UNIVERSITÁ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI TECNICA E GESTIONE DEI SISTEMI INDUSTRIALI

Corso di laurea triennale in Ingegneria Gestionale

TESI DI LAUREA IN RICERCA OPERATIVA

Scheduling ottimo delle sale operatorie

Relatore: Ch.mo Prof. Romanin Jacur Giorgio

Laureando

Marco Zecchetto

Anno Accademico 2011/2012

a Sebastiano e Angela

INDICE

INTRODUZIONE

Le attività direttamente collegate alle sale operatorie interessano il 9-10% del bilancio di un ospedale (Viapiano e Ward 2000). Le statistiche ospedaliere dal 2004 e 2005, sono disponibili attraverso Hospital Cost & Utilization Project of the Agency for Healthcare Research and Quality, e indicano che gli interventi chirurgici possono essere responsabili fino al 52% di tutti i ricoveri ospedalieri. Questa stima è ottenuta sottraendo la somma dei pazienti ricoverati che vengono dal pronto soccorso, da un altro ospedale o da cliniche a lunga degenza, al numero totale di cure a lungo termine che interessano l'ospedale.

La gestione delle operazioni delle sale operatorie(Operation Management OM) (Fig.1) influenza i costi, il flusso di pazienti e le risorse utilizzate in tutto l' ospedale. Si vuole identificare una serie di problemi decisionali che vengono affrontati dai dirigenti.

Figura 1. Sala operatoria

GENERALITA'

Una ricerca finalizzata ad affrontare il tema delle sale operatorie richiede una profonda comprensione dei vari tipi di decisioni che influiscono sui costi e sulle entrate di un ospedale, è quindi opportuno descrivere l'ambiente in cui la gestione di interventi chirurgici si verifica al fine di incoraggiare ulteriori ricerche su questi temi.

Gli ospedali eseguono interventi chirurgici sia in risposta ad un bisogno *urgente o emergente,* sia su una base *programmabile*.

Il termine *programmabile* non significa che la procedura è facoltativa, ma piuttosto che la sua programmazione può essere ritardata per accogliere pazienti e offrire il servizio a casi più urgenti. Nella modalità programmata, la necessità per un intervento chirurgico è spesso definita dopo una visita con il medico di base del paziente ed eventualmente, uno o più specialisti.

Al contrario, i bisogni urgenti possono insorgere a seguito di un episodio improvvisamente serio, ad esempio una manifestazione di un ictus che può necessitare di cure specifiche immediate.

Le emergenze, infine, sono di solito associate a incidenti automobilistici, traumi acuti o condizioni mediche (ad esempio attacco di cuore) che creano una situazione di pericolo di vita per il paziente.

Gli interventi chirurgici possono essere eseguiti sia con pazienti che già si trovano nella struttura ospedaliera (*in-patient*) sia con pazienti esterni che giungono in ospedale unicamente per sottoporsi all'intervento (*out-patient*).

Con la configurazione di paziente già in ospedale, i pazienti sono ricoverati prima dell'intervento chirurgico e viene assegnato loro un posto letto, mentre nella configurazione *out-patient* rimangono in ospedale solo fino alla completa guarigione, successivamente torneranno alle loro residenze abituali.

La pianificazione degli interventi comprende la gestione della cura chirurgia del paziente, sia nella fase pre-operatoria e post-operatoria.

Nella fase pre-operatoria, i pazienti possono ricevere istruzioni su come prendersi cura di se stessi dopo l'intervento, oltre a qualsiasi analisi medica necessaria.

Diversa è la fase di pre-ospedalizzazione, dove l'ospedale ha lo scopo di preparare i pazienti all'intervento chirurgico mediante l'esecuzione di tutte

quelle indagini finalizzate all'intervento stesso (cliniche, laboratoristiche, strumentali e radiologiche), e/o per mezzo di altre metodologie talvolta necessarie, come ad esempio il pre-deposito di sangue. La pre-ospedalizzazione può essere fatta in Day surgery o con ricovero ordinario.
Per i pazienti in Day surgery, il vantaggio è di eseguire tutti i vari accertamenti propedeutici in una sola sede, senza recarsi in vari centri diagnostici esterni.
Per i secondi il vantaggio è di eseguire tutti gli accertamenti nel più breve tempo possibile, senza lunghe e stressanti degenze preoperatorie in reparto.
La pre-ospedalizzazione, è importante ricordare, risulta unicamente finalizzata all'ottenimento del nulla osta dell'anestesia e non alla formulazione della diagnosi. Tale servizio è quindi un bene, sia per i pazienti, che vedono così ridotti i tempi e migliorate le condizioni di ricovero, sia per i medici, per una migliore e più attenta gestione della fase pre-operatoria che per l'ospedale, per la riduzione dei costi e un'aumentata disponibilità dei posti letto.
I vantaggi che derivano dalla pre-ospedalizzazione sono:

- risparmiare la degenza preoperatoria, utilizzare al meglio il reparto di degenza
- risparmiare sul tempo del paziente che viene ricoverato meno a lungo

Inoltre, un buon lavoro in questa fase permette di pianificare i ricoveri conoscendo già le esigenze assistenziali del singolo paziente, che quindi verrà ricoverato solo se esistono le condizioni per eseguire il suo intervento (da appunti, Dott. Sergio Pierdominici).
Le cure successive all'intervento avvengono o in un' unità di cura post anestesia (PACU), o un reparto rianimazione (ICU), oppure in zona di guarigione post-intervento. Diversi livelli di cura sono forniti in queste aree di recupero secondo la gravità dei bisogni dei pazienti. A recupero completato i pazienti interni che erano già stati ricoverati tornano ai loro letti d'ospedale dove saranno curati fino a quando il medico curante autorizza la loro deambulazione, mentre gli esterni sono dimessi subito dopo il recupero.
Un ruolo fondamentale viene svolto dalla *Recovery room* (Fig.2), la quale interessa le fasi subito prima e subito dopo l'intervento chirurgico assumendo

un ruolo fondamentale in termini di sicurezza e di gestione del *turn–over* dei pazienti.
Inizialmente ha lo scopo di accogliere e preparare il paziente per l'intervento chirurgico. Nella sala vengono eseguiti controlli e manovre di competenza medica e infermieristica (es. dialogo con il paziente, venipuntura, controllo documentazione clinica); nelle fasi immediatamente successive all'intervento la *Recovery room* ha lo scopo di accogliere e sorvegliare il paziente, con un adeguato monitoraggio, fino a quando non diventa stabile e trasferibile.

Figura 2. Recovery Room

La gestione degli interventi chirurgici richiede una comprensione della gerarchia delle decisioni che incidono su tali operazioni.
Al più alto livello strategico di questa gerarchia, gli amministratori dell'ospedale devono decidere quali tipi di interventi chirurgici saranno eseguiti presso il reparto, il numero di sale operatorie che dovranno essere costruite, le tipologie e le quantità di attrezzature e strumenti diagnostici che saranno acquistati.
Le sale operatorie, possono essere diversificate sia dal tipo di procedura (derivante dalla specialità medica) che dal livello d'urgenza.

Il primo tipo è più comune, perché molte procedure richiedono attrezzature mediche specialistiche. Strettamente legate a queste decisioni di capacità sono le decisioni relative al numero di personale infermieristico e al grado in cui essi dovrebbero essere addestrati per affrontare diversi tipi di interventi chirurgici. L'accesso alla capacità per la chirurgia programmabile è gestito in uno dei due modi seguenti:

- pianificazione a blocchi (Fig.3)(*block-scheduling*)
- pianificazione libera (*open-scheduling*)

La scelta tra *open* o *block* è una decisione strategica. Con l'approccio *block*, i chirurghi vengono assegnati precisamente alla sala operatoria con un programma periodico (settimanale o mensile) e possono prenotare appuntamenti nello spazio loro assegnato secondo le condizioni dei pazienti. Inoltre l' 8-10% degli spazi operatori è lasciato disponibile.
La durata media dell'operazione chirurgica (ottenuta da documenti storici) è tipicamente usata per decidere se il paziente può essere inserito o meno in un blocco. Nel caso in cui non si possa inserire, il chirurgo deve richiedere un'aggiunta di tempo per le "prenotazioni eccedenti". In questo tipo di gestione è necessario che le sale operatorie, il personale infermieristico e il personale medico siano intercambiabili per ottimizzare la distribuzione dei carichi di lavoro (da appunti, Dott. Sergio Pierdominici).
I meccanismi sono complessi e non sempre prevedibili, ma in un'organizzazione che prevede l'interscambiabilità di sale e personale sono possibili spostamenti anche all'ultimo minuto per consentire di ottimizzare le risorse disponibili.
In un sistema di programmazione *open* i chirurghi danno le richieste di tempo per gli interventi il giorno prima dell'intervento e il calendario specifica, al rispettivo chirurgo, la sala operatoria e l'orario d'inizio.
In molti casi gli ospedali utilizzano un misto di questi due approcci in cui i gruppi chirurgici (o assistenti) possono dividere il tempo in funzione della richiesta che viene posta nel loro tempo assegnato.
Durante l'attività operatoria programmata si rende necessaria una pianificazione (settimanale/mensile) delle liste, dei materiali necessari e dei posti letto; c'è quindi una gestione continua dei flussi di pazienti, ottimizzando

gli spazi che si rendono liberi. Questa condivisione è ottenuta mediante la definizione di una scadenza (un numero particolare di giorni prima dell'intervento chirurgico) nel momento in cui il tempo di blocco non utilizzato da un gruppo chirurgico diventa disponibile per l'uso di altri gruppi.

		Annual Elective Cases	Avarage Chases/Wk 48 Wks/Yr	Elective Weekday Hours/Yr	Elective Hrs/Wk 48 Wks/Yr	Turnover Minutes per case	Hrs/Wk+ Turnover	Requested Block Time/Wk	Allocated Block Time/Wk
Orthopedics	Aisle	395	8	699	15	20	17	19	18
	Fish	369	8	689	14	20	17	16	15,5
	Pits	215	5	422	9	20	10	12	10
	Bergdorf	323	7	499	10	20	13	18	11,5
General Surgery	Washer	365	8	536	11	15	13		
	Buryle ather	362	8	433	9	15	11		
	Pillow	220	5	327	7	15	8		
	General Surgery						32	14	13
Gynecology	Lay	226	5	311	6	15	8		
	Dillow	65	1	99	2	15	2		
	Lay-Dillow							4,5	5
	Quisling	164	3	217	5	15	5		
	Bernhart	30	1	43	1	15	1		
	Quisling-bernhart							5,5	5
Urology	Ghengis	417	9	537	11	15	13	9,5	10
	Feeling	46	1	50	1	15	1		
ENT	Peregrine	471	10	295	6	10	8	5,5	5
	Blender	27	1	19	0	10	0		
EYE	Hoddkins	373	8	196	4	10	5	8	6,3
	Samson	291	6	192	4	10	5	7,5	5
Other	Gone	60	1	72	2	15	2		
	Daveson	34	1	40	1	15	1		
	Ramirez	50	1	39	1	15	1		
	Foster	25	1	33	1	15	1		
	Ringo	6	0	21	0	15	0		
	Montague	12	0	14	0	15	0		
	Eppley	8	0	12	0	15	0		
	Realty	12	0	11	0	15	0		
	All Others	217	5	41	1	15	2		
Grand Total			103		121		176	119,5	104,3
Total Available								180	

Figura 3. Tabella riassuntiva per i dati necessari all'inizializzazione dei blocchi

Al livello successivo di decisione i manager devono determinare la quantità di capacità operativa da attribuire ad ogni medico o ad ogni specialista. Comunque questi problemi, che sorgono solo quando sono utilizzate tecniche *block*, sono comunemente riscontrati negli ospedali statunitensi e canadesi a causa dell'uso diffuso di questo tipo di tecnica.

Le assegnazioni dei blocchi vengono periodicamente riviste quando cambia la capacità, cioè quando c'è un cambiamento nel numero di chirurghi o quando innovazioni medico-tecnologiche alterano la capacità di utilizzo di alcuni tipi di procedure.

Prendere decisioni tattiche coinvolge anche la decisione su quanti interventi chirurgici programmare in un specifico blocco e l'assegnazione degli orari di inizio di ciascun intervento chirurgico programmato, in modo che i membri della staff chirurgico (ad esempio anestesisti e chirurghi) possano sapere quando e dove presentarsi per ogni intervento.

Le decisioni di pianificazione tattica si dividono in due fasi.

Nella prima fase vengono decisi il livello di priorità medica del paziente e le varie regole di prenotazione usate per determinare un arco di tempo approssimativo (diciamo, una settimana) per l'esecuzione di ogni procedura.

Nella seconda i blocchi di interventi chirurgici vengono assegnati alle specifiche sale operatorie e vengono fissati i tempi per ogni procedura.

Si noti che quest'ultima fase è divisa in altre due parti:

- la sequenza dell'intervento
- la selezione dei tempi di inizio.

Questa procedura avendo due passi offre maggiore flessibilità nella programmazione delle analisi e, quando necessario, nella cancellazione dei pazienti come candidati per l'intervento chirurgico.

Una complicazione che sorge quando la stessa sala operatoria è utilizzata anche per trattare i casi di urgenza o emergenti è che i manager devono decidere la quantità di capacità da dedicare a questi ultimi.

Il livello operativo del processo decisionale sopporta degli slittamenti rispetto alle condizioni ipotizzate. Tali slittamenti si verificano quando gli interventi richiedono più tempo del previsto, ovvero quando:

- le apparecchiature non sono disponibili all'orario di inizio previsto
- il personale non è disponibile all'orario di inizio previsto
- vi è un gran numero di casi di emergenza
- le risorse a valle non sono disponibili per prendersi cura di pazienti in uscita dalle sale operatorie.

E' responsabilità dei manager cercare di rispettare i tempi, garantendo nel contempo la sicurezza dei pazienti.
In questo scritto si descrivono le sfide nella modellazione e nelle soluzioni per risolvere tre esempi di problemi.
I problemi sono i seguenti:

1. Come si dovrebbe distribuire il tempo delle sale operatorie alle sottospecialità chirurgiche? Dopo l'espansione di capacità o su base periodica?
2. Come si dovrebbero gestire le prenotazioni degli interventi programmati?
3. A parte i vincoli medici, qual è la sequenza ottimale per l'esecuzione degli interventi chirurgici?
4. Qual è la perdita di efficienza seguendo le regole euristiche per la sequenzialità degli interventi chirurgici?

La lista sopra nominata non è un elenco completo dei problemi importanti che devono affrontare i manager delle sale operatorie, i quali loro utilizzano una varietà di ipotesi euristiche per trovare soluzioni pratiche per questi ed altri problemi (Magerlein e Martin 1978; Blake e Carter 1997 per maggiori dettagli).
Questi problemi sono al centro dell'attenzione perché non c'è sufficiente ricerca sulla modellazione degli stessi.
L'obiettivo è quello di proporre modelli preliminari ed esplorare le possibili soluzioni. Alcuni di questi problemi possono essere discussi solo dopo aver specificato una particolare procedura. In tali casi si devono assumere le caratteristiche che potrebbero rappresentare una situazione ricorrente. Questi risultati sono, tuttavia, dipendenti dalla struttura assunta e non possono essere generalizzati a tutte le situazioni.

In questa tesi, vengono usate le seguenti convenzioni:

- le lettere maiuscole indicano le variabili casuali, per evitare confusione tra il numero 1 e L minuscola
- i vettori sono scritti con “ → ".

CAPITOLO 1

Allocazione della capacità per interventi programmabili

Si andrà a considerare una situazione in cui l'ospedale deve decidere di aumentare o il numero delle sale operatorie o il numero di ore di funzionamento delle stesse per l'attività chirurgica, in quanto il tempo previsto è già completamente utilizzato e la domanda si prevede aumenterà in futuro.

Di conseguenza, l'amministrazione dell'ospedale deve decidere quanto tempo (in blocchi) assegnare a chirurghi e sottospecialisti. Un approccio simile è importante anche per la revisione periodica del tempo di allocazione delle sale operatorie.

Il tempo deve essere assegnato ben in anticipo per attrezzare le sale operatorie con il giusto equipaggiamento e di assumere personale adeguato per sostenere l'ulteriore capacità delle sottospecialità. Il problema può essere risolto anche senza un'analisi attraverso un modello, ad esempio, l'ospedale può ritenere necessario espandersi in alcune aree di servizio, per una questione di strategia oppure per far fronte all'aumento delle eccedenze temporali nelle sale operatorie (Fig.4).

Tuttavia, in molte situazioni, l'amministrazione dell'ospedale non ha idea di come il tempo dovrebbe essere assegnato nelle sale operatorie (Oliver 2004). In questi casi, può essere opportuno allocare una capacità extra per massimizzare il contributo che l'ospedale è in grado di erogare, sempre mantenendo certi vincoli.

Un insieme di possibili vincoli potrebbero essere, il fatto di non penalizzare nessuna sottospecialità, in termini di assegnazione della capacità. In alternativa, durante una revisione periodica, potrebbe essere necessario imporre vincoli minimi di stanziamento per ciascuna sottospecialità per mantenere la vitalità clinica.

Una formulazione del problema risultante che deve essere affrontato dell'amministrazione dell'ospedale è stato recentemente proposto, in

determinate ipotesi semplificatrici, in un articolo di Dexter, Ledolter e Wachtel (2005) dove viene descritta anche la letteratura in materia.
Tuttavia, Dexter fa certe ipotesi che restringono il campo di applicabilità della sua formulazione a scenari specifici. In quello che segue, innanzitutto viene fornita una formulazione generale del problema.
Si discute una procedura di soluzione, le varianti di questa formulazione e le sfide pratiche saranno discusse nella sezione 1.2.

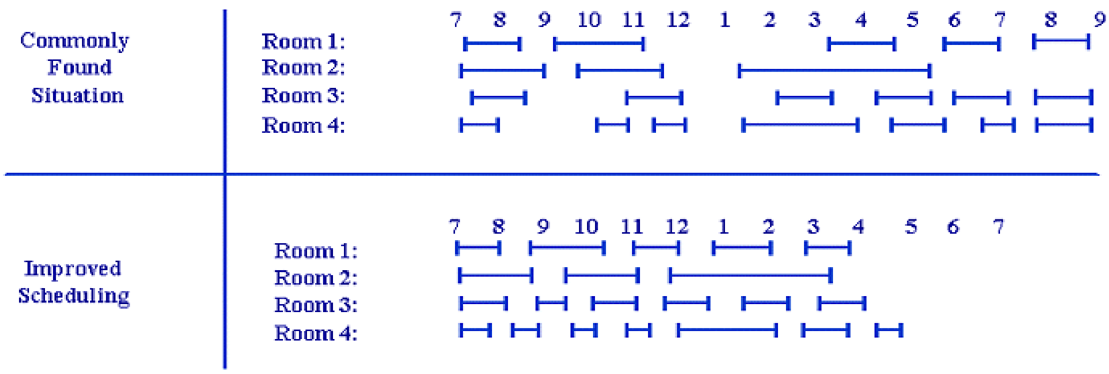

Figura 4. Decremento totale del tempo necessario per gli interventi chirurgici

1.1 Formulazione del modello

Si supponga $i = 1 \ldots m$ sottospecialità (o chirurghi se i blocchi sono assegnati a singoli chirurghi). $\vec{T}=(T_1,\ldots T_m)$ rappresentano le domande di tempo in sala operatoria per ognuna delle *m* sottospecialità.
T_i si presume essere un numero reale non negativo, il guadagno dell'ospedale per ogni unita di tempo assegnata è $\vec{r}$, le assegnazioni esistenti sono $\vec{q}_0$ e la capacità attuale è $k_0 = \sum_{i=1}^{m} q_i^0$.
L'ospedale sta aggiungendo nuova capacità, in modo che, la capacità disponibile totale dopo questa modifica sarà $k \geq k_0$. Il problema che si trova davanti l'ospedale è quello di decidere come allocare questa capacità per massimizzare il suo guadagno totale, pur rispettando determinati vincoli.
Se le nuova allocazione di capacità è indicata con $\vec{q}$, allora, come discusso in precedenza, un possibile vincolo posto sulla possibile decisione di assegnazione può essere che $q_i^0 \leq q_i$ per ogni *i*.
Ulteriori vincoli possono venire dalla disponibilità di valle PACU o dalla capacità di letti. Si organizzano gli indici delle sottospecialità come segue
$r_1 \geq r_2 \geq, \ldots, r_m$. Almeno una delle disuguaglianze precedenti devono essere fissate, altrimenti tutte le sottospecialità possono essere considerate come una specialità singola per il modello.
Partiamo dal presupposto che tutto il tempo inutilizzato delle sale operatorie può essere assegnato a procedure miste che hanno un contributo medio di r per unità di tempo. Inoltre, $r < r_1$ perché altrimenti, sarebbe ottimale destinare tutte le capacità in più a procedure miste.
L'obbiettivo dei manager delle sale operatorie è quello di massimizzare il ricavo *z(q)*

$$\begin{aligned} z(q) &= \sum_{i=1}^{m} E[q_i(t_i \wedge q_i) + r(q_i - t_i)^+] \qquad (1) \\ &= \sum_{i=1}^{m} r_i q_i - \sum_{i=1}^{m} (r_i - r) E(q_i - T_i)^+ \end{aligned}$$

A condizione dei seguenti vincoli:

$$\sum_{i=1}^{m} q_i \leq k_i \quad (2)$$

$$q_i^0 \leq q_i \qquad \textit{per ogni i= 1,……,m} \quad (3)$$

Verrà presentata una tecnica di soluzione nella sezione 1.2, che generalizza l'approccio specificato da Dexter, Ledolter, e Wachtel (2005).

Questa formulazione può fornire una soluzione immediata al problema dell'ospedale. Tuttavia, ci sono una serie di problemi non affrontati nella formulazione che potrebbero richiedere un approccio diverso. Tali questioni sono discusse nella sezione 1.2.

La grandezza del problema (il numero di variabili decisionali) che deve affrontare un ospedale di medie dimensioni dipende dal fatto, se le assegnazioni sono fatte su chirurghi individuali o sulle sottospecialità.

Nel primo caso, la formulazione può avere decine di diverse variabili decisionali, una per ogni chirurgo che ha competenze nell'ospedale mentre nel secondo scenario, il numero di variabili possono essere meno della metà, una per ogni servizio o sottospecialità.

1.2 Approcci di soluzione, estensioni, e sfide aperte

Nell' equazione più importante (Luenberger 1984), esistono q_i, λ, μ_i,….,μ_m con $\lambda \geq 0$, $\mu_i \geq 0$ e

$$r_i \vec{F} i(q_i) + rF_i(q_i) + \mu_i - \lambda = 0 \quad (4)$$

$$\mu_i(q_i^0 - q_i) = 0 \quad \text{per ogni } i=1,\ldots,m \quad (5)$$

$$\lambda(\textstyle\sum_{i=1}^{m} q_i\, k) = 0 \quad (6)$$

in questa equazione di ottimo $F_i\ e \vec{F_i}$ indicano rispettivamente una funzione di distribuzione cumulativa e la complementare, del dominio delle sottospecialità i, con $F_i(0) = 0$.

Dalla (4) otteniamo $\lambda = r_i - (r_i - r)F_i(q_i) + \mu_i$ per ogni $i=1,\ldots,m$. Inoltre siccome $F_i \leq 1$ per ogni i e $r_1 \geq r$ segue che $\lambda \geq r + \mu_1 \geq r > 0$. Un' immediata conseguenza di questa osservazione e dell'equazione (6) è che $\sum_{i=1}^{m} q_i$ deve essere uguale a k. In altre parole, i manager dovrebbero assegnare tutta la capacità tra m sottospecialità, questo è ottimale perché $r_1 > r$.

Ora si svilupperà un algoritmo da risolvere per ottenere il q_i ottimo.

L' algoritmo è iterativo, in cui, si affina sequenzialmente il valore di λ.

Il valore di λ al k-esima iterazione è denotata con la scrittura λ_k e le rispettive assegnazioni di tempo per le sale operatorie sono indicate $q_i^{(k)}s$.

All'inizio della procedura si sceglie un $\lambda_i \in [r_{l1+1}, r_{l1}]$ per l_1 tale che $r \in [r_{l1+1}, r_{l1})$, inoltre $\lambda_1 > r$. Cioè per gli indici $i = 1, \ldots, l_1$, $r_1 > \lambda_1 > r$ e per gli indici $i = l_1 + 1, \ldots, m$, $r_i \leq r < \lambda_i$. Le sottospecialità saranno indicizzate $i > l_i$.

Si considerino le seguenti operazioni sull'equazione (4)

$$\begin{aligned} \mu_i &= -r_i \vec{F_i}(q_i) - rF_i(q_i) + \lambda_i \\ &= (\lambda_i - r_i) - (r - r_i)F_i(q_i) \geq (\lambda_i - r_i) - (r - r_i) > 0. \end{aligned} \quad (7)$$

perciò, dalla (5), $q_i^1 = q_i^0$ per $i = l_1 + 1, \ldots, m$.

Per le restanti sottospecialità, si assume che, una nuova capacità venga aggiunta alla loro assegnazione precedente. Questo si ottiene risolvendo

$$q_i^{(1)} = F_i^{-1}[\frac{(r_i - \lambda_i)}{r_i - r}]$$

per ogni $i = 1, \dots, l_1$.
Se noi troviamo $\lambda_1 \in [r_{l1+1}, r_{l1}]$ tale che

$$\sum_{i=1}^{l_1} F_i^{-1}(r_i - \lambda_i)/(r_i - r). \quad (8)$$

Si avrà l'assegnazione migliore quando le sottospecialità $1, \dots, l_1$ ricevono ulteriori assegnazioni. Questa procedura può essere ripetuta per ogni gamma di valori possibili di λ. Nello specifico questo significa impostare λ nell'intervallo $[r_{l1}, r_{l1-1}), \dots, [r_2, r_1)$, in successione.
Argomenti simili a quelli che si trovano nella disuguaglianza (7) possono essere utilizzati per mostrare che quando $\lambda \in [r_{l+1}, r_l)$ per ogni $l < l_1$, solo le sottospecialità $i = 1, \dots, l$ ricevono assegnazioni di capacità aggiuntiva. Per vedere questo bisogna considerare le variabile nella (7) per $i > l$ tale che $r_i > r$:

$$\begin{aligned} \mu_i &= -r_i \vec{F}_i(q_i) - r_i F_i(q_i) + \lambda \\ &= (\lambda - \mathrm{r_i}) + (\mathrm{r_i} - \mathrm{r})\mathrm{F_i}(\mathrm{q_i}) > 0 \end{aligned} \quad (9)$$

Per $i > l$ tale che $r_i \leq r$, gli argomenti della (7) sono validi. Perciò, la procedura per la soluzione completa comporta un numero finito di passi, ciascuno dei quali può essere risolto in tempi relativamente brevi. La soluzione migliore in assoluto è quella che massimizza il profitto dell'ospedale, $z(\vec{q})$, tra tutte le soluzioni fattibili.

Ci sono una varietà di modi in cui la formulazione nella sezione 1.1 può essere migliorata per soddisfare le realtà che devono affrontare i manager delle sale operatorie. Si presenta prima, un formulazione alternativa che può essere risolta con una variante della metodologia illustrata sopra. Si discutono anche altre formulazioni che potrebbero costituire la base del lavoro futuro in questo ambiente. Supponiamo che i gestori delle sale operatorie possano utilizzare gli avanzi di capacità di qualsiasi sottospecialità, ma solo quando esiste una domanda complementare ad un'altra, capace di approfittare di questa disponibilità lasciata inutilizzata. Quando questo accade, non è generalmente possibile utilizzare utilizzarla a pieno (a causa della perdita di sequenzialità), i manager possono aver bisogno di programmare alcuni straordinari di utilizzo delle sale operatorie, per ottimizzare la resa.

Questo accade quando un intervento chirurgico che di solito occupa 1 ora deve essere inserito in un intervallo di tempo di 45 minuti. L'effetto complessivo di questa inefficienza può essere modellato utilizzando dati storici per stimare il contributo per unità di tempo, r, da interventi chirurgici che vengono programmati in questo modo.

Ha senso pensare che $r \leq r_i$ per ogni i, altrimenti, sarebbe vantaggioso per l'ospedale non assegnare alcun tempo alle sottospecialità e guadagnare un contributo più alto dando loro il tempo di assegnazione su una base a doc.

La capacità in eccesso in un determinato periodo è $k - \sum_{i=1}^{m}(q_i \wedge T_i)$, mentre la domanda di straordinaria complessiva è $\sum_{i=1}^{m}(T_i - q_i)^+$. Usando $T = \sum_{i=1}^{m} T_i$ per denotare la convoluzione sull'uso delle sale operatorie da parte delle sottospecialità, si ottiene la seguente formulazione per i problemi

$$
\begin{aligned}
z(\vec{q}) \quad &= E[\sum_{i=0}^{m} r_i (T_i \wedge q_i)] + rE\{\sum_{i=1}^{m} (T_i - q_i)^+ \wedge (k - \sum_{i=1}^{m} (q_i \wedge T_i)\} \qquad (10) \\
&= \sum_{i=1}^{m} (r_i - r) E(q_i - t_i)^+ + rE(T \wedge k)
\end{aligned}
$$

Influenzata dalle condizioni (2) e (3).

La formulazione qui sopra scritta è simile a quella presentata in precedenza che può essere adattata per risolvere quest'ultima. Tuttavia, poiché la

capacità rimasta risulta minore (perché $r \leq r_i$ per tutte le i, non è possibile utilizzare tutta la capacità in eccesso), ci si aspetta che l'allocazione dei blocchi di tempo sia relativamente elevata. La formulazione (10) si basa sul presupposto che il contributo in eccesso delle sale operatorie è r, indipendentemente da quale sottospecialità fa uso di questa capacità.
Questo rende più facile scrivere la formulazione (10), perché $z(\vec{q})$ è indipendente dalle diversità delle sottospecialità che utilizzano la capacità in eccesso e dalle regole utilizzate per gestire la perdita di sequenzialità. L'approccio è ragionevole dal punto di vista pratico, perché gli ospedali potrebbero liberare lo spazio in eccesso per la chirurgia basata sulla filosofia, primo-arrivato primo-servito, per non dare l'impressione che avvengano favoritismi di una sottospecialità rispetto alle altre.
Tuttavia, aggregando i contributi delle diverse sottospecialità, la formulazione potrebbe potenzialmente nascondere le inefficienze nella gestione degli eccessi delle sale operatorie.
Successivamente, si considerino alcuni miglioramenti che richiedono un diverso approccio di formulazione e soluzione.
In primo luogo, l'ospedale incorre in alcuni costi non ricorrenti riguardanti attrezzature e personale che sono dipendenti dalle assegnazioni di capacità. Tuttavia, la formulazione (1) - (3) ignora tali costi. I costi fissi possono essere importanti per modellare alcune situazioni. I modelli con costi fissi per problemi di ottimizzazione sono significativamente più difficili da risolvere.
In secondo luogo, la nozione di contributo fisso per unità di tempo assegnato è una convenienza di modellazione. In realtà, i servizi chirurgici sono programmati in unità discrete, quindi, l'i-esima sottospecialità chirurgica che può utilizzare un incremento di q_i^0 può influenzare r_i. Ciò significa che per unità, l'entrata è funzione della ripartizione.
Allo stesso modo, l'assegnazione minima può anche dipendere dai ritardi massimi clinicamente accettabili nella programmazione dei casi ad alta priorità. Le funzioni devono essere modellate utilizzando dati storici sui tempi di procedura e sui ritardi della programmazione chirurgica. In terzo luogo, può essere difficile ottenere dati affidabili sulla vera richiesta per ogni sottospecialità.

Stimare la vera domanda è particolarmente difficile nel campo sanitario, ambiente in cui l'offerta indotta dalla domanda è un fenomeno ben documentato (Manning et al 1987.; Evans 1973).

Dexter, Ledolter e Wachtel (2005) affrontano il problema di stima della domanda assumendo che solo le sottospecialità che sono interessate da una più alta media di richiesta di capacità possono beneficiare del contributo di capacità aggiuntiva. Poi, presumono che la domanda per le restanti sottospecialità è uniformemente distribuita tra T_j e $2T_j$, dove $r_j > \vec{r}$ e $\vec{r} = (\frac{1}{m})\sum_{i=1}^{m} r_j$.

Si potrebbe considerare una varietà di approcci diversi per risolvere questo problema. Per esempio, investendo in attrezzature dedicate, flessibili o sull'addestramento del personale; questo potrebbe dare l'opportunità di capire la vera richiesta durante un periodo di prova dopo il quale gli stanziamenti possono essere regolati per trovare una corrispondenza migliore tra domanda e offerta. Problemi di questo tipo possono essere formulati con programmi stocastici ricorsivi a due livelli. Un'ulteriore alternativa può essere progettare sistemi che rivelano “la verità”, dove i chirurghi, che probabilmente hanno informazioni più precise sulle richieste, sono incentivati per segnalare la vera domanda. Gli incentivi non sono monetari. L'ospedale può premiare le segnalazioni veritiere, dando priorità alla sottospecialità che ha una domanda verosimile, oppure togliendo capacità alle sottospecialità che non sono in grado di utilizzare costantemente e completamente il tempo a loro dedicato. La redistribuzione di tempo delle sale operatorie può essere effettuata a intervalli predeterminati (revisione periodica). In alcune strutture ospedaliere, i chirurghi non sono ritenuti responsabili per il sotto utilizzo del blocco di tempo a loro assegnato, con il risultato che non hanno un motivo forte per segnalare la vera domanda. I manager delle sale operatorie possono utilizzare uno schema fornito da un agente-capo con il quale si scambiano informazioni (esempi di questo approccio possono essere trovati in Laffont e Martimort 2002; Bolton e Dewatripont 2005) per modellare il problema e scegliere gli incentivi. In questo approccio, ad ogni chirurgo viene offerto un menù di assegnazioni, con incentivi per un'efficace utilizzo del tempo assegnato. Il problema del manager è quello di trovare un menù di capacità e funzioni di costo corrispondenti, per ogni chirurgo, in modo che sia nel suo interesse scegliere la migliore assegnazione di capacità complessiva, cioè, quello che massimizza il beneficio combinato dell' ospedale, dei chirurghi e dei pazienti. Il problema dell'assegnazione della capacità delle sale operatorie include una domanda multi-prioritaria, preferenze sul fornitore e vincoli di capacità, che spingono la costruzione di nuovi modelli che catturano la realtà del dominio di applicazione.

CAPITOLO 2

Controllo della prenotazione della chirurgia programmata

Quando i chirurghi prenotano spazi di capacità nelle sale operatorie per diverse procedure, molto spesso non esiste un meccanismo per garantire che ci siano adeguate risorse a valle (PACU, ICU, capacità di letti) per la fase post operatoria. Ciò può causare una serie di problemi nel flusso dei pazienti. Per esempio, può portare al bloccaggio delle sale operatorie, all'aumento del tempo di attesa del paziente, all'uso eccessivo del tempo in eccesso e incapacità di gestire le emergenze, che causano la deviazione delle ambulanze in altri ospedali (Litvak et al. 2001, 2005). Descriviamo un modello stilizzato per il controllo delle prenotazioni chirurgiche, che tiene conto della capacità limitata delle risorse critiche a valle.

2.1 Formulazione del modello

Per modellare la risposta del modello di controllo si deve ipotizzare che i chirurghi abbiano deciso di dividere tutti i tipi di interventi chirurgici in n classi. Una classe può essere la combinazione di urgenza, caratteristiche del paziente, entrate, oppure su un'altra base di classificazione dei pazienti. Per le classi di pazienti *i* c'è un ritardo massimo τ_i, le classi sono indicizzate in modo che $1 = \tau_1 \leq \tau_2 \leq \cdots \tau_n$.

Usiamo il giorno come unità di tempo per convenienza. La classe 1 è la classe emergente, cioè i pazienti in questa classe sono serviti il giorno che la loro richiesta viene ricevuta. Uno approccio più sofisticato è descritto da Alter et al. (1999), che utilizza determinanti clinici e non clinici per consigliare un tempo massimo di attesa per una coronoscopia.

Per determinare il giorno in cui prenotare un i-esimo tipo di intervento chirurgico, il modello effettua un compromesso tra il costo di avere piena disponibilità di risorse a valle e il costo di servire in futuro un caso ad alta priorità. Il modello definisce che l'ammontare di capacità a valle (PACU, ICU, o posti letto) dovrebbe essere riservata esclusivamente per i casi più urgenti. Quindi, una richiesta non può essere programmata sulla prima data disponibile. Tuttavia, si assume che l'ospedale ha l'obbligo di soddisfare questa domanda entro il tempo τ_i, anche se questo significa pianificare straordinari a valle. Un tipo *i* di procedura guadagna un contributo medio di r_i. La tempistica media per una cura post-operatoria per l'ospedale è indicata con k (misurata in minuti).

Un modello del controllo del problema di prenotazione chirurgia è difficile da formulare perché, dopo l' intervento chirurgico, i pazienti possono necessitare di cure per diversi giorni in unità a valle e le tempistiche di permanenza non si conoscono sufficientemente. Quindi, ogni intervento chirurgico prenotato consuma una quantità di risorse a valle indefinita. Si semplifica il problema in vari modi. L'obbiettivo non è quello di catturare ogni dettaglio degli interventi chirurgici, piuttosto si vuole dimostrare che i modelli in grado di catturare le caratteristiche che hanno effetto rilevante sulle politiche operative permettano un miglioramento delle politiche finanziarie ed operative.

Una simulazione dettagliata al computer è spesso necessaria per studiare attentamente l'implicazione di tali politiche prima della loro implementazione. Si assuma che $D_{i,t}$ sia la domanda della classe-*i* (misurata in minuti della risorsa critica a valle) che arriva nel periodo *t*. Si supponga una domanda regolare, specificatamente, qualsiasi porzione di $D_{i,t}$ può essere programmata nei giorni $t, t+1, \dots, t+\tau_{\mathrm{n}}$. Inoltre, non si modella la domanda per un periodo multiplo. Questa ipotesi è ragionevole per le zone di guarigione come PACU, ma non per le unità a valle con letti dove i pazienti possono essere curati per diversi giorni. I modelli con domanda multi-periodo rappresentano un problema di ricerca ancora aperto.

Il problema di controllo della prenotazione è modellato su una base di orizzonte *rolling* con una pianificazione dell'orizzonte temporale di $\tau = \tau_n$ all'inizio di ogni iterazione di questo processo. All'inizio di ogni nuovo giorno un giorno è aggiunto alla fine dell'orizzonte di pianificazione e il primo giorno viene cancellato. La situazione di prenotazione dell'ospedale all'inizio di ogni giorno è rappresentata da $s = \{\vec{S}_1, \dots, \vec{S}_{\tau-1}\}$ dove $\vec{S}_t = (S_{1,t}, \dots, S_{n,t})$ e $S_{i,t}$ sono l'ammontare della capacità del giorno-t in precedenza impegnati dal tipo-i di intervento. Imponiamo $s_t = \sum_{i=1}^{n} s_{i,t}$. Poi $(k - s_t)$ è l'ammontare della capacità disponibile per *t* giorni a partire da oggi.

Poiché il $\tau - esimo$ giorno viene aggiunto all'inizio di ogni nuovo giorno, la capacità disponibile su quel giorno è sempre *k*. In altre parole, $\vec{s}_\tau = (0, \dots, 0)$. Pertanto, non è necessario includere $\vec{s}_\tau$ nel sistema vettoriale.

Le variabili decisionali sono $\{a_{i,t}\}, i = 2, \dots, n$ e $t = 1, \dots, \tau$ dove, $a_{i,t}$ è l'ammontare della domanda dell'i-esimo tipo che è programmata per il giorno *t*. Si noti che le domande di tipo-1 devono essere soddisfatte nello stesso periodo in cui arrivano. Pertanto, $a_{11} = d_1$ e $a_{1t} = 0$ per ogni $t > 2$.

La domanda di allocazione deve soddisfare i seguenti vincoli:

- $a_{i,t} \geq 0$ per ogni *i*
- $\sum_{t=1}^{\tau_i} a_{i,t} = D_{i,t}$

A causa dei forti vincoli imposti dai ritardi medici accettabili, è possibile avere molti straordinari in certi giorni. Non si mette un limite alla quantità di straordinario disponibile. Tuttavia, per evitare un eccessivo ricorso al lavoro straordinario, i costi straordinari, indicati con $h(\cdot)$, si assume siano una funzione crescente e convessa.

Vale a dire che il costo unitario del lavoro straordinario è proporzionale alla quantità di straordinario utilizzato. I manager dell'ospedale possono ottenere una buona politica di prenotazione scegliendo una funzione di costo del tempo straordinario che corrisponda allo stato reale delle sale operatorie.

Siano **s** e **s'** indicatori sullo stato di 2 giorni consecutivi, dove **s'** segue **s**; il sistema si evolve secondo le equazioni seguenti,

$$s'_{i,t} = s_{i,t+1} + a_{i,t+1}, \qquad \forall\, i = 1, \ldots, n \quad e \quad t < \tau \tag{11}$$

dove $s_{i,\tau} = 0 \ \forall\, i$. Si è ora in grado di scrivere l'equazione di ottimo di Bellman. In questa equazione *j* denota in numero dell'iterazione ($j = 1,2, \ldots$), $v(\cdot)$ indica la funzione di valore ottimale, e β il tasso di sconto.

$$v_i(s) = E\{\underbrace{max}_{\{a_{j,i}\}}(r_1 D_1 + \sum\nolimits_{i=2}^{n} r_i\left(s_{i,1} + a_{i,1}\right) - \\ h\left(\left(D_1 + \sum\nolimits_{1=2}^{n}\left[s_{i,1} + a_{i,1}\right] - k\right)^{+}\right) + \beta v_{i+1}(s'))\} \tag{12}$$

i termini sul lato destro della (12) possono essere spiegati come segue. I primi due termini rappresentano le attuali entrate di periodo.

Si noti che le entrate delle assegnazioni dei giorni futuri e la domanda futura, al netto degli eventuali costi straordinari, sono inclusi nel termine $v_{j+1}(\cdot)$, che rappresenta il massimo beneficio netto di tutti i futuri interventi chirurgici.

Il terzo termine è la penalità del tempo straordinario del periodo corrente. Risolvere per l'ottimo $a_{i,t}$, soggetto ai vincoli $\sum_{t=1}^{\tau_i} a_{i,t} = D_{i,t}$, può essere difficile.

Per alcune procedure il ritardo massimo ammissibile può essere in mesi, quindi τ, il massimo numero di periodi nella formulazione, può essere in centinaia. Un tipico ospedale esegue centinaia di procedure chirurgiche e molti usano tre livelli di priorità per ogni tipo di procedura, indifferenti, urgenti ed emergenti. Così, il numero delle classi n, può essere da centinaia a migliaia per un problema comune.

2.2 Approcci di soluzione, estensioni, e sfide aperte

Nonostante la complessità della formulazione nella Sezione 2.1, è possibile ricavare alcune utili proprietà di una politica ottimale che forniscono indicazioni sui tipi di implementazione euristica per valutare modelli dettagliati (ad esempio simulazione al computer). Si dedicherà la maggior parte di questa sezione per illustrare come determinare le proprietà utili di una politica ottima. Prima di ciò si presenta una revisione della relativa letteratura.

Bruin, Koole, e Visser (2005) utilizzano un modello di coda per studiare il flusso di pazienti cardiopatici e trovano che l'origine del collo di bottiglia è la mancanza di disponibilità di posti letto a valle nella catena di cura.

Tuttavia, essi non proporranno un metodo per affrontare questo problema.

Chase (2005) sviluppa un modello di simulazione per Vancouver Coastal Healt per ridurre i tempi di attesa, per migliorare l'utilizzo delle risorse e per determinare il giusto numero di risorse necessarie a valle.

Questo modello di simulazione fornisce un base effettiva per testare varie strategie per gestire la domanda e la variabilità dei tempi di procedura.

Tuttavia, non contiene indicazioni sui tipi di politiche di controllo di prenotazione che dovrebbero essere considerate.

Rohleder, Sabapathy e Schorn (2005) usano la programmazione di un obbiettivo per trovare la misura ottimale degli orari dei blocchi per minimizzare la somma dei costi di sotto o sovra-utilizzo delle sale operatorie. Questo è un metodo *trial-and-error* accoppiato con un modello di simulazione ad eventi discreti con l'obbiettivo di trovare la misura ottimale del blocco.

Il problema di controllo di prenotazione non è modellato. In effetti, ci sono un numero limitato di articoli che si occupano della ripartizione della capacità dell'ospedale tra flussi di domanda distinti.

Gerchak, Gupta, e Henig (1996) considerano il problema di riservare capacità chirurgica per casi di emergenza quando le stesse sale operatorie sono usate anche per interventi programmati e per interventi con durata imprecisa. Loro hanno formulato il problema come un programma dinamico stocastico e definiscono la struttura di una politica ottimale. Ai pazienti non vengono dati particolari appuntamenti per la chirurgia, anzi, questi permangono in una coda fino a quando non giunge il loro turno di intervento. Al contrario,nel modello qui descritto si configurano diversi livelli di emergenza e una procedura in cui le domande devono essere assegnate nel giorno in cui arrivano.

Verde, Savin, e Wang (2006) analizzano il problema di pianificazione della domanda per una struttura medica che è condivisa da pazienti esterni, ricoverati e pazienti d'emergenza. La loro analisi si concentra sulla creazione di regole di priorità di un servizio ottimale per l'ammissione di pazienti in un particolare giorno per sostenere gli interventi. Al contrario, nel modello qui descritto l'obiettivo è quello di determinare le migliori modalità per assegnare le richieste di interventi chirurgici che giungono nei giorni futuri. Si illustra un possibile approccio considerando un caso particolare in cui $n = 2$, $\tau_i = 1$ e $D_{i,t} = D_i$; queste sono variabili aleatorie. Da ciò si ottiene che la domanda di tipo-1 deriva da casi urgenti, mentre il tipo-2 si riferisce a tutti i casi non urgenti che devono essere serviti in $\tau_2 = \tau$ giorni dalla richiesta fatta dal chirurgo. Diversi parametri del sistema possono essere semplificati, come risulta avendo un modello a due classi di urgenza. a_t ora rappresenta l'ammontare della domanda di tipo-2 assegnata al giorno t-esimo, mentre tutte le domande di tipo-1 devono essere soddisfatte il giorno in cui arrivano.

Lo stato del sistema può essere rappresentato dal vettore $\vec{s} = (s_1, \dots, s_{\tau-1})$ perche solo la domanda di tipo-2 può essere posticipata.

Infine, si può scrivere (12) come

$$v_i(s) = E\{\underbrace{max}_{\{a_{j,i}\}}(\sum_{i=1}^{n} r_1 D_1 + (s_1 + a_1) - h((D_1 + s_1 + a_1 - k)^+) + \beta v_{i+1}(\vec{s}'))\}. \tag{13}$$

Con il valore dell'iterazione dell'algoritmo si può dimostrare che quando $\beta < 1$, il valore delle funzione converge a $j \to \infty$ (vedi dettagli in Puterman 1994).

Si indica con $v(\vec{s})$ il valore limite della funzione e si lavora solo in questo campo applicativo. Si definisce $s_\tau = 0$ e $b_t = s_t + a_t$.

Successivamente il valore limitato della funzione si può scrivere come segue:

$$v(\vec{s}) = (r_1E(D_1) + E(\underbrace{max}_{\{b_l \geq s_l\}} \{r_2b_1 - h\big((b_1 + D_1 - k)^+\big)+\beta_v(b_2, \dots, b_\tau)\big)\big\}.$$

Questa massimizzazione è limitata dai vincoli $\sum b_t = \sum s_t + D_2$.

$\Phi(\vec{b}) = r_2b_1 - h(b_1 + d_1 - k)^+) + \beta v(b_2, \dots, b_\tau)$, si può dimostrare che $v(\cdot) e\, \Phi(\cdot)$ sono funzioni concave.

Il problema dei gestori delle sale operatorie è di trovare una serie di controlli di prenotazione, rappresentata da $\{b_t\}$, che massimizza le entrate dell'ospedale complessive nel rispetto dei limiti di accesso imposti da τ. Questo modello si comporta bene perché $\Phi(\vec{b})$ è concava e i vincoli sono lineari in b_t. Si noti che $\Phi(\vec{b})$ è in aumento in b_1 finché $b_1 + d_1 \leq k$. Ciò significa che se $s_1 + d_1 < k$, la domanda di tipo-2 dovrebbe essere assegnata al giorno corrente fino a quando il limite di capacità viene raggiunto o tutte le nuove domande di tipo-2 vengono soddisfatte.

In altre parole, si può sempre fissare $b_1 = min\{d_2, k - (s_1 + d_1)\}$ quando $s_1 + d_1 < k$. Questo è intuitivo, perché qualsiasi capacità non utilizzata durante il giorno sparisce alla fine dello stesso. Il problema della prenotazione non è banale solo quando $d_2 > k - (s_1 + d_1)$.

Quando si presenta ciò diventa un problema di massimizzazione di una funzione concava soggetta a vincoli lineari. Pertanto le tecniche standard di ottimizzazione non lineare possono essere applicate allo studio della soluzione ottimale. Utilizzando le condizioni ottimali del primo ordine (che sono necessarie e sufficienti in questo caso; si veda in dettaglio Luenberger; 1984, Sezione 10.8), si ottiene la seguente equazione simulativa

$$\mu_1 = \lambda - (r_2 - \nabla_{b_1} h((b_1 + d_1 - k)^+))$$

$$\mu_k = \lambda - \beta \nabla_{b_k} v(\vec{b})$$

per ogni $k = 2, \dots, \tau$

∇_x indica il gradiente collegato a x, μ, λ. λ non ha restrizioni, $\mu_k \geq 0$ per ogni $k = 1, \dots, \tau$, E almeno una delle seguenti uguaglianze vale per ogni k: $\mu_k = 0$ $b_k = s_k$.

Si suppone una domanda d_2 assegnata ai giorni appartenenti alla serie di indice J. Allora, per ogni *j* e *k* appartenenti alla serie J,$\nabla_{b_j} v(\vec{b}) = \nabla_{b_k} v(\vec{b})$. In altre parole, le allocazioni ottimali sono tali che il beneficio incrementale di assegnare una porzione infinitesimale di d_2 a tutti i giorni futuri è uguale.

Per capire come questo può essere d'aiuto ad identificare una buona politica di controllo delle prenotazioni, si consideri l'entità del beneficio di assegnare una porzione di d_2 al *k*-esimo giorno futuro. Sappiamo che questo beneficio è negativo quando b_k supera una certa soglia e perciò un'ulteriore assegnazione comporta l'uso di straordinari costosi. Si conosce anche, che il vantaggio è positivo quando b_k è piccolo e può essere applicato al giorno k+1con la differenza che il beneficio è scontato dal fattore β. Perciò, quando s_k è piccolo e $s_k = s_{k+1}$ il beneficio incrementale di assegnare al giorno *k* è più grande.

Allo stesso modo, quando, s_k è grande e $s_k = s_{k+1}$ il beneficio incrementale di assegnarlo al giorno *k* è più piccolo. Pertanto, per pareggiare i benefici marginali, si assegna più domanda al $(k+1)$giorno quando s_kè grande.

Le proprietà del criterio di prenotazione ottimale descritte sopra possono essere spiegate a livello intuitivo. Se gli impegni precedenti di capacità sono relativamente piccoli, si assegnano le nuove richieste ai precedenti giorni. Se gli impegni precedenti sono relativamente grandi, si spingerà la domanda nel futuro. La prima ci permette di sfruttare al meglio la capacità disponibile che altrimenti non viene utilizzata. Quest'ultima ci permette di spingere i costi straordinari nel futuro e salvaguardare la capacità in un prossimo futuro per arrivi urgenti. Il secondo approccio fornisce anche più tempo di reazione per regolare i livelli del personale, il quale può aiutare ancor di più a ridurre i costi straordinari. Il modello aiuta dimostrando che il controllo di prenotazione nel giorno t per il $(t+k)$ giorno dovrebbe variare con lo stato attuale delle prenotazioni per quel giorno, in relazione alla quantità di capacità che normalmente è prenotata al tempo t. Un controllo di prenotazione fisso potrebbe non essere ottimale. In particolare, la quantità di capacità disponibile per un intervento chirurgico, con un particolare livello di priorità,

cambia nel tempo ed è possibile che la disponibilità aumenti mentre ci si avvicina al giorno dell'intervento.
Uno svantaggio di questa formulazione è la mancanza di preoccupazione per l'equità. Può succedere che un paziente che arriva dopo, ma che ha identico livello di urgenza rispetto ad altri pazienti venga servito prima. Questo succede quando non c'è domanda per i casi più urgenti, che rilascia a breve termine la capacità per eventuali ritardi. In pratica, questo problema può essere superato informando i pazienti della possibilità di una disponibilità anticipata, consentendo loro di essere inseriti in una lista chiamata "volontaria". Quando la capacità precedente diventa disponibile, si può offrire loro la possibilità di un trattamento anticipato.
Alcuni semplici miglioramenti del modello comprendono l'analisi per il caso in cui $n > 2$. I compromessi (*trade-off)* sono meno trasparenti in questi casi e un analisi numerica dettagliata può essere necessaria per identificare la politica euristica che funziona bene. Il modello sopra scritto può essere visto come un modello di pianificazione della capacità aggregata perché ignora l'interezza, la tempistica e la natura di una domanda multi - periodo di risorse a valle.
All'interno di ogni periodo la natura discreta della domanda e il raggruppamento delle dimissioni dalle sale operatorie può causare picchi ingestibili di domanda delle risorse a valle, anche quando la capacità complessiva nel periodo è sufficiente. Un'altro problema importante è quello di collegare la domanda con i piani di assunzione. Il numero del personale può essere modificato solo in unità discrete(intere). La pianificazione del personale è periodica e spesso gli orari dei turni sono fissati per 4-6 settimane alla volta, questo limita il grado di flessibilità a disposizione dei gestori delle operazioni di rispondere alle mutevoli condizioni di fabbisogno del personale. Collegando il controllo della chirurgia programmabile con i modelli del personale, si può ottenere un sostanziale miglioramento dell'efficienza.

CAPITOLO 3

Sequenzialità interventi programmabili

In questo capitolo verrà affrontato il problema della sequenzialità delle prenotazioni programmabili in un determinato periodo. Questo problema sorge dopo la decisione da parte del manager su quanti interventi programmabili prenotare. E' irrilevante se questo periodo corrisponde ad una prenotazione a blocchi o ad una prenotazione libera.
Nella formulazione che segue si assume che tutte le sequenze possibili sono medicalmente fattibili; quando i vincoli medici sono articolati, questi possono essere aggiunti alla formulazione come vincoli del caso specifico.

3.1 Formulazione del modello

Dato che *p* interventi chirurgici, con durate $\vec{Z} = (Z_1, \ldots, Z_p)$, devono essere eseguiti durante una sessione di lunghezza fissata *d*, un problema concomitante è quello di determinare gli orari di inizio in cui è stato stabilito alle equipe mediche e ai pazienti di presentarsi in sala operatoria.

Partiamo dal presupposto che l'amministratore dell'ospedale abbia una procedura per determinare gli orari di inizio ottimale per ogni *p*, *d* ed una sequenza in cui gli interventi saranno eseguiti (Denton e Gupta, 2003).

Si assume anche che i pazienti e le squadre di servizio arrivino puntualmente all'orario previsto per gli interventi.

Si impone che i_k sia l'indice del k-esimo intervento nella sequenza e x_{i_k} il tempo di tolleranza corrispondente: cioè gli interventi chirurgici vengono eseguiti nella sequenza $\{i_k\}_{k=1}^{p} = (i_1, \ldots, i_p)$, dove la successione $\{i_k\}$ è una permutazione sull'insieme *{1,. . . ,P}*.

Si usa $\Im_p$ per indicare l'insieme di tutte le sequenze possibili quando si hanno *p* procedure chirurgiche. Si noti che occorre specificare i tempi permessi solo per la prima $(p - 1)$ procedura.

Supponendo che la prima procedura inizi puntuale, il tempo d'inizio di a_i può essere determinato da x_i come segue: $a_{i_1} = 0$ e $a_{i_k} = \sum_{i=1}^{k-1} x_{i_j}$ per $k = 2, \ldots p$.

Dato $(\{i_k\}_{k=1}^{p}, \vec{x}, \vec{Z})$, i punti fondamentali di misura della prestazione sono $\vec{W}$, i tempi di attesa sono $\vec{S}$ (tempi inutilizzati delle sale operatorie) e *L* il ritardo rispetto a *d*, che può essere determinato dalle ricorsioni

$$W_{i_k} = (W_{i_{k-1}} + Z_{i_{k-1}} - x_{i_{k-1}})^+, \; k = 2, \ldots p \tag{14}$$

$$S_{i_k} = (-W_{i_{k-1}} - Z_{i_{k-1}} + x_{i_{k-1}})^+, \; k = 2, \ldots p \tag{15}$$

$$L = (W_{i_p} + Z_{i_p} + \sum_{i=1}^{p-1} x_{i_k} - d), \tag{16}$$

dove si settano $W_{i_1} = S_{i_1} = 0$ come conseguenza della puntualità di arrivo di chirurghi e pazienti. Si noti che il tempo di attesa e inutilizzo soddisfano l'equazione $W_{i_1} S_{i_1} = 0$, con $k = 1, \dots p$. I vettori $\vec{c}^w$ e $\vec{c}^s$ specificano i coefficienti di costo per l'attesa da parte del paziente o degli sfaff e i tempi inutilizzati delle sale operatorie. Similmente c_l è l'unità di costo unitario del ritardo rispetto a d.
L'obbiettivo è quello di determinare la sequenza ottimale (al minimo costo), che si denoterà con $(i_1^*, \dots, i_p^*)$. $\Phi(\vec{x}; \{i_1, \dots, i_p\})$, rappresenta il costo totale stimato quando siamo nella sequenza $i_1, \dots, i_p$ e le tolleranze sono definite dal vettore $\vec{x}$. Successivamente

$$\Phi(\vec{x}; \{i_1, \dots, i_p\}) = \sum_{k=1}^{p} c_{i_k}^{w} E[W_{i_k}] + \sum_{k=1}^{p} c_{i_k}^{s} E[S_{i_k}] + c_l E[L]. \quad (17)$$

Il primo compito per i manager delle sale operatorie è quello di trovare le tolleranze $\vec{x}^* = [i_1, \dots, i_p]$ che minimizzano il costo totale presunto $\Phi(\vec{x}; \{i_1, \dots, i_p\})$, per ogni $\{i_1, \dots, i_p\}$. $\Phi(\vec{x}^*[i_1, \dots, i_p])$, indica il minimo costo totale presunto quando la sequenza $\{i_1, \dots, i_p\}$ è rispettata

$$\Phi(\vec{x}^*[i_1, \dots, i_p]) = \underbrace{min}_{\vec{x}} \{\Phi(\vec{x}; \{i_1, \dots, i_p\})\} \quad (18)$$

Infine, il problema di trovare la sequenza ottima può essere espressa come segue

$$\{i_k^*\}_{k=1}^{p} = \arg \underbrace{min}_{\{i_k\}} \Phi(\vec{x}^+ [\{i_k\}] \quad (19)$$

Questo problema è difficile a causa di due ragioni:

- l'insieme $\mathfrak{I}_p$ cresce rapidamente all'aumentare di *p* (si noti che la dimensione di $\mathfrak{I}_p$ è *p!*)
- le tolleranze ottimali dipendono dalla sequenza in cui gli interventi vengono eseguiti.

La procedura per determinare le tolleranze ottimali è un'algoritmo computazionale (vedi Denton e Gupta 2003 per i dettagli). Il numero di interventi chirurgici che possono essere eseguiti in una sessione particolare variano molto da una sottospecialità ad un altra. Tuttavia, per molte sottospecialità, una tipica sessione può essere in grado di ospitare interventi chirurgici al massimo in doppia cifra.

CAPITOLO 4

Esempi numerici

Di seguito vengono proposti due esempi numerici modellati con l'ausilio del programma informatico *Gams*.

Il primo esempio riguarda il modello di massimizzazione del profitto in funzione delle assegnazioni (1) mentre il secondo riguarda la scelta del giorno ottimo in cui prenotare un tipo di intervento chirurgico nell'ambito della prenotazione programmata in riferimento ai costi (12).

4.1 Esempio numerico 1

```
SETS
i sottospecialità  / sp1 * sp5 /;

PARAMETERS
T(i) domande di tempo in sala operatoria
                / sp1 10, sp2 9,   sp3 9,  sp4  8,  sp5  7 /

ri(i) guadagno dell'ospedale per ogni unità di tempo
                / sp1 8, sp2 7, sp3 6, sp4 4, sp5 2 /

qo(i) assegnazioni iniziali
                / sp1 3, sp2*sp3 6, sp4 5, sp5 2/;

SCALARS
K   capacità nuova / 30 /
r   contributo per tempo inutilizzato nelle sale operatorie / 3 /
Ko  capacità iniziale;
Ko:=sum(i,qo(i));

VARIABLES
q(i) variabili di allocazione di capacità
z funzione obiettivo: ricavo totale ;

POSITIVE VARIABLES q(i);

EQUATIONS
assegnazione(i)   vincolo sulla decisione di assegnazione
capacita   vincolo sulla massima capacità consentita
funz_ob ;


assegnazione(i)..  q(i) =G= qo(i)    ;

capacita..  sum(i, q(i)) =L= K ;

funz_ob ..  z =E= sum ( i , ri(i)*q(i)) - sum(i, (ri(i)-r)*(q(i)-T(i)) ) ;


MODEL AllocazioneCapacita  /all/ ;

option optcr = 0;

SOLVE AllocazioneCapacita using MIP maximizing z;


DISPLAY q.L;
```

```
E x e c u t i o n


----      49 VARIABLE q.L  variabili di allocazione di capacità

sp1  3.000,    sp2  6.000,    sp3  6.000,    sp4  5.000,    sp5 10.000


----      49 VARIABLE z.L                  =      204.000  funzione obiettivo: r
                                                           icavo totale


EXECUTION TIME       =        0.000 SECONDS      3 Mb  WIN232-232 Aug 11, 2009
```

```
               S O L V E      S U M M A R Y

     MODEL   AllocazioneCapacita   OBJECTIVE  z
     TYPE    MIP                   DIRECTION  MAXIMIZE
     SOLVER  CPLEX                 FROM LINE  46

**** SOLVER STATUS     1 Normal Completion
**** MODEL STATUS      1 Optimal
**** OBJECTIVE VALUE              204.0000

 RESOURCE USAGE, LIMIT          0.015      1000.000
 ITERATION COUNT, LIMIT         0    2000000000

ILOG CPLEX       Aug 14, 2009 23.2.1 WIN 12361.12582 VIS x86/MS Windows
Cplex 12.1.0, GAMS Link 34

LP status(1): optimal
Optimal solution found.
Objective :         204.000000
```

4.2 Esempio numerico 2

Si intenda con z, la funzione obbiettivo v nella formulazione(12).

```
set i /1*3/
t /1*17/
l(t) /1*7/
j(l) /1*6/
;
scalar n /5/
k /480/;

*classi: 1 rosso, 2 giallo, 3 verde

parameters classi(i) durate interventi /1 1, 2 2, 3 3/
ritardo_tau(i) / 1 1, 2 2, 3 4/
r(i) contributo medio per l'i esimo tipo di procedura /1 10, 2 15, 3 20/;

table domande(i,t)
        1       2       3       4       5       6       7       8       9       10
1       5       3       1       1       7       10      1       6       1       1
2       12      40      34      29      36      36      37      43      43      42;

variable a(i,t) domanda di tipo h programmata per il giorno l;
variable s(i,j) ammontare della capacità del giorno t;
variable s1(i,j);
variable z;
*positive variable a;
positive variable s, s1;
s.up(i,j)=480;
equations eq1(t), eq2(t), eq3(i,j), ob;

eq1(t).. a('1',t)=E= domande('1',t);
eq2(t).. a('2',t)+a('2',t+1)+a('2',t+2)=E= domande('2',t);
eq3(i,j).. s(i,j)=E=s(i,j-1)+a(i,j);
ob.. z=E=  r('1')*(sum(t, domande('1',t)))+ (r('2')*(s('2','1')+a('2','1')))
-(sum(t, domande('1',t))+(s('2','1')+a('2','1'))-480);

model esempio2/all/;
solve esempio2 using nlp maximizing z;
display a.l, s.l;
```

```
E x e c u t i o n


----      36 VARIABLE a.L  domanda di tipo h programmata per il giorno 1

            1           2           3           4           5           6

1       5.000       3.000       1.000       1.000       7.000      10.000
2       1.000       6.000       5.000      29.000

+           7           8           9          10

1       1.000       6.000       1.000       1.000
2      36.000                   1.000      42.000


----      36 VARIABLE s.L  ammontare della capacità del giorno t

            1           2           3           4           5           6

1       5.000       8.000       9.000      10.000      17.000      27.000
2       1.000       7.000      12.000      41.000      41.000      41.000


----      36 VARIABLE z.L                   =      832.000
```

```
               S O L V E      S U M M A R Y

     MODEL   esempio2            OBJECTIVE  v
     TYPE    NLP                 DIRECTION  MAXIMIZE
     SOLVER  CONOPT              FROM LINE  35

**** SOLVER STATUS     1 NORMAL COMPLETION
**** MODEL STATUS      1 OPTIMAL
**** OBJECTIVE VALUE              832.0000

 RESOURCE USAGE, LIMIT          0.090      1000.000
 ITERATION COUNT, LIMIT         4         10000
 EVALUATION ERRORS              0             0


    C O N O P T 3   x86/MS Windows version 3.14R-017-061
    Copyright (C)   ARKI Consulting and Development A/S
                    Bagsvaerdvej 246 A
                    DK-2880 Bagsvaerd, Denmark

 Using default options.


 ** Optimal solution. There are no superbasic variables.
```

CONCLUSIONI

Questa ricerca ha portato all'individuazione dello scenario clinico critico, nel quale modelli matematici e tecniche operative hanno un ruolo fondamentale per l'ottimizzazione di prestazioni e costi dell'unità ospedaliera. Si può così ottenere un servizio sempre più professionale e mirato alle varie esigenze del paziente. Sono state descritte una serie di decisioni che possono influenzare l'efficienza delle sale operatorie, i tempi di attesa e la sicurezza del paziente.

In particolare sono stati trattati 3 diversi modelli con finalità di seguito riportate.

Attraverso l'utilizzo di ipotesi semplificative e vincoli è stato possibile scrivere un modello per massimizzare il profitto in riferimento all'allocazione di capacità per interventi programmabili e quindi ottenere informazioni utili al raggiungimento di tale scopo.

In secondo luogo, sempre nell'ambito della chirurgia programmabile viene fornito un modello stilizzato per il controllo delle prenotazioni chirurgiche il quale tiene conto della capacità limitata delle risorse critiche a valle, effettuando un compromesso tra il costo di avere piena disponibilità di risorse a valle e il costo di servire in futuro un caso ad alta priorità.

Inoltre è stato descritto un modello per trovare la sequenza ottima (al minimo costo) in relazione a vari interventi chirurgici, attraverso una realtà ospedaliera semplificata.

Nell'ultimo capitolo vengono proposti due semplici esempi numerici riguardanti i modelli dei capitoli 1 e 2, con lo scopo di completare una visione generale sui problemi affrontati.

La letteratura delle sale operatorie è in gran parte assente in relazione a questa classe di problemi, anche se ci sono stati alcuni tentativi di capire il problema attraverso analisi statistiche(vedi, ad esempio, Dexter et al. 2004).

Tentativi per fornire modelli di soluzioni a tali problemi possono avere un impatto significativo sul miglioramento di utilizzo e di efficienza delle sale operatorie, potenziando la qualità e riducendo i costi.

RINGRAZIAMENTI

Ringrazio il Professor Giorgio Romanin Jacur per la sincera disponibilità e cortesia mostratami durante lo sviluppo di questo lavoro e l' Ingegnere Marzia Viali dell'ospedale di Verona, per avermi fornito materiale utile ai fini dell'elaborato.

Un ringraziamento speciale a mamma Annamaria e papà Paolo che con il loro sostegno fondamentale mi sono sempre stati vicini nei momenti di gioia e di difficoltà offrendomi l'opportunità di giungere a questo traguardo, dandomi sempre motivo di guardare avanti.
Ringrazio tutti i miei familiari in particolare i nonni che mi hanno sempre offerto validi consigli con la maturità di una vita alle spalle.

Ringrazio gli amici di sempre e di università con i quali condivido ogni giorno le mie emozioni, i quali mi hanno sempre trasmesso molto affetto e autenticità e che hanno sempre avuto il tempo di consigliarmi nei momenti di difficoltà. Un grazie particolare a Pietro per tutto quello che ha fatto per me in questi anni, un grazie a Sebastiano che con il suo sorriso mi lascerà sempre un ricordo indelebile di quelle giornate trascorse condividendo la nostra giovinezza e la nostra spensieratezza.

Un ringraziamento all'insegnante Angela per l'aiuto e la disponibilità per farmi comprendere al meglio la lingua Inglese, salutandomi sempre con quel "Hello Marco!" che non dimenticherò.

Ringrazio le signore della segreteria di Vicenza per la loro professionalità, e le signore della mensa per la loro simpatia e abbondanza nelle porzioni.

BIBLIOGRAFIA e SITOLOGIA

- Blake, J. T., M. W. Carter. 1997. Surgical process scheduling: A structured review. Journal of the Society for Health Systems 5(3) 17–30.

- Bolton, P. and Dewatripont, M. (2005) Contract Theory, The MIT Press, Cambridge, MA.

- Bruin, A. M., G. M. Koole, M. C. Visser. 2005. Bottleneck analysis of emergency cardiac in-patient flow in a university setting: An application of queueing theory. Clinical and Investigative Medicine 28(6) 316 –317.

- Chase, M. 2005. Beginning patient flow modeling in Vancouver Coastal Health. Clinical and Investigative Medicine 28(6) 323–325.

- Dott. Sergio Pierdominici, Complesso ospedaliero S. Giovanni addolorata, Roma. Appunti "Esperienza triennale nella gestione settimanale delle liste operatorie elettive di un blocco operatorio". Roma, 21 aprile 2010, dipartimento Anestesia e Rianimazione.

- Dott. Sergio Pierdominici, Appunti. Istituto internazionale di ricerca, Milano 2011.

- D. Gupta and B. Denton. Appointment scheduling in health care: Challenges and opportunities. IIE Transactions (2008) 40, 800–819

- D. Gupta B. and Denton. A sequential bounding approach for optimal appointment scheduling. IIE Transactions, 35(11) 1003–1016.

- Evans, R. G. 1973. Modeling the economic objectives of the physicians. Health Economics Symposium: Proceedings of the First Canadian Conference. Queens University Press, Kingston, Ontario, 33– 46.

- F. Dexter, J. Ledolter, and R.E. Wachtel. Tactical decision making for selective expansion of operating room resources incorporating ¯financial criteria and uncertainty in subspecialties' future workloads. Anesthesia and Analgesia, 100(5) 1425–1432.

- F. Dexter, R. Epstein, R. Traub, Y. Xiao. 2004. Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. Anesthesiology 101(6) 1444 –1453.

- Gerchak, Y., Gupta, D. and Henig, M. (1996) Reservation planning for elective surgery under uncertain demand for emergency surgery. Management Science, 42(3), 321–334.

- Green, L. V., S. V. Savin, B. Wang. 2006. Managing patient service in a diagnostic medical facility. Operations Research 54(1) 11–25.

- http://www.surgerymanagement.com/presentations/operating-room-scheduling.php.

- Laffont, J. J., D. Martimort. 2002. The theory of incentives: The principal-agent model. Princeton University Press, Princeton, New Jersey.

- Litvak, E., P. I. Buerhaus, F. Davidoff, M. C. Long, M. L. McManus, D. M. Berwick. 2005. Managing unnecessary variability in patient demand to reduce nursing stress and improve patient safety. Joint Commission Journal on Quality & Patient Safety 31(6) 330 –338.

- Luenberger, D. G. 1984. Linear and Nonlinear Programming. 2nd ed. Addison-Wesley, Reading, Massachusett, 314 -315.

- Luenberger, D. G. 1984. Linear and Nonlinear Programming. 2nd ed. Addison-Wesley, Reading, Massachusett.

- Magerlein, J. M., J. B. Martin. 1978. Surgical demand scheduling: A review. Health Services Research 13(4) 418 – 433.

- Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, A. Leibowitz, M. S. Marquis. 1987. Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review 77(3) 251–277.

- Oliver, K. 2004. OR expansion—A journey into the unknown. AORN Journal 79(2), 369 –372.

- Brecht Cardoen, Erik Demeulemeester and Jeroen Beliën, Operating room planning and scheduling: A literature review. Department of decision sciences and information management (KBI), Faculty of Business and Economics, Katholieke Universiteit Leuven.

- Puterman, Martin L. 1994. Markov decision processes: Discrete stochastic dynamic programming. Wiley, New York.

- D. Gupta, Surgical Suites Operations Management, Production and operations management, Vol. 16, No. 6, 2007, pp. 689-700.

- T.R. Rohleder, D. Sabapathy, and R. Schorn. An operating room block allocation model to improve hospital patient. Clinical and Investigative Medicine, 28 (6): 353-355, 2005.

- Viapiano, J., D. S. Ward. 2000. Operating room utilization: The need for data. International Anesthesiology Clinics 38(4) 127-140.