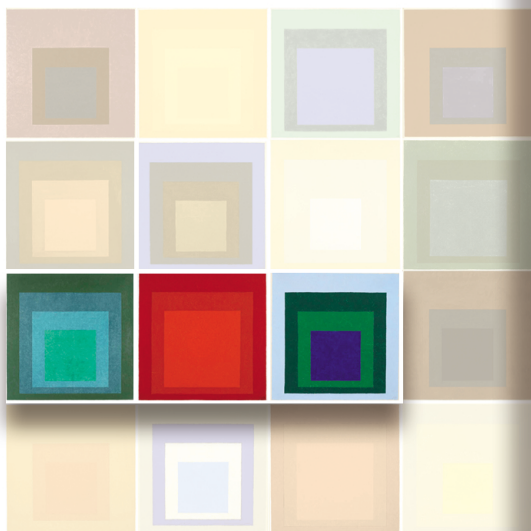




APPROPRIATEZZA. UNA GUIDA PRATICA

A CURA DI
STEFANIA RODELLA
DAVIDE BOTTURI



Ministero della Salute

Agenzia Sanitaria e Sociale Regionale della Regione Emilia-Romagna



Grazie a un finanziamento del Ministero della Salute
nell'ambito dell'Azione A del settore
"Programma nazionale per lo sviluppo del governo clinico
e della qualità del Servizio sanitario nazionale"
"Produzione di un manuale di riferimento per il governo clinico
e la qualità in forma integrata nonché la sperimentazione dello stesso"



**Dipartimento della programmazione e dell'ordinamento
del Servizio sanitario nazionale
DIREZIONE GENERALE DELLA PROGRAMMAZIONE SANITARIA
EX UFFICIO III**

In collaborazione con l'Azienda USL di Imola



**SERVIZIO SANITARIO REGIONALE
EMILIA-ROMAGNA**
Azienda Unità Sanitaria Locale di Imola

PRESENTAZIONE GENERALE

I contributi presentati in questi sette volumi hanno avuto storia e genesi diversa, ma sono accomunati dall'avere il ruolo dei professionisti nelle aziende sanitarie come oggetto di interesse. Alcuni sono stati elaborati nel contesto di un progetto finanziato dal Ministero della Salute e finalizzato alla produzione di un "manuale per il governo clinico", fornendo, nel contesto di quel progetto, i documenti di riferimento per la produzione di materiale utilizzato dalla Federazione Nazionale degli Ordini dei Medici e da Ipasvi in iniziative di formazione a distanza. Altri sono l'esito di specifici progetti di ricerca, alcuni finanziati dalla Regione Emilia-Romagna nel suo Programma di Ricerca Regione-Università, altri sviluppati nel contesto delle iniziative della Agenzia Sanitaria e Sociale di quella stessa Regione.

Tutti vedono, sia pure partendo da prospettive diverse, il ruolo dei professionisti nelle organizzazioni sanitarie come centrale, sia esaminando alcuni dei problemi principali su cui i professionisti dovrebbero poter dare il proprio contributo nell'esercizio delle loro responsabilità cliniche ed organizzative, sia indagando empiricamente con specifiche iniziative di ricerca alcuni dei molteplici aspetti che attengono alla partecipazione professionale, a quei processi decisionali che nelle aziende sanitarie riguardano la qualità dei servizi ed il loro sviluppo. In questo senso, questi contributi nel loro insieme intendono riproporre la questione del governo clinico cogliendone l'aspetto essenziale.

Probabilmente il riproporre oggi la questione del governo clinico merita una qualche giustificazione. Questo concetto si è rapidamente diffuso anche nel nostro contesto a partire dalla fine degli anni '90 per poi eclissarsi progressivamente.¹ Nel frattempo, ha avuto modo di vedersi assimilato rapidamente a tematiche limitrofe e pure, in qualche modo almeno, ad esso collegate, delle quali è diventato una sorta di elegante sinonimo. In questo modo, negli scorsi anni, parlare di governo clinico ha voluto dire parlare di audit clinico e *quality assurance*, di linee-guida e di *evidence-based medicine*, di *technology assessment*. In breve, ha significato parlare essenzialmente di strumenti e metodi. Di fatto, varie comunità specialistiche hanno “preso possesso” del governo clinico assimilandolo alle proprie specifiche metodologie ed al proprio ambito di interesse e quel termine ha finito per essere un modo diverso, ed almeno in quella fase più “a la page”, per parlare delle “solite cose”, piuttosto che rappresentare la ricerca di un modo diverso ed innovativo per affrontare i problemi di sempre (quelli della qualità e della sicurezza delle cure, dell'appropriatezza, di un cambiamento tecnologico, ma anche clinico ed organizzativo dentro le aziende da stimolare e governare).

In questo modo, come chi scrive ha già avuto modo in passato di segnalare,² si è finito per concentrare l'attenzione su strumenti e metodi (quelli della valutazione della qualità e della sicurezza delle cure, quelli della EBM o del *technology assessment*, a seconda dei casi), ma non sulle condizioni necessarie per fare penetrare efficacemente quegli strumenti nella quotidiana operatività delle organizzazioni sanitarie, tanto meno sulle condizioni necessarie per ottenere che il loro impiego fosse l'esito di un processo di acquisizione di responsabilità da parte dei professionisti clinici, rendendoli soggetto elaborante politiche aziendali a presidio della qualità dei servizi ed efficacemente propositivo relativamente allo sviluppo innovativo. Eppure, il senso del governo clinico e delle sue possibilità innovative stava esattamente nel cercare di cogliere il tema della partecipazione professionale allo sviluppo strategico delle organizzazioni sanitarie, nella ricerca di forme e modalità che stimolassero una diretta responsabilizzazione delle professionalità cliniche nella formulazione delle politiche aziendali direttamente rilevanti per la qualità delle cure, per indirizzare

lo sviluppo delle professionalità e le scelte di innovazione clinica, organizzativa e tecnologica di cui le organizzazioni sanitarie hanno costantemente bisogno.^{2 3}

Oggi, a più di un decennio di distanza dal momento in cui il governo clinico è stato introdotto nel dibattito italiano ed ha cominciato ad essere seriamente considerato dalle politiche, i problemi sembrano altri ed apparentemente molto distanti da queste questioni. Il SSN si trova a dover affrontare significativi problemi di sostenibilità e di riduzione delle risorse a cui si cerca di dare risposta anche con politiche finalizzate ad una robusta razionalizzazione nella struttura dell'offerta assistenziale, ad una revisione complessiva della rete dei servizi ospedalieri e territoriali. Tutte questioni assolutamente cogenti ed urgenti nella contingenza della crisi che investe il Paese e che riduce le risorse disponibili per la sanità, al punto da far apparire il tema della partecipazione professionale nelle aziende sanitarie come non particolarmente rilevante, alla luce dei problemi di sostenibilità dei servizi che il contesto odierno pone alla nostra attenzione.

Tuttavia, crediamo che anche in queste specifiche circostanze il tema della partecipazione professionale mantenga intatta, anzi per certi aspetti veda rafforzata, la propria attualità e rilevanza. Infatti, qualità delle cure ed innovazione sono, in estrema sintesi, le dimensioni della vita di un'organizzazione sanitaria su cui si esprime il governo clinico. Ora, se oggi la preoccupazione principale è rappresentata dalla sostenibilità dei servizi, possiamo ben dire che quelle dimensioni hanno davvero molto a che vedere con il fondamentale problema del mantenere, ed anzi per molti aspetti migliorare, i servizi sanitari in un contesto di risorse limitate. Queste ultime pongono la necessità di profondi e radicali cambiamenti finalizzati alla ricerca di forme più avanzate di razionalizzazione ed efficienza nel modo in cui sono affrontati bisogni assistenziali più complessi ed aspettative dei cittadini più elevate. Si tratta di cambiamenti che riguardano la riconfigurazione delle reti ospedaliere e territoriali, il ridisegno degli ospedali in funzione dell'intensità dei bisogni assistenziali dei pazienti, come pure l'elaborazione e l'implementazione di politiche di concentrazione delle casistiche chirurgiche in funzione dei volumi di attività dei punti di erogazione, per citare solo alcuni degli esem-

più possibili. Ebbene, tutti questi cambiamenti, per essere possibili ed avere gli effetti desiderati, hanno bisogno di essere supportati da un elevato livello di partecipazione professionale, pongono la necessità di acquisire la loro attiva partecipazione al ridisegno dei servizi che questi cambiamenti richiedono, ad esempio in termini di ridefinizione di percorsi assistenziali dei pazienti, di riformulazione dei ruoli e delle competenze professionali.

In questo senso, parlare di partecipazione professionale oggi significa parlare di uno dei determinanti della sostenibilità del SSN, almeno nella misura in cui la partecipazione professionale viene riconosciuta come uno degli ingredienti indispensabili per attivare e sviluppare quei cambiamenti richiesti da una migliore e più sostenibile configurazione delle reti dei servizi.

Non solo. La contingenza attuale finisce per rendere, per così dire, più vistosi i problemi di sempre, quelli che riguardano le concrete modalità di erogazione di servizi e prestazioni. Pensiamo ai temi della variabilità e inappropriatezza nel ricorso a procedure ed interventi nella pratica clinica, a quello di una maggiore sicurezza nei contesti clinico-assistenziali. In altri termini, il tema della qualità, l'esigenza di indirizzare i comportamenti professionali contrastando il ricorso ad interventi, laddove non siano clinicamente giustificati, sono oggi, in un quadro di riduzione delle risorse, se possibile ancora più acuti e di fatto mantengono come centrale la questione del come fare in modo che i comportamenti professionali, le decisioni e le scelte concretamente adottate dai clinici nelle modalità di assistenza siano pienamente coerenti con le finalità istituzionali di assicurare benefici di salute ai pazienti, attraverso interventi efficaci ed appropriati, in condizioni di sicurezza e di efficienza operativa.

Oggi, anche questi “vecchi problemi” sono alla ricerca di nuove possibili soluzioni, o quantomeno di strade innovative. Da questo punto di vista, una riflessione su come alcuni problemi – la variabilità della pratica clinica, l'inappropriatezza – siano stati sin qui affrontati e con quali risultati appare necessaria.^{4,5} In questo contesto, nel corso degli ultimi due decenni almeno è stata dispiegata una pluralità di approcci diversi, non necessariamente mutualmente esclusivi, che, riassumendoli in modo sin troppo schematico e sintetico, possiamo

dire essere stati rappresentati dal ricorso a forme di incentivazione economica, dalla comunicazione al pubblico di informazioni sulla qualità dei servizi nel tentativo di stimolare una maggiore attenzione alla qualità facendo leva sul patrimonio reputazionale degli operatori e delle organizzazioni o indirizzando le scelte dei cittadini. Da questo punto di vista è significativo che siano proprio quei sistemi sanitari che più intensamente e con maggiore determinazione sono ricorsi in questi anni agli approcci citati a manifestare oggi un esplicito interesse verso il ricorso al coinvolgimento ed alla partecipazione professionale, come ulteriore possibile percorso per costruire e consolidare nelle organizzazioni sanitarie un'attenzione sistematica e continuativa alla qualità, interesse che si esprime nell'attenzione elaborativa ed analitica rivolta al tema di quello che viene oggi definito come "clinical engagement".⁶⁻²⁰ Questo concetto si riferisce ad un attivo e positivo contributo fornito dai clinici nell'esercizio delle loro responsabilità al miglioramento delle performance delle organizzazioni, le quali a loro volta riconoscono questo impegno nel sostenere ed incoraggiare un'assistenza di alta qualità.¹¹ Rimanda quindi ad un'idea di partecipazione professionale fortemente propositiva verso i processi decisionali di un'organizzazione che attengono alla qualità ed allo sviluppo dei servizi, una partecipazione pienamente disponibile all'assunzione di responsabilità, che trova il proprio presupposto anche in uno spiccato senso di appartenenza all'organizzazione stessa.⁹ L'interesse per il tema del *clinical engagement*, che, come si vede, ha un'evidente affinità con un'idea di governo clinico fortemente orientata alla partecipazione professionale, trova la propria giustificazione nella consapevolezza che la varietà di strumenti e politiche sin qui adottata per controllare/orientare i comportamenti professionali (ricorrendo al monitoraggio attraverso indicatori, a motivazioni economiche, reputazionali, ecc.) ha inevitabili limiti di applicabilità, effetti collaterali negativi e risultati in generale non risolutivi. Sono quindi strumenti certamente impiegabili e persino imprescindibili oggi per molti aspetti, ma necessitano di un ulteriore ingrediente, per così dire, rappresentato appunto dalla capacità di acquisire la disponibilità delle professioni cliniche a riconoscere i temi della qualità come propri, come parte integrante del proprio ruolo dentro un'azienda, ed a mette-

re in gioco le proprie capacità e competenze al servizio dello sviluppo delle proprie organizzazioni.

In questo senso, quindi, parlare oggi di governo clinico può essere ancora decisamente attuale, a patto che il tema venga riproposto cogliendone l'essenza e spostando l'attenzione dagli strumenti ai contesti in cui tali strumenti dovrebbero essere impiegati, vedendolo quindi come problema non di strumenti e metodologie, magari da perfezionare e sviluppare ulteriormente, ma di *governance*, vale a dire di relazioni interne alle organizzazioni sanitarie, in particolare di relazioni tra clinici e manager o, se si preferisce, tra i clinici e le loro aziende. Diciamo questo evidentemente non per negare o sminuire il valore e le potenzialità di quegli strumenti, ma soltanto per sottolineare come quelle potenzialità rischino di restare largamente inesprese in assenza di una contestuale riflessione sulle aziende sanitarie, sulle loro concrete modalità di funzionamento per quanto riguarda mantenimento e miglioramento della qualità e della sicurezza delle cure, sul tipo di competenze che sono in grado di mettere in campo su questi temi, su come sono allocate le responsabilità per quanto attiene al *quality management* e sul ruolo dei clinici in questo contesto.

Una simile riflessione potrebbe probabilmente portarci a ragionare su alcuni significativi aspetti problematici. Pur tenendo conto della variabilità esistente anche su questi aspetti a livello nazionale, possiamo affermare che, in generale, le nostre aziende sanitarie sono caratterizzate dalla presenza di una significativa articolazione di tecnologie dedicate a presidiare ambiti definiti e relativamente circoscritti (la qualità e l'accreditamento, la formazione professionale, la gestione del rischio e, almeno in alcune realtà come quella dell'Emilia-Romagna, la ricerca e l'innovazione). Una almeno apparente, quindi, ricchezza di dotazione specialistica cui spesso non pare corrispondere un'effettiva efficacia operativa e, soprattutto, un'effettiva pervasività nei contesti clinico-assistenziali delle attività direttamente finalizzate a monitorare e migliorare la qualità dei servizi. In larga misura, qualità e sicurezza delle cure, nelle loro diverse sfaccettature, permangono come temi di interesse delle specifiche tecnologie di riferimento, con un coinvolgimento dei clinici relativamente limitato, per intensità e continuità, e generalmente passivo.

Questa tipologia di assetto aziendale sul *quality management* rende particolarmente pertinente la constatazione che una delle principali contraddizioni che caratterizzano i sistemi sanitari moderni è rappresentata dall'essere contesti spesso avanzatissimi per quanto riguarda la dotazione tecnologica disponibile, ma assolutamente antiquati per quanto attiene invece ai loro assetti organizzativi e di management.²¹ Infatti, non soltanto la parcellizzazione delle tecnostrutture sopra citata rimanda alla necessità di una capacità di sintesi e di *governance* complessiva – generalmente affidata alla direzione sanitaria aziendale – ma risulta fortemente evocatrice di assetti tipici della tradizionale impresa fordista, una delle caratteristiche della quale è rappresentata dalla presenza di una netta separazione tra la linea produttiva e gli ambiti che su tale linea esercitano funzioni di sviluppo, verifica e controllo.

Tuttavia, questo assetto è da tempo già stato messo in discussione e superato nel mondo delle imprese, in particolare laddove l'impresa si confronta con processi produttivi poco o nulla standardizzabili, per la cui realizzazione è indispensabile l'apporto di competenze – anche sofisticate – da parte degli operatori e che richiedono un'azione di costante sorveglianza e manutenzione, esercitata non da soggetti collocati al di fuori del processo, ma direttamente dagli operatori “in linea”.²² Queste caratteristiche generali che hanno portato al superamento del tradizionale modello fordista sono esattamente alcuni dei connotati che caratterizzano la peculiarità delle organizzazioni sanitarie, sotto il profilo della complessità e dell'elevato grado di specializzazione delle competenze degli operatori, dell'elevato grado di flessibilità dei processi di cura, della necessità di una continua azione di sorveglianza, “non delegabile”, almeno non completamente, a soggetti esterni alla linea produttiva. Tuttavia, questo è proprio quel che in larga misura è accaduto nei nostri contesti assistenziali, con la “delega” ad occuparsi di qualità, formazione, gestione del rischio, ricerca&innovazione, attribuita a tecnici, relativamente “distanti” dagli ambiti di produzione, vale a dire dai dipartimenti. Queste considerazioni peraltro acquisiscono un ulteriore significato alla luce delle evidenze empiriche che segnalano come, nel nostro contesto nazionale, siano ancora largamente prevalenti nelle aziende sanita-

rie culture organizzative di tipo gerarchico ascrivibili a quel tipo di paradigma produttivo.²³ Il governo clinico, così come l'abbiamo sin qui conosciuto, non è di fatto riuscito a sanare questa contraddizione tra caratteristiche dei processi produttivi e rigidità delle modalità di management adottate.

Ora, se il governo clinico viene visto a partire dalle premesse che ci siamo dati, non si tratta solo di fornire istruzioni sull'uso di tecniche, ma di rendere possibile e praticabile una responsabile partecipazione professionale. Quest'ultima non è una tecnica, codificabile e trasmissibile con un manuale. È il risultato di relazioni che devono essere costruite e di ruoli e responsabilità che devono essere esercitati. Per questo i contributi di questa collana non si limitano ad illustrare tecniche e metodologie, ma si sforzano di affrontare problemi ed esaminare alcuni temi centrali per le politiche di sviluppo dei servizi ed il miglioramento della qualità e della sicurezza delle cure.

Detto questo, va anche detto che se il tema è il perseguimento di una responsabile partecipazione dei professionisti clinici ai processi aziendali di mantenimento e sviluppo della qualità dei servizi, questi volumi sono di certo vistosamente insufficienti allo scopo. La partecipazione richiede una molteplicità di condizioni per potersi realizzare, il cui determinarsi è evidentemente al di fuori delle possibilità di semplici contributi scritti. Tuttavia, la partecipazione professionale ha anche la necessità di avere come premessa un sufficiente grado di condivisione e consapevolezza circa le questioni potenzialmente rilevanti per lo sviluppo di un'organizzazione sanitaria ed i principi ed i valori generali con cui tali questioni dovrebbero poter essere affrontate. È a questo livello che si posizionano i contributi di questa raccolta, con l'auspicio che possano essere di una qualche utilità, immaginando il loro impiego anche nel contesto di iniziative di formazione e discussione da avviare nelle aziende sanitarie, come pure come materiale didattico a supporto di iniziative formative rivolte agli studenti di medicina e ai giovani medici in formazione.

BIBLIOGRAFIA

1. Walshe K. Pseudoinnovation: the development and spread of healthcare quality improvement methodologies. *Int J Qual Health Care* 2009; 21: 153-9. [Link alla risorsa]
2. Grilli R, Taroni F. Governo clinico. Governo delle organizzazioni sanitarie e qualità dell'assistenza. Roma: Il Pensiero Scientifico Editore, 2004.
3. Grilli R. Le opportunità offerte dal governo clinico. Il difficile rapporto tra discrezionalità clinica ed autorità manageriale. *Salute e Territorio* 2005; 152: 306-12. [Link alla risorsa]
4. Moynihan R, Heneghan C, Godlee F. Too much medicine: from evidence to action. *BMJ* 2013; 347: 7141. [Link alla risorsa]
5. Appleby J, Raleigh V, Frosini F, Bevan G, Gao H, Lyscom T. Variations in health care: the good, the bad and the inexplicable. London: The King's Fund, 2011. [Link alla risorsa]
6. Bethune R, Soo E, Woodhead P, Van Hamel C, Watson J. Engaging all doctors in continuous quality improvement: a structured, supported programme for first-year doctors across a training deanery in England. *BMJ Qual Saf* 2013; 22: 613-7. [Link alla risorsa]
7. Denis JL, Baker GR, Black C et al. Exploring the dynamics of physician engagement and leadership for health system improvement. Prospects for Canadian healthcare systems. Ottawa: Canadian Institute of Health Research, 2013. [Link alla risorsa]
8. Brown B, Ahmed-Little Y, Stanton E. Why we cannot afford not to engage junior doctors in NHS leadership. *J R Soc Med* 2012; 105: 105-10.
9. Clark J. Medical engagement. Too important to be left to chance. London: The King's Fund, 2012. [Link alla risorsa]
10. The King's Fund. Leadership and engagement for improvement in the NHS: together we can. London: The King's Fund, 2012. [Link alla risorsa]
11. Spurgeon P, Mazelan PM, Barwell F. Medical engagement: a crucial underpinning to organisational performance. *Health Serv Manage Res* 2011; 24: 114-20.
12. Taitz JM, Lee TH, Sequist TD. A framework for engaging physicians in quality and safety. *BMJ Qual Saf* 2012; 21: 722-8. [Link alla risorsa]
13. The Health Foundation. Are clinicians engaged in quality improvement? A review of the literature on healthcare professionals' views on quality

- improvement initiatives. London: The Health Foundation, 2011. [Link alla risorsa]
14. Parand A, Burnett S, Benn J, Iskander S, Pinto A, Vincent C. Medical engagement in organisation-wide safety and quality-improvement programmes: experience in the UK Safer Patients Initiative. *Qual Saf Health Care* 2010; 19: e44. [Link alla risorsa]
 15. Macleod D, Clarke N. Engaging for success: enhancing performance through employee engagement. London: Department for Business, Innovation and Skills, 2009. [Link alla risorsa]
 16. Conway J. Getting boards on board: engaging governing boards in quality and safety. *Jt Comm J Qual Patient Saf* 2008; 34(4): 214-20. [Link alla risorsa]
 17. Dickinson H, Ham C. Engaging doctors in leadership: review of the literature. Academy of Medical Royal Colleges, University of Birmingham, Institute for Innovation and Improvement, 2008. [Link alla risorsa]
 18. Spurgeon P, Barwell F, Mazelan P. Developing a medical engagement scale (MES). *International Journal of Clinical Leadership* 2008; 16: 213-23.
 19. Reinertsen JL, Gosfield AG, Rupp W, Whittington JW. Engaging physicians in a shared quality agenda. IHI Innovation Series White Paper. Cambridge, MA: Institute for Healthcare Improvement, 2007. [Link alla risorsa]
 20. Neale G, Vincent C, Darzi SA. The problem of engaging hospital doctors in promoting safety and quality in clinical care. *J R Soc Promot Health* 2007; 127:87-94. [Link alla risorsa]
 21. Doctor innovation. Shaking up the health system. London: The Economist Intelligence Unit Limited, 2009. [Link alla risorsa]
 22. Fumagalli A. Bioeconomia e capitalismo cognitivo. Roma: Carocci Editore, 2007.
 23. Calciolari S, Prenestini A, Lega F, Grilli R. Possono i “valori” fare la differenza? Evidenze sull’influenza della cultura organizzativa sul governo clinico. *Mecosan* 2012; 21(82): 9-24.

CONTRIBUTI PER IL GOVERNO CLINICO

Collana a cura di Roberto Grilli

**A cura di
Stefania
Rodella
Davide
Botturi**

APPROPRIATEZZA

Una guida pratica

Volume 3



Il Pensiero Scientifico Editore

Stefania Rodella

Area Valutazione e sviluppo dell'assistenza e dei servizi
Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna

Davide Botturi

Area Valutazione e sviluppo dell'assistenza e dei servizi
Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna

*Contributi di***Luciana Ballini**

Osservatorio Regionale per l'Innovazione (ORI)
Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna

Vittorio Maio

Jefferson School of Population Health
Thomas Jefferson University, Philadelphia (PA) USA

Lucia Nobilio

Area Valutazione e sviluppo dell'assistenza e dei servizi
Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna

Luca Vignatelli

Osservatorio Regionale per l'Innovazione (ORI)
Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna

Progetto editoriale realizzato da Il Pensiero Scientifico Editore
per conto dell'Agenzia Sanitaria e Sociale Regionale
della Regione Emilia-Romagna,
grazie a un finanziamento del Ministero della Salute
e in collaborazione con l'Azienda USL di Imola

Prima edizione: dicembre 2014
© 2014 Agenzia Sanitaria e Sociale Regionale,
Ministero della Salute e Azienda USL di Imola

Prodotto editoriale liberamente fruibile per fini non commerciali;
chiunque è autorizzato per fini informativi, di studio o didattici,
a utilizzare e duplicare i contenuti di questa pubblicazione,
purché ne venga citata la fonte

Progetto editoriale: Il Pensiero Scientifico Editore,
Via San Giovanni Valdarno 8, 00138 Roma
Tel. (+39) 06 862821 - Fax: (+39) 06 86282250
pensiero@pensiero.it - www.pensiero.it
www.facebook.com/PensieroScientifico

Impaginazione: Doppiosegno, Roma
Progetto grafico: Studio Rosa Pantone, Roma
Illustrazione di copertina: Josef Albers, *Omaggio al Quadrato*
Coordinamento redazionale: Silvana Guida

ISBN 978-88-490-0511-0

INDICE

INTRODUZIONE	V
<i>P a r t e p r i m a</i> DALLA DEFINIZIONE ALLA VALUTAZIONE	1
1 • BACKGROUND, DEFINIZIONI, MODELLI CONCETTUALI <i>Stefania Rodella, Davide Botturi</i>	3
2 • VARIAZIONI DELLA PRATICA CLINICA: CONCETTI, METODI ED ESEMPI <i>Stefania Rodella, Davide Botturi</i>	19
3 • LA MISURA DELL'APPROPRIATEZZA <i>Stefania Rodella, Davide Botturi</i>	39
<i>P a r t e s e c o n d a</i> APPROFONDIMENTO	63
4 • THE NHS ATLAS OF VARIATION IN HEALTHCARE: IL PROGRAMMA RIGHT CARE DEL DEPARTMENT OF HEALTH (UK) <i>Davide Botturi</i>	65
5 • THE DARTMOUTH ATLAS PROJECT <i>Davide Botturi</i>	79
6 • L'APPROPRIATEZZA D'USO DELLE TECNOLOGIE SANITARIE: IL CASO DELLA FDG-PET IN AMBITO ONCOLOGICO <i>Luciana Ballini, Luca Vignatelli</i>	93
7 • L'APPROPRIATEZZA PRESCRITTIVA NELLA POPOLAZIONE ANZIANA <i>Vittorio Maio, Davide Botturi</i>	103

8 • I RICOVERI OSPEDALIERI IN EMILIA-ROMAGNA 2001-2008. ATLANTE DELL'APPROPRIATEZZA ORGANIZZATIVA	115
<i>Lucia Nobile</i>	
9 • L'INIZIATIVA AMERICANA CHOOSING WISELY E IL PROGETTO ITALIANO	133
<i>Luciana Ballini, Davide Botturi</i>	
APPENDICE NORMATIVA	143
<i>Davide Botturi</i>	
BIBLIOGRAFIA E SITOGRAFIA	155

INTRODUZIONE

Questo studio si propone di riassumere e organizzare le principali conoscenze a 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 volume si articola in due parti. 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, ecc.; esplora il problema della variabilità (geografica, organizzativa) frequentemente osservata nelle attività clinico-assistenziali e nei comportamenti professionali e la sua relazione (ancora irrisolta) 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; sviluppa infine il tema della misura dell'appropriatezza, con riferimento prevalente alla metodologia messa a punto dalla RAND.

La seconda parte, di approfondimento, propone, in alcuni schematici capitoli, schede esemplificative tratte dalla letteratura internazionale e da esperienze condotte in Italia.

Conclude il lavoro un'appendice costituita da un elenco selezionato e ragionato di riferimenti normativi – nazionali e regionali – che

negli ultimi venti anni hanno introdotto, intorno al concetto di appropriatezza, proposte e azioni di politica sanitaria.

Il testo è corredato da una bibliografia-sitografia essenziale.

P a r t e p r i m a

DALLA DEFINIZIONE ALLA VALUTAZIONE

1 • BACKGROUND, DEFINIZIONI, MODELLI CONCETTUALI

Stefania Rodella, Davide Botturi

Questo primo capitolo 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, ecc.

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

IL CONCETTO DI APPROPRIATEZZA

Secondo il dizionario della lingua italiana (Zingarelli, 2012) 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”, come si legge nel dizionario Garzanti (Stoppelli, 2009), 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 ai quali un paziente va incontro quando accede a una, o più, 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 (Sanmartin et al., 2008).

L'appropriatezza quindi porta con sé implicazioni dirette e indirette per la procedura corretta, per la persona giusta, al momento opportuno e nel setting più adatto. Allo stesso tempo, l'attenzione all'appropriatezza solleva la domanda "in quale prospettiva?" (paziente, medico, *policy-maker*, ecc.).

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.

L'APPROPRIATEZZA DEGLI INTERVENTI SOCIO-SANITARI: DEFINIZIONI

Una revisione di letteratura pubblicata nel 2008 ed estesa a un periodo di osservazione molto ampio, compreso tra il 1966 e il 2006 (Sanmartin et al., 2008), 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 (Research

ANd Development), che hanno introdotto i concetti di “rischi” e “benefici”.

Secondo questa prospettiva, si definisce appropriata una procedura in cui il beneficio atteso (ad esempio, un aumento della *life expectancy*, il sollievo dal dolore, la riduzione dell’ansia, il miglioramento della capacità funzionale) supera le (eventuali) conseguenze negative attese (ad esempio, mortalità, morbidità, ansietà, dolore, tempo lavorativo perso) con un margine sufficientemente ampio, tale da ritenere che valga la pena effettuarla (cfr. Brook et al., 1986; Park et al., 1986). Questa definizione si applica a un paziente “medio” che si presenta ad un medico “medio”. Al contrario, viene considerata inappropriata una procedura il cui rischio sia superiore ai benefici (attesi). Lo stesso Robert H. Brook¹ ha enfatizzato la necessità di disporre di uno strumento capace di misurare l’appropriatezza delle cure, motivandola 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. È importante considerare come, al momento in cui queste definizioni vengono proposte (siamo tra gli anni ’80 e gli anni ’90), il problema dei costi non sia ancora tenuto in considerazione.

La definizione proposta dalla RAND è, ad oggi, la più ampiamente usata a livello internazionale.

Con il passare del tempo, la variabile “costi” entra però a far parte del concetto di appropriatezza per l’urgenza di porre un contenimento al peso crescente dei costi in sanità, particolarmente negli Stati Uniti; questo cambiamento di direzione è stato ben descritto da Sharpe (1997). Prima degli anni ’80, negli Stati Uniti, l’interpretazione del termine “appropriato” era sostanzialmente condivisa, esente da controversie, e stava a significare l’insieme di cure in grado di meglio rispondere ai bisogni del singolo paziente. In seguito, in un contesto economico in cui la precedente percezione di abbondanza si trasformò gradualmente in una consapevolezza di scarsità/limitatezza, emerse la

¹ Robert Brook, docente di Health Care Services presso la RAND Corporation, dove ha prestato servizio per 19 anni come vicepresidente e direttore di RAND Health, è uno dei nomi più conosciuti, a livello internazionale, nello studio della qualità dei servizi sanitari (<https://www.rand.org/search.html?query=Robert%20Brook>)

necessità, per i *third-payers* (i finanziatori pubblici e privati delle cure sanitarie) di porre un freno al rimborso dei servizi e delle prestazioni fornite. Ecco dunque che, da un lato l'appropriatezza venne collocata nel contesto di una relazione costo-efficacia, dall'altro l'offerta di servizi fu inquadrata in una rete di collegamenti condizionati da risorse limitate. La definizione di appropriatezza divenne quindi più complessa e precisa, configurandosi 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" (Buetow et al., 1997).

Altri autori sostengono che, nonostante gli sforzi compiuti per definire l'appropriatezza in termini di rapporto tra rischi e benefici, il concetto continua a restare elusivo e ambiguo, difficile da chiarire. A seguito di una breve review delle evidenze disponibili riguardo ai metodi 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" (Naylor, 1998).

In Europa il concetto di appropriatezza delle cure compare in documenti istituzionali nel 1997 (cfr. Balduzzi, 2004). Nella Raccomandazione n. 17 del Comitato dei Ministri del Consiglio d'Europa (1997) essa viene elencata tra "le componenti potenziali della qualità dell'assistenza sanitaria" e nella Convenzione di Oviedo sui Diritti dell'Uomo e la biomedicina (Consiglio d'Europa, 1997) gli Stati della Comunità Europea (e anche altri Stati non ad essa appartenenti) firmatari della Convenzione si impegnano a garantire un "accesso equo a cure della salute di qualità appropriata" (art. 3).

Nel 2000 l'Ufficio regionale europeo della World Health Organization (WHO) organizza un workshop, in collaborazione con il Ministero della Salute tedesco, proprio dedicato allo scambio di informazioni e conoscenze, a livello internazionale e nazionale, sul significato del termine appropriatezza nell'assistenza sanitaria, sulle metodologie

per identificare i trattamenti appropriati e non appropriati e sulle strategie per aumentare la condivisione di interventi e trattamenti appropriati (Regional Office for Europe, WHO, 2000). Nel rapporto del workshop si riconosce la definizione di appropriatezza proposta dalla RAND Corporation come un utile punto di partenza per ragionare intorno al significato di tale concetto, sebbene ne vengano riconosciute la genericità e i limiti.

Questi ultimi vengono ricondotti alla mancata considerazione della “individualità del paziente” e “della disponibilità di risorse per l’assistenza sanitaria”, secondo un report inglese preparato per il Director of Research and Development of the Department of Health (Hopkins et al., 1993). Il gruppo di lavoro propone la seguente definizione (che viene qui tradotta liberamente e riportata per intero):

Una cura appropriata consiste nella selezione, tra gli 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 (LEA).² In particolare, l’attenzione dei *policy-makers* italiani si è focalizzata sull’appropriatezza organizzativa, *in primis*, dell’assistenza ospedaliera, di cui le liste di DRG (Diagnosis Related Group) a rischio di inappropriatezza³ rappresentano una delle principali traduzioni operative. All’appropriatezza clinica è stato dedicato uno specifico progetto, il Sistema nazionale per le linee-guida, lanciato dal Piano Sanitario Nazionale 1998-2000 e coordinato dall’Istituto Superiore di Sanità.

Anche il Piano Sanitario Nazionale 2011-2013 (bozza pubblicata il 21 gennaio 2011), dedica attenzione al concetto di appropriatezza (cfr. *Appendice normativa*).

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 e dell’Agenzia nazionale per i servizi sanitari:⁴ “L’appropriatezza definisce un intervento 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 qui di seguito le principali definizioni proposte nella letteratura e nella manualistica italiana.

- 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, 2007).

² Cfr. D.Lgs. 229/1999 nell’*Appendice normativa* a questo volume. Cfr. anche France, Taroni e Donatini (2005).

³ Si consulti a tal riguardo l’elenco dei DRG ad alto rischio di inappropriatezza presente nel *Patto per la Salute 2010-2012*, approvato nella seduta del 3 dicembre 2009 dalla Conferenza permanente per i rapporti tra Stato, Regioni e Province Autonome di Trento e Bolzano.

⁴ http://ape.agenas.it/documenti/6.glossario_ecm.pdf

- [L'appropriatezza] 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, o specifica, si gioca nel processo decisionale e chiama in causa dunque il giudizio clinico (Morosini e Perraro, 2001).
- L'appropriatezza organizzativa o generica è riferita al livello di effettuazione dell'intervento (ad esempio, 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 e Perraro, 2001).
- L'appropriatezza è una 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 et al., 1996).
- 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 et al., 2004).⁵
- 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, includendo, ad esempio, dimensioni come la qualità di vita. Inoltre,

⁵ La definizione di appropriatezza è stata adattata dalla Joint Commission on Accreditation of Healthcare Organizations nella *Guide to Quality Assurance*.

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 et al., 2005).

- 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, 2005).

Nella tabella 1.1 vengono presentate schematicamente le principali definizioni di appropriatezza proposte negli Stati Uniti, in Europa e in Italia.

APPROPRIATEZZA E ALTRI CONCETTI CORRELATI

Data la complessità e la multidimensionalità del concetto di appropriatezza, molti sono, comprensibilmente, i termini ad essa correlati, come ad esempio: efficacia, efficienza, equità, necessità (clinica), variabilità geografica della pratica clinica (espressa anche dall'ambito di indagini conosciute come *small-area variation*).

Per una trattazione estesa del concetto di *efficacia*, alla quale è stata (e continua ad essere) dedicata una mole consistente di letteratura, si rimanda a fonti specifiche;⁶ in questa sede è sufficiente menzionare una definizione di “efficacia nella pratica”, espressa come “grado in cui i miglioramenti di salute raggiungibili sono, in effetti, raggiunti” e associata a quella di “efficacia teorica”, espressa come “l'abilità della scienza e della tecnologia sanitaria di produrre miglioramenti nella salute quando usate nelle condizioni più favorevoli” (Donabedian, 2010).

⁶ Per una trattazione “classica” del concetto di efficacia, cfr. l'opera di Archibald L. Cochrane (1999).

TABELLA 1.1 - DEFINIZIONI DI APPROPRIATEZZA PROPOSTE DALLA LETTERATURA

STUDIO/FONTE	PAESE	ELEMENTI CHIAVE DELLA DEFINIZIONE	PROSPETTIVA
Woodward et al. (1984)	USA	Effetti positivi attesi per il paziente	Paziente
RAND (varie fonti)	USA, WHO Europe	Benefici attesi superiori alle possibili conseguenze negative	Paziente
Sharpe e Faden (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
Buetow et al. (1997)	USA	Massimizzazione del beneficio netto per la salute individuale nell'ambito delle risorse disponibili	Paziente Società
Kelley, Hurst (2006)	OECD	Rilevanza rispetto ai bisogni clinici	Paziente
Hopkins et al. (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
Ministero Salute (cfr. Appendice normativa e Progetto Mattoni, 2007)	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

Segue

TABELLA 1.1 - SEQUE

STUDIO/FONTE	PAESE	ELEMENTI CHIAVE DELLA DEFINIZIONE	PROSPETTIVA
Zanetti et al. (1996)	Italia	Accettabilità e pertinenza rispetto a persone, circostanze e luogo, conoscenza	Paziente Sistema sanitario Società
Morosini e Perraro (2001)	Italia	Benefici attesi superiori alle conseguenze negative Deve tenere conto dei costi Inerente al 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
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 variabili non strettamente cliniche (qualità di vita) Erogazione dell'intervento nel momento giusto e secondo il regime organizzativo adeguato Deve tenere conto sia del contesto culturale e organizzativo del paziente e dell'operatore sia delle risorse	Paziente Operatore Sistema sanitario Società
Cinotti (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
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

In quanto includono la dimensione “efficacia”, anche altre dimensioni della qualità dell’assistenza, quali l’*efficienza* e l’*equità* sono senz’altro correlate a quella di appropriatezza; ma anche per questi concetti, non strettamente oggetto della presente pubblicazione, si rimanda a fonti specifiche (Donabedian, 2010; cfr capitoli 4 e 6), almeno per un inquadramento generale.

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 (Berghthold, 1995; Caulfield, 1996; Jacobson et al., 1997). 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” (Charles et al. 1997).

Analogamente, il concetto di appropriatezza è stato collegato a quello di *servizi necessari* o *critici*, intesi come quelli destinati a un “paziente che si presenta con determinate indicazioni, tali da suggerire che il medico sia obbligato a raccomandare un certo trattamento, nella convinzione che il non farlo costituirebbe una scelta deleteria per la salute del paziente” (Kahan et al., 1994). Ma è possibile definire i criteri in base ai quali un servizio o una prestazione possano essere considerati “critici” o “necessari”? Eccone alcuni, sempre proposti dalla letteratura americana e dalla lettura dei quali risulta evidente come il concetto di “necessità” rappresenti un’estensione (o una specificazione) del concetto di “appropriatezza”:

- la procedura in questione è appropriata (secondo il metodo RAND);
- sarebbe scorretto non raccomandarla;
- esiste una ragionevole probabilità che il paziente possa trarne beneficio;
- il beneficio per il paziente non è piccolo.

In definitiva, si può dire che, laddove l’appropriatezza rimanda logicamente a criteri di esclusione (orientando l’attenzione verso il non effettuare prestazioni inappropriate), la necessità è meglio riconducibile a criteri di inclusione (raccomandando di effettuare la procedura

se ritenuta critica). È intuitiva la connessione tra questo concetto anglosassone di “necessità” e quello introdotto da alcuni anni nel nostro Paese e conosciuto come “livelli essenziali di assistenza” (cfr. oltre, *Appendice normativa*).

Alla *variabilità della pratica clinico-assistenziale*, non solo geografica (ad esempio, tra le aree di residenza di determinate popolazioni) ma anche organizzativa (ad esempio, tra ospedali o tra unità cliniche) e professionale (ad esempio, tra professionisti pur appartenenti alla stessa specialità), viene dedicato il capitolo 2. Qui basti ricordare come la constatazione di un’elevata variabilità nel ricorso a una gamma estesa di procedure mediche e chirurgiche piuttosto comuni (come, ad esempio, l’angioplastica coronarica, la sostituzione protesica dell’anca, la tonsillectomia, ecc.) sia stata (e sia tuttora) oggetto di numerosi studi e pubblicazioni. Una delle ipotesi per spiegare questa osservazione è stata formulata intorno al possibile sovrautilizzo di procedure inappropriate nelle aree con più elevata frequenza. Tuttavia, nonostante la consistente mole di ricerche al riguardo, si può dire che l’evidenza ad oggi disponibile indichi la mancanza di una chiara e riproducibile relazione tra *small-area variation* e appropriatezza. Le differenze riscontrate nel ricorso ai servizi possono essere, al contrario, il risultato di diverse combinazioni di fattori: le caratteristiche dei pazienti (socio-demografiche, 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 (cfr. Wennberg, 2011).

MODELLI CONCETTUALI E COMPONENTI DELL’APPROPRIATEZZA

La revisione di letteratura di Sanmartin et al. (2008) ha evidenziato due fondamentali *frameworks* concettuali inerenti all’appropriatezza. Il primo è stato introdotto da Lavis e Anderson (1996), a fine anni ’90, con il duplice concetto di *appropriateness of a service*, o appropriatezza clinica, e *appropriateness of setting*, solitamente espressa in italiano come 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, Lavis e Anderson 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 indicazioni". I costi, in questo caso, non sono ancora considerati.

L'appropriatezza organizzativa viene chiamata in gioco quando si intenda valutare "se il tipo di assistenza (teoricamente) richiesta dalle caratteristiche cliniche del paziente corrisponda al tipo di assistenza concretamente offerta". In questo caso, il setting è usato come *proxy* dell'entità di risorse utilizzate per fornire assistenza; di conseguenza, l'appropriatezza organizzativa prende in considerazione la relazione costi-efficacia.

Un secondo e più ampio modello concettuale, riferito in particolare alle prestazioni chirurgiche, è stato introdotto nello stesso periodo da Sharpe e Faden (1996), 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 sottolinearono 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.⁷ In termini di *over-use* e *under-use* sono stati espressi i risultati di numerosi studi e rapporti pubblicati dalla RAND. A titolo di esempio citiamo in questa sede uno studio del 1999 dedicato ai test diagnostici per la malattia coronarica,⁸ uno studio del 2000 dedicato alla colonscopia in Europa⁹ e uno studio del 2001 che affronta i problemi metodologici nella rilevazione dell'*over-use* di rivascolarizza-

⁷ Vale la pena ricordare un terzo concetto correlato, il *mis-use*, che si colloca sostanzialmente nell'ambito dell'errore medico e della gestione del rischio.

⁸ Pubblicato su *American Journal of Medicine* (106, 4, Apr. 1999, pp. 391-8).

⁹ Pubblicato su *Gastrointestinal Endoscopy* (52, 5, Nov. 2000, pp. 593-9).

zione coronarica e isterectomia.¹⁰ Per quanto riguarda il fenomeno dell'*over-use* l'esperienza più importante degli ultimi anni consiste nell'iniziativa americana "Choosing Wisely" e nell'analogo progetto avviato in Italia (cfr. più avanti, il capitolo 9).

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 (1989), 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 almeno un RCT o da una revisione sistematica o da altri studi di buona qualità, il livello di evidenza è inferiore, ma è ancora elevato e i corrispondenti interventi sanitari sono raccomandati (o controindicati). Infine, il più basso livello di evidenza è offerto dalla sola disponibilità di case report o studi clinici non controllati e può costituire l'unico riferimento per interventi potenzialmente benefici (o dannosi) o il cui equilibrio tra benefici e rischi sia incerto.

Una seconda chiave di lettura è riconducibile alla *prospettiva dei pazienti*, i quali possono contribuire alla determinazione dell'appropriatezza sia quando vengano considerati come "paziente medio" sia quando venga prestata attenzione al "singolo paziente". Nel primo caso le informazioni empiriche sugli outcome di cura – risultati clinici, misure di qualità della vita o gradi di soddisfazione, stato di salute percepito o riferito¹¹ – contribuiscono alle stime di efficacia dei diversi studi associati ai livelli di evidenza già menzionati.

Nel secondo caso entra in gioco l'accettabilità della cura per la quale sono stati dimostrati determinati outcome; l'accettabilità è una dimensione definibile come "conformità ad aspirazioni, desideri e attese dei pazienti e dei membri responsabili delle loro famiglie" (Do-

¹⁰ Pubblicato su *Journal of Clinical Epidemiology* (54, 10, Oct. 2001, pp. 1004-10).

¹¹ È il caso dei *patient reported outcomes* (cfr. più avanti, capitolo 2, nota 7).

nabedian, 2010, p. 27); in altre parole, gli outcome clinici legati a un particolare intervento devono essere valutati dal singolo paziente, che sarà orientato dai propri valori, dalla propria inclinazione al rischio e da preferenze, condizioni finanziarie e sociali. Un intervento può essere considerato accettabile – o desiderabile – quando sia “liberamente accettato da un paziente informato o da un suo valido rappresentante” (Sharpe e Faden, 1996).

Infine, nella definizione di appropriatezza è necessario considerare 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 infatti prendere in considerazione il tipo di obiettivi di salute che si è, o non si è, disposti a perseguire; e questo ha a che vedere non solo con le evidenze scientifiche, ma anche con i valori etici condivisi dalla società stessa e condizionati da elementi di giudizio relativi, ad esempio, alla distribuzione sociale o alla necessità di contenere i costi. In questa prospettiva un dibattito sull’appropriatezza diventa necessariamente una discussione sulla *cost-worthiness* (letteralmente, rapporto costi-valore, costi-merito) di un trattamento sanitario, cioè sull’opportunità di mettere in atto un determinato intervento, dati i costi che esso comporta.¹²

¹² Il dibattito su questo tema è comprensibilmente ricco e complesso, difficilmente riconducibile a voci bibliografiche selezionate. Vale la pena tuttavia ricordare in questa sede la controversa esperienza condotta negli anni '90 dallo Stato americano dell'Oregon, sulla quale è possibile documentarsi a partire da un articolo di Alakeson (2008).

2 • VARIAZIONI DELLA PRATICA CLINICA: CONCETTI, METODI ED ESEMPI

Stefania Rodella, Davide Botturi

Le variazioni delle pratiche clinico-assistenziali per problemi di salute simili, osservabili quando si mettano a confronto aree geografiche o organizzazioni sanitarie, ha rappresentato (e continua a rappresentare) il punto di partenza per molti studi sull'appropriatezza. Questo capitolo ne propone un'interpretazione complessiva, attraverso l'analisi e le conclusioni della letteratura e di esperienze condotte soprattutto negli Stati Uniti e in Gran Bretagna.

LE VARIAZIONI DELLA PRATICA CLINICA: QUALE RELAZIONE CON L'APPROPRIATEZZA?

L'esistenza di considerevoli variazioni nel ricorso/accesso a una vasta gamma di servizi socio-sanitari è un dato di fatto, messo in evidenza da numerosi studi, a partire già dagli anni '40-'50 (cfr. Glover, 2011), e successivamente confermato, fino ad oggi, da una formidabile produzione di letteratura (The King's Fund, 2011), soprattutto per quanto riguarda le prestazioni chirurgiche (ma non solo). Un rapporto pubblicato dal King's Fund (Ham, 1988) aveva già tracciato la storia delle variazioni nell'assistenza e nelle politiche sanitarie a partire dagli anni '70, allo scopo di esplorare le differenze nell'allocazione delle risorse, le variazioni di performance tra organizzazioni socio-sanitarie locali e la variabilità geografica nell'accesso alle cure e nell'utilizzo dei servizi da parte delle popolazioni residenti in determinate

aree. Fu proprio questo report a mettere in risalto l'impatto di queste differenze su equità, efficienza e sicurezza per i pazienti. Le cause – ad esempio, il ruolo della domanda, dell'offerta e del processo decisionale dei professionisti – rimangono le stesse ancora oggi; così come, purtroppo, sembrano sostanzialmente invariate la persistenza e l'entità delle differenze stesse.

Il problema – o il quesito – è in realtà più complesso e potrebbe essere così formulato: quali sono le differenze – o la proporzione della variabilità descritta – desiderabili, e quindi “buone”, e quali sono quelle indesiderate, e quindi “cattive”? Come è stato rimarcato da Evans (1990):

[...] se le differenze indicano l'esistenza di cure inappropriate, quali sono esattamente queste cure? Si tratta di un problema di sovra-utilizzo [*over-use*] da parte delle regioni (aree, strutture, professionisti) dove le frequenze sono più elevate? O piuttosto si è di fronte a un sotto-utilizzo [*under-use*] nel caso delle frequenze più basse? Oppure la frequenza ottimale sta da qualche parte nel mezzo? O persino al di là dell'uno o dell'altro estremo?

Come già sottolineato nel capitolo precedente, le evidenze raccolte ad oggi indicano, in effetti, che non esiste una relazione chiara e ripetibile tra l'appropriatezza e la variabilità geografica: le differenze di utilizzo di servizi o prestazioni sanitarie sembrano essere il risultato di diverse combinazioni di fattori riguardanti i pazienti, i professionisti, l'organizzazione e il contesto. Come ha ben detto Mulley (2010):

Se tutte le variazioni osservate fossero “cattive”, le soluzioni sarebbero semplici. La difficoltà sta 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. 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, e producendo errori di sovrautilizzo generalmente più costosi rispetto a quelli legati al sottoutilizzo.

LE DIFFERENZE NEL RICORSO ALLE CURE. LE CAUSE POSSIBILI

Quanti e quali sono i fattori che stabiliscono il ricorso a una determinata prestazione sanitaria da parte di un paziente che soffre di una particolare condizione clinica? Comprensibilmente numerosi e diversificati.

Possono essere fattori *riconducibili al paziente*, come la gravità delle sue condizioni cliniche, le sue convinzioni e percezioni riguardo alla malattia, la fiducia nei servizi sanitari, la sua inclinazione al rischio e a prendere decisioni, il tipo di comunità e di famiglia a cui appartiene, l'influenza del medico curante, i suggerimenti delle persone che lo circondano. Può trattarsi di *fattori inerenti al sistema dell'offerta*, come la disponibilità di servizi nella zona di residenza del paziente, la distanza geografica tra casa e ospedale, le modalità con cui sono organizzati gli ambulatori (inclusi gli orari di accesso), i tempi d'attesa, il numero di posti letto, il sistema di finanziamento delle organizzazioni sanitarie, gli incentivi ai professionisti.

O ancora, possono entrare in gioco *fattori riguardanti lo stesso professionista*, come la sua competenza, la scuola a cui si è formato, la sua adesione a linee-guida autorevoli e indipendenti.

Infine, devono essere presi in considerazione i problemi di *qualità dei dati*, la loro completezza e rappresentatività, l'incertezza nelle stime.

E naturalmente le possibili interazioni tra queste diverse categorie di fattori sono infinite, come tenta di rappresentare la figura 2.1 (tratta dal già citato report del King's Fund), che viene qui proposta soprattutto per la sua efficacia iconografica nell'evocare la complessità del problema, piuttosto che per il suo contenuto specifico, che è strettamente pertinente alla situazione del National Health Service inglese.

L'identificazione di fattori rilevanti e plausibili in grado di spiegare le differenze osservate costituisce una sfida concreta, che implica la costruzione e la validazione di ipotesi e di modelli statistici. Sia la ricerca statistica che la ricerca qualitativa hanno messo in evidenza come la variabilità della pratica clinica sia uno dei fattori più importanti (The King's Fund, 2011).

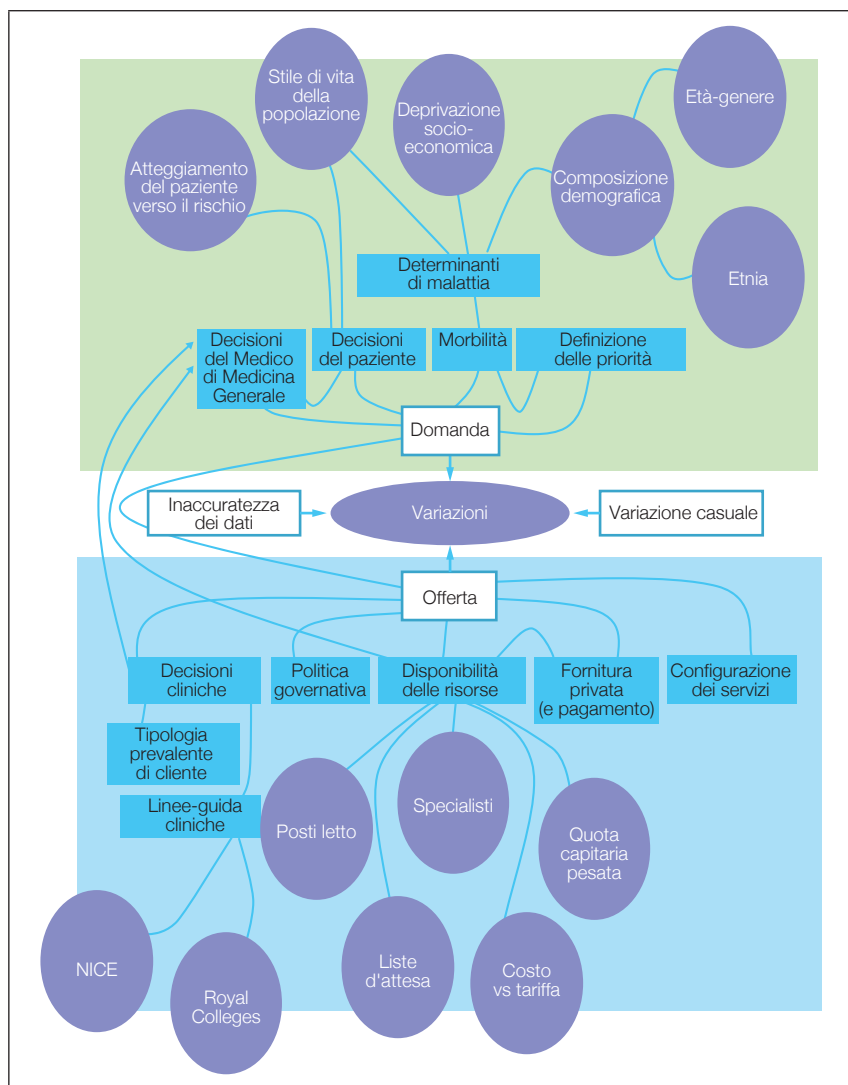


Figura 2.1 • Una possibile mappa delle cause di variazioni nel ricorso ai servizi sanitari (modificata da The King's Fund, 2011)

METODI E MISURE

Le già citate ricerche di Glover negli anni '30 possono essere considerate un'esperienza pionieristica di quella che viene chiamata *small-area variation analysis* (SAV), applicata alla pratica clinica.

La SAV è una metodologia consolidata di ricerca epidemiologica sulla salute e sui servizi sanitari, che ha avuto una particolare diffusione negli anni '80 e '90. 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 o la frequenza di una patologia o condizione clinica) e di spiegarle in termini, ad esempio, di attitudini dei professionisti, di caratteristiche dei pazienti o, ancora, di fattori ambientali (Donabedian, 2010, p. 115; Parchman, 1995).

Le differenze di pratica clinica possono essere analizzate anche in relazione a singoli ospedali o servizi, anziché secondo aree geografiche di residenza di una popolazione (a volte in gran parte sovrapponibili alle aree di competenza di organizzazioni sanitarie complesse, come le nostre USL o come gli inglesi Primary Care Trust – PCT); in questo caso è necessario affrontare specifici problemi di metodo inerenti, ad esempio, al *referral bias*,¹ al *case-mix*,² ecc. Quest'area

¹ Il *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 indicatori 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 scelta dei soggetti che parteciperanno a un determinato studio o avranno accesso a un servizio o a una prestazione; o ancora, quando solo un sottogruppo di una popolazione più ampia utilizza un certo servizio, ad esempio, nel caso di un centro di cura specializzato, universitario, di terzo livello. 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 (cfr. Sica, 2006).

² Il *case-mix* è una caratteristica misurabile di gruppi o popolazioni di pazienti assistiti dalle strutture sanitarie, che consente di descrivere e quantificare il fatto che diversi ospedali, o diversi reparti, trattino casistiche diverse, che includono pazienti con condizioni cliniche di diversa gravità e quindi con probabilità diverse di insorgenza di esiti, complicanze e, in generale, di eventi di interesse per

di ricerca tuttavia, pur essendo sempre pertinente, sul piano concettuale, al tema delle differenze di assistenza sanitaria, nella pratica è generalmente riconducibile all'ambito della valutazione di performance dei servizi sanitari. Non sempre la distinzione tra studi della performance e studi delle variazioni nell'assistenza sanitaria risulta chiara e motivata (a volte sembra essere in gioco un filone di ricerca o una disciplina scientifica, piuttosto che una reale e fondata differenza di approccio); tuttavia il tema della performance esula dagli obiettivi di questo documento e, per ulteriori approfondimenti, si rimanda alla imponente bibliografia e sitografia dedicata.³

Le variazioni di ricorso alle cure possono essere misurate in diversi modi. Il metodo più semplice consiste nel confrontare gli estremi (un minimo e un massimo); il metodo più complesso consiste invece nell'analizzare a fondo l'intera distribuzione del fenomeno. Alcuni indicatori di uso più comune vengono qui di seguito definiti e messi a confronto (figura 2.2) secondo il criterio della complessità, traendo sempre spunto dal documento del King's Fund.

la qualità dell'assistenza. Il non tenere conto di questo fenomeno porta inevitabilmente a distorsioni nei risultati delle osservazioni e a una loro errata interpretazione, con conseguenze potenzialmente scorrette o addirittura dannose sulle eventuali decisioni correlate. La componente del *case mix* viene trattata, negli studi di valutazione di performance o di qualità dell'assistenza, attraverso procedure di *case mix adjustment* (cfr. Donabedian, 2010, pp. 67-9; Iezzoni et al., 1996).

³ Dato l'enorme sviluppo a livello internazionale, soprattutto negli ultimi 10 anni, di sistemi di indicatori per il monitoraggio della performance, a titolo puramente esemplificativo citiamo qui alcune fonti consultabili via web.

Stati Uniti e Canada: l'Agency for Healthcare Research and Quality mantiene una linea di attività specifica sugli indicatori di performance e pubblica periodicamente l'Healthcare Quality Report (<http://www.ahrq.gov>); la Joint Commission sostiene da molti anni un programma dedicato a iniziative di *performance measurement*, che si è evoluto nel tempo. In Canada l'esperienza di maggior rilievo viene dalla provincia dell'Ontario, che ha pubblicato annualmente fino al 2007 un rapporto sulla performance degli ospedali, basato sul modello concettuale della balanced scorecard: <http://secure.cihi.ca/estore/productSeries.htm?locale=en&pc=PCC219>

Europa: Health Care Quality Indicators Project dell'OECD (<http://www.oecd.org/els/health-systems/36262363.pdf>). Il Department of Health inglese conduce da anni un programma di valutazione delle performance delle strutture sanitarie.

Italia: in diverse Regioni sono stati sviluppati sistemi di monitoraggio della qualità dell'assistenza. Lombardia: <http://www.sanita.regione.lombardia.it> (sezione "Qualità e controllo", "Valutazione Aziende"); Lazio: Progetto P.R.E.Val.E (http://151.1.149.72/vislazio/vis_index.php); Toscana: <http://www.meslab.sssup.it/it/index.php?page=sistema-di-valutazione-del-sistema-sanitario-toscana>

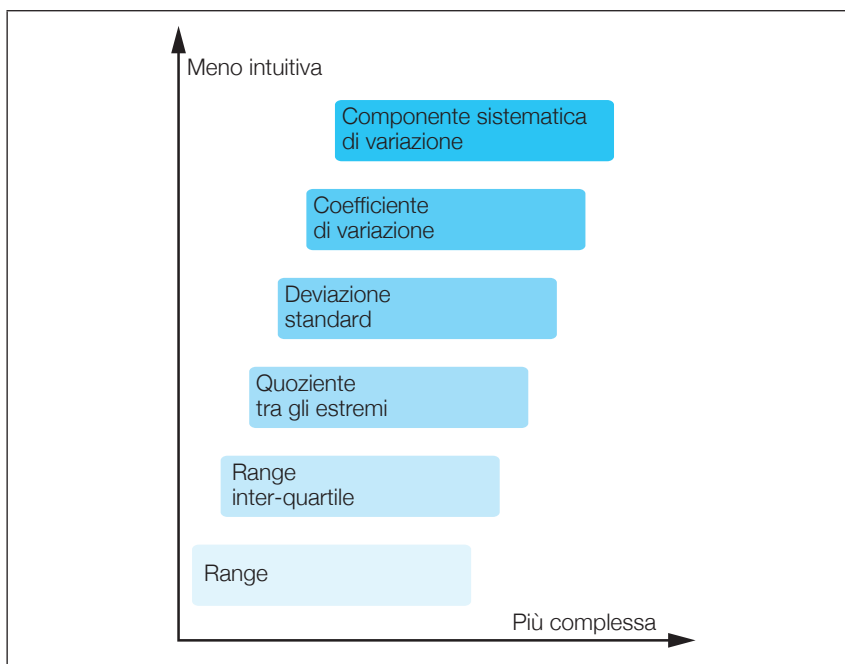


Figura 2.2 • Rappresentazione schematica delle principali misure usate nel descrivere le differenze geografiche di pratica clinico-assistenziale (modificata da The King's Fund, 2011)

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

Range inter-quartile: 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 esempio, 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.

ESEMPI

IL CASO DELLA 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 pionieristico di Glover (1938) a Londra, che osservò come la frequenza del ricorso alla tonsillectomia variesse fino a 20 volte tra un quartiere e un altro. Più o meno nello stesso periodo fu condotto a New York uno studio su mille bambini in età scolare, con risultati sorprendenti (box 2.1; Bakwin, 1958).

Le evidenze sulla tonsillectomia fornite da una revisione Cochrane (Burton e Glasziou, 2009) suggeriscono che l'intervento sia efficace nei bambini con tonsillite grave e ricorrente, ma che l'intervento chirurgico sia associato a rischi e, in definitiva, i bambini possano “uscire da soli” dal problema; per i bambini con malattia più lieve le indicazioni all'intervento sono ancora più modeste; eppure, alcuni

BOX 2.1 - 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. Ma i risultati più sorprendenti furono altri: 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 ($n = 175$). Ancora, 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% che era già tonsillectomizzato all'inizio dello studio – sarebbe stato pari a 935 su 1000!

studi (Suleman et al., 2010) dimostrano ancora differenze di frequenza fino a 7 volte nelle regioni inglesi.

In Italia, un documento di indirizzo del PNLG è stato diffuso nel 2003 (ISS-PNLG, 2003): negli stessi anni il numero di interventi di chirurgia tonsillare è gradualmente diminuito, così come le variazioni geografiche dei tassi standardizzati di tonsillectomia. Nel complesso 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 10,7 (2000) e 10,5 (2002), a 9,1 nel 2003 e 9,4 nel 2004. Contemporaneamente anche il tasso di adenoidectomia ha mostrato un progressivo decremento (figura 2.3).

La differenza tra i tassi standardizzati di tonsillectomia osservati a livello regionale si è ridotta, tra il 2000 e il 2004, da 19 a 16,6 in Piemonte (la regione con il tasso più elevato) e da 3,5 a 4,2 in Basilicata, la regione con il tasso più basso (tabella 2.1). È naturalmente difficile attribuire i cambiamenti osservati alla diffusione delle linee-guida;

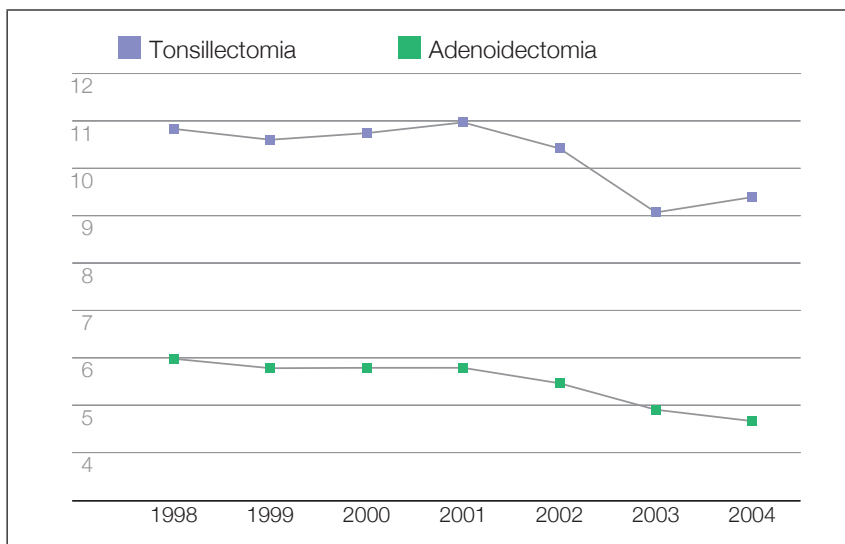


Figura 2.3 • Andamenti temporali dei tassi di tonsillectomia e adenoidectomia in Italia, 1998-2004 (modificata da ISS-SNLG, 2008, figura 1)

più verosimilmente la stesura e la diffusione delle LG è avvenuta in un momento di tensione, da parte del mondo professionale, verso una revisione critica di una pratica tanto consolidata quanto poco scientificamente fondata. Un'analoga esperienza era maturata in Canada e negli Stati Uniti (Vermont), dove era stato osservato un decremento dei tassi di tonsillectomia a seguito della diffusione di linee-guida per la pratica clinica (ISS-SNLG, 2008).

TABELLA 2.1 – TASSI STANDARDIZZATI DI TONSILLECTOMIA, PER REGIONE E PROVINCIA AUTONOMA. ITALIA, CONFRONTO 2000-2004 (MODIFICATA DA ISS-SNLG, 2008)

REGIONE	2000	2004
Abruzzo	13,06	11,67
Basilicata	3,47	4,22
Bolzano (Alto Adige)	15,98	14,83
Calabria	5,15	6,04
Campania	9,26	7,31
Emilia-Romagna	11,62	11,06
Friuli Venezia Giulia	12,58	11,26
Lazio	10,89	9,76
Liguria	1,35	15,65
Lombardia	10,60	9,45
Marche	14,30	10,43
Molise	6,10	8,19
Piemonte	19,04	16,64
Puglia	5,41	4,53
Sardegna	7,28	8,60
Sicilia	10,23	9,69
Toscana	9,23	9,03
Trento	8,49	10,20
Umbria	11,29	8,63
Valle d'Aosta	18,75	16,31
Veneto	12,36	11,85

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 bypass aorto-coronarico – hanno dimostrato frequenze variabili tra un PCT e l'altro (il PCT può essere considerato sovrapponibile alla nostra ASL), espresse, ad esempio, in termini di “quoziente tra estremi”, cioè rapporto tra il tasso di ricorso massimo e il minimo (figura 2.4).

Un'altra osservazione sorprendente riguarda la *persistenza nel tempo* delle variazioni, così come emerge, sempre in Inghilterra, da confronti effettuati tra due periodi temporali successivi, 2005-2006 e 2009-2010, per tre procedure chirurgiche comuni: sostituzione protesica dell'anca, rimozione di cataratta e tonsillectomia (The King's Fund, 2011). Inoltre, al di là delle variazioni annuali osservabili per i singoli PCT, nel tempo è stato possibile constatare come il ricorso alle rispettive procedure sia rimasto 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 (Wennberg e Gittelsohn, 1973), le differenze nel ricorso agli interventi sanitari sono direttamente correlate al grado di consenso professionale riguardo al trattamento indicato per una deter-

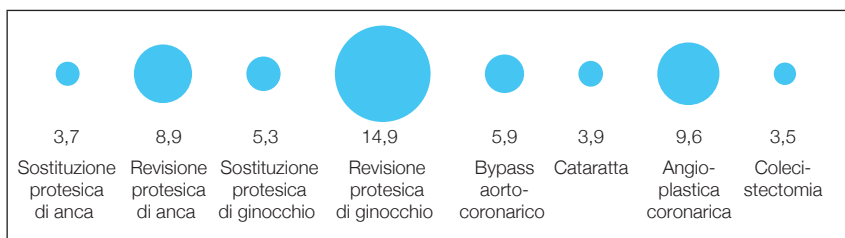


Figura 2.4 • Ricorso a diverse procedure invasive di comune utilizzo in Inghilterra: rapporti tra valore massimo e minimo del tasso dei corrispondenti ricoveri (modificata da The King's Fund, 2011)

minata 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, presentato nelle pagine precedenti, ne è un esempio; nonostante siano state recentemente definite indicazioni rigorose sull'appropriatezza dell'intervento (Suleman et al., 2010), 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 (sempre in Inghilterra) differenze comprese tra 3 e 8 volte, misurate anche in questo caso con il rapporto tra tasso massimo e minimo (figura 2.5). Le procedure considerate erano: isterectomia per via addominale, miringotomia, tonsillectomia, isteroscopia, isterectomia per via vaginale, puntura lombare.

Poiché queste osservazioni sono generalmente basate su fonti informative correnti disponibili per la sanità pubblica, nell'interpretazione di queste differenze è ragionevole tener conto anche dell'eventuale quota di prestazioni che confluiscono sul settore privato (abitudine più diffusa per le procedure “minori”) e che quindi non vengono rilevate, sebbene alcuni studi non abbiano riscontrato un peso determinante di questo fattore (cfr. Suleman et al., 2010).

Il ricorso ad alcune di queste procedure sembra associato alle condizioni socio-economiche, ed è maggiormente elevato per le persone con condizioni più disagiate (misurate da uno specifico indice di

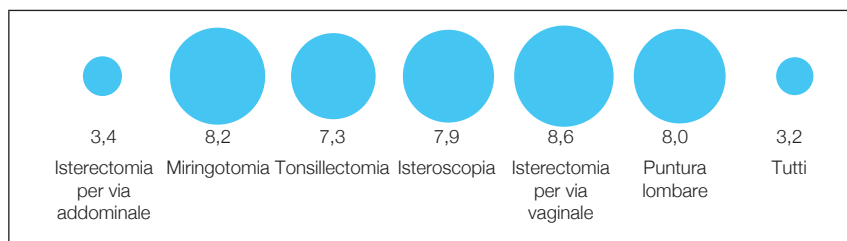


Figura 2.5 • Ricorso a sei procedure di scarsa efficacia in Inghilterra: rapporti tra valore massimo e minimo del tasso dei corrispondenti ricoveri (modificata da The King's Fund, 2011)

deprivazione).⁴ 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 del Department of Health: un rapporto del 2006 infatti sostiene che, nei tempi in cui la procedura era approvata dalla comunità medica e, in un certo senso, popolare e in espansione, erano soprattutto i bambini di famiglie più benestanti a ricorrervi con più frequenza (Department of Health, 2006); più tardi, con la diffusione di maggiori evidenze e con l'acquisizione, da parte dei clinici, di una più ampia consapevolezza e di un atteggiamento più critico, proprio la stessa popolazione – probabilmente più istruita, meglio informata e con un grado più alto di autonomia nelle decisioni sulla salute – ha ridotto il ricorso all'intervento, mentre le persone socialmente più svantaggiate hanno continuato ad essere esposte a un rischio maggiore di *over-use* 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 disuguaglianze ingiustificate di accesso alle cure, determinate non tanto dalle condizioni cliniche dei pazienti o dalle loro preferenze, quanto dalle condizioni socio-economiche. Questa situazione viene ulteriormente sviluppata nel paragrafo successivo.

Per molte condizioni cliniche possono esistere ragionevoli alternative al trattamento chirurgico vero e proprio: ad esempio, una procedura interventistica, uno o più trattamenti medici, un cambiamento negli stili di vita o l'atteggiamento denominato *watchful waiting* (attesa "vigile"). In questi casi paziente e medico si trovano dunque a considerare pro e contro delle diverse possibili decisioni, che comportano rispettive valutazioni in termini di complessità di trattamento,

⁴ Il concetto di deprivazione è frequentemente utilizzato da diverse discipline (demografia, epidemiologia, scienze economiche) ed è 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 (cfr., ad esempio, Morris e Carstairs, 1991; Cadum et al., 1999; Testi e Ivaldi, 2004).

qualità di vita, benefici a breve e a lungo termine. Sono questi i casi in cui il livello di informazione del paziente e il rispetto delle sue preferenze dovrebbero avere il peso maggiore.

In questa categoria possono essere incluse condizioni cliniche quali: il cancro della prostata in stadio precoce, l'osteoartrite dell'anca e del ginocchio, i calcoli della colecisti. La stessa malattia coronarica è trattabile con procedura chirurgica – il bypass aorto-coronarico (BPAC) – con procedura interventistica – l'angioplastica coronarica (PCI) – ma anche con terapia medica; naturalmente 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 comorbilità e le condizioni generali di salute; ma spesso l'entità delle differenze osservate non sembra giustificabile in questo senso. La letteratura internazionale ha ripetutamente documentato differenze nel ricorso a BPAC o PCI fino a 5, 10 volte in diversi ospedali o aree geografiche⁵ (figura 2.6). Inoltre, nonostante la PCI sia da considerarsi spesso come un'alternativa al più invasivo BPAC, sia negli Stati Uniti sia in Inghilterra è stata osservata una correlazione positiva tra le due procedure.

È possibile identificare le ragioni delle differenze osservate? La letteratura (Wennberg, 2011; cfr., più avanti, il capitolo 5 sul Dartmouth Atlas Project) suggerisce due spiegazioni fondamentali:

1. la soglia per la scelta del trattamento è variabile, soprattutto negli stadi precoci della malattia; esiste inoltre una certa dose di incertezza riguardo al possibile beneficio, per quale tipo di paziente e in quale stadio nella storia naturale del problema clinico affrontato;
2. 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; in effetti, è stato dimostrato come pazienti pienamente e correttamente informati riguardo alle alternative disponibili, spesso tendano a scegliere in modo diverso dai rispettivi medici e siano 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.

⁵ Cfr. The Dartmouth Atlas of Health Care (<http://www.dartmouthatlas.org>).

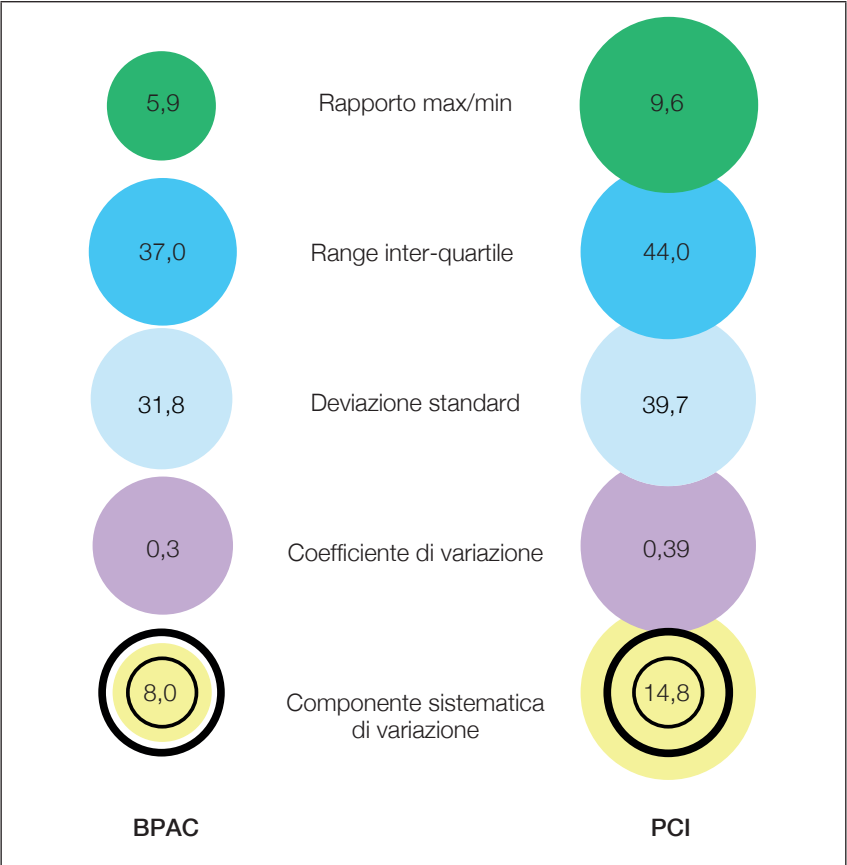


Figura 2.6 • Ricorso a BPAC e angioplastica coronarica (PCI) in Inghilterra: differenze osservate, misurate con diversi indicatori (modificata da The King's Fund, 2011)

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):⁶ in altre parole, i pazienti più poveri hanno minore probabilità di essere sottoposti all'intervento (The King's Fund, 2011; Judge et al., 2010). Analoghe osservazioni sono state riportate da studi condotti negli Stati Uniti e in Canada (Hawker et al., 2002; Fisher et al., 2010).

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),⁷ rilevazioni condotte dal National Health Service inglese a partire dal 2009, in particolare:

⁶ L'MDI viene usato in molte analisi prodotte nel Regno Unito e diffuso, ad esempio, dai Department for Communities and Local Government. L'indice, la cui definizione concettuale e la cui metodologia di calcolo vengono periodicamente aggiornate, tiene in considerazione diverse componenti di stato socio-economico. Ad esempio, la versione 2007 utilizza sette ambiti: 1. reddito; 2. occupazione; 3. condizioni di salute e disabilità; 4. educazione, competenze e training; 5. barriere relative alla condizione abitativa e all'uso dei servizi; 6. criminalità; 7. ambiente di vita (cfr. The English Indices of Deprivation 2007 al link: geoconvert.mimas.ac.uk/help/imd-2007-manual.pdf).

⁷ L'espressione "Patient reported outcome" (PRO) identifica una terminologia "ombrello", che include 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. Questi possono essere *general*, rivolti soprattutto alla valutazione della qualità della vita o alla valutazione soggettiva dello stato di salute, come l'SF-36 Health Survey, l'SF-12 Health Survey, il Sickness Impact Profile, il Nottingham Health Profile, l'Health Utilities Index, il Quality of Well-Being Scale, l'EuroQol (EQ-5D) e il Consumer Assessment of Healthcare Providers and Systems (CAHPS). Altri strumenti sono *specifici per patologia*, come l'Adult Asthma Quality of Life Questionnaire (AQLQ) per i pazienti adulti con asma, il Kidney Disease Quality of Life Instrument per i pazienti con malattie renali, l'Oxford Hip Score per i pazienti sottoposti a intervento di protesi d'anca e molti altri (per una ricerca interattiva di voci bibliografiche inerenti ai PROMs cfr. http://phi.uhce.ox.ac.uk/perl/phig/phidb_search.pl).

l'Oxford Hip Score⁸ e l'indice generico di qualità della vita EQ-5D.⁹ Da rilevazioni effettuate nel 2009-2010 emerge come entrambe queste misure si attestino su valori più bassi, cioè più sfavorevoli – che denotano quindi una peggiore salute riferita dal paziente – per le persone con indice di privazione più elevato. Inoltre, i pazienti con peggiore salute riferita tendevano a provenire maggiormente da aree geografiche con minor frequenza di ricovero per sostituzione protesica dell'anca.

La relazione tra privazione e intervento di protesi d'anca si gioca 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: potrebbe riflettere ad esempio, un grado più elevato di comorbidità, un maggior ritardo nell'arrivare all'intervento (e quindi in condizioni più avanzate, per disabilità o per dolore) o altri fattori riconducibili, anche con modalità più complesse, alla condizione socio-economica. In definitiva, esistono evidenze della diseguità di accesso alla sostituzione protesica dell'anca in Inghilterra, i cui determinanti principali risultano essere: età, genere, privazione, residenza in aree rurali ed etnia (Dixon et al., 2004; Judge et al., 2010).

Osservazioni analoghe a quelle qui presentate per l'implantologia protesica dell'anca sono state fatte per altre procedure, come la rivascolarizzazione coronarica e l'intervento di cataratta (Weinstein et al., 2004).

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

⁸ Il sito <http://www.orthopaedicscore.com/> riporta tutti i tipi di questionari utilizzati per il calcolo degli score in pazienti ortopedici, sia dal punto di vista del professionista sia dal punto di vista del paziente, incluso l'Oxford Hip Score.

⁹ Cfr. a questo proposito il sito dell'EuroQol Group: <http://www.euroqol.org/>

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 (una notevole riduzione dei giorni di degenza).

Nel National Health Service inglese la percentuale di interventi effettuati in DS si avvicina ora al 70%:¹⁰ tuttavia sono ancora elevate le differenze tra 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 moltissime procedure chirurgiche potenzialmente effettuabili in DS (circa 1 milione), ma oltre il 20% (220.000 interventi) di esse non ha avuto luogo in questo *setting*; la proporzione di interventi in DS ha mostrato differenze significative tra un PCT e l'altro, con valori compresi tra il 67 e l'87%.

Per alcuni interventi specifici, come la chirurgia della cataratta, le variazioni sono state minime (88-100% di interventi in DS), ma in altri casi è stato osservato un considerevole sottoutilizzo in alcuni PCT, come, ad esempio, per la tonsillectomia e per la colecistectomia laparoscopica. Per alcune procedure è stata inoltre dimostrata un'associazione negativa con indicatori di svantaggio socio-economico (deprivazione); in particolare, nei PCT con popolazioni "più povere", gli interventi di circoncidione, resezione transuretrale di tumori vescicali, decompressione del tunnel carpale, artroscopia, chirurgia della cataratta e alcuni altri, sono stati eseguiti con maggior frequenza in regime di degenza ordinaria. Secondo la Audit Commission (2001) questa osservazione non è determinata dalle caratteristiche dei pazienti, ma da diverse caratteristiche dell'offerta, come la disponibilità di infrastrutture o l'attitudine clinica dei professionisti.

Tuttavia, la proporzione di interventi eligibili effettivamente eseguiti in DS ha mostrato un aumento costante negli ultimi anni pas-

¹⁰ Misurato come rapporto tra il numero di pazienti trattati in DS e il numero totale di pazienti operati con chirurgia elettiva, sia in regime di degenza tradizionale sia in DS.

sando, complessivamente, da un 56% nel 1996-1997 a un 78% nel 2009-2010.

Per quanto riguarda il nostro Paese, molti e complessi sono stati i cambiamenti culturali, istituzionali e organizzativi che hanno caratterizzato, negli ultimi 20 anni, il graduale sviluppo di setting assistenziali più snelli ed economicamente meno onerosi, sia in ambito medico che in ambito chirurgico (ad esempio, day hospital, day service, day surgery, chirurgia ambulatoriale); una buona sintesi per orientarsi può essere offerta dal documento finale prodotto nell'ambito del Progetto Mattoni, promosso dal Ministero della Salute (2007). Un'interessante narrazione del percorso evolutivo del concetto e del modello organizzativo di day surgery e chirurgia ambulatoriale in Italia è quella firmata nel 2006 da Elio Guzzanti e Isabella Mastrobuono, dal titolo *La day surgery e la chirurgia ambulatoriale in Italia: la storia dei primi quindici anni di attività scientifica, istituzionale ed organizzativa e le prospettive di sviluppo*.¹¹

¹¹ http://www.daysurgeryitalia.it/pdf/articoli/organizzativi/storia_day_surgery.pdf

3 • LA MISURA DELL'APPROPRIATEZZA

Stefania Rodella, Davide Botturi

La misura dell'appropriatezza continua ad essere una sfida per i clinici, per i ricercatori, per i manager e per i soggetti finanziatori dei servizi sanitari.

Prerequisito per la messa in atto di un processo concreto di misurazione/monitoraggio dell'appropriatezza è naturalmente la definizione di criteri espliciti, scientificamente fondati, ai quali poter ricondurre valutazioni, individuali o collettive, effettuate da esperti.

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.

LA VALUTAZIONE DI APPROPRIATEZZA DEI RICOVERI OSPEDALIERI: DAL PROTOCOLLO PRUO AI DRG INAPPROPRIATI

L'USO DELLA DOCUMENTAZIONE CLINICA

I protocolli per la valutazione di appropriatezza del ricovero, nel paziente adulto e nel bambino, hanno avuto in Italia, soprattutto negli anni '90 e nei primi anni 2000, una certa popolarità concretizzatasi

nella diffusione dei cosiddetti studi PRUO (Protocolli di Revisione d'Uso dell'Ospedale), che in seguito si è molto ridimensionata, fino a rivestire oggi probabilmente un interesse storico-documentale, piuttosto che pratico.

Il primo esemplare della “famiglia PRUO” può essere considerato 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, la gravità della malattia che ha condotto al ricovero, l'intensità dell'assistenza fornita, le prestazioni effettuate in ciascun giorno di degenza, l'assistenza infermieristica, le condizioni dei pazienti e a definire, di conseguenza, sulla base di una griglia di riferimento, 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% (Payne, Ash e Restuccia, 1991). 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 Regionale San Gerardo di Monza, cfr. Apolone et al., 1993) al 28% (in uno studio condotto all'Ospedale di Aosta: cfr. Pesenti Compagnoni, 1991).

L'AEP ha rappresentato la matrice di riferimento per l'elaborazione del Protocollo di revisione d'uso dell'ospedale (PRUO) (Gruppo di lavoro PRUO, 1994), promosso da 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. Il PRUO conserva una struttura sostanzialmente simile all'AEP, basandosi su 18 criteri espliciti, diagnosi-indipendenti, relativi alla condizione del paziente e all'intensità/complessità del trattamento. Il giudizio di congruità del ricovero è formulato qualora venga soddisfatto almeno un criterio. Rispetto all'AEP, il PRUO contiene una dettagliata lista che permette di individuare i motivi delle ammissioni non giustificate.

La complessità e la varietà delle motivazioni (e delle catene di eventi) che conducono a un ricovero ospedaliero rendono praticamente impossibile la definizione di gold standard a cui rapportarsi

e difficilmente risolvibile il problema del giudizio di appropriatezza d'uso dell'ospedale.

Il PRUO consente di formulare giudizi tramite criteri espliciti, obiettivi, verificabili, riproducibili, indipendenti dalla diagnosi. Strumenti quali l'AEP o il PRUO:

...sono buoni ed accettabili per l'uso in quanto se ne riconosce il valore definitorio e di esplicitazione di regole e criteri. Si deve, in altre parole, cercare un difficile equilibrio tra ragionevolezza e arbitrarietà trovando una concordanza sufficientemente ampia tra compiti assegnati all'ospedale, esplicitazione di regole di comportamento clinico ed amministrativo e monitoraggio della compliance a queste regole (Gruppo di lavoro PRUO, 1994).

Nella conduzione di studi basati sul PRUO è necessario considerare con attenzione alcune fasi, a partire dalle *modalità di campionamento*.

Il campione può essere scelto in modo casuale (*random*) o sulla base di valutazioni specifiche e contingenti (campione di "convenienza") o in modo tale da risultare rappresentativo della tipologia di ricoveri indagati.

I ricoveri possono essere selezionati in funzione della giornata di degenza, ad esempio, della totalità o di un sottogruppo di pazienti presente in uno o più reparti in una o più specifiche giornate ("giornata-indice"; studio trasversale). In questo caso si risale a ritroso al giorno dell'ingresso in ospedale per valutare – sulla base di ciò che è riportato nella documentazione clinica – anche l'appropriatezza dell'ammissione. Si deve tenere presente che, in campioni così selezionati, i ricoveri a più lunga durata di degenza avranno maggiore probabilità di essere rappresentati.

In alternativa, i ricoveri possono essere selezionati sulla base delle nuove ammissioni, seguite poi nel corso della degenza (studio longitudinale). In questo caso l'appropriatezza dell'ammissione viene valutata subito (e a questo scopo possono essere utilizzate anche informazioni ricavate dal medico accettante e/o di reparto) procedendo in seguito a valutare una o più successive giornate di degenza. Campioni

così selezionati saranno maggiormente rappresentativi dell'effettiva distribuzione delle degenze.

Altri aspetti devono essere attentamente valutati nella pianificazione, conduzione e interpretazione di studi basati sul PRUO. Innanzitutto 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, ecc.) e l'insieme (mix) di reparti considerati, poiché le percentuali e le ragioni della non appropriatezza secondo PRUO possono variare da reparto a reparto; in secondo luogo il case-mix dei pazienti; e infine il timing della rilevazione, poiché la raccolta di informazioni può essere concorrente o retrospettiva. Nel primo caso (rilevazione concorrente) i dati vengono raccolti nel momento stesso in cui il paziente è presente in ospedale, nello stesso giorno o nel giorno successivo. È evidente come in questo caso la raccolta delle informazioni sia più ricca e dettagliata e possa contare, oltre che sulla cartella clinica, su fonti di informazione dirette quali il paziente stesso, il curante, i familiari. Nel secondo caso (rilevazione retrospettiva) i dati vengono raccolti a posteriori e la cartella clinica, il referto di accettazione, la cartella infermieristica rappresentano l'unica fonte di informazione.¹

L'USO DEI DATABASE AMMINISTRATIVI

I DRG AD ELEVATO RISCHIO DI INAPPROPRIATEZZA

Il DPCM del 29 novembre 2001 (*Definizione dei livelli essenziali di assistenza*) riporta, nell'Allegato 2C, l'elenco dei 43 DRG ad elevato rischio di inappropriatezza (cfr. *Appendice normativa*, alla fine del presente volume). Questa prima lista è stata riesaminata nell'ambito del "Progetto Mattoni. Misura dell'appropriatezza. 1.2 Analisi appropriatezza organizzativa" (Ministero della Salute, 2006), che si è proposto di sviluppare, sperimentare e validare una metodologia,

¹ Per ulteriori approfondimenti su AP e PRUO e sulle loro applicazioni, soprattutto in Italia, cfr. Payne et al., 1991; Pesenti Compagnoni, 1991; Apolone et al., 1993; Gruppo di lavoro PRUO, 1994; Fellin et al., 1995; Lattuada et al., 1998; Smeets et al., 2000; Lattuada et al., 2001; Materia, 2003.

statisticamente robusta e riproducibile, per la stima del livello di rischio di inappropriatezza organizzativa associato ai singoli DRG, sia a livello nazionale sia a livello regionale.

Nel progetto è stata adottata la seguente definizione di DRG ad elevato rischio di inappropriatezza: “DRG che presentano un valore ‘anomalo’ (rispetto al corrispondente valore nazionale) della percentuale di ricoveri effettuati in regime diurno oppure (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 del 29 novembre 2001, andando a creare una nuova lista nazionale di 55 DRG potenzialmente inappropriati dal punto di vista organizzativo.

Il metodo, che può essere usato tanto a livello nazionale quanto a livello regionale, si propone di rappresentare uno strumento di applicazione periodica e ricorsiva; in ragione dei continui cambiamenti clinici e organizzativi dell'assistenza, è utile infatti, soprattutto per le Regioni, disporre di riferimenti sempre aggiornati per le stime di inappropriatezza dei ricoveri, basati su strumenti e metodi semplici, adattabili al contesto regionale e basati su dati amministrativi di facile reperibilità e la cui raccolta risulti relativamente economica.

Si raccomanda anche, alle singole Regioni, di integrare la lista proposta con una valutazione di tipo clinico multidisciplinare, che possa meglio tener conto degli specifici contesti organizzativi e gestionali (cfr. anche più avanti, capitolo 8).

ESPERIENZE IN ITALIA

I metodi di misurazione dell'appropriatezza basati sull'esame delle cartelle cliniche, indubbiamente onerosi e costosi nella pratica, si sono sviluppati nell'assunto che non sia possibile valutare l'appropriatezza dei ricoveri ospedalieri a partire dalle informazioni contenute nella scheda di dimissione ospedaliera; quest'ultima infatti, essendo un documento di natura amministrativa, soffre di problemi di qualità formale e sostanziale ormai 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 delle 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 inappropriately su insiemi anche piuttosto ampi di ricoveri, monitorandone l'andamento nel tempo (Baglio et al., 2001).

Lo studio dei ricoveri ospedalieri inappropriati – perché 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* (Taroni et al., 1997), o i DRG o sistemi più complessi e più o meno commerciali, come MEDISGRPS e altri, 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. Tra le esperienze realizzate, vale la pena ricordare quella condotta ormai diversi anni fa

nel Lazio, che ha adottato il sistema APR-DRG² per una simulazione condotta sulle dimissioni ospedaliere di ospedali pubblici e privati accreditati, nell'anno 1999. La metodologia e i limiti dello studio sono ben descritti nell'articolo di Baglio et al. (2001); qui basti ricordare che, oltre alla simulazione sui ricoveri, è stata effettuata anche una simulazione di possibili sanzioni correlate all'inappropriatezza, che ha condotto a una stima della diminuzione di spesa regionale pari a 44 miliardi di vecchie lire.

L'uso degli APR-DRG (successivamente applicati anche in altre realtà ospedaliere o, meno frequentemente, regionali) 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.

Ancora sull'utilizzo degli APR-DRG si basa un'esperienza condotta a livello nazionale, secondo il cosiddetto metodo APPRO (Fortino et al., 2002): lo studio è stato condotto sulle dimissioni ospedaliere avvenute nelle 21 Regioni e Province autonome nel corso dell'anno 2000, assegnate ai 43 DRG ad elevato rischio di inappropriatezza (Allegato 2C del DPCM 29 novembre 2001) e trasmesse al Ministero della Salute. Lo scopo principale di questo esercizio consisteva nel calcolare le soglie di ammissibilità nazionale verso cui misurare le performance regionali di appropriatezza, misurando i valori del-

² 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 (Baglio et al., 2001).

le singole Regioni e stimandone sia la quota di ricoveri certamente inappropriati sia la quota di ricoveri inappropriati rispetto alle migliori performance, generando quindi un terreno utile a iniziative di *benchmarking*. La metodologia messa a punto intendeva proporsi come soluzione “aperta”, a disposizione delle singole Regioni.

Infine, merita di essere segnalata l’esperienza della Regione Emilia-Romagna, alla quale viene dedicato un apposito capitolo di approfondimento (capitolo 8).

IL METODO RAND PER LA VALUTAZIONE DI APPROPRIATEZZA

BACKGROUND

Il metodo RAND/UCLA per la valutazione di appropriatezza (RAND/UCLA Appropriateness Method, RAM) è stato sviluppato alla metà degli anni ’80 come parte di uno studio più ampio,³ sostanzialmente come strumento di misura del sovrautilizzo e del sottoutilizzo di procedure mediche e chirurgiche.

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 esempio, un aumento dell’aspettativa di vita, il sollievo dal dolore, la riduzione dell’ansia, il miglioramento della capacità funzionale) superino le previste conseguenze negative (ad esempio, mortalità, morbidità, ansietà, dolore, tempo sottratto al lavoro) con un margine sufficientemente ampio tale da giustificare la procedura, senza tener conto dei costi” (cfr. capitolo 1).

Il razionale che sostiene il metodo (ideato principalmente dal ricercatore Robert Brook) consiste nel prendere atto dell’insufficienza dei *randomised clinical trials* nel fornire evidenze a un dettaglio tale

³ RAND Corporation/University of California Los Angeles (UCLA), Health Services Utilization Study (<http://www.rand.org/health/feature/forty/health-services-utilization-study.html>).

da trovare applicazione alla vastissima gamma di situazioni in cui i pazienti vengono trattati ogni giorno.

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 alle opportunità e alle indicazioni per il 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 all'appropriatezza di esecuzione di una procedura in situazioni specifiche di pazienti con determinati sintomi, storia clinica e risultati di test clinici.

Molte procedure sono state oggetto di studi sull'appropriatezza negli Stati Uniti, ad esempio: angiografia coronarica, bypass aorto-coronarico, angioplastica coronarica, endoarterectomia carotidea, chirurgia dell'aneurisma aortico addominale, endoscopia diagnostica del tratto gastrointestinale superiore, chirurgia della cataratta, colonscopia, colecistectomia, isterectomia, timpanostomia e manipolazione del rachide per il mal di schiena.

Dai tempi delle prime applicazioni il metodo si è successivamente diffuso in molti Paesi – Canada, Israele, Italia, Olanda, Spagna, Svezia, Svizzera, Regno Unito e altri Paesi, soprattutto in Europa Occidentale – ed è stato applicato a ulteriori e numerosi problemi clinici, come l'iperplasia prostatica benigna, la laminectomia, il cancro della mammella e la sostituzione protesica dell'anca (Kahan e Van Het, 1999).

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 (Chassin et al., 1987). Poiché una mole consistente di ricerche aveva documentato variazioni importanti nell'uso di determinate procedure tra regioni geografiche ed entro tali regioni, i ricercatori della RAND avanzarono l'ipotesi che le differenze potessero essere in parte spiegate da un maggior livello di inapproprietezza nelle aree con maggiori volumi. Sorprendentemente, 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.

In una testimonianza di fronte al Comitato del Senato nel marzo 1999, Robert Brook citò i risultati di uno studio condotto nel Regno Unito, secondo il quale i medici eseguivano un numero di procedure cardiache pari a un settimo del numero effettuato negli Stati Uniti e i pazienti erano spesso costretti a lunghe attese per ricevere la rivascolarizzazione coronarica. Lo studio dimostrò quindi un importante “sottoutilizzo di procedure in persone che ne avevano necessità” e un altrettanto importante “sovrautilizzo in persone che non ne avevano bisogno”. Per le definizioni di sovrautilizzo (*over-use*), sottoutilizzo (*under-use*) e cattivo utilizzo (*mis-use*) delle procedure e degli interventi sanitari cfr. il box 3.1.

Per misurare il sottoutilizzo, il metodo RAND/UCLA fu esteso alla misura della *necessità* di procedure cliniche (Kahan, et al., 1994). La necessità è un criterio più stringente dell’appropriatezza e si riferisce a procedure che devono essere offerte ai pazienti che corrispondono a determinati profili clinici. Criteri di necessità sono stati sviluppati ad esempio, per l’angiografia coronarica e la rivascolarizzazione coronarica, per misurare il sottoutilizzo o il “bisogno non soddisfatto”. Tuttavia, la necessità è più difficile da misurare rispetto all’appropriatezza, perché implica l’identificazione di gruppi di pazienti che avrebbero potuto beneficiare di una certa procedura, o di un trattamento, che non hanno ricevuto.

COME FUNZIONA IL METODO RAND

Il razionale e le modalità di utilizzo del metodo RAND sono descritti ampiamente in un famoso manuale (Fitch et al., 2000) pubblicato in versione aggiornata e articolato in modo da costituire una guida pratica (sebbene non propriamente un “libro di ricette”) per i nuovi utilizzatori, basata sull’ormai consolidata ed estesa esperienza del metodo nel corso degli anni. Il manuale è articolato in 13 capitoli, che ripercorrono le tappe operative del metodo, dalla selezione del problema di interesse alla revisione della letteratura, alla costituzione

BOX 3.1 - SOVRAUTILIZZO, SOTTOUTILIZZO ED ERRATO UTILIZZO DELLE CURE MEDICHE. DEFINIZIONI ED ESEMPI**Sovrautilizzo (*over-use*)**

Si verifica quando un farmaco o un trattamento vengono somministrati senza giustificazione medica.

Esempio tipico di sovrautilizzo è il trattamento con antibiotici in caso di infezioni semplici (soprattutto di origine virale) che non lo richiedono. Gli antibiotici vengono prescritti nell'80% dei casi in modo inappropriato nei bambini con infezioni dell'orecchio, nonostante la ricerca abbia dimostrato che queste infezioni si risolvono senza antibiotici entro tre giorni.

Un altro esempio tipico 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, configurando così un uso e un costo non necessario della tecnologia.

L'*over-use* di cure in ragione dell'offerta si verifica anche nel trattamento di malattie croniche (ad esempio, l'ospedalizzazione di pazienti con diabete, piuttosto che il loro trattamento a casa o in ambulatori territoriali). La causa di questo fenomeno può risiedere in una dipendenza eccessiva dalle strutture di cura per acuti e in una mancanza di infrastrutture necessarie a sostenere l'assistenza ai malati cronici in altri tipi di setting.

Sottoutilizzo (*under-use*)

Si verifica quando i medici o gli ospedali trascurano di somministrare ai pazienti cure necessarie o di mettere in atto pratiche di provata efficacia.

Esempi tipici sono quelli della terapia con beta-bloccanti alle persone con infarto miocardico acuto, o dell'esame della retina per i pazienti diabetici, ma esistono infiniti, allarmanti esempi di sottoutilizzo.

Si stima che oltre 90.000 americani muoiano ogni anno per non aver ricevuto la corretta terapia, di efficacia documentata, per condizioni cliniche croniche come l'ipertensione arteriosa, il diabete e la malattia coronarica o per omissione di pratiche di prevenzione secondaria, come lo screening per il cancro della mammella o la vaccinazione antinfluenzale nelle persone di età uguale o superiore a 65 anni.

Le cause di *under-use* includono la discontinuità nelle cure (che tende a peggiorare quando più medici sono coinvolti nella cura dello stesso paziente) e la mancanza di procedure e sistemi (anche proattivi) in grado di facilitare l'accesso ai servizi necessari.

Uso errato (*mis-use*)

Con questo termine si descrive di fatto l'errore medico, che si verifica quando un paziente va incontro a problemi prevenibili con determinate pratiche mediche o quando soffre conseguenze indesiderate di diversa entità a causa di un trattamento medico. Un esempio tipico è quello della somministrazione di un antibiotico a un paziente con allergia nota e documentata.

Un altro caso di *mis-use* riguarda le condizioni “sensibili alle preferenze dei pazienti” (*preference sensitive care*), in cui esistono significativi *trade-off* tra soluzioni alternative allo stesso problema. La scelta dell'opzione più appropriata dovrebbe tener conto delle preferenze del paziente, come nel caso della scelta tra nodulectomia e quadrantectomia per un cancro della mammella in stadio precoce. Il *mis-use* può derivare da una mancata o carente comunicazione in merito alle opzioni disponibili e/o da un'insufficiente attenzione per le preferenze della persona direttamente coinvolta nella scelta.

del panel, via via fino all'applicazione dei risultati a supporto delle decisioni. In accordo con gli scopi del presente volume, si ritiene sufficiente proporre una sintesi ragionata del metodo RAND (e del manuale), rimandando i lettori eventualmente interessati ad approfondire l'argomento alla bibliografia e sitografia dedicata. Le tappe fondamentali di applicazione del metodo RAND sono rappresentate nella figura 3.1 e riprese nei box 3.2, 3.3 e 3.4.

Prima di tutto, viene condotta un'accurata revisione della letteratura volta a sintetizzare le più recenti evidenze scientifiche disponibili in merito alla procedura/intervento che deve essere valutato.

Allo stesso tempo, un elenco di specifici “scenari” o “indicazioni” viene elaborato, in forma di matrice, per classificare i pazienti che potrebbero essere candidati all'intervento in questione in termini di: sintomi, storia clinica, risultati dei test diagnostici più rilevanti.

Queste indicazioni vengono raggruppate in “capitoli”, basati sui sintomi fondamentali che conducono il paziente a presentarsi o ad essere preso in considerazione per la procedura in questione. Un esempio di specifica indicazione per la rivascolarizzazione coronarica nel capitolo dedicato all'angina stabile cronica è il seguente:

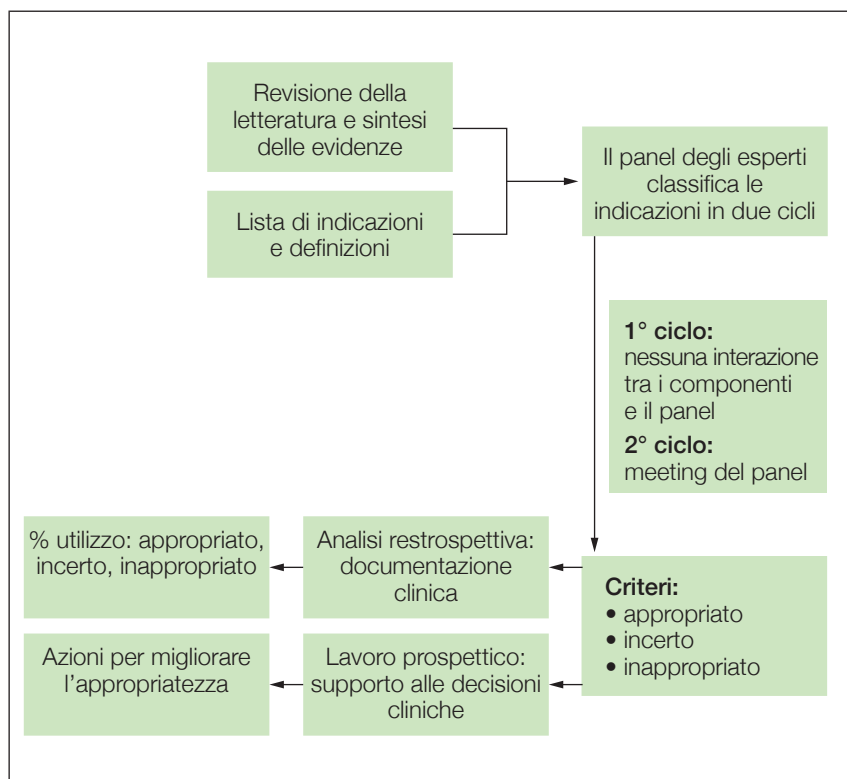


Figura 3.1 • Il metodo RAND/UCLA per la valutazione dell'appropriatezza (modificata da Fitch et al., 2000)

Paziente con angina grave (classe III/IV) nonostante l'assunzione di terapia medica ottimale; con malattia dei 2 vasi senza coinvolgimento della arteria discendente prossimale sinistra anteriore; con frazione di eiezione (FE) compresa tra 30 e 50%, stress test francamente positivo, rischio chirurgico basso-moderato.

Un panel di esperti (generalmente tra 9 e 11) viene costituito, sulla base di criteri o raccomandazioni proprie del contesto in cui lo studio si svolge. La revisione della letteratura e la lista di indicazioni (ovvia-

mente prodotte da un gruppo di lavoro tecnico, con esperienza specifica), insieme a un elenco delle definizioni usate nella lista di indicazioni, vengono inviate ai membri del panel. Per ciascuna indicazione, i membri del panel esprimono un giudizio sul rapporto costi-benefici della procedura, usando un punteggio su una scala da 1 a 9, dove 1 significa che i rischi attesi superano di gran lunga i benefici e 9 significa esattamente l'opposto, con una netta preponderanza dei benefici attesi rispetto ai rischi. Un punteggio intermedio pari a 5 può significare che benefici e rischi vengono pesati allo stesso modo oppure che il valutatore non è in grado di esprimere un giudizio per il paziente descritto in quella specifica indicazione. I componenti del panel esprimono un giudizio sulle indicazioni proposte in due tempi, in un processo a due cicli che rappresenta una sorta di Delphi⁴ modificato. Nel *primo ciclo* le valutazioni vengono fatte individualmente, da ciascun esperto del panel, senza alcuna interazione con gli altri componenti; nel *secondo ciclo* i membri del panel si incontrano per 1-2 giorni, con la guida di un moderatore esperto nell'uso del metodo. Ciascun componente del panel riceve un dossier individuale che mostra la distribuzione dei giudizi formulati da tutti i componenti del panel nel corso del primo ciclo, insieme con il suo giudizio personale. Durante il meeting, i componenti del panel discutono i punteggi, si concentrano sulle aree di disaccordo e hanno la possibilità, se lo ritengono opportuno, di modificare la lista originaria di indicazioni e/o definizioni. Dopo aver completato la discussione, essi ripetono la valutazione con una nuova attribuzione di punteggio, sempre individuale. Nessun tentativo viene messo in atto per forzare il panel al consenso. Piuttosto, la valutazione condotta in due cicli è disegnata per far emergere la differenza tra discordanze di punteggio dovute a reale disaccordo clinico in merito all'uso della procedura (disaccordo "reale") e discordanze determinate da fatica o fraintendimenti (disaccordo "artefattuale").

Infine, ciascuna indicazione viene classificata come "appropriata", "incerta" o "inappropriata" per la procedura in esame, in accordo con lo score mediano e con il livello di disaccordo tra i componenti del panel. Le indicazioni che hanno riportato valori mediani di score com-

⁴ La tecnica Delphi è uno dei metodi più diffusi per raggiungere il consenso (cfr. anche il box 3.2).

presi tra 1 e 3 vengono classificate come “inappropriate”; quelle con score mediano tra 4 e 6 sono giudicate “incerte” e quelle nel range 7-9 vengono classificate come “appropriate”. Tuttavia, tutte le indicazioni classificate “con disaccordo”, qualunque sia il punteggio mediano, vengono classificate come “incerte”.

In questi casi “disaccordo” significa sostanzialmente una mancanza di consenso, dovuta ad esempio, a una polarizzazione del gruppo o a una distribuzione degli score su tutta la scala di punteggi, tra 1 e 9.⁵

Nel caso in cui si intendano sviluppare anche *criteri di necessità* (cfr. capitolo 1), viene effettuato un terzo ciclo di valutazioni, che si svolge generalmente via e-mail, in cui ai componenti del panel viene chiesto di valutare il grado di necessità di quelle indicazioni che sono state già giudicate appropriate dal panel stesso. La definizione di necessità, secondo la metodologia RAND, si basa sui seguenti punti (Kahan et al., 1994):

- la procedura è appropriata (secondo la definizione RAND più volte citata);
- non offrire la procedura al paziente configurerebbe un’assistenza impropria;
- esiste una ragionevole probabilità che la procedura porti beneficio al paziente;
- l’entità del beneficio atteso non è piccola.

Tutti i criteri precedenti devono essere soddisfatti perché una procedura sia considerata necessaria in una determinata indicazione. Anche la valutazione della necessità viene fatta sulla base di una scala da 1 a 9, dove il punteggio = 1 identifica la procedura chiaramente non necessaria e il punteggio = 9 la procedura sicuramente necessaria. Se i componenti del panel sono in disaccordo nella valutazione della

⁵ Definizioni alternative di “disaccordo” sono state usate lungo la storia evolutiva del metodo RAND. Gli studi di appropriatezza talvolta classificano ulteriormente i livelli di accordo, per identificare le indicazioni valutate “con accordo” e quelle con “accordo indeterminato” (né accordo né disaccordo). In ragione degli utilizzi previsti per il risultato del processo di rating e i relativi criteri di appropriatezza, può essere conveniente talvolta identificare le indicazioni valutate con maggiore o minore grado di accordo.

necessità, o se la mediana del punteggio è inferiore a 7, allora la procedura viene giudicata “appropriata ma non necessaria”. Solo le indicazioni appropriate con uno “score di necessità” uguale o superiore a 7 senza disaccordo possono essere classificate come “necessarie”.

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). *Le assunzioni del metodo*, è utile ricordarlo, sono sostanzialmente tre:

1. il giudizio di appropriatezza consiste essenzialmente nel valutare il profilo rischio-beneficio di una procedura (così come testimoniato dalla stessa definizione di appropriatezza adottata dalla RAND). Questo lascia ampio spazio all’interpretazione soggettiva dei singoli esperti che compongono il panel, chiamati, e in un certo modo forzati, a valutare l’impatto complessivo di una procedura, non soltanto questo o quell’aspetto specifico;
2. la seconda assunzione riguarda l’incompletezza e l’insufficienza delle informazioni derivabili dalla letteratura scientifica circa l’efficacia di un trattamento. È indispensabile, quindi, che le evidenze scientifiche vengano integrate dall’opinione di esperti;
3. la terza assunzione è che ciascun clinico nel valutare gli scenari clinici faccia riferimento alla propria esperienza, filtrata attraverso le evidenze della letteratura scientifica. Si tratta quindi di un metodo centrato sugli aspetti tecnico-clinici, che permette di definire criteri di appropriatezza molto dettagliati per individuare i pazienti che possono trarre beneficio da uno specifico intervento.⁶

Se da una parte, come è già stato ricordato, 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:

⁶ Questo metodo è stato applicato anche nel nostro Paese, in particolare nell’ambito del Progetto “Oncologia femminile”, al fine di produrre linee-guida in oncologia ginecologica.

- si può verificare una 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. Questo fatto diventa evidente quando si osservano le differenze di risultati cui sono giunti panel diversi di clinici in presenza di uguali indicazioni cliniche;
- si possono incontrare difficoltà nel 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 diretta parte in causa di qualunque scelta assistenziale.

In definitiva, benché le valutazioni di appropriatezza ottenibili dal metodo RAND possano sembrare oggettive ed esplicite, il processo attraverso il quale vengono prodotte rimane per molti aspetti poco trasparente. La possibilità di migliorare questa metodologia, rendendo più espliciti i criteri di valutazione di appropriatezza e verificando la possibilità di applicarla anche con la partecipazione di figure non mediche è oggetto di ulteriore e futura ricerca. D'altra parte, uno degli aspetti importanti del metodo è quello di consentire non soltanto di elaborare raccomandazioni che indichino in quali specifici pazienti l'uso di una procedura sia o meno appropriato, ma anche di delimitare le aree – rappresentate dagli scenari in disaccordo o da quelli per i quali vi è stato un consenso nel giudicare dubbia l'opportunità di utilizzare la procedura – di incertezza entro la professione medica.

BOX 3.2 - IL METODO RAND E LA FAMIGLIA DI METODI PER LA COSTRUZIONE DI VALUTAZIONI COLLETTIVE (MODIFICATO DA FITCH ET AL., 2000)

Il metodo RAND per la valutazione di appropriatezza (RAND Appropriateness Method, RAM) è solo uno dei numerosi metodi che sono stati sviluppati per organizzare le opinioni collettive di esperti (Fink et al., 1984). 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), piuttosto che ottenere un consenso tra di loro. Il RAM è una variazione del metodo Delphi, anch'esso sviluppato presso la RAND negli anni '50 e applicato a temi politico-militari, tecnologici ed economici, come strumento per prevedere il futuro. (Linstone et al., 1975). Da allora, il processo Delphi 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 ai quesiti 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 i cui partecipanti vengono radunati in uno stesso luogo e viene loro chiesto di scrivere tutte le loro idee in merito a un determinato soggetto. Il moderatore (o conduttore) chiede a ciascuna persona di descrivere brevemente l'idea a suo giudizio più rilevante tra quelle che ha elencato; il giro di tavolo continua fino a quando tutti non si siano espressi. Dopo la discussione di ogni argomento, a ciascuno dei partecipanti viene chiesto di ordinare su una scala numerica, secondo importanza, le proprie valutazioni relative ai temi trattati. Infine, diverse tecniche matematiche vengono utilizzate per aggregare e analizzare i risultati.

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 (Kanouse, 1989). In pratica, la CC raccoglie una gamma eterogenea di soggetti interessati al tema oggetto di studio (medici,

ricercatori, cittadini, società scientifiche, ecc.) ai quali viene chiesto di sviluppare, in risposta a quesiti predefiniti sul tema affrontato, dichiarazioni che esprimano un accettabile livello di consenso reciproco.

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.

In conclusione il RAM consiste in una versione modificata del metodo Delphi ma, a differenza di esso, lascia ai componenti del panel l'opportunità di discutere le loro valutazioni tra un ciclo e l'altro. Contrariamente ai timori espressi da coloro che originariamente svilupparono il Delphi, l'esperienza condotta con RAM e la letteratura contemporanea sui processi di gruppo indicano entrambe che il rischio di distorsioni determinate da processi di comunicazione *face-to-face* (ad esempio, prevaricazione da parte di soggetti più esuberanti, mancata espressione di soggetti più remissivi, conflitti, ecc.) può essere largamente controllato da un'efficace conduzione del gruppo (Kahan et al., 1994). Quindi, mentre gli esperti del panel ricevono il feedback relativo alle risposte del gruppo (come nel classico metodo Delphi), hanno al contempo la possibilità di discutere direttamente le loro valutazioni, così come accade con la tecnica NGT e nella CC secondo il metodo dei NIH.

È importante che queste aree possano essere riconosciute, dal momento che esse rappresentano tutte quelle situazioni in cui mancano valide e sufficienti conoscenze per decidere se un intervento possa essere “sicuramente raccomandato” o “sicuramente non raccomandato”. In questo senso, una delle possibili implicazioni del metodo RAND è il suo impiego anche per identificare quelle aree verso le quali la ricerca medica dovrebbe essere indirizzata.

BOX 3.3 - LE TAPPE DEL PROCESSO DI VALUTAZIONE DELL'APPROPRIATEZZA SECONDO IL METODO RAND/UCLA (I) (MODIFICATO DA FITCH ET AL., 2000)

Identificare/selezionare il problema clinico. Il problema deve essere di tipo medico-tecnico. Aspetti di carattere economico e/o di organizzazione sanitaria, ad esempio, sono dichiaratamente esclusi. Nell'identificazione 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, morbidità), importanti variazioni geografiche, procedure il cui uso è controverso, buone probabilità terapeutiche o diagnostiche. Un altro importante aspetto da considerare è la qualità dell'evidenza scientifica disponibile: laddove la qualità sia bassa, l'affidabilità del lavoro del panel sarà inevitabilmente inferiore.

Revisione della letteratura scientifica. Molto è stato scritto sui metodi di revisione della letteratura scientifica e innumerevoli sono i riferimenti bibliografici. L'avvento della Cochrane Collaboration ha senz'altro introdotto sostanziali miglioramenti della qualità delle revisioni disponibili (Chalmers, 1993). Tuttavia è utile sottolineare come una revisione finalizzata alla valutazione di appropriatezza possa/debba essere basata su criteri di inclusione meno stringenti rispetto alle tradizionali revisioni Cochrane e comprendere, laddove siano assenti trial controllati, anche 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. Pur essendo ogni membro del panel libero di utilizzare o meno questa revisione, è ad essa che si farà riferimento in tutti i casi di disaccordo tra i vari componenti.

Costruzione degli scenari clinici (o "indicazioni"). È questa la caratteristica principale del metodo RAND; la lista delle "indicazioni" rappresenta il documento di lavoro principale per il panel. Consiste in una serie di descrizioni dettagliate di specifiche tipologie di pazienti, per ciascuno dei quali viene esplicitamente richiesto un giudizio sulla appropriatezza o meno dell'uso della procedura/intervento. Nella stesura della lista è importante assicurarsi che le indicazioni siano:

- esaustive, così da includere un'ampia popolazione di pazienti potenzialmente candidati alla procedura in esame;

- mutualmente esclusive, cosicché nessun paziente possa essere ricondotto a due scenari diversi;
- omogenee, così da consentire una uguale applicabilità della valutazione di appropriatezza a tutti i pazienti classificati in uno stesso scenario;
- gestibili, così da consentire ai componenti del panel di valutare tutte le indicazioni in un ragionevole intervallo di tempo. In diversi studi, il numero di indicazioni affrontate da un panel ha oscillato tra 200 (colecistectomia) e oltre 3000 (chirurgia della cataratta). Si concorda comunque sulla scarsa fattibilità di un processo di valutazione per un numero di indicazioni superiore a 2000.

Il panel di esperti. La costituzione del panel è naturalmente una fase cruciale e i criteri della selezione/nomina dei suoi componenti possono variare molto in ragione del contesto in cui il processo di valutazione si svolge. Negli studi condotti dal gruppo RAND negli Stati Uniti, gli esperti (in genere in numero compreso tra 9 e 11) sono stati spesso identificati attraverso il coinvolgimento delle Società scientifiche. Nella selezione dei partecipanti è comunque necessario seguire alcuni criteri di seguito indicati.

- *Multidisciplinarietà.* Per ottenere una valutazione equilibrata sull'appropriatezza dell'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 nei partecipanti disponibilità e motivazione ad eseguire i compiti loro affidati. È altrettanto importante che i clinici coinvolti siano disponibili a lavorare in gruppo.

BOX 3.4 - LE TAPPE DEL PROCESSO DI VALUTAZIONE DELL'APPROPRIATEZZA SECONDO IL METODO RAND/UCLA (II) (MODIFICATO DA FITCH ET AL., 2000)

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 appropriato (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 discussi 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 però la reciproca segretezza) e con il giudizio medio. La discussione si concentra sulle aree di disaccordo eventualmente emerse (disaccordo reale? Problemi di comprensione/interpretazione degli scenari?). 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 seconda valutazione degli scenari clinici.

Valutazione (eventuale) di necessità. In questo caso è necessario un terzo (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 prescrivesse, 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. Della classificazione nei tre livelli "appropriato", "inappropriato", "incerto" si è già detto nel testo. 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 si presentano spesso con situazioni complesse, per l'analisi delle quali è stata sviluppata una metodologia con basi statistiche che esulano dagli obiettivi di questo testo (il lettore interessato può consultare la versione originale del RAND/UCLA appropriateness method user's manual).

L'utilizzo dei risultati. I criteri di appropriatezza che si rendono disponibili alla fine del processo di valutazione descritto, possono essere usati in molti modi: per misurare in modo retrospettivo il sovrautilizzo di una procedura, per indagare in modo prospettico sia il sovrautilizzo sia il sottoutilizzo, per l'elaborazione di linee-guida, per la predisposizione di strumenti di supporto alle decisioni "sul campo" dei clinici, per l'identificazione di aree di incertezza che meritano l'avvio di ricerche cliniche mirate.

Il *RAND/UCLA appropriateness method user's manual* propone un interessante approfondimento metodologico per ciascuna di queste possibili applicazioni.

P a r t e s e c o n d a

APPROFONDIMENTO

4 • THE NHS ATLAS OF VARIATION IN HEALTHCARE: IL PROGRAMMA RIGHT CARE DEL DEPARTMENT OF HEALTH (UK)

Davide Botturi

“Nel luglio 2010 è stato pubblicato il White Paper inglese *Equity and Excellence: Liberating the NHS*, che dichiara un impegno esplicito ad aumentare il valore delle risorse allocate nel servizio sanitario nazionale, impegno che impone di affrontare con determinazione il problema dell’appropriatezza.”¹

L’*NHS atlas of variation in healthcare* si configura come uno strumento a supporto di coloro che sono coinvolti nella programmazione locale dei servizi sanitari (*local commissioner*) e nell’erogazione dell’assistenza sanitaria (*clinician*) per la comprensione delle variazioni, in particolare “ingiustificate”, riferite alla qualità dell’assistenza sanitaria e agli esiti delle cure. L’atlante presenta una serie di istogrammi e mappe, per Primary Care Trust (PCT), indicative di variazioni importanti che interessano il territorio nazionale, relative a problemi clinico-assistenziali identificati come rilevanti da un gruppo di esperti provenienti da diverse specialità.

¹ Dalla Prefazione di Sir Muir Gray and Philip DaSilva [Editors e Co-Leaders del QIPP Right Care Programme (www.rightcare.nhs.uk/atlas/qipp-nhsAtlas-LOW_26110c.pdf)] a *The atlas of variation in healthcare* (2010) (traduzione nostra).

OBIETTIVI E CONTENUTI DELL'ATLANTE

L'*NHS atlas of variation in healthcare* si colloca nell'ambito del programma nazionale inglese Right Care,² il cui obiettivo primario e generale per gli anni 2013-2014 consiste nel massimizzare il valore³ che:

- il singolo paziente trae dall'assistenza e dal trattamento ricevuti;
- l'intera popolazione trae dall'investimento in assistenza sanitaria.

Questo obiettivo viene perseguito attraverso una serie di azioni, tra cui il supporto a coloro che, coinvolti nella programmazione locale dei servizi sanitari (*local commissioners*) e nell'erogazione dell'assistenza sanitaria (*clinicians*), si trovano a dover comprendere le variazioni, in particolare "ingiustificate", osservate nelle misure di qualità dei servizi, di esiti delle cure e di spesa.

Le variazioni "ingiustificate" nel ricorso ai servizi sanitari vengono definite, in accordo con John Wennberg (Wennberg e Gittelsohn, 1973, 1982; Wennberg e Cooper, 1988; Wennberg, 2011), come 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".

² Right Care rientra, a sua volta, in un più ampio programma, il Quality Innovation Productivity and Prevention (QIPP).

³ La definizione di "valore" adottata è quella proposta da Michael Porter (2010). Lo studioso definisce il concetto di "valore" nell'assistenza sanitaria come gli esiti di salute raggiunti per dollaro speso. Secondo l'autore questo sarebbe ciò che conta per i pazienti e ciò che unifica gli interessi di tutti gli attori coinvolti nell'assistenza sanitaria. In particolare, gli esiti, il numeratore dell'equazione, sono da riferire ad uno specifico problema di salute e sono di natura multidimensionale; i costi, il denominatore dell'equazione, si riferiscono ai costi complessivi dell'intero percorso di cura del paziente, e non solo ai costi di un singolo servizio. Porter (ivi, p. 15) sostiene come sia fuorviante considerare in maniera separata il valore per l'individuo e quello per la popolazione. 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.

L'*NHS atlas of variation in healthcare* rappresenta il principale strumento utilizzato per descrivere e rappresentare la variabilità nell'assistenza sanitaria.⁴

La prima edizione è stata pubblicata nel 2010, la seconda nel 2011; dal 2012 l'orientamento è stato quello di approfondire e aggiornare alcuni specifici temi contenuti nelle precedenti edizioni (ad esempio, infanzia e adolescenza, diabete, disturbi respiratori, servizi diagnostici). I dati e le informazioni sono consultabili gratuitamente;⁵ è stata anche realizzata una piattaforma web interattiva.⁶

Numerose organizzazioni hanno collaborato e collaborano alla realizzazione dell'*NHS atlas of variation in healthcare*, tra cui: l'NHS PCT Network, l'NHS Alliance, il National Prescribing Centre e diversi osservatori di salute pubblica.

Di seguito vengono illustrate le caratteristiche principali della metodologia dell'*NHS atlas of variation in healthcare*, sostanzialmente identica per le varie edizioni. La principale differenza riguarda il numero e il tipo di misure incluse; in particolare, l'edizione 2011 presenta un numero superiore di misure rispetto all'edizione 2010 (71 vs 34) e, in alcuni casi, anche diversi tipi di misura riferita al medesimo problema di salute. Ad esempio, la misura "proporzione di persone diabetiche incluse nei nove processi assistenziali chiave raccomandati", dell'edizione 2010, è stata distinta sulla base delle due tipologie di diabete, I e II; o ancora, la "proporzione di decessi avvenuti in ospedale sul totale dei decessi avvenuti in una determinata area geografica" è stata sostituita con la misura complementare, la "proporzione di decessi avvenuti al domicilio o in una struttura residenziale".

La fonte della metodologia brevemente descritta di seguito è rappresentata dall'edizione 2010 dell'atlante.

⁴ 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.

⁵ Si rimanda all'indirizzo web: <http://www.rightcare.nhs.uk/index.php/nhs-atlas/>

⁶ Per accedere alla piattaforma web relativa all'edizione 2011 si rimanda all'indirizzo web: <http://www.sepho.org.uk/extras/maps/NHSAAtlas2011/atlas.html>

METODI: ALCUNI CENNI⁷

La selezione delle aree tematiche e delle rispettive misure è avvenuta 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 (ad esempio, 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, 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-ostruttiva), 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 (vedi figura 4.1);
- 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.

⁷ Per una trattazione più dettagliata degli aspetti metodologici si rimanda direttamente alle diverse edizioni dell'atlante.

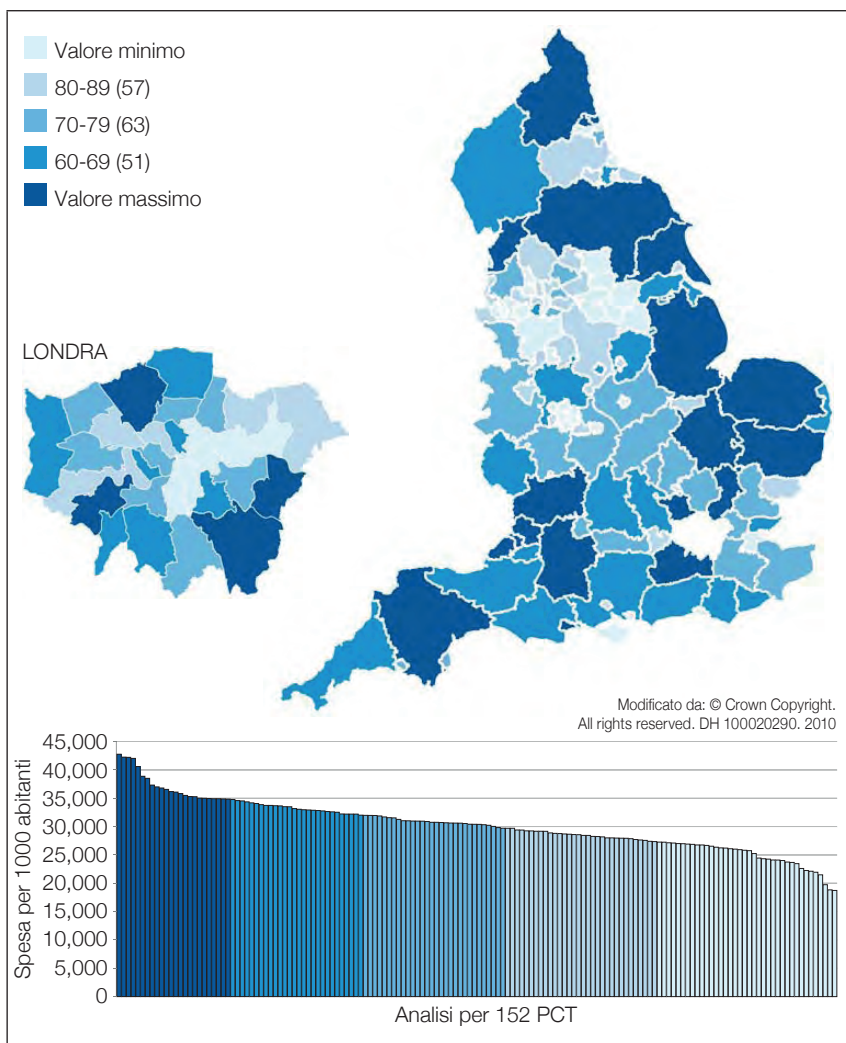


Figura 4.1 • Rappresentazione degli indicatori nelle singole schede dell'NHS Atlas of variation in healthcare. Esempio: spesa ospedaliera per pazienti oncologici per 1000 abitanti. Analisi per PCT. Aggiustamento per età, sesso e bisogni (modificata da NHS atlas of variation in healthcare, 2010)

Negli indicatori dell'atlante il denominatore è rappresentato dalla popolazione residente in una determinata area geografica. La motivazione di tale scelta è riconducibile alla tipologia di destinatari a cui il lavoro si rivolge, soprattutto i committenti dei servizi e coloro che pianificano e allocano le risorse dedicate all'assistenza sanitaria, oltre che i clinici.

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.

Nelle mappe, i confini delle aree geografiche corrispondono nella maggior parte dei casi a quelli dei PCT (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 viene dedicato alla città di Londra.

Negli istogrammi ogni barra corrisponde ad un PCT, 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 (ad esempio, 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

Le misure sono state selezionate per questa pubblicazione sulla base della potenziale rilevanza informativa e decisionale per il medico di medicina generale, o perché di interesse generale, o per la pertinenza a specifici programmi sviluppati o in corso a diversi livelli istituzionali del sistema sanitario nazionale italiano. In particolare:

- proporzione di persone con diabete di tipo 2, nel National Diabetes Audit, che hanno ricevuto i nove processi assistenziali chiave raccomandati;
- proporzione 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 *1000 abitanti;
- proporzione di decessi avvenuti in ospedale (sul totale dei decessi avvenuti in una Local Authority).

Ciascuna scheda ripropone la struttura e i contenuti delle schede-indicatore (contesto, ampiezza della variazione, alternative per l'azione) dell'atlante. Per semplicità, le schede da noi esaminate nei box 4.1-4.4 riportano soltanto l'istogramma.

La fonte dei dati è rappresentata dall'edizione 2011 dell'atlante (per gli indicatori 1 e 2) e dall'edizione 2010 (per gli indicatori 3 e 4).

BOX 4.1 - PROPORZIONE DI PERSONE CON DIABETE DI TIPO 2, NEL NATIONAL DIABETES AUDIT, CHE HANNO RICEVUTO I NOVE PROCESSI ASSISTENZIALI CHIAVE RACCOMANDATI. ANALISI PER PRIMARY CARE TRUST (PCT) (GENNAIO 2009-MARZO 2010)

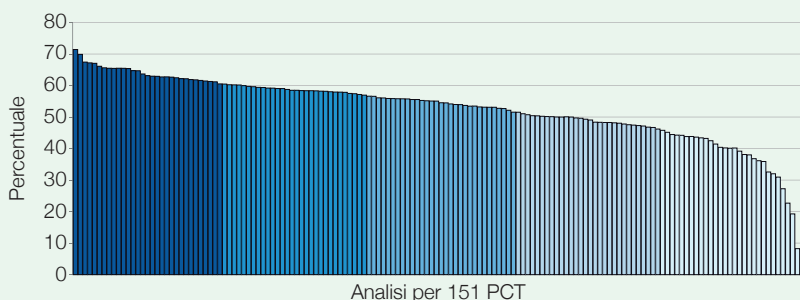
Contesto. In accordo con le raccomandazioni del National Institute for Health and Clinical Excellence (NICE), le persone sofferenti di diabete, tipo 1 e tipo 2, dovrebbero essere sottoposte a nove controlli chiave per la gestione e il monitoraggio del diabete: HbA1c, colesterolo, creatinina, micro-albumina, pressione arteriosa, peso corporeo, abitudine al fumo, occhio, piede.

Si stima che in Inghilterra solo il 52,9% delle persone con diabete tipo 2 abbia ricevuto i nove controlli nel periodo gennaio 2009-marzo 2010.

La fonte informativa dell'indicatore è rappresentata dal National Diabetes Audit che, per gli anni 2009-2010, include 1.997.646 persone con diabete: 1.762.184 con tipo 2 e 185.962 con tipo 1. Negli stessi anni, 6700 GP practice hanno partecipato all'Audit, con un tasso di partecipazione del 76% (6700 su 8842 GP practice)

Ampiezza della variazione. Tra i PCT, il valore massimo della proporzione di diabetici di tipo 2 che hanno ricevuto tutti e nove i processi di cura raccomandati dal NICE è pari al 71,4%, percentuale 10 volte superiore a quella del valore minimo, pari al 7%. Se si escludono i 5 PCT con le percentuali più elevate e i 5 PCT con le più basse, l'ampiezza diminuisce a 2,1 volte. Nelle persone con diabete di tipo 1 tale ampiezza è lievemente superiore, pari a 2,6 volte.

Si rileva una forte associazione (coefficiente di correlazione = 0,798) tra la proporzione di diabetici di tipo 2 che hanno ricevuto tutti e nove i controlli raccomandati nel 2008-2009 e quella riferita al 2009-2010, suggerendo che la variazione è persistente nel tempo. Non si rileva una correlazione statisticamente significativa tra l'indicatore



e il livello di deprivazione per PCT. Entrambi questi risultati consentono di ipotizzare che il livello di variabilità osservato possa essere legato alle modalità di organizzazione dei servizi.

Indicazioni per l'azione. Le azioni necessarie per ridurre tale variabilità riguardano, ad esempio: l'impiego di sistemi amministrativi in grado di invitare le persone ai controlli annuali; il ricorso a processi di follow-up; l'organizzazione di modalità di accesso ai controlli che siano convenienti per le persone; la registrazione accurata dei risultati dei controlli al fine di avere la piena garanzia che il controllo prenotato sia stato effettuato.

BOX 4.2 - PROPORZIONE DI PERSONE RICOVERATE IN OSPEDALE A SEGUITO DI STROKE, CHE HANNO TRASCORSO IL 90% DEL TEMPO IN UNA STROKE UNIT. ANALISI PER PCT (GENNAIO-MARZO 2011)

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 PCT, il valore massimo delle persone colpite da stroke che hanno trascorso il 90% del tempo presso una stroke unit, è pari al 100%, percentuale che è 3,2 volte superiore al valore minimo, pari al 31,5%. Se si escludono i 5 PCT con i valori più elevati e i 5 PCT con i valori più bassi, l'ampiezza diminuisce a 1,8 volte. Essa si presenta inferiore rispetto all'ampiezza rilevata negli anni 2009-2010 pari ad almeno 3 volte.

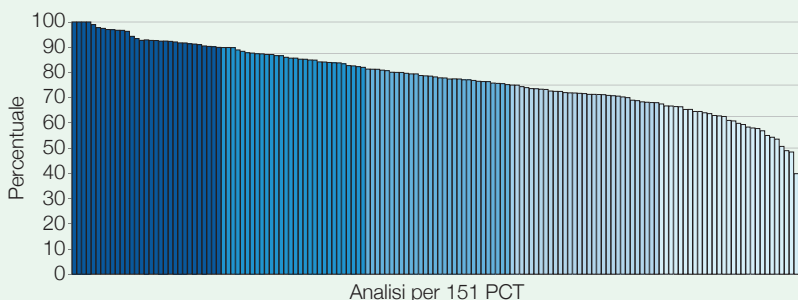
Sulla base dei risultati emersi dal National Sentinel Audit of Stroke (anno 2010), il Royal College of Physicians riferisce che la maggioranza dei pazienti (il 57%) non viene ancora ricoverata presso stroke unit, ma presso altri reparti, e ritiene davvero deludente il fatto che solo il 36% dei pazienti venga ricoverato direttamente presso una stroke unit e soltanto il 38% entro le prime 4 ore dall'arrivo in ospedale.

Indicazioni per l'azione. Le azioni chiave necessarie a migliorare l'accesso alle stroke unit consistono in:

- monitoraggio continuo (24 ore) dei parametri fisiologici;
- 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.

Secondo lo Stroke Improvement Programme, gli interventi efficaci includono:

- vincolare l'utilizzo dei posti letto nelle stroke unit solamente ai pazienti colpiti da stroke;
- lavorare in collaborazione con le ambulanze;
- passare ad una terapia di sei giorni settimanali.



BOX 4.3 - TASSO STANDARDIZZATO (METODO DIRETTO) DI RICOVERO IN URGENZA DI ANZIANI >75 ANNI *1000 ABITANTI. ANALISI PER PCT (ANNI 2008-09)

Contesto. Agli anziani di età >75 anni, specialmente se sofferenti di due o più patologie croniche, sono riconducibili:

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

Alcuni ricoveri programmati possono portare dei benefici agli anziani "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 in alcuni casi a una perdita irreversibile 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 1000 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 e i 5 PCT con i tassi più bassi, questo ordine di grandezza si mantiene inalterato.

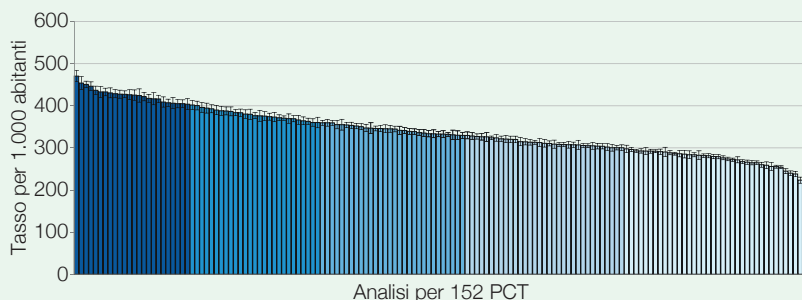
Si può osservare un ulteriore elemento di variabilità esaminando i dati della Care Quality Commission relativi agli esiti dei ricoveri in urgenza: la variazione nella percentuale di anziani >85 anni presi in carico con l'assistenza domiciliare dopo un ricovero in urgenza è dell'ordine di 1:4, ovvero la percentuale più elevata di anziani inseriti nel programma di assistenza domiciliare, pari al 20%, è quattro volte superiore alla 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 caregiver,

al coinvolgimento dei medici di medicina generale. A questo 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.



BOX 4.4 - PROPORZIONE DI DECESSI AVVENUTI IN OSPEDALE, SUL TOTALE DEI DECESSI RIFERITI A UN'AREA GEOGRAFICA DEFINITA, NELLO STESSO INTERVALLO DI TEMPO. ANALISI PER LOCAL AUTHORITY (LA) (ANNI 2006-08)

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 possibile, di coloro che affrontano la fase terminale della vita.

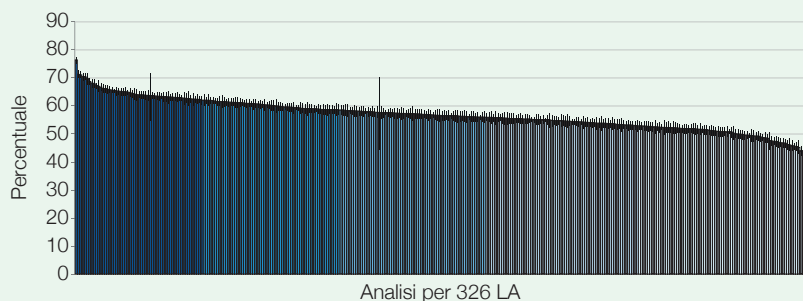
Ampiezza della variazione. Nel triennio 2006-2008 il 58% dei decessi area geografica-specifici è avvenuto in ospedale. L'intervallo di variazione tra le LA è compreso tra un valore massimo pari al 76% di decessi avvenuti in ospedale sul numero totale di decessi in quella determinata area e un valore minimo pari al 44%. Nella metà di tutte le LA 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

l'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 diversi soggetti e agenzie, incluso il Servizio sanitario nazionale, i servizi sociali delle LA, gli hospice, le charities. I committenti dovrebbero valutare il modo migliore di coordinare questa tipologia di assistenza.

Più dell'80% dei decessi coinvolge gli anziani (≥ 65 anni) che, con maggiore probabilità, soffrono al momento della morte di patologie multiple e croniche. È necessario assicurare supporto agli anziani sofferenti di problemi di salute multipli in contesti assistenziali diversi dall'ospedale, come le case protette e le strutture residenziali (oltre, naturalmente, alla propria dimora nei casi in cui questo è possibile). Ogni persona per la quale la morte non rappresenti un evento inatteso dovrebbe poter contare su un piano di assistenza che tenga conto dei suoi bisogni e delle sue preferenze riguardo alla fine della vita. Il 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.



5 • THE DARTMOUTH ATLAS PROJECT

Davide Botturi

Da più di vent'anni il Dartmouth Atlas Project documenta le variazioni nella distribuzione e nell'utilizzo delle risorse sanitarie (spesa, strutture, professionisti, interventi) all'interno degli Stati Uniti. I risultati di questi studi hanno supportato e supportano tutt'ora i *policy-makers*, gli analisti, i *media* e tanti altri nell'accrescere la comprensione della efficienza ed efficacia nel funzionamento del sistema sanitario americano e rappresentano una base per molte delle azioni in corso finalizzate a migliorare la salute e il sistema sanitario in America.

Il Progetto ha sede presso il Dartmouth Institute for Health Policy and Clinical Practice (Lebanon, New Hampshire) ed è finanziato da un ampio gruppo di fondazioni no profit guidate dalla Robert Wood Johnson Foundation e da enti nazionali, quali il National Institutes of Health e il National Institute of Aging. Il fondatore dell'iniziativa è John Wennberg, che dal 1973¹ ha dedicato buona parte della sua attività di ricerca alla comprensione della variabilità nella pratica clinica e i cui lavori sono stati e sono tutt'ora fonte di ispirazione a livello internazionale. Tra questi si cita l'*NHS atlas of variation in healthcare* (Regno Unito), a cui è dedicato l'approfondimento nel precedente capitolo 4.

L'assunto da cui prende le mosse il Dartmouth Atlas Project consiste nel riconoscere come la spesa sanitaria stia assorbendo una

¹ In quell'anno è stato pubblicato il suo primo articolo dedicato al tema delle variazioni nell'erogazione dell'assistenza sanitaria (Wennberg e Gittelsohn, 1973).

proporzione crescente del PIL, passando, ad esempio, dal 15,8% nel 2005 al 17,7% nel 2011 (la più elevata dei Paesi OCSE); a fronte di tale aumento di spesa sono scarse le evidenze circa eventuali benefici aggiuntivi per la salute della popolazione (OCSE, 2013). A questo proposito, i ricercatori di diverse organizzazioni (Dartmouth Institute for Health Policy and Clinical Practice, New England Healthcare Institute, McKinsey, Thomson Reuters) stimano che il 30% circa della spesa sanitaria corrente sia “non necessaria”.²

In particolare, il gruppo di ricerca del Dartmouth Atlas Project ha come obiettivo principale quello di spiegare le cause e le conseguenze della variabilità nell’assistenza sanitaria, offrendo un ampio panorama di informazioni e di analisi a livello nazionale, regionale e locale, così come a livello di singolo ospedale e singolo medico. A titolo esemplificativo, alcune domande a cui i ricercatori tentano di rispondere sono:

- Perché alcuni medici prescrivono un numero di tomografie assiali computerizzate doppio rispetto ad altri colleghi?
- Perché i tassi di impianto degli stent coronarici sono tre volte superiori in Elyria (Ohio) rispetto alla vicina Cleveland, dove si trova la famosa Cleveland Clinic?
- Che cosa significano queste differenze per i pazienti?

L’intento è quello di perfezionare la capacità di descrivere e capire che cosa attualmente succede ai pazienti e che cosa può essere fatto per migliorare l’assistenza, impegno che oggi deve essere preso alquanto “seriamente” (Pronovost e Goeschel, 2011).³

² L’approccio Dartmouth consiste nello stimare il risparmio possibile se tutte le Regioni erogassero lo stesso livello di assistenza fornito dalle Regioni che, pur con una spesa contenuta, garantiscono un’assistenza equa e di qualità. I ricercatori del Dartmouth Atlas Project ritengono, inoltre, che la stima della spesa “non necessaria”, pari a circa il 20-30%, rappresenti una sottostima del fenomeno.

³ Secondo Pronovost e Goeschel è di primaria importanza lo sviluppo della ricerca sull’erogazione dell’assistenza per diverse ragioni: i pazienti ricevono la metà delle terapie raccomandate; il costante incremento dei costi dell’assistenza non è più sostenibile; sono sempre più noti i rischi prevenibili per i pazienti; alle organizzazioni che producono assistenza viene chiesta una maggiore responsabilità sul modo in cui impiegano le risorse economiche per garantire determinati esiti ai pazienti; l’esplosione nella ricerca di base e clinica (vengono pubblicati ogni anno circa 18.000 clinical trial) rende praticamente impossibile per un clinico ricordare tutte le conoscenze più recenti sulle terapie.

La popolazione di riferimento degli studi è rappresentata dai beneficiari (età ≥ 65 aa.) di *Medicare*, programma federale di assicurazione per l'assistenza sanitaria. La popolazione anziana, e in particolare il suo impatto sulla sostenibilità dei sistemi di *welfare* e specialmente per l'assistenza sanitaria e sociale pubblica, sono peraltro al centro del dibattito internazionale di questi ultimi anni.⁴

Una delle principali ragioni che ha spinto gli studiosi a concentrarsi su questo target di popolazione è l'entità e la qualità dei dati disponibili. Il Centers for Medicare and Medicaid Services, l'agenzia federale statunitense che raccoglie i dati di ciascuna persona e di ciascun erogatore di assistenza inseriti nel programma Medicare, rende infatti disponibile un database nazionale e uniforme a fini di ricerca. Non ci sono database corrispettivi per la popolazione <65 anni. Peraltro, dai risultati di studi condotti sempre dai ricercatori del Dartmouth Atlas Project in alcuni Stati americani (Michigan, Pennsylvania, Virginia) emerge come le differenze nell'assistenza sanitaria erogata alla popolazione <65 anni siano simili a quelle che interessano la popolazione ≥ 65 anni.

BREVI CENNI SU METODO, STRUMENTI E PRODOTTI

Analizziamo sinteticamente alcuni degli aspetti salienti presenti nel sito web del Dartmouth Atlas Project.

Un primo aspetto riguarda la metodologia di analisi della variabilità e di rappresentazione dei dati. Essa è comunemente conosciuta come *small area analysis* (Wennberg, 1973) e si focalizza sull'espe-

Pertanto, è necessaria una scienza che, con un punto di vista sistemico, cerchi di legare maggiormente i risultati della ricerca clinica a quelli di salute dei pazienti, garantendo che questi ultimi ricevano le terapie raccomandate al minor rischio. Questo implica un cambiamento nel comportamento dei clinici, rimuovere le barriere che ostacolano tali cambiamenti, lavorare sulla cultura e sul teamwork, misurare e offrire il feedback dei risultati raggiunti, creare gli incentivi per i clinici ed i pazienti affinché assumano i comportamenti desiderati.

⁴ Si citano a titolo esemplificativo: European Commission. European Innovation Partnership on Active and Healthy Aging (http://ec.europa.eu/research/innovation-union/index_en.cfm?section=active-healthy-ageing); OECD (2011); US Department of Health and Human Services; Healthy People 2020: Older Adults (<http://www.healthypeople.gov/2020/topicsobjectives/topic/older-adults?topicid=31>).

rienza della popolazione che risiede in una definita area geografica o che riceve i servizi di uno specifico ospedale (sia esso generale o punto di riferimento per il trattamento di determinati problemi di salute).

Un secondo aspetto riguarda gli strumenti (*tools*) interattivi offerti dal sito web, tra cui la possibilità di effettuare azioni di *benchmarking*. Questo strumento permette di selezionare l'ambito di interesse, il tipo di servizio, l'indicatore, il riferimento temporale (singolo anno, arco temporale), le aree geografiche o gli ospedali che si intende confrontare, e il/le aree o il/gli ospedali che fungono da *benchmark*.

Un terzo punto riguarda la serie di innumerevoli pubblicazioni consultabili gratuitamente dal sito. Esse comprendono i celebri Dartmouth Atlas of Health Care (prodotti dal 1996), diversi report dedicati a tematiche specifiche (tra le più recenti: assistenza ai malati cronici, assistenza di fine vita, *patient decision-making*, riammissioni) e articoli scientifici pubblicati su riviste internazionali. Uno dei vari report viene qui di seguito sinteticamente discusso, a scopo esemplificativo.

RIFLESSIONI SULLA VARIABILITÀ

I dubbi e la confusione che alcuni risultati⁵ dei lavori del gruppo di ricerca del Dartmouth Atlas Project hanno generato specialmente nel dibattito pubblico, dove può essere difficile sintetizzare in maniera comprensibile ricerche complesse, hanno condotto gli studiosi a raccogliere e riassumere le principali criticità emerse, cercando di formulare sintetiche risposte di chiarimento. Nelle pagine che seguono si riporta una selezione di quesiti e relative risposte.⁶ Si tratta di osservazioni e riflessioni sulla variabilità dell'assistenza derivanti dalla lunga esperienza maturata dai ricercatori del Dartmouth Atlas Project e che, per la rilevanza delle questioni affrontate, trascendono i confini

⁵ Altri risultati degli studi condotti nell'ambito del Dartmouth Atlas Project sono stati, invece, ampiamente accettati e già tradotti nella pratica.

⁶ La fonte delle informazioni è rappresentata dalle sezioni del sito web del Dartmouth Atlas of Health Care (aggiornato al 2014) dedicate a: *reflections on variations* (<http://www.dartmouthatlas.org/keyissues/issue.aspx?con=1338>); *frequently asked questions* (<http://www.dartmouthatlas.org/tools/faq/>).

nazionali americani e si pongono come trasversali ai singoli ambiti di studio indagati.

Quali fattori sono in grado di spiegare la variabilità geografica nella spesa sanitaria (di Medicare)? Qual è il contributo del reddito e dello stato di salute?

I risultati delle ricerche, non solo quelle condotte dagli studiosi del Dartmouth Atlas Project, sottolineano come una parte rilevante della variabilità geografica nella spesa sanitaria procapite, nell'allocazione delle risorse e nell'utilizzo dei servizi, non sia spiegabile sulla base della condizione di salute e delle preferenze del paziente e della popolazione. In particolare, tale variabilità sarebbe caratterizzata da un sovrautilizzo di *supply* e *preference sensitive care* e da un sottoutilizzo di *effective care*. Wennberg (2011) fornisce una definizione delle tre tipologie di assistenza. La *supply sensitive care* comprende quelle attività cliniche, come le visite mediche, gli esami diagnostici e i ricoveri ospedalieri, la cui frequenza d'uso è legata alla capacità di offerta del sistema sanitario locale, in termini di numero di strutture, posti letto, strumenti e professionisti.

La *preference sensitive care* riguarda le situazioni in cui è disponibile più di un'opzione di trattamento accettata scientificamente – come nella chirurgia elettiva – e dove la frequenza d'uso dovrebbe dipendere dalla decisione informata del paziente, piuttosto che dalle opinioni del professionista.

La *effective care* riguarda tutti quegli interventi nei quali i benefici per la salute del paziente superano i rischi e che pertanto dovrebbero essere offerti al 100% dei pazienti che presentano le caratteristiche indicate dalle linee-guida.

Secondo i risultati di uno studio pubblicato sul *New England Journal of Medicine* (Sutherland et al., 2009), che ha coinvolto un campione di 15 mila beneficiari (≥ 65 aa.) del programma Medicare, una condizione economica di povertà spiegherebbe soltanto il 4% della variabilità geografica nella spesa sanitaria. Queste evidenze hanno sollevato diverse critiche, secondo le quali il reddito avrebbe un peso maggiore. Uno dei limiti, tuttavia, di tali osservazioni è il mancato controllo di alcuni importanti fattori che condizionano le

variazioni e che sono spesso associati alla povertà, quali l'età molto avanzata (grandi anziani) e lo stato di salute. A questi fattori si deve aggiungere, inoltre, la difficoltà nell'accesso ai servizi, e quindi una minore capacità di influenzare la spesa, che le persone con reddito relativamente basso possono incontrare.

Sempre secondo i risultati dello studio pubblicato sul *New England Journal of Medicine* lo stato di salute, rilevato mediante questionario (salute percepita), è il fattore che più del reddito e dell'etnia di appartenenza contribuisce a spiegare la variabilità geografica nella spesa sanitaria, con una percentuale pari al 18%. Un altro studio condotto mediante MedPAC analysis (Skinner e Fisher, 2010), suggerirebbe che il contributo della condizione di salute/malattia è maggiore di quello precedentemente rilevato. La metodologia MedPAC utilizza una misura sviluppata da Medicare, basata sul numero di differenti diagnosi ricevute da un paziente nel corso di un anno, assumendo che, più numerose sono le diagnosi, pesate a seconda del significato di ciascuna, e peggiore è lo stato di salute. Tale metodologia presenta, tuttavia, un bias ampiamente conosciuto. La misura utilizzata nella MedPAC analysis risente infatti dello stile lavorativo del medico e della struttura: nelle aree geografiche dove ospedali e medici praticano un'assistenza maggiormente "aggressiva" aumentano le probabilità che a una persona venga diagnosticato un determinato disturbo, e una delle caratteristiche delle aree geografiche con la spesa sanitaria più elevata consiste proprio nella tendenza ad effettuare maggiori esami, aumentando la probabilità di ricevere una determinata diagnosi, rispetto alle aree con una spesa inferiore.

Una possibile strada da percorrere nel cercare di comprendere il contributo dello stato di salute nello spiegare la variabilità osservata consiste nell'esplorare metodi diversi di *risk adjustment*. Uno dei più promettenti sembra essere quello di focalizzarsi su coorti di pazienti con diagnosi simili.

In conclusione, con le metodologie attuali, oltre il 70% delle differenze geografiche nella spesa sanitaria sembra non trovare spiegazione.

Ci sono evidenze che supportano la tesi secondo cui spendere di più conduce a migliori risultati di qualità dell'assistenza e di salute dei pazienti?

Una notevole mole di studi condotti dai ricercatori del Dartmouth Atlas Project dimostra come una maggiore spesa sanitaria non comporti necessariamente una migliore salute per i pazienti. Tra questi vale la pena citare lo studio di Fisher et al., (2003a, 2003b), condotto su dati nazionali e su un campione di milioni di persone, nel quale si è cercato di studiare che cosa accadeva ai pazienti ricoverati in ospedale per una delle seguenti tre condizioni: infarto miocardico acuto, frattura dell'anca e cancro del colon retto (in questa ultima categoria sono stati inclusi solo i pazienti ricoverati per sottoporsi ad intervento chirurgico).

Le principali conclusioni della ricerca sono state le seguenti:

- i pazienti trattati nelle aree geografiche con la spesa e l'utilizzo dei servizi più elevati avevano ricevuto circa il 60% di interventi assistenziali in più dopo il primo anno successivo al loro iniziale ricovero, non avendo ottenuto migliori esiti di salute rispetto ai pazienti trattati in aree con una spesa e un utilizzo minori. In altre parole “di più non è necessariamente meglio”;
- nelle aree geografiche con la spesa più elevata, la proporzione di spesa aggiuntiva era ampiamente assorbita da un insieme di ricoveri ospedalieri, consulenze ed esami specialistici che i medici delle aree con una spesa più bassa non avevano prescritto.

Altri studi dimostrano, invece, come, in determinate condizioni e situazioni, “di più è meglio”. Ad esempio, dai risultati di un recente studio del MIT Sloan School of Management (Doyle, 2010) è emerso come i turisti che avevano avuto accesso al pronto soccorso di un'area geografica con un'elevata spesa sanitaria presentavano un maggiore tasso di sopravvivenza rispetto ai turisti con accesso al pronto soccorso di un'area geografica con una spesa sanitaria meno elevata.

Se si prendono in considerazione alcuni problemi di salute, come l'infarto miocardico acuto o lo scompenso cardiaco congestizio, la letteratura scientifica appare contrastante; alcuni studi infatti avva-

lorano la tesi secondo cui una maggiore spesa e consumo di servizi porta a migliori esiti di salute, mentre altri la smentiscono.

In definitiva, una conclusione ragionevole consiste nel riconoscere come un'ampia parte della variabilità nella spesa e nell'utilizzo dell'assistenza non sia associata ad esiti positivi di salute e come vi siano aree geografiche e ospedali che garantiscono eccellenti risultati a basso costo, rappresentando importanti esperienze da cui apprendere.

Com'è possibile convincere le persone a non richiedere assistenza “aggiuntiva”, ovvero non necessaria rispetto alla loro condizione di salute? E come è possibile fare in modo che non siano i medici a raccomandarla?

Dai risultati delle ricerche condotte dagli studiosi del Dartmouth Atlas Project emerge come tre quarti degli americani riferiscano di avere rifiutato interventi consigliati dai rispettivi medici, in quanto li ritenevano non necessari o ritenevano che i benefici non superassero i rischi o i possibili effetti avversi. Altri studi hanno evidenziato che i pazienti informati preferiscono una quantità di interventi chirurgici inferiore, in media, a quella che i chirurghi sono inclini a praticare o proporre. Migliorare la comprensione e la consapevolezza dei pazienti circa i rischi ed i *trade-off* associati alle scelte di trattamento rappresenta un “buon” modo per ridurre la domanda, ad esempio, di ricoveri ospedalieri, di esami diagnostici non necessari e di interventi chirurgici, quando esistono altre opzioni.

Nel caso dei medici, il sistema di incentivi può essere una strategia con cui orientare il loro comportamento professionale verso la riduzione di quelle prescrizioni di esami e interventi non necessari.

Dai risultati delle ricerche emerge come un'assistenza “maggiormente aggressiva” nei confronti di pazienti sofferenti di patologie croniche non necessariamente prolunghi l'aspettativa di vita o ne migliori la qualità. Non si dovrebbe fare tutto ciò che si può per salvare una vita?

Dalle ricerche emerge come interventi particolarmente “aggressivi” su pazienti che soffrono di patologie croniche riducono l'aspettativa di vita, probabilmente a causa dei rischi associati all'ospedalizzazione.

zazione. Ciò suggerisce che la migliore strategia per prolungare una vita di qualità alle persone con disturbi cronici consista nell'orientarsi verso quelle attività che producono un beneficio dimostrato per la sopravvivenza – ad esempio un migliore controllo della pressione sanguigna nelle persone con il diabete – piuttosto che verso una “eroica” assistenza di fine vita.

ESEMPIO

A scopo esemplificativo viene, di seguito, sinteticamente descritto uno studio dedicato alla qualità dell'assistenza primaria (Goodman et al., 2010) e viene analizzato uno degli indicatori utilizzati. Lo studio è stato selezionato sulla base della potenziale rilevanza per il medico di medicina generale e, più in generale, per il sistema sanitario italiano ai diversi livelli istituzionali (locale, regionale e nazionale). La pubblicazione di Goodman et al. si articola in 5 parti:

1. Nel *summary* sono riportati i principali risultati emersi, tra cui:
 - una marcata variabilità tra le diverse aree geografiche in termini di accesso e utilizzo dell'assistenza primaria da parte dei pazienti, così come della loro probabilità di essere ricoverati in ospedale. Il contributo principale nello spiegare tale variabilità deriverebbe dalle differenze geografiche, piuttosto che dall'etnia (bianchi/neri) dei pazienti, dimostrando in altre parole che “il luogo in cui i pazienti vivono ha un'influenza maggiore sull'assistenza ricevuta rispetto al colore della pelle”;
 - una relazione complessa tra offerta di medici e accesso all'assistenza; in altre parole, non necessariamente a una maggiore concentrazione di medici corrisponde un maggiore accesso dei pazienti alle cure e viceversa.
2. La seconda parte esplora in maniera complessiva le eventuali relazioni nell'ambito dell'assistenza primaria tra offerta (misurata principalmente come numero di professionisti dedicati), accesso (mi-

surato come proporzione della popolazione che ha effettuato almeno una visita ambulatoriale), qualità e outcome, dimensioni che vengono approfondite nella terza e quarta parte.

3. La terza parte analizza due tematiche relative alla qualità dell'assistenza: il diabete e lo screening per la prevenzione del cancro alla mammella, misurati rispettivamente come proporzione di diabetici 65-75 aa. beneficiari di Medicare sottoposti al controllo dell'emoglobina A1c e proporzione media annua di donne 67-69 aa. beneficiarie di Medicare sottoposte ad almeno una mammografia nel corso di due anni.

4. La quarta parte approfondisce due tematiche relative agli outcome: l'amputazione della gamba – come esito avverso di una non appropriata gestione territoriale di patologie quali il diabete e il disturbo vascolare periferico – e i ricoveri ospedalieri per condizioni mediche rispetto alle quali un'appropriata gestione territoriale può prevenirne l'ospedalizzazione (*ambulatory care-sensitive conditions*), misurati rispettivamente come tasso medio annuo di amputazione della gamba e tasso di ospedalizzazione per *ambulatory care-sensitive conditions*,⁷ entrambi calcolati per 1000 beneficiari di Medicare (cfr. box 5.1).

5. La quinta parte illustra un caso-studio condotto ad Atlanta (Georgia).

La popolazione di riferimento dello studio è rappresentata dal totale dei beneficiari (età ≥ 65 aa.) del programma pubblico assicurativo per l'assistenza Medicare o da una sotto-popolazione di essi, come nel caso dei diabetici. L'area geografica è prevalentemente definita come Hospital Referral Region (HRR). Si tratta di un'area che circo-

⁷ L'americana Agency for Healthcare Research and Quality (12 marzo 2007) ha identificato, attraverso un set di misure denominate "Prevention Quality Indicators", una serie di condizioni mediche per le quali l'ospedalizzazione può essere prevenuta, ad esempio: diabete, disturbo polmonare cronico-ostruttivo, polmonite batterica, asma, disturbo cardiaco congestizio, ipertensione, infezione del tratto urinario, altro. La Regione Emilia-Romagna ha, peraltro, assegnato alle proprie aziende sanitarie il monitoraggio di alcuni di questi indicatori come obiettivi per gli anni 2010-2011.

scrive il percorso dei pazienti verso un ospedale specialistico (specializzato nel trattamento di una casistica complessa come, ad esempio, i disturbi cardiovascolari o neurologici), collocabile al terzo livello dell'assistenza (se per primo livello si definisce l'assistenza di base e per secondo livello l'assistenza ospedaliera generale). Negli Stati Uniti sono state identificate in tutto 306 Hospital Referral Regions e ciascuna include almeno un ospedale specialistico e diversi ospedali generali (vedi box 5.1).

**BOX 5.1 – OSPEDALIZZAZIONE PER AMBULATORY CARE-SENSITIVE CONDITIONS.
ANALISI PER HOSPITAL REFERRAL REGIONS (HRR) (ANNI 2003-2007)**

Contesto. Molti ricoveri ospedalieri avvengono per condizioni mediche – come ad esempio diabete poco controllato o peggioramento successivo allo scompenso cardiaco – che possono essere trattate in un setting assistenziale diverso da quello ospedaliero (assistenza territoriale) e per le quali l'ospedalizzazione può essere evitata con una migliore gestione territoriale. Sebbene il ricovero in ospedale possa apparire più sicuro per il paziente, possa facilitare il medico nel tenere traccia del paziente, oppure possa essere l'unica opzione nel caso di una scarsa disponibilità di alternative di cura (ad esempio, domicilio, rete di servizi territoriali), nel caso di determinate condizioni cliniche esso non dovrebbe rappresentare la primaria e migliore scelta. L'ospedale sottopone il paziente a diversi rischi, incluse le infezioni e gli errori, e rappresenta per i familiari del paziente e la società un sostanziale costo, non solo di natura economica. Per determinate condizioni cliniche il ricorso all'ospedalizzazione è quindi parzialmente determinato o sensibile alla disponibilità e qualità dell'assistenza ambulatoriale (territoriale).

Variabilità geografica. Negli Stati Uniti il tasso di ospedalizzazione per condizioni sensibili all'assistenza ambulatoriale ha subito un decremento pari a circa l'11% nel periodo compreso tra il 2003 ed il 2007, come conseguenza di un maggiore investimento del sistema sanitario sull'assistenza ambulatoriale. Tale diminuzione non è tuttavia uniforme, si rileva un'ampia variabilità tra le HRR. Considerando la media degli anni 2003-2007, la HRR con il tasso di ospedalizzazione più elevato è risultata essere Monroe (Louisiana), con 136 dimessi ogni mille beneficiari di Medicare, tasso quattro volte superiore a quello più basso, pari a circa 30 dimessi ogni mille beneficiari di Medicare, corrispondente alla HRR di Honolulu (figura 1). Sempre rispetto agli anni 2003-2007 è stata osservata una media nazionale pari a 76 dimessi ogni mille beneficiari di Medicare.

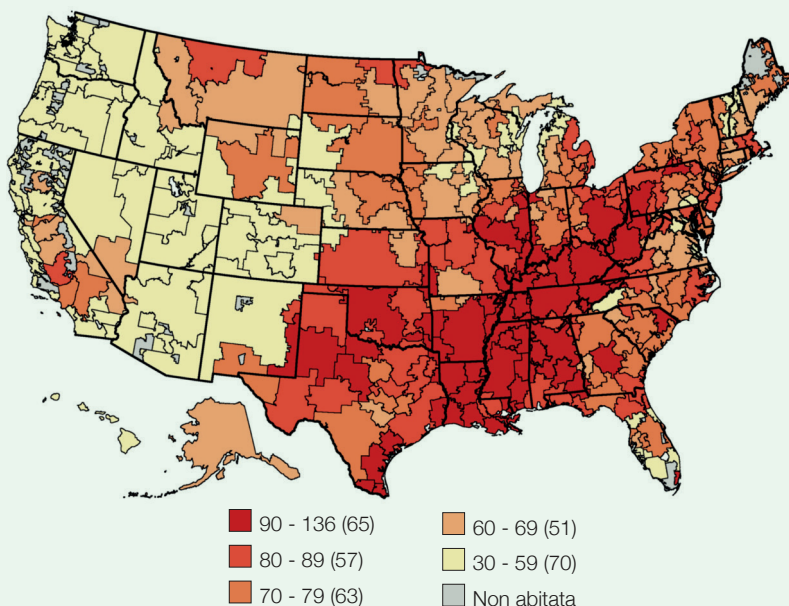


Figura 1 • Tasso di ospedalizzazione per *ambulatory care-sensitive conditions* *1.000 beneficiari di Medicare. Il tasso è standardizzato per età, sesso ed etnia attraverso metodo indiretto. La popolazione standard è rappresentata dalla popolazione di età compresa tra 65 e 99 anni beneficiaria di *US Medicare*.
Analisi per Hospital Referral Regions (HRR), anni 2003-2007
(modificata da Goodman et al., 2010)

Spiegazione. Una precedente ricerca ha dimostrato come il tasso di ospedalizzazione per tali condizioni mediche (*ambulatory care-sensitive conditions*), così come per molte altre, sia ampiamente correlato all'offerta locale di posti letto ospedalieri (Wennberg e Cooper, 1998). Maggiore è il numero di posti letto procapite e maggiore è la probabilità che i medici decidano di ricoverare i pazienti sofferenti di condizioni mediche che sarebbero invece trattabili in un setting assistenziale diverso dall'ospedale.

È tuttavia complesso stimare il contributo delle diverse variabili (stato di salute, offerta e accesso ai servizi) nello spiegare le differenze tra le aree geografiche. Ad esempio, non è stata rilevata alcuna associazione tra la percentuale di beneficiari di Medicare che hanno ricevuto almeno una visita ambulatoriale ($R^2 = 0,00$) (figura 2)

o la quantità di lavoro svolto nell'assistenza primaria ($R^2 = 0,00$) e il tasso di ospedalizzazione per *ambulatory care-sensitive conditions*. L'esercizio è stato ripetuto su aree geografiche di dimensioni inferiori rispetto all'HRR senza, tuttavia, ottenere risultati esplicativi ($R^2 = 0,01$).

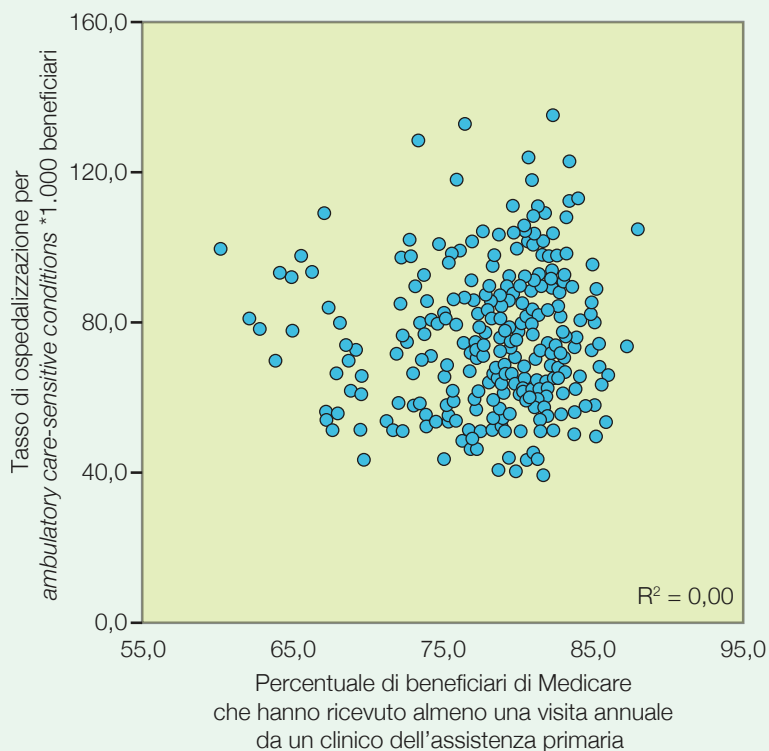


Figura 2 • Relazione tra la percentuale di beneficiari di Medicare che hanno ricevuto almeno una visita annuale da un clinico dell'assistenza primaria e il tasso di ospedalizzazione per ambulatory care-sensitive conditions *1.000 beneficiari di Medicare. Per "clinico dell'assistenza primaria" (*primary care clinician*) si intendono figure come: *family physician, general internist, nurse practitioner*. Ogni punto del grafico a dispersione rappresenta una HRR. Non si osserva alcuna relazione tra le due variabili ($R^2 = 0,00$).

Analisi per HRR, anni 2003-2007 (modificata da Goodman et al., 2010)

6 • L'APPROPRIATEZZA D'USO DELLE TECNOLOGIE SANITARIE: IL CASO DELLA FDG-PET IN AMBITO ONCOLOGICO

Luciana Ballini, Luca Vignatelli

L'Agenzia Sanitaria e Sociale Regionale (ASSR) dell'Emilia-Romagna ha iniziato ad occuparsi dell'uso appropriato della FDG-PET in ambito oncologico nel 2002. Nel 2010, l'ASSR ha avviato un nuovo progetto di sviluppo di criteri di appropriatezza nell'uso della FDG-PET, coinvolgendo oltre 60 esperti di diverse discipline, organizzati in 6 panel relativi a 6 tipi di tumore. La principale novità del progetto, rispetto ai precedenti, è consistita nella metodologia, che ha cercato di combinare il metodo RAND e l'approccio Grading of Recommendations, Assessment, Development and Evaluation (GRADE). Dai risultati del lavoro è emerso come la metodologia utilizzata sia risultata funzionale alla formulazione di raccomandazioni sull'appropriatezza dei test diagnostici, nonostante la sua complessità.

GLI ANTECEDENTI (DAL 2002 AL 2007)

La tomografia ad emissione di positroni (Positron Emission Tomography, PET) è una tecnologia di medicina nucleare che, utilizzando radiofarmaci a breve emivita, permette una diagnostica per immagini non invasiva. La PET studia i processi metabolici e le loro alterazioni conseguenti a diverse patologie. L'ambito di utilizzo principale e maggiormente studiato è quello oncologico.

L'Agenzia Sanitaria e Sociale Regionale (ASSR) dell'Emilia-Romagna ha iniziato ad occuparsi dell'uso appropriato della FDG-PET

nel 2002. L'attività svolta dall'Agenzia ha portato alla pubblicazione di un primo Dossier nel 2003,¹ che ha preso in considerazione 16 tipi di tumore per un totale di 47 indicazioni cliniche. I risultati del rapporto furono utilizzati per la realizzazione, nel 2002, di un audit clinico su un campione consecutivo di casi sottoposti a PET, presso l'unico Servizio PET allora presente nelle Aziende Sanitarie dell'Emilia-Romagna. Complessivamente, è emerso che dei 452 pazienti per i quali era stato richiesto un esame PET nel periodo oggetto dello studio, circa un terzo (38,7%) rientrava in una delle indicazioni giudicate appropriate, mentre un quarto dei casi (26,1%) rientrava nelle indicazioni inappropriate (figura 6.1). Nel 2007 è stato pubblicato un nuovo rapporto (Liberati et al., 2007), che ha aggiornato e al tempo stesso integrato il precedente, prendendo in considerazione 18 tipi di tumore (due in più rispetto al precedente) e un totale di 65 indicazioni cliniche. L'elaborazione del rapporto ha offerto, inoltre, l'occasione per realizzare un secondo audit clinico, nel 2006, presso i 6 Servizi PET presenti nelle Aziende Sanitarie dell'Emilia-Romagna (nel 2002 ne era presente 1). Complessivamente, è emerso che dei 588 pazienti per i quali era stato richiesto un esame PET nel periodo oggetto dello studio, più della metà (56%) rientrava in una delle indicazioni giudicate appropriate, mentre solo il 3,1% rientrava nelle indicazioni inappropriate (figura 6.2). Confrontando questo secondo audit (2006) con il primo (2002) emergono due risultati degni di attenzione:

1. incremento nell'uso appropriato della FDG-PET (38,7% nel 2002 vs 56% nel 2006) e decremento dell'uso inappropriate (26,1% nel 2002 vs 3,1% nel 2006);
2. incremento dell'uso della FDG-PET (8,4% nel 2002 vs 17,3% nel 2006) per indicazioni cliniche non incluse nei rapporti (vedi "altre indicazioni" nelle figure 6.1 e 6.2).

Questo secondo risultato mette in evidenza come nella valutazione di appropriatezza d'uso della FDG-PET condotta nel 2006 non sia

¹ Liberati et al. (2003). Il rapporto è stato realizzato dal CeVEAS (Centro per la valutazione dell'efficacia dell'assistenza sanitaria) su incarico dell'ASSR.

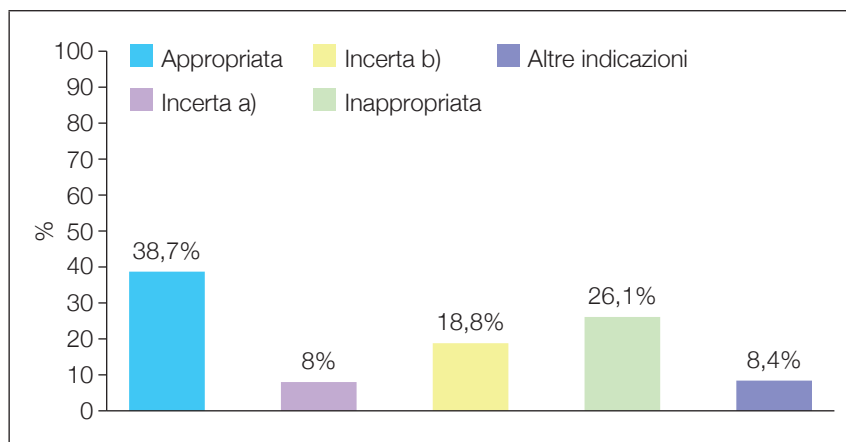


Figura 6.1 • La distribuzione della proporzione di pazienti per i quali era stato richiesto un esame PET, secondo i criteri di appropriatezza (modificata da Ballini et al., 2011. Audit clinico, anno 2006)

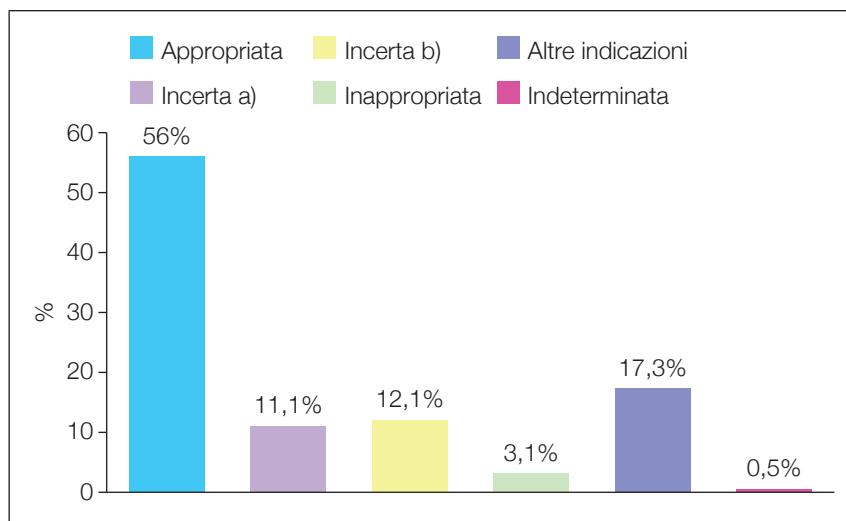


Figura 6.2 • La distribuzione della proporzione di pazienti per i quali era stato richiesto un esame PET, secondo i criteri di appropriatezza (modificata da Ballini et al., 2011. Audit clinico, anno 2006)

stata presa in considerazione una parte rilevante di quesiti clinici e diagnostici affrontati nella pratica clinica.

IL PROGETTO 2010-2012

Nel 2010, nell'ambito del Programma integrato Oncologia 5/2006, finanziato dal Ministero della Salute, è stato commissionato all'Osservatorio Regionale per l'Innovazione (ORI) dell'Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna un terzo aggiornamento (rispetto ai due precedenti, del 2003 e 2007). A tal fine, oltre 60 esperti di diverse discipline (medicina nucleare, radiologia, radioterapia, chirurgia, oncologia, pneumologia, ematologia, medicina interna, rappresentanti delle Direzioni sanitarie) sono stati coinvolti in 6 panel per lo sviluppo dei criteri di appropriatezza nell'uso appropriato della FDG-PET in 6 tipi di tumore (mammella, esofago, polmone, testa e collo, colon-retto, linfomi).

Il principale obiettivo del progetto è stato quello di sperimentare e valutare una nuova metodologia di lavoro in gruppi multidisciplinari, finalizzata alla formulazione di raccomandazioni diagnostiche, che combina assieme il metodo RAND e l'approccio GRADE (Schünemann et al., 2008). Nel box 6.1 viene riportata la definizione dei criteri di appropriatezza, adottata nel progetto regionale.

LA METODOLOGIA

Il processo di formulazione delle raccomandazioni per l'uso appropriato della FDG-PET è stato suddiviso in tre fasi principali.

1. *Definizione dei quesiti clinici e revisione sistematica delle evidenze.* Sulla base delle linee-guida internazionali, per ogni tumore è stato costruito il relativo percorso assistenziale. L'utilizzo della FDG-PET è stato ipotizzato nei punti del percorso di snodo decisionale, ovvero i risultati possibili del test sono stati collegati sequenzialmente ad un atto terapeutico (figura 6.3).

BOX 6.1 - LA DEFINIZIONE DEI CRITERI APPROPRIATEZZA

Appropriato. Scenari clinici per cui sussiste il razionale del *change in management* per un esito clinico importante, vi è un livello di evidenza moderato/alto di accuratezza diagnostica dell'esame PET e il valore del presunto beneficio è maggiore del valore del presunto danno conseguente all'esame PET.

Incerto. Scenari clinici per cui sussiste il razionale del *change in management* per un outcome clinico importante, ma vi è un livello di evidenza basso/molto basso sull'accuratezza diagnostica dell'esame PET e il rapporto rischio/beneficio è incerto.

Inappropriato. Scenari clinici per cui non sussiste il razionale del *change in management* per un outcome clinico importante; scenari clinici per cui sussiste il razionale del *change in management* per un outcome clinico importante, vi è un livello di evidenza alto/moderato di scarsa accuratezza diagnostica dell'esame PET e/o il valore del presunto danno è maggiore del valore del presunto beneficio conseguente all'esame PET.

Indeterminato. scenari clinici per cui sussiste il razionale del *change in management* per un outcome clinico importante, ma non vi sono dati di accuratezza diagnostica dell'esame PET.

Mappando il percorso assistenziale dei 6 tumori considerati sono stati formulati 55 quesiti clinici (*research questions*), relativi a: diagnosi di tumore primitivo, staging iniziale, risposta alla terapia, durante e dopo trattamento adiuvante, definizione del volume target per la radioterapia, follow-up, e staging della recidiva.

2. *Metodo GRADE per valutare l'importanza degli esiti.* In ambito diagnostico, un test viene definito appropriato quando consente di modificare la conoscenza di uno stato clinico (accuratezza diagnostica) e quando a tale conoscenza segue una modifica del percorso clinico e della scelta del trattamento ritenuto efficace (efficacia clinica). Un test viene, invece, definito inappropriato quando all'accuratezza diagnostica non segue la scelta di un trattamento efficace, rispetto al nuovo stato clinico osservato. La misurazione e valutazione dell'appropri-

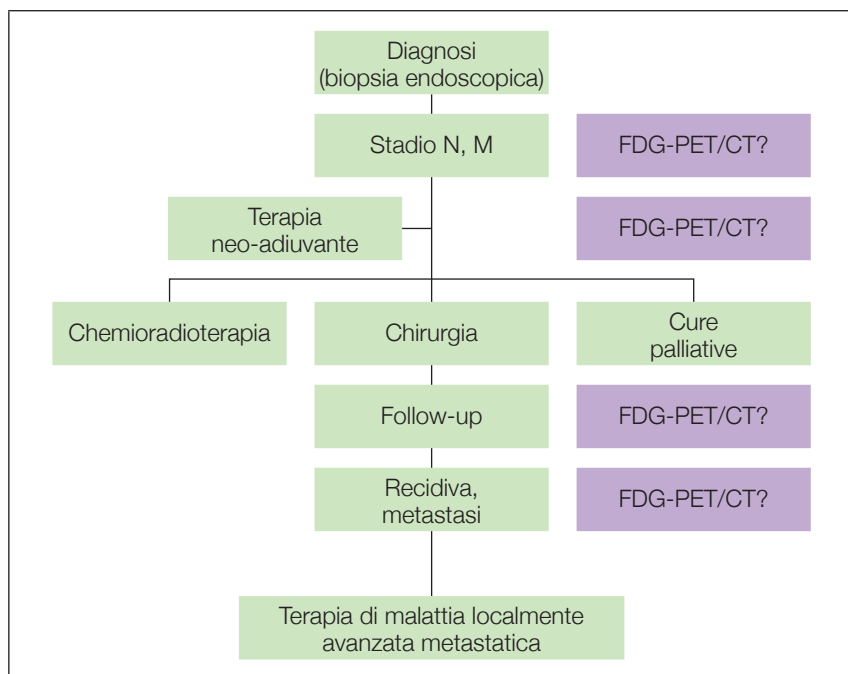


Figura 6.3 • Il percorso assistenziale nel tumore all'esofago ed il possibile utilizzo della FDG-PET (modificata da Ballini et al., 2011)

tezza di un test (ovvero del collegamento tra accuratezza diagnostica ed efficacia clinica) è, tuttavia, possibile solo indirettamente, attraverso una prospettiva teorica, più che empirica, a causa delle difficoltà metodologiche e pratiche nel condurre studi adeguati al quesito. Ciò significa legare in maniera teorica il risultato di un test diagnostico (positivo/negativo) all'azione conseguente (applico/non applico il trattamento) e agli esiti attesi per ciascuna categoria di pazienti (veri positivi, veri negativi, falsi positivi e falsi negativi). In altri termini, per valutare se l'accuratezza di un test – in termini di sensibilità e specificità (desumibili dagli studi) – incide sulle diverse categorie di pazienti (veri positivi, veri negativi, falsi positivi e falsi negativi) occorre individuare gli esiti possibili e giudicarne l'importanza clinica.

Questa operazione è possibile applicando il metodo GRADE, il quale richiede ad ogni componente del panel di definire e votare (con scala da 1 a 9) l'importanza degli esiti possibili per ogni conseguenza di ogni scenario clinico.

3. *Votazioni: metodo RAND.* Per ogni quesito clinico, ad ogni componente del panel è stata presentata la sintesi delle evidenze in termini di qualità e accuratezza diagnostica, con la classificazione quantitativa dei pazienti secondo il risultato del test. A ciascuno è stato poi richiesto di esprimere una votazione, sulla base del materiale proposto, che comprendeva anche l'importanza attribuita ad ogni esito. Il giudizio finale di appropriatezza è stato ottenuto secondo il metodo RAND (RAND/UCLA Method of Appropriateness Voting), il quale prevede l'assegnazione di un punteggio, compreso tra 1 e 9, e il calcolo del livello di accordo tra votanti. Per raggiungere un accordo, dopo l'eliminazione dei 2 voti estremi, i restanti voti devono cadere nella stessa regione di punteggio (da 1 a 3 per inappropriato, da 4 a 6 incerto, da 7 a 9 appropriato). In caso di disaccordo, dopo discussione, si procede a un secondo voto. Qualora l'accordo non venga raggiunto, l'indicazione clinica rimane incerta.

RISULTATI

Nella tabella 6.1 vengono presentati i risultati generali del progetto, consultabili in maniera dettagliata nei dossier pubblicati dall'Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna.² Dall'analisi della figura emerge come l'uso della FDG-PET nei vari tipi di tumore venga valutato dai clinici come tendenzialmente appropriato nella stadiazione iniziale e anche in alcuni casi nella stadiazione delle ricadute; mentre viene valutato sempre inappropriato nel follow-up.

² Per consultare i singoli dossier, impostare una ricerca per parole chiave nel seguente indirizzo web: <http://assr.regione.emilia-romagna.it/it/servizi/pubblicazioni/dossier>

TABELLA 6.1 - I CRITERI DI APPROPRIATEZZA D'USO DELLA FDG-PET IN AMBITO ONCOLOGICO

DOMANDA TUMORE	DIAGNOSI	STADIO N	STADIO M	DEFINIZIONE TV
Mammella	Inappropriato	Inappropriato	Disaccordo	N/A
Esofago	N/A	Incerto	Appropriato	Inappropriato
Coloretto	Inappropriato	Inappropriato	Appropriato	Inappropriato
Testa-collo	Appropriato (sconosciuto T-C primitivo)	Appropriato	Appropriato	Disaccordo
Polmone	Appropriato (SPN)	Appropriato (NSCLC)		Disaccordo
		Disaccordo		
		Inappropriato (BAC)		
Linfoma di Hodgkin	N/A	Appropriato		Indeterminato
Linfoma non Hodgkin	N/A	Appropriato		Indeterminato

In conclusione, la metodologia utilizzata è risultata funzionale alla formulazione di raccomandazioni sull’appropriatezza dei test diagnostici. In particolare, si evidenzia come:

- tutti e 60 gli esperti coinvolti nel progetto abbiano deciso di adottare e valutare la metodologia proposta, per quanto complessa;
- l’approccio GRADE abbia rappresentato un’opportunità concreta per prendere in considerazione l’impatto della FDG-PET sugli esiti di salute dei pazienti, tematica scarsamente indagata nei trial diagnostici;
- il focus della discussione degli esperti si sia ri-orientato, dall’accuratezza diagnostica della FDG-PET all’uso dei risultati del test per l’assunzione di decisioni cliniche;

RISPOSTA DURANTE IL TRATTAMENTO	FINE DELLA RISPOSTA AL TRATTAMENTO	FOLLOW-UP	STADIAZIONE DELLE RICADUTE
Incerto	Inappropriato	Inappropriato	Disaccordo
Inappropriato	Disaccordo	Inappropriato	Disaccordo
Indeterminato	Inappropriato (rettale)	Inappropriato	Appropriato
	Disaccordo (colon)		
Indeterminato	Disaccordo	Inappropriato	Appropriato
Inappropriato	Disaccordo (NSCLC)	Inappropriato	Disaccordo
	Inappropriato (NSCLC)		
Appropriato	Appropriato	Inappropriato	Appropriato
Inappropriato	Appropriato	Inappropriato	Appropriato

- la metodologia adottata, per quanto complessa, abbia permesso di completare l'intero progetto in 15 mesi, un tempo sufficientemente breve.

7 • L'APPROPRIATEZZA PRESCRITTIVA NELLA POPOLAZIONE ANZIANA

Vittorio Maio, Davide Botturi

Nel 2007, in Emilia-Romagna, un gruppo di esperti composto da ricercatori della Thomas Jefferson University (Philadelphia, USA) e professionisti dell'Azienda USL e dell'Azienda Ospedaliero-Universitaria di Parma ha sviluppato una serie di criteri, definiti nella letteratura internazionale come “Maio Criteria”, per la classificazione dei farmaci potenzialmente inappropriati nella popolazione anziana (≥ 65 anni). I farmaci potenzialmente inappropriati sono stati suddivisi in tre categorie: farmaci che dovrebbero essere evitati (ad esempio, pentazocina, disopiramide), farmaci raramente appropriati (ad esempio, ticlopidina, fluoxetina) e farmaci da usare solo per indicazioni specifiche (ad esempio, amiodarone, risperidone, doxazosina). Questi criteri hanno trovato applicazione in Emilia-Romagna a livello sia locale sia regionale, per il monitoraggio e la valutazione dell'appropriatezza nel comportamento prescrittivo dei medici di medicina generale.

LA METODOLOGIA

Da tempo la letteratura scientifica dedica notevole attenzione all'individuazione di criteri espliciti per valutare l'appropriatezza della terapia farmacologica nell'anziano a partire dai dati sul consumo farmaceutico. Tra gli strumenti presenti in letteratura, i criteri di Beers sono quelli maggiormente adottati e studiati (American Geriatrics Society, 2012). I criteri di Beers individuano farmaci o classi di farmaci

che, in generale, dovrebbero essere evitati negli anziani in quanto “il rischio di reazioni avverse alla terapia è tale da diminuirne il potenziale beneficio” (Beers, 1997). Secondo uno studio condotto in Emilia-Romagna (cfr. box 7.1), nell’anno 2001 il 18% della popolazione anziana (≥ 65 anni) residente utilizzava abitualmente almeno un farmaco incluso nella lista di Beers. Si tratta di un valore di prevalenza sostanzialmente elevato, ma comunque sottostimato, perché la lista di

**BOX 7.1 - SINTESI DI UNO STUDIO CONDOTTO IN EMILIA ROMAGNA
DA MAIO ET AL. (2006)**

Obiettivo. Valutare la prevalenza della prescrizione di farmaci potenzialmente inappropriati nella popolazione anziana (≥ 65 anni) e indagare i fattori predittivi.

Metodologia. Il disegno dello studio è retrospettivo di coorte. Per lo studio sono state utilizzate le prescrizioni di farmaci presenti nel flusso amministrativo corrente dell’Assistenza Farmaceutica Territoriale (AFT), con riferimento al periodo 1° gennaio-31 dicembre 2001. Le prescrizioni sono state collegate, attraverso record *linkage*, alla banca dati Anagrafe degli Assistiti, contenente una serie di informazioni demografiche (ad esempio, genere, età, residenza). Sono stati arruolati nella coorte gli anziani residenti in Emilia-Romagna che hanno ricevuto almeno una prescrizione di farmaci nel periodo di osservazione; gli arruolati sono stati 849.425.

È stata valutata l’associazione tra la prevalenza di farmaci potenzialmente inappropriati, come definiti nei criteri di Beers (Fick et al., 2003), e una serie di variabili relative alle persone arruolate, tra cui il genere, l’età, il luogo di residenza (area urbana/rurale), il livello economico, il numero di patologie croniche, il numero di farmaci assunti.

Risultati. Complessivamente, 152.641 anziani residenti in Emilia-Romagna (il 18% degli arruolati) hanno ricevuto almeno una prescrizione di farmaci potenzialmente inappropriati. In particolare, all’11,5% sono stati prescritti due farmaci potenzialmente inappropriati, e all’1,7% tre o più. Il farmaco (principio attivo) potenzialmente inappropriato più prescritto è risultato essere la doxazosina (prescritto al 23% degli anziani), seguito da ketorolac (20,5%), ticlopidina (18,3%) e amiodarone (12,6%). L’età avanzata e la presenza di politerapia e di più patologie croniche rappresentano fattori associati ad una maggiore probabilità di ricevere almeno un farmaco potenzialmente inappropriato, mentre il genere femminile, la residenza in un’area urbana e un livello economico elevato sono fattori associati ad una minore probabilità.

farmaci usata per lo studio non ha incluso alcuni farmaci presenti nei criteri di Beers, ma non commercializzati o rimborsati in Italia.

Per questo motivo, nel 2007, in Emilia-Romagna un gruppo di esperti composto da ricercatori della Thomas Jefferson University e professionisti dell'Azienda USL e dell'Azienda Ospedaliero-Universitaria di Parma¹ ha sviluppato, partendo dai criteri di Beers, una serie di criteri espliciti che tengono conto sia del prontuario farmaceutico italiano che delle abitudini prescrittive dei medici italiani (cfr. anche Maio et al., 2010). La lista è stata definita nella letteratura internazionale come Maio Criteria (Kaufmann et al., 2014) e include farmaci/classi di farmaci potenzialmente inappropriati, suddivisi in tre categorie² (tabella 7.1).

Inoltre, la lista distingue i farmaci rimborsabili da quelli non rimborsabili dall'SSN. Tra i farmaci rimborsabili dall'SSN sono stati inclusi 25 farmaci o classi di farmaci; 16 di essi appartengono alla categoria di "farmaci che dovrebbero essere sempre evitati", 3 sono considerati come raramente appropriati, e 6, pur avendo alcune indi-

TABELLA 7.1 - LE TRE CATEGORIE DI FARMACI POTENZIALMENTE INAPPROPRIATI NELLA POPOLAZIONE ANZIANA (≥ 65 ANNI) (MODIFICATA DA MAIO ET AL., 2010)

Farmaci che dovrebbero essere evitati	Farmaci che la letteratura riporta ad alto rischio per l'anziano. Esistono alternative più sicure
Farmaci raramente appropriati	Farmaci efficaci ma non di prima scelta, con un profilo rischio/beneficio e/o beneficio/costo non favorevole
Farmaci da usare solo per indicazioni specifiche	Farmaci che hanno alcune indicazioni, ma che spesso devono essere usati sotto stretto controllo. Tali farmaci sono potenzialmente soggetti ad un uso non appropriato

¹ Il panel era composto da 9 esperti: 3 medici di medicina generale, 1 geriatra, 1 farmacista clinico, 2 psichiatri, 1 cardiologo, e 1 direttore di una struttura ospedaliera di lungodegenza.

² Nella lista sono presenti alcuni principi attivi raramente prescritti oggi, solo perchè sono ancora inseriti nel prontuario terapeutico nazionale.

cazioni specifiche, spesso risultano erroneamente usati. Alcuni farmaci sono considerati potenzialmente inappropriati a determinate dosi o durata di somministrazione (tabella 7.2).

TABELLA 7.2 - L'ELENCO DELLE CLASSI DI FARMACI E DEI PRINCIPI ATTIVI RIMBORSABILI DALL'SSN E POTENZIALMENTE INAPPROPRIATI NELLA POPOLAZIONE ANZIANA (≥ 65 ANNI), SECONDO I MAIO CRITERIA (MODIFICATA DA AMOS ET AL., 2014)

FARMACI DA EVITARE	<i>Analgesici</i>	Pentazocina
	<i>Antiaritmici</i>	Disopiramide
	<i>Antinfiammatori</i>	FANS (>15 giorni), indometacina, ketorolac iniettabile (max 2 gg)
	<i>Antiparkinson</i>	Orfenadrina cloridrato
	<i>Sistema cardiovascolare</i>	Clonidina (comprese), metildopa, nifedipina a breve durata d'azione, spironolattone >25 mg/die
	<i>Sistema endocrino</i>	Estrogeni (sistemici), metiltestosterone
	<i>Sistema nervoso centrale</i>	Amitriptilina, citalopram >20 mg/die, clorpropamide, escitalopram >10 mg/die
FARMACI RARAMENTE APPROPRIATI	<i>Antiaggreganti</i>	Ticlopidina
	<i>Antisecretivi</i>	Inibitori di pompa (≥ 12 mesi)
	<i>Antidepressivi</i>	Fluoxetina
FARMACI DA USARE SOLO PER ALCUNE CONDIZIONI	<i>Antiaritmici</i>	Amiodarone, diidrochinidina, chinidina, dronedarone, flecainide, propafanone, sotalolo
	<i>Antipsicotici atipici</i>	Risperidone, quetiapina, olanzapina, clozapina, aripipazolo, paliperidone, ziprasidone
	<i>Antipsicotici convenzionali</i>	Aloperidolo, amisulpiride, clorpromazina, clotiapina, flufenazina, perfenazina, pimozide, promazina, sulpiride, trifluoperazina
	<i>Sistema cardiovascolare</i>	Clonidina (cerotti), digossina, doxazosina

Per i farmaci non rimborsabili dall'SSN, la lista include 6 farmaci appartenenti alla categoria di quelli che dovrebbero essere sempre evitati: benzodiazepine a lunga emivita, benzodiazepine a breve emivita, antistaminici, lassativi stimolanti, nitrofurantoina e metoclopramide (orale).

Il panel di esperti ha inoltre sviluppato una lista di farmaci alternativi, ove possibile, a quelli ritenuti inappropriati, con l'obiettivo di fornire ai medici di medicina generale uno strumento concreto a cui rivolgersi per migliorare la qualità della prescrizione per i pazienti anziani nella pratica ambulatoriale quotidiana.

Per ogni farmaco ritenuto potenzialmente inappropriato è stata predisposta una specifica scheda contenente il rationale per cui il farmaco è ritenuto inappropriato, eventuali farmaci alternativi e le relative motivazioni. A titolo esemplificativo si riporta nella tabella 7.3 la scheda riferita alla pentazocina (analgesico).

TABELLA 7.3 - ESEMPIO DI SCHEDA PREDISPOSTA PER OGNI FARMACO POTENZIALMENTE INAPPROPRIATO NELLA POPOLAZIONE ANZIANA (≥ 65 ANNI) (MODIFICATA DA SERVIZIO ASSISTENZA DISTRETTUALE REGIONE EMILIA-ROMAGNA ET AL., 2012)

PENTAZOCINA	
Perché	Analgesico narcotico con effetti avversi severi, inclusi confusione ed allucinazione. Inoltre, ha proprietà sia agoniste sia antagoniste e può indurre sintomi da astinenza, incluso dolore, in pazienti dipendenti da altri oppioidi.
Quali alternative?	Nella terapia del dolore da moderato a grave è preferibile usare altri farmaci, incluso l'oxycodone e la morfina.
Quali motivazioni?	Presenta un effetto su pressione e frequenza cardiaca mediato dalle catecolammine.

IL PROGETTO DI RICERCA A PARMA CON I MEDICI DI MEDICINA GENERALE (2007-2009)³

Nel periodo 2007-2009 i ricercatori della Thomas Jefferson University (Philadelphia, USA)⁴ hanno coordinato e realizzato un progetto finalizzato a ridurre l'inappropriatezza prescrittiva negli anziani (≥ 65 anni) residenti nella Azienda USL di Parma, attraverso un complesso intervento rivolto a tutti i medici di medicina generale (MMG) operanti nel medesimo territorio. La ricerca triennale è stata progettata come un "multi-phase, multi-factorial, proof-of-concept intervention" (figura 7.1).

Dall'analisi della figura 7.1 emerge come il progetto sia stato suddiviso in tre fasi, ognuna corrispondente ad un'annualità, e come per ciascuna fase siano state realizzate specifiche azioni e utilizzati specifici strumenti, tra cui:

- creazione della lista di farmaci potenzialmente inappropriati nella popolazione anziana (come già riportato nel primo paragrafo, *La metodologia*, di questo capitolo), utilizzando come metodologia per la valutazione collettiva tra esperti la *nominal group technique*;
- creazione di un modello formativo "attivo", ovvero il più possibile orientato a stimolare la partecipazione degli MMG, che ha previsto l'utilizzo di materiale informativo relativo alla lista dei farmaci potenzialmente inappropriati, il coinvolgimento di esperti e opinion leader locali, l'organizzazione di incontri *ad hoc* direttamente nelle varie sedi di lavoro degli MMG, la diffusione delle informazioni attraverso un approccio di *academic detailing* e *case study*.⁵

Una parte integrante del progetto di ricerca è rappresentata dalla misurazione e valutazione delle conseguenze del complesso interven-

³ Il progetto è descritto dettagliatamente in Keith et al. (2013) e Lopatto et al. (2014), a cui si rimanda per l'approfondimento.

⁴ Nell'ambito di un rapporto di collaborazione con l'Azienda USL di Parma, l'Agenzia Sanitaria e Sociale Regionale e la Direzione Generale Sanità e Politiche Sociali della Regione Emilia-Romagna.

⁵ Tra gli strumenti utilizzati, vi è anche un questionario con vignette, somministrato ai medici di medicina generale per rilevare il livello di appropriatezza nella prescrizione. Per approfondimenti cfr. Maio et al. (2011).

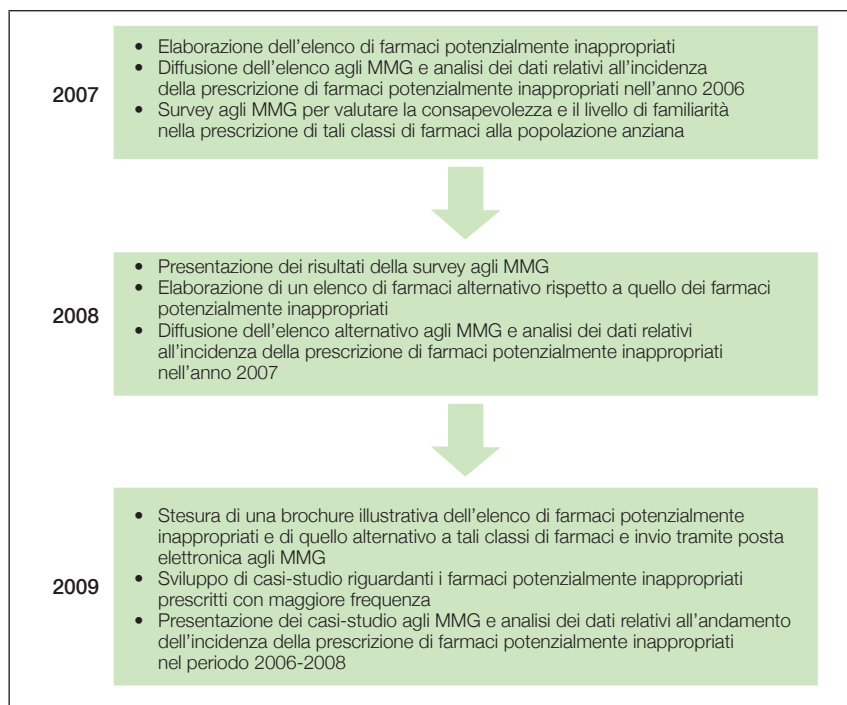


Figura 7.1 • La descrizione del progetto (2007-2009) (modificata da Keith et al., 2013)

to formativo sui comportamenti prescrittivi dei medici di medicina generale. Nelle figure 7.2 e 7.3 vengono illustrati alcuni risultati, rimandando agli specifici articoli scientifici (Keith et al., 2013; Lopatto et al., 2014) la loro trattazione dettagliata.

Nella figura 7.2 viene rappresentato l'andamento nell'incidenza della proporzione di anziani (≥ 65 anni) a cui è stato prescritto almeno un farmaco potenzialmente inappropriato negli anni 2007-2009⁶ residenti nell'Azienda USL di Parma ("Parma LHA") e nell'Azien-

⁶ Il numeratore è rappresentato dal numero di anziani (≥ 65 anni) residenti a cui è stato prescritto almeno un farmaco potenzialmente inappropriato; il denominatore dal numero di anziani residenti a cui è stata prescritta qualsiasi tipologia di farmaco.

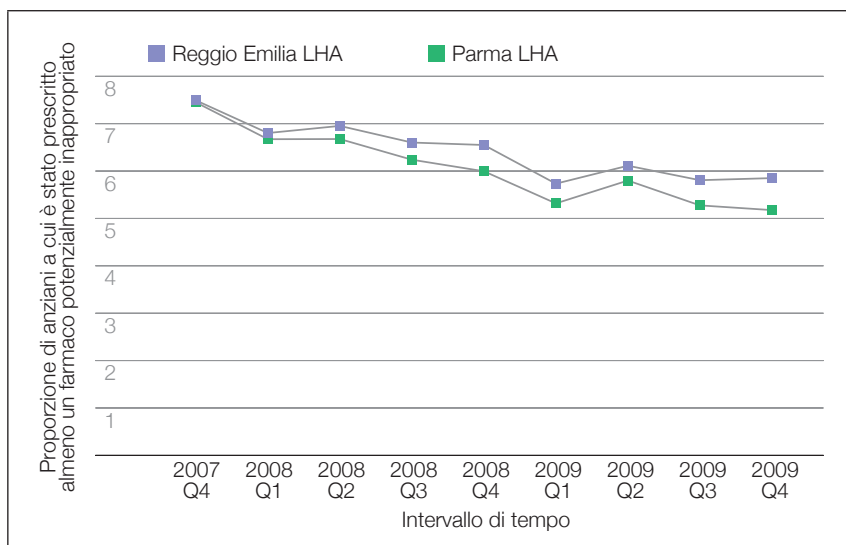


Figura 7.2 • Andamento temporale nell'incidenza della proporzione di anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato. Aziende USL (di residenza) di Parma e Reggio Emilia. Anni 2007-2009 (modificata da Keith et al., 2013)

da USL di Reggio Emilia ("Reggio Emilia LHA"), identificata come controllo (i medici di medicina generale non sono stati esposti all'intervento formativo). Il periodo compreso tra il quarto quadrimestre ("Q4") del 2007 e il quarto quadrimestre del 2009 rappresenta il periodo di esposizione all'intervento formativo dei medici di medicina generale operanti nel territorio dell'Azienda USL di Parma.

Dall'analisi della figura si osserva come nell'Azienda USL di Parma la proporzione di anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato:

- sia progressivamente diminuita, passando dal 7,8% (6098 anziani su 78.482) nel quarto quadrimestre del 2007 (etichetta "Q4"), al 5,3% (4277 anziani su 80.229) nel quarto quadrimestre del 2009;

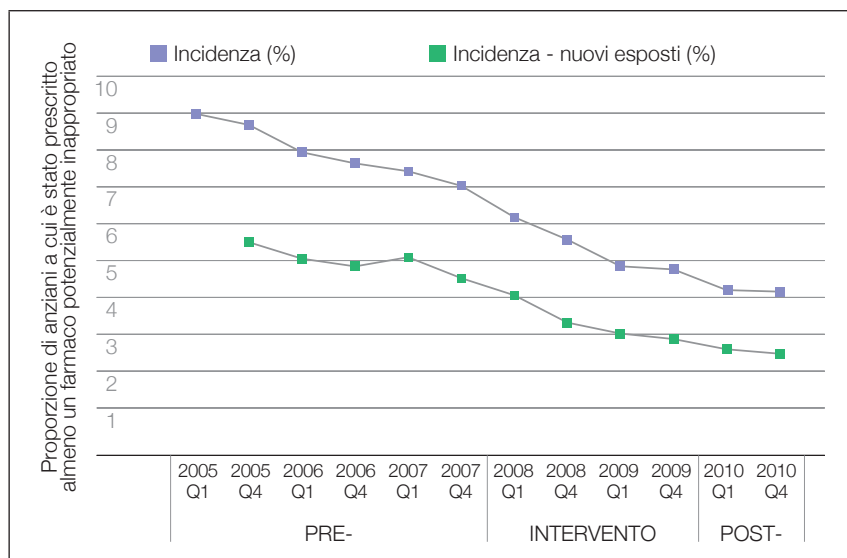


Figura 7.3 • Andamento temporale nell'incidenza di anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato. Azienda USL (di residenza) di Parma. Anni 2005-2010 (modificata da Lopatto et al., 2014)

- si sia mantenuta costantemente al di sotto della proporzione dell'Azienda USL di Reggio Emilia, anch'essa caratterizzata da un trend in diminuzione;
- presenti una variazione (2009 Q4 a confronto con 2007 Q4) significativamente superiore a quella dell'Azienda Usl di Reggio Emilia (-31,4% Azienda USL di Parma vs -21,6% Azienda USL di Reggio Emilia, $p < 0,001$).

Nella figura 7.3 viene rappresentato l'andamento nell'incidenza della proporzione di anziani (≥ 65 anni) a cui è stato prescritto almeno un farmaco potenzialmente inappropriato, residenti nel territorio su cui insiste l'Azienda USL di Parma, negli anni 2005-2010. L'arco temporale di osservazione è stato suddiviso in tre periodi, corrispondenti alla fase pre-intervento formativo, alla fase dell'intervento, e a quella del post-intervento. L'unità temporale di analisi è rappresentata dal primo

e quarto quadrimestre di ogni anno (Q1 e Q4). Per il calcolo dell'incidenza sono state identificate due tipologie di esposizione: gli anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato in un determinato quadrimestre (etichetta "incidenza" nella figura 7.3); gli anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato in un determinato quadrimestre, ma a cui non è stato prescritto alcun farmaco potenzialmente inappropriato nei 90 giorni precedenti (etichetta "incidenza - nuovi esposti" nella figura 7.3).

Dall'osservazione della figura si rileva un progressivo decremento, durante l'intero periodo di analisi, delle due misure di incidenza di anziani a cui è stato prescritto almeno un farmaco potenzialmente inappropriato. Ciò permette di proporre almeno due considerazioni di carattere generale,⁷ rispetto al trend di diminuzione dell'inappropriatezza prescrittiva:

- sembra essere pre-esistente alla realizzazione dell'intervento formativo; il decremento appare, tuttavia, più marcato in corrispondenza del periodo di realizzazione dell'intervento, a supporto dell'ipotesi che quest'ultimo possa aver contribuito a modificare l'attitudine prescrittiva dei medici di medicina generale;
- si mantiene anche nell'anno 2010, per quanto l'ampiezza di tale variazione appaia inferiore rispetto a quella osservata negli anni 2007-2009, a supporto dell'ipotesi che i cambiamenti nelle abitudini prescrittive degli MMG possano essere sufficientemente stabili, e non temporanei.

I PROFILI DEI NUCLEI DI CURE PRIMARIE IN EMILIA-ROMAGNA (2010)

In Emilia-Romagna questa metodologia è stata adottata per il calcolo di una serie di misure incluse nel report *Profili dei Nuclei di cure primarie*, un report annuale che dal 2010 viene messo a disposizione

⁷ Si rimanda direttamente all'articolo di Lopatto et al. (2014) per approfondire la metodologia, come i risultati e la loro discussione.

delle Aziende Unità Sanitarie Locali e contiene un insieme di dati riferiti ad ogni singolo Nucleo di cure primarie (NCP).⁸ In particolare, nei profili sono contenute informazioni riguardanti:

- le caratteristiche dei medici di medicina generale (ad esempio, genere, classi d'età, anni di convenzionamento, forma associativa);
- le caratteristiche demografiche (ad esempio, genere, classi d'età) ed epidemiologiche (stima della prevalenza di malattie croniche) della popolazione assistita;
- il ricorso ai servizi sanitari (i ricoveri ospedalieri, tra cui i ricoveri potenzialmente evitabili, gli accessi al pronto soccorso, il consumo di farmaci e di prestazioni specialistiche, l'assistenza domiciliare);
- l'appropriatezza dell'assistenza erogata a persone sofferenti di malattie cardiovascolari (IMA, scompenso cardiaco congestizio), diabete e asma. Le misure si riferiscono al trattamento farmacologico (prescrizione) e alla gestione (ad esempio, visite, test effettuati) delle singole patologie;
- alcuni ambiti specifici della terapia farmacologica, tra cui l'*appropriatezza della terapia farmacologica nell'anziano* (≥ 65 anni).

L'obiettivo generale consiste nell'offrire ai medici di medicina generale di ogni singolo NCP uno strumento utile per l'autovalutazione e di stimolo per l'avvio di pratiche di miglioramento. A tal fine, i dati e le informazioni riferiti al singolo Nucleo vengono presentati in modo da consentire un confronto spaziale (con la media distrettuale, aziendale, e regionale) e temporale (serie storica). La costruzione del report e la definizione degli indicatori sono il risultato dei lavori di un gruppo costituito da collaboratori della Regione Emilia-Romagna,

⁸ I Nuclei di cure primarie sono "le unità organizzative di base del sistema delle cure primarie" (DGR 2011/2007) a cui aderiscono, *in primis*, i medici di medicina generale. Nel 2014, in Emilia-Romagna, ci sono 215 NCP rispetto ad un totale di 3079 medici di medicina generale, in media 14 medici di medicina generale per NCP. Per approfondimenti consultare: a) la DGR 2011/2007, *Indirizzi per l'organizzazione dei Dipartimenti di Cure primarie, di Salute mentale e Dipendenze patologiche, di Sanità Pubblica*; b) i risultati dell'Osservatorio delle Cure primarie, che dal 2010 si occupa di monitorare, attraverso la somministrazione annuale di un questionario, l'evoluzione di diversi aspetti, clinici e organizzativi, relativi agli NCP.

ricercatori della Thomas Jefferson University (Philadelphia, USA), professionisti delle Aziende USL, e medici di medicina generale.⁹

Nel 2012, in Emilia-Romagna, il numero totale di anziani (≥ 65 anni) a cui è stato prescritto almeno un farmaco ritenuto potenzialmente inappropriato era pari a 234.288,¹⁰ corrispondente al 27,6% degli anziani a cui è stato prescritto almeno un farmaco nello stesso anno (847.642) e al 23% del totale della popolazione anziana residente in regione (1.018.053).¹¹

⁹ Per approfondimenti consultare l'indirizzo web: <http://www.saluter.it/siseps/sanita/prim/progetto-criteri-di-calcolo>

Questo intervento è stato anche oggetto di un Progetto dell'Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna, dal titolo "Valutazione dell'impatto sui comportamenti clinici dei medici di medicina generale della messa a disposizione di informazioni sulla qualità dell'assistenza erogata", finanziato dal Ministero della Salute, nell'ambito del programma CCM, anno 2010.

¹⁰ Banca dati Cure primarie – PRIM, Regione Emilia-Romagna. Il dato è aggiornato al 1° gennaio 2013.

¹¹ Statistica Self-Service Regione Emilia-Romagna. Il dato è aggiornato al 1° gennaio 2013, ed è provvisorio.

8 • I RICOVERI OSPEDALIERI IN EMILIA-ROMAGNA 2001-2008. ATLANTE DELL'APPROPRIATEZZA ORGANIZZATIVA

Lucia Nobilio

La regione Emilia-Romagna ha messo in atto, già a partire dagli anni '90, specifiche azioni mirate alla riduzione dei ricoveri ospedalieri inappropriati, sviluppandole e rafforzandole successivamente, anche in adempimento agli indirizzi normativi nazionali.

L'*Atlante dell'appropriatezza organizzativa 2001-2004* (Fiorini et al., 2006) descrive la metodologia messa a punto dall'Agenzia Sanitaria Sociale per la valutazione di appropriatezza dei ricoveri, basata sull'analisi delle informazioni contenute nelle SDO e sul giudizio dei clinici. Viene inoltre qui presentato un aggiornamento dei dati per gli anni 2001-2008.

BACKGROUND

Negli ultimi 10 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* (il cosiddetto “Decreto LEA”) definiva i DRG ad elevato rischio di inappropriately; successivamente la legge 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.¹

¹ Per un approfondimento cfr. più avanti l'Appendice normativa.

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 (Philadelphia, USA) 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.²

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 novembre 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 ricoveri riferiti agli anni 2001-2004, oggetto di una pubblicazione tecnica, l'*Atlante dell'appropriatezza organizzativa* (Fiorini et al., 2006).

Il corrispondente percorso normativo si è concretizzato in alcune deliberazioni regionali:

- la DGR 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 DGR 2126/2005 (Allegato 6) che, facendo proprio l'elenco di DRG definito dalla ASR per il progetto *Atlante dell'appropriatezza organizzativa*, portava a 62 il numero di DRG “potenzialmente inappropriati”;
- successivamente la DGR 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

² La metodologia utilizza i dati provenienti dalle SDO per valutare la gravità della diagnosi principale, la presenza di comorbidità e il loro livello di gravità, la natura delle procedure chirurgiche effettuate.

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 2126/2005³ 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 dal 1° gennaio 2009, presso tutte le Aziende Sanitarie pubbliche e private è stata adottata la versione italiana 2007 dell'*International Classification of Diseases 9th revision – Clinical Modification* (ICD9-CM) 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.⁴

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

³ Dai 62 DRG dell'elenco originario è stato rimosso il DRG 127 (scompenso cardiaco congestizio).

⁴ Per *comorbidità significative* si intendono le patologie, segnalate tra le diagnosi secondarie, che possono giustificare la necessità di un ricovero in regime ordinario e che contribuiscono al giudizio complessivo di appropriatezza del ricovero.

– 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 (Fiorini et al., 2006) 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 inappropriatezza in regime ordinario” e comprendenti i 43 DRG definiti, nell'Allegato 2C, dal DPCM del 29 novembre 2001 (“Decreto LEA”) 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 (box 8.1).

BOX 8.1 - DEFINIZIONE DEI CRITERI DI APPROPRIATEZZA
(42 DRG MEDICI E 20 DRG CHIRURGICI)**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.

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 (ad esempio, emicrania, diabete mellito in stadio 1 senza complicazioni).

Ricovero appropriato tempestivo. Ricovero necessario, per una patologia che richiede il ricovero in regime ordinario (ad esempio, infarto miocardico acuto)/segni o sintomi (ad esempio, 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 (ad esempio, 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 delle procedure eseguite in termini di effettuabilità in *day surgery* o in ambulatorio. La necessità dell'intervento chirurgico o della procedura rispetto alle carat-

* Il *disease staging* è un 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 fisio-patologiche e al grado di diffusione locale e di interessamento sistemico. Per un approfondimento cfr. Taroni (1996).

teristiche 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 esempio, 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).

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 esempio, gli interventi di cataratta o di tunnel carpale).

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 inappropriatezza 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 inappropriatezza 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:

- 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;
- pazienti residenti in altre regioni italiane (la cosiddetta mobilità interregionale attiva) per evitare distorsioni sul livello di gravità dei casi trattati;⁵
- pazienti residenti in Emilia-Romagna e ricoverati in altre regioni (la mobilità passiva);
- 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;
- 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 (figura 8.1).

⁵ I pazienti provenienti da fuori Regione, e quindi disposti ad affrontare lunghe distanze, potrebbero presentare mediamente un livello di gravità della patologia superiore rispetto a quello osservato per i flussi di mobilità infra regionali; allo stesso tempo, l'indicazione al ricovero potrebbe essere determinata proprio dalla distanza della struttura sanitaria dal domicilio del paziente.

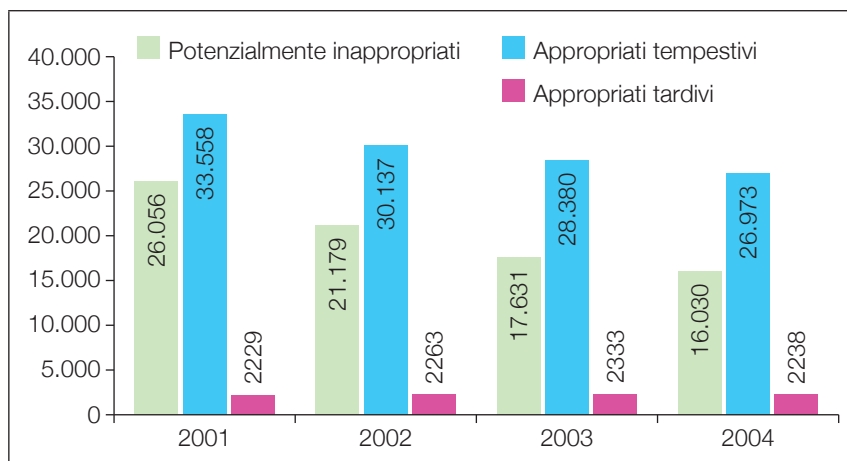


Figura 8.1 • Profili di appropriatezza dei ricoveri ordinari >1 giorno attribuiti ai DRG selezionati (popolazione in studio, anni 2001-2004)
(modificata da Fiorini et al., 2006, figura 27)

Il trend di riduzione dei ricoveri potenzialmente inappropriati differisce però per la casistica medica e per quella chirurgica (figura 8.2).

DRG chirurgici. In questo caso, nell'arco dei 4 anni la riduzione dell'incidenza di potenziale inappropriatezza 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 8.3, l'esempio della chirurgia del cristallino, 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.

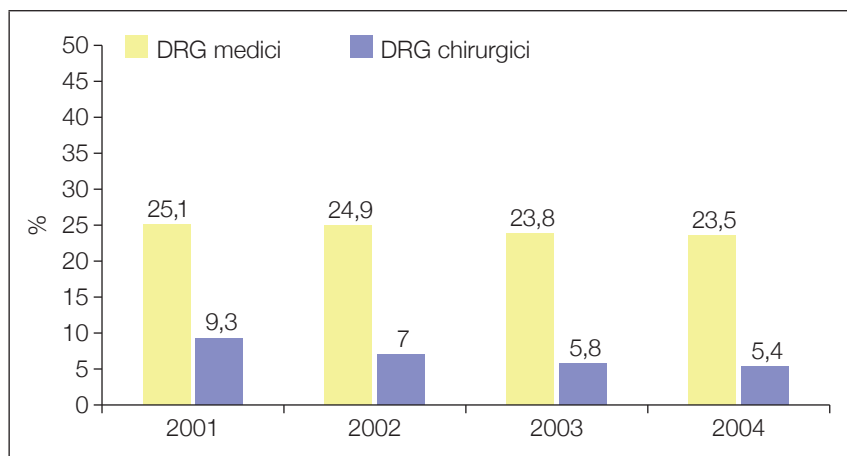


Figura 8.2 • Ricoveri potenzialmente inappropriati attribuiti ai DRG selezionati (%sul totale dei ricoveri ordinari e in day hospital; DRG medici vs DRG chirurgici) (modificata da Fiorini et al., 2006, figura 29)

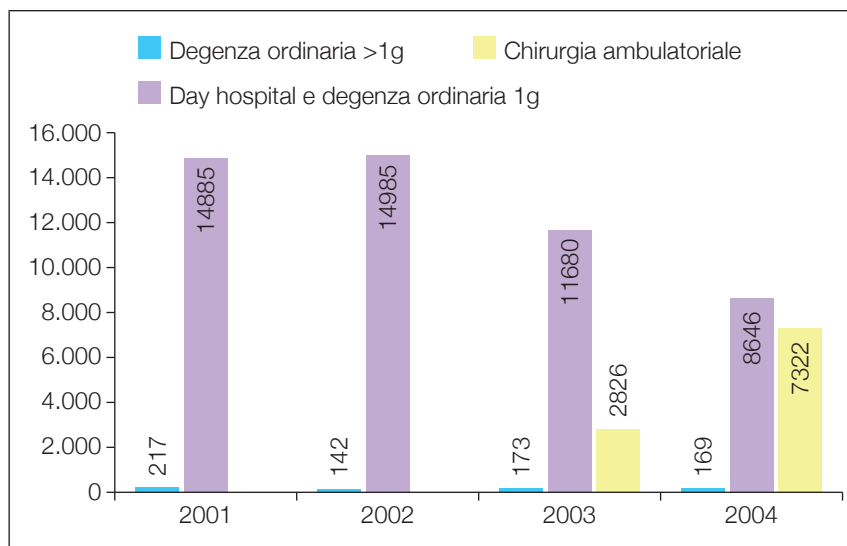


Figura 8.3 • Intervento sul cristallino, DRG 39 (anni 2001-2004; popolazione 18-74 anni) (modificata da Fiorini et al., 2006, figura 32)

DRG medici. In questo caso (figura 8.4), 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 (2169 casi su 45.490), mentre nel 2004 ne costituiscono il 6,4% (2175 casi su 34.063).

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 extraospedaliero. Come è stato detto, essi rappresentano una piccola quota della casistica analizzata ($n =$

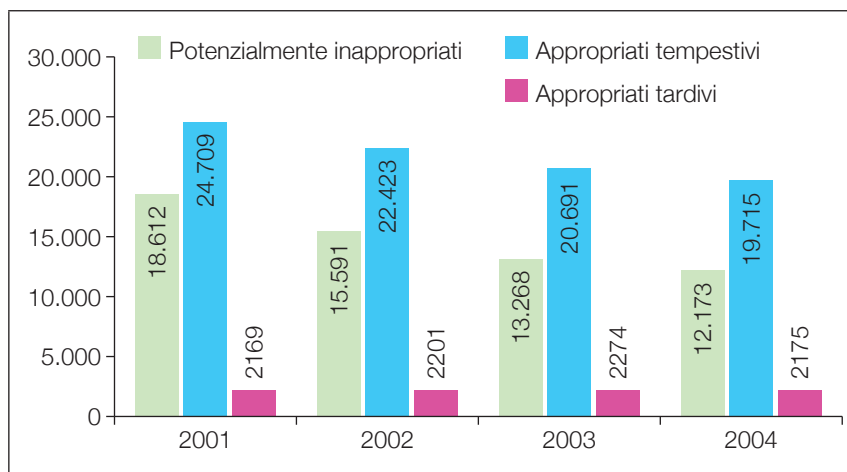


Figura 8.4 • Profili di appropriatezza dei ricoveri ordinari >1 giorno attribuiti ai DRG medici selezionati (popolazione in studio, anni 2001-04)
(modificata da Fiorini et al., 2006, figura 36)

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 inappropriatazza", esplora la distribuzione di questi ricoveri nell'intera casistica ospedaliera. Dall'osservazione dei 62 DRG, nel solo anno 2004,

TABELLA 8.1 - 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) (MODIFICATA DA FIORINI ET AL., 2006)

DRG	TOTALE RICOVERI ORD. >1 G		RICOVERI APPROPRIATI TARDIVI		
	N	N	% SUI RICOVERI 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
Sub-totale		2.115		94,5	

emerge come il 94,5% di questi ricoveri sia riconducibile ad alcuni DRG, come riportato nella tabella 8.1.

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 nella tabella 8.2.

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. Nella tabella 8.3 vengono presentati i risultati.

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 1947 a 4074. La loro distribuzione tra i DRG ad elevato rischio di inappropriatazza 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.

TABELLA 8.2 - CATEGORIA DIAGNOSTICA PRINCIPALE DEL 94,5% DEI CASI APPROPRIATI TARDIVI (RICOVERI ORDINARI >1 GIORNO; ANNO 2004) (MODIFICATA DA FIORINI ET AL., 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 colelitiasi	313	14,8	60,1
CV02 Miocardiopatia	228	10,8	70,9
RS30 Polmonite batterica	210	9,9	80,9
ND15 Iperitiroidismo	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

TABELLA 8.3 - 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 (MODIFICATA DA FIORINI ET AL., 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	484	56,6	1116	27,4
Colecistite e colelitiasi	313	609	33,9	922	22,6
Diabete mellito	327	246	57,1	573	14,1
Polmonite batterica	210	302	41,0	512	12,6
Miocardipatia	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	3,2
Ipotiroidismo	49	41	54,4	90	2,2
Disturbi vie G-U NAC	51	8	86,4	59	1,4
TOTALE (9 categorie diagnostiche)	1947	2127	47,8	4074	100,0

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 richi amino 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 (tabella 8.4).

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 8.5) (Louis et al., 2008).

Gli autori dell'*Atlante* invitano a 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.

TABELLA 8.4 - ELENCO DEI PREVENTION QUALITY INDICATORS SECONDO L'AHQ:
INDICATORI OSPEDALIERI "SENSIBILI ALLE CURE AMBULATORIALI" (MODIFICATA DA AHQ
GUIDE TO PREVENTION QUALITY INDICATORS, 2007)

PQI NUMERO	PREVENTION QUALITY INDICATORS (PQI)
1	Tasso di ospedalizzazione per complicanze a breve termine del diabete
2	Tasso di ospedalizzazione per appendicite perforata
3	Tasso di ospedalizzazione per complicanze a lungo termine del diabete
5	Tasso di ospedalizzazione per broncopneumopatia cronica ostruttiva
7	Tasso di ospedalizzazione per ipertensione
8	Tasso di ospedalizzazione per scompenso cardiaco congestizio
9	Basso peso alla nascita
10	Tasso di ospedalizzazione per disidratazione
11	Tasso di ospedalizzazione per polmonite batterica
12	Tasso di ospedalizzazione per infezione delle vie urinarie
13	Tasso di ospedalizzazione per angina senza procedure
14	Tasso di ospedalizzazione per diabete non controllato
15	Tasso di ospedalizzazione per asma (popolazione adulta)
16	Tasso di ospedalizzazione per amputazione degli arti inferiori nelle persone con diabete

AGGIORNAMENTI 2001-2008

Nell'aggiornamento delle misure di appropriatezza dei ricoveri potenzialmente inappropriati per gli anni dal 2005 al 2008, alcuni criteri di inclusione della casistica considerata sono stati modificati, risalendo con l'analisi anche al periodo 2001-2004, per uniformare i dati.⁶

⁶ Ulteriori esclusioni: la stadiazione (*disease staging*) di pochissimi record del database SDO non è stata possibile; non sono considerati i ricoveri per polmonite batterica e per epilessia rinviando tale casistica ad un'analisi mirata; dall'anno 2005 gli interventi chirurgici per emorroidi sono considerati potenzialmente inappropriati; dall'anno 2005 i DRG 221-222 (procedure sul ginocchio) sono stati sostituiti con il DRG 503; le strutture con codice COD_PRES = 080402-080412.

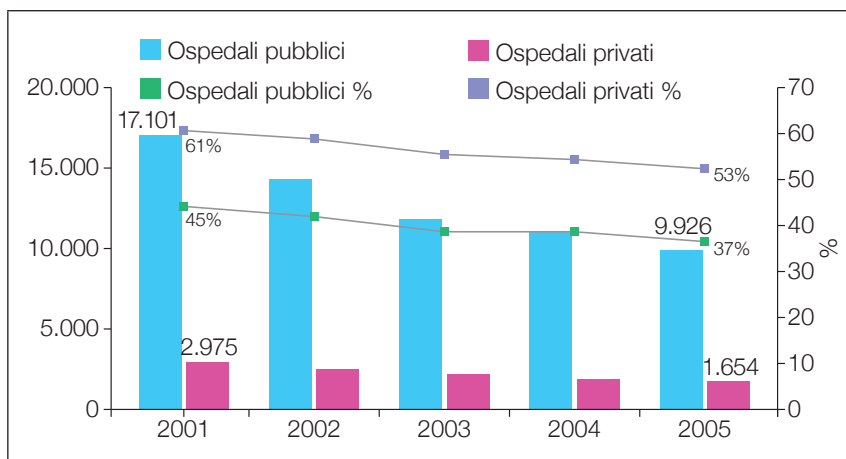


Figura 8.5 • Numero e percentuale di ricoveri inappropriati, Regione Emilia-Romagna, anni 2001-2004, ospedali pubblici e privati (modificata da Louis et al., 2008)

Queste correzioni non hanno inciso sui risultati e sulle conclusioni discusse e hanno influito solo marginalmente sulla numerosità dei casi riportati rispetto a quanto precedentemente pubblicato nell'*Atlante*.

Nella figura 8.6 sono quindi riportati i valori relativi ai profili di potenziale inappropriatezza dei ricoveri ordinari e day-hospital effettuati per i 62 DRG selezionati, per la popolazione compresa tra i 18 e i 75 anni, rilevati nel periodo dal 2001 al 2008 e stimati con i criteri di selezione aggiornati.

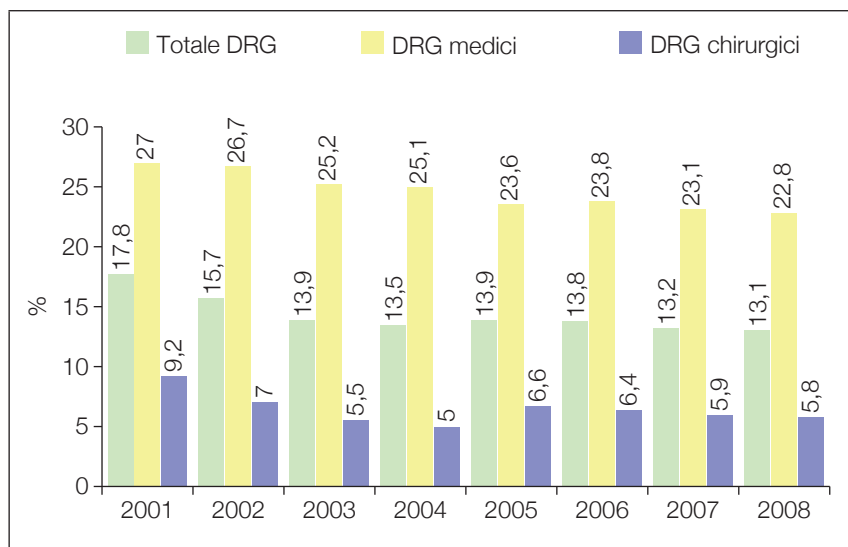


Figura 8.6 • Ricoveri potenzialmente inappropriati attribuiti ai DRG selezionati (% sul totale dei ricoveri ordinari e day-hospital, DRG medici e chirurgici), trend 2001-2008 (elaborazioni ASSR, dati non pubblicati)

9 • L'INIZIATIVA AMERICANA CHOOSING WISELY E IL PROGETTO ITALIANO

Luciana Ballini, Davide Botturi

L'INIZIATIVA AMERICANA CHOOSING WISELY E IL PROGETTO ITALIANO

Da alcuni anni si assiste, a livello internazionale, allo sviluppo di iniziative intese a fronteggiare il fenomeno del sovrautilizzo (*over-use*) di prestazioni sanitarie,¹ possibile fattore di rischio per la salute delle persone – in termini di esposizione a prestazioni sanitarie che, oltre a non offrire alcun beneficio, possono addirittura risultare dannose – e, in secondo luogo, fattore concorrente all'incremento della spesa sanitaria e all'uso non efficiente delle risorse.

Le soluzioni per contrastare il sovrautilizzo possono essere classificate in due macrocategorie: “quelle che cercano di disincentivare l'uso delle prestazioni sanitarie (come i sistemi di *co-payment* o l'introduzione di tetti di spesa) e quelle che cercano di intercettare e correggere l'inappropriatezza d'uso”, attraverso la definizione di criteri di appropriatezza (riferiti, ad esempio, a trattamenti, test diagnostici, percorsi assistenziali, altro) e lo sviluppo di line-guida e raccomandazioni a cui professionisti e pazienti sarebbero chiamati ad aderire (Ballini, 2012).

¹ Per un breve approfondimento della definizione di sovrautilizzo, si rimanda, tra le diverse fonti, al capitolo 3, box 3.1, del presente volume. Per approfondire, invece, le metodologie e gli strumenti di rilevazione si rimanda, tra le diverse fonti, a Chan et al. (2013). Tra i concetti strettamente connessi a quello di *over-use* vi è quello di *over-diagnosis*, recentemente oggetto di una pubblicazione tradotta in lingua italiana (Welch et al., 2013).

Le iniziative, di seguito descritte, appartengono a questa seconda macrocategoria. Il tratto distintivo di queste esperienze consiste nell'aver adottato un approccio "dal basso" (*bottom-up*) alla definizione dei criteri di appropriatezza. In altri termini, si tratta di iniziative promosse e coordinate da associazioni, società scientifiche, organizzazioni dei professionisti sanitari e dei cittadini o pazienti.

A livello internazionale, una delle più celebri iniziative è quella statunitense denominata "Choosing Wisely",² la quale ha rappresentato un esempio per lo sviluppo di iniziative simili in altri Paesi, come ad esempio in Canada³ e in Italia. L'iniziativa americana attinge ad un più ampio dibattito politico e scientifico, avviato qualche tempo prima, i cui aspetti più rilevanti sono stati sintetizzati nel box 9.1.

Choosing Wisely è stata avviata nel corso dell'anno 2012 dalla fondazione statunitense American Board of Internal Medicine (ABIM), in collaborazione con Consumer Reports, organizzazione no profit e indipendente di consumatori. Nove società scientifiche americane hanno aderito, fin dall'inizio, al progetto, in rappresentanza di circa 375.000 medici.⁴ A ciascuna società scientifica è stato affidato l'incarico di redigere un elenco delle 5 prestazioni più in uso nella propria disciplina medica ("Top five list") che possono essere ritenute a più alto rischio di inappropriata, in determinate situazioni. Ogni elenco cerca di offrire informazioni riguardanti le condizioni cliniche in cui un determinato test o una determinata prestazione potrebbero risultare maggiormente appropriati (cfr. le tabelle 9.1 e 9.2), e rappresenta una base informativa a supporto della discussione tra medico e paziente per la scelta del trattamento più appropriato.⁵

² Indirizzo web: <http://www.choosingwisely.org/>

³ L'iniziativa Choosing Wisely Canada è stata avviata ufficialmente nell'aprile 2014, ed organizzata dalla Canadian Medical Association e dall'University of Toronto. Per approfondimenti consultare l'indirizzo web: <http://www.choosingwiselycanada.org/>

⁴ American Academy of Allergy, Asthma & Immunology; American Academy of Family Physicians; American College of Cardiology; American College of Physicians; American College of Radiology; American Gastroenterological Association; American Society of Clinical Oncology; American Society of Nephrology; American Society of Nuclear Cardiology.

⁵ Un aspetto meritevole di riflessione è rappresentato dal fatto che 34 delle 45 prestazioni elencate (5 prestazioni per 9 società scientifiche) sono test diagnostici.

BOX 9.1 - L'INIZIATIVA CHOOSING WISELY: GLI ASPETTI PIÙ RILEVANTI DEL DIBATTITO

I fatti più salienti dell'ampio dibattito politico e scientifico che ha preceduto l'iniziativa americana sono:

- Il dibattito suscitato dalla riforma sanitaria del presidente Obama che esigeva una profonda revisione dei costi della sanità, al fine di estenderne la copertura alla maggior parte dei cittadini americani.
- La Carta della professionalità medica per il nuovo millennio, proposta nel 2002 dalla fondazione American Board of Internal Medicine, dalla fondazione American College of Physicians e dalla Federazione Europea di Medicina Interna. La Carta invita i medici ad assumersi la responsabilità dell'allocazione appropriata delle risorse e dell'evitare scrupolosamente test e procedure definibili come superflue, in quanto fornire servizi non necessari non solo espone i pazienti a rischi e costi evitabili ma riduce anche le risorse disponibili per gli altri.
- L'appello lanciato nel 2010 da Howard Brody, attraverso le pagine del *New England Journal of Medicine*, a medici e società scientifiche, affinché predisponessero un elenco di cinque procedure diagnostiche o terapeutiche comunemente utilizzate nella pratica clinica, e particolarmente costose, per le quali vi fosse la dimostrazione che non determinavano alcun beneficio significativo per la salute della persona ("Top five list").
- La serie "Less is more", avviata nel 2010 dagli *Archives of Internal Medicine*, nella quale sono pubblicati articoli scientifici che documentano casi in cui meno (*less*) assistenza sanitaria può equivalere a più (*more*) salute.
- Il progetto della National Physician Alliance, denominato "Promoting Good Stewardship in Clinical Practice", che ha portato alla definizione e pubblicazione nel 2011 negli *Archives of Internal Medicine* di un elenco di dodici attività cliniche, riferite alla medicina generale, alla medicina interna e alla pediatria, che potevano essere ritenute *over-used*. In una successiva analisi, pubblicata sempre sugli *Archives of Internal Medicine* con l'evocativo titolo "Top 5 Lists Top \$5 Billion", è stato, inoltre, stimato che se tali pratiche fossero state eliminate, il risparmio sarebbe consistito in almeno 5 miliardi di dollari all'anno.*

* Le fonti bibliografiche degli elementi sopra sintetizzati sono consultabili in Ballini (2012), Zuppiroli (2012). Inoltre La storia di *Choosing Wisely* è consultabile nel sito di *Slow Medicine* all'indirizzo web: <http://www.slowmedicine.it/fare-di-piu-non-significa-fare-meglio/la-storia-di-choosing-wisely.html>

TABELLA 9.1 - LA "TOP FIVE LIST" DELL'AMERICAN ACADEMY OF FAMILY PHYSICIANS (MODIFICATA DA GOOD STEWARDSHIP WORKING GROUP, 2011; BALLINI, 2012)

AMERICAN ACADEMY OF FAMILY PHYSICIANS

1. Non eseguire diagnostica per immagini nella lombalgia per le prime sei settimane, salvo in caso di concomitante grave sintomatologia.
2. Non prescrivere routinariamente antibiotici nella sinusite lieve o moderata, a meno che i sintomi non siano presenti da sette o più giorni o peggiorino dopo un iniziale miglioramento clinico.
3. Non eseguire test di screening DEXA per l'osteoporosi nelle donne al di sotto dei 65 anni di età e negli uomini al di sotto dei 70 anni in assenza di fattori di rischio.
4. Non eseguire elettrocardiogramma annualmente né altri test cardiaci di screening in pazienti asintomatici a basso rischio.
5. Non eseguire Pap Test nelle donne di età inferiore ai 21 anni o che hanno avuto un'isterectomia per malattia non oncologica.

TABELLA 9.2 - LA "TOP FIVE LIST" DELL'AMERICAN COLLEGE OF RADIOLOGY (MODIFICATA DA ACR, 2012 - [HTTP://WWW.CHOOSINGWISELY.ORG](http://www.choosingwisely.org); BALLINI, 2012)

AMERICAN COLLEGE OF RADIOLOGY

1. Non eseguire test diagnostici per emicranie non complicate.
2. Non eseguire diagnostica per immagini per sospetta embolia polmonare (PE) in assenza di moderata o elevata probabilità pre-test di PE.
3. Evitare Rx torace toracici all'ammissione o come esame pre-operatorio in pazienti ambulatoriali con anamnesi ed esami fisici nella norma.
4. Non effettuare tomografia computerizzata (TAC) per la valutazione di sospetta appendicite in bambini prima di aver considerato l'opzione dell'esame ecografico.
5. Non raccomandare diagnostica di follow-up per cisti ovariche clinicamente irrilevanti.

A fine febbraio 2014, 52 società scientifiche statunitensi hanno pubblicato le loro liste (alcune anche più di una); complessivamente, sono stati individuati 280 test e trattamenti ad alto rischio d'inappropriatezza.

In Italia, in analogia all'iniziativa americana, il movimento Slow Medicine⁶ ha lanciato in dicembre 2012 il progetto "Fare di più non significa fare meglio".⁷

Il presupposto è che, come è avvenuto negli Stati Uniti, la spinta all'utilizzo appropriato e senza sprechi delle risorse disponibili non possa che partire da una assunzione di responsabilità da parte dei professionisti della salute e in primo luogo dei medici, in alleanza con pazienti e cittadini. Lo scopo del progetto è di migliorare la qualità e la sicurezza dei servizi sanitari attraverso la riduzione di pratiche (esami diagnostici e trattamenti) che, secondo le conoscenze scientifiche disponibili, non apportano benefici significativi ai pazienti ai quali sono generalmente prescritte, ma possono, al contrario, esporli a rischi.⁸

Il progetto, a differenza di Choosing Wisely, intende coinvolgere non solo i medici ma anche gli altri professionisti della salute, in una prospettiva multidisciplinare. È inoltre prevista la partecipazione attiva dei cittadini e dei pazienti. Dall'avvio del progetto a marzo 2014 hanno aderito più di 20 società scientifiche mediche e infermieristiche e organizzazioni di professionisti⁹ e numerose associazioni, enti e istituti di ricerca.¹⁰ Le società scientifiche e le organizzazioni aderenti

⁶ Il movimento italiano Slow Medicine, rete di professionisti e cittadini, è stato fondato nel 2011 a Torino da parte di alcuni membri del direttivo della Società Italiana per la Qualità dell'Assistenza Sanitaria (SIQuAS), alcuni dei fondatori dell'Istituto CHANGE, agenzia formativa specializzata nella formazione alla comunicazione e al counseling sistemico in ambito sanitario, educativo e sociale, e altri professionisti.

⁷ Per approfondimenti consultare l'indirizzo: <http://www.slowmedicine.it/fare-di-piu-non-significa-fare-meglio/48-fare-di-piu-non-significa-fare-meglio.html>

⁸ Citazione testuale da "Il disegno del progetto – Sintesi" disponibile all'indirizzo <http://www.slowmedicine.it/fare-di-piu-non-significa-fare-meglio/il-disegno-del-progetto.html>

⁹ Ad esempio: Società Italiana di Radiologia Medica, Associazione Italiana di Neuroradiologia, Associazione Nazionale Medici Cardiologi Ospedalieri, Associazione Medici Diabetologi, Società Italiana di Cure Palliative.

¹⁰ Un elenco esemplificativo, e non esaustivo, è il seguente: Federazione Nazionale degli Ordini

hanno il compito di definire, con la collaborazione di esperti e di pazienti/cittadini e a partire dalle pratiche già individuate nell'iniziativa Choosing Wisely, una lista di cinque pratiche (esami diagnostici o trattamenti) che:

- sono effettuate molto comunemente in Italia;
- non apportano benefici significativi ai pazienti ai quali vengono generalmente prescritte;
- possono esporre i pazienti al rischio di subire effetti dannosi.

Come nell'iniziativa Choosing Wisely questi test e trattamenti ad alto rischio d'inappropriatezza dovranno essere oggetto di aperto dialogo nella relazione tra medico e paziente, per facilitare scelte informate e condivise.

A giugno 2014 otto società scientifiche hanno pubblicato le prime liste di cinque esami diagnostici e trattamenti a rischio d'inappropriatezza:

- Associazione italiana di dietetica e nutrizione clinica;
- Associazione italiana di radioterapia oncologica;
- Cochrane Neurological Field;
- Collegio italiano dei primari medici oncologi ospedalieri;
- Società italiana di radiologia medica (tabella 9.3);
- Società italiana di medicina generale (tabella 9.4);
- Società scientifiche di IPASVI (AICO, AIOSS, AIUC, ANIMO).
- Associazione Nazionale Medici Cardiologi Ospedalieri (tabella 9.5).

dei Medici Chirurghi e degli Odontoiatri (FMOMCeO); Federazione Nazionale Collegi Infermieri professionali, Assistenti sanitari, Vigilatrici d'infanzia (IPASVI), Progetto PartecipaSalute (promosso dall'IRCCS - Istituto di Ricerche Farmacologiche Mario Negri, dal Centro Cochrane Italiano e dall'agenzia Zadig di editoria scientifica), Altroconsumo (associazione di consumatori indipendente e senza fini di lucro).

TABELLA 9.3 - SOCIETÀ ITALIANA DI RADIOLOGIA MEDICA (SIRM):
LE CINQUE PRATICHE A RISCHIO DI INAPPROPRIATEZZA
(MODIFICATA DA [HTTP://WWW.SLOWMEDICINE.IT](http://www.slowmedicine.it))

1. Non eseguire risonanza magnetica (RM) del rachide lombosacrale in caso di lombalgia nelle prime sei settimane in assenza di segni/sintomi di allarme (semafori rossi o *red flags*).
2. Non eseguire di routine risonanza magnetica (RM) del ginocchio in caso di dolore acuto da trauma o di dolore cronico.
3. Non eseguire risonanza magnetica (RM) dell'encefalo per cefalea non traumatica in assenza di segni clinici di allarme.
4. Non eseguire radiografie del torace preoperatorie in assenza di sintomi e segni clinici che facciano sospettare patologie capaci di influire sull'esito dell'intervento.
5. Non eseguire di routine radiografia del cranio nel trauma cranico lieve.

TABELLA 9.4 - SOCIETÀ ITALIANA DI MEDICINA GENERALE (SIMG):
LE CINQUE PRATICHE A RISCHIO DI INAPPROPRIATEZZA
(MODIFICATA DA [HTTP://WWW.SLOWMEDICINE.IT](http://www.slowmedicine.it))

1. Non richiedere di routine esami di diagnostica per immagini in caso di lombalgia senza segni o sintomi di allarme (bandierine rosse o *red flags*).
2. Non prescrivere di routine antibiotici a pazienti affetti da infezioni acute delle vie aeree superiori. Valutarne l'opportunità nei pazienti a rischio di infezioni delle vie aeree inferiori o in caso di peggioramento del quadro clinico dopo qualche giorno.
3. Non prescrivere di routine inibitori di pompa protonica (IPP) a pazienti senza fattori di rischio per malattia ulcerosa. Nella malattia da reflusso gastroesofageo prescriverli alla più bassa dose in grado di controllare i sintomi, educando il paziente ad auspicabili periodi di sospensione.
4. Non prescrivere terapie con antinfiammatori non steroidei (FANS) senza valutare inizialmente e riconsiderare periodicamente la reale indicazione clinica e il rischio di effetti collaterali in quel momento e in quello specifico paziente.
5. Non prescrivere di routine in prima istanza benzodiazepine o farmaci Z* nei pazienti anziani in caso di insonnia. Raccomandarne comunque l'uso intermittente e non continuare cronicamente la terapia senza rivalutare periodicamente l'indicazione e l'eventuale comparsa di effetti indesiderati.

* I farmaci-Z (così denominati per le iniziali dei principi attivi, zaleplon, zolpidem e zopiclone) sono una classe di farmaci chimicamente diversi dalle benzodiazepine, ma che presentano un meccanismo d'azione simile.

TABELLA 9.5 - ASSOCIAZIONE NAZIONALE MEDICI CARDIOLOGI OSPEDALIERI (ANMCO):
LE CINQUE PRATICHE A RISCHIO DI IN APPROPRIATEZZA
(MODIFICATA DA BOBBIO ET AL., 2014)¹¹

1. Non richiedere ecocardiografia di controllo in pazienti con valvulopatia lieve-moderata o con disfunzione ventricolare sinistra, in assenza di nuovi sintomi, segni o eventi clinici.
2. Non richiedere di routine prova elettrocardiografica da sforzo di controllo in pazienti asintomatici dopo rivascolarizzazione chirurgica o percutanea.
3. Non richiedere registrazione Holter in pazienti con dolore toracico da sforzo che siano in grado di eseguire prova da sforzo, a meno che non vi sia anche il sospetto di aritmie.
4. Non richiedere test di imaging associato a test provocativo in fase di valutazione iniziale di sospetta cardiopatia ischemica.
5. Non richiedere prova elettrocardiografica da sforzo per screening di cardiopatia ischemica in pazienti asintomatici a basso rischio cardiovascolare.

Per lo sviluppo futuro del progetto alcune azioni sono da considerarsi prioritarie:¹²

- il coinvolgimento dei professionisti e la loro formazione, particolarmente centrata sull'EBM e sulla relazione medico-paziente;
- lo sviluppo di metodologie appropriate per l'identificazione delle prestazioni dismissibili e la misurabilità delle rispettive proporzioni di disinvestment realisticamente perseguibili;
- la messa a punto di strumenti di supporto alle decisioni cliniche relativi alle diverse pratiche, anche per contrastare l'esecuzione di prestazioni a scopo di "medicina difensiva";

¹¹ Per approfondimenti si invita a consultare anche Cherubini et al. (2014).

¹² Fonte: Documento congiunto elaborato a seguito dell'incontro, tenutosi il 4 marzo 2014 a Roma, tra Slow Medicine e il Comitato tecnico delle Regioni e Province Autonome per la Sicurezza delle cure, centrato su Choosing Wisely e il progetto italiano "Fare di più non significa fare meglio". Il documento è consultabile all'indirizzo: <http://www.slowmedicine.it/fare-di-piu-non-significa-fare-meglio/l-evoluzione-del-progetto/210-documento-congiunto-4-marzo-2014.html>

- l'attivazione di strumenti e progetti che si propongono di valutare l'impatto economico della riduzione delle pratiche ad alto rischio di inappropriatezza;
- l'adeguata informazione di cittadini e pazienti anche attraverso specifico materiale informativo, preparato in collaborazione tra professionisti e cittadini.

Non ultimo, nell'ambito del progetto viene anche curata la traduzione in lingua italiana dei vari elenchi di test e trattamenti ad alto rischio d'inappropriatezza costantemente proposti dall'iniziativa *Choosing Wisely*.¹³

Nel momento in cui questo volume viene pubblicato, il dibattito internazionale e nazionale sull'iniziativa *Choosing Wisely* è vivace e in continua evoluzione; si raccomanda quindi al lettore di esplorare autonomamente la letteratura per i necessari aggiornamenti. Nell'intento di offrire alcuni spunti per riflessioni necessariamente critiche e caute sulle potenzialità e sulle possibili ricadute concrete e future delle iniziative in corso, può essere utile tuttavia citare alcune posizioni recenti e autorevoli.

In particolare, nell'aprile 2014, sulla rivista *Academic Medicine*, Daniel Wolfson e John Santa, rispettivamente *executive vice-president* dell'American Board of Internal Medicine Foundation e *medical director* di Consumer Reports Health, con Lorie Slass, affermano che il fine ultimo dell'iniziativa *Choosing Wisely* consiste nel favorire il confronto tra medici e pazienti rispetto alle cure "realmente necessarie" (Wolfson et al., 2014). Ancora, nel giugno 2014 si è tenuta ad Amsterdam l'International Roundtable on *Choosing Wisely*, con la partecipazione di 12 Paesi, europei (tra gli altri Germania, Italia, Olanda, Regno Unito) ed extra-europei (tra gli altri Australia, Canada, Giappone, Stati Uniti). In questa occasione è stata posta l'enfasi sugli obiettivi principali delle campagne in atto, riconducibili soprattutto al miglioramento dei processi di *shared decision making* tra medici e pazienti. Secondo Daniel Wolfson le *top five list* non sono impor-

¹³ Per approfondimenti consultare l'indirizzo: <http://www.slowmedicine.it/fare-di-piu-non-significare-meglio/la-traduzione-della-pratiche.html>

tanti di per sé, ma solo come strumento per sostenere il ragionamento (Hurley, 2014). Durante la tavola rotonda è stata anche ribadita l'opportunità, per le nuove campagne, di concentrarsi prevalentemente su appropriatezza e sicurezza dell'assistenza, piuttosto che sui costi, pur meritevoli di attenzione (Hurley, 2014).

È evidente la molteplicità e la complessità dei temi (ad esempio, che cosa si intende per cure “realmente necessarie”?) e delle prospettive di analisi (etica, politica, economica, *evidence-based*) sottese alla cornice Choosing Wisely. Una loro trattazione rigorosa ed esaustiva esula dagli obiettivi di questo volume.

A chiusura di questo approfondimento si propongono alcuni interrogativi tuttora aperti, tra gli innumerevoli che sarebbe possibile elencare (Domenighetti e Vernerio, 2013):

- “...al di là dell'adesione dichiarata dalle società scientifiche (...) quanti saranno poi i medici realmente disposti a ridurre/evitare, laddove sia effettivamente necessario, le pratiche identificate come inappropriate?”;
- “...in che misura i cittadini-pazienti, che non possiedono competenze tecniche specifiche, saranno in grado di esprimere preferenze circa l'utilità o la futilità di tali prestazioni, quando direttamente interessati? E in che misura saranno in grado di non percepire l'invito a limitare l'uso di una prestazione inappropriata come un ‘razionamento implicito’?”.
- “Quali risultati dimostrabili è possibile ottenere?” e “in quanto tempo?”, “con quali strumenti organizzativi e professionali?”.

Quesiti, questi, a cui non è semplice rispondere nell'ambito di iniziative intese a modificare cultura e comportamenti e quindi inevitabilmente proiettabili su medio-lungo termine.

APPENDICE NORMATIVA

Davide Botturi

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.*

Nel D.Lgs. 229/1999 l'appropriatezza viene riconosciuta come uno dei principi da rispettare, da parte del Servizio sanitario nazionale, nel garantire i livelli essenziali di assistenza (LEA), nonché come criterio sulla cui base valutare l'inclusione/esclusione nei LEA delle tipologie di assistenza, dei servizi e delle prestazioni sanitarie. In particolare, si afferma nell'art. 1, *Tutela del diritto alla salute, programmazione sanitaria e definizione dei livelli essenziali e uniformi di assistenza*, al comma 2:

Il Servizio sanitario nazionale assicura [...] i livelli essenziali e uniformi di assistenza definiti dal Piano Sanitario Nazionale nel rispetto dei principi della dignità della persona umana, del bisogno di salute, dell'equità nell'accesso all'assistenza, della qualità delle cure e della loro appropriatezza riguardo alle specifiche esigenze, nonché dell'economicità nell'impiego delle risorse.

Sempre nell'art. 1, al comma 7:

Sono posti a carico del Servizio sanitario le tipologie di assistenza, i servizi e le prestazioni sanitarie che presentano, per specifiche condizioni cliniche o di rischio, evidenze scientifiche di un significativo beneficio in termini di salute, a livello individuale o collettivo, a fronte delle risorse impiegate. Sono esclusi dai livelli di assistenza erogati a carico del Servizio sanitario nazionale le tipologie di assistenza, i servizi e le prestazioni sanitarie che:

- a) non rispondono a necessità assistenziali tutelate in base ai principi ispiratori del Servizio sanitario nazionale di cui al comma 2;
- b) non soddisfano il principio dell'efficacia e dell'appropriatezza, ovvero la cui efficacia non è dimostrabile in base alle evidenze scientifiche disponibili o sono utilizzati per soggetti le cui condizioni cliniche non corrispondono alle indicazioni raccomandate;
- c) in presenza di altre forme di assistenza volte a soddisfare le medesime esigenze, non soddisfano il principio dell'economicità nell'impiego delle risorse, ovvero non garantiscono un uso efficiente delle risorse quanto a modalità di organizzazione ed erogazione dell'assistenza.

Inoltre, con riferimento alla dirigenza medica, viene sancito nell'art.15-decies *l'obbligo di appropriatezza*, affermando nei commi 1 e 2 che:

1. I medici ospedalieri e delle altre strutture di ricovero e cura del Servizio sanitario nazionale, pubbliche o accreditate, quando prescrivono o consigliano medicinali o accertamenti diagnostici a pazienti all'atto della dimissione o in occasione di visite ambulatoriali, sono tenuti a specificare i farmaci e le prestazioni erogabili con onere a carico del Servizio sanitario nazionale. Il predetto obbligo si estende anche ai medici specialisti che abbiano comunque titolo per prescrivere medicinali e accertamenti diagnostici a carico del Servizio sanitario nazionale.
2. In ogni caso, si applicano anche ai sanitari di cui al comma 1 il divieto di impiego del ricettario del Servizio sanitario nazionale per la prescrizione di medicinali non rimborsabili dal Servizio, nonché le

disposizioni che vietano al medico di prescrivere, a carico del Servizio medesimo, medicinali senza osservare le condizioni e le limitazioni previste dai provvedimenti della Commissione Unica del Farmaco, e prevedono conseguenze in caso di infrazione.

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

Il decreto definisce, ai sensi del D.Lgs. 30 dicembre 1992, n. 502, e successive modificazioni e integrazioni, e conformemente agli accordi sanciti in sede di Conferenza permanente per i rapporti tra lo Stato, le Regioni e le Province autonome di Trento e Bolzano, i livelli essenziali di assistenza sanitaria.

Il testo del decreto è integrato dai seguenti allegati:

Allegato 1 – Classificazione dei livelli (assistenza sanitaria collettiva in ambiente di vita e di lavoro, assistenza distrettuale, assistenza ospedaliera) e ricognizione della normativa vigente, con l'indicazione delle prestazioni erogabili, delle strutture di offerta e delle funzioni.

Allegato 2 – a) Prestazioni totalmente escluse dai LEA; b) Prestazioni parzialmente escluse, in quanto erogabili secondo specifiche indicazioni cliniche; c) Elenco di DRG ad alto rischio di non appropriatezza in regime di ricovero ordinario.

Allegato 3 – Indicazioni particolari per l'applicazione dei livelli in materia di assistenza ospedaliera, farmaceutica, specialistica e integrazione socio-sanitaria (l'allegato 3.1 contiene, inoltre, un elenco di indicatori – organizzativi e clinici – di livello ospedaliero).

Allegato 4 – Linee-guida alle Regioni in materia di LEA.

Conferenza permanente per i rapporti tra lo Stato, le Regioni e le Province Autonome di Trento e Bolzano. Patto per la Salute 2010-2012. Rep. n. 243 del 3 dicembre 2009.

Nel Patto per la Salute 2010-2012 viene proposto il monitoraggio di un ampio elenco di indicatori (allegato 3), prestazioni (allegato A) e DRG (allegato B) relativi all'appropriatezza delle cure. In particolare, si tratta di:

- 8 indicatori di appropriatezza organizzativa (ad esempio, degenza media preoperatoria, percentuale di interventi per frattura di femore operati entro due giorni, percentuale di dimessi da reparti chirurgici con DRG medici);
- più di 20 prestazioni ad alto rischio di non appropriatezza in regime di *day surgery*, trasferibili in regime ambulatoriale (ad esempio, liberazione del tunnel carpale, intervento di cataratta con o senza impianto di lente intraoculare, riparazione monolaterale di ernia inguinale diretta o indiretta, amputazione e disarticolazione di dita della mano);
- più di 100 DRG ad alto rischio di non appropriatezza in regime di degenza ordinaria (ad esempio, decompressione del tunnel carpale, interventi sulla retina, tonsillectomia e/o adenoidectomia età >17 anni, ipertensione, diabete età <36 anni, ipertrofia prostatica benigna senza CC, anomalie dei globuli rossi età <18 anni, radioterapia, assistenza riabilitativa con/senza anamnesi di neoplasia maligna come diagnosi secondaria).

Ministero della Salute. *Schema di Piano Sanitario Nazionale 2011-2013 (bozza). Anno 2011.*

Nella bozza del Piano Sanitario Nazionale (PSN) 2011-2013, l'appropriatezza di alcune prestazioni viene riconosciuta come una delle principali criticità del sistema sanitario italiano e, pertanto, inclusa tra le prospettive di miglioramento, attraverso il coinvolgimento dei professionisti.

Nel capitolo 1 (*Il Servizio Sanitario Nazionale nel contesto europeo e strategie d'internazionalizzazione*), paragrafo 1.5 (*Criticità del sistema e prospettive di miglioramento*), si afferma:

[...] non si può nascondere che esso [il Servizio sanitario italiano] presenti seri elementi di criticità, quali l'inappropriatezza di alcune prestazioni, come l'utilizzo improprio dei ricoveri ospedalieri e dei servizi di pronto soccorso, dovuto all'organizzazione ancora insufficiente della medicina generale e al livello medio dei servizi territoriali e di assistenza domiciliare integrate [...]

E inoltre:

La valutazione preliminarmente condivisa è che gradi sempre più elevati di appropriatezza clinica ed organizzativa possano essere conseguiti con un processo che va sviluppato con il coinvolgimento dei professionisti operanti nel campo della tutela della salute.

In particolare, il PSN si concentra sul miglioramento dell'appropriatezza dell'assistenza specialistica ambulatoriale (cfr. capitolo 1, paragrafo 1.5, sottoparagrafo 1.5.1 *Miglioramento dell'appropriatezza*), attraverso un'azione che riguarda più leve:

- [...] promozione dell'appropriatezza nel ricorso ai ricoveri ospedalieri, e quindi il passaggio di nuove prestazioni dal ricovero diurno

all'assistenza in regime ambulatoriale (cfr. Patto per la Salute 2010-2012);¹

- [...] diffusione [...] adeguata di assistenza domiciliare e residenziale sul territorio nazionale per erogare prestazioni, specialmente terapeutiche, [...] in *setting* assistenziali più appropriati, [...] *rispetto* al regime ambulatoriale;
- [...] corretto uso del *day hospital* rispetto all'ambulatorio e attivazione [...] della modalità di *day service*;
- azione sistematica di *technology assessment* [...], che includa i processi di pianificazione, acquisto e gestione delle apparecchiature biomediche e che sia prioritariamente orientata [...] a perseguirne l'appropriato utilizzo;
- [...] realizzazione di adeguati strumenti di supporto alla programmazione sanitaria, tenuto conto anche di quanto reso disponibile attraverso il Sistema Tessera Sanitaria, che consentano di misurare l'appropriatezza dell'assistenza specialistica ambulatoriale e, conseguentemente, di mettere in atto opportuni piani di governo della domanda che agiscano direttamente sul comportamento prescrittivo.

Una trattazione più approfondita del tema dell'appropriatezza viene effettuata nel capitolo 3 (*Monitoraggio appropriatezza e uniformità dei LEA*), paragrafo 3.3 (*Appropriatezza*), in cui vengono distinti due ambiti dell'appropriatezza, clinico e organizzativo, e per ciascuno identificate le aree di lavoro; viene inoltre richiamata la normativa nazionale di riferimento e vengono dichiarati gli obiettivi per il triennio 2011-2013:

Nelle politiche dell'SSN l'appropriatezza rimane un obiettivo costante ed in continua evoluzione, identificandosi con la maggior parte degli obiettivi della qualità di un sistema sanitario, in rapporto al mutamento del quadro epidemiologico, delle scoperte scientifiche, dell'evoluzione tecnologica, della situazione economica e politica. All'interno dei due tradizionali ambiti dell'appropriatezza, quello clinico e quello organiz-

¹ Provvedimento 3 dicembre 2009 della Conferenza Permanente per i Rapporti tra lo Stato, le Regioni e le Province Autonome di Trento e Bolzano. Patto per la Salute per gli anni 2010-2012. Pubblicato sulla Gazzetta Ufficiale, Serie Generale n. 3 del 5 gennaio 2010.

zativo, si individuano anche delle importanti aree di lavoro, come di seguito riportate.

Appropriatezza clinica

- *Clinical governance*
- Appropriatezza prescrittiva (governo della domanda) che riguarda anche l'utilizzo delle tecnologie diagnostiche e terapeutiche (interventi, farmaci, ecc.) ed il contenimento del fenomeno del "difensivismo medico" mediante l'adozione di procedure prestabilite.

Appropriatezza organizzativa

- Luogo di cura che si associa all'appropriatezza di utilizzo dei regimi assistenziali sia a livello ospedaliero (attraverso lo spostamento progressivo dei 148 DRG dal regime di ricovero ordinario a quello di DH/DS e ambulatoriale) che territoriale (anche attraverso il concreto avvio del Punto Unico di Accesso verso le prestazioni residenziali, semiresidenziali e domiciliari) come anche previsto dall'art. 6 del Patto della Salute 2010-2012.
- Gestionale/organizzativa.

In particolare l'art. 6 del Patto della Salute 2010-2012 prevede una razionalizzazione della rete ospedaliera ottenuta attraverso l'incremento dell'appropriatezza dei ricoveri, e la contestuale promozione del passaggio dal ricovero ordinario al ricovero diurno, dal ricovero diurno all'assistenza in regime ambulatoriale e, in generale, dal ricovero all'assistenza residenziale e domiciliare. I rimanenti ricoveri saranno destinati a casi complessi a cui si fornirà un'assistenza altamente qualificata. Nel comma 5 dello stesso articolo si prevede l'estensione delle prestazioni ad alto rischio di inappropriately, con passaggio da 43 DRG, di cui all'allegato 2C del D.P.C.M. 29 novembre 2001, alle liste contenute negli elenchi A e B allegati allo stesso Patto della Salute. Le Regioni dovranno, a tal fine, stabilire i valori soglia entro cui ammettere i ricoveri ordinari, prevedere le tariffe da attribuire alle prestazioni e definire le misure da adottare per le prestazioni extrasoglia. Gli indirizzi sopra indicati comportano l'esigenza di una riprogettazione organizzativa assistenziale, finalizzata allo sviluppo progressivo di

risposte che privilegino altri livelli assistenziali sia presso l'ospedale (l'osservazione breve, il *day service*, i percorsi ambulatoriali complessi nei *day service*, le prestazioni ambulatoriali), sia in sede territoriale (strutture residenziali e semiresidenziali, case della salute, ospedali di comunità) e, infine, presso il domicilio del paziente con l'attivazione delle cure domiciliari di complessità appropriata al bisogno espresso. La riduzione dei ricoveri e delle prestazioni non appropriate di Pronto soccorso (codici bianchi e, in parte, codici verdi) presso le strutture pubbliche è un presupposto per la concentrazione e la ricostruzione della rete ospedaliera pubblica. Riducendo i ricoveri non appropriati è possibile utilizzare le Aziende ospedaliere, gli *Hub* e i centri *Spoke* per i ricoveri appropriati, anche per i pazienti provenienti dai piccoli ospedali da riconvertire.

Obiettivi nel triennio sono:

- predisporre a livello regionale un *report* riepilogativo annuale delle azioni e dei risultati ottenuti per l'aumento dell'appropriatezza nel campo di utilizzo delle tecnologie e dei regimi assistenziali;
- implementare a livello regionale logiche e prassi dei processi assistenziali sulle patologie a maggior impatto *quoad vitam* (infarto miocardico acuto, ictus cerebrale) e *quoad valetudinem* (scompenso cardiaco, frattura del femore negli anziani);
- inserire il raggiungimento di obiettivi misurabili di appropriatezza in termini di interventi sulla salute, di soddisfazione dei cittadini e di utilizzo delle risorse tra gli aspetti più rilevanti nella valutazione dei direttori generale e dei dirigenti.

**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.*²
Febbraio 2011 (sintesi)**

OBIETTIVO E FASI DI LAVORO

L'obiettivo dello studio consiste nella predisposizione di un rapporto di analisi dell'appropriatezza di alcune prestazioni ospedaliere di ricovero, determinate da un'offerta presumibilmente inadeguata di servizi territoriali, attraverso l'utilizzo di indicatori specifici calcolabili con i dati delle schede di dimissione ospedaliera (SDO). Le fasi di lavoro sono tre:

1. individuazione degli indicatori attraverso una raccolta bibliografica e una rassegna critica della normativa nazionale e delle principali pubblicazioni realizzate in tema di misura dell'appropriatezza;
2. elaborazione degli indicatori attraverso l'utilizzo delle informazioni raccolte nelle SDO, finalizzata a misurare alcuni aspetti dell'appropriatezza;
3. analisi descrittiva del livello di appropriatezza raggiunto per ciascuna Regione italiana.

METODOLOGIA

Indicatori. Sono stati selezionati 19 indicatori di appropriatezza, sulla base dei set di indicatori proposti da organizzazioni internazionali, agenzie o istituzioni, ad esempio: OCSE, AHRQ, Nuovo Sistema Indicatori di aggiornamento del Decreto Ministeriale del 12 dicembre 2001, Progetto di ricerca finalizzata, ex art. 12 D.Lgs. 502/1992, co-

² Consultabile al link: http://www.salute.gov.it/imgs/C_17_pubblicazioni_1492_allegato.pdf

ordinato dall'Agenzia Nazionale per i Servizi Sanitari Regionali negli anni 2002-2003 sul tema degli indicatori di qualità.

Il gruppo di indicatori individuato è stato ritenuto idoneo a valutare in particolare l'ospedalizzazione evitabile per alcune patologie croniche, la cui gestione attiene prevalentemente alle competenze territoriali. Ciò al fine di individuare e caratterizzare le eventuali carenze nell'offerta dei servizi extraospedalieri. Tutti gli indicatori individuati sono stati descritti in modo analitico, attraverso una scheda tecnica, contenente la definizione, il significato, le modalità di calcolo (comprese eventuali tecniche di aggiustamento), la validità e i limiti.

Fonti. Le fonti informative utilizzate sono le schede di dimissione ospedaliera relative ai ricoveri effettuati nel periodo 2001-2009, e i dati sulla popolazione residente al 1° gennaio dell'anno di riferimento, per il periodo 2001-2009, strutturati per sesso ed età, pubblicati dall'ISTAT. Le standardizzazioni, ove effettuate, sono riferite alla popolazione del Censimento 2001.

Analisi dei dati. La sperimentazione degli indicatori di appropriatezza è stata effettuata, con dettaglio regionale, sulle strutture pubbliche e private accreditate, escludendo quindi le dimissioni dalle case di cura private non accreditate, in linea con quanto previsto dal decreto ministeriale 12 dicembre 2001 sul sistema di monitoraggio del rispetto dei livelli di assistenza.

DISCUSSIONE E CONCLUSIONI

Il tema della potenziale inappropriata ospedalizzazione determinata da carenze nel livello di assistenza territoriale è di complessa valutazione per la difficoltà di individuare indicatori che misurino direttamente la relazione fra questi diversi livelli di assistenza sanitaria. Nel presente lavoro è stato compiuto lo sforzo di individuare gli indicatori fattibili, secondo i flussi informativi correnti, e capaci di esplorare, seppure in modo indiretto, la relazione tra ospedalizzazione e potenziale inadeguatezza dei servizi territoriali.

Dal punto di vista globale, i risultati evidenziano un miglioramento sostanziale della situazione nazionale nel lungo periodo considerato. Se si analizzano i risultati è ragionevole ipotizzare che:

- nelle Regioni del Centro-Nord, dove esistono reti territoriali più sviluppate, venga trattata nel territorio una quota significativa di casistica;
- nelle Regioni meridionali, invece, [la casistica] viene presa in carico per lo più dall'ospedale, al fine di assicurare comunque, anche in condizioni di inadeguatezza della rete dei servizi territoriali, la gestione complessiva dell'iter diagnostico-terapeutico dei singoli pazienti assistiti.

È necessario mettere a punto metodologie e strumenti di valutazione per favorire lo sviluppo delle reti territoriali e promuovere l'integrazione e la continuità fra ospedale e territorio al fine di sostenere gli obiettivi di appropriatezza dell'assistenza sanitaria.

È inoltre importante monitorare nel tempo alcuni indicatori, e studiare possibili approfondimenti per descrivere in modo più accurato le caratteristiche di appropriatezza delle prestazioni erogate nei vari livelli di assistenza.

BIBLIOGRAFIA E SITOGRAFIA

- AHRQ (2007). Guide to prevention quality indicators. [Link alla risorsa]
- Alakeson V (2008). Why Oregon went wrong. *BMJ* 337: a2044. [Link alla risorsa]
- American Geriatrics Society (2012). American Geriatrics Society updates Beers criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc* Apr 60 (4): 616-31. [Link alla risorsa]
- Amos TB, Keith SW, Del Canale S et al. (2014). Inappropriate prescribing in a large community-dwelling older population: a focus on prevalence and how it relates to patient and physician characteristics. *J Clin Pharm Ther*, Oct 1: doi: 10.1111/jcpt.12212.
- Andersen TF, Mooney G (eds) (1990). The challenges of medical practice variations. London: Macmillan Press.
- Apolone G, Tampieri A, Bevilacqua L et al. (1993). L'appropriatezza dell'uso dell'ospedale: i risultati di uno studio all'ospedale S. Gerardo di Monza. *Fed Med* 8: 27-32.
- Audit Commission (2001). Day surgery. [Link alla risorsa]
- Baglio G, Materia E, Vantaggiato G, Perucci CA (2001). Valutare l'appropriatezza dei ricoveri con i dati amministrativi. *Tendenze Nuove* 1: 51-70. [Link alla risorsa]
- Bakwin H (1958). The tonsil-adenoidectomy enigma. *Journal of Pediatrics* 52 (3): 339-61.
- Balduzzi R (2004). 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. Bologna: Il Mulino.

- Ballini L (2012). L'iniziativa Choosing Wisely® rilancia la scelta ragionata di medici e pazienti. *Politiche sanitarie* 13, 4, ottobre-dicembre. [Link alla risorsa]
- Ballini L, Vignatelli L, Negro A, Maltoni S, Longo G (2011). Criteria for appropriate use of FDG-PET in esophageal cancer. Dossier 209. Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna. [Link alla risorsa]
- Beers MH (1997). Explicit criteria for determining potentially inappropriate medication use by the elderly. An update. *Arch Intern Med* 157: 1531-6. [Link alla risorsa]
- Berghthold LA (1995). Medical necessity: do we need it? *Health Aff (Millwood)* 14: 180-90. [Link alla risorsa]
- Berti E, Casolari L, Cisbani L et al. (2004). Il monitoraggio delle performance cliniche. In: Grilli R, Taroni F (eds). *Governo clinico. Governo delle organizzazioni sanitarie e qualità dell'assistenza*. Roma: Il Pensiero Scientifico Editore.
- Bobbio M, Abrignani MG, Caldarola P et al. (2014). Fare di più non significa fare meglio. Le proposte dell'ANMCO. *G Ital Cardiol* 15 (4): 244-52. [Link alla risorsa]
- Brook RH (1999). Ensuring delivery of necessary care in the United States. Testimony presented to the Senate Committee on health, education, labor, and pensions. [Link alla risorsa]
- Brook RH, Chassin MR, Fink A et al. (1986). A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 2: 53-63. [Link alla risorsa]
- Buetow SA, Sibbald B, Cantril JA et al. (1997). Appropriateness in health care: applications to prescribing. *Soc Sci Med* 45: 261-71. [Link alla risorsa]
- Burton MJ, Glasziou PP (2009). Tonsillectomy or adeno-tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis. *Cochrane Database Syst Rev* 1: CD001802. [Link alla risorsa]
- Cadum E, Costa G, Biggeri A, Martuzzi M (1999). Deprivazione e mortalità: un indice di deprivazione per l'analisi delle disuguaglianze su base demografica. *Epidemiol Prev* 23: 175-87.
- Caulfield TA (1996). Wishful thinking: defining "medically necessary" in Canada. *Health Law I* 4: 63-85. [Link alla risorsa]
- Chalmers I (1993). The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci* 31: 703-63. [Link alla risorsa]

- Chan KS, Chang E, Nassery N, Chang H-Y, Segal JB (2013). The state of overuse measurement: a critical review. *Medical Care Research and Review* 70 (5): 473-96. [Link alla risorsa]
- Charles C, Lomas I, Giacomini I et al. (1997). Medical necessity in Canadian health policy. Four meanings and... a funeral? *Milbank Q* 75: 365-94. [Link alla risorsa]
- Chassin MR, Kosecoff J, Park RE et al. (1987). Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA* 258 (18): 2533-7. [Link alla risorsa]
- Cherubini A, Mureddu GF, Temporelli PL et al. (2014). Appropriately delle procedure diagnostiche in prevenzione cardiovascolare: di che cosa possiamo fare a meno? *G Ital Cardiol* 15 (4): 253-63. [Link alla risorsa]
- Cinotti R (2005). Glossario della qualità e dell'accreditamento. Bologna: Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna.
- Cochrane AL (1999). Efficienza ed efficacia. Riflessioni sui servizi sanitari. Seconda edizione italiana. Roma: Il Pensiero Scientifico Editore.
- Comitato dei Ministri del Consiglio d'Europa (1997). Sullo sviluppo e l'attivazione dei sistemi di miglioramento della qualità dell'assistenza sanitaria. Raccomandazione n. 17, 30 settembre. [Link alla risorsa]
- Consiglio d'Europa (1997). Convenzione sui Diritti dell'Uomo e la biomedicina. Oviedo, 4 aprile. [Link alla risorsa]
- Department of Health (2006). On the state of the public health. Annual report of the Chief Medical Officer 2005. [Link alla risorsa]
- Dixon T, Shaw M, Ebrahim E, Dieppe P (2004). Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need. *Ann Rheum Dis* 63 (7): 825-30. [Link alla risorsa]
- Domenighetti G, Vernero S (2013). Fare di più non significa fare meglio. *Salute Internazionale*. [Link alla risorsa]
- Donabedian A (2010). Il Maestro e le margherite. La qualità dell'assistenza sanitaria secondo Avedis Donabedian. Edizione italiana. Roma: Il Pensiero Scientifico Editore.
- Doyle JJ (2010). Returns to local-area healthcare spending: evidence from health shocks to patients far from home. [Link alla risorsa]
- Economic Policy Committee and European Commission (DG ECFIN) (2006). The impact of ageing on public expenditure: projections for the EU25 Member States on pensions, health-care, long term care, education and unemployment transfers (2004-2050). Special report, 1. [Link alla risorsa]

- Evans RG (1990). The dog in the night-time: medical practice variations and health policy. In: Andersen TF, Mooney G (eds). The challenges of medical practice variations. London: Macmillan Press.
- Fellin G, Apolone G, Tampieri A et al. (1995). Appropriateness of hospital use: an overview of Italian studies. *Int J Qual Health Care* 7 (3): 219-25. [Link alla risorsa]
- Fick DM, Cooper JW, Wade WE, Waller JL, Maclean JR, Beers MH (2003). Updating the Beers criteria for potentially inappropriate medication use in older adults: results of a US consensus panel of experts. *Arch Intern Med* 163: 2716-24.
- Fink A, Kosecoff J, Chassin M et al. (1984). Consensus methods: characteristics and guidelines for use. *Am J Public Health* 74: 979-83.
- Fiorini M, Vizioli M, Melotti R, Louis DZ, Grilli R (2006). Atlante dell'appropriatezza organizzativa. I ricoveri ospedalieri in Emilia-Romagna. Dossier ASR 132/2006. Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna. [Link alla risorsa]
- Fisher ES, Bell JE, Tomek I, Esty A, Goodman D (2010). Trends and regional variation in hip, knee and shoulder replacement. *Dartmouth Atlas of Health Care*. [Link alla risorsa]
- Fisher ES, Wennberg D, Stukel T, Gottlieb DJ, Lucas FL, Pinder EL (2003a). The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Ann Intern Med* 138 (4): 273-87. [Link alla risorsa]
- Fisher ES, Wennberg D, Stukel T, Gotthieb DJ, Ducas FL, Pinder EL. (2003b). The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Ann Intern Med* 138 (4): 288-98. [Link alla risorsa]
- Fitch K, Bernstein FJ, Aguilar MD et al. (2000). The RAND/UCLA appropriateness method user's manual. RAND Corporation, Sept. [Link alla risorsa]
- Fortino A, Lispi L, Materia E, Di Domenicantonio R, Baglio G (2002). La valutazione dei ricoveri ospedalieri in Italia con il metodo APPRO. Ministero della Salute e Agenzia di Sanità Pubblica del Lazio. [Link alla risorsa]
- France G, Taroni F, Donatini A (2005). The Italian health-care system. *Health Economics* 14: S187-202. [Link alla risorsa]
- Glover AJ (1938). The incidence of tonsillectomy in school children. *Proceedings of the Royal Society of Medicine* 31 (10): 1219-36. [Link alla risorsa]

- Good Stewardship Working Group (2011). The “Top 5” lists in primary care. *Arch Intern Med* 171 (15): 1385-90. [Link alla risorsa]
- Goodman D, Mick SS, Bott D et al. (2010). Regional and racial variation in primary care and the quality of care among medicare beneficiaries. Report of the Dartmouth Atlas Project. September 9. [Link alla risorsa]
- Gruppo di lavoro PRUO (1994). Revisione e valutazione dell’uso dell’ospedale. *Ricerca e Pratica* 58: 99-117. [Link alla risorsa]
- Ham C (ed) (1988). *Health care variations: assessing the evidence*. London: King’s Fund Institute.
- Hawker GA, Wright JG, Glazier RH et al. (2002). The effect of education and income on need and willingness to undergo total joint arthroplasty. *Arthritis Rheum* 46 (12): 3331-9. [Link alla risorsa]
- Hopkins A, Fitzpatrick R, Foster A et al. (1993). What do we mean by appropriate health care? Report of a working group prepared for the Director of Research and Development of the Department of the NHS Management Executive. *Qual Health Care* 2: 117-23.
- Hurley R (2014). Can doctors reduce harmful medical overuse worldwide? *BMJ* 349: g4289, doi: 10.1136/bmj.g4289.
- Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan YD (1996). Judging hospitals by severity-adjusted mortality rates: the influence of the severity-adjustment method. *Am J Public Health* 86: 1379-87. [Link alla risorsa]
- Istituto Superiore di Sanità. Programma nazionale per le linee guida (ISS-PNLG) (2003). *Appropriatezza clinica e organizzativa degli interventi di tonsillectomia e/o adenoidectomia*. Documento di indirizzo. PNLG 4. [Link alla risorsa]
- Istituto Superiore di Sanità. Sistema nazionale per le linee guida (ISS-SNLG) (revisione 2008). *Appropriatezza e sicurezza degli interventi di tonsillectomia e/o adenoidectomia*. Documento 15. [Link alla risorsa]
- Jacobson PD, Asch S, Glassman PA, Model KE, Hernandez JB (1997). Defining and implementing medical necessity in Washington State and Oregon. *Inquiry* 34: 143-54. [Link alla risorsa]
- Joint Commission on Accreditation of Healthcare Organizations (1998). *Guide to quality assurance*. Chicago: JCAHO.
- Judge A, Welton NJ, Sandhu J, Ben-Shlomo Y (2010). Equity in access to total joint replacement of the hip and knee in England: cross sectional study. *BMJ* 341, c4092. [Link alla risorsa]
- Kahan JP, Bernstein SJ, Leape LL et al. (1994). Measuring the necessity of medical procedures. *Med Care* 32: 357-65. [Link alla risorsa]

- Kahan JP, Van Het LM (1999). Defining appropriate health care. *Eurohealth* 5 (3): 16-8. [Link alla risorsa]
- Kanouse DE, Brook RH, Winkler JD, et al. (1989). Changing medical practice through technology assessment: an evaluation of the NIH Consensus Development Program. RAND Report No. R-3452-NIH/RC.
- Kaufmann CP, Tremp R, Hersberger KE, Lampert ML (2014). Inappropriate prescribing: a systematic overview of published assessment tools. *Eur J Clin Pharmacol* 70 (1): 1-11. [Link alla risorsa]
- Keith SW, Maio V, Dudash K, Templin M, Del Canale S (2013). A physician-focused intervention to reduce potentially inappropriate medication prescribing in older people: a 3-year, Italian, prospective, proof-of-concept study. *Drugs Aging* 30 (2): 119-27. [Link alla risorsa]
- Kelley E, Hurst J (2006). Health care quality indicators project: conceptual framework paper. Paris: Organisation for Economic Co-operation and Development Publishing (OECD Health working paper, n.23). [Link alla risorsa]
- Lattuada L, Francescutti C, Borgnolo G, Simon G, Rizzi L (1998). La valutazione dell'appropriatezza dei ricoveri nell'esperienza del Friuli-Venezia Giulia. *Tendenze Nuove* 2: 4-8. [Link alla risorsa]
- Lattuada L, Francescutti C, Sampognaro E (2001). L'appropriatezza dei ricoveri: come rilevarla e migliorarla con il Protocollo di Revisione dell'Utilizzo dell'Ospedale (PRUO). Marenco di Piave (TV): Edizioni Vega-Q&S.
- Lavis JN, Anderson GM (1996). Appropriateness in health care delivery. Definitions, measurement and policy implications. *CMAJ* 154: 321-8. [Link alla risorsa]
- Liberati A, Longo G, Formoso G, Ballini L (2003). Indicazioni all'utilizzo della FDG-PET in oncologia. Analisi critica della letteratura scientifica. Dossier n. 81. Bologna: Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna. [Link alla risorsa]
- Liberati A, Longo G, Ballini L, De Palma R (2007). FDG-PET in oncologia. Criteri per un uso appropriato. Dossier n. 157. Bologna: Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna. [Link alla risorsa]
- Linstone, HA, Turoff M (1975). The Delphi method: techniques and applications. Reading (MA): Addison-Wesley.
- Lopatto J, Keith SW, Del Canale S, Templin M, Maio V (2014). Evaluating sustained quality improvements: long-term effectiveness of a physician-focused intervention to reduce potentially inappropriate medication prescribing in an older population. *J Clin Pharm Ther* 39: 266-71. [Link alla risorsa]

- Louis D, Taroni F, Melotti R et al. (2008). Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 13 (4): 202-8. [Link alla risorsa]
- Maio V, Yuen EJ, Novielli K, Smith KD, Louis DZ (2006). Potentially inappropriate medication prescribing for elderly outpatients in Emilia-Romagna, Italy: a population-based cohort study. *Drugs Aging* 23 (11): 915-24. [Link alla risorsa]
- Maio V, Del Canale S, Abouzaid S; GAP Investigators (2010). Using explicit criteria to evaluate the quality of prescribing in elderly Italian outpatients: a cohort study. *J Clin Pharm Ther* 35: 219-29. [Link alla risorsa]
- Maio V, Jutkowitz E, Herrera K, Abouzaid S, Negri G, Del Canale S (2011). Appropriate medication prescribing in elderly patients: how knowledgeable are primary care physicians? A survey study in Parma, Italy. *J Clin Pharm Ther* 36: 468-80. [Link alla risorsa]
- Materia E (2003). Appropriatezza: origini, implicazioni, valutazione. *Tendenze Nuove* 4-5: 343-54. [Link alla risorsa]
- Ministero della Salute (2006). Progetto Mattoni. Misura dell'appropriatezza. 1.2 Analisi appropriatezza organizzativa. "Proposta metodologia per la modifica DPCM LEA elenco DRG". Regione Umbria. [Link alla risorsa]
- Ministero della Salute (2007). Progetto Mattoni. Evoluzione del sistema DRG nazionale. Milestone 1.2.1. Forme alternative alla degenza: definizione (day hospital, day surgery, day service) e relativi flussi informativi. Regione Emilia-Romagna. [Link alla risorsa]
- Morosini PL, Perraro F (2001). *Enciclopedia della gestione di qualità in sanità*. 2a edizione ampliata e riveduta. Torino: Centro Scientifico Editore.
- Morris R, Carstairs V (1991). Which deprivation? A comparison of selected deprivation indexes. *Public Health Med* 13 (4): 318-26. [Link alla risorsa]
- Mulley AJ (2010). Improving productivity in the NHS. *BMJ* 341 (7766): c3965. [Link alla risorsa]
- Naylor CD (1998). What is appropriate care? *N Engl J Med* 338: 1918-20. [Link alla risorsa]
- OECD (2011). Health reform: meeting the challenge of ageing and multiple morbidities. Conference. Wednesday 22 June. [Link alla risorsa]
- Parchman ML (1995). Small area variation analysis: a tool for primary care research. *Fam Med* 27 (4): 272. [Link alla risorsa]
- Park RE, Fink A, Brook RH et al. (1986). Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health* 76: 766-72. [Link alla risorsa]

- Payne SM, Ash A, Restuccia JD (1991). The role of feed-back in reducing medically unnecessary hospital use. *Medical Care* 29 (Suppl) AS 91-AS 106. [Link alla risorsa]
- Pesenti Compagnoni M (1991). Appropriately dei ricoveri ospedalieri: valutazione di un servizio di pronto soccorso secondo i criteri AEP. *QA* 1-2: 32-8.
- Porter ME (2010). What is value in health care. *N Engl J Med* 363 (26): 2477-81. [Link alla risorsa]
- Pronovost PJ, Goeschel CA (2011). Time to take health delivery research seriously. *JAMA* 306 (3): 301-11. [Link alla risorsa]
- Regional Office for Europe, WHO (2000). Appropriateness in health care services. Report. Koblenz, Germany, 23-25 March. [Link alla risorsa]
- Roos NP, Freeman JL (1989). Potential for inpatient-outpatient substitution with diagnosis-related groups. *Health Care Financ Rev* 10: 31-8. [Link alla risorsa]
- Sackett DL (1989). Roles of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 95 (Suppl 2): 2S-4S. [Link alla risorsa]
- Sanmartin C, Murphy K, Choptain N et al. (2008). Appropriateness of healthcare interventions: concepts and scoping of the published literature. *Int J Technol Assess Health Care* 24 (3): 342-9. [Link alla risorsa]
- Servizio Assistenza Distrettuale - Regione Emilia-Romagna, Thomas Jefferson University, Philadelphia; Azienda USL di Parma (2012). Appropriately prescrittiva nella popolazione anziana. Elenco farmaci potenzialmente inappropriati ed eventuali alternative terapeutiche. [Link alla risorsa]
- Sharpe VA (1997). The politics, economics and ethics of "appropriateness". *Kennedy Inst Ethics J* 7: 337-43. [Link alla risorsa]
- Sharpe VA, Faden AI (1996). Appropriateness in patient care: a new conceptual framework. *Milbank Q* 74: 115-38. [Link alla risorsa]
- Schünemann HJ, Oxman AD, Brozek J et al. (2008). Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 336 (7653): 1106-10. [Link alla risorsa]
- Sica GT (2006). Bias in research studies. *Radiology* 238: 780-9. [Link alla risorsa]
- Skinner J, Fisher E (2010). Reflections on Geographic Variations in U.S. Health Care. Dartmouth Institute for health policy and clinical practice. May 12. [Link alla risorsa]
- Smeets PM, Verheggen FW, Pop P, Panis LJ, Carpay JJ (2000). Assessing the necessity of hospital stay by means of the appropriateness evaluation

- protocol: how strong is the evidence to proceed? *Int J Qual Health Care*; 12 (6): 483-93. [Link alla risorsa]
- Stoppelli P (ed.) (2009). Grande dizionario dei sinonimi e contrari. CD-ROM. Milano: Garzanti.
- Suleman M, Clark MPA, Goldacre M, Burton M (2010). Exploring the variation in paediatric tonsillectomy rates between English regions: a 5-year NHS and independent sector data analysis. *Clin Otolaryngol* 35 (2): 111-17. [Link alla risorsa]
- Sutherland JM, Fisher ES, Skinner JS (2009). Getting past denial: the high cost of health care in the United States. *N Engl J Med* 361 (13): 1227-30. [Link alla risorsa]
- Taroni F, Rossi F, Braga M et al. (1997). Funzione di controllo nel finanziamento dell'attività ospedaliera. *Ricerca sui Servizi Sanitari* 1: 74-7.
- Taroni F (1996). DRG/ROD e nuovo sistema di finanziamento degli ospedali. Roma: Il Pensiero Scientifico Editore.
- Testi A, Ivaldi E (2004). Una proposta di indicatore di deprivazione. Relazione al Convegno nazionale AIES "I livelli di governo nella sanità. Finanziamento, assetti istituzionali e management". 4-5 novembre, Milano: Università Bocconi.
- The King's Fund (2011). Variations in health care. The good, the bad and the inexplicable. [Link alla risorsa]
- Tonelli S (2008). La qualità nelle organizzazioni sanitarie. Appunti per formazione e consulenza [Documento non pubblicato].
- Vasselli S, Filippetti G, Spizzichino L (2005). Misurare la performance del sistema sanitario. Proposta di una metodologia. Roma: Il Pensiero Scientifico Editore.
- Weinstein JN, Bronner KK, Morgan TS, Wennberg JE (2004). Trends and geographic variations in major surgery for degenerative diseases of the hip, knee and spine. *Health Aff (Millwood)*. Suppl Variation: VAR81-9. [Link alla risorsa]
- Welch HG, Schwartz LM, Woloshin S (2013). Sovradiagnosi. Come gli sforzi per migliorare la salute possono renderci malati. Edizione italiana a cura di Amato L, Davoli M. Roma: Il Pensiero Scientifico Editore. [Link alla risorsa]
- Wennberg J (2011). Time to tackle unwarranted variations in practice. *BMJ* 342: 687-90. [Link alla risorsa]
- Wennberg J, Cooper M (1998). The Dartmouth Atlas of Health Care. 2nd edition. Chicago: American Hospital Association. [Link alla risorsa]

- Wennberg J, Gittelsohn A (1973). Small area variations in health care delivery. *Science* (14 December) 182: 1102-08. [Link alla risorsa]
- Wennberg JE, Gittelsohn A (1982). Variations in medical care among small areas. *Scientific American* 246 (4): 120-34. [Link alla risorsa]
- WHO Europe (2009). How can health systems respond to population ageing? Policy Brief 19. [Link alla risorsa]
- Wolfson D, Santa J, Slass L (2014). Engaging physicians and consumers in conversations about treatment overuse and waste: a short history of the Choosing Wisely campaign. *Acad Med* 89 (7): 990-5. [Link alla risorsa]
- Woodward RS, Warren-Boulton F (1984). Considering the effects of financial incentives and professional ethics on 'appropriate' medical care. *J Health Econ* 3: 223-37. [Link alla risorsa]
- Zanetti M, Montaguti U, Ricciarelli G et al. (1996). Il medico e il management. Guida ragionata alle funzioni gestionali. Genova: Accademia Nazionale di Medicina.
- Zingarelli N (2012). Lo Zingarelli. Vocabolario della lingua italiana. Bologna: Zanichelli (versione on line).
- Zuppiroli A (2012). Da "Spending Review" a "Choosing Wisely". *SaluteInternazionale.info*, 9 luglio. [Link alla risorsa]

SITOGRAFIA*

- Agenzia nazionale per i servizi sanitari regionali
www.agenas.it
- Agency for Healthcare Research and Quality
www.ahrq.gov
- Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna
<http://assr.regione.emilia-romagna.it/it>
- Canadian Institute for Health Information
<http://www.cihi.ca/CIHI-ext-portal/internet/EN/Home/home/cihi000001>
- Choosing Wisely. An initiative of the American Board of Internal Medicine (ABIM) Foundation
<http://www.choosingwisely.org>

* La data di ultimo accesso ai siti qui elencati (così come ai link riportati nel testo) è il 23 ottobre 2014.

- The Dartmouth Atlas of Health Care
<http://www.dartmouthatlas.org>
- Department of Health (UK)
<https://www.gov.uk/government/organisations/departement-of-health>
- EuroQol Group
<http://www.euroqol.org/home.html>
- Istituto Superiore di Sanità
<http://www.iss.it>
- The Joint Commission - Performance Measurement
http://www.jointcommission.org/performance_measurement.aspx
- NHS Atlas (UK)
<http://www.rightcare.nhs.uk/index.php/nhs-atlas>
- OECD Health Care Quality Indicators
<http://www.oecd.org/els/health-systems/health-care-quality-indicators.htm>
- Patient Reported Outcomes Measurement Group – University of Oxford
<http://phi.uhce.ox.ac.uk/home.php>
- RAND Corporation – Health and Health Care
<http://www.rand.org/topics/health-and-health-care.html>
- Regione Lazio - Sanità
http://www.regione.lazio.it/rl_sanita
- Regione Lombardia – Salute
<http://www.sanita.regione.lombardia.it/>
- Regione Toscana – Servizio Sanitario Toscana
<http://www.regione.toscana.it/sst>
- Slow Medicine. La medicina sobria, rispettosa, giusta
<http://www.slowmedicine.it>

ALTRI TITOLI DELLA COLLANA “CONTRIBUTI PER IL GOVERNO CLINICO”

Volume 1

Le organizzazioni sanitarie

A cura di Roberto Grilli

Volume 2

Gestione del rischio e del contenzioso

A cura di Martina Brini e Francesco Taroni

Volume 4

Valutazione delle performance cliniche

Roberto Grilli

Volume 5

Il governo dell'innovazione nei sistemi sanitari

Luciana Ballini

Volume 6

La formazione per il governo clinico

Corrado Ruozi, Francesca Terri, Maria Barbara Lelli

Volume 7

Audit clinico: uno strumento per favorire il cambiamento

Dante Baronciani, Enrica Perrone, Nicola Magrini

