



*DIPARTIMENTO DELLA PROGRAMMAZIONE E DELL'ORDINAMENTO
DEL SERVIZIO SANITARIO NAZIONALE*

*DIREZIONE GENERALE DELLA PROGRAMMAZIONE SANITARIA
Ufficio III ex D.G.PROGS*

Manuale di formazione per il governo clinico: Appropriatezza



Luglio 2012

©

Indice

Presentazione	3
1. Background, definizioni, modelli concettuali	5
2. Variazioni della pratica clinica: concetti e metodi	14
3. La misura dell'appropriatezza	26
4. Approfondimenti	37
The NHS Atlas of Variation in Healthcare	37
L'appropriatezza dei ricoveri in Emilia-Romagna.....	47
Normativa	57
Bibliografia	58

Presentazione

Il presente corso costituisce il secondo modulo del programma formativo predisposto dall'Ufficio III della Direzione generale della Programmazione sanitaria del Ministero della salute, in condivisione con Fnomceo e Ipasvi, predisposto sulla base di un documento prodotto dall'Agenzia Sanitaria e Sociale regionale della Regione Emilia-Romagna..

Il *buon funzionamento* dei sistemi sanitari tradotto in termini di efficacia, efficienza e equità e la loro tenuta, sempre più spesso, si rapporta alla capacità di determinare e identificare le cure necessarie, minimizzando fenomeni di inappropriatazza.

La responsabilità di tutelare il servizio sanitario nazionale, pone, infatti, la necessità di focalizzare alcune priorità tra cui la valutazione dell'efficacia, la riorganizzazione delle modalità e dei setting di erogazione delle cure, la realizzazione di meccanismi di integrazione e coordinamento ai vari livelli del sistema nonché la continua interazione con la collettività e la partecipazione attiva dei pazienti.

L'aumento dei costi sanitari, inoltre, contestualmente all'esigenza di promuovere qualità e sicurezza, colloca la valutazione dell'appropriatezza al centro delle politiche sanitarie nazionali, regionali e locali.

È ormai diffusa la cognizione che l'appropriatezza delle cure può essere misurata e costituire quindi la base per compiere le scelte migliori sia per il singolo paziente sia per la collettività.

Tuttavia il ricorso inappropriato ad alcune prestazioni, come l'utilizzo improprio dei ricoveri ospedalieri è ancora molto esteso e ciò, sommato ai differenti livelli di qualità dei servizi sanitari sul territorio nazionale, rappresenta un fattore di criticità da non sottovalutare ai fini della sostenibilità del sistema e della sua equità. Il ruolo dei professionisti che operano sul campo è fondamentale per raggiungere livelli sempre più elevati di appropriatezza clinica ed organizzativa e pertanto va promosso e sviluppato. In tal senso, il presente corso si pone come supporto agli operatori sanitari per approfondire le proprie conoscenze sul tema dell'appropriatezza clinica ed organizzativa.

Francesco BEVERE
DIRETTORE GENERALE

Introduzione

Il corso si propone di riassumere e organizzare le principali conoscenze ad oggi disponibili in merito a un aspetto della qualità dell'assistenza che, sia pure comunemente citato e discusso, anche nell'ambito di leggi e norme, non sempre viene compreso e trattato con sufficiente chiarezza, sia per quanto riguarda l'inquadramento concettuale sia per quanto attiene a definizioni operative e interventi di miglioramento collegati.

Il corso si articola in quattro sezioni. La prima affronta gli aspetti semantici e chiarisce definizioni e modelli concettuali, anche attraverso l'analisi comparativa di dimensioni correlate all'appropriatezza, come l'efficacia, l'equità, l'accettabilità.

La seconda sezione esplora il problema della variabilità (geografica, organizzativa) frequentemente osservata nelle attività clinico-assistenziali e nei comportamenti professionali e la sua relazione con il concetto di appropriatezza; esamina inoltre l'appropriatezza secondo alcune specifiche chiavi di lettura: la qualità delle evidenze disponibili, il punto di vista dei pazienti, i valori della società.

La terza sezione sviluppa il tema della misura dell'appropriatezza, con riferimento prevalente alla metodologia messa a punto dalla RAND.

Infine, la quarta sezione propone alcune schede esemplificative, tratte dalla letteratura internazionale e da alcune esperienze, anche condotte in Italia.

1. Background, definizioni, modelli concettuali

Evoluzione e inquadramento del concetto di appropriatezza

Questa prima sezione affronta gli aspetti semantici e chiarisce definizioni e modelli concettuali, attraverso l'analisi comparativa di dimensioni correlate all'appropriatezza.

Poiché l'appropriatezza è un concetto complesso e multidimensionale, è possibile individuare alcune prospettive di lettura che ne identificano altrettante componenti e dimensioni, come il livello di evidenza scientifica, il punto di vista del paziente, i valori della società.

Il concetto di appropriatezza

Secondo il dizionario Zingarelli della lingua italiana è “appropriato” ciò che risulta adeguato (ad esempio a una circostanza), preciso (ad esempio in risposta a determinati bisogni), calzante (soprattutto con riferimento a espressioni verbali o linguistiche), opportuno (aggiungendo ai significati precedenti una sfumatura di buon senso o buon gusto).

In generale, al centro del concetto risiede dunque l'idea che un'espressione verbale o un comportamento, ma anche un'azione, una scelta, una decisione, sia appropriata quando risulti “adatta, conveniente, giusta” per la situazione in cui si verifica.

Il concetto di appropriatezza che si è andato sviluppando nell'ambito dei servizi e dell'assistenza sanitaria trova ovvie radici in questi significati; in termini più specifici, una cura può considerarsi appropriata quando sia associata a un beneficio netto o, più precisamente, quando è in grado di massimizzare il beneficio e minimizzare il rischio al quale un paziente va incontro quando accede a determinate prestazioni o servizi. Inoltre, nella storia naturale di qualunque condizione o trattamento, è possibile fissare un punto *prima del quale* è inappropriato intervenire o fornire un servizio e, analogamente, è possibile fissare un punto *oltre il quale* un servizio diventa inappropriato, in gran parte a causa della perdita di benefici in quella determinata fase di malattia o di bisogno di assistenza.

L'appropriatezza

comporta implicazioni dirette e indirette riguardanti
la procedura corretta sul paziente *giusto* al momento opportuno e nel *setting* più adatto

Non si tratta, in definitiva, di una dimensione semplice e riconducibile a un'unica

interpretazione, valida in assoluto; al contrario, indica condizioni complesse, contesto-dipendenti, che devono essere di volta in volta collocate nel rispettivo ambito di riferimento, definite e articolate in termini operativi, riferite ai diversi soggetti coinvolti e misurate con metodi specifici.

Definizioni di appropriatezza

Una revisione di letteratura pubblicata nel 2008 ed estesa a un periodo di osservazione molto ampio, compreso tra il 1966 e il 2006, ha identificato quattordici articoli che proponevano una definizione di appropriatezza, generalmente intesa in termini di effetti positivi per il paziente.

La nozione di effetti positivi è stata declinata con maggior precisione dai ricercatori dell'organizzazione americana **RAND**, che hanno introdotto i concetti di “rischi” e “benefici”. La definizione proposta dalla **RAND** è ad oggi la più ampiamente usata a livello internazionale.

Definizione RAND:

una procedura è **appropriata se**:

il beneficio atteso (ad es. un aumento della aspettativa di vita, il sollievo dal dolore, la riduzione dell'ansia, il miglioramento della capacità funzionale) supera le eventuali conseguenze negative (ad es. mortalità, morbosità, ansia, dolore, tempo lavorativo perso) con un margine sufficientemente ampio, tale da ritenere che valga la pena effettuarla

Al contrario, viene considerata **inappropriata** una procedura il cui rischio sia superiore ai benefici attesi.

Robert H. Brook ha enfatizzato la necessità di disporre di uno strumento capace di misurare l'appropriatezza delle cure e l'ha motivata con “*la preoccupazione che la crescente complessità delle cure mediche si traduca, per alcuni pazienti, nel mancato godimento di cure necessarie e, in altri, nel sottoporsi a procedure inutili*”.

Queste definizioni, proposte tra gli anni '80 e gli anni '90, non tengono in considerazione il problema dei costi; con il passare del tempo, la necessità di contenere il peso crescente dei costi in sanità, ha indotto a considerare la variabile “costi” parte integrante del concetto di appropriatezza.

Negli Stati Uniti questo cambiamento di direzione è stato descritto da Sharpe.

Prima degli anni '80, l'interpretazione del termine “appropriato” era sostanzialmente condivisa e indicava l'insieme di cure in grado di rispondere adeguatamente ai bisogni del singolo paziente.

In seguito, il mutato contesto economico ha richiesto ai *third-payer* (finanziatori pubblici e privati delle cure sanitarie) di porre un freno al rimborso dei servizi e delle prestazioni fornite.

Alla luce di ciò, è stata proposta una definizione di appropriatezza più complessa intesa come “*i risultati di un processo decisionale che assicura il massimo beneficio netto per la salute del paziente, nell’ambito delle risorse che la società rende disponibili*”.

Nonostante gli sforzi compiuti per definire l’appropriatezza in termini di rapporto tra rischi e benefici, il concetto continua a restare difficile da chiarire. A seguito di una breve *review* delle evidenze disponibili per definire l’appropriatezza, Naylor concluse che il giudizio su di essa *dipende dai medici a cui viene rivolto il quesito, dal luogo dove essi vivono e lavorano, dal peso attribuito alle diverse tipologie di evidenze e di end-point indagati, dal fatto che le preferenze di pazienti e famiglie entrino o meno a far parte del ragionamento, dall’entità delle risorse in un determinato sistema sanitario e infine dai valori prevalenti sia nel sistema stesso sia nella società in cui si è sviluppato*.

In Europa il concetto di appropriatezza delle cure compare in documenti istituzionali nel 1997. Nella Raccomandazione n. 17 del Comitato dei Ministri del Consiglio d’Europa, essa viene elencata tra “*le componenti potenziali della qualità dell’assistenza sanitaria*”; nella Convenzione di Oviedo sui Diritti dell’Uomo e la biomedicina, all’art.3, gli Stati firmatari della Convenzione si impegnano a garantire un “*accesso equo a cure della salute di qualità appropriata*”

Nel 2000, a seguito del workshop organizzato dall’Ufficio regionale europeo della *World Health Organization* (WHO) in collaborazione con il Ministero della Salute tedesco, la definizione di appropriatezza proposta dalla **RAND Corporation** viene riconosciuta come un utile punto di partenza per ragionare intorno al significato di tale concetto, sebbene ne vengano sottolineate la genericità e i limiti.

Secondo un *report* inglese preparato per il *Director of Research and Development of the Department of Health*, i limiti di tale definizione vengono ricondotti alla mancata considerazione della individualità del paziente e della disponibilità di risorse per l’assistenza sanitaria.

Viene proposta la seguente definizione:

Una cura appropriata consiste nella selezione, sulla base degli interventi di cui è stata dimostrata l'efficacia per un determinato disturbo, dell'intervento con la maggiore probabilità di produrre gli esiti di salute attesi da quel singolo paziente. Un intervento può essere appropriato soltanto quando vengono soddisfatti determinati criteri. Devono essere disponibili le competenze tecniche e tutte le altre risorse necessarie alla realizzazione di una prestazione con uno standard sufficientemente elevato. Le modalità con cui l'intervento viene svolto devono essere accettabili dal paziente. Ai pazienti dovrebbero essere fornite adeguate informazioni sul range degli interventi di provata efficacia pratica. Le loro preferenze sono centrali nella scelta dell'intervento appropriato tra quelli conosciuti come efficaci. Le loro preferenze rifletteranno non solo il primario esito di salute che sperano di raggiungere, ma anche i potenziali effetti avversi che si potrebbero verificare. Ne consegue che il paziente deve essere totalmente coinvolto nella discussione riguardante la probabilità dei differenti esiti, con e senza l'intervento, ed i disagi e gli eventi avversi che si potrebbero verificare. L'appropriatezza degli interventi sanitari deve anche essere considerata all'interno dell'attuale contesto sociale e culturale e rispetto alla giustizia nell'allocazione delle risorse

Per quanto riguarda il nostro Paese, il termine “appropriatezza” ha acquistato una rilevanza normativa con il Piano Sanitario Nazionale 1998-2000, divenendo uno dei criteri per la definizione dei Livelli Essenziali di Assistenza (D.Lgs 229/99). In particolare, l'attenzione si è focalizzata sull'appropriatezza organizzativa, *in primis* dell'assistenza ospedaliera, di cui le liste di DRG a rischio di inappropriata rappresentano una delle principali traduzioni operative.

L'appropriatezza clinica è oggetto di una specifica attività da parte del Sistema Nazionale Linee Guida (SNLG) già Programma nazionale linee guida, previsto dal Piano Sanitario Nazionale 1998-2000 e coordinato dall'Istituto Superiore di Sanità.

Nei documenti legislativi e di programmazione risulta, tuttavia, particolarmente complesso rintracciare una chiara definizione di appropriatezza delle cure che ne identifichi le specificità rispetto ad altri concetti collegati, quali l'efficacia e l'efficienza. Essa è recuperabile, invece, dal Glossario a cura del Ministero della Salute.

L'appropriatezza definisce un intervento sanitario (preventivo, diagnostico, terapeutico, riabilitativo) correlato al bisogno del paziente (o della collettività), fornito nei modi e nei tempi adeguati, sulla base di standard riconosciuti, con un bilancio positivo tra benefici, rischi e costi

Riportiamo di seguito le principali definizioni proposte nella letteratura e nella manualistica italiana:

- a) si considera appropriata qualunque prestazione sia in grado di racchiudere in sé due qualità primarie, come *l'efficacia e l'efficienza*, e al tempo stesso possa essere considerata accettabile sia da chi la riceve che da chi la eroga (*Tonelli S*)
- b) l'appropriatezza clinica o specifica è la misura in cui un particolare intervento è sia efficace sia indicato per la persona che lo riceve. Perché un intervento sia appropriato è necessario che i benefici attesi siano superiori ai possibili effetti negativi e ai disagi eventualmente connessi alla sua messa in atto. Occorre anche tenere conto dei costi: tra due interventi per la stessa condizione, un intervento sarebbe da considerarsi meno appropriato se, a fronte di un'efficacia di poco superiore, il suo costo fosse decisamente troppo alto. L'appropriatezza clinica si gioca nel processo decisionale e chiama in causa dunque il giudizio clinico. L'appropriatezza organizzativa o generica è riferita al livello di effettuazione dell'intervento (es. ospedale, domicilio, ambulatorio); per appropriatezza del livello di effettuazione si intende se un intervento (che può essere più o meno appropriato dal punto di vista clinico o specifico) viene erogato al livello più gradito dall'utente e/o meno costoso (*Morosini PL*)
- c) [appropriatezza] componente della qualità dell'assistenza che fa riferimento a validità tecnico-scientifica, accettabilità e pertinenza (rispetto a persone, circostanze e luogo, stato corrente delle conoscenze) delle prestazioni sanitarie (*Zanetti M., Montaguti U., Ricciarelli G., et al*)
- d) per appropriatezza clinica si intende l'utilizzo di un intervento sanitario efficace in pazienti che ne possono effettivamente beneficiare in ragione delle loro condizioni cliniche, mentre per appropriatezza organizzativa si intende l'erogazione di un intervento/prestazione in un contesto organizzativo idoneo e congruente, per quantità di risorse impiegate, con la complessità dell'intervento e con le caratteristiche cliniche del paziente (*Berti E., Casolari L., Cisbani L. et al*)
- e) una prestazione [appropriata] dovrebbe essere erogata rispettando il quadro clinico del paziente e le indicazioni per le quali si è dimostrata efficace, nel momento giusto e secondo il regime organizzativo più adeguato. Le variabili da tenere in considerazione vanno oltre le caratteristiche specificamente cliniche del paziente, come ad esempio la qualità di vita. Inoltre, fanno sempre parte della definizione di appropriatezza la considerazione del contesto culturale e organizzativo sia del paziente che degli operatori e il quadro delle compatibilità *hic et nunc* (*Vasselli S., Filippetti G., Spizzichino L*)
- f) [l'appropriatezza] è la misura della capacità di fornire, al livello di assistenza più idoneo, un intervento efficace per uno specifico soggetto, nel rispetto delle sue preferenze individuali e dei suoi valori culturali e sociali. È utile distinguere tra appropriatezza clinica, intesa come la capacità di fornire l'intervento più efficace, date le caratteristiche e le aspettative di un determinato paziente, e appropriatezza organizzativa, intesa come capacità di scegliere il livello assistenziale più idoneo alla erogazione delle cure (*Cinotti R. et al*).

La tabella sottostante presenta le principali definizioni di appropriatezza proposte dalla letteratura internazionale

Studio/Fonte	Paese	Elementi chiave della definizione	Prospettiva
Woodward et al (1984)	USA	Effetti positivi attesi per il paziente	Paziente
Hopkins (1993)	UK	Maggiore probabilità di ottenere gli esiti di salute attesi dal singolo paziente Accettabilità delle modalità dell'intervento da parte del paziente Scelta dell'intervento appropriato coinvolgendo il paziente e tenendo conto non solo degli esiti di salute ma anche dei rischi Deve tenere conto delle risorse disponibili, del contesto sociale e culturale	Paziente Sistema sanitario Società
Sharpe et al. (1996)	USA	Benefici clinici attesi per il paziente. Rapporto costi-benefici tale da consentire l'erogazione dei servizi in uno (specifico) contesto di risorse limitate.	Paziente Società
Zanetti et al. (1996)	Italia	Accettabilità e pertinenza rispetto a persone, circostanze e luogo, conoscenza	Paziente Sistema sanitario Società
Morosini (1999) Morosini, Perraro (2001)	Italia	Benefici attesi superiori alle conseguenze negative. Deve tenere conto dei costi. Inerente il processo decisionale, chiama in gioco il giudizio clinico Livello organizzativo di effettuazione dell'intervento più gradito dall'utente e/o meno costoso	Paziente Sistema sanitario
RAND (2000)	USA, WHO Europe	Benefici attesi superiori alle possibili conseguenze negative	Paziente
Berti et al. (2004)	Italia	Beneficio effettivo per il paziente in ragione delle condizioni cliniche Quantità di risorse impiegate Contesto organizzativo dell'intervento	Paziente Sistema sanitario
Vasselli et al. (2005)	Italia	Erogazione dell'intervento secondo il quadro clinico del paziente e anche variabili non strettamente cliniche (qualità di vita) Erogazione dell'intervento nel momento giusto e secondo il regime organizzativo adeguato Deve tenere conto del contesto culturale e organizzativo del paziente e dell'operatore; deve tenere conto delle risorse	Paziente Operatore Sistema sanitario Società
Cinotti et al. (2005)	Italia	Intervento efficace per il singolo paziente Rispetto delle preferenze individuali e dei valori culturali e sociali Livello organizzativo di assistenza più idoneo	Paziente Sistema sanitario Società
Kelley, Hurst (2006)	OECD	Rilevanza rispetto ai bisogni clinici	Paziente
Ministero salute (2006)	Italia	Intervento correlato al bisogno del paziente o della collettività Erogazione dell'intervento secondo modi e tempi adeguati Bilancio positivo tra benefici, rischi e costi	Paziente Collettività Sistema sanitario
Tonelli (2008)	Italia	Compendia in sé efficacia ed efficienza ed è al tempo stesso considerata accettabile da chi riceve e da chi eroga le cure.	Paziente Sistema sanitario

Concetti correlati all'appropriatezza

L'appropriatezza è una dimensione della qualità dell'assistenza e, data la sua complessità e multidimensionalità, sono molti i termini ad essa correlati, quali efficacia, efficienza, equità, necessità clinica, variabilità geografica della pratica clinica.

Al concetto di **efficacia** è stata dedicata una notevole quantità di letteratura; in questa sede si riporta una definizione di “efficacia nella pratica clinica” (*effectiveness*), espressa come “grado in cui i miglioramenti di salute raggiungibili sono effettivamente raggiunti” e di “efficacia teorica” (*efficacy*), “l'abilità della scienza e della tecnologia sanitaria di produrre miglioramenti nella salute quando usate nelle condizioni più favorevoli”.

L'equità è la capacità del sistema di rispondere ai bisogni di gruppi e singole persone, sulla base dei possibili benefici, indipendentemente da fattori quali l'età, l'etnia, il genere, la disabilità, il livello socioeconomico e la scolarità.

L'efficienza considera il rapporto tra risorse impiegate e intervento erogato, considerando anche le modalità di applicazione delle risorse.

La **necessità clinica** rappresenta uno dei modi in cui l'appropriatezza è stata spesso descritta, soprattutto da parte dei soggetti finanziatori che, negli Stati Uniti e in Canada, hanno provveduto a definire i criteri per il rimborso dei servizi garantiti dal finanziamento pubblico. Una definizione molto pratica è proposta, a riguardo, da Charles: *i servizi clinicamente necessari sono “quelli di cui un paziente ha bisogno per sottrarsi a conseguenze negative sulla propria salute”*.

Analogamente, il concetto di appropriatezza è stato collegato a quello di **servizi necessari o critici**, intesi come quelli destinati a un “paziente in determinate condizioni tali da indurre il medico a effettuare un certo trattamento, poiché *il non farlo* costituirebbe una scelta deleteria per la salute del paziente”.

La **variabilità della pratica clinico-assistenziale**, geografica, organizzativa e professionale, è affrontata in una specifica sezione. Le differenze riscontrate nel ricorso ai servizi possono essere il risultato di diverse combinazioni di fattori: le caratteristiche dei pazienti (sociodemografiche, gravità clinica), le attitudini dei professionisti (secondo la scuola di riferimento, l'orientamento personale ma anche il metodo di pagamento vigente), il contesto istituzionale e organizzativo.

Modelli concettuali e componenti dell'appropriatezza

La revisione di letteratura di Sanmartin ha evidenziato due fondamentali *framework* concettuali inerenti l'appropriatezza. Il primo è stato introdotto da Lavis ed Anderson, a fine anni '90, con il duplice concetto di “*appropriateness of a service*”, o appropriatezza clinica, e

“*appropriateness of setting*”, solitamente intesa come determinante della appropriatezza organizzativa.

L’appropriatezza clinica identifica di fatto il livello di efficacia di una prestazione o procedura per un particolare paziente ed è determinata sulla base sia delle informazioni cliniche relative alle manifestazioni patologiche del paziente sia delle conclusioni diagnostiche che orientano verso quel preciso intervento sanitario, dal quale ci si attende un beneficio per il paziente. In pratica, gli autori hanno adottato una definizione molto simile a quella proposta dalla RAND, secondo la quale è appropriata una procedura da cui “*ci si attende più benefici che effetti negativi in un paziente che presenta un set preciso di manifestazioni cliniche*”. I costi, in questo caso, non sono considerati.

L’appropriatezza organizzativa identifica la situazione in cui l’intervento viene erogato in condizioni tali (ambito assistenziale, professionisti coinvolti) da “consumare” un’appropriata quantità di risorse (efficienza operativa). L’appropriatezza organizzativa prende quindi in considerazione la relazione costi-efficacia.

L’appropriatezza organizzativa viene presa in considerazione quando si intende valutare “*se il tipo di assistenza teoricamente richiesta dalle caratteristiche cliniche del paziente corrisponda al tipo di assistenza concretamente offerta*”.

Un secondo e più ampio modello concettuale, riferito in particolare alle prestazioni chirurgiche, è stato introdotto nello stesso periodo da Sharpe e Faden, secondo i quali l’appropriatezza non può essere considerata un costrutto dicotomico (appropriato vs inappropriato) ma un *continuum* di situazioni a diversa “intensità” di appropriatezza.

Gli stessi autori sottolineano anche la necessità di declinare il concetto generico di appropriatezza nel contesto più specifico del “sottoutilizzo” (*under-use*) e “sovrautilizzo” (*over-use*) delle prestazioni sanitarie. Se la definizione di appropriatezza proposta dalla RAND si fonda sull’equilibrio tra benefici e rischi, è vero che questo elemento può essere letto secondo almeno tre prospettive.

La prima chiave di lettura è offerta dal **livello di evidenza**. Secondo l’approccio originariamente proposto da Sackett, gli interventi sanitari sono **fortemente** raccomandabili (o **fortemente** controindicati) quando la superiorità dei benefici rispetto ai rischi e rispetto a terapie alternative è dimostrata da *trial* clinici randomizzati (RCT) conclusivi o replicati, considerati il più alto livello di evidenza raggiungibile.

Quando il rapporto costi-benefici è dimostrato da un RCT o da una revisione sistematica o altri studi di buona qualità, il livello di evidenza è ancora elevato, seppure inferiore al livello precedente e i corrispondenti interventi sanitari **sono raccomandati (o controindicati)**.

Infine, la sola disponibilità di *case report* o studi clinici non controllati rappresenta il più basso livello di evidenza e i corrispondenti interventi sono **potenzialmente** benefici (o dannosi) o l'equilibrio tra benefici e rischi è incerto.

Una seconda chiave di lettura è riconducibile alla **prospettiva dei pazienti**, che può contribuire alla determinazione dell'appropriatezza sia rispetto ad un "paziente medio" sia al "singolo paziente". Nel primo caso le informazioni sugli *outcome* clinici, su misure di qualità della vita o gradi di soddisfazione, stato di salute percepito o riferito contribuiscono alle stime di efficacia.. Nel secondo caso il focus è sull'accettabilità della cura, dimensione definita come "rispondenza ad aspirazioni, desideri e attese dei pazienti e delle loro famiglie"; un intervento può essere considerato accettabile o desiderabile quando sia "liberamente accettato da un paziente informato o da un suo rappresentante".

Infine, la terza chiave di lettura considera nella definizione di appropriatezza la **prospettiva della società**, che ha assunto un'importanza crescente in molti Paesi, come conseguenza dell'incremento dei costi dell'assistenza sanitaria.

Una società deve definire gli obiettivi di salute che si è disposti a perseguire; questo si correla non solo con le evidenze scientifiche, ma anche con i valori etici, condivisi dalla società stessa e condizionati da elementi di giudizio relativi alla distribuzione sociale o alla necessità di contenere i costi.

2. Variazioni della pratica clinica: concetti e metodi

Molti studi sull'appropriatezza si sono focalizzati sulle variazioni delle pratiche clinico-assistenziali, tra le diverse aree geografiche o organizzazioni sanitarie, nella gestione di pazienti che presentano condizioni patologiche e bisogni assistenziali simili. L'esistenza di considerevoli variazioni è evidenziata da numerosi studi, già a partire dagli anni '40 soprattutto per quanto riguarda le prestazioni chirurgiche.

Un rapporto pubblicato dal King's Fund Institute nel 1988, ha tracciato la storia delle variazioni nell'assistenza e nelle politiche sanitarie dagli anni '70, allo scopo di esplorare le differenze nell'allocazione delle risorse, le variazioni di *performance* tra organizzazioni socio-sanitarie e la variabilità geografica nell'accesso alle cure e nell'utilizzo dei servizi da parte di popolazioni residenti in determinate aree. Questo report ha evidenziato l'impatto di queste differenze sull'equità, l'efficienza e la sicurezza per i pazienti. Ad oggi rimangono sostanzialmente invariate l'entità delle differenze e le cause rilevate, quali il ruolo della domanda e dell'offerta e il processo decisionale dei professionisti.

Il problema della variabilità è tuttavia molto complesso e pone in primo luogo la questione di definire la differenza o la proporzione di variabilità desiderabile e quindi "positiva" rispetto a quella indesiderata e quindi "negativa".

Se le differenze indicano l'esistenza di cure inappropriate:

- quali sono le cure alle quali attribuire le differenze?
- dove la frequenza di cure inappropriate è più elevata c'è un problema di sovra-utilizzo (over-use) da parte di regioni e/o strutture, o di sovra utilizzo da parte dei professionisti?
- dove la frequenza di cure inappropriate è più bassa c'è un problema di sotto-utilizzo (under-use)?

Oppure la frequenza ottimale sta da qualche parte nel mezzo o persino al di là dell'uno o dell'altro estremo?"

modificato da Evans

Come già sottolineato nel capitolo precedente, le evidenze raccolte indicano che non esiste una relazione chiara e ripetibile tra l'appropriatezza e la variabilità geografica. Le differenze di

utilizzo di servizi o di prestazioni sanitarie sembrano essere il risultato della combinazione di diversi fattori riguardanti i pazienti, i professionisti, l'organizzazione e il contesto.

Siamo di fronte a un fallimento quando offriamo servizi a pazienti che non ne hanno bisogno o non li sceglierebbero, negando gli stessi servizi alle persone che ne beneficerebbero. La difficoltà consiste nel ridurre la quota di variabilità "cattiva", che riflette i limiti delle conoscenze professionali e i fallimenti nelle rispettive applicazioni e, nel contempo, preservare la variabilità "buona", che mantiene l'attenzione sulle esigenze dei singoli pazienti

Modificato da Mulley

Le possibili cause delle variazioni della pratica clinica

Quanti e quali sono i fattori che determinano il ricorso a una determinata prestazione sanitaria da parte di un paziente affetto da una particolare condizione clinica?

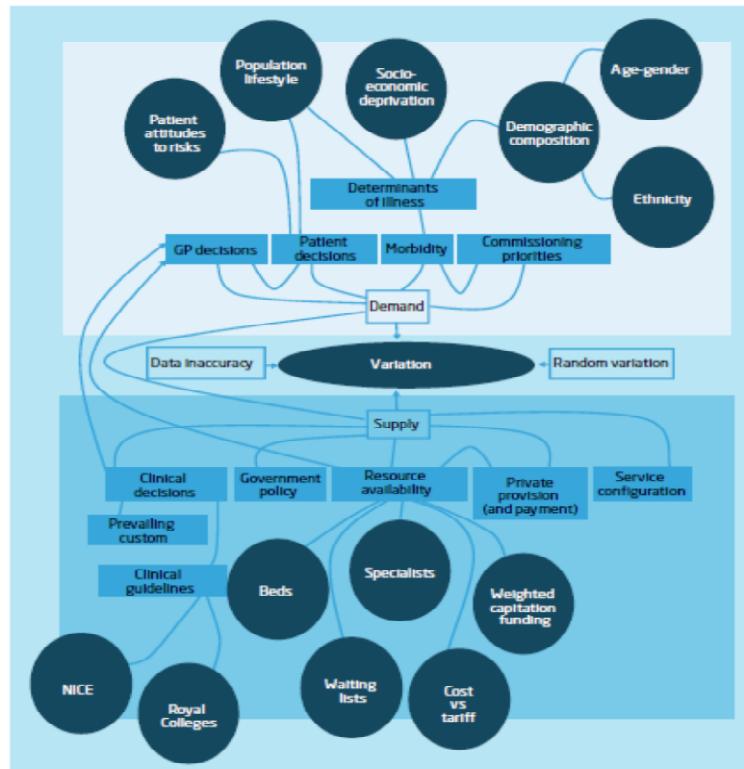
I fattori che comportano il ricorso a determinate prestazioni possono essere così schematizzati:

- **fattori riconducibili al paziente:** gravità delle condizioni cliniche, convinzioni e percezioni riguardo la malattia, fiducia nei servizi sanitari, inclinazione al rischio e a prendere decisioni, tipo di comunità e di famiglia a cui appartiene, influenza del medico curante, suggerimenti delle persone che lo circondano.
- **fattori inerenti il sistema dell'offerta:** disponibilità di servizi nella zona di residenza del paziente, distanza geografica tra casa e ospedale, modalità con cui sono organizzati gli ambulatori (inclusi gli orari di accesso), tempi d'attesa, numero di posti letto, sistema di finanziamento delle organizzazioni sanitarie, incentivi ai professionisti.
- **fattori riguardanti il professionista:** competenza, adesione a linee guida, esperienza.

Le interazioni tra queste diverse categorie di fattori sono infinite; la figura sottostante (Figura 1), seppure attinente alla situazione del *National Health Service* inglese, è comunque indicativa della complessità del problema.

Figura 1 mappa delle cause di variazioni nel ricorso ai servizi sanitari (Fonte: *Variations in Health Care. The King's Fund, 2011*)

Figure 1 Mapping causes of variation



Metodi e misure

L'identificazione dei fattori che possono spiegare le differenze di utilizzo di servizi o di prestazioni sanitarie, costituisce una sfida concreta, che implica la costruzione e la validazione di ipotesi e di modelli statistici.

Tra i metodi applicati per la valutazione delle differenze nelle pratiche clinico-assistenziali, si fa riferimento alla *small-area variation analysis*.

La ***small-area variation analysis (SAV)*** è una metodologia di analisi dei servizi sanitari, che ha avuto una particolare diffusione negli anni '80 e '90, anche se le prime ricerche sull'applicazione di tale metodologia alla pratica clinica vennero effettuate da Glover negli anni '30. Il suo scopo è quello di descrivere le differenze che si verificano tra aree geografiche nella distribuzione di un fenomeno, ad esempio il ricorso a un determinato servizio o prestazione sanitaria, la frequenza di una patologia o condizione clinica e di spiegarlo in termini di attitudini dei professionisti, di caratteristiche dei pazienti o di fattori ambientali.

Le differenze di pratica clinica, anziché secondo aree geografiche di residenza, possono essere analizzate in relazione a singoli ospedali o servizi; in questo caso è necessario affrontare specifici problemi di metodo inerenti, ad esempio, il *referral bias*, il *case-mix*.

Referral bias

è una delle possibili distorsioni da selezione delle osservazioni (selection bias), che si verificano, in generale, quando si confrontano gruppi differenti come se fossero comparabili; accade così di interpretare le differenze osservate (ad esempio in termini di outcome, come la mortalità o la guarigione) come effetto di variabili in studio (espresse, ad esempio, da misure della qualità dei servizi), mentre, al contrario, sono attribuibili alle differenze intrinseche ai gruppi di soggetti messi a confronto. In particolare, il referral bias si verifica quando le preferenze degli individui o le pratiche locali influiscono sulla selezione dei soggetti arruolati in un determinato studio o che accedono a un servizio o a una prestazione; Una situazione compatibile con il referral bias può anche verificarsi quando gruppi distinti di professionisti applicano criteri diversi per inviare i pazienti a un determinato accertamento diagnostico o percorso terapeutico. Ad esempio, se si intendesse studiare il ruolo della angio-TAC polmonare nella valutazione dei pazienti con difficoltà respiratoria, si dovrebbe tenere conto del fatto che i pazienti inviati dal Pronto Soccorso saranno diversi (quanto a durata e gravità dei sintomi) rispetto a quelli inviati dai Medici di Medicina Generale (Fonte: *Sica GT. Bias in research studies. Radiology 2006;238:780-789*).

Case-mix

è una caratteristica misurabile di gruppi o popolazioni di pazienti che consente di descrivere e quantificare la diversità della casistica trattata dai diversi ospedali o diversi reparti e che includono pazienti con condizioni cliniche di diversa gravità e quindi con probabilità diverse di insorgenza di esiti, complicanze e di eventi correlati con la qualità dell'assistenza. Non tener conto di questo fenomeno porta inevitabilmente a distorsioni nei risultati delle osservazioni e a una loro errata interpretazione, con decisioni ed interventi potenzialmente scorretti o addirittura dannosi. La variabilità del case mix viene trattata, negli studi di valutazione di performance o di qualità dell'assistenza, attraverso procedure di case mix adjustment (Fonte: *Parchman; Iezzoni LI et al*)

Le variazioni di ricorso alle cure possono essere misurate con diversi metodi, di differente complessità, dal semplice confronto degli estremi sino a metodi più complessi che analizzano l'intera distribuzione del fenomeno. Alcuni metodi di misura di uso più comune vengono qui di seguito definiti:

Range differenza tra il valore più alto e il valore più basso osservati. È una misura intuitiva ma fortemente influenzata dai valori estremi.

Range interquartile distanza tra il primo quartile (che corrisponde al 25° percentile) e il terzo quartile (il 75° percentile). Anche questa misura è abbastanza intuitiva e meno influenzata dagli estremi: tuttavia è basata su due sole osservazioni.

Quoziente tra gli estremi rapporto tra il valore più alto e il più basso. Facilmente comprensibile, ha gli stessi limiti delle due precedenti misure di *range*.

Deviazione standard (DS) misura il grado di dispersione dei dati rispetto alla media. Relativamente intuitiva e basata su tutte le osservazioni. Non fornisce alcuna informazione sul *pattern* di variazione.

Coefficiente di variazione (CV) è il rapporto tra la DS e la media. Ha il vantaggio di poter essere usata per confrontare le variazioni tra dati espressi con differenti unità di misura. Tuttavia, non consente l'aggiustamento per le variazioni intra-area (ad es. le variazioni casuali), è sensibile a piccoli cambiamenti se il valore della media è molto vicino allo 0 e non è intuitiva come le misure più semplici.

Componente sistematica di variazione (SCV) è fondata su un modello che considera il numero di eventi osservati in relazione al numero atteso secondo la distribuzione per età e genere della popolazione in studio. Consente l'aggiustamento per le variazioni intra-area.

Sulla base del documento prodotto dal *King's Fund Institute*, vengono messi a confronto, secondo il criterio della complessità, i principali metodi di misura (Figura 2).

Figura 2 Rappresentazione schematica delle principali misure usate nel descrivere le differenze di pratica clinico-assistenziale (Fonte: Variations in health Care, The King's Fund 2011)



1 Esempio

La tonsillectomia

Un discreto numero di studi volti a indagare le differenze di pratica clinica si è concentrato, storicamente, sull'intervento chirurgico di tonsillectomia, a partire dal lavoro di Glover (1938) il quale osservò che la frequenza della tonsillectomia variava fino a 20 volte tra i vari quartieri di Londra. Più o meno nello stesso periodo fu condotto a New York uno studio su mille bambini in età scolare, con risultati sorprendenti (*Box A*).

Box A

L'enigma della tonsillectomia – Uno studio a New York

Dove: la città di New York

Quando: anni '50

Chi: 1000 bambini in età scolare con tonsilliti acute e ricorrenti

Risultati: il 61% dei bambini (n=610) aveva già subito una tonsillectomia; questa osservazione dimostrava già di per sé la scarsa efficacia della procedura. Il risultato più sorprendente fu che il rimanente 39% dei bambini ancora non tonsillectomizzati (n=390) fu sottoposto alla visita da parte di un gruppo di medici scolastici, che raccomandarono l'intervento nel 45% dei casi (n=175). Il rimanente 55% dei bimbi (n=215) fu visitato da un altro gruppo di medici, che raccomandarono l'intervento nel 46% dei casi (n=99). Infine, i bambini ai quali, per due volte di seguito, non era stata consigliata la tonsillectomia (n=116), furono a loro volta inviati a un terzo gruppo di medici, che consigliarono l'intervento nel 44% dei casi (n=51). In definitiva, se tutti i bambini a cui era stata raccomandata la procedura l'avessero effettivamente subita, il numero totale di bambini operati, incluso quel 61% già tonsillectomizzato all'inizio dello studio, sarebbe stato pari a 935 su 1000.

Le evidenze più recenti sulla tonsillectomia, fornite da una revisione Cochrane, suggeriscono che l'intervento sia efficace nei bambini con tonsillite grave e ricorrente, per i bambini con malattia più lieve le indicazioni all'intervento sono modeste, ma, studi recenti dimostrano ancora differenze di frequenza fino a 7 volte nelle regioni inglesi.

In Italia, nel 2003 è stato elaborato il documento di indirizzo del PNLG "Appropriatezza e sicurezza degli interventi di tonsillectomia e/o adenoidectomia", cui ha fatto seguito, nel 2008 una specifica Linea guida. Il numero di interventi di chirurgia tonsillare, in quegli anni, è gradualmente diminuito, così come le variazioni geografiche dei tassi standardizzati di tonsillectomia. Gli interventi di tonsillectomia sono diminuiti da 59.916 nel 2002 a 51.983 nel 2003; a livello nazionale, il tasso per 10.000 di tonsillectomia è diminuito da 10,5 nel 2002 a 9,4 nel 2004.

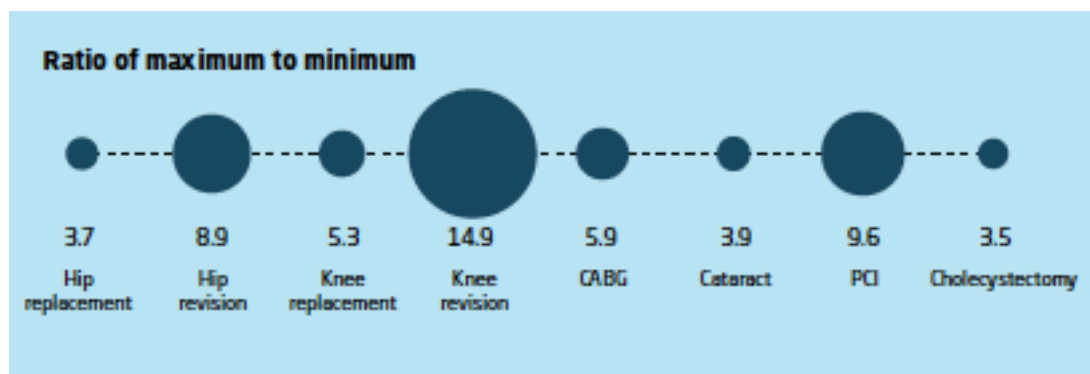
Contemporaneamente anche il tasso di adenoidectomia ha mostrato un progressivo decremento

2 Esempio

Variabilità nel ricorso a procedure chirurgiche con diversi livelli di efficacia

Osservazioni condotte in Inghilterra negli anni 2009-2010 su alcune procedure chirurgiche di comune esecuzione e di provata efficacia: **sostituzione o revisione protesica di anca e ginocchio, rimozione di cataratta, angioplastica coronarica, colecistectomia e by pass aorto-coronarico**, hanno dimostrato frequenze variabili tra un *Primary Care Trust* (PCT) e l'altro (il PCT può essere considerato sovrapponibile alla nostra Azienda Sanitaria Locale), espresse, ad esempio, in termini di “quoziente tra estremi”, cioè rapporto tra il tasso di ricorso massimo e il minimo (Figura 3).

Figura 3 Ricorso a diverse procedure invasive di comune utilizzo in Inghilterra: rapporti tra valore massimo e minimo del tasso dei corrispondenti ricoveri (Fonte: Variations in HealthCare 2011, cit)



Un'altra osservazione riguarda la persistenza nel tempo delle variazioni, così come emerge, dal confronto effettuato tra due periodi temporali successivi, 2005-2006 e 2009-2010, per tre procedure chirurgiche comuni: sostituzione protesica dell'anca, rimozione di cataratta e tonsillectomia.

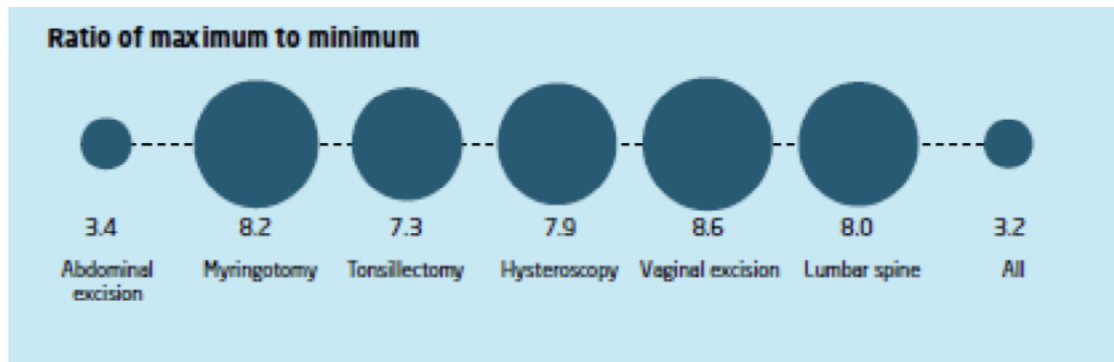
Inoltre, al di là delle variazioni annuali osservabili per i singoli PCT, nel tempo è stato possibile constatare come il ricorso alle rispettive procedure si sia mantenuto sostanzialmente invariato: i centri con tassi di ricorso più elevati si sono mantenuti tali e così pure è accaduto per i centri con tassi più bassi.

Secondo alcuni autori, le differenze nel ricorso agli interventi sanitari sono direttamente correlate al grado di consenso professionale riguardo il trattamento indicato per una determinata condizione clinica: minore è il consenso (o più ampia la distribuzione di opinioni dei clinici), maggiore è la probabilità di riscontrare differenze anche notevoli nell'utilizzo di una determinata procedura.

Il caso della tonsillectomia ne è un esempio; nonostante siano state recentemente definite indicazioni rigorose sull'appropriatezza dell'intervento, la disparità di convinzioni da parte dei clinici è tuttora considerevole.

L'osservazione del ricorso a un insieme di sei procedure, inclusa la tonsillectomia, classificate a bassa efficacia, ha mostrato differenze comprese tra 3 e 8 volte, misurate anche in questo caso con il rapporto tra tasso massimo e minimo (Figura 4).

Figura 4 Ricorso a sei procedure di scarsa efficacia in Inghilterra: rapporti tra valore massimo e minimo del tasso dei corrispondenti ricoveri (Fonte: *Variations in Health Care* 2011, cit.)



Le procedure considerate erano: isterectomia per via addominale, miringotomia, tonsillectomia, isteroscopia, isterectomia per via vaginale, puntura lombare.

Il ricorso ad alcune di queste procedure sembra associato alle condizioni socio-economiche, con un ricorso più elevato per le persone di condizioni più disagiate misurate da uno specifico **indice di deprivazione**.

Il concetto di **deprivazione** è frequentemente utilizzato da diverse discipline (demografia, epidemiologia, scienze economiche) e associato all'utilizzo di misure o indicatori in grado di quantificare lo stato di bisogno degli individui che abitano in una certa zona. Esiste al riguardo una lunga tradizione quasi esclusivamente anglosassone, sviluppatasi dagli anni '90 anche nel nostro Paese, con finalità prevalentemente epidemiologiche e quindi riferita generalmente a singole cause di mortalità o singole patologie. Diverse tipologie di indicatori di deprivazione sono stati introdotti per esaminare se e in che misura condizioni socioeconomiche diverse comportino differenze nella mortalità e nell'efficacia dei trattamenti.

La stessa tonsillectomia costituisce ancora una volta un esempio interessante, anche per la disponibilità di studi condotti in un periodo di tempo molto ampio, dalle prime osservazioni di Glover ai dati più recenti del *Department of Health*.

Un rapporto del 2006 sostiene che, nei tempi in cui la procedura era approvata dalla comunità medica, erano soprattutto i bambini di famiglie più benestanti a ricorrervi con maggior frequenza; successivamente, con la diffusione di maggiori evidenze circa l'appropriatezza della

procedura, venne acquisita una maggiore consapevolezza da parte dei clinici e la popolazione più istruita, meglio informata e con un grado più alto di autonomia nelle decisioni sulla salute, ridusse il ricorso all'intervento, mentre le persone socialmente più svantaggiate continuarono ad essere esposte a un rischio maggiore di *overuse* e di ricorso inappropriato alla procedura.

L'esempio della tonsillectomia mette in evidenza la connessione esistente tra le differenze di utilizzo degli interventi sanitari e il rischio di ingiustificate disuguaglianze nell'accesso alle cure, determinate non tanto dalle condizioni cliniche dei pazienti o dalle loro preferenze, quanto dalle condizioni socio-economiche.

Un altro esempio della variabilità di accesso ad una procedura è rappresentata dalla malattia coronarica che può essere trattata con procedura chirurgica (*by pass* aorto-coronarico-BPAC), con procedura interventistica (angioplastica coronarica-PCI), ma anche con terapia medica. La scelta dell'una o dell'altra opzione si basa, nel singolo paziente, su fattori come la gravità e l'estensione della malattia, l'età, la comorbidità e le condizioni generali di salute.

La letteratura internazionale ha tuttavia documentato, in diversi ospedali o aree geografiche, differenze nel ricorso a BPAC o PCI fino a 5, 10 volte.

È possibile identificare le ragioni delle differenze osservate? La letteratura suggerisce due spiegazioni fondamentali:

- a) la soglia per la scelta del trattamento è variabile, soprattutto negli stadi precoci della malattia; esiste inoltre una certa dose di incertezza riguardo a: possibile beneficio, tipo di paziente, stadio della malattia
- b) il processo decisionale è imperfetto: i pazienti spesso delegano la scelta al medico, oppure il medico non è in grado di inquadrare correttamente il paziente e comprenderne le preferenze. È stato dimostrato come pazienti correttamente informati riguardo le alternative disponibili, spesso tendono a scegliere in modo diverso dai rispettivi medici e sono meno inclini a scegliere la chirurgia rispetto a un gruppo di controllo in cui l'informazione sia stata fornita in modo tradizionale e meno strutturato.

Differenze e disuguaglianze: il caso della protesi d'anca

Alcuni studi hanno dimostrato l'esistenza di un'associazione negativa tra ricoveri per intervento di sostituzione protesica dell'anca e misure di bisogno socio-sanitario, come il *Multiple Deprivation Index* (MDI): in altre parole, i pazienti più poveri hanno minore probabilità di essere sottoposti all'intervento.

Il **MDI** viene usato in molte analisi prodotte nel Regno Unito e diffuse dai *Department for Communities and Local Government*. L'indice, la cui definizione concettuale e metodologia di calcolo sono periodicamente aggiornate, tiene in considerazione diverse componenti di stato socio-economico. La versione 2007 utilizza sette ambiti: 1) reddito 2) occupazione 3) condizioni di salute e disabilità 4) istruzione, competenze e *training* 5) barriere relative alla condizione abitativa e all'uso dei servizi 6) criminalità 7) ambiente di vita. (Fonte: *The English Indices of Deprivation 2007*)

Agli interventi di protesi d'anca è legato anche un altro tipo di osservazione: tra le persone in procinto di essere sottoposte all'intervento, i pazienti in condizioni di svantaggio socio-economico riferiscono uno stato di salute peggiore, così come emerge da due *Patient Reported Outcomes Measurements* (PROMs), rilevazioni condotte dal *National Health Service* inglese a partire dal 2009, in particolare: l'*Oxford Hip Score* e l'indice generico di qualità della vita EQ-5D.

Da rilevazioni effettuate nel 2009-2010 emerge come entrambe queste misure si attestino su valori più bassi che denotano quindi uno stato di salute peggiore per i pazienti con indice di deprivazione più elevato. Inoltre, questi provengono più frequentemente da aree geografiche con minor tasso di sostituzione protesica dell'anca.

La relazione tra deprivazione e intervento di protesi d'anca agisce quindi su due livelli: i pazienti "più poveri" hanno meno probabilità di essere sottoposti all'intervento pur avendone necessità e, nel momento in cui vi accedono, dichiarano di "stare peggio" rispetto ai pazienti più benestanti. Quest'ultima osservazione potrebbe trovare diverse spiegazioni: grado più elevato di comorbidità, ritardo nell'accesso all'intervento e quindi condizioni più gravi, o altri fattori riconducibili alla condizione socio-economica.

Esistono quindi chiare evidenze della disuguaglianza di accesso alla sostituzione protesica dell'anca in Inghilterra, i cui determinanti principali risultano essere l'età, il genere, la deprivazione, la residenza in aree rurali e l'etnia.

Osservazioni analoghe sono state fatte per altre procedure, come la rivascolarizzazione coronarica e l'intervento di cataratta.

L'espressione *Patient Reported Outcome (PRO)* identifica una terminologia “ombrello”, che identifica un approccio concettuale e una metodologia specificamente orientati a cogliere il punto di vista dei pazienti, piuttosto che adottare la tradizionale prospettiva clinica o professionale sugli *outcome*. I PRO possono essere raccolti attraverso questionari auto-compilati o attraverso interviste. Al concetto di PRO è quindi strettamente connesso quello di PROMs (*Patient Reported Outcomes Measurements*), che identifica una vasta gamma di strumenti di rilevazione e misurazione.

Strumenti di rilevazione e misurazione

Generali se rivolti alla valutazione della qualità di vita o alla valutazione soggettiva dello stato di salute (*SF-36 Health Survey, SF-12 Health Survey, Sickness Impact Profile, Nottingham Health Profile, Health Utilities Index, Quality of Well-Being Scale, EuroQol-EQ-5D, Consumer Assessment of Healthcare Providers and Systems-CAHPS*)

Specifici per patologia (*Adult Asthma Quality of Life Questionnaire-AQLQ* per i pazienti adulti con asma, *Kidney Disease Quality of Life Instrument* per i pazienti con malattie renali, *Oxford Hip Score* per i pazienti sottoposti a intervento di protesi d'anca)

Il setting assistenziale: il caso della *day surgery*

Uno dei cambiamenti più importanti nell'organizzazione dei servizi sanitari degli ultimi 20 anni è stato lo sviluppo della *day surgery* (DS), cioè la possibilità di effettuare un numero elevato di interventi chirurgici in regime diurno, evitando il tradizionale ricovero ospedaliero, con notevoli vantaggi per i pazienti (maggiore tempestività, minor rischio di infezioni, minor rischio di rinvii e ritardi, più sereno e rapido rientro a casa) e per l'efficienza del sistema (notevole riduzione dei giorni di degenza).

Nel *National Health Service* inglese la percentuale di interventi effettuati in DS si avvicina al 70%, tuttavia sono ancora elevate le differenze tra le diverse aree geografiche, come dimostrano i rapporti della *Audit Commission* a partire dagli anni '90, con particolare riferimento a un gruppo di 25 procedure per le quali il regime di DS risulta fortemente *cost-effective*.

Nel 2009-2010 in Inghilterra sono state eseguite circa 1 milione di procedure chirurgiche potenzialmente effettuabili in DS, ma oltre il 20% (220.000 interventi) non ha avuto luogo in questo

setting; la proporzione di interventi in DS variava tra 67% e l'87% con differenze significative tra le varie strutture.

Per alcuni interventi specifici, come la chirurgia della cataratta, le variazioni sono state minime (88%-100% di interventi in DS), mentre in altri casi quali la tonsillectomia e la colecistectomia laparoscopica è stato osservato un considerevole sotto-utilizzo. Per alcune procedure è stata inoltre dimostrata un'associazione negativa con indicatori di svantaggio socio-economico, in particolare, nei territori con popolazioni "più povere", gli interventi di circoncisione, resezione transuretrale di tumori vescicali, decompressione del tunnel carpale, artroscopia, chirurgia della cataratta, sono stati eseguiti con maggior frequenza in regime di degenza ordinaria.

Secondo la *Audit Commission* questo riscontro non è determinato dalle caratteristiche dei pazienti, ma dalle caratteristiche dell'offerta, come la disponibilità di infrastrutture o l'attitudine clinica dei professionisti.

3. La misura dell'appropriatezza

La misura dell'appropriatezza continua ad essere una sfida per i clinici, per i ricercatori, per i manager e per i *policymaker*.

I risultati di processi formali e condivisi di valutazione dell'appropriatezza possono essere utilizzati per diversi obiettivi: la valutazione di qualità dell'assistenza, l'elaborazione di raccomandazioni per la pratica clinica, la predisposizione di strumenti di supporto pratico al lavoro dei clinici, l'identificazione di aree meritevoli di ulteriore ricerca.

Tali processi richiedono adeguati strumenti e metodi di misurazione/monitoraggio, individuati sulla base di criteri espliciti, condivisi, scientificamente fondati, ai quali poter ricondurre valutazioni, individuali o collettive, effettuate da esperti.

La valutazione di appropriatezza dei ricoveri ospedalieri attraverso l'uso delle cartelle cliniche e dei database amministrativi : PRUO, DRG ad elevato rischio di inappropriatezza

PRUO

I protocolli per la valutazione di appropriatezza del ricovero del paziente adulto e del bambino, hanno avuto in Italia una certa diffusione soprattutto negli anni '90 e nei primi anni 2000, e si sono concretizzati nei cosiddetti studi PRUO (*Protocolli di Revisione d'Uso dell'Ospedale*); in seguito l'uso di tali strumenti si è molto ridimensionato ed oggi riveste più un interesse storico-documentale che pratico.

La matrice di riferimento per l'elaborazione del **Protocollo di Revisione d'Uso dell'Ospedale (PRUO)** è stato il *Pediatric Appropriateness Evaluation Protocol (AEP)*, uno strumento per la raccolta guidata di informazioni dalla cartella clinica, mirate a valutare, attraverso set di criteri e standard ben definiti.

Sulla base dell'*AEP* un gruppo di ricercatori del Laboratorio di Epidemiologia clinica dell'Istituto Mario Negri di Milano, dell'Ospedale San Gerardo di Monza e dell'Ospedale Niguarda di Milano, ha prodotto i PRUO, elaborati tenendo conto della gravità della malattia che ha condotto al ricovero, dell'intensità dell'assistenza fornita, delle prestazioni effettuate in ciascun giorno di degenza, dell'assistenza infermieristica, delle condizioni dei pazienti. E' stata quindi elaborata una griglia di riferimento in base alla quale si definiva l'inappropriatezza dell'ammissione in ospedale e/o delle giornate di degenza.

In diversi studi condotti negli ospedali americani le percentuali di ammissioni inappropriate sono risultate comprese tra 8% e 25%.

In alcuni studi condotti in Italia a fine anni '80, che si sono avvalsi dell'AEP, le ammissioni in ospedale giudicate non appropriate variavano dal 26% (in uno studio condotto all'Ospedale San Gerardo di Monza) al 28% (Ospedale di Aosta, 1989).

Il PRUO conserva una struttura sostanzialmente simile all'AEP, basandosi su diciotto criteri espliciti, indipendenti dalla diagnosi, relativi alla condizione del paziente e alla intensità/complessità del trattamento. Il giudizio di congruità del ricovero è formulato qualora sia soddisfatto almeno un criterio.

Nella conduzione di studi basati sul PRUO è necessario considerare con attenzione alcune fasi, in primo luogo il campionamento. Il campione può essere scelto in modo casuale o sulla base di valutazioni specifiche e contingenti (campione di "convenienza").

I ricoveri possono essere selezionati in funzione della giornata di degenza; in questo caso si risale a ritroso al giorno dell'ingresso in ospedale per valutare sulla base della documentazione clinica l'appropriatezza dell'ammissione.

In alternativa, i ricoveri possono essere selezionati sulla base delle nuove ammissioni, seguite poi nel corso della degenza.

Gli altri aspetti che devono essere valutati nella pianificazione, conduzione e interpretazione di studi basati sul PRUO sono:

- il **setting** nel quale lo studio viene condotto (tipologia e caratteristiche dell'ospedale, grado di specializzazione, presenza di specifici reparti e servizi diagnostici, area geografica etc.)
- il **case-mix** dei pazienti
- il **timing** della rilevazione, poiché la raccolta di informazioni può essere concorrente o retrospettiva.

DRG ad elevato rischio di in appropriatezza

Il DPCM del 29/11/2001 (“Definizione dei livelli essenziali di assistenza”) riporta, in allegato 2.C, l’elenco dei 43 DRG ad elevato rischio di in appropriatezza. Questa prima lista è stata riesaminata nell’ambito del Progetto Mattoni “Misura dell’appropriatezza-Identificazione di DRG ad elevato rischio di in appropriatezza” ed è stata formulata la seguente definizione di DRG ad elevato rischio di in appropriatezza: *“DRG che presentano un valore “anomalo” rispetto al corrispondente valore nazionale della percentuale di ricoveri effettuati in regime diurno e/o in regime ordinario con degenza inferiore a 2 giorni “.*

La metodologia messa a punto nel corso del progetto ha consentito di individuare nuovi DRG, da affiancare a livello nazionale alla lista dei 43 DRG previsti dal Dpcm 29/11/2001, andando a creare una nuova lista nazionale di 55 DRG potenzialmente inappropriati dal punto di vista organizzativo.

Il metodo consente una stima di in appropriatezza dei ricoveri, basata su dati amministrativi di facile reperibilità e la cui raccolta risulta relativamente economica forniti dalla Scheda di dimissione ospedaliera (SDO). Tuttavia, essendo questa un documento di natura amministrativa, soffre di problemi di qualità formale e sostanziale ampiamente studiati e misurati. In altri termini, la rappresentazione/codifica delle condizioni cliniche e dell’assistenza erogata durante il ricovero possono essere condizionate dal loro utilizzo per il pagamento della prestazione e non vi è garanzia che le informazioni riportate siano complete, accurate e non distorte. Inoltre, nella scheda di dimissione manca qualunque riferimento al contesto nel quale vengono prese le decisioni cliniche, ad esempio relativamente alle condizioni sociali del paziente che, in alcuni casi, sono in grado di giustificare il prolungamento della degenza e/o il ricorso a livelli più intensivi di assistenza.

D’altra parte, la valutazione di appropriatezza a partire da dati amministrativi è interessante per l’immediata disponibilità della informazioni, relative al processo e all’esito dell’assistenza, raccolte attraverso il sistema informativo ospedaliero in modo relativamente uniforme anche per quanto riguarda i problemi di qualità della codifica. L’utilizzo dei dati amministrativi non si propone di valutare l’appropriatezza del singolo ricovero, ma piuttosto di effettuare stime di frequenza e di in appropriatezza su insiemi anche piuttosto ampi di ricoveri, monitorandone l’andamento nel tempo.

Lo studio dei ricoveri ospedalieri inappropriati, effettuati in regime ordinario per prestazioni potenzialmente eseguibili a livello ambulatoriale o in ricovero diurno, è stato condotto in diverse esperienze del Nord America, utilizzando le diagnosi principali di dimissione, come quelle proposte dal *Medicare core editor*, o i DRG o sistemi più complessi, come MEDISGRPS, che, nati nella realtà americana, presentano non pochi problemi di trasferibilità in Paesi con una diversa

organizzazione dei servizi socio-sanitari e diverse attitudini alla codifica amministrativa dei dati sanitari. La loro diffusione in Italia è stata scarsa o nulla. In pochi casi, come è avvenuto per il sistema APR-DRG (*All Patient Refined DRG*) e per il “*Disease Staging*”, l’utilizzo è stato circoscritto ad alcune esperienze.

Anche in Italia comunque, con l’entrata a regime del sistema informativo ospedaliero, si sono susseguiti, da un certo momento in avanti, i tentativi di ottenere stime di appropriatezza dei ricoveri attraverso i dati riportati sulla scheda di dimissione.

L’uso degli APR-DRG consente di prendere in considerazione, oltre alla complessità assistenziale, anche la gravità clinica, superando dunque uno dei principali problemi delle stime di appropriatezza basate sui dati amministrativi, cioè la difficoltà di identificare gli episodi di ricovero di gravità clinica tale da giustificare la degenza ordinaria, anche se attribuibili a DRG a bassa complessità assistenziale. D’altra parte, un importante limite di questo metodo, che assegna un ruolo determinante alle diagnosi secondarie, risiede nella sua forte dipendenza dalla completezza e accuratezza di compilazione della scheda di dimissione ospedaliera.

APR-DRG

La prima versione degli APR-DRG (12.0) risale al 1995. A differenza dei DRG, il sistema degli APR-DRG si basa, per la definizione dei gruppi, su criteri relativi alla gravità clinica dei pazienti, oltre che alla loro complessità assistenziale. Consta di 341 DRG non differenziati in base all’età o alla presenza di patologie concomitanti e/o complicanze; e di 43 categorie finali per pazienti in età neonatale (0-28 giorni al ricovero), con l’utilizzo del peso alla nascita quale variabile classificatoria. Ogni APR-DRG è poi articolato in quattro sottoclassi di severità e quattro sottoclassi di rischio di morte (minori o assenti, moderate, maggiori, estreme), assegnate sulla base di un complesso algoritmo che tiene conto principalmente delle diagnosi secondarie presenti. Queste sottoclassi descrivono l’impatto della gravità clinica sull’intensità di assorbimento delle risorse nel processo assistenziale e sulla mortalità, rispettivamente (Fonte: *Baglio G. et al.*)

Il metodo RAND

Il metodo RAND/UCLA per la valutazione di appropriatezza (*RAND/UCLA Appropriateness Method*, RAM), è stato sviluppato alla metà degli anni ‘80, nell’ambito di uno studio più ampio, il *RAND Corporation/University of California Los Angeles - UCLA Health Services Utilisation Study*, come strumento di misura del sovrautilizzo e del sottoutilizzo di procedure mediche e chirurgiche

Il RAM è una variazione del “metodo Delphi; sebbene sia spesso considerato un “metodo per la costruzione del consenso”, in realtà non appartiene a questa categoria, perché il suo obiettivo è quello di identificare le situazioni in cui gli esperti si trovano d'accordo o in disaccordo e lascia ai componenti l'opportunità di discutere le loro valutazioni tra un ciclo e l'altro.

Metodo Delphi

Sviluppato presso la RAND negli anni '50 e applicato a temi politico-militari, tecnologici ed economici, come strumento per prevedere il futuro, è stato utilizzato in una vasta gamma di situazioni pertinenti alla salute e alla medicina.

Il Delphi si articola generalmente in cicli ripetuti, nei quali un questionario viene inviato a un gruppo di esperti che rispondono in modo anonimo. I risultati di questa prima survey vengono quindi sintetizzati in tabelle e restituiti al gruppo; a ciascun esperto si chiede quindi di rispondere nuovamente al questionario. Il processo viene reiterato fino a quando non si verifichi una convergenza di opinioni sul tema affrontato o fino a quando nessun sostanziale cambiamento sia osservabile nelle risposte degli esperti.

Il RAM è talvolta erroneamente considerato come un esempio della Tecnica del Gruppo Nominale (Nominal Group Technique, NGT).

La NGT consiste in un processo molto strutturato in cui ai partecipanti viene chiesto di scrivere tutte le idee in merito a un determinato oggetto. Il facilitatore chiede a ciascuna persona di descrivere brevemente l'idea a suo giudizio più rilevante; dopo la discussione di ogni argomento, a ciascuno dei partecipanti viene chiesto di ordinare su una scala numerica, le proprie valutazioni relative ai temi trattati. Infine, i risultati vengono aggregati e analizzati con diverse tecniche matematiche.

Il RAM, a differenza del metodo NGT, inizia con una lista molto strutturata di indicazioni cliniche e la discussione è strettamente collegata alle misurazioni dell'appropriatezza.

Un terzo gruppo di metodi per la valutazione collettiva è la Consensus Development Conference, o Conferenza di Consenso (CC). Il metodo è stato sviluppato con lo scopo di sostenere il mandato dei National Institutes of Health (NIH) americani, cui è affidata la valutazione e la disseminazione di informazioni relative alle tecnologie sanitarie e alla ricerca biomedica. In pratica, rispetto al tema oggetto di studio la CC raccoglie una gamma eterogenea di stakeholder (medici, ricercatori, cittadini, società scientifiche ecc.) ai quali viene chiesto di sviluppare, in risposta a

quesiti pre-definiti sul tema affrontato, dichiarazioni che esprimano un accettabile livello di consenso. Anche la CC prevede una revisione della letteratura, la sintesi delle conoscenze disponibili, presentazioni di esperti e soggetti impegnati in azioni di advocacy e una discussione plenaria. Una CC solitamente si prolunga per 2 o più giorni e non si conclude fino a quando i partecipanti non abbiano trovato un accordo su un documento scritto. Molti Paesi europei hanno sviluppato proprie versioni modificate del metodo originario.

Secondo il RAM, il concetto di appropriatezza viene ricondotto al peso relativo di rischi e benefici **degli interventi medici o chirurgici**. Viene definita “appropriata” una procedura in cui “*i benefici attesi per la salute (ad es. l’aumento dell’aspettativa di vita, il sollievo dal dolore, la riduzione dell’ansia, il miglioramento della capacità funzionale) superi le previste conseguenze negative (es. mortalità, morbosità, ansia, dolore, tempo sottratto al lavoro) con un margine sufficientemente ampio tale da giustificare la procedura, senza tener conto dei costi*”.

Il rationale che sostiene il metodo consiste nel prendere atto della insufficienza dei *trial clinici randomizzati*, nel fornire evidenze tanto dettagliate da consentirne l’applicazione alla vastissima gamma di situazioni clinico assistenziali.

Anche in assenza di robuste evidenze scientifiche circa i benefici di determinate procedure, i clinici si trovano tuttavia ogni giorno a dover prendere decisioni in merito all’opportunità del loro utilizzo. Si è sentita quindi la necessità di un metodo che potesse combinare la migliore evidenza scientifica disponibile con il giudizio clinico collettivo degli esperti, per rendere possibile una presa di posizione esplicita riguardo l’appropriatezza di esecuzione di una procedura in situazioni specifiche. Molte procedure sono state oggetto di studi sull’appropriatezza negli Stati Uniti, ad esempio l’angiografia coronarica, il *bypass* aorto-coronarico, l’angioplastica coronarica, l’endoarterectomia carotidea, la chirurgia dell’aneurisma aortico addominale, l’endoscopia diagnostica del tratto gastrointestinale superiore, la chirurgia della cataratta, la colonscopia, la colecistectomia, l’isterectomia, la timpanostomia e la manipolazione del rachide per il mal di schiena.

Dai tempi delle prime applicazioni, il metodo si è diffuso in molti Paesi compresa l’Italia ed è stato applicato a ulteriori problemi clinici, come l’iperplasia prostatica benigna, la laminectomia, il cancro della mammella e la sostituzione protesica dell’anca.

Un risultato importante dei primi studi condotti dalla RAND è stata l’osservazione di una mancata relazione tra il volume delle procedure effettuate in diverse aree geografiche e i livelli di appropriatezza. Numerose ricerche avevano documentato variazioni importanti nell’uso di determinate procedure nelle varie aree geografiche; i ricercatori della RAND avanzarono l’ipotesi che le differenze potessero essere in parte spiegate da un maggior livello di inappropriatezza nelle

aree con maggiori volumi, ma questa ipotesi non venne confermata. La proporzione di procedure inappropriate non era associata al numero di procedure eseguite, suggerendo che sia il sottoutilizzo, sia il sovrautilizzo possono verificarsi contemporaneamente nella stessa area.

Di seguito si riportano le definizioni di **sovra utilizzo e sottoutilizzo delle cure mediche**

Sovrautilizzo (*overuse*)

Si verifica quando un farmaco o un trattamento vengono somministrati senza giustificazione medica. Esempio tipico di sovrautilizzo è il trattamento con antibiotici in condizioni che non lo richiedono. Un altro esempio è il ricorso alla Risonanza Magnetica (RM) o ad altro esame di diagnostica per immagini in casi in cui il risultato di tali indagini non modifichi il trattamento o non migliori i risultati di salute del paziente, comportando un potenziale danno al paziente e un costo non necessario.

L'*overuse* di cure è particolarmente evidente nel trattamento di malattie croniche, ad es. l'ospedalizzazione di pazienti con diabete, piuttosto che il loro trattamento a casa o in ambulatori territoriali. La causa di questo fenomeno è una dipendenza eccessiva dalle strutture di cura per acuti e una mancanza di infrastrutture necessarie a sostenere l'assistenza ai malati cronici in altri tipi di *setting*.

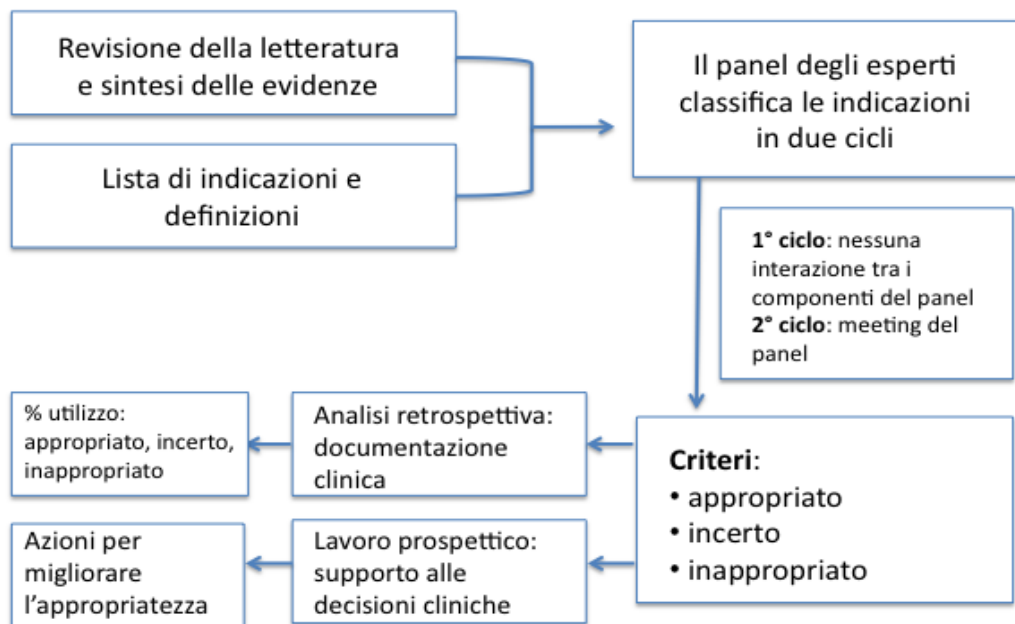
Sottoutilizzo (*underuse*)

Si verifica quando non vengono somministrate le cure necessarie o non vengono messe in atto pratiche di provata efficacia. Esempi sono la terapia antitrombotica nella chirurgia addominale ed ortopedica, l'esame della retina nei pazienti diabetici, lo screening per il cancro della mammella o la vaccinazione antinfluenzale nelle persone di età uguale o superiore a 65 anni.

Una delle cause di *underuse* è la discontinuità nelle cure quando più medici sono coinvolti nella cura dello stesso paziente e la mancanza di procedure e sistemi in grado di facilitare l'accesso ai servizi necessari.

Il rationale e le modalità di utilizzo del metodo RAND sono descritte nel manuale "Appropriateness Method User's Manual" (Fitch K, Bernstein FJ, Aguilar MD, Burnand B, Lacalle JR, Lazaro P), pubblicato in versione aggiornata nel 2001 e articolato in modo da costituire una guida pratica per gli utilizzatori del metodo. Le fasi fondamentali di applicazione del metodo RAND sono rappresentate in Figura 5

Figura 5 Il metodo RAND/UCLA per la valutazione dell'appropriatezza: schema (Fonte: The RAND/UCLA Appropriateness Method User's Manual 2000, modif.)



Le fasi del processo di valutazione dell'appropriatezza secondo il metodo RAND/UCLA

Identificazione e selezione del problema clinico: devono essere identificati problemi di tipo clinico, mentre gli aspetti di carattere economico e/o di organizzazione sanitaria vanno esclusi. Nella selezione del problema clinico giocano un ruolo fondamentale alcuni fattori che contribuiscono alla formulazione di un giudizio di rilevanza: elevata frequenza di utilizzo della procedura, costi elevati, rischio elevato di complicanze (mortalità, morbosità), importanti variazioni geografiche, procedure il cui uso è controverso, buone probabilità terapeutiche o diagnostiche e la qualità dell'evidenza scientifica disponibile.

Revisione della letteratura scientifica: la revisione delle letteratura scientifica finalizzata alla valutazione di appropriatezza può essere effettuata sulla base delle tradizionali revisioni Cochrane o essere basata su criteri di inclusione meno stringenti e comprendere, laddove siano assenti *trial* controllati, evidenze provenienti da studi di coorte o da *case series*. Nel contesto del metodo RAND, la revisione delle evidenze scientifiche disponibili, oltre a fornire il supporto necessario per la formulazione dei giudizi degli esperti, costituisce anche il punto di partenza per la definizione degli scenari clinici che verranno sottoposti alla valutazione di appropriatezza. Una volta prodotta, la revisione della letteratura, riassunta in tavole e tabelle di facile lettura, deve essere consegnata agli esperti del panel e ad essa si farà riferimento in tutti i casi di disaccordo tra i vari componenti.

Costruzione degli scenari clinici o “indicazioni”: la lista delle “indicazioni” rappresenta il documento di lavoro principale per il panel. Viene elaborato, in forma di matrice, un elenco di specifici “scenari” o “indicazioni, per classificare, in base a sintomi, storia clinica, risultati dei test diagnostici, i pazienti che potrebbero essere candidati all’intervento in questione.

Nella stesura della lista è importante assicurarsi che le indicazioni siano: a) esaustive, così da includere un’ampia popolazione di pazienti potenzialmente candidati alla procedura in esame; b) mutualmente esclusive, cosicché nessun paziente possa essere ricondotto a due scenari diversi; c) omogenee, così da consentire una uguale applicabilità della valutazione di appropriatezza a tutti i pazienti classificati in uno stesso scenario; d) gestibili, così da consentire ai componenti del panel di valutare tutte le indicazioni in un ragionevole intervallo di tempo.

Il panel di esperti: la costituzione del panel è una fase cruciale e i criteri della loro selezione possono variare molto in ragione del contesto in cui il processo di valutazione si svolge. Negli studi condotti dal gruppo RAND negli USA, gli esperti, in numero generalmente compreso tra 9 e 11, sono stati spesso identificati attraverso il coinvolgimento delle Società Scientifiche. Nella selezione dei partecipanti è necessario seguire alcuni criteri:

- Multidisciplinarietà. Per ottenere una valutazione equilibrata dell’appropriatezza nell’uso di una procedura, il giudizio non deve essere affidato solo a coloro che di tale procedura sono gli esecutori diretti, ma devono essere coinvolte anche altre figure mediche che sono parte in causa, anche se indirettamente, nel problema clinico considerato.
- Credibilità. Gli esperti selezionati devono godere di indiscusso prestigio professionale ed avere alcune caratteristiche specifiche, quali una buona familiarità con la letteratura scientifica.
- Disponibilità. Per la riuscita del lavoro, che si protrae nelle sue diverse fasi per alcuni mesi, è essenziale che vi sia disponibilità e motivazione da parte dei partecipanti ad eseguire i compiti loro affidati. È altrettanto importante che i clinici coinvolti siano disponibili a lavorare in gruppo.

Prima valutazione di appropriatezza. Il primo passo della vera e propria attribuzione del giudizio di appropriatezza consiste nella valutazione individuale, da parte di ogni membro del gruppo di esperti, per ciascuno degli scenari proposti. Il giudizio viene espresso su una scala da 1 a 9, dove 1 = sicuramente inappropriato, 9 = sicuramente appropriato, 5 = incerto. Valori intermedi corrispondono a differenti modulazioni rispettivamente del giudizio di inappropriato (2-3), di incertezza (4-6), e di appropriatezza (7-8). Nel valutare ogni singola indicazione, ogni membro del panel deve fare riferimento sia alla propria esperienza e giudizio clinico, sia alle evidenze scientifiche disponibili.

Seconda valutazione di appropriatezza. Si svolge nel corso di una riunione collegiale (*meeting*), coordinata da un moderatore esperto. Nel corso dell'incontro vengono analizzati i risultati della prima valutazione, con l'obiettivo di discutere le motivazioni delle diversità di opinioni eventualmente emerse. Ogni membro del panel riceve un report che illustra i risultati della prima valutazione, scenario per scenario. Il report è strutturato in modo tale da consentire a ciascun partecipante di confrontare il proprio giudizio con quello espresso dagli altri componenti del panel, mantenendo la reciproca segretezza, e con il giudizio medio. La discussione si concentra sulle aree di disaccordo eventualmente emerse. È importante ricordare che il metodo non implica la ricerca di un consenso ad ogni costo; al contrario, è inevitabile che vi siano scenari, cioè indicazioni cliniche, sulle quali possono persistere divergenze di opinione, che riflettono una reale incertezza circa l'impatto della procedura, verosimilmente a causa della incompletezza delle evidenze disponibili. Al termine del *meeting*, i componenti del panel effettuano una ulteriore valutazione degli scenari clinici.

Valutazione di necessità. In questo caso è necessario un terzo e, talvolta, anche un quarto ciclo di valutazioni, che avviene generalmente via *mail*. I componenti del panel valutano solo il subset di indicazioni già giudicate appropriate nel secondo *round*, cioè con punteggio mediano di 7, 8 o 9 senza disaccordo. La necessità è quindi un criterio più stringente dell'appropriatezza; una procedura viene considerata necessaria quando i benefici superano i rischi con un margine talmente ampio che il medico deve offrirla o prescriverla. Si potrebbe dire che *“la cura in questione appare così chiaramente la cosa giusta da fare che il medico considererebbe non etico il non prescriverla e, nel caso non la prescrivere, potrebbe prevedere una denuncia da parte del paziente”* (Brook, 1999). La valutazione di necessità avviene con la stessa scala di punteggi, dove 1=non necessaria (benché appropriata) e 9=necessaria (non raccomandare la procedura configura un giudizio clinico improprio).

L'analisi dei risultati. Il punto cruciale della classificazione degli scenari, in tutti i cicli di valutazione, risiede nel livello di “accordo” e “disaccordo”. In particolare, vengono giudicati “in accordo” gli scenari in cui, dopo aver escluso dall'analisi i due giudizi estremi, i rimanenti giudizi cadono in una qualunque delle tre “regioni” del punteggio (1-3, 4-6, 7-9), corrispondenti ai tre livelli di valutazione. Vengono giudicati “in disaccordo” gli scenari in cui, dopo rimozione dei due giudizi estremi, la distribuzione dei rimanenti giudizi sia estesa a tutte e tre le “regioni”, anche con un solo giudizio. Vengono infine giudicate “dubbie” le situazioni in cui la distribuzione dei giudizi non rientra nei due casi citati. Nella pratica, la distribuzione dei risultati e le situazioni di disaccordo danno origine a situazioni complesse, per l'analisi delle quali vengono utilizzate metodologie di tipo statistico.

L'utilizzo dei risultati. I criteri di appropriatezza che si rendono disponibili alla fine del processo di valutazione descritto, possono essere usati per misurare in modo retrospettivo il sovrautilizzo di una procedura, indagare in modo prospettico sia il sovrautilizzo sia il sottoutilizzo, elaborare linee guida, predisporre strumenti di supporto alle decisioni “sul campo” dei clinici, identificare aree di incertezza che meritano l'avvio di ricerche cliniche mirate.

In conclusione, il metodo RAND consente di produrre criteri utilizzabili per pazienti “reali” e non indicazioni generali per pazienti “teorici”, criteri che possono essere sia uno strumento di valutazione e monitoraggio della qualità dell'assistenza sia la base per la produzione di raccomandazioni cliniche di comportamento (linee-guida).

Se da una parte il metodo RAND ha avuto una diffusa applicazione come strumento di valutazione dell'appropriatezza dell'assistenza, non sono mancate le critiche, la sostanza delle quali può essere così riassunta:

- a. La soggettività della valutazione:** medici esperti chiamati a giudicare l'appropriatezza di una determinata procedura possono basare le proprie scelte su aspetti tra loro diversi, intendere il concetto di beneficio secondo diverse priorità assegnate ai vari aspetti di un intervento (impatto sulla salute fisica, psichica, ecc.). In sostanza, il significato che ciascun medico attribuisce al concetto di rapporto rischio/beneficio può essere assai diverso: per alcuni sono più importanti i risultati tecnici di una determinata procedura, mentre per altri è più importante l'impatto della procedura sulla qualità della vita del paziente o le complicanze e gli effetti collaterali che ne conseguono.
- b. La difficoltà di coinvolgere** altre figure professionali e sociali, come amministratori sanitari, economisti e rappresentanti dell'utenza o delle associazioni di pazienti. È infatti opinione di molti che un giudizio di appropriatezza non possa non tenere conto delle opinioni di chi, in quanto fruitore di un servizio, è parte in causa di qualunque scelta assistenziale.

In definitiva, benché le valutazioni di appropriatezza clinica ottenibili dal metodo RAND possano sembrare oggettive ed esplicite, il processo attraverso il quale vengono prodotte rimane per molti aspetti “poco oggettivo”. La possibilità di migliorare questa metodologia, rendendo più espliciti i criteri di valutazione e verificando la possibilità di applicarla anche con la partecipazione di figure non mediche è oggetto di ulteriore e futura ricerca.

4. Approfondimenti

The NHS Atlas of Variation in Healthcare Regno Unito – Il Programma Right Care del Department of Health Ridurre le variazioni ingiustificate e indesiderate dell'assistenza sanitaria

Obiettivi e contenuti dell'Atlante

Nell'ambito del programma nazionale *Quality Innovation Productivity and Prevention* (QIPP) è stato attivato il programma *Right Care* (<http://www.rightcare.nhs.uk>) che si propone di migliorare la qualità dei servizi e dell'assistenza affrontandone in particolare le variazioni ingiustificate; il *NHS Atlas of Variation in Healthcare* (novembre 2010) è uno degli strumenti messi a punto al fine di facilitare la comprensione di questo problema. Più precisamente, l'Atlante viene proposto come complementare a un insieme di “pacchetti” di dati predisposti su misura, e disponibili per ciascuno dei 152 *Primary Care Trust* (PCT) inglesi, che includono set di informazioni cliniche ed economiche ed esplorano, per un gruppo definito di malattie, la posizione del singolo PCT a confronto con il dato nazionale e con la posizione di altri PCT con caratteristiche simili.

Le variazioni, cosiddette ingiustificate, nell'utilizzo di assistenza sanitaria, vengono definite, in accordo con quanto sostiene *Jack Wennberg*, come quelle variazioni “non spiegabili sulla base delle malattie e delle preferenze dei pazienti, come anche dei bisogni della popolazione, ma legate maggiormente alle differenze nelle caratteristiche del sistema di servizi, in particolare all' offerta locale di strutture/servizi/professionisti e al comportamento dei singoli professionisti”.

Alla predisposizione del documento hanno collaborato 15 organizzazioni, tra cui: il *NHS Primary Care Trust Network*, la *NHS Alliance*, il *National Prescribing Centre* e diversi osservatori di salute pubblica. L'Atlante è articolato sostanzialmente in quattro parti.

La **prima parte** offre un inquadramento generale dei problemi relativi allo studio delle variazioni nell'assistenza, e un approfondimento sulle variazioni di qualità, di esiti delle cure, di spesa, di attività, e di “**valore**”, attingendo dai risultati di ricerche e dall'esperienza del Servizio sanitario inglese.

La definizione di “valore” è soggettiva e complessa. Diversi gruppi di persone, come ad esempio i committenti dei servizi, i fornitori, i clinici, i manager e infine gli stessi cittadini possono attribuire un valore diverso allo stesso servizio o a una decisione riguardante un investimento o una sottrazione di risorse. Per un paziente, un intervento può essere di grande valore, mentre per un altro il valore dello stesso intervento è basso. Per una popolazione, il “valore” è meglio misurabile in termini di relazione tra outcome e costi. Tuttavia, è fuorviante considerare il valore per l’individuo e il valore per la popolazione come problemi separati. Se gli interventi sanitari cambiano, così si modifica anche la relazione tra rischi e benefici per una popolazione e ugualmente cambia l’equilibrio tra probabilità di beneficio e probabilità di rischio per il singolo paziente

Nell’Atlante le variazioni di qualità e di esiti vengono presentate in relazione alla popolazione e non, come avviene più frequentemente, in relazione al fornitore. La motivazione di tale scelta è riconducibile alla tipologia di destinatari a cui il lavoro si rivolge, soprattutto i clinici, i committenti dei servizi e coloro che pianificano e allocano le risorse dedicate all’assistenza sanitaria. Nelle conclusioni vengono proposti alcuni *step* per affrontare tale sfida, dalla necessità da parte dei committenti e dei clinici di chiedersi quali variazioni esistano nella loro popolazione locale e quali siano le possibili implicazioni sulla spesa sanitaria e sugli esiti di salute fino alla produzione di *report* annuali.

Nella **seconda parte** vengono presentati sinteticamente alcuni strumenti di lavoro (per ciascuno di essi viene indicato l’indirizzo web di riferimento per ulteriori approfondimenti), a supporto dei *Primary Care Trusts*, per la lettura e comprensione delle variazioni di spesa, dei rispettivi *drivers* e degli esiti di salute.

Ad esempio, il *Programme Budgeting Benchmarking Tool* (PBBT) consente di identificare l’entità della spesa di un PCT per ogni programma di assistenza, a confronto con la media nazionale e con quella di altri PCT simili. Il PBBT contiene dati di spesa relativi ai 23 principali programmi di intervento sanitario in Inghilterra, (dedicati, ad es., ai tumori, alla salute mentale, ecc.) per tutti i PCT. L’uso di questo strumento consente inoltre a ciascun PCT di monitorare i propri dati di spesa nel tempo.

Lo *Spend and Outcome Tool* (SPOT) combina invece dati di spesa e dati inerenti gli esiti di salute per identificare i programmi di assistenza di un PCT potenzialmente *outliers* se confrontati con altri PCT. Questo strumento consente a ciascun PCT di studiare la relazione tra risultati e spesa, soprattutto per quanto riguarda i programmi principali; e di capire per quali programmi è necessario

ottenere ulteriori informazioni per identificare i *driver* di spesa. Il limite principale di questo strumento consiste nel fatto che, per alcuni programmi, non sono disponibili dati di *outcome*.

Una volta identificati i programmi meritevoli di ulteriori approfondimenti, i *Programme Budgeting Atlases* (PBA) consentono di identificare i possibili *driver* di spesa – ad es. i ricoveri in emergenza, i tassi di prescrizione di farmaci - attraverso la combinazione dei dati di spesa con un ampio insieme di dati di attività e di esito. Ulteriori informazioni in questo senso sono offerte dai dati dei *NHS comparators*, uno strumento di *benchmarking* che rende disponibili dati disaggregati per fornitore e per medico di medicina generale (*GP practice*) e consente di monitorare attività, *outcome* e costi nel tempo.

Un altro strumento, l'*Inpatient Variation Expenditure Tool* consente ai PCT di misurare e confrontare la spesa ospedaliera per le aree ad elevata frequenza di malattia e ad elevato volume di ricorso ai servizi. Lo strumento utilizza un aggiustamento dei dati di spesa secondo i bisogni della popolazione.

La **terza parte** descrive brevemente la metodologia impiegata (vedi capitolo successivo, Metodi).

Nella **quarta parte** vengono presentati gli indicatori, raggruppati in 17 aree tematiche, selezionate attraverso un processo di consultazione a cui hanno partecipato i *National Clinical Directors*, diversi *Department of Health teams*, diversi Osservatori di salute pubblica e Osservatori su tematiche specifiche (es. obesità, salute materna e infantile). La scelta dei temi è stata guidata dall'intento di rappresentare un ampio insieme di specialità cliniche e di problemi di salute, includendo in particolare alcune aree già oggetto di specifici programmi di miglioramento della qualità e di aumento del valore, come ad esempio: i problemi neurologici (epilessia), i problemi vascolari (*stroke*, attacco ischemico transitorio, malattia coronarica), i problemi del sistema respiratorio (asma, malattia polmonare cronico-ostruttivo), del sistema muscolo-scheletrico (ginocchio, anca), la salute materna e riproduttiva (parti cesarei), gli *screening* (dell'udito per i nuovi nati), i servizi diagnostici (risonanza magnetica, tomografia assiale computerizzata), le prescrizioni farmaceutiche (antibatterici) e altre aree riguardanti gli anziani (ricoveri di urgenza) e l'assistenza di fine vita (decessi avvenuti in ospedale).

Per ciascun indicatore, appartenente ad una delle aree tematiche sopra esemplificate, è stata predisposta una scheda, contenente:

- rappresentazione dei valori dell'indicatore mediante mappa e istogramma;
- inquadramento dell'argomento mediante la misura del fenomeno e la citazione di programmi dedicati;
- sintetica analisi della grandezza della variazione e, in alcuni casi, spiegazione delle possibili cause;

- indicazioni per l'azione tratte da programmi dedicati;
- bibliografia e sitografia per l'approfondimento.

Metodi

Il capitolo dedicato ai metodi chiarisce:

- le modalità di selezione dei temi rappresentati (ne è stato fatto cenno nel capitolo precedente, paragrafo dedicato alla quarta parte dell'Atlante);
- i metodi di calcolo delle stime;
- il metodi di conversione dei dati in mappe e i *software* utilizzati;
- i metodi di standardizzazione e di aggiustamento secondo i “bisogni” della popolazione.

Nelle mappe, i confini delle aree geografiche corrispondono nella maggior parte dei casi a quelli dei *Primary Care Trusts* (approssimativamente sovrapponibili alle nostre aziende USL), in alcuni casi a quelli delle *Local Health Authorities* (sovrapponibili ai Comuni italiani) e in limitatissimi casi a quelli delle *Strategic Health Authorities* (corrispondenti all'incirca alle Regioni italiane).

Un approfondimento specifico è stato dedicato alla città di Londra.

I valori degli indicatori sono presentati in quintili. Ciascun quintile include un *range* di valori compreso entro la quinta parte del totale cumulato di tutti i valori misurati. I quintili non necessariamente contengono numeri uguali.

Negli istogrammi ogni barra corrisponde ad un *Primary Care Trusts*, in tutto 152. In alcuni istogrammi il valore degli indicatori è espresso con gli intervalli di confidenza, al fine di fornire un orientamento sul livello di incertezza delle stime: minore è l'ampiezza dell'intervallo e maggiore è la probabilità che l'indicatore sia preciso.

L'ampiezza della variazione è espressa solitamente mediante il rapporto tra il valore maggiore e il valore minore della serie. Al fine di tenere sotto controllo i possibili artefatti (es. omissione accidentale di parte del *dataset*, scarsa accuratezza nella restituzione dei dati) la variazione viene calcolata sia rispetto all'insieme totale di valori, sia all'insieme dopo sottrazione:

- dei 5 valori più elevati e dei 5 più bassi, quando i dati sono disaggregati per *Primary Care Trust*;
- dei 10 valori più elevati e dei 10 più bassi quando i dati sono disaggregati per *Local Authority* (LA).

Per rendere confrontabili le popolazioni è stata utilizzata la standardizzazione diretta dell'indicatore o la pesatura secondo il “bisogno”, definito in diversi modi a seconda della fonte informativa utilizzata, ma per lo più basato su variabili quali sesso, età, disabilità (*Disability Living Allowance*), deprivazione economica, distanza dai servizi.

Esempi

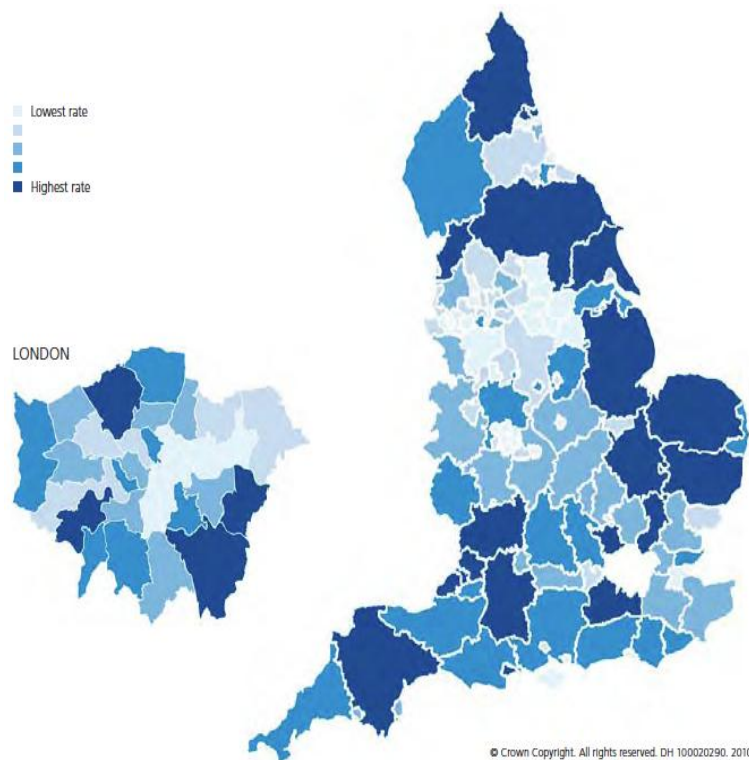
A scopo esemplificativo vengono di seguito descritti alcuni indicatori dell'Atlante, selezionati sulla base della potenziale rilevanza per il medico di medicina generale o di interesse generale, anche sulla base della presenza di specifici programmi a diversi livelli istituzionali del sistema sanitario nazionale italiano. In particolare:

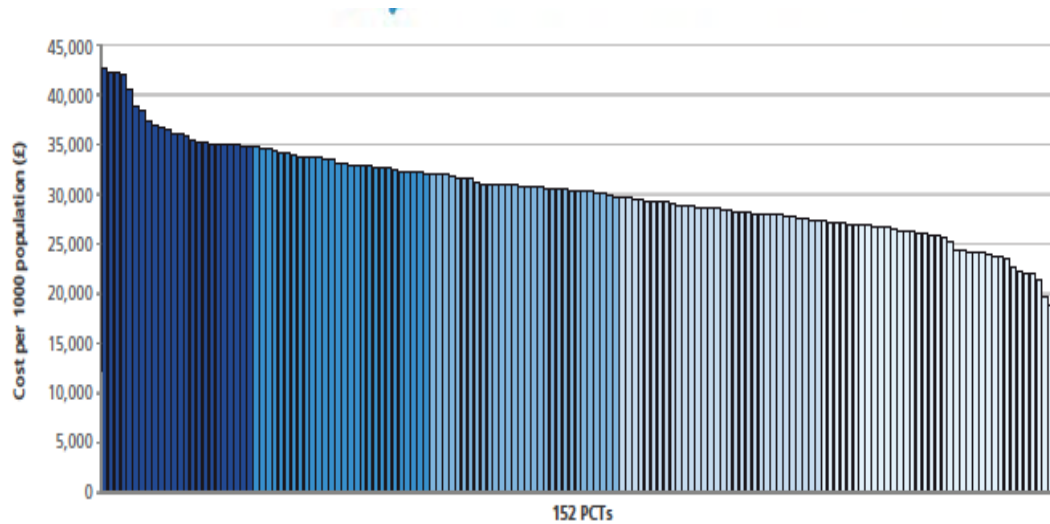
- percentuale di persone diabetiche incluse nei nove processi assistenziali chiave raccomandati;
- percentuale di persone ricoverate in ospedale a seguito di *stroke* che hanno trascorso il 90% del tempo in una *stroke unit* ;
- tasso standardizzato di ricovero in urgenza di anziani >75 anni *1.000 abitanti;
- percentuale di decessi avvenuti in ospedale (sul totale dei decessi avvenuti in una *Local Authority*).

Ciascuna delle schede proposte ripropone la struttura ed i contenuti delle schede-indicatore dell'Atlante (contesto, ampiezza della variazione, alternative per l'azione). Nell'Atlante, ciascun indicatore è rappresentato su mappa e con istogramma (vedi figure qui sotto). Per semplicità, le schede da noi proposte riportano soltanto l'istogramma.

Rappresentazione degli indicatori nelle singole schede del NHS Atlas of Healthcare.

Esempio: Spesa ospedaliera per pazienti oncologici per 1000 abitanti . Analisi per PCT. Aggiustamento per età, sesso e bisogni).





Il *NHS Atlas of Variation in Healthcare* (novembre 2010) è disponibile sia nella versione on-line e interattiva sia nella versione.pdf all'indirizzo web: <http://www.rightcare.nhs.uk/atlas/index.html>.

Percentuale di persone diabetiche che ricevono i nove processi assistenziali chiave raccomandati. Analisi per *Primary Care Trust*, anni 2008/09.

Contesto

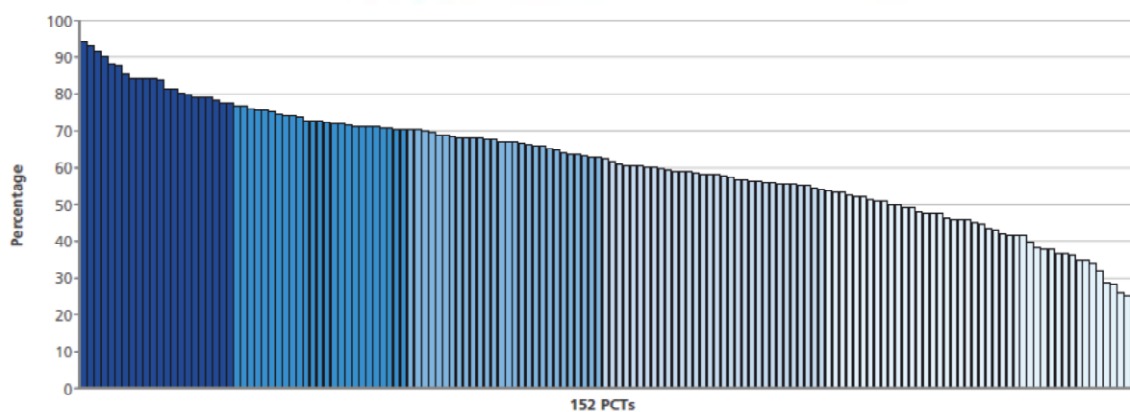
Secondo un *framework* nazionale per lo sviluppo dei servizi e il miglioramento degli esiti di salute delle persone diabetiche, e in accordo alle raccomandazioni del *National Institute for Health and Clinical Excellence* (NICE), nove sono i controlli chiave per la gestione e il monitoraggio del diabete, tra cui: peso corporeo, pressione arteriosa, glucosio, colesterolo, rapporto albumina/creatinina nelle urine, salute del piede. Si stima che in Inghilterra l'11,6% delle morti di persone di età 20-79 anni siano attribuibili al diabete, con un intervallo di variabilità nei *Primary Care Trusts* compreso tra un minimo del 9% ed un massimo del 17% circa. La fonte informativa dell'indicatore è rappresentata dal Sistema informativo nazionale sul diabete, nell'ambito del quale viene svolta regolarmente attività di *audit*, giunta con gli anni 2008/09 alla sesta edizione. In particolare, viene fatto riferimento all'*audit* delle persone diabetiche adulte, che include i *patient records* di 1 milione e 620 mila individui, pari al 75% circa dei 2,2 milioni di persone con diagnosi di diabete. Dai risultati dell'*audit* emerge come negli anni 2008/09 il 50.8% dei diabetici di tipo 2 e il 32.2% dei diabetici di tipo 1 abbiano ricevuto tutti e nove i controlli. Tra questi, il meno frequente è la misurazione del rapporto albumina/creatinina nelle urine, mentre il più frequente è la misurazione della pressione arteriosa.

Ampiezza della variazione

Tra i PCT, il valore massimo della proporzione di diabetici che hanno ricevuto tutti e nove i processi di cura raccomandati dal NICE è pari al 70%, percentuale 35 volte più grande del valore minimo, pari a circa il 2%. Se si escludono i 5 PCT con i valori più elevati ed i 5 PCT con i valori più bassi, tale ordine di grandezza è pari a 5 volte.

Indicazioni per l'azione

Le azioni necessarie per ridurre tale variabilità vengono proposte nel *Primary Care Quality and Outcomes Framework* (2004), uno schema di incentivi per i medici di medicina generale, ad adesione volontaria, che premia il professionista in base alla capacità di raggiungere determinati risultati, valutati rispetto ad un set predefinito di indicatori, suddivisi per aree cliniche.



Percentuale di persone ricoverate in ospedale a seguito di *stroke*, che hanno trascorso il 90% del tempo in una *stroke unit*. Analisi per Primary Care Trust, anni 2009/10.

Contesto

Nella strategia nazionale per lo *stroke* vengono indicati i cambiamenti necessari per migliorare gli esiti di salute delle persone che ne sono colpite e le Linee guida del *National Institute for Health and Clinical Excellence* (NICE) indicano un chiaro standard: tutte le persone con sospetto di *stroke* devono essere ricoverate direttamente presso una *stroke unit*, dopo l'iniziale valutazione dei servizi territoriali oppure dei dipartimenti di emergenza-urgenza. Il ricovero presso una *stroke unit* diminuirebbe per il paziente la probabilità di morte ed aumenterebbe la probabilità di recupero della autonomia.

Ampiezza della variazione

Tra i *Primary Care Trusts*, il valore massimo delle persone colpite da *stroke* che hanno trascorso il 90% del tempo presso una *stroke unit*, è pari a circa il 92%, percentuale che è 4 volte più grande del valore minimo, pari a circa il 23%. Se si escludono i 5 PCT con i valori più elevati ed i 5 PCT con i valori più bassi, tale ordine di grandezza è di circa 3 volte. Secondo i risultati emersi dagli studi del *Royal College of Physicians - National Sentinel Audit of Stroke*, nel 2009, il numero di posti letto nelle *stroke unit* è pari al numero di pazienti colpiti da *stroke* e ricoverati negli ospedali inglesi; tuttavia, il 75% dei pazienti colpiti da *stroke* non è stato ricoverato direttamente presso una *stroke unit*. Tra le ragioni del fenomeno, quasi la metà degli ospedali riferiscono di aver ricoverato tali pazienti in unità non specialistiche in quanto nella *stroke unit* il posto letto non era immediatamente disponibile al momento dell'evento.

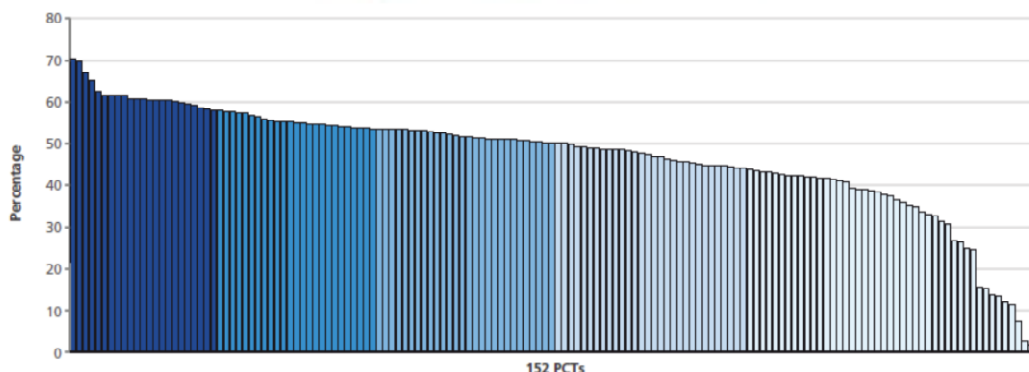
Indicazioni per l'azione

Le azioni chiave necessarie a migliorare l'accesso alle *stroke unit* consistono in:

- accesso immediato alla tomografia computerizzata;
- ricovero diretto dal dipartimento di emergenza-urgenza o dall'ambulanza;
- specialisti di reparto in turno sette giorni su sette;
- infermieri formati per la gestione dello *stroke*.

Nell'ambito dei casi studio dello *Stroke Improvement Programme*, si sono dimostrati efficaci:

- vincolare l'utilizzo dei posti letto nelle *stroke unit* solamente ai pazienti colpiti da *stroke*;
- sviluppare una forza lavoro specializzata nel trattamento e gestione dello *stroke* e flessibile;
- lavorare in collaborazione con le ambulanze;
- passare ad una terapia di sei giorni settimanali.



Tasso standardizzato (metodo diretto) di ricovero in urgenza di anziani >75 anni *1.000 abitanti, per *Primary Care Trust*. Anni 2008/09.

Contesto

Gli anziani >75 anni, specialmente coloro che soffrono di due o più patologie croniche, sono responsabili:

- del 40% dei ricoveri in urgenza;
- del 50% delle giornate di degenza;
- di più del 50% dei re-ricoveri;
- di più dell'80% dei ritardi nel trasferimento ad altro luogo di cura.

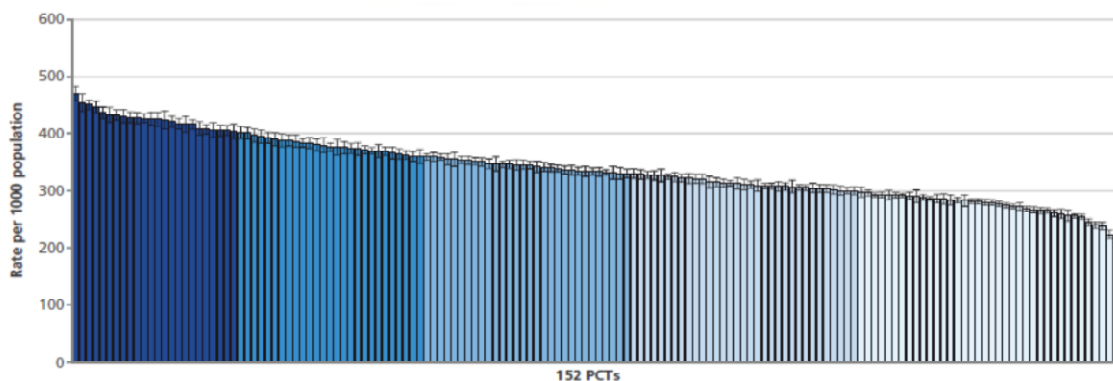
Alcuni ricoveri programmati possono portare dei benefici agli anziani, cosiddetti fragili, con condizioni multiple di cronicità. I problemi acuti di salute che insorgono in un quadro di disturbi pre-esistenti possono infatti comportare una significativa e permanente perdita di funzioni e disabilità, qualora non siano accuratamente diagnosticati ed efficacemente trattati. Tuttavia, un ricovero inappropriato può avere gravi effetti avversi per l'anziano e condurre ad una irreversibile perdita di abilità fisiche e mentali.

Grandezza della variazione

Tra i PCT, il valore massimo di questo indicatore è pari a circa 500 ricoveri di anziani > 75 anni ogni 1.000 abitanti, valore 2 volte più grande del valore minimo, pari a circa 250 ricoveri. Se si escludono i 5 PCT con i tassi più elevati ed i 5 PCT con i tassi più bassi, tale ordine di grandezza si mantiene inalterato. Si può osservare, inoltre, un ulteriore elemento di variabilità esaminando i dati della *Care Quality Commission* relativi agli esiti dei ricoveri in urgenza. Dai risultati emerge infatti come la variazione nella percentuale di anziani >85 anni presi in carico con l'assistenza domiciliare dopo un ricovero in urgenza sia dell'ordine di grandezza di quattro volte, ovvero la percentuale più elevata di anziani inseriti nel programma di assistenza domiciliare, pari al 20%, è quattro volte più grande della percentuale minima, pari al 5%.

Indicazioni per l'azione

Per i PCT che presentano i tassi di ricovero in urgenza più bassi è necessario verificare che un atteggiamento "a priori" dei professionisti nei confronti dell'età non comporti il trascurare la necessità di un ricovero tempestivo in ospedale come strategia assistenziale per rendere reversibile una condizione di progressivo deterioramento. Per i PCT che presentano i tassi più elevati è necessario verificare la presenza di ricoveri inappropriati e prevenirli attraverso lo sviluppo di servizi di assistenza integrati, orientati al supporto delle persone nelle strutture residenziali, ai loro *caregivers*, al coinvolgimento dei medici di medicina generale. A tal riguardo, nel 2009 è stato avviato il programma biennale *Integrated Care* con l'obiettivo di esplorare le modalità con cui l'assistenza sanitaria e sociale possano essere meglio erogate a livello locale soprattutto nei confronti degli anziani fragili. Tra l'altro, alcune tipologie di interventi per le persone anziane con problemi multipli e complessi sono simili a quelle utilizzate per persone sofferenti di una singola condizione cronica, come ad esempio l'*empowerment*, la *self-care* e il *case management*.



Percentuale di decessi avvenuti in ospedale, sul totale dei decessi in un'area definita, nello stesso intervallo di tempo. Analisi per *Local Authority*, anni 2006-2008.

Contesto

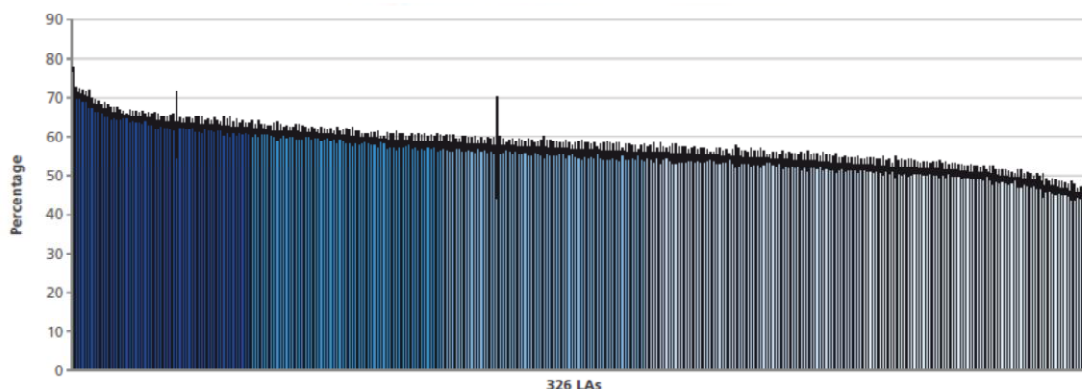
In Inghilterra si verificano circa 450 mila decessi ogni anno, oltre la metà dei quali avviene in ospedale. Le persone dovrebbero avere l'opportunità di scegliere dove morire e dovrebbero essere evitate le ospedalizzazioni, quando possibili, di coloro che sono in condizioni terminali.

Ampiezza della variazione

Nel triennio 2006-08 il 58% dei decessi area geografica-specifici sono avvenuti in ospedale. L'intervallo di variazione tra le *Local Authorities* (LAs) è compreso tra un valore massimo pari al 76% di morti avvenuti in ospedale sulle morti totali in quella determinata area e un valore minimo pari al 44%. Nella metà di tutte le LAs la proporzione di decessi in ospedale è compresa tra il 53% e il 61%. I dati suggeriscono una carenza di servizi territoriali in grado di offrire alle persone la opportunità di scegliere dove morire, con la conseguenza di un utilizzo inappropriato dei servizi ospedalieri (per acuti). Il ricorso al ricovero ospedaliero sarebbe da evitare se clinicamente non appropriato e se contrario ai desideri della persona.

Indicazioni per l'azione

L'assistenza di fine vita è erogata da diverse agenzie, incluso il Servizio sanitario nazionale, i servizi sociali delle *Local Authorities*, gli *hospice*, le *charities*. I committenti dovrebbero valutare il modo migliore di coordinare questa tipologia di assistenza. Più dell'80% dei decessi coinvolgono gli anziani (≥ 65 anni), i quali, con maggiore probabilità, soffrono al momento della morte di patologie multiple e croniche. Bisognerebbe porre attenzione al supporto degli anziani sofferenti di problemi multipli in contesti assistenziali diversi dall'ospedale. Una molteplicità di luoghi possono rappresentare una casa per l'anziano, non solo la propria dimora, come ad esempio le case protette e le strutture residenziali. Ciascuna persona per la quale la morte non rappresenti un evento inatteso dovrebbe avere un piano di assistenza che tenga conto dei suoi bisogni e delle sue preferenze riguardo alla fine della vita. Un tale piano dovrebbe essere conosciuto dai medici di medicina generale, dai servizi sanitari territoriali e ospedalieri e dai servizi sociali. È possibile ottenere un maggiore valore per l'individuo e per la popolazione spostando le risorse dall'assistenza ospedaliera a quella domiciliare.



L'appropriatezza dei ricoveri in Emilia-Romagna

Atlante dell'appropriatezza organizzativa

I ricoveri ospedalieri in Emilia-Romagna, 2001-2004

Negli ultimi dieci anni la normativa nazionale ha indicato alcuni obiettivi precisi alle regioni in materia di appropriatezza organizzativa: nel 2001 il DPCM “Definizione dei livelli essenziali di assistenza”, definiva i DRG a elevato rischio di inappropriatezza; successivamente la L. 133/2008 prescriveva i relativi controlli sulle cartelle cliniche; infine il Nuovo Patto per la Salute 2010-2012 proponeva in due Allegati un nuovo elenco di DRG “inappropriati”, a integrazione di quelli già esistenti.

Gli indirizzi della regione Emilia-Romagna

Allo scopo di assicurare una risposta agli indirizzi normativi nazionali, l’Agenzia Sanitaria Regionale dell’Emilia-Romagna, in collaborazione con il *Jefferson Medical College*, ha sviluppato fin dal 2001 una propria metodologia, basata su criteri espliciti di appropriatezza del ricovero in regime ordinario (per i pazienti di età compresa tra 18 e 74 anni) e sulla verifica di fattibilità del processo di sostituzione del ricovero ordinario con altri *setting* assistenziali, garantendo comunque uguali livelli di qualità e sicurezza.

Questa metodologia è stata inizialmente applicata a 62 DRG, comprendenti i 43 DRG ritenuti “ad alto rischio di inappropriatezza in regime ordinario” dall’Allegato 2C del DPCM del 29/11/2001 e anche quelli ad essi correlati o già oggetto di specifiche politiche di promozione dell’appropriatezza attivate in Emilia-Romagna. Questo ha consentito la valutazione dell’appropriatezza dei ricoveri¹ riferiti agli anni 2001-2004, oggetto di una pubblicazione tecnica, “L’Atlante della appropriatezza organizzativa”.

Il corrispondente percorso normativo si è concretizzato in alcune deliberazioni regionali: la **DGR n. 1872/2004** con cui, attraverso la metodologia definita dall’Agenzia Sanitaria Regionale il gruppo iniziale di 43 DRG veniva esteso a 51 DRG a rischio di inappropriatezza (con la relativa definizione delle soglie di ammissibilità); la deliberazione **n. 2126/2005** che, facendo proprio l’elenco di DRG definito dalla ASR per il progetto “Atlante della appropriatezza organizzativa”, portava a 62 il numero di DRG “potenzialmente inappropriati”; successivamente la **DGR n. 1920/2007** dava conto dei risultati certificati dal progetto, rilevando come, anche grazie alle politiche sanitarie e tariffarie regionali messe in atto dal 1999 al fine di disincentivare il ricorso al ricovero ordinario per attività e problemi clinici trattabili con pari efficacia e sicurezza in regimi

assistenziali meno onerosi, la regione Emilia-Romagna fosse riuscita ad ottenere un decremento dei ricoveri ad alto rischio di inappropriately in regime ordinario pari al 19,4%.

Infine, la **DGR 1890/2010** stabilisce l'integrazione dei 61 DRG di cui alla DGR 2126/2005 con i DRG medici presenti nelle tabelle A e B del "Nuovo patto per la salute per gli anni 2010-2012" per un totale complessivo di 89 DRG, che sono da considerarsi potenzialmente inappropriati in relazione al regime di erogazione.

È necessario inoltre ricordare che, a decorrere dall'1 gennaio 2009, presso tutte le Aziende Sanitarie pubbliche e private è stata adottata la versione italiana 2007 della *International Classification of Diseases 9th revision - Clinical Modification (ICD9CM)* e la corrispondente versione 24 della Classificazione *Diagnosis Related Groups*. Il nuovo sistema di classificazione ha previsto, tra l'altro, la soppressione di 24 DRG e la contestuale introduzione di 54 nuovi DRG. Per questo nuovo set di DRG è stato quindi necessario sviluppare nuovi criteri di appropriatezza, coinvolgendo anche *panel* di clinici dell'Emilia-Romagna, per contestualizzarli alla realtà organizzativa e professionale italiana, in particolare per quanto riguarda i DRG chirurgici.

Questo ha comportato la valutazione di ulteriori 179 categorie diagnostiche; in aggiunta, 27 ulteriori diagnosi secondarie sono state identificate come comorbidità significative

I contenuti dell'Atlante: dati 2001-2004.

L'Atlante, frutto della collaborazione tra l'Agenzia sanitaria regionale dell'Emilia-Romagna, il *Center for Research in Medical Education and Health Care del Jefferson Medical College* (Philadelphia, USA) e un panel di clinici di sette Aziende sanitarie della Regione, presenta un metodo di approccio al problema della valutazione dell'appropriatezza organizzativa che integra le informazioni desumibili dalla scheda di dimissione ospedaliera (SDO) con il giudizio clinico.

La valutazione di appropriatezza è stata condotta sui ricoveri complessivamente prodotti dalla rete ospedaliera regionale, nel periodo 2001-2004, per pazienti di età compresa tra 18 e 74 anni e residenti in Emilia-Romagna. In particolare, sono stati selezionati i ricoveri riconducibili ai 62 DRG ritenuti "ad alto rischio di inappropriately in regime ordinario" e comprendenti i 43 DRG definiti dal DPCM del 29/11/2001, ("Decreto LEA", Allegato 2C) e quelli ad essi correlati o già oggetto di specifiche politiche di promozione dell'appropriatezza attivate in Emilia-Romagna.

I criteri sviluppati per valutare la tempestività e l'appropriatezza dei ricoveri, inizialmente messi a punto dal *Jefferson Medical College*, sono stati successivamente validati da un panel di clinici dell'Emilia-Romagna, per contestualizzarli alla realtà organizzativa e professionale italiana. I criteri sono basati sulla diagnosi principale di dimissione - e relativo stadio di gravità - e sulla

presenza di specifiche comorbidità, anch'esse con relativa gravità; per quanto riguarda i ricoveri chirurgici, si è tenuto anche conto delle procedure effettuate.

I risultati derivanti dall'applicazione della metodologia così sviluppata hanno evidenziato una riduzione, dal 2001 al 2004, dei casi potenzialmente inappropriati attribuiti ai 62 DRG selezionati, sia in termini assoluti sia di incidenza percentuale: nel 2001 i ricoveri ordinari potenzialmente inappropriati costituivano il 17,8% del complesso dei ricoveri ordinari e diurni esaminati, mentre nel 2004 l'incidenza di potenziale inappropriatezza risultava ridotta al 13,5%.

Nell'analisi dei *trend* temporali si deve anche considerare il graduale passaggio di parte delle prestazioni in esame al regime ambulatoriale (come ad es. gli interventi di cataratta o di tunnel carpale).

I criteri di appropriatezza

I ricoveri effettuati in regione sono stati classificati in gruppi di appropriatezza, combinando le definizioni dei DRG e la classificazione del *disease staging*, a partire dal 2001, con successive validazioni e revisioni.

Disease Staging: sistema di classificazione che individua gruppi di pazienti omogenei rispetto al livello di gravità clinica (espressa come probabilità di morte o di disabilità permanente), caratterizzati da prognosi e bisogni di assistenza simili e quindi anche da profili di cura analoghi. Il sistema identifica 596 categorie diagnostiche, in base alla localizzazione e all'eziologia della condizione clinica di base. Per ognuna di queste categorie sono previsti quattro livelli di gravità gerarchicamente ordinati in base alla prognosi e differenziati in ordine alle manifestazioni fisiopatologiche e al grado di diffusione locale e di interessamento sistemico.

DRG medici (n=42)

Ricovero inappropriato

Segni o sintomi non specifici oppure condizioni croniche che possono essere assistite in *day hospital* o in ambito ambulatoriale (es. emicrania, diabete mellito in stadio 1 senza complicazioni).

Ricovero appropriato tempestivo

Ricovero necessario, per una patologia che richiede il ricovero in regime ordinario (ad es. infarto miocardico acuto)/ segni o sintomi (es. sincope) che potrebbero giustificare un ricovero ospedaliero per escludere la presenza di eventuali problemi più gravi.

Ricovero appropriato tardivo

Ricovero ordinario necessario, ma la condizione che richiede il ricovero costituisce una complicanza o la progressione di una malattia che avrebbe potuto essere evitata mediante un più efficace e tempestivo trattamento ambulatoriale o domiciliare (es. nefrolitiasi con idronefrosi).

In base a questi criteri, i ricoveri ordinari con degenza superiore a un giorno possono essere considerati potenzialmente inappropriati (assistenza ospedaliera in degenza ordinaria non necessaria).

DRG chirurgici (n=20)

Sono state considerate non solo la gravità delle condizioni cliniche, ma anche la natura della procedura eseguita in termini di effettuabilità in *day surgery* o in ambulatorio. La necessità dell'intervento chirurgico o della procedura rispetto alle caratteristiche cliniche del paziente (la cosiddetta appropriatezza clinica) non è oggetto di valutazione.

Ricovero inappropriato

Limitata gravità della malattia e basso rischio della procedura. È possibile il trattamento in ambito ambulatoriale a condizione che siano disponibili le risorse necessarie (ad es. tunnel carpale).

Ricovero appropriato tempestivo

Il ricovero in regime ordinario è appropriato in considerazione del tipo di intervento o della gravità della malattia, oppure l'intervento chirurgico potrebbe essere svolto in regime ambulatoriale, ma a causa della presenza di patologie correlate (che riguardano ad esempio il sistema nervoso centrale, il sistema cardiovascolare o quello respiratorio) può essere necessario un ricovero in regime ordinario superiore a 1 giorno.

Ricovero appropriato tardivo

Il ricovero in regime ordinario è richiesto per la presenza di complicanze che avrebbero potuto essere prevenute con un trattamento tempestivo in regime ambulatoriale o di *day surgery* (ad esempio, ernia inguinale strozzata).

L'analisi dei profili di appropriatezza è stata condotta differenziando la casistica chirurgica da quella medica; in particolare, per la prima è stata osservata una percentuale di potenziale inappropriata più bassa di quella osservata per la casistica medica (5,4% vs 23,5%), a conferma del notevole sforzo effettuato nel corso degli ultimi anni dalle strutture per fornire le prestazioni in regime diurno o ambulatoriale.

Nonostante il quadro complessivo soddisfacente, è stata rilevata una significativa variabilità nella frequenza di inappropriata tra le diverse tipologie di soggetti erogatori e tra le singole strutture. I risultati di questo studio sembrano quindi indicare nella promozione di forme di audit clinico, già in parte attivate a livello locale, le future azioni per migliorare l'efficacia nell'utilizzo del regime di ricovero ordinario e per incentivare l'appropriatezza organizzativa.

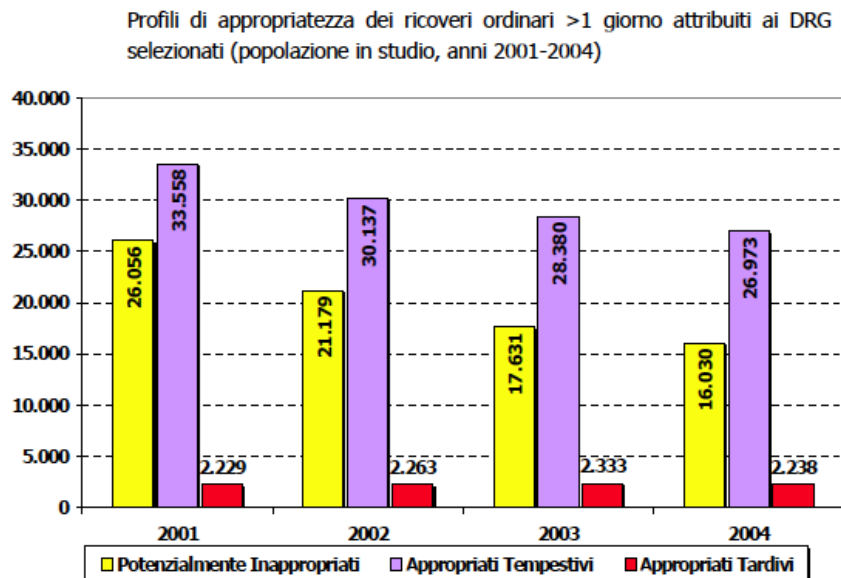
I ricoveri (ordinari e diurni) complessivamente analizzati nel quadriennio in studio per la definizione dei *profili di appropriatezza* sono stati 2.008.993, di cui 550.732 (il 27,4%) appartenenti ai 62 DRG selezionati.

Dall'analisi sono stati esclusi i ricoveri di: a) soggetti con età inferiore ai 18 anni e superiore ai 74 anni, in ragione dei numerosi fattori demografici, clinici e socio-economici che incidono sulla decisione di ricovero di questi pazienti; b) pazienti residenti in altre regioni italiane (la cosiddetta mobilità interregionale attiva) per evitare distorsioni sul livello di gravità dei casi trattati; c) i ricoveri di pazienti residenti in Emilia-Romagna e ricoverati in altre regioni (la mobilità passiva); d) pazienti ammessi e dimessi dai reparti di lungodegenza, recupero e riabilitazione funzionale, neuroriabilitazione, unità spinale e psichiatria, in quanto non riconducibili ai ricoveri per acuti, anche dal punto di vista tariffario; e) pazienti trattati in strutture poliambulatoriali chirurgiche.

I profili di appropriatezza 2001-2004

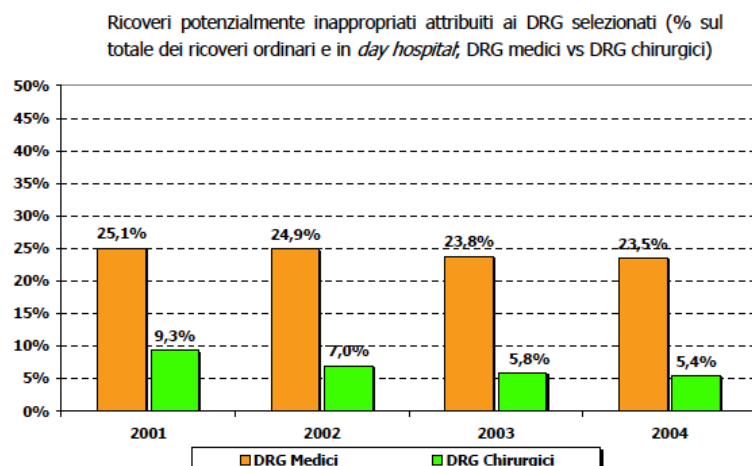
Nell'arco dei quattro anni, i casi potenzialmente inappropriati sono diminuiti sia in termini assoluti (circa 26.056 nel 2001 e 16.030 nel 2004) sia in termini di incidenza sui ricoveri complessivi (v. Figura 1. Fonte. Dossier ASR 132/2006).

Figura 1



Il trend di riduzione dei ricoveri potenzialmente inappropriati differisce però per la casistica medica e per quella chirurgica (v. Figura 2. Fonte. Dossier ASR 132/2006).

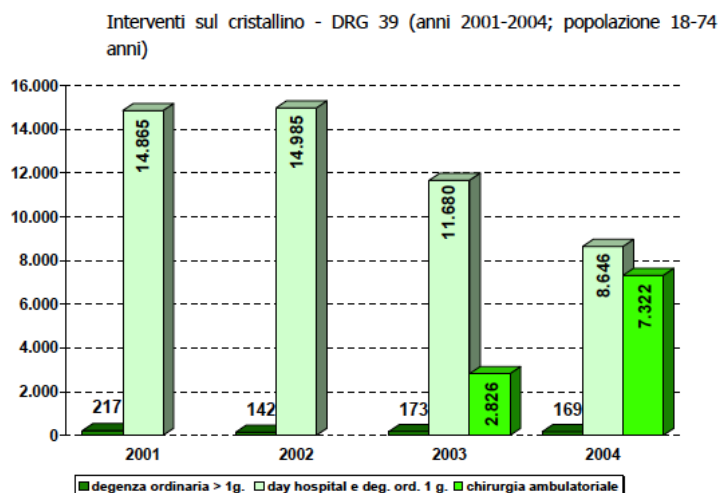
Figura 2



DRG chirurgici. In questo caso, nell'arco dei 4 anni la riduzione dell'incidenza di potenziale inapproprietezza è stata pari a circa 3,9 punti percentuali (dal 9,3% nel 2001 al 5,4% nel 2004). La riduzione ha interessato la quasi totalità dei 20 DRG considerati. Casi particolarmente esemplificativi, anche per la relazione con il *setting* della chirurgia ambulatoriale, sono rappresentati dagli interventi sul tunnel carpale, sul cristallino, sul ginocchio, per ernia inguinale e femorale e

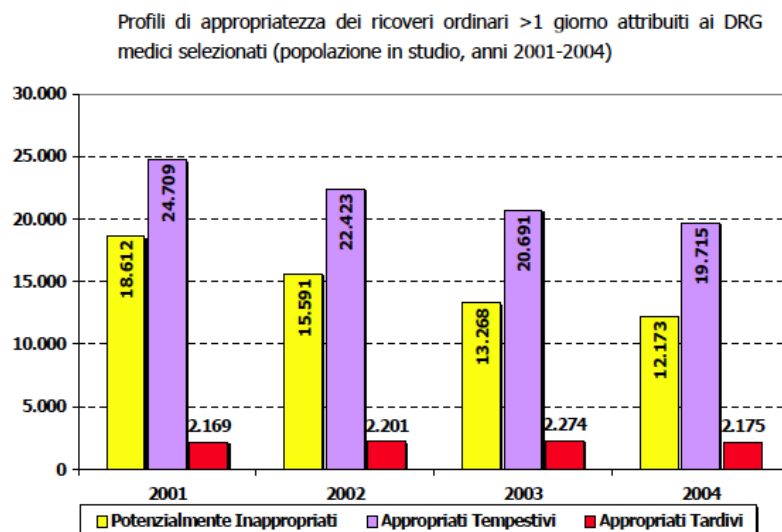
artroscopia). Nella figura 3, l'esempio della chirurgia del cristallino (Fonte: Dossier ASR 132/2006), per la quale si può osservare il numero molto basso di ricoveri ordinari con degenza >1 giorno e l'aumento, dal 2003, del ricorso alla chirurgia ambulatoriale, in concomitanza con una speculare riduzione del ricorso alla *day surgery*.

Figura 3



DRG medici. In questo caso (v. Figura 4), nell'arco dei 4 anni la riduzione dell'incidenza di potenziale inappropriatezza è stata pari a 1,6 punti percentuali (dal 25,1% nel 2001 al 23,5% nel 2004). In numero assoluto, dal 2001 al 2004 i ricoveri potenzialmente inappropriati si riducono del 34% (da 18.612 casi a 12.173); ma si riducono anche i ricoveri qualificati come appropriati. Il caso dei ricoveri appropriati ma potenzialmente tardivi è particolare: mentre il volume è stabile in tutti gli anni considerati, la loro proporzione è in aumento: nel 2001 costituivano circa il 4,8% della casistica medica in studio erogata in regime ordinario maggiore di un giorno (2.169 casi su 45.490), mentre nel 2004 ne costituiscono il 6,4% (2.175 casi su 34.063).

Figura 4



I **ricoveri appropriati tardivi** meritano un approfondimento specifico in quanto rappresentano, secondo i criteri che li definiscono, una quota di ricoveri sostanzialmente prevenibili con una migliore assistenza sanitaria prestata in ambito extra-ospedaliero. Come è stato detto, essi rappresentano una piccola quota della casistica analizzata (n=2238 nel 2004, relativamente ai 62 DRG selezionati), ma colpisce il loro andamento nel periodo osservato, percentualmente in aumento. Per una migliore comprensione del fenomeno, è stata quindi inclusa nell'Atlante un'ulteriore analisi che, sulla base di alcune indicazioni offerte dall'osservazione sui DRG "ad elevato rischio di inappropriatezza", esplora la distribuzione di questi ricoveri nell'intera casistica ospedaliera. Dall'osservazione dei 62 DRG, nel solo anno 2004, emerge come il 94,5% di questi ricoveri sia riconducibile ad alcuni DRG, come riportato in Figura 5.

Figura 5 - Ricoveri attribuiti ai 62 DRG selezionati con almeno 1 caso appropriato tardivo (ricoveri ordinari >1 giorno (anno 2004, totale ricoveri appropriati tardivi pari a 2238 casi). Fonte: Dossier ASR 132/2006.

DRG	Totale ricoveri ord. >1 g N	Ricoveri appropriati tardivi			
		N	% sui ric. ord. >1 g	% sul totale di 2.238 casi	% cumulata
323-324 M - Calcolosi urinaria con/senza CC	2.355	687	29,2	30,7	30,7
127 M - Insufficienza cardiaca e shock	4.081	386	9,5	17,2	47,9
207-208 M - Malattie delle vie biliari con/senza CC	1.894	315	16,6	14,1	62,0
294 M - Diabete, età >35	1.067	289	27,1	12,9	74,9
089 M - Polmonite semplice e pleurite, età >17 con CC	1.738	224	12,9	10,0	84,9
130-131 M - Malattie vascolari periferiche con/senza CC	1.260	110	8,7	4,9	89,9
300-301 M - Malattie endocrine con/senza CC	672	104	15,5	4,6	94,5
<i>sub-totale</i>		2.115		94,5	

A partire dai questi primi gruppi di DRG, sono state dunque analizzate le diagnosi principali (o meglio le categorie diagnostiche ad esse concettualmente associate), come risulta in Figura 6.

Figura 6 – Categoria diagnostica principale del 94,5% dei casi appropriati tardivi (ricoveri ordinari > 1 giorno; anno 2004). Fonte: Dossier ASR 132/2006.

Categoria diagnostica principale	Ricoveri appropriati tardivi		
	N	% sul totale	% cumulata
RN06 Calcoli delle vie urinarie	632	29,9	29,9
ND12 Diabete mellito	327	15,5	45,3
HB07 Colecistite e coleditiasi	313	14,8	60,1
CV02 Miocardiopatia	228	10,8	70,9
RS30 Polmonite batterica	210	9,9	80,9
ND15 Ipertiroidismo	95	4,5	85,3
RN77 Disturbi delle vie genito-urinarie, non altrimenti classificati	51	2,4	87,8
ND04 Ipotiroidismo	49	2,3	90,1
VS02 Malattie dell'arteria tibiale, iliaca, femorale, popliteale	42	2,0	92,1

A questo punto, allo scopo di proporre una riflessione più ampia, per ciascuna delle prime categorie diagnostiche sono stati analizzati i ricoveri relativi alla casistica generale, indipendentemente dal DRG generato. In tabella I vengono presentati i risultati.

Tabella I – Distribuzione dei ricoveri “appropriati tardivi” identificati nell’intera casistica di ricoveri ospedalieri > 1 giorno (anno 2004) secondo categoria diagnostica principale e gruppo di DRG di appartenenza, in ordine decrescente di volume totale. Fonte: Dossier ASR 132/2006.

Categoria diagnostica principale	Ricoveri Appropriati Tardivi (RAT)				
	DRG selezionati	Altri DRG (non inclusi nei 62 DRG)	% (DRG selezionati su Totale)	Totale	% (sui RAT)
Calcoli delle vie urinarie	632	498	55,4	1116	27,4
Colecistite e colelitiasi	313	609	33,9	922	22,6
Diabete Mellito	327	46	57,1	573	14,1
Polmonite batterica	210	302	41,0	512	12,6
Miocardiopatia	228	180	55,9	408	10,0
Malattie arteria tibiale, iliaca, femorale, poplitea	42	223	15,8	265	6,5
Ipertiroidismo	95	34	73,6	129	3,1
Ipotiroidismo	49	41	54,4	90	2,2
Disturbi vie G-U NAC	51	8	86,4	59	1,5
TOTALE (9 categorie diagnostiche)	1933	1941	47,4	4.074	100,0

Come si può osservare, passando ad esaminare l’intera casistica ospedaliera del 2004, il numero di ricoveri “appropriati tardivi” (riferiti alle principali categorie diagnostiche) si modifica da un totale di 1933 a 4074. La loro distribuzione tra i DRG ad elevato rischio di inappropriatezza e gli altri DRG è complessivamente, di poco inferiore al 50%, ma questa proporzione si modifica, in una direzione o nell’altra, per le singole categorie diagnostiche.

L’esame dei singoli DRG (che qui viene omessa per brevità) può offrire ovviamente ulteriori suggerimenti in merito al tipo di deficit assistenziale extra-ospedaliero eventualmente associato al ricovero.

Questo tipo di analisi, condotta in via preliminare su un solo anno di ricovero, merita senz’altro di essere ripetuta (anche con modalità campionaria) periodicamente ed eventualmente affinata sul piano metodologico. Il suggerimento che ne emerge riguarda soprattutto la possibilità

che i ricoveri “appropriati tardivi” costituiscano un’interessante e promettente categoria di prestazioni ospedaliere rispetto alle quali organizzare attività di revisione della documentazione clinica e/o di audit laddove si intenda indagare la qualità dell’assistenza offerta sul territorio per determinate patologie.

È altresì interessante constatare come alcune delle diagnosi relative a questi ricoveri richiamino le patologie tradizionalmente indagate dagli indicatori cosiddetti *Prevention* (secondo la definizione dell’*Agency for Healthcare Research and Quality – AHRQ* che li ha originariamente proposti), mirati a valutare, attraverso l’attività ospedaliera, il funzionamento dei servizi extra-ospedalieri (v. Figura 7).

Figura 7 – Elenco dei *Prevention Indicators* secondo l’*AHRQ*: indicatori ospedalieri “sensibili alle cure ambulatoriali”. Fonte: *AHRQ Guide to Prevention Quality Indicators*, 2007

PQI Number	Prevention Quality Indicators
1	Diabetes short-term complication admission rate
2	Perforated appendix admission rate
3	Diabetes long-term complication admission rate
5	Chronic obstructive pulmonary disease admission rate
7	Hypertension admission rate
8	Congestive heart failure admission rate
9	Low Birth Weight
10	Dehydration admission rate
11	Bacterial pneumonia admission rate
12	Urinary tract infection admission rate
13	Angina admission without procedure
14	Uncontrolled diabetes admission rate
15	Adult asthma admission rate
16	Rate of lower-extremity amputation among patients with diabetes

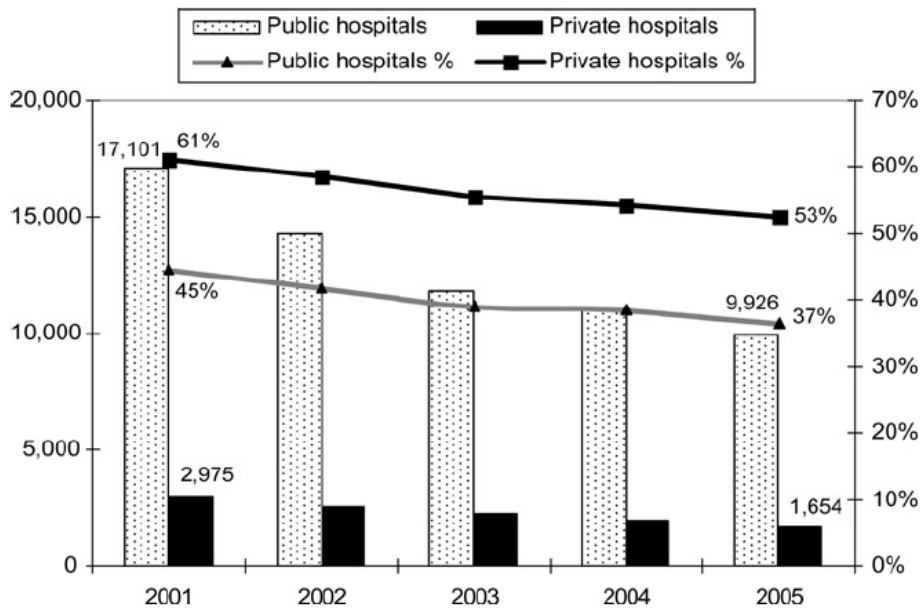
In conclusione, i principali messaggi offerti dall’Atlante dell’appropriatezza ospedaliera 2001-2004 sono i seguenti:

- riduzione, nel periodo esaminato, della quota di inappropriatezza ospedaliera
- persistenza di una maggiore quota di inappropriatezza per i DRG medici rispetto ai chirurgici
- incremento percentuale, nel periodo esaminato, dei ricoveri “appropriati tardivi”

Un ulteriore dato emerso riguarda la maggior quota di inappropriatezza e variabilità a carico delle strutture private (Figura 8)

Gli autori dell’Atlante invitano ad interpretare con cautela i risultati del lavoro, sia per i limiti notoriamente riconducibili all’affidabilità dello strumento SDO, sia per gli inevitabili limiti dell’analisi, che fotografa il solo episodio di cura ospedaliero, in mancanza di qualunque informazione specifica e di altra fonte sul percorso assistenziale che l’ha preceduto

Figura 8 – Numero e percentuale di ricoveri inappropriati, Regione Emilia-Romagna, anni 2001-2004, ospedali pubblici e privati. Fonte: Louis D et al, 2008.



Normativa

Decreto Legislativo 19 giugno 1999, n. 229. **Norme per la razionalizzazione del Servizio sanitario nazionale, a norma dell'articolo 1 della legge 30 novembre 1998, n. 419.** Pubblicato nella Gazzetta Ufficiale n. 165 del 16 luglio 1999 - Supplemento Ordinario n. 132.

Decreto del Presidente del Consiglio dei Ministri 29 novembre 2001. **Definizione dei livelli essenziali di assistenza.** Pubblicato nella Gazzetta Ufficiale Serie Generale n. 33 del 8 febbraio 2002.

Ministero della Salute. **Schema di Piano Sanitario Nazionale 2011-2013.** Anno 2011. Disponibile all'indirizzo: http://www.salute.gov.it/imgs/C_17_pubblicazioni_1454_allegato.pdf

Direzione generale della programmazione sanitaria, dei livelli di assistenza e dei principi etici di sistema–Ministero della Salute. **Programma Operativo Appropriatezza. Studio e sperimentazione di indicatori per determinare il grado di appropriatezza delle prestazioni di ricovero dovute a carenze delle altre forme di assistenza.** Febbraio 2011. Disponibile all'indirizzo: http://www.salute.gov.it/imgs/C_17_pubblicazioni_1492_allegato.pdf

Programma Nazionale Valutazione Esiti (PNE),II Ed. 2011, SDO 2005-2010

Bibliografia

- Agency for Healthcare Research and Quality, *Healthcare Quality Report*.
- Alakeson V. *Why Oregon went wrong*. British Medical Journal 2008;337:a2044. doi: 10.1136/bmj.a2044.
- Apolone G. et al, "L'appropriatezza dell'uso dell'ospedale: i risultati di uno studio all'ospedale S. Gerardo di Monza", Fed. Med. 1993; 8: 27-32.
- Baglio G, Materia E, Vantaggiato G, Perucci CA. *Valutare l'appropriatezza dei ricoveri con i dati amministrativi*. Tendenze Nuove 2001;1: 51-70; 2001.
- Balduzzi R. *L'appropriatezza in sanità: il quadro di riferimento legislativo*. In Falcitelli N., Trabucchi M., Vanara F. *Rapporto Sanità 2004. L'appropriatezza in sanità: uno strumento per migliorare la pratica clinica*. 2004. Il Mulino, Bologna.
- Brook RH, Chassin MR, Fink A. et al. *A method for the detailed assessment of the appropriateness of medical technologies*. Int j Technol Assess Health Care. 1986; 2:53-63. Park RE, Fink A, Brook RH, et al. *Physician ratings of appropriate indications for six medical and surgical procedures*. Am J Public Health. 1986;76:766-772
- Buetow SA, Sibbald B, Cantril JA et al. *Appropriateness in health care: applications to prescribing*. Soc Sci Med 1997;45:261-271.
- Burton MJ, Glasziou PP. *Tonsillectomy or adeno-tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis* (Cochrane Review), 2009.
- Chalmers I. *The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care*. Ann N Y Acad Sci 1993;31:703-63
- Chassin MR, Koseoff J, Park RE, et al. *Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures*. JAMA 1987; 258 (18): 2533-2537.
- Cinotti R. *et al. Glossario della qualità e dell'accreditamento*. 2005. Agenzia sanitaria e sociale regionale dell'Emilia-Romagna, Bologna.
- Dixon T, Shaw M, Ebrahim E, Dieppe P (2004). *Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need*. Annals of the Rheumatic Diseases, 63;7: 825-30.
- Donabedian A. *Il Maestro e le margherite. La qualità dell'assistenza sanitaria secondo Avedis Donabedian*. Edizione italiana. Roma: Il Pensiero Scientifico Editore, 2010.
- Doyle, J. *Returns to Local-Area Healthcare Spending: Using Shocks to Patients far from Home*. 2010. Consultabile all'indirizzo: http://www.mit.edu/~jjdoyle/doyle_vacation_feb2010.pdf.
- Evans RG. *The dog in the night-time: Medical practice variations and health policy*. in Andersen TF, Mooney G (eds), *The Challenges of Medical Practice Variations*, pp 117-152. London: Macmillan Press, 1990.
- Fellin G, Apolone G, Tampieri A, Bevilacqua, Meregalli G, Minella C, Liberati A. *Appropriateness of hospital use: an overview of italian studies*. Int J Qual HealthCare 1995;7(3):219-25

- Fisher ES, et al. *The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care.* Ann Intern Med. 2003; 138(4): pp. 273-87.
- Fisher ES, et al. *The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care.* Ann Intern Med. 2003; 138(4): pp. 288-98.
- Fitch K, Bernstein FJ, Aguilar MD, Burnand B, Lacalle JR, Lazaro P. *The RAND/UCLA Appropriateness Method User's Manual.* RAND Corporation, Sept. 2000
- France G., Taroni F., Donatini A. *The Italian health-care system.* Health Economics. 2005; 14: S187-202.
- Glover AJ. *The incidence of tonsillectomy in schoolchildren.* Proceedings of the Royal Society of Medicine. 2011;31(10):1219-1236.
- Ham C (Ed.). *Health Care Variations: Assessing the evidence.* London: King's Fund Institute, 1988.
- ISS-SNLG. *Appropriatezza e sicurezza degli interventi di tonsillectomia e/o adenoidectomia.* Documento 15, 2003, revisione 2008.
- Kahan JP, Loo M van het. *Defining appropriate health care.* Eurohealth 1999; 5(3): 16-18.
- Lavis JN, Anderson GM. *Appropriateness in health care delivery. Definitions, measurement and policy implications.* CMAJ. 1996;154:321-328.
- Materia E. *Appropriatezza: origini, implicazioni, valutazione.* Tendenze Nuove 2003;4-5:343-354
- Mulley AJ. *Improving productivity in the NHS.* Br Med J. 2010;341(7766):p c3965, 2010.
- Pronovost P.J., Goeschel C.A. *Time to Take Health Delivery Research Seriously.* JAMA. 20 luglio 2011; Vol. 306, No. 3: pp.301-11.
- Sanmartin C et al. *Appropriateness of healthcare interventions: concepts and scoping of the published literature.* Int J Technology Assessment in Health Care (2008)
- Sharpe VA, Faden AI. *Appropriateness in patient care: A new conceptual framework.* Milbank Q. 1996;74:115-138.
- Sharpe VA. *"The politics, economics and ethics of "appropriateness".* Kennedy Inst Ethics J. 1997;7:337-343.
- Suleman M, Clark MPA, Goldacre M, Burton M. *Exploring the variation in paediatric tonsillectomy rates between English regions: a 5-year NHS and independent sector data analysis.* Clinical Otolaryngology. 2010;35(2):111-17.
- Sutherland, J.M., E.S. Fisher, and J.S. Skinner. *Getting past denial--the high cost of health care in the United States.* N Engl J Med. 2009; 361(13): pp. 1227-30.
- The King's Fund. *Variations in Health Care. The good, the bad and the inexplicable,* 2011. www.kingsfund.org.uk/publications
- Tonelli S. *La qualità nelle organizzazioni sanitarie. Appunti per formazione e consulenza,* 2007.
- Vasselli S., Filippetti G., Spizzichino L. *Misurare la performance del sistema sanitario. Proposta di una metodologia.* 2005. Pensiero Scientifico Editore, Roma.
- Wennberg J. *Time to tackle unwarranted variations in practice.* Br Med J. 2011;342:687-690.

Ringraziamenti

Il presente manuale è stato realizzato sulla base del documento “*Appropriatezza Una guida pratica*” elaborato da: Stefania Rodella, Davide Botturi, Lucia Nobile (Agenzia Sanitaria e Sociale Regionale dell’Emilia-Romagna, Area Valutazione e sviluppo dell’assistenza e dei servizi), Vittorio Maio (Jefferson School of Population Health, Thomas Jefferson University, Philadelphia, PA, USA).

La revisione del documento ai fini della predisposizione del corso FAD è stata effettuata a cura di un gruppo di lavoro costituito presso il Ministero della salute, Direzione Generale della Programmazione Sanitaria, Ufficio III (Alessandro Ghirardini, Rosetta Cardone, Angela De Feo), in collaborazione con la Federazione Nazionale Ordini Medici Chirurghi e Odontoiatri – FNOMCeO – (Luigi Arru, Sergio Bovenga, Valerio Brucoli, Luigi Conte), la Federazione Nazionale Collegi Infermieri - IPASVI – (Giuseppina Grugnetti, Barbara Mangiacavalli, Emma Martellotti), con il contributo di Roberta Andrioli Stagno, Susanna Ciampalini, Antonietta Colonna, Daniela Furlan, Lucia Guidotti, Giorgio Leomporra, Claudio Seraschi (Ministero della Salute, Ufficio III).