinical trials. Additional patient-level data could also be used to validate clinical algorithms and clinical trial-reported clinical effectiveness of treatment.

The representation of glaucoma services and booking systems could also be further developed. Outpatient activity includes the conduct of relevant tests as well as consultation time. The Humphrey Visual Field Analyzer is a specific piece of equipment, for which a separate booking process is used. The availability of test results for visual fields will affect clinical decisions and so the use of this resource needs to be explicitly modelled. Surgery is an option for patients who do not respond to medical intervention. The treatment algorithm should include surgical options, and so the model will need to represent the availability of inpatient services and the prioritisation of patients requiring surgery.

A more detailed algorithm will be required to describe the booking processes for services provided (clinics, tests, and inpatient admissions). For clinic bookings, the challenge is to represent the informal rules that are currently used so that space is left for urgent patients who may require repeat referrals at short notice (e.g. a week or a fortnight). In particular:

- How do the booking clerks assign patients to clinics as a function of the intended time to next clinic (as specified by the clinician)?
- What circumstances justify the expansion of clinic capacity beyond the intended capacity of 15 patients?

In terms of model calibration, it is important to calibrate the full model probabilistically to ensure that it predicts observed outputs, given inputs that reflect current service parameters. Probability weights could be attached to different input parameter sets that reflect the accuracy with which the model predicts the following outputs over a defined time period (e.g. 1 year):

- patients in active contact with glaucoma services (i.e. on waiting or booking lists)
- patients receiving the different glaucoma services
- disease progression, e.g. change in visual fields by IOP, CDR and MD status

In conclusion, existing studies of service delivery and organisation have not assessed the effect of alternative approaches to SDO on patient outcomes, rather such studies have focussed on measures of throughput. The evaluative framework presented in this paper combines simulation models of the natural history of disease with models of patient pathways of hospital-based care to evaluate the impact of SDO on patient outcomes. The results illustrate and quantify the potential for health gains from adjusting the process of health service delivery, in the absence of the introduction of new technologies.

References

Burr JM, Kilonzo M, Vale L, Ryan M, Developing a Preference-Based Glaucoma Utility Index Using a Discrete Choice Experiment, Optom Vis Sci 2007;84:E797–E809

Carlton J, Karnon J, Czoski-Murray C, Smith KJ, Marr J, The clinical effectiveness and cost-effectiveness of screening programmes for amblyopia and strabismus in children up to the age of 4-5 years: a systematic review and economic evaluation, Health Technol Assess. 2008 Jun;12(25):iii, xi-194

Fielden JM, Cumming JM, Horne JG, Devane PA, Slack A, Gallagher LM, Waiting for hip arthroplasty: economic costs and health outcomes, J Arthroplasty. 2005 Dec;20(8):990-7.

Gravelle H, Siciliani L, Ramsey waits: Allocating public health service resources when there is rationing by waiting, J Health Econ (2008), 27(5): 1143-1154

Heijl A, Leske C, Bengisson B, Hyman L, Bengisson B, Hussein M, Reduction of intraocular pressure and glaucoma progression. Results from the early manifest Glaucoma Trial, Arch Ophthalmol. 2002;120:1268-79.

Karnon J, Czoski-Murray C, Smith KJ, Brand C, A Hybrid Cohort Individual Sampling Natural History Model of Age-Related Macular Degeneration: Assessing the Cost-Effectiveness of Screening Using Probabilistic Calibration, Medical Decision Making 2009,

Karnon J, Jones R, Czoski-Murray C, Smith KJ, Cost-utility analysis of screening high-risk groups for anal cancer, J Public Health (Oxf). 2008 Sep;30(3):293-304

Kass MA, Heuer DK, Higginbotham EJ, Johnson CA, Kelner JL, Miller JP, et al, The ocular hypertension treatment study. A Randomized Trial Determines That Topical Ocular Hypotensive Medication Delays or Prevents the Onset of Primary Open-Angle Glaucoma, Arch Ophthalmol. 2002;120:701-713.

Koopmanschap MA, Brouwer WBF, Hakkaart-van Roijen L, van Exel NJA, Influence of waiting time on cost-effectiveness, Social Science & Medicine 2005; 60: 2501–2504

Kwon YH, Kim YI, Pereira MLM, Montague PR, Zimmerman MB, Alward WLM, Rate of optic disc cup progression in treated primary open-angle glaucoma, Journal of Glaucoma 2003;12(5):409-16

Núñez M, Núñez E, Segur JM, Maculé F, Sanchez A, Hernández MV, Vilalta C, Health-related quality of life and costs in patients with osteoarthritis on waiting list for total knee replacement, Osteoarthritis Cartilage. 2007 Mar;15(3):258-65.

Quan H, Lafreniere R, Johnson D, Health service costs for patients on the waiting list, Can J Surg. 2002 Feb;45(1):34-42.

Ratcliffe J, Young T, Buxton M, Eldabi T, Paul R, Burroughs A, Papatheodoridis G, Rolles K, A Simulation Modelling Approach to Evaluating Alternative Policies for the Management of the Waiting List for Liver Transplantation, Health Care Management Science 4, 117–124, 2001

Rath EZ, Shin DH, Kim C, Tsai CS, Zeiter JH, Hong YJ, Relationship between optic disc cupping change and intraocular pressure control in adult glaucoma patients, Graefe's Archive for Clinical and Experimental Ophthalmology 1996, 234(7):434-39

Rein DB, Wirth KE, Johnson CA, Lee PP, Estimating quality-adjusted life year losses associated with visual field deficits using methodological approaches. Ophthalmic Epidemiology 2007, 14

(4):258-264.

Salmon JF, Glaucoma, in Vaughan & Asbury's General Ophthalmology, 17th edition (eds Riordan-Eva P, Whitcher JP), McGraw Hill

Schmier JK, Halpern MT, Covert DW, Robin A. Travoprost versus latanoprost combinations in glaucoma: Economic evaluation based on visual field deficit progression. Curr Med Res Opin 2006; 22(9):1737–1743

Stahl JE, Rattner D, Wiklund R, Lester J, Beinfeld M, Gazelle GS, Reorganizing the System of Care Surrounding Laparoscopic Surgery: A Cost-Effectiveness Analysis Using Discrete-Event Simulation, Med Decis Making 2004; 24; 461