



manuale metodologico

Come produrre, diffondere
e aggiornare linee guida
per la salute pubblica

Redazione

Lisa Triscioglio, Zadig - Milano

Impaginazione

Giovanna Smiriglia

Il presente documento è stato realizzato con il contributo del Centro nazionale per la prevenzione e il controllo delle malattie (Ccm) del Ministero della Salute, nell'ambito del progetto «Costituzione di un network per la ricerca dell'efficacia dei programmi di prevenzione rivolti alla popolazione e la diffusione della cultura scientifica nelle strutture di sanità pubblica».

Finito di stampare nel mese di settembre 2011 presso Arti Grafiche Vertemati, Vimercate

manuale metodologico

Come produrre, diffondere
e aggiornare linee guida
per la salute pubblica

Data di pubblicazione: Settembre 2011

Gruppo di lavoro

Questo manuale è stato prodotto nell'ambito del progetto del Centro nazionale per la prevenzione e il controllo delle malattie (Ccm) del Ministero della Salute «Costituzione di un network per la ricerca dell'efficacia dei programmi di prevenzione rivolti alla popolazione e la diffusione della cultura scientifica nelle strutture di sanità pubblica» (iscritto al capitolo 4393 - anno 2007), affidato all'Agenzia Regionale di Sanità (ARS) della Toscana, Osservatorio di Epidemiologia (Responsabile scientifico: Dr.ssa E. Buiatti; dal gennaio 2009: Dr. F. Cipriani), con la collaborazione della Cattedra di Igiene dell'Università degli Studi del Piemonte Orientale «Amedeo Avogadro» con sede in Novara (Responsabile scientifico: Prof. F. Faggiano) e dell'Istituto di Igiene dell'Università Cattolica del Sacro Cuore con sede in Roma (Responsabile scientifico: Prof. G. Ricciardi).

Il manuale è stato elaborato anche grazie a un accordo di collaborazione con il Sistema Nazionale Linee Guida - Istituto Superiore di Sanità (Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute - Reparto Epidemiologia Clinica e Linee Guida - Direttore: Dr. A. Mele).

HANNO COLLABORATO

Fabrizio Faggiano

Università degli Studi del Piemonte Orientale «Amedeo Avogadro», Novara e Osservatorio Epidemiologico delle Dipendenze (OED) della Regione Piemonte

Alberto Baldasseroni

Centro Regionale Infortuni e Malattie Professionali CeRIMP, Regione Toscana, Firenze

Antonio Giulio de Belvis

Università Cattolica del Sacro Cuore, Roma

Sara Franchi

Agenzia Regionale Sanitaria Toscana, Firenze

Elena Gelormino

Servizio di Epidemiologia - ASL T03, Grugliasco (TO)

Patrizia Brigoni

Centro Regionale di Documentazione per la Salute, Grugliasco (TO)

Salvatore Demasi

Istituto Superiore di Sanità, Roma

Alfonso Mele

Istituto Superiore di Sanità, Roma

Si ringrazia la Dr.ssa **Marta Marino** per l'attenta rilettura del testo finale

Presentazione

Questo manuale, elaborato dal Network per la ricerca dell'efficacia dei programmi di prevenzione, si inserisce nell'ambito delle attività del Sistema Nazionale Linee Guida (SNLG). Il lavoro dell'SNLG ha avuto inizio da oltre 10 anni a seguito di una precisa volontà politica che nel 1999 volle avviare, presso l'Istituto Superiore di Sanità, un gruppo di lavoro che si occupasse di linee guida. Successivamente, le attività hanno avuto alterne vicende per la discontinuità dovuta all'intermittenza e alla scarsità dei finanziamenti. Nonostante le difficoltà, sono stati realizzati diversi progetti grazie alla collaborazione entusiastica di molti professionisti, di società scientifiche e di strutture del Servizio Sanitario Nazionale. Tutto questo ha creato una rete di competenze che, a ragione, può essere definita «sistema». È importante che questi elementi vengano colti dalle autorità politiche al fine di dare un assetto stabile e indipendente a un centro nazionale di coordinamento. L'indipendenza, accanto alla stabilità, è un elemento cruciale per un'attività fortemente esposta ai conflitti di interesse.

Ritornando a questo manuale, è necessario precisare che esso affianca quello redatto nel 2002 relativo alla stesura di linee guida cliniche, e affronta tutti i punti salienti, dando a ciascuno il giusto approfondimento. Va però segnalato che le linee guida in sanità pubblica che trattano di interventi su gruppi, oltre che su singoli individui, sono un argomento meno rodato rispetto alle linee guida cliniche. Sappiamo tutti che le linee guida nascono come strumento di governo clinico e ben si prestano a trattare quesiti strettamente terapeutici, che a loro volta si avvalgono di studi clinici randomizzati e controllati. Tutto ciò non è facilmente reperibile nei quesiti di prevenzione.

La sfida è quindi quella di procedere nell'approfondimento della metodologia di stesura delle linee guida in sanità pubblica. In tal senso, un campo aperto e meritevole di ulteriori approfondimenti potrebbe essere l'allestimento di check-list di valutazione degli studi più utilizzati, come quelli ecologici e before-after, e della generabilità degli studi disponibili.

Alfonso Mele
Istituto Superiore di Sanità

Indice

1. Introduzione	7
1.1 A cosa servono le raccomandazioni per la tutela della salute pubblica	7
1.2 Perché un manuale	8
1.3 Come leggere questo manuale	9
2. L'integrazione delle prove scientifiche e del parere degli esperti	11
3. Gli obiettivi, i destinatari e il trasferimento di conoscenza	12
3.1 I criteri per la scelta degli argomenti e le chiavi di lettura	13
3.1.1 <i>Il modello dei bisogni o epidemiologico</i>	13
3.1.2 <i>Il modello delle opportunità o politico</i>	14
3.1.3 <i>Il modello delle conoscenze o basato sulle migliori prove di efficacia</i>	15
3.1.4 <i>Il modello delle risorse o economico</i>	15
3.1.5 <i>Il modello basato sulla percezione del rischio</i>	15
4. Metodi	16
4.1 Requisiti essenziali	16
4.2 Approccio alle conoscenze pubblicate	18
4.3 Costruzione del <i>logical framework</i>	20
4.4 Elaborazione dei quesiti	20
4.5 La ricerca delle prove di efficacia	23
4.5.1 <i>La ricerca di linee guida</i>	23
4.5.2 <i>Adattamento di linee guida internazionali</i>	24
4.5.3 <i>La ricerca di revisioni sistematiche</i>	25
4.5.4 <i>La ricerca di studi primari</i>	26
4.6 La funzione di documentazione	27
4.7 La valutazione critica delle prove di efficacia	28
4.7.1 <i>Valutazione delle linee guida</i>	30
4.7.2 <i>Valutazione delle revisioni sistematiche</i>	31
4.7.3 <i>Valutazione degli studi primari</i>	34
4.8 Caratteristiche dell'intervento e gerarchia delle prove di efficacia	34
4.9 Qualità degli studi	37
4.9.1 <i>Qualità degli studi randomizzati</i>	37
4.9.2 <i>Qualità degli studi non randomizzati</i>	37

4.10	Definizione dei livelli di evidenza	38
4.11	Il parere degli esperti	39
4.12	La formulazione delle raccomandazioni e il grado di forza	41
5.	Gli indicatori per la valutazione	44
5.1	La valutazione d'impatto	44
6.	Il conflitto d'interesse nella tutela della salute pubblica	46
7.	La disseminazione	47
7.1	Materiali	47
7.2	Formazione	48
8.	L'implementazione	49
9.	L'aggiornamento	50
10.	Allegati	51
	Bibliografia	58

1. Introduzione

1.1 A cosa servono le raccomandazioni per la tutela della salute pubblica

Il movimento culturale che mira a fondare le scienze della salute sulle prove scientifiche (la cosiddetta Evidence Based Medicine, EBM) data ormai circa vent'anni dal suo esordio nelle comunità scientifiche di tutto il mondo. Due erano i concetti chiave del nuovo atteggiamento culturale in medicina: l'evidenza scientifica e la valutazione critica che, insieme, costituivano la base per una pratica medica animata dal cosiddetto "illuminato scetticismo" (EBM Working Group 1992). La strategia proposta prevedeva la ricerca delle prove dell'efficacia degli atti medici e una critica sistematica delle stesse e delle consolidate abitudini di categoria.

Ci si muoveva, dunque, affinché la ricerca scientifica servisse la pratica clinica quotidiana di ogni medico.

Come ben noto, il processo ha avuto in questi anni, nonostante le prime resistenze, un imponente sviluppo e, poco dopo il suo riconoscimento e la sua validazione metodologica, cominciò a essere chiaro che lo stesso approccio poteva e doveva applicarsi anche ad altri campi del sapere medico, differenti dalla medicina clinica. Uno fra questi, principale per noi, è la sanità pubblica.

Questo campo è di per sé particolarmente complesso: gli argomenti cui rivolge la propria attenzione sono infatti oggetto di studio e intervento di comunità scientifiche differenti (sociologica, psicologica, medica, antropologica e molte altre) e gli interventi sono operati da attori molto diversi in quanto a competenze, responsabilità e interessi.

Le differenze che distinguono la sanità pubblica dalla clinica sono state sintetizzate da Fineberg nel 1990 e sono presentate nella tabella 1 (Fineberg 1990).

	Sanità pubblica	Clinica
Principale ambito	Popolazioni	Individuo
Enfasi	Prevenzione	Diagnosi
	Promozione della salute	Trattamento
	Approccio olistico alla comunità	Approccio olistico al paziente
Paradigma	Interventi rivolti all'ambiente, agli stili di vita, ai sistemi sanitari, ecc.	Cura medica
Criteri di specializzazione	Analitico (epidemiologia)	Organico (cardiologia, ecc.)
	Per tipologia di popolazione e setting (es. salute occupazionale)	Gruppi di pazienti (es. pediatria)
	Per funzioni di valutazione, di sviluppo di politiche, ecc.	Per funzioni tecniche (es. radiologia)

Tabella 1. Differenze fra approccio di sanità pubblica e approccio clinico ai problemi di salute (Fineberg 1990)

La ricerca e la pratica di sanità pubblica si rivolgono dunque alle comunità nel loro complesso; l'approccio basato sulle prove di efficacia richiede, in questo contesto, una comprensione della complessità delle strutture sociali e organizzative, delle interazioni e delle dinamiche che si sviluppano ogni giorno, influenzando tutti i livelli decisionali e complicandone la comprensione. La salute pubblica basata sulle prove di efficacia (o Evidence Based Prevention, EBP, come è chiamato tale movimento in Italia) consiste, dunque, nell'uso informato, esplicito e giudizioso dell'evidenza derivata dai vari campi della ricerca scientifica, da quello biomedico a quello sociale, e dalle valutazioni empiriche (Rychetnik 2004).

Proprio le radicali differenze dalla medicina clinica e la complessità che operatori e ricercatori incontrano quando si confrontano con la realtà hanno spinto il Network per la ricerca dell'efficacia dei programmi di prevenzione ad affrontare l'elaborazione di questo manuale, nel quale le difficoltà metodologiche della ricerca si aggiungono a quelle che si incontrano nel trasferimento delle conoscenze al decisore politico e amministrativo.

Nella consapevolezza che l'eshaustività non possa essere raggiunta in questa sede, gli autori si sono adoperati per includere in questo lavoro gli elementi di multidisciplinarietà indispensabili almeno al fine di stimolare una nuova discussione e un approfondimento da parte dei numerosi gruppi che in Italia si occupano di tutela della salute pubblica.

1.2 Perché un manuale

Il 7 luglio 2005 il Centro nazionale per la prevenzione e il controllo delle malattie (Ccm) ha affidato all'ARS Toscana un progetto per la diffusione della cultura dell'EBP nelle strutture di prevenzione delle regioni e delle ASL italiane. A questa iniziativa è seguita la costituzione di un apposito sottocomitato scientifico, d'ora in poi denominato comitato EBP, con decreto del Ministro della Salute del 26 novembre 2006 (Ccm 2008).

Il comitato EBP ha proposto di modificare il progetto, mirando a costituire una vera e propria rete composta da centri di ricerca e servizi di prevenzione in grado di sviluppare valutazioni di programmi o interventi attualmente in uso, ma la cui efficacia sia ritenuta dubbia, e di programmi o interventi utili per contrastare problemi di salute e verosimilmente dotati di prove di efficacia, ma la cui diffusione è ancora insufficiente.

I problemi di salute per cui è necessario identificare prioritariamente i programmi di intervento più efficaci sono quelli identificati da atti istituzionali come l'Intesa Stato-Regioni del 23 marzo 2005 o il DPCM del 4 maggio 2007, che riporta il programma «Guadagnare Salute» (Ccm 2010). In queste aree esistono interventi che, pur essendo verosimilmente efficaci, non sono correntemente implementati.

La rete, denominata Network per la ricerca dell'efficacia dei programmi di prevenzione, viene costituita con l'apporto di tre centri di ricerca individuati nell'ARS Toscana, nella Cattedra di Igiene dell'Università del Piemonte Orientale con sede in Novara e nell'Istituto di Igiene dell'Università Cattolica di Roma.

Al Network viene affidato il compito di sviluppare tre linee guida di prevenzione, oltre che un manuale metodologico, oggetto di questa pubblicazione.

Successivamente, a partire dal 2009, il Network ha avviato una collaborazione con l'Istituto Superiore di Sanità e il Sistema Nazionale Linee Guida (SNLG) al fine di valorizzare al meglio l'esperienza e le metodologie sviluppate nel campo dell'elaborazione di linee guida.

Il presente manuale nasce da questa collaborazione, nell'alveo delle azioni promosse dall'SNLG. Il documento è stato impostato secondo il modello già adottato nel 2002 dal Piano Nazionale Linee Guida (PNLG) per l'analogo manuale dedicato all'elaborazione di raccomandazioni per la pratica medica (PNLG 2002), consapevoli che la Evidence Based Prevention deve ispirazione all'esperienza di sviluppo della EBM nella clinica, pur con la necessità di adattarne i principi cardine alla multiforme realtà della salute pubblica.

In linea con le premesse, questo manuale metodologico ha l'obiettivo di proporre un metodo uniforme per lo sviluppo di linee guida evidence-based in campo di sanità pubblica. A causa delle complessità metodologica esso si rivolge quindi soprattutto alla produzione di linee guida di salute pubblica di rilevanza nazionale. Le istituzioni locali che avvertono l'importanza di linee guida adattate al proprio contesto avranno quindi il compito di adeguare le raccomandazioni alla propria realtà sociale e organizzativa.

1.3 Come leggere questo manuale

Questo manuale è un documento operativo che vuole proporre i metodi per sviluppare linee guida di prevenzione. Vengono quindi affrontati brevemente l'importanza dell'interazione fra conoscenza scientifica e parere di chi lavora sul campo, e quanti e quali siano gli ostacoli in questo incontro (vedi capitolo 2 *L'integrazione delle prove scientifiche e del parere degli esperti* a pag. 11). Gli strumenti sono ben noti e furono già descritti nel Manuale PNLG (PNLG 2002). Qui ci limitiamo ad approfondire le peculiari applicazioni alla sanità pubblica, quando queste siano possibili, e soprattutto a descrivere l'integrazione dei diversi metodi: in questo campo è molto più utile, infatti, sfruttare al meglio le potenzialità di metodi diversi preservandone il valore scientifico, piuttosto che costringere a scelte rigide fra mezzi diversi.

Nel capitolo 3 *Gli obiettivi, i destinatari e il trasferimento di conoscenza* a pag. 12 affrontiamo invece il tema dei destinatari delle raccomandazioni, argomento parti-

colarmente impegnativo che riguarda il trasferimento delle conoscenze a un'arena decisionale animata da interessi talvolta contrastanti e competenze complesse. Questo argomento si collega direttamente a quello affrontato nel capitolo 4 *Metodi* a pag. 16 che, pur riprendendo alcuni temi del Manuale PNLG del 2002, li investiga nell'ottica del difficile percorso che le conoscenze di tutela della salute pubblica devono percorrere per informare decisioni corrette. Questo capitolo è certamente il più ricco e quello che, partendo dalle esigenze della pratica di sanità pubblica, si inoltra nei temi raffinati della metodologia della ricerca epidemiologica e sociale.

Una breve riflessione sui temi etici e sui possibili interessi che, contrastando con il benessere della collettività, possono influenzare i contenuti di un testo di raccomandazioni per la pratica di sanità pubblica, è presentata al capitolo 6 *Il conflitto d'interesse nella tutela della salute pubblica* a pag. 46.

Infine, si accenna al progetto che ha animato all'origine questo documento, e che è stato promosso per l'elaborazione delle prime linee guida nazionali italiane di salute pubblica.

2. L'integrazione delle prove scientifiche e del parere degli esperti

La maggiore critica che l'approccio basato sulle prove di efficacia proposto da Sackett e collaboratori (EBM Working Group 1992) ricevette a suo tempo riguardava nel campo clinico due aspetti principali: il timore che il medico perdesse la propria autonomia in relazione alle caratteristiche personali di ciascun paziente, e l'evidente povertà di risposte che la ricerca clinica riusciva a dare a fronte del grande numero di domande cui un medico clinico deve rispondere nel percorso di cura di ciascun paziente. L'approccio veniva perciò vissuto come riduttivo della professionalità del medico e inadeguato alla realtà soggettiva del paziente.

Questo avveniva in un contesto nel quale la ricerca ha raggiunto, tutto sommato, buoni livelli sperimentali, con una validità interna degli studi spesso molto elevata. Eppure, nonostante la solidità delle conoscenze disponibili, esse risultavano spesso poco adeguate alla realtà clinica.

A maggior ragione, dunque, possiamo prevedere che l'applicazione dei principi di una sanità pubblica basata sulle prove di efficacia trovi alcuni ostacoli sul suo cammino e susciti perplessità: le prove di efficacia sono molto più carenti, sia quantitativamente sia qualitativamente, e i contesti di applicazione assai difformi.

Per queste ragioni il presente manuale insiste molto sull'integrazione fra prova scientifica, parere d'esperto e interesse dell'attore coinvolto. Il capitolo 4 dedicato ai metodi (pag. 16) si concentra sia sull'approccio alle prove scientifiche secondo la tradizionale linea di condotta dell'EBM (costruzione del quesito, ricerca e sintesi delle prove di efficacia, loro valutazione critica), sia sui metodi per la raccolta del consenso di esperti. Lo sforzo maggiore e più innovativo consiste nella fusione di questi due elementi in una raccomandazione che, forte del contributo di entrambe le fonti, sia la più adeguata possibile alla realtà e la più affidabile possibile. Per questa ragione il paragrafo 4.12 *La formulazione delle raccomandazioni e il grado di forza* a pag. 41, dedicato alla formulazione delle raccomandazioni e all'attribuzione di un grado, è logicamente successivo ai metodi utili per ottenere sia le prove di efficacia, sia il consenso degli esperti: in sanità pubblica, cioè, la prova scientifica e l'esperienza sul campo sono necessariamente inscindibili.

3. Gli obiettivi, i destinatari e il trasferimento di conoscenza

Gli obiettivi principali di una linea guida di sanità pubblica sono il miglioramento del benessere collettivo (riduzione dell'incidenza di malattie e inabilità prevenibili, miglioramento degli esiti di malattia nella popolazione, miglioramento della distribuzione delle opportunità di salute fra gruppi sociali differenti) e il miglior utilizzo delle risorse.

La consapevolezza circa la destinazione delle raccomandazioni è indispensabile per una formulazione comprensibile e per la definizione di obiettivi che incontrino i processi decisionali di coloro che dovranno utilizzare sul campo le conoscenze trasmesse dalle raccomandazioni.

Questo vale sia per chi lavora al trasferimento di conoscenze nel campo clinico, sia per quanti trattano argomenti di tutela della salute collettiva.

In quest'ultimo caso la situazione è complicata dalla molteplicità di livelli e aree decisionali che hanno competenze diverse e parlano linguaggi spesso radicalmente differenti, perseguendo interessi e fini magari contrapposti.

Il trasferimento di conoscenze è oggetto da alcuni anni di analisi da parte delle comunità scientifiche che si occupano di sanità pubblica e di politica sanitaria.

Lavis ha ben descritto gli elementi fondamentali che è necessario identificare con chiarezza per un trasferimento di conoscenze che abbia qualche ambizione di cambiare la realtà (Lavis 2003). L'approccio di questo gruppo di ricerca prevede uno schema concettuale per la strategia di trasferimento che, ovviamente, implica una scrittura dei contenuti mirata alla situazione. Lavis identifica cinque questioni fondamentali:

- 1) qual è il messaggio da trasmettere;
- 2) a chi deve essere trasferita la conoscenza (i destinatari o il target);
- 3) chi deve trasmettere la conoscenza (il messaggero);
- 4) in che modo deve avvenire il trasferimento (il processo e le infrastrutture comunicative);
- 5) quale effetto viene sortito dal trasferimento (la valutazione).

Gli stessi autori ordinano i destinatari delle raccomandazioni: i cittadini/pazienti, i fornitori dei servizi sanitari, i decisori amministrativi, i decisori politici. L'elemento «destinatari», o *audience*, condiziona tutti gli altri.

D'altro canto, il cosiddetto paradigma razionale delle politiche ha rivelato la sua debolezza nel corso del secondo dopoguerra negli Stati Uniti (Regonini 2001); la comunità epidemiologica deve dunque prestare attenzione a non cadere in errori, già commessi in passato, dovuti all'eccessiva semplificazione della realtà.

Le raccomandazioni di salute pubblica sono rivolte ad attori diversi: dai professionisti della salute, che attuano i programmi sul campo, ai dirigenti sanitari, che pianificano gli interventi sulle popolazioni di competenza, agli amministratori, che scelgono gli obiettivi e orientano i finanziamenti, ai detentori di carica elettiva, che vengono chiamati a pronunciarsi su strategie di salute e tutela dei cittadini. A questi vanno aggiunti le rappresentanze della società civile, che possono farsi carico della promozione di certi impegni e interventi o possono esercitare pressioni sulle sedi di potere politico e amministrativo per la salvaguardia del benessere collettivo, e i detentori di interessi economici e finanziari direttamente e/o indirettamente legati alle politiche di salute.

I destinatari, dunque, sono numerosi e quasi mai neutrali rispetto né ai temi proposti, né a interessi particolari spesso conflittuali. Un presupposto fondamentale per l'implementazione della linea guida è che essa sia formulata avendo chiarito chi sia responsabile dell'applicazione delle singole raccomandazioni. In effetti, l'esplicitazione dei destinatari aiuta gli autori delle raccomandazioni ad adattare i testi al linguaggio da loro utilizzato e ad acquisire consapevolezza circa i reali ruoli degli attori coinvolti.

3.1 I criteri per la scelta degli argomenti e le chiavi di lettura

Come già evidenziato nel Manuale PNLG del 2002, un programma che si ponga l'obiettivo di migliorare l'efficacia degli interventi attraverso la produzione e l'implementazione di linee guida deve necessariamente operare scelte di priorità. Per questa ragione lo stesso documento passa in rassegna i criteri principali con cui scegliere gli argomenti da trattare con una linea guida, e a esso si rimanda. In questa sede le categorie concettuali per la definizione delle priorità vengono approfondite con l'occhio della sanità pubblica.

Ovviamente, la scelta degli argomenti da trattare dovrà tenere in considerazione anche le future possibilità di implementazione.

3.1.1 Il modello dei bisogni o epidemiologico

Questo punto di vista è promosso soprattutto nella Evidence Based Medicine. Si parte, cioè, dai bisogni dei pazienti.

Nel caso della salute pubblica, il criterio si rivolge al bisogno di salute delle popolazioni che, per come è tradizionalmente descritto dall'epidemiologia delle popolazioni, viene compreso con difficoltà dai decisori, soprattutto quando si tratta in concreto di operare delle scelte fra riduzioni di rischi attribuibili molto diversi, che comportano esiti di salute altrettanto diversi e costi difficilmente confrontabili. Bisogna anche considerare che spesso il decisore segue logiche assai differenti da quelle dell'epidemiologo o del medico di sanità pubblica. Il processo decisionale, infatti, deve

includere elementi spesso sconosciuti allo scienziato, ma egualmente legittimi in un sistema democratico.

È interessante, però, notare come, da un punto di vista analitico quantitativo, la costruzione di nuovi strumenti di misura (le cosiddette *summary measure* dell'Organizzazione Mondiale della Sanità) (van der Maas 2003) possa avvicinare il decisore a questo modello di definizione delle priorità, esplicitando in modo univoco il peso che la malattia, o eventi anche molto diversi, hanno sulla salute della popolazione, e permettendo valutazioni economiche di costo-beneficio comparative. Il Disability Adjusted Life Years (DALY), per esempio, rappresenta una misura univoca del carico imposto dalle diverse malattie alla popolazione in termini di anni di vita perduti prematuramente e standardizzati per carico di disabilità. Il DALY ha due grandi vantaggi: è monetizzabile, permettendo in tal modo confronti di costo-beneficio fra interventi diversi, e riporta il danno subito dai cittadini a un'unica unità di misura. In questo modo il decisore politico e quello amministrativo sono facilitati nella scelta, potendo includere fra i parametri decisionali anche quello epidemiologico (Lopez 2006).

Nel momento in cui tali misure saranno disponibili sistematicamente per la maggior parte dei fenomeni di salute, è probabile che il criterio epidemiologico potrà diventare più utile e maneggevole per la scelta della prioritarizzazione degli argomenti.

Nell'esperienza del progetto Evidence Based Prevention le priorità emergono dal tavolo tecnico interregionale per la prevenzione. Questo tavolo, organo della Conferenza Stato-Regioni, è la sede in cui le politiche regionali sulla prevenzione si confrontano, e dove vengono definite le priorità per affrontare le quali tutte le regioni si impegnano a investire risorse economiche e professionali.

3.1.2 Il modello delle opportunità o politico

La letteratura dedicata all'analisi delle politiche pubbliche ha da tempo descritto le cosiddette *policy window*. In un processo decisionale spesso di difficile comprensione, che include interessi molteplici e, secondo certi autori, una buona dose di casualità, il cosiddetto imprenditore politico è colui che riesce a cogliere le "finestre" di opportunità che si aprono, anche inaspettatamente, al suo ingresso in un processo altrimenti chiuso, affinché l'agenda politica includa anche gli argomenti che stanno a cuore alla sanità pubblica (*agenda setting*) (Kingdon 1984). Questo significa, in pratica, che coloro che elaborano linee guida e si prefiggono di fornire raccomandazioni al decisore dovrebbero essere in grado di comprendere quali siano le priorità del decisore stesso, quelle cui potrebbe essere particolarmente interessato e attento, mediando questo tipo di interesse con la tutela del benessere collettivo.

L'economicità delle scelte compete al decisore sulla base degli interessi in gioco e delle informazioni a propria disposizione: in questa sede includiamo concettualmente questo elemento nel criterio dell'opportunità, mentre consideriamo le spese per la produzione della linea guida e il trasferimento di conoscenze come criterio a sé stante.

3.1.3 Il modello delle conoscenze o basato sulle migliori prove di efficacia

È forse il modello più consolidato e agevole per la comunità scientifica medica. Esso si basa sul fatto che siano elaborate raccomandazioni per quegli interventi, programmi e politiche per i quali sia stata condotta una valutazione d'efficacia di buona qualità metodologica che abbia condotto a risultati positivi per la salute collettiva.

Questo modello, pur nel fascino del suo rigore scientifico, lascia spazio a molti campi privi di studi di buona qualità, anche nella medicina clinica e nell'assistenza, e a maggior ragione in sanità pubblica, dove, come vedremo nella sezione dedicata ai metodi, la conduzione di studi aderenti al modello sperimentale è molto più complessa. Questo significa che, adottando questo criterio per la definizione delle priorità, si offrirebbero raccomandazioni in ambiti di intervento estremamente ristretti. Pur adottando la metodologia consolidata dell'Evidence Based Medicine, la sanità pubblica deve necessariamente adattare le proprie esigenze metodologiche al campo di intervento e porre attenzione non solo ai livelli di compromesso accettabili in termini di validità interna, ma anche alla validità esterna degli studi e al contesto locale nel quale gli interventi verrebbero messi in atto.

3.1.4 Il modello delle risorse o economico

I costi di produzione di una linea guida sono in relazione all'impegno del personale dedicato e alla tipologia di fonti informative che contribuiscono alla formulazione delle raccomandazioni. Grazie al rapido e poco costoso accesso alle risorse bibliotecarie, le linee guida che fondano le proprie affermazioni solo sulle evidenze in senso stretto (cioè pubblicate nella letteratura scientifica internazionale e in quella grigia, quali per esempio tesi di dottorato, atti di congressi, rapporti di enti internazionali) sono relativamente poco costose.

I lavori si complicano e le spese aumentano nel caso di un'integrazione fra conoscenza scientificamente validata ed esperienza, soprattutto quando gli esperti di settore o gli attori più rappresentativi provengono e rappresentano realtà geografiche distanti e, ulteriormente, quando alla produzione si debba affiancare la valutazione di impatto, l'implementazione, l'aggiornamento, e la diffusione dello strumento.

3.1.5 Il modello basato sulla percezione del rischio

Esso riveste un peso e un interesse crescente, soprattutto per i temi di sanità pubblica che interessano gli ambienti di vita e di lavoro dei cittadini e i comportamenti di alcuni gruppi di popolazione a rischio. Si pensi al dibattito che spesso coinvolge l'emotività di intere popolazioni riguardo la nocività del trattamento dei rifiuti (AIE 2008, D'Argenio 2008) o la percezione del rischio che alcuni gruppi di popolazione (per esempio i giovani neoparentati) hanno su comportamenti e abitudini rischiose per l'incolunità propria e degli altri.

4. Metodi

4.1 Requisiti essenziali

Le raccomandazioni di comportamento pratico sono i tasselli fondamentali di una linea guida e devono essere basate sulle prove di efficacia; un'esplicita descrizione delle prove scientifiche alla base delle raccomandazioni, inclusa una valutazione della qualità delle prove stesse, deve essere resa disponibile per ogni raccomandazione. Il metodo e il processo che è stato utilizzato per la costruzione della raccomandazione deve inoltre essere accuratamente definito. Le prove di efficacia che accompagnano la raccomandazione dovrebbero esprimere in sintesi le caratteristiche dello studio citato, i suoi risultati e la possibile rilevanza clinica. Il metodo e il processo che è stato utilizzato per la costruzione della raccomandazione deve inoltre essere ben stabilito. Studi empirici hanno dimostrato che la presenza di questi elementi favorisce una migliore implementazione delle linee guida stesse da parte degli operatori (Grol 1998).

Per loro natura le raccomandazioni raccolte in una linea guida sono affermazioni operative: in quanto tali devono essere, oltre che basate sulle migliori conoscenze disponibili, anche chiare, non passibili di interpretazioni divergenti, sintetiche e applicabili nel contesto operativo di destinazione. La realtà è complessa e per questo all'interno di raccomandazioni apparentemente semplici è necessario raccogliere l'accordo di professionalità e interessi diversi, a volte contrastanti. Per questa ragione la multidisciplinarietà e la rappresentatività dei gruppi coinvolti nell'elaborazione e nella revisione dei testi sono fondamentali: le conoscenze sanitarie, economiche, politiche ed etiche devono integrarsi con il consenso delle parti in causa (rappresentanti delle istituzioni, dei beneficiari degli interventi, dei gruppi minoritari, ecc.).

In pratica, per rendere operativi i principi sopra esposti, si suggerisce agli operatori che si accingono a elaborare una linea guida su uno dei molteplici argomenti della sanità pubblica di identificare i gruppi di lavoro prima di ogni altro passaggio.

Gli autori dovranno da un lato garantire la qualità delle conoscenze messe a disposizione nella linea guida, e dall'altro essere sufficientemente rappresentativi della realtà nella quale le linee guida dovranno essere messe in pratica; il tutto garantendo una certa linearità e celerità dei lavori (si considera ragionevole la durata di un anno per i lavori di una linea guida che non preveda revisioni sistematiche *ex novo*).

Per salvaguardare il valore scientifico del documento e il suo legame costitutivo con il contesto di applicazione, precedenti esperienze (Faggiano 2007) suggeriscono di definire ambiti concentrici di competenza nei quali identificare i soggetti più rispondenti alle esigenze di quell'ambito (vedi figura 1 a pag.17).

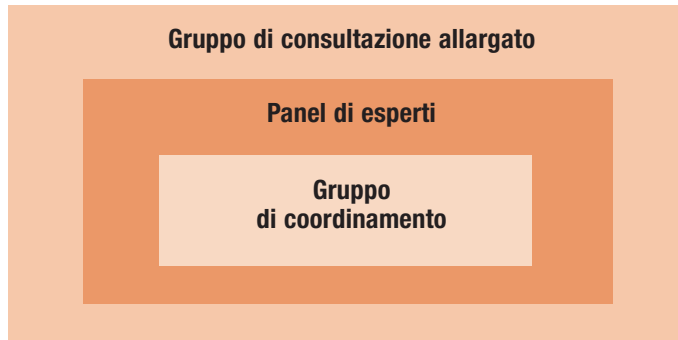


Figura 1. L'esperienza piemontese nella definizione dei gruppi di elaborazione della linea guida (Faggiano 2007)

Il **gruppo di coordinamento** ha la responsabilità della conduzione dei lavori. Deve coordinare tutto lo svolgimento dei lavori, essere guidato da un responsabile istituzionale e includere:

- uno o più esperti di metodo di ricerca, di Evidence Based Medicine, di revisione sistematica di letteratura e di trasferimento delle conoscenze;
- almeno un esperto dell'argomento trattato nella linea guida;
- almeno un esperto di organizzazione dei livelli istituzionali o amministrativi cui la linea guida si rivolge.

Al fine di rendere efficiente il lavoro del gruppo di coordinamento, alcune funzioni del gruppo potrebbero essere svolte da un soggetto con competenza da documentalista. Il gruppo di coordinamento ha quindi il compito di identificare i due livelli di elaborazione concettuale e consenso successivi.

Il **panel di esperti** include tecnici competenti degli argomenti trattati dalla linea guida, arricchisce la lista dei quesiti con problemi derivanti dalla pratica, supervisiona il recupero delle evidenze scientifiche, elabora insieme al gruppo di coordinamento i quesiti e coordina la consultazione del gruppo allargato. È pertanto necessario che in questo gruppo siano rappresentate tutte le professionalità coinvolte dalla linea guida e siano incluse le competenze economiche, politologiche ed etiche eventualmente necessarie. Per consentire una corretta dinamica del gruppo, esso non dovrebbe superare i 12 componenti.

Il **gruppo di consultazione allargato** rappresenta gli interessi e i punti di vista sia degli operatori sia della popolazione. I suoi compiti sono la revisione delle raccomandazioni, l'aggiunta di quesiti eventualmente mancati alla lista iniziale, l'approvazione delle raccomandazioni. È importante però che il gruppo allargato venga coinvolto nel lavoro di elaborazione della linea guida sin dal principio del lavoro, e che venga aggiornato sull'avanzamento dei lavori. In questo modo i componenti potranno eventualmente contribuire fin dalle prime fasi anche in azioni diverse da quelle a loro ri-

servate. I suoi componenti saranno poi i principali riferimenti del percorso di implementazione della linea guida a livello delle singole regioni.

I cittadini, attraverso loro rappresentanti o associazioni, devono essere coinvolti direttamente. Questa componente è di fondamentale importanza sia per il contributo derivante da un loro diretto coinvolgimento, sia perché raccoglie un'esigenza di maggiore partecipazione e di trasparenza che viene sempre più chiaramente rivolta ai sistemi sanitari. Rappresentanti dei cittadini devono essere quindi accuratamente identificati e invitati a collaborare nell'ambito del gruppo allargato.

Gli utenti di una linea guida di sanità pubblica si differenziano sostanzialmente da quelli di un analogo documento clinico. Si tratta infatti sia di professionisti che lavorano in diversi livelli del Servizio sanitario, dal Dipartimento di prevenzione, ai Distretti, alle Direzioni aziendali, sia di professionisti che operano al di fuori della sanità, come policy-maker nazionali, amministratori di enti locali, operatori della polizia locale. Per questo è importante garantire la rappresentatività di questi soggetti nell'ambito del gruppo allargato, coinvolgendo le società scientifiche e i professionisti degli enti coinvolti dalla linea guida. A questo fine il gruppo potrà essere costituito da 20-40 soggetti.

L'integrazione del punto di vista dei valori degli utenti, accanto a quello degli erogatori e degli amministratori, è fra gli obiettivi principali di questo gruppo. La partecipazione dei cittadini e dei destinatari finali è importante anche per il destino delle raccomandazioni, poiché una partecipazione attiva alla loro elaborazione e una conoscenza dettagliata del loro contenuto può essere importante nella fase di implementazione.

4.2 Approccio alle conoscenze pubblicate

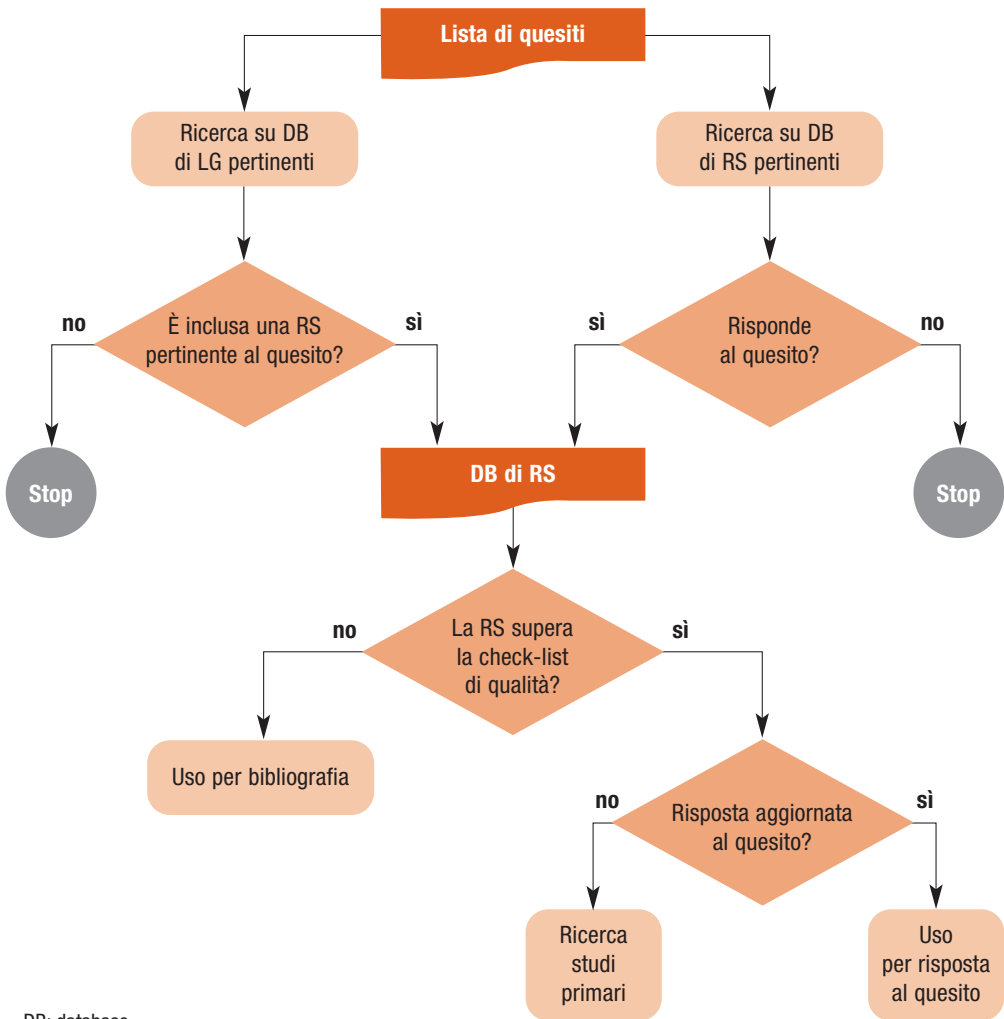
L'approccio alla conoscenza già acquisita dalla comunità scientifica differisce concettualmente da quello dell'Evidence Based Medicine.

La linea guida è un documento che deve sintetizzare la migliore evidenza disponibile per elaborare raccomandazioni di pratica professionale. La conoscenza può essere acquisita sotto forma di singoli studi primari (studi randomizzati, studi osservazionali che confrontano soggetti esposti all'intervento in studio con controlli), oppure sotto forma sintetica. È auspicabile nell'elaborazione di una linea guida utilizzare conoscenze già sintetizzate da altri (in revisioni sistematiche e linee guida) per ragioni di parsimonia e per aumentare l'affidabilità dei risultati riportati: secondo alcuni studi, in sanità pubblica può accadere che un singolo studio riporti un effetto poi contraddetto da una revisione di studi (Egger 1997, Devereaux 2002). Naturalmente i documenti di sintesi devono includere evidenze derivanti da studi validi scientificamente, cioè poveri o completamente privi di distorsioni. L'alternativa di risalire agli studi pri-

mari deve essere considerata per i quesiti che non abbiano trovato una soddisfacente risposta in sintesi elaborate da altri o per aggiornare sintesi ormai obsolete.

Il percorso logico della ricerca delle prove scientifiche segue quindi la flow-chart riportata nella figura 2.

Si ribadisce in questa sede che le fasi di produzione della linea guida non devono essere considerate, un po' scolasticamente, in sequenza; al contrario, spesso si lavorerà in parallelo, come accade nelle situazioni in cui il parere di esperti ridefinisce i quesiti iniziali e reindirizza la ricerca di letteratura.



DB: database
RS: revisione sistematica
LG: linee guida

Figura 2. Flow-chart della ricerca delle evidenze a partire dai quesiti

La ricerca delle prove scientifiche deve essere guidata da una definizione del quadro concettuale in cui si colloca il problema di sanità pubblica da affrontare (il *logical framework*) e dalla conseguente elaborazione dei quesiti di ricerca.

4.3 Costruzione del *logical framework*

Per poter formulare correttamente i quesiti è necessario identificare il modello logico (*logical framework*), ossia compiere l'analisi logica dell'oggetto di valutazione. Il *logical framework* serve a dimostrare la relazione esistente tra particolari interventi e gli outcome relativi (NICE 2009a), oltre alla valutazione dell'applicabilità del programma e all'analisi del rapporto costi-benefici.

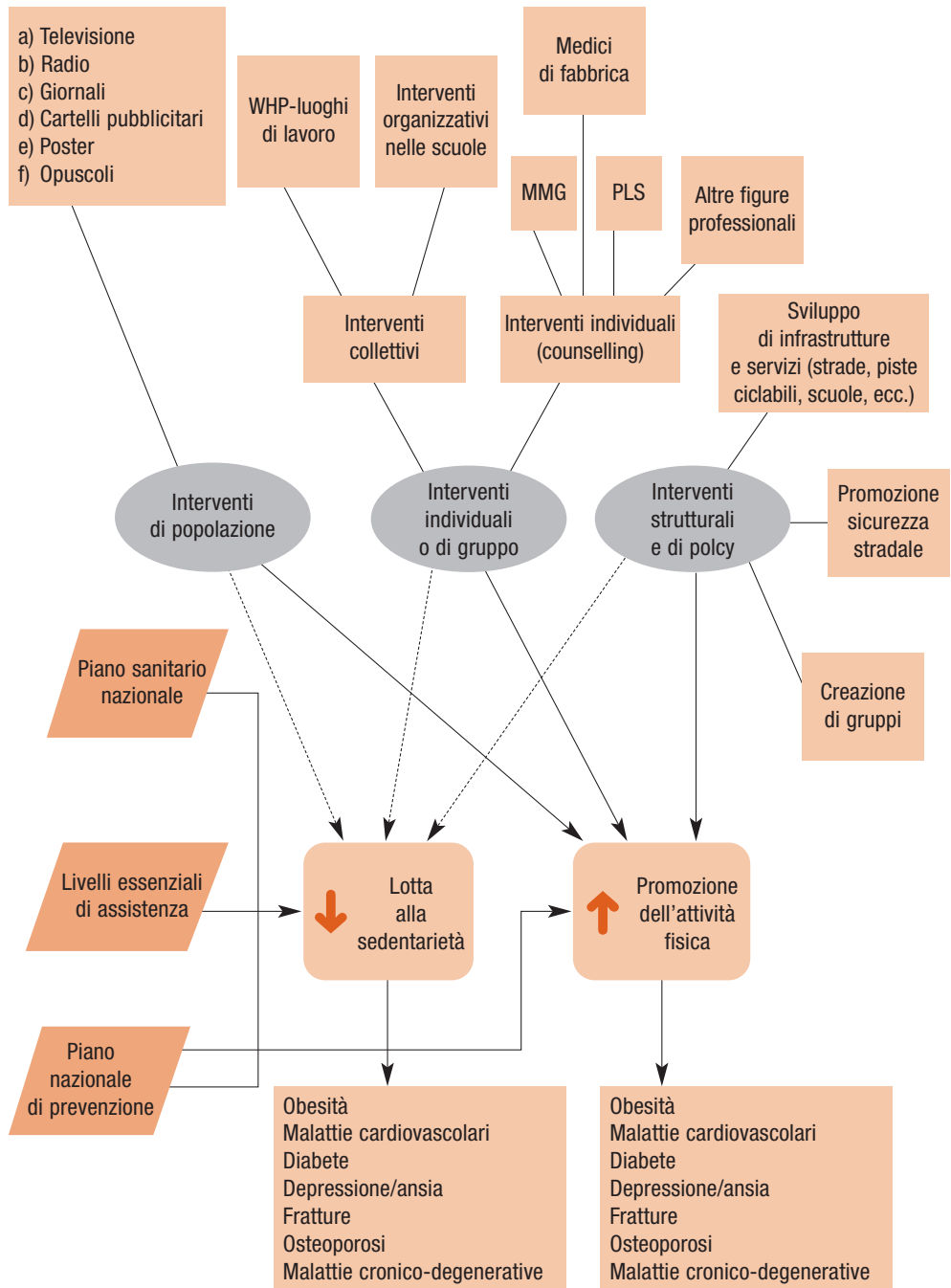
Se il programma non supera la fase di analisi logica, perché non esiste coerenza logica tra azione e outcome, o perché sono venuti meno i presupposti storici ed epidemiologici, è verosimile affermare che esso sia inefficace. Se invece le condizioni che hanno portato alla strutturazione del programma sono ancora attuali, o se esiste coerenza tra azioni e outcome, allora si prosegue verso la fase successiva, che consiste nella valutazione delle prove di efficacia. Nella figura 3 a pagina 21 è riportato un esempio di *framework* logico nel campo della promozione dell'attività fisica.

4.4 Elaborazione dei quesiti

La formulazione dei quesiti richiede chiarezza: essi devono definire in modo inequivocabile i limiti dell'argomento da ricercare. I quesiti possono essere suddivisi in:

- *background question*, quando vengono formulati sulla base delle conoscenze disponibili, ottenute anche grazie a una ricerca preliminare di letteratura (NICE, 2009b) e preliminarmente esaminate (il gruppo di coordinamento è in grado, pur non essendo competente dello specifico settore, di derivare la prima serie di quesiti dalla ricerca già disponibile);
- *foreground question*, che originano dagli esperti di settore, in misura maggiore e più sistematica se si fornisce una prima serie di quesiti (*background*) su cui ragionare.

I quesiti, infatti, sviluppati in prima battuta dal gruppo di coordinamento, dovrebbero essere affinati e integrati da tutti i membri del panel ed eventualmente sottoposti a esperti esterni. Nel caso in cui la conoscenza tra i membri del gruppo di coordinamento non sia specialistica, è preferibile eseguire dapprima una ricerca rapida della letteratura, in modo da inquadrare il problema e vagliare tutte le possibili soluzioni, e in seguito la strutturazione del *logical framework* e la formulazione dei quesiti. In questa fase sono un supporto importante le linee guida eventualmente



WHP: workplace health promotion
 MMG: medici di medicina generale
 PLS: pediatri di libera scelta

Figura 3. Elaborazione del quadro logico di riferimento del problema di salute oggetto della linea guida

già elaborate sull'argomento. Il fatto che il panel sia composto da figure professionali differenti assicura che vengano identificati i quesiti nel modo corretto. Spesso i quesiti principali necessitano di essere ridefiniti nuovamente dopo che è stata condotta la ricerca di letteratura: in questo modo si possono generare anche dei sottoquesiti.

I quesiti hanno un duplice scopo: da un lato servono come punto di partenza per la revisione sistematica di letteratura, dall'altro come guida per facilitare lo sviluppo delle raccomandazioni dal panel. Il numero di quesiti cui una linea guida dovrebbe rispondere non è codificabile; dipende comunque dall'argomento della linea guida. È importante assicurarsi che il numero totale di quesiti sia ben gestibile dal gruppo di coordinamento e dal panel, e appropriato per il budget assegnato.

I quesiti dovrebbero coprire tutte le aree che fanno parte dello scopo della linea guida, senza aggiungere altri aspetti. Inoltre, i quesiti dovrebbero occuparsi di efficacia (*effectiveness*), sicurezza, costo-efficacia, fattibilità e accettabilità.

Un approccio utile che consente di strutturare in modo corretto il quesito è il PICOT framework, approccio mutuato dalla Cochrane Collaboration e in parte modificato per le peculiarità delle linee guida di sanità pubblica. Su questa base, ogni quesito dovrebbe contenere al suo interno le seguenti componenti:

- la popolazione target (età, sesso, condizioni cliniche, ecc.);
- l'intervento di interesse;
- l'intervento utilizzato come confronto (intervento alternativo o nessun intervento);
- l'outcome di interesse;
- il tipo di studio.

Mentre normalmente il confronto e l'outcome sono poco utili ad aumentare la sensibilità del quesito, a volte può essere utile, soprattutto per gli interventi di prevenzione, aggiungere il setting dell'intervento (scuola, ambiente di lavoro, ecc.) e il/i destinatario/i del quesito stesso. Ai fini operativi la specificazione dell'outcome e del tipo di studio possono non essere necessari, in quanto la selezione delle evidenze su questa base può avvenire in un secondo tempo, soprattutto per i quesiti che presentano poche evidenze. In generale si può dire che il grado di completezza del quesito deve tener conto della quantità di articoli disponibili. Minore è la disponibilità di prove di efficacia, più elevato è il rischio che un quesito troppo dettagliato applicato alla strategia di ricerca per parole chiave causi la perdita di testi importanti.

Se volessimo ricercare gli interventi per la prevenzione dell'obesità nei soggetti di 6-13 anni, potrebbero essere strutturati diversi quesiti. Alcuni esempi possono essere:

- «È in grado la promozione del consumo di frutta e verdura, attraverso interventi come il counselling dei genitori, la formazione degli insegnanti, la formazione degli addetti alla somministrazione dei pasti, il parere sulle caratteristiche nutrizionali delle mense, le campagne sui mass-media, di diminuire il numero di bambini sovrappeso nella fascia d'età 6-13 anni?»

- «È in grado la promozione del riordino dei pasti giornalieri, attraverso il counseling sulla famiglia, la formazione degli insegnanti e le campagne sui mass-media locali, di diminuire il numero di bambini sovrappeso nella fascia d'età 6-13 anni?»»
- «È in grado la promozione dell'attività fisica attraverso la formazione dei genitori e degli insegnanti, le campagne sui mass-media locali e il miglioramento dell'offerta sportiva a livello scolastico, di diminuire il numero di bambini sovrappeso nella fascia d'età 6-13 anni?»»

Sulla base dei quesiti viene quindi eseguita la ricerca delle prove di efficacia. La formulazione accurata del quesito è fondamentale per rispondere appropriatamente assicurando un livello adeguato di sensibilità e specificità alla ricerca. Domande troppo generiche lasciano inoltre troppo spazio alle opinioni, a conoscenze già acquisite e agli interessi particolari del rispondente. Una elaborazione accurata dei singoli quesiti permette il riconoscimento delle parole chiave più utili nell'approccio ai diversi archivi oggi disponibili.

4.5 La ricerca delle prove di efficacia

Partendo dai quesiti elaborati dal gruppo, i documentalisti in un primo tempo ricercano linee guida e revisioni sistematiche già pubblicate sull'argomento (vedi figura 2 a pagina 19). L'elaborazione della strategia di ricerca dovrebbe essere eseguita almeno in doppio.

4.5.1 La ricerca di linee guida

Le **linee guida** sono una fonte rilevante di revisioni sistematiche; infatti, ogni raccomandazione dovrebbe essere sostenuta da una revisione sistematica che in molti casi viene elaborata nell'ambito della linea guida stessa (vedi per esempio AHRQ 2008). Queste revisioni possono essere utilizzate di per sé nell'elaborazione della nuova linea guida: ciò non toglie che la raccomandazione stessa di un'altra linea guida possa essere presa in considerazione per una sua inclusione in quella nuova, seppure sia consigliabile ripercorrere il processo che dalle evidenze porta alla produzione della raccomandazione.

Le linee guida non vengono usualmente pubblicate nei periodici scientifici, seppure questo non sia un evento da escludere. Vengono invece pubblicate, oltre che in forma cartacea, anche in formato elettronico sui siti web degli enti produttori. Inoltre, vengono raccolte in siti web che hanno lo scopo di collezionare e permettere l'accesso a materiali simili, come per esempio www.guidelines.gov, gestito dalla Agency for Healthcare Research and Quality statunitense.

Nella tabella 2 a pag. 24 sono elencati i principali siti web che vengono siste-

Base dati	Ente	Indirizzo web	Note
National Library of Guidelines	NICE-NHS	www.library.nhs.uk/guidelinesfinder/	gratuita
US National Guidelines Clearing House	US Agency for Healthcare Research and Quality (AHRQ)	www.guideline.gov/	gratuita
Guidelines Library	The New Zealand Guidelines Group	www.nzgg.org.nz/	gratuita
Clinical Practice Guidelines	The Australian National Health and Medical Research Council	www.nhmrc.gov.au/	gratuita
SIGN	Scottish Intercollegiate Guidelines Network	www.sign.ac.uk	gratuita

Tabella 2. Basi dati per la ricerca delle linee guida di sanità pubblica

maticamente ricercati dagli enti nazionali e internazionali che producono linee guida (NICE, AHRQ, SIGN, ecc). L'identificazione delle linee guida utili non richiede particolari strategie di ricerca. Una volta identificata la linea guida attinente, deve essere ricercata l'eventuale presenza di revisioni sistematiche prodotte per elaborare le raccomandazioni. Una volta identificate le revisioni, queste verranno valutate con le stesse procedure delle revisioni sistematiche pubblicate autonomamente (vedi paragrafo successivo).

Se vengono individuate **linee guida**, le stesse vengono esaminate da più ricercatori che ne valutano la qualità (vedi sottoparagrafo 4.7.1 *Valutazione delle linee guida* a pag. 30) e verificano la congruità degli argomenti trattati (ovvero, a quali quesiti effettivamente rispondono).

Nell'ambito della sanità pubblica le linee guida sono scarse. Per questa ragione è quasi sempre necessario accedere ad altre fonti secondarie o primarie.

4.5.2 Adattamento di linee guida internazionali

In futuro sarà sempre più frequente che sui grandi argomenti della sanità pubblica organismi internazionali formulino raccomandazioni sugli interventi migliori da adottare, talvolta redigendo vere e proprie linee guida con l'intento di orientare l'azione dei professionisti. È il caso dell'OMS che, per esempio, a partire dalla risoluzione *WHA 57.17: Global Strategy on Diet, Physical Activity and Health* nell'ambito della 57° World Health Assembly del maggio 2004 (http://apps.who.int/gb/ebwha/pdf_files/WHA57/A57_R17-en.pdf), ha avviato la produzione di linee guida relative alla lotta all'obesità, dal punto di vista sia della dieta sia dell'attività fisica. Ma anche altri enti autorevoli hanno imboccato questa strada, come il Department of Health degli Stati Uniti e il National Institute for Health and Clinical Excellence in Gran Bretagna (NICE, www.nice.org.uk). La metodologia di costruzione di linee guida da parte di questi organismi internazionali è spesso rigorosa, basandosi su revisioni sistematiche del-

la letteratura scientifica più aggiornata, con la collaborazione dei più eminenti esperti della materia, e spesso prevedendo complesse ricerche di consenso tra tutti i soggetti interessati.

In presenza di materiali di tale qualità esiste un'alternativa allo sforzo di ripetere a livello nazionale, o addirittura locale, la redazione di una linea guida: la traduzione e l'adattamento al contesto delle raccomandazioni formulate da questi organismi. È quanto peraltro suggerisce anche l'OMS nel presentare le sue *Global Recommendations on Physical Activity for Health* del 2010 (www.who.int/dietphysicalactivity/factsheet_recommendations/en/index.html), sottolineando l'importanza di adattare al contesto locale le raccomandazioni generali, al fine di tener conto di «norme sociali, valori religiosi, sicurezza pubblica a livello locale, disponibilità di spazi sicuri per lo svolgimento di attività fisica, caratteristiche meteorologiche e climatiche, problemi di genere, ruolo delle autorità locali, accesso alle scuole e ai luoghi di lavoro soprattutto per le donne, infrastrutture di trasporto, sportive e ricreative, caratteristiche urbanistiche, pattern di partecipazione in tutti i campi dell'attività fisica», avendo in mente, per esempio, quanto questi fattori possano condizionare l'adozione di raccomandazioni, pur fortemente sostenute da evidenze scientifiche.

Il processo di adattamento si giova da qualche anno di un nuovo strumento costruito proprio per agevolare l'adozione di un percorso standardizzato. Si tratta del manuale ADAPTE (www.adapte.org), che vuole accompagnare coloro i quali vogliono adottare raccomandazioni a livello locale derivanti da linee guida di altri organismi (The ADAPTE Collaboration 2009).

Il rationale del manuale di ADAPTE cita: «*The development and updating of high-quality clinical practice guidelines require substantial time, expertise and resources. There is pressure for organisations to produce guidelines rapidly, to ensure that medical practice is consistent with current, emerging medical knowledge and with increasingly limited resources. Guideline adaptation recognises and responds to legitimate differences in organisational, regional or cultural circumstances that could lead to variations in recommendations supported by the same evidence. We postulate that existing high quality guidelines as a resource may be an alternative to de novo development to reduce duplication of effort, enhance efficiency, and promote local uptake of quality guideline recommendations.*».

4.5.3 La ricerca di revisioni sistematiche

La **revisione sistematica (RS)**, o *systematic review*, è uno strumento di ricerca secondario il cui obiettivo è quello di riassumere dati provenienti da strumenti di ricerca primari. Questi documenti vengono pubblicati su periodici, su basi dati specifiche (Cochrane Library) oppure in documenti degli enti produttori.

La tabella 3 a pagina 26 riporta le principali basi dati nelle quali, oltre che in Medline, possono essere identificate revisioni sistematiche di interesse. Quando le revisioni sistematiche risultassero non adeguate a rispondere al quesito, per conte-

Base dati	Ente	Indirizzo web	Note
Cochrane Database of Systematic Reviews	The Cochrane Collaboration	www.cochrane.org/reviews	A pagamento
Database of Abstracts of Reviews of Effects (DARE)	Centre for review and dissemination University of York - NHS	www.crd.york.ac.uk/crdweb	Gratuita
Health Technology Assessment Database (HTA)	Centre for review and dissemination University of York - NHS	www.crd.york.ac.uk/crdweb	Gratuita
NHS Economic Evaluation Database	Centre for review and dissemination University of York - NHS	www.crd.york.ac.uk/crdweb	Gratuita
The Campbell Library	The Campbell Collaboration	www.campbellcollaboration.org/	Gratuita
Database of promoting health effectiveness reviews (DoPHER)	EPPI-Centre	www.eppi.ioe.ac.uk/cms/	Gratuita
Health Evidence	Public Health Agency of Canada	http://health-evidence.ca/	Gratuita

Tabella 3. Basi dati per la ricerca delle revisioni sistematiche

nuti o qualità metodologica, o qualora non fossero sufficientemente aggiornate, il gruppo di coordinamento e i documentalisti dovranno ricorrere all'analisi degli studi primari.

4.5.4. La ricerca di studi primari

La ricerca degli **studi primari** deve essere sistematica e si compone di due fasi: nella prima vengono ricercati gli studi pubblicati nelle riviste scientifiche. Per questa ricerca vanno consultati i principali database di studi primari, che sono elencati nella tabella 4 a pagina 27. La strategia di ricerca, che deve originare dai quesiti, non dovrebbe avere limitazioni di tempo o di lingua imposte dal ricercatore.

La seconda fase consiste nella ricerca di studi e documenti non pubblicati su riviste scientifiche. Questo materiale è costituito da rapporti di ricerca, documenti programmatici ed esperienze, atti di convegno, normative, banche dati specialistiche, e da ogni risorsa prodotta da istituzioni internazionali, nazionali e locali, enti di ricerca, associazioni, gruppi di interesse, che possa essere utile per inquadrare il contesto delle linee guida, per conoscere esperienze e studi che per loro natura, o per altre ragioni, non sono state pubblicate attraverso i canali dell'editoria tradizionale. Le risorse non pubblicate si trovano oggi principalmente sul web (alcuni esempi sono riportati in tabella 5 a pagina 27), o possono essere ricercate attraverso un contatto diretto con gli autori.

Base dati	Ente	Indirizzo web	Note
Medline	National Library of Medicine (USA)	www.ncbi.nlm.nih.gov/pubmed/	Gratuita
CINAHL	EBSCO	www.ebscohost.com/	A pagamento
EMBASE	Elsevier	www.embase.com	A pagamento
Cochrane Central Register of Controlled Trials	The Cochrane Collaboration	www.mrw.interscience.wiley.com	A pagamento
TRIP Database	TRIP Database Ltd	www.tripdatabase.com/	Gratuita
PsycINFO/PscyLIT	American Psychological Association	www.apa.org/psycinfo/	Gratuita
Sociological Abstracts	ProQuest e CSA	www.csa.com/factsheets/socioabs-set-c.php	A pagamento
ERIC	Educational Resources Information Center	www.eric.ed.gov/	Gratuita
Web of science	Institute of Scientific Information	www.thomsonreuters.com/products_services/scientific/Web_of_Science	A pagamento
Econlit	American Economic Association	www.econlit.org	A pagamento
ESDS Qualidata	UK Data Archive	www.esds.ac.uk/qualidata/	Gratuita

Tabella 4. Basi dati per la ricerca degli studi primari

Base dati	Ente	Indirizzo web	Note
The National Research Register	The National Institute for Health Research-UK	www.nihr.ac.uk/	Gratuita
Healthy People 2010 Information Access Project	U.S. Department of Health and Human Services	www.wonder.cdc.gov	Gratuita
CHID	Agenzie Federali US	www.cehn.org/archives	Gratuita
Dissertation Abstracts	Ann Arbor	www.library.dialog.com/bluesheets/html/bl0035.html	A pagamento
Conference Papers Index	ProQuest-CSA	www.csa.com/factsheets/cpi-set-c.php	A pagamento
Googlescholar.com	Google	www.scholar.google.it/	Gratuita

Tabella 5. Alcune basi dati per la ricerca degli studi non pubblicati

4.6 La funzione di documentazione

La funzione di documentazione ha un ruolo fondamentale nel processo di produzione di una linea guida di sanità pubblica. A partire dai quesiti elaborati dal panel sarà necessario:

- costruire un database (consultabile attraverso una community) che permetta l'accesso al testo integrale di tutti i documenti. Il gruppo di lavoro potrà consultare direttamente gli studi per approfondire singoli temi, e proporre modifiche o commenti alla linea guida in corso di elaborazione;
- individuare i documenti pertinenti a ciascun quesito. Verranno effettuate diverse ricerche sulle banche dati biomediche e su altre fonti, per individuare studi e materiali utili all'elaborazione della linea guida. Dagli elenchi dei materiali trovati, completi di abstract, si procederà a una prima selezione del materiale pertinente rispetto ai singoli quesiti;
- contribuire all'analisi in doppio degli studi e descrivere gli stessi in tabelle;
- elaborare una sintesi descrittiva degli studi e della loro qualità, completa di una proposta di raccomandazione (o non raccomandazione) rispetto a ogni singolo quesito. La relazione dovrà contenere il numero di studi selezionati, gli studi esclusi e le motivazioni per l'esclusione, una prima valutazione della qualità degli studi inclusi, una sintesi dei risultati, una conclusione preliminare.

4.7 La valutazione critica delle prove di efficacia

La valutazione critica di ogni rapporto di ricerca (pubblicato o non pubblicato) rappresenta un passaggio fondamentale per il corretto utilizzo delle conoscenze nella pratica di sanità pubblica. Ogni tipo di documento deve essere sottoposto a un modello di valutazione proprio, ma, secondo il modello proposto da Sackett, i tre pilastri della Evidence Based Medicine, nel momento della valutazione della conoscenze pubblicate in letteratura, sono il giudizio sulla qualità dello studio o della revisione di studi (validità interna), quello sulla dimensione dell'effetto dell'intervento (importanza) e quello della sua applicabilità alla situazione locale (generalizzabilità o validità esterna) (Sackett 2000).

Se proviamo ad adattare le domande fondamentali dell'EBM alla pratica della sanità pubblica possiamo formulare una serie di quesiti secondo lo schema che segue:

- questo studio/revisione è valido?
- questo studio/revisione valido ha un impatto sulla popolazione degno di considerazione per il decisore?
- questo studio/revisione valido e rilevante per il decisore, è applicabile al contesto (sociale, politico organizzativo, ecc.) cui si rivolge la raccomandazione?

Queste domande formano la trama di ogni strumento di valutazione di buona qualità. Ognuna di esse, ovviamente, deve essere sviluppata nei suoi elementi fondamentali, anche se, di solito, il livello di evidenza è attribuito in base al solo disegno dello studio.

Le prove di efficacia vengono classificate secondo livelli di evidenza che riflettono il grado di validità (assenza di distorsioni sistematiche) dei risultati e delle conclusioni

relative all'efficacia dei trattamenti riportati nei singoli studi. Il presupposto su cui si basano è che i disegni di studio dei livelli superiori siano in grado di fornire evidenze più valide di quelli ai livelli più bassi. Infatti, affinché i risultati di uno studio possano essere attribuiti con relativa certezza all'intervento sperimentale, è necessario che il disegno dello studio abbia rispettato una serie di requisiti metodologici. In caso contrario, possono verificarsi errori sistematici che possono portare a concludere erroneamente per l'efficacia o la non efficacia di un intervento. Le prove di efficacia reperite negli studi vengono interpretate tenendo conto dei livelli di evidenza, della numerosità e della qualità metodologica degli studi, della consistenza dei risultati e della rilevanza clinica delle misure di risultato utilizzate. Tale interpretazione, alla luce anche dell'esperienza clinica degli autori e del contesto locale in cui le linee guida andranno implementate, porta alla formulazione del grading, cioè della forza delle raccomandazioni.

Il metodo utilizzato per valutare la qualità delle evidenze dall'OMS e da altre importanti agenzie di produzione di linee guida, quali per esempio il NICE, è quello descritto dal GRADE Working Group (Atkins 2004, Guyatt 2008a). Questo approccio permette di valutare la qualità delle prove rispetto alle singole specifiche domande poste nelle linee guida, tenendo in considerazione la dimensione dell'effetto, la rilevanza degli esiti utilizzati rispetto al quesito posto, la numerosità del campione degli studi considerati, la metodologia utilizzata e la consistenza dei risultati tra i diversi studi.

Preliminarmente però al processo di identificazione e valutazione delle prove, il panel decide l'importanza relativa delle varie misure di esito, collocandole in una scala da 1 a 9. Convenzionalmente la valutazione dell'importanza degli esiti è la seguente (Guyatt 2008b):

- punteggi tra 1 e 3: gli esiti non sono considerati importanti e generalmente non verranno utilizzati per formulare le raccomandazioni;
- punteggi tra 4 e 6: gli esiti sono considerati importanti, ma non critici;
- punteggi tra 7 e 9: gli esiti sono considerati critici, cioè particolarmente rilevanti e indispensabili per la formulazione delle raccomandazioni.

Una volta stabiliti da parte del panel gli esiti da considerare e la loro relativa importanza, un gruppo di metodologi valuterà la qualità delle prove disponibili per ogni esito da considerare. Nel sistema GRADE, la qualità delle prove viene valutata come segue (Guyatt 2008b):

- livello alto: è molto improbabile che ulteriori ricerche modifichino la nostra fiducia nella stima dell'effetto;
- livello intermedio: è probabile che ulteriori ricerche possano modificare la nostra fiducia nella stima dell'effetto;
- livello basso: è molto probabile che ulteriori ricerche modifichino la nostra fiducia nella stima dell'effetto;
- livello molto basso: la stima dell'effetto è molto incerta.

Questo metodo permette una valutazione esaustiva, obiettiva, e soprattutto riproducibile ed esplicita, delle prove disponibili. Essa è però basata sui principi dell'EBM che tradizionalmente considera gli studi controllati randomizzati il gold standard su cui confrontare gli altri tipi di disegno di studio; il suo utilizzo può pertanto essere considerato problematico in campi, come quello della salute pubblica, dove gli studi randomizzati controllati (RCT) non sono il disegno di studio più frequentemente utilizzato. Tuttavia, il metodo GRADE ammette l'esistenza di specifiche condizioni in cui studi diversi dagli RCT possono raggiungere un punteggio di qualità elevato, quali per esempio la dimensione della stima dell'effetto o la confidenza sull'assenza di un importante confondimento (Guyatt 2008c). Inoltre, vi è al momento un intenso dibattito tra i membri del gruppo GRADE che coinvolge anche tutti coloro che a vario titolo vogliono utilizzare questo metodo per la valutazione delle prove nei processi di stesura delle linee guida per individuare procedure che consentano il suo utilizzo anche in casi in cui le prove disponibili provengano da differenti disegni di studio.

L'uso di GRADE in un campo in cui le evidenze provengono spesso da studi di bassa qualità può contribuire ad aumentarne il rigore scientifico e la trasparenza del processo di costruzione delle raccomandazioni. Peraltro, per produrre le raccomandazioni, il panel di esperti non considera solo le evidenze valutate con il metodo GRADE, ma anche gli aspetti tecnici, la disponibilità di risorse e il rapporto rischi-benefici delle varie alternative possibili. La forza delle raccomandazioni si baserà quindi su tutte queste considerazioni.

Nei paragrafi che seguono verranno sintetizzati alcuni suggerimenti relativi alla valutazione della qualità delle evidenze provenienti dai diversi strumenti che possono contribuire alla produzione di linee guida di prevenzione.

4.7.1 Valutazione delle linee guida

Le linee guida elaborate da altri centri sono uno strumento utile a tre scopi:

- 1) per contribuire alla costruzione del *logical framework*;
- 2) come fonte di revisioni sistematiche;
- 3) come traccia per la costruzione delle raccomandazioni.

A questo fine è però necessario che le linee guida da utilizzare siano documenti validi scientificamente. Si suggerisce di utilizzare lo strumento AGREE (Appraisal of Guidelines Research and Evaluation) sviluppato come parte di un progetto finanziato dal Programma EU Biomedicine and Health Research (BIOMED 2) 1994-1998 (The AGREE Collaboration 2003). La traduzione italiana della seconda versione originale è consultabile sul sito web www.gimbe.org/agree/AGREE_IT.pdf.

AGREE è un modello utile a valutare le componenti principali di una linea guida, esaminando sia il processo di sviluppo, sia quello di stesura. Esso non fornisce alcun elemento utile per giudicare i contenuti specifici della linea guida o la qualità del-

le prove di efficacia portate a giustificazione delle raccomandazioni. Lo strumento comprende sei domini teorici:

- 1) obiettivi della linea guida;
- 2) coinvolgimento degli attori;
- 3) rigore metodologico nello sviluppo;
- 4) chiarezza e presentazione degli argomenti;
- 5) applicabilità;
- 6) indipendenza editoriale.

Questi domini sono articolati in 23 quesiti. L'accettabilità dello strumento è stata giudicata elevata, risultando di semplice applicazione; anche l'affidabilità è buona con una riproducibilità al test Cronbach alpha che oscilla nei vari domini fra 0,64 e 0,88.

4.7.2 Valutazione delle revisioni sistematiche

Sia le revisioni sistematiche pubblicate come tali, sia quelle pubblicate come parte integrante di una linea guida devono essere valutate criticamente. Qualora fossero in grado da sole di rispondere a un quesito (vedi figura 2 a pag. 19), potrebbero essere infatti l'unica fonte di evidenze scientifiche di una raccomandazione.

Esistono numerosi strumenti di cui possiamo servirci a questo scopo:

- la check-list elaborata da CASP (Critical Appraisal Skills Programme), e riportata anche come Glasgow check-list, che trae origine dal documento di Oxman, Cook e Guyatt del 1994 (Oxman 1994);
- la AMSTAR (Assessment of Multiple Systematic Reviews);
- la ARIF (Aggressive Research Intelligence Facility), lo strumento messo a punto da SIGN (Scottish Intercollegiate Guideline Network);
- la check-list elaborata dal Canadian Institutes of Health Research per la valutazione di qualità delle revisioni da includere nel database di Health-Evidence (che verrà presentata più avanti insieme al database);
- PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), uno sviluppo di QUOROM (QUALity Of Reporting Of Meta-analyses), seppure mirato alla qualità del reporting più che alla qualità dello studio.

Lo strumento messo a punto dal **CASP** prevede 10 domande, a ognuna delle quali è possibile rispondere con sì, no o non so. Le prime due domande prendono in considerazione il fatto che il quesito sia formulato in modo chiaro, e se la revisione utilizzi un appropriato disegno di studio. Se la risposta a entrambe queste domande è positiva, è possibile procedere oltre, altrimenti il giudizio viene considerato a priori negativo. Le domande successive prendono in considerazione la qualità della ricerca di letteratura condotta e la valutazione di qualità degli studi inclusi. Inoltre, largo spa-

zio è dedicato ai risultati: viene chiesto se i risultati di più studi siano stati combinati, quali siano i risultati principali, intendendo con quale misura sono espressi e con che precisione, e se questi possano essere applicati alla popolazione locale. Infine, si chiede se vengano presi in considerazione i principali outcome, e se una politica o una pratica sanitaria possa cambiare in base alle evidenze raccolte dalla revisione stessa. La checklist per la valutazione della qualità degli studi è disponibile presso il sito web www.sph.nhs.uk/what-we-do/public-health-workforce/resources/critical-appraisals-skills-programme.

Lo strumento **AMSTAR** prevede 11 item per valutare la qualità metodologica di revisioni sistematiche. È stato costruito empiricamente dopo aver valutato oltre 150 revisioni sistematiche con strumenti già in uso e nuove variabili; secondo gli autori sono necessari studi volti a valutare la sua riproducibilità, prima che l'utilizzo di questo strumento possa essere raccomandato (Shea 2007).

Infine, lo strumento messo a punto da **SIGN**, utilizzabile per le revisioni sistematiche e le metanalisi che prendono in considerazione le varie tipologie di studio, è suddiviso in tre sezioni: la prima, relativa alla validità interna, considera se il quesito sia espresso in modo chiaro e adeguato, se sia presente una descrizione della metodologia utilizzata, quale siano il rigore della ricerca di letteratura, la valutazione della qualità degli studi analizzati dalla revisione, la confrontabilità degli studi; la seconda sezione riguarda la valutazione della presenza di bias e in che modo, se presenti, questi possano condizionare i risultati; l'ultima sezione riguarda la tipologia degli studi presi in considerazione e le conclusioni principali della revisione (www.sign.ac.uk/guidelines/fulltext/50/checklist1.html).

Lo strumento sviluppato da **ARIF**, Università di Birmingham, è semplice, basato su 14 item, corredato da spiegazioni chiare sul suo utilizzo, che mira a valutare la qualità della revisione, a partire dai suoi metodi (www.arif.bham.ac.uk/critical-appraisal-checklist.shtml).

Lo strumento **PRISMA**, già denominato QUOROM, messo a punto per il miglioramento del *quality reporting* di supporto per gli autori di revisioni sistematiche e metanalisi, può anche essere utilizzato per la valutazione critica di questi documenti; differisce dai precedenti perché valuta in maniera dettagliata le varie sezioni che compongono una revisione, partendo dal titolo della stessa, analizzando l'abstract, l'introduzione, i metodi, i risultati e la parte relativa alla discussione; infine, chiede agli autori di dare informazioni circa i metodi impiegati per finanziare la produzione del documento (Moher 2009).

Due enti internazionali effettuano sistematicamente valutazioni di revisioni sistematiche: il Centre for Reviews and Dissemination (CRD) della Università di York, che alimenta il database DARE (Database of Abstracts of Reviews of Effects), della Cochrane Collaboration (www.crd.york.ac.uk/crdweb), e il Canadian Institutes of Health Research, che produce Health-Evidence (<http://health-evidence.ca>).

Per quanto riguarda il **DARE** database, ogni mese vengono vagliate e valutate decine di revisioni da parte di due ricercatori che lavorano in modo indipendente, che ne riassumono il contenuto in schede di sintesi che descrivono i metodi utilizzati, i risultati e le conclusioni, e le valutano dal punto di vista della qualità metodologica con un breve commento allo scopo di metterne in luce i punti deboli e i punti di forza. A oggi sono state trattate circa 15.000 revisioni; il commento valutativo è presente per più di 6.000 di queste.

I criteri che consentono a una revisione sistematica di essere inclusa nel database sono i seguenti:

- sono specificati i criteri di inclusione e di esclusione degli studi primari;
- la ricerca è stata condotta in maniera sistematica, esaustiva e rigorosa;
- vi è una sintesi degli studi presi in considerazione;
- gli studi inclusi sono stati valutati;
- vi sono sufficienti dettagli relativi ai singoli studi inclusi.

Per essere inclusa nel database la revisione sistematica deve rispettare obbligatoriamente il primo e il secondo criterio e almeno altri due tra i restanti tre. Le revisioni incluse possono trattare temi prettamente clinici o affrontare problemi di sanità pubblica: l'attenzione è rivolta a tutti gli interventi che possono avere un effetto sulla salute umana. La ricerca viene effettuata esplorando sia la letteratura scientifica pubblicata, sia quella grigia.

Health-evidence, Promoting evidence-informed decision making è un database canadese costruito per fornire le evidenze ai decisori di sanità pubblica, consentendo di risparmiare il tempo speso nella ricerca, nella selezione e nella valutazione delle revisioni ([http://health-evidence.ca/downloads/QA tool_Doc 4.pdf](http://health-evidence.ca/downloads/QA_tool_Doc_4.pdf)). Il sito ha censito oltre 500.000 titoli, dei quali qualche migliaia sono stati considerati rilevanti, valutati qualitativamente e inclusi nel database. Il sito nasce da un progetto di ricerca del Canadian Institutes of Health Research, al fine di avere una fonte accreditata di revisioni sulle evidenze, che i decisori di sanità pubblica potessero consultare con facilità.

Per essere considerate rilevanti, e quindi essere incluse nel database, le revisioni devono rispettare cinque criteri, esposti, insieme allo strumento, nell'allegato 1 *Check-list per la valutazione critica di revisioni sistematiche (Health-Evidence)* a pag. 51.

Per le revisioni considerate rilevanti viene condotta una valutazione di qualità che fornisce un giudizio supportato dal commento di *strong*, *moderate* o *weak*, oltre a un punteggio da 1 a 10. La valutazione viene condotta prendendo in considerazione 10 item (vedi allegato 1, pag. 51).

La check-list utilizzata da Health-evidence viene ritenuta particolarmente adeguata agli scopi di questa guida, per cui si raccomanda di utilizzarla per la valutazio-

ne di qualità delle revisioni sistematiche usate per le linee guida di prevenzione. Questa scelta permette di fruire dei giudizi di qualità elaborati dal Canadian Institute of Health Research per le revisioni incluse nel database.

4.7.3 Valutazione degli studi primari

I disegni di studio adottati per la valutazione di interventi e programmi di sanità pubblica sono parzialmente diversi da quelli della medicina clinica. In particolare, la randomizzazione, per mezzo della quale si vogliono eliminare le differenze fra il gruppo dei trattati e il gruppo di controllo, a parte ovviamente l'esposizione all'intervento, non è frequentemente applicabile. Per esempio, l'esposizione a interventi, quali le normative o le campagne informative, non può essere assegnata in modo casuale, né il ricercatore può esercitare alcun controllo sull'attuazione dell'intervento, del programma o della politica.

La qualità della prova, o la sua posizione nella gerarchia degli studi di efficacia, dipende dunque dal tipo di intervento e dalla popolazione cui è rivolto e, proprio in quanto *ranking* relativo, si deve adattare a questi fattori.

Il paragrafo 4.8 *Caratteristiche dell'intervento e gerarchia delle prove di efficacia* viene dedicato alla qualità delle prove di efficacia nel campo della prevenzione.

4.8 Caratteristiche dell'intervento e gerarchia delle prove di efficacia

L'intervento può essere oggetto di randomizzazione? Cioè, può essere assegnato a un individuo o per lo meno a un gruppo di individui, senza che questo coinvolga eventuali individui o gruppi di controllo? Questo è il caso di molti interventi di prevenzione quali le vaccinazioni, il counselling individuale per adolescenti a rischio di abuso di droga, gli interventi scolastici e familiari.

Quando l'intervento può essere randomizzato, lo studio randomizzato (o il *cluster-randomised* nel caso l'intervento sia somministrabile a gruppi) deve essere il gold standard su cui si costruisce la gerarchia dell'evidenza.

Il disegno di studio varierà a seconda che si tratti di un intervento erogabile a livello individuale o a livello di gruppo. In quest'ultimo caso, come per esempio un intervento di promozione di una pratica vaccinale in un distretto o un programma scolastico di prevenzione, una scelta efficiente è quella di randomizzare le unità a cui viene somministrato l'intervento (classi, scuole, distretti), ma di misurare l'outcome a livello individuale. L'adozione di disegni di studio tipo *cluster-randomised* in questi casi è necessaria al fine di tenere conto della perdita di potenza dovuta alla minore variabilità *intracluster* (Murray 2004), rispetto a quella fra *cluster*. Per esempio, entro una classe di una scuola superiore i comportamenti degli studenti tendono a uniformarsi, riducendo la variabilità *intracluster*, e a differenziarsi da quelli di altre classi, aumentando la variabilità

fra *cluster*. Questo fenomeno deve essere corretto in fase di disegno inflazionando il *sample size*, e in fase di analisi utilizzando analisi di tipo multilivello.

Oltre al **disegno dello studio** (randomizzato/non randomizzato, controllato/non controllato) e alla sua validità interna, è necessario valutare anche i soggetti inclusi, o **target** (individui/popolazioni), e i **fattori di contesto** che influenzano i risultati, sia nel gruppo sottoposto a intervento sia, quando è disponibile, nel gruppo di controllo non sottoposto a intervento o sottoposto a interventi diversi (o di intensità diversa). Per lo studio dei fattori di contesto è di fondamentale utilità la valutazione della plausibilità dell'effetto ottenuto nella popolazione sulla quale è stato condotto l'intervento (Habicht 1999, Victora 2004).

Dovremo poi valutare l'entità del beneficio ottenuto mediante l'intervento, sia in termini di popolazione o gruppi, sia considerando i costi dello stesso. Da ultimo è necessario capire quanto l'intervento valutato dallo studio sia applicabile al nostro contesto (per ragioni culturali, economiche, ambientali, politiche, ecc.), cioè la sua validità esterna, e in questo gli esperti di settore che partecipano all'elaborazione della linea guida potranno essere particolarmente utili.

Quando l'intervento non è randomizzabile si possono utilizzare solo studi **ecologici** basati sul paradigma del *before&after* (Cook 1979). In particolare, Eccles identifica *controlled/uncontrolled before and after study* e *time series design* (Eccles 2003). West include anche la *discontinuity regression analysis*, per quanto questa costituisca un esperimento naturale poco frequente (West 2008).

Vi sono situazioni in cui l'intervento può essere assegnato individualmente, ma non è fattibile la randomizzazione, come per esempio per la valutazione di efficacia di una vaccinazione, qualora non sia eticamente accettabile la randomizzazione per il livello di confidenza sull'efficacia dell'intervento, oppure quando si vogliono valutare effetti rari o a lungo termine di un intervento, come per esempio gli effetti collaterali di una vaccinazione. In questi casi vengono condotti studi osservazionali di tipo longitudinale. Tra questi, mentre è riconosciuto il ruolo degli studi di coorte per valutare l'eventuale associazione tra interventi ed effetti su individui di determinate popolazioni, alcune guide metodologiche sull'Evidence Based Public Health (EBPH) sconsigliano disegni di studio di tipo caso-controllo per valutare gli effetti di interventi di sanità pubblica, mentre ne giustificano l'impiego per evidenziarne gli effetti indesiderati (SNIPH 2001).

Considerando anche i target dello studio possiamo allora classificare i disegni come rappresentato in tabella 6 a pagina 36.

Viene di seguito presentato un elenco di definizioni operative dei disegni di studio che verranno presi in considerazione in questo manuale.

- Definiamo **randomized controlled trial** (RCT) lo studio il cui disegno prevede la presenza di un gruppo di controllo e di quello sperimentale, e l'allocazione casuale in uno dei due gruppi di singoli individui.

Target	Sperimentali	Non sperimentali
Individuo o gruppo (intervento randomizzabile)	RCT	Caso-controllo Coorte
Popolazione (intervento non randomizzabile a livello individuale)	<i>Cluster randomized trial</i> (quando l'intervento è assegnabile a gruppi)	<i>Before and after</i> <i>Controlled/uncontrolled</i> <i>Time series</i> <i>Discontinuity regression analysis</i>

Tabella 6. Una sintesi degli studi valutativi in sanità pubblica

- Definiamo ***cluster randomized controlled trial*** (CRCT) lo studio il cui disegno prevede l'allocazione casuale di unità organizzative o amministrative a un gruppo di intervento e a uno di controllo. In esso l'assunzione di indipendenza fra le osservazioni è violata e bisogna tenerne conto nell'analisi. La valutazione di esito può invece essere effettuata a livello degli individui che compongono il gruppo.
- Definiamo **studio caso-controllo** lo studio il cui disegno prevede il confronto di soggetti, detti casi, che hanno subito l'evento in studio (per esempio una diagnosi di malattia) con uno o più controlli che invece non lo hanno subito: i diversi fattori a cui sono stati esposti casi e controlli, raccolti retrospettivamente, potranno identificare i determinanti della malattia in questione.
- Definiamo **studio di coorte** l'osservazione di una coorte di esposti e di una di non esposti al fattore sottoposto a valutazione, fino alla comparsa della condizione in studio; anche in questo caso si tratta di studi a target individuale.
- Definiamo **studio *before and after* controllato o non controllato** (CB&A e B&A) lo studio di valutazione di efficacia di un intervento il cui disegno prevede una misurazione dell'outcome a livello di popolazione (ecologico) prima e dopo l'intervento. Il controllo è costituito da un gruppo in cui l'outcome viene misurato nello stesso tempo, ma l'intervento non è applicato.
- Definiamo ***time series design*** (ITS) lo studio il cui disegno prevede molteplici misure dell'outcome prescelto, sia prima sia dopo l'intervento, a livello di popolazione (ecologico). In questo modo è possibile evidenziare l'effetto di andamenti "natural" dei fenomeni (trend storico), altrimenti non misurati negli studi *before and after*. Come questi ultimi, rappresentano la sola tipologia di studi praticabile per interventi ad ampio spettro sulla popolazione, che impediscono l'identificazione di gruppi di controllo.
- Definiamo ***discontinuity regression analysis*** lo studio che analizza una popolazione di individui posti in graduatoria con una soglia che divide la stessa in due gruppi: coloro che beneficeranno dell'intervento (per esempio in base al reddito o all'età) da coloro che non ne hanno diritto. Supponendo che coloro che si trovano in prossimità della soglia, alcuni dei quali quindi inclusi fra i beneficiari e altri esclu-

si, abbiano caratteri (definiti dai criteri di ammissibilità all'intervento) simili, si considera che gli ultimi ammessi costituiscano il gruppo di intervento e i primi esclusi quello di controllo.

4.9 Qualità degli studi

4.9.1 Qualità degli studi randomizzati

La valutazione della qualità degli studi randomizzati ha una lunga tradizione nell'ambito dell'EBM e in Italia, per cui non verrà qui affrontata in dettaglio. La griglia suggerita per questa valutazione è quella del CASP (consultabile su www.phru.nhs.uk/pages/phd/casp.htm), modificata da Guyatt (Guyatt 1993, Guyatt 1994) e riportata nell'allegato 2 *Check-list per la valutazione critica degli studi randomizzati e controllati* a pag. 53. Questa scala è stata giudicata un ottimo compromesso fra validità e applicabilità (Deeks 2003). Le scale sperimentate per la valutazione di *cluster randomised trial* sono invece rare. Per questo viene invece consigliata la scala CASP per gli RCT accoppiata con le osservazioni riportate da Eldridge (Eldridge 2008). Può essere di aiuto anche la consultazione delle griglie di valutazione elaborate nell'ambito di CONSORT (Consolidated Standards of Reporting Trials) e disponibili nel sito web www.equator-network.org. Si tratta di check-list per la valutazione del reporting, e quindi non direttamente della qualità metodologica, ma sono comunque spesso utili strumenti di valutazione. Inoltre, l'Handbook della Cochrane Collaboration consente di confrontare i criteri utilizzati da questa organizzazione per la valutazione di qualità degli studi (Higgins 2008).

4.9.2 Qualità degli studi non randomizzati

Gli studi non randomizzati che verranno presi in considerazione da questo manuale sono i seguenti, elencati secondo una gerarchia dal meno valido al più valido:

- B&A non controllati (B&A);
- B&A controllati (CB&A);
- ITS non controllate;
- ITS controllate.

Gli studi *before and after*, prevedendo una sola misura della variabile di outcome prima, e una dopo, l'intervento, non sono in grado di controllare l'eventuale esistenza di una tendenza di lungo periodo nella prevalenza della variabile in studio. Per esempio, durante gli anni '90 e il primo decennio del 2000 la prevalenza di fumatori in Italia è sempre diminuita. Uno studio B&A per la valutazione di un intervento di prevenzione avrebbe mostrato una riduzione della prevalenza, che non poteva essere attribuita all'intervento. La presenza di una coppia di misure analoghe, in una popola-

zione di controllo (CB&A) non esposta all'intervento (per esempio una regione), avrebbe potuto suggerire l'esistenza di un trend storico, e quindi di una spiegazione alternativa, ma sovente le popolazioni di controllo sono sufficientemente lontane da quelle di intervento da non permettere di escludere spiegazioni alternative nei fattori di contesto che differenziano le popolazioni.

Gli studi di serie temporali interrotte (ITS) sono indicati per controllare per il trend storico. Una serie di misure dell'esito in studio viene effettuata prima e dopo l'intervento, e viene valutata la discontinuità eventualmente provocata dall'intervento nella retta interpolata alle misure (Ramsay 2003). La qualità dell'evidenza che deriva da un tale studio dipende principalmente dalla qualità del controllo del trend temporale di base e dalla possibilità di escludere che un'eventuale discontinuità evidenziata possa avere spiegazioni alternative. Il controllo del trend temporale dipende dal numero delle misure effettuate prima e dopo. Nonostante la letteratura sull'argomento non sia concorde, il gruppo Effective Practice and Organisation of Care della Cochrane Collaboration (Bero 2002) suggerisce che siano necessarie almeno 3 misure prima e 3 dopo l'intervento (<http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/inttime.pdf>).

Oltre al criterio del numero di misure che compongono le serie, un elemento chiave dello studio deve essere l'accurata disamina delle possibili spiegazioni alternative della discontinuità eventualmente osservata, che richiede un approccio di tipo sociale più che statistico-epidemiologico. Nell'allegato 3 a pag. 56 è riportata una checklist proposta dal gruppo EPOC ed estratta da Ramsay per la valutazione degli studi ITS (Ramsay 2003).

4.10 Definizione dei livelli di evidenza

Alla luce di quanto detto finora, la gerarchia delle evidenze, che viene normalmente adottata nel giudizio sulla validità scientifica delle prove durante l'elaborazione di linee guida, è applicabile completamente solo agli studi di valutazione degli interventi che ammettono un'esposizione a livello individuale. Per questi interventi possono quindi essere applicati i criteri in uso da parte dell'SNLG per la classificazione delle prove (tabella 7 a pagina 39). Per gli interventi non randomizzabili potrebbe essere utile adottare una diversa gerarchia delle prove, nell'ambito della quale possano contribuire a produrre raccomandazioni forti anche studi di qualità che hanno però una posizione più bassa nella scala tradizionale della validità (Ogilvie 2005, Faggiano 2006). Tale approccio permette di temperare la portata della *inverse evidence law* secondo la quale gli interventi dotati di maggiore impatto sulla salute della popolazione, quelli cioè che hanno la popolazione come target (quali le normative, le leggi, ecc.), sono quelli che presentano meno evidenze di efficacia (Main 2005).

La possibile gerarchia suggerita per questo tipo di studi, sulla base di Cook (Cook 1979) e di Eccles (Eccles 2003), è riportata nella tabella 8 a pagina 39.

Livelli di prova per gli interventi randomizzabili	
I	Prove ottenute da più studi clinici controllati randomizzati e/o da revisioni sistematiche di studi randomizzati
II	Prove ottenute da un solo studio randomizzato di disegno adeguato
III	Prove ottenute da studi di coorte non randomizzati con controlli concorrenti o storici o loro metanalisi
IV	Prove ottenute da studi retrospettivi tipo caso-controllo o loro metanalisi
V	Prove ottenute da studi di casistica (serie di casi) senza gruppo di controllo
VI	Prove basate sull'opinione di esperti autorevoli o di comitati di esperti come indicato in linee guida o consensus conference, o basata su opinioni dei membri del gruppo di lavoro responsabile di queste linee guida

Tabella 7. Livelli di prova di efficacia per gli studi randomizzabili, in base alle indicazioni dell'SNLG

Livelli di prova per gli interventi non-randomizzabili	
I	Prove ottenute da più studi ITS con popolazione di controllo e/o da revisioni sistematiche di studi ITS
II	Prove ottenute da un solo studio ITS con popolazione di controllo di disegno adeguato
III	Prove ottenute da studi B&A controllati o loro metanalisi
IV	Prove ottenute da studi B&A senza popolazione di controllo o loro metanalisi
V	Prove ottenute da studi di casistiche senza gruppo di controllo
VI	Prove basate sull'opinione di esperti autorevoli o di comitati di esperti come indicato in linee guida o consensus conference, o basata su opinioni dei membri del gruppo di lavoro responsabile di queste linee guida

Tabella 8. Proposta di livelli di prove di efficacia per gli studi di interventi non randomizzabili

4.11 Il parere degli esperti

Abbiamo già visto nel capitolo 2 *L'integrazione delle prove scientifiche e del parere degli esperti* a pag. 11 che lo scopo di questo manuale è esplicito: i risultati della ricerca scientifica spesso non sono sufficienti o adatti a dare risposte specifiche per le scelte di sanità pubblica, e dunque riteniamo necessaria una stretta, ma rigorosa, integrazione di questi con l'opinione di coloro che lavorano quotidianamente a vari livelli nel campo.

Solitamente il metodo più utilizzato per conoscere e sintetizzare il parere degli esperti di un certo settore consiste nell'assumere il parere del singolo opinion leader, oppure mettere a confronto persone competenti, permettendo loro di scambiare le proprie idee in modo cosiddetto informale.

Dagli anni '50 hanno cominciato a prendere corpo i metodi formali di raccolta

del parere di esperti e di recente questi hanno trovato un'applicazione nell'elaborazione di linee guida cliniche (Murphy 1998).

Per quale ragione è meglio utilizzare metodi formali rispetto agli informali? Il motivo principale è costituito dal fatto che, in una discussione, la presenza di altri può essere un ostacolo e non un aiuto nella genesi e nell'espressione di nuove idee. In particolare, alcuni autori hanno evidenziato come, in un gruppo, la produttività ideativa di coloro che hanno capacità riconosciute è potenziata, mentre viene ostacolata l'espressione dei nuovi adepti in quel settore. La ricerca sociologica sul conformismo ha evidenziato come i comportamenti conformisti di adeguamento alla maggioranza o all'idea dell'opinione leader impoveriscano la discussione e la generazione di nuove idee. Il bisogno di raggiungere un accordo può superare quello di essere accurati e di considerare tutte le alternative possibili.

L'approccio formale è dunque giustificato per diverse ragioni:

- più persone hanno meno probabilità di assumere una decisione errata rispetto a quello che può fare un singolo individuo;
- un gruppo ben selezionato di esperti garantisce maggiore autorevolezza alle raccomandazioni prodotte;
- la razionalità della scelta è aumentata dal processo più strutturato e riflessivo della semplice discussione;
- il processo decisionale supera gli aspetti negativi della decisione assunta in gruppo;
- i metodi formali di consenso aumentano la credibilità scientifica del prodotto.

Esistono diversi metodi per la raccolta formale del consenso; le differenze principali sono rappresentate da alcuni fattori:

- sono utilizzati dei questionari, anche molto sintetici, attraverso la posta elettronica;
- i soggetti coinvolti nel processo devono esprimere separatamente e in modo personale il proprio parere;
- fra una consultazione individuale e l'altra ci sono dei momenti di feedback nei quali ogni esperto può rendersi conto del complessivo parere del gruppo in quel momento;
- possono esserci dei momenti di incontro e scambio fra i componenti del gruppo di esperti;
- i risultati delle varie fasi di consultazione sono analizzati e aggregati in modo esplicito.

La tabella 9 a pagina 41 descrive i metodi di consenso di possibile applicazione nell'elaborazione di linee guida e le caratteristiche principali.

La tecnica del gruppo nominale si basa sull'idea che gli aspetti non interattivi permettano il lavoro comune di soggetti con opinioni e competenze diverse e la genera-

	Questionari postali	Decisioni individuali	Offerta di feedback	Scambi diretti fra esperti	Interazioni strutturate fra esperti	Metodi di aggregazione espliciti
Metodi informali	no	no	no	si	no	no
Metodo di Delphi	si	si	si	no	si	si
Gruppo nominale (RAND)	si	si	si	si	si	si
Conferenza di consenso	no	no	no	si	no	no

Tabella 9. Caratteristiche dei metodi di consenso (Murphy 1998)

zione di idee. Lo scopo della tecnica è di creare interazione all'interno di un gruppo. In un primo momento ogni membro del gruppo deve esprimere per iscritto la propria idea sul tema proposto in modo privato e indipendente; nei gruppi di esperti l'indipendenza è spesso difficile da raggiungere, soprattutto quando i soggetti abbiano avuto esperienze di lavoro comune. Le tecniche di comunicazione attraverso la rete facilitano moltissimo questo tipo d'interazione, con l'esclusione delle web community. Il passaggio seguente prevede la condivisione delle idee anonime emerse in una seduta comune. La raccolta individuale consente di affinare il testo delle raccomandazioni vere e proprie. L'ultima fase di condivisione e voto sulle affermazioni conclude il processo, cui segue l'aggregazione statistica dei pareri (Murphy 1998). Da quest'ultima espressione dei pareri deve emergere il grado di rilevanza di ogni raccomandazione che, con il livello delle prove di efficacia sottostanti, determina il grado di forza di ogni raccomandazione. In questa sede sembra opportuno raccomandare la tecnica del gruppo nominale nella sua forma proposta dalla RAND Corporation, che meglio sembra adattarsi alle esigenze degli autori di linee guida. Questo è anche il metodo attualmente più utilizzato a questo scopo nell'esperienza internazionale.

4.12 La formulazione delle raccomandazioni e il grado di forza

Definire il grado di validità delle informazioni scientifiche poste alla base di una serie di raccomandazioni e chiarire quanta importanza si attribuisce alla loro applicazione costituisce uno dei caratteri distintivi del processo di produzione di linee guida basate sulle prove di efficacia.

È in corso da parte dell'SNLG un lavoro di valutazione critica e di sistematizzazione dei diversi metodi utilizzati da agenzie e gruppi internazionali. In particolare, è in discussione l'adozione dell'approccio proposto dal GRADE Working Group, che comporterebbe un aumento del rigore e della trasparenza della costruzione delle raccomandazioni, ma anche un sostanziale aumento dell'investimento necessario alla stesura di una linea guida (Atkins 2004).

In attesa che questa elaborazione porti a un metodo condiviso, e in analogia con il manuale PNLG del 2002, riteniamo utile definire alcuni punti fermi dei sistemi di grading sia per quanto riguarda la qualità delle prove (che verrà chiamata in italiano «livello delle prove», derivandolo dal termine di riferimento inglese *level of evidence*) sia per quanto riguarda la forza delle raccomandazioni che da esse possono essere derivate (in inglese il termine di riferimento è *strength of recommendation*).

Il livello di prova (LdP) si riferisce alla probabilità che le conoscenze alla base delle raccomandazioni siano derivate da studi pianificati e condotti in modo tale da produrre informazioni valide e prive di errori sistematici; ha quindi a che fare con il grado di confidenza che si ha nelle conoscenze che hanno condotto alla raccomandazione. **La forza della raccomandazione** (FdR) si riferisce invece alla probabilità che l'applicazione nella pratica di una raccomandazione determini un miglioramento dello stato di salute della popolazione, obiettivo cui la raccomandazione è rivolta. Il manuale PNLG introduceva un altro criterio, **la rilevanza delle decisioni** (RdD), che si riferisce all'entità delle conseguenze prevedibili. In altre parole, l'LdP dipende dalla confidenza che si può avere sull'effetto di un intervento, mentre la FdR riguarda il grado di convinzione con cui si ritiene che una certa raccomandazione debba essere implementata. È chiaro quindi che il livello di prova condiziona, con intensità variabile a seconda del tipo e complessità del quesito, ma non determina completamente la FdR. Quest'ultima, a sua volta, dipenderà dall'LdP, e da quanto maggiore è la specificità del quesito e l'applicabilità diretta del disegno di studio del trial controllato e randomizzato (essendo il trial il livello apicale dei differenti sistemi di grading degli LdP).

Complessivamente, quindi, l'LdP dipende da:

- l'appropriatezza del disegno di studio utilizzato (qualità metodologica);
- il rigore e la qualità di conduzione e analisi dello studio (controllo dell'errore sistematico);
- la precisione dell'effetto misurato (controllo dell'errore casuale);
- l'appropriatezza e la rilevanza degli indicatori di esito utilizzati (rilevanza clinico-epidemiologica).

La FdR invece dipende da:

- la valutazione complessiva dell'LdP secondo quanto sopra definito;
- la valutazione dell'applicabilità dello specifico comportamento contenuto nella raccomandazione;
- il tipo di difficoltà e le barriere previste nell'applicazione della raccomandazione;
- l'accettabilità culturale e sociale della specifica raccomandazione, e quindi i valori e le preferenze.

La maggior parte dei sistemi di classificazione dei livelli di prova deriva la FdR dall'LdP. Si va da sistemi di classificazione che fanno discendere in modo rigido la FdR dal relativo LdP (vedi il sistema di grading dell'LdP e delle FdR utilizzato nelle linee guida scozzesi del SIGN), a sistemi nei quali, pur tenendo conto dell'LdP, la deter-

minazione della FdR viene modulata considerando la complessità del quesito, l'effettiva possibilità di condurre studi randomizzati sull'argomento e la percezione socio-culturale del problema (Weightman 2005). Il sistema GRADE mostra l'indubbio vantaggio di esplicitare e rendere trasparenti questi diversi fattori che stanno alla base della FdR, assegnando a ognuno anche un proprio peso (Guyatt 2008a). Assumono una posizione intermedia i sistemi di grading che affiancano alla considerazione della validità a priori del disegno di studio, la valutazione della precisione dei risultati e della omogeneità (consistenza) tra risultati di studi diversi. Il manuale PNLG del 2002 presenta in appendice i sistemi di grading messi a punto da diversi organismi nazionali e internazionali fino alla sua pubblicazione (PNLG 2002).

Nell'ambito dell'elaborazione di linee guida di prevenzione, adottando una classificazione semplificata derivata da quella in uso all'SNLG, il sistema di *grading* suggerito è riportato nella tabella 10.

Forza della raccomandazione	
A	L'esecuzione di quella particolare procedura o intervento è fortemente raccomandata. Indica una particolare raccomandazione sostenuta da prove scientifiche di buona qualità, anche se non necessariamente di tipo I o II
B	Si nutrono dei dubbi sul fatto che quella particolare procedura o intervento debba sempre essere raccomandata, ma si ritiene che la sua esecuzione debba essere attentamente considerata
C	Esiste una sostanziale incertezza a favore o contro la raccomandazione di eseguire la procedura o l'intervento

Tabella 10. Sistema di grading suggerito per l'elaborazione delle linee guida di prevenzione

5. Gli indicatori per la valutazione

La produzione e l'uso di una linea guida di pratica medica ha l'obiettivo esplicito di migliorare l'attività professionale e la salute della popolazione. La letteratura scientifica presenta numerosi esempi di questo effetto (Bo 2007). Ogni linea guida dovrebbe prevedere quindi, a seguito di ogni raccomandazione, una proposta di indicatore di processo, di output e di outcome che permetta l'istituzione di un sistema di monitoraggio e di valutazione.

La scelta degli indicatori dovrebbe essere effettuata all'interno di quelli già routinariamente collezionati da parte del Sistema sanitario, per evitare di dover costituire un sistema informativo *ad hoc*. Se ciò non viene ritenuto possibile, la linea guida dovrebbe suggerire anche la costruzione dello strumento di raccolta delle informazioni mancanti nei sistemi correnti, se necessario distinguendo il livello nazionale da quello di regione e di azienda sanitaria. Il flusso informativo dovrebbe entrare al più presto nel debito informativo delle aziende sanitarie nei confronti della regione.

5.1 La valutazione d'impatto

L'applicabilità nella pratica delle linee guida e l'impatto sull'organizzazione dei servizi devono essere sottoposti a una verifica a partire dalla valutazione degli indicatori di monitoraggio utilizzati. In questo modo si misurano gli effetti sul campo in termini di accettabilità e praticabilità da parte di operatori e pazienti, oltre che di compatibilità con gli assetti organizzativi e gestionali dei servizi.

Già nella fase di produzione del documento devono essere identificate le raccomandazioni principali e definiti appropriati indicatori di monitoraggio di queste ultime. Questi indicatori devono riguardare aspetti chiave del processo di sanità pubblica.

Insieme all'impatto organizzativo, e quindi al monitoraggio dell'applicazione, è da prevedere un sistema di valutazione dell'impatto sulla salute della popolazione. Questo verrà costruito al livello appropriato al tipo di intervento raccomandato: nazionale se si tratta di interventi implementabili a quel livello (per esempio variazioni di prezzo di prodotti a rischio per la salute), oppure locale, quando si tratta di interventi implementabili anche a livello di azienda sanitaria. Sulla selezione e la metodologia relativa a indicatori che vanno a misurare l'impatto di interventi di sanità pubblica a livello di popolazione, il dibattito si sta orientando verso sistemi di indicatori a supporto delle decisioni in sanità pubblica, come, per esempio, le misure di impatto sulla popolazione (Population Impact Measures, PIMs) che hanno il vantaggio di fornire misure di rischio/beneficio per le popolazioni interessate, e rigorose e praticabili per i policy-maker, anche a livello locale (Torun 2009).

La possibilità di realizzare un processo continuo di audit e di restituzione delle informazioni relative alla effettiva adozione delle raccomandazioni è strettamente funzionale alla disponibilità di un buon sistema informativo corrente sulle attività e i profili di attività. Questo rappresenta nel panorama italiano un ostacolo maggiore alle attività di audit ed è verosimilmente un fattore di ostacolo importante anche all'implementazione di linee guida.

6. Il conflitto d'interesse nella tutela della salute pubblica

È ormai abitudine ricercare, nella valutazione di un articolo scientifico o di una linea guida clinica, la dichiarazione di possibili conflitti di interesse da parte di uno o più autori. La pratica di pubblicazione chiede, infatti, di dichiarare propri eventuali legami di natura economica con finanziatori che potrebbero avere l'interesse a modificare o anche solo a dare un'interpretazione parziale dei risultati.

Nel campo della sanità pubblica l'identificazione di conflitti di interessi è più complessa; accanto a interventi che possono presentare tipologie di conflitti simili a quelli presenti nella valutazione di farmaci, come per esempio per linee guida sugli approcci vaccinali, vi sono interventi che non prevedono nessun coinvolgimento diretto di interessi economici, come per esempio la prevenzione dell'uso di droghe illegali. In mezzo ci sono situazioni che presentano il coinvolgimento di interessi più indiretti, come per esempio gli interventi contro il fumo o l'alcool, oppure quelli sulla nutrizione.

Oltre a questo tipo di conflitto di interesse "classico", è importante identificare altri tipi di conflitto che per esempio possono coinvolgere materiali o interventi di prevenzione posti sul mercato, oppure istituzioni specifiche che operano nel campo con interventi propri.

I membri dei gruppi di lavoro dovranno costruire una scheda di dichiarazione di conflitto di interesse specifica per ogni linea guida, che tenga conto delle specificità citate sopra. Tutti i membri dei gruppi coinvolti nell'elaborazione della linea guida, nel coordinamento, nel panel e nel gruppo allargato, dovranno compilare tale modulo di dichiarazione di conflitti di interesse, e le informazioni relative verranno raccolte e pubblicate nel documento finale.

7. La disseminazione

Il processo di disseminazione di una linea guida è cruciale al fine di promuoverne l'adozione. Questo si compone di due diverse tappe: la produzione di materiali adeguati e la formazione.

7.1 Materiali

Per quanto riguarda i materiali, l'esperienza dei lavori clinici suggerisce diversi strumenti in grado di esaltare di volta in volta l'aspetto scientifico, quello formativo-esplorativo e la semplicità di utilizzo. A questo fine si propongono i seguenti strumenti.

- **Linea guida completa:** la pubblicazione di un volume di testo di concezione tradizionale appare indispensabile. Questa versione, infatti, non contiene solo le raccomandazioni e i suggerimenti delle linee guida, ma anche la presentazione del processo decisionale da cui esse sono scaturite e la necessaria bibliografia. Gli aspetti culturali e formativi, che tradizionalmente giustificano questa modalità di presentazione, fanno delle linee guida un vero e proprio testo di studio aggiornato. Per essere facilmente consultato è opportuno che non superi le 60 pagine.
- **Quick reference guide:** si tratta di un formato adottato da molti istituti internazionali di sviluppo di linee guida al fine di mettere a disposizione dei professionisti uno strumento semplice, di accesso rapido e di facile portabilità. Il formato è quello del pieghevole, utile per estrapolare alcune raccomandazioni, indirizzate a specifici utilizzatori e per creare del materiale di informazione sanitaria da distribuire a cittadini o pazienti. Un modo efficace di rappresentare la complessità del flusso decisionale è il diagramma di flusso o flow-chart, una serie di affermazioni a cascata, che si susseguono secondo un filo logico, esprimendo un algoritmo decisionale basato su variabili categoriche, del tipo sì/no, vero/falso, lieve/medio/grave. Graficamente esse possono anche essere riportate in un poster da appendere in studio oppure in fogli plastificati, da tenere sulla scrivania.
- **Versione online della linea guida completa,** con i capitoli indicizzati in modo da rendere rapido e semplice l'accesso ai singoli contenuti.
- **Presentazione digitale della linea guida:** questa può aver un supporto fisico, come CD-ROM o DVD, oppure essere basata in rete. Può contenere i testi completi, le versioni ridotte, eventuali filmati, registrazioni audio, animazioni grafiche, biblioteche di articoli e altro. Le raccomandazioni potrebbero essere presentate come set di diapositive. La presentazione digitale, però, rispetto a quella cartacea, consente di:
 - rappresentare una iper flow-chart che, come un ipertesto, può essere navigato, personalizzato e può rimandare, ove possibile, alle diverse versioni della linea gui-

da (per esempio, quella per il medico e quella per il paziente), a seconda della tipologia di utente collegato;

- rappresentare in modo compatto la forza della raccomandazione e la rilevanza delle decisioni (per esempio usando simboli convenuti, con colori diversi, collegati a riferimenti bibliografici o a stralci di testo);
- integrare dizionari di termini standardizzati o glossari.

Oltre ai vantaggi sopra indicati, il formalismo informatizzato facilita la distribuzione e l'utilizzo della linea guida poiché consente, laddove esiste una cartella clinica elettronica, di collegarla ai dati del paziente, e di offrire quindi, in tempo reale, un supporto decisionale all'utente.

7.2 Formazione

La formazione è uno strumento importante di disseminazione. In tempi di crediti formativi e di restrizione del budget per la formazione, i corsi di formazione a distanza (FAD), con attribuzione di crediti formativi ECM, sono quelli senza dubbio più indicati. Un esempio di un corso FAD adattato alla disseminazione di una linea guida sulla cessazione del tabagismo è accessibile all'indirizzo www.saepe.it (è necessaria un'iscrizione preliminare sul sito web www.ecmpiemonte.it).

8. L'implementazione

Le linee guida sono uno degli interventi della politica del Servizio Sanitario Nazionale per la promozione dell'efficacia e dell'appropriatezza nella pratica di sanità pubblica.

L'implementazione consiste nell'introdurre nella pratica corrente i cambiamenti di comportamento proposti per mezzo delle raccomandazioni, utilizzando strategie di intervento appropriate. Un obiettivo non secondario dell'implementazione è quello di superare barriere di tipo strutturale e organizzativo che sono spesso fonte di disuguaglianze sociali e geografiche nella salute.

Implementazione significa anche agire attivamente rispetto alla soluzione di un problema; le esperienze relative all'implementazione delle linee guida cliniche indicano che non basta semplicemente presentare e diffondere le linee guida; è necessario adottare strategie attive per promuovere il loro ingresso nella pratica e per radicarle in un cambiamento culturale complessivo.

L'esperienza suggerisce che i seguenti elementi sono indispensabili al successo dell'implementazione di una linea guida:

- il coinvolgimento fin dalle fasi precoci di elaborazione della linea guida di rappresentanti di tutti i professionisti e decision-maker che possono avere un ruolo nell'implementazione. Questo è lo spirito della costruzione del gruppo allargato di consultazione e del suo coinvolgimento in tutte le fasi dell'elaborazione del documento (vedi paragrafo. 4.1 *Requisiti essenziali* a pag. 16);
- un mandato formalizzato di implementazione da parte dell'istituzione più vicina del sistema, come per esempio la definizione di obiettivi per i Direttori generali di azienda sanitaria, o l'inserimento in programmi nazionali, quali per esempio quelli Ccm;
- la formazione di un gruppo multidisciplinare rappresentativo delle professionalità coinvolte nell'applicazione della linea guida che, al livello adeguato del sistema (nazionale, regionale o locale), sia responsabile della sua implementazione;
- la pubblicizzazione dei progressi nell'implementazione e dei risultati della valutazione di impatto. Il monitoraggio della pratica corrente permette di identificare fattori ostacolanti o favorenti l'implementazione e incoraggiare la disseminazione.

Il ruolo e l'importanza relativa degli elementi elencati varia ovviamente a seconda degli argomenti della linea guida. Si raccomanda pertanto che ogni linea guida contenga un capitolo specifico di proposte di implementazione operative e realistiche; un esempio di proposte di implementazione è quello della Linee guida clinico-organizzative per la cessazione del tabagismo della Regione Piemonte (Faggiano 2007).

9. L'aggiornamento

Ogni documento di raccomandazioni dovrebbe specificare una data entro la quale il documento stesso dovrebbe essere oggetto di ulteriore analisi e valutazione alla luce di informazioni divenute nel frattempo disponibili. Questa data dipende dalla velocità del processo di produzione di nuove evidenze, che è specifico di ogni settore, e dall'aspettativa di risultato di studi in corso al momento della chiusura della linea guida. Oltre a incorporare eventuali nuove informazioni, il processo periodico di revisione potrebbe permettere l'aggiunta di sezioni precedentemente non considerate e la modifica di specifici indicatori di monitoraggio e, infine, l'ulteriore rifinitura di un aspetto problematico e controverso del documento.

Nel momento di affrontare la revisione di un documento di raccomandazioni possono verificarsi tre possibilità:

- il documento è diventato del tutto obsoleto e deve essere completamente riconsiderato;
- il documento necessita solo di minori aggiustamenti che possono essere apportati dai responsabili dell'iniziale produzione con la sola consultazione esterna degli esperti originariamente coinvolti, ed eventualmente pubblicato solo come *addendum*;
- il problema che aveva originato la produzione di raccomandazioni si sia di fatto risolto e che pertanto non vi sia più necessità dell'implementazione di specifiche raccomandazioni.

10. Allegati

Allegato 1. Check-list per la valutazione critica di revisioni sistematiche (Health-Evidence)

Per l'inclusione nel database Health-Evidence, la valutazione si attua in due fasi: nella prima viene decisa l'inclusione della revisione, nella seconda viene valutata. La prima fase può essere svolta anche senza la formalizzazione suggerita dalla check-list, mentre per la seconda fase è essenziale una sua accurata applicazione.

Vi ricordiamo che molti articoli includono nel titolo i termini «revisione sistematica» o «metanalisi». Altri contengono questi termini all'interno dell'articolo. In ogni caso le revisioni sistematiche devono presentare la strategia di ricerca; quando invece manca una strategia di ricerca, l'articolo non può essere classificato come revisione. È importante che l'articolo includa delle tabelle dettagliate sulle caratteristiche degli studi e sugli outcome. Una strategia di ricerca descritta in maniera chiara dovrebbe essere la caratteristica delle revisione stessa e nella maggior parte dei casi la si ritrova nei metodi. Idealmente il lettore dovrebbe essere in grado di replicare la ricerca utilizzando le informazioni fornite dall'autore.

Check-list di inclusione

Criteri di rilevanza	Si	No
Si tratta di una revisione?		
La revisione include tabelle dettagliate sulle caratteristiche degli studi e sugli outcome?		
La revisione tratta rilevanti problemi di sanità pubblica o di promozione sanitaria?		
La revisione si occupa di promozione, protezione o prevenzione (3P)?		
La revisione ha come argomento principale l'efficacia di un intervento, di un programma, di un servizio o di una politica?		
La revisione prende in esame l'efficacia sul campo (effectiveness) di uno specifico intervento di sanità pubblica, che potrebbe essere un programma, un servizio o una policy?		
La revisione esamina altri aspetti quali il nesso di causalità, la correlazione e il costo-efficacia?		
Vengono individuati in modo chiaro gli outcome della revisione?		
Gli outcome di ciascun studio, sia positivi sia negativi, vengono descritti in modo chiaro?		
Viene descritta la strategia di ricerca utilizzata?		
Viene descritta la decisione del revisore?		
La revisione è inclusa nel registro? (Se sì, compilare la scheda di valutazione di qualità)		
Ci sono referenze rilevanti? (Se sì, segnare le referenze in bibliografia)		

Check-list di valutazione (<http://health-evidence.ca>)

Criterion	Yes	No
1. Did the authors have a clearly focused question [population, intervention (strategy), and outcome(s)]?		
2. Were appropriate inclusion criteria used to select primary studies?		
3. Did the authors describe a search strategy that was comprehensive? (circle all strategies used) health databases handsearching psychological databases key informants social science databases reference lists educational databases unpublished other _____		
4. Did search strategy cover an adequate number of years?		
For questions 5, 6, and 8, please choose the column relating to the appropriate methodology. Strike a line through the column that does not apply.		
5. <i>Quantitative reviews:</i> Did the authors describe the level of evidence in the primary studies included in the review? Level I (RCTs only) Level II (non-randomized, cohort, case-control studies) Level III (uncontrolled studies)	5. <i>Qualitative reviews:</i> Do the authors provide a clear description of the range of methods in each of the primary studies included in the review?	
6. <i>Quantitative reviews:</i> Did the review assess the methodological quality of the primary studies, including: (minimum requirement: 4/7 of the following) • research design • study sample • participation rates • sources of bias (confounders, respondent bias) • data collection (measurement of independent/dependent variables) • follow-up/attrition rates • data analysis	6. <i>Qualitative reviews:</i> Did the review assess the methodological quality of the primary studies, including: (minimum requirement: 4/7 of the following) • suitability of methodology/paradigm to the research question • sampling (selection of participants/settings/documentation) • clear description of context, data collection and data analysis • rigor i) audit trail ii) some coding by 2 or more cover, if appropriate iii) deviant case analysis (negative cases) iv) respondent validation (member checking) • triangulation • reflexivity (researcher and research process) • relevance (credibility, consistency, applicability, transferability)	
7. Are the results of the review transparent?		
8. <i>Quantitative review:</i> Was it appropriate to combine the findings of results across studies?	8. <i>Qualitative review:</i> Is there a description of how reviewers determined results were similar enough across studies to compare or combine them?	
9. Were appropriate methods used for combining or comparing results across studies?		
10. Do the data support the author's interpretation?		
QUALITY ASSESSMENT RATING (circle one) (total score 7-10) STRONG (total score 5-6) MODERATE TOTAL SCORE: _____ (total score 4 or less) WEAK		

Allegato 2. Check-list per la valutazione critica degli studi randomizzati e controllati (www.phru.nhs.uk/pages/phd/casp.htm)

Critical Appraisal Skills Programme (CASP) - making sense of evidence 10 questions to help you make sense of randomised controlled trials

Screening Questions

1. Did the study ask a clearly focused question?

Yes Can't tell No

Consider if the question is 'focused' in terms of:

- the population studied
- the intervention given
- the outcomes considered

2. Was this a randomised controlled trial (RCT) and was it appropriately so?

Yes Can't tell No

Consider:

- why this study was carried out as an RCT
- if this was the right research approach for the question being asked

Is it worth continuing?

Detailed questions

3. Were participants appropriately allocated to intervention and control groups?

Yes Can't tell No

Consider:

- how participants were allocated to intervention and control groups. Was the process truly random?
- whether the method of allocation was described. Was a method used to balance the randomisation, e.g. stratification?
- how the randomisation schedule was generated and how a participant was allocated to a study group
- if the groups were well balanced. Are any differences between the groups at entry to the trial reported?
- if there were differences reported that might have explained any outcome(s) (confounding)

4. Were participants, staff and study personnel 'blind' to participants' study group?

Yes Can't tell No

Consider:

- the fact that blinding is not always possible
- if every effort was made to achieve blinding
- if you think it matters in this study
- the fact that we are looking for 'observer bias'

5. Were all of the participants who entered the trial accounted for at its conclusion?

Yes Can't tell No

Consider:

- if any intervention-group participants got a control-group option or vice versa
- if all participants were followed up in each study group (was there loss-to-follow-up?)
- if all the participants' outcomes were analysed by the groups to which they were originally allocated (intention-to-treat analysis)
- what additional information would you liked to have seen to make you feel better about this





6. Were the participants in all groups followed up and data collected in the same way?

Yes Can't tell No

Consider:

– if, for example, they were reviewed at the same time intervals and if they received the same amount of attention from researchers and health workers. Any differences may introduce performance bias.

7. Did the study have enough participants to tell minimise the play of chance?

Yes Can't tell No

Consider:

– if there is a power calculation. This will estimate how many participants are needed to be reasonably sure of finding something important (if it really exists and for a given level of uncertainty about the final result).

8. How are the results presented and what is the main result?

Consider:

- if, for example, the results are presented as a proportion of people experiencing an outcome, such as risks, or as a measurement, such as mean or median differences, or as survival curves and hazards
- how large this size of result is and how meaningful it is
- how you would sum up the bottom-line result of the trial in one sentence

9. How precise are these results?

Consider:

- if the result is precise enough to make a decision
- if a confidence interval were reported. Would your decision about whether or not to use this intervention be the same at the upper confidence limit as at the lower confidence limit?
- if a p-value is reported where confidence intervals are unavailable

10. Were all important outcomes considered so the results can be applied?

Yes Can't tell No

Consider whether:

- the people included in the trial could be different from your population in ways that would produce different results
- your local setting differs much from that of the trial
- you can provide the same treatment in your setting

Consider outcomes from the point of view of the:

- individual
- policy maker and professionals
- family/carers
- wider community

Consider whether:

- any benefit reported outweighs any harm and/or cost. If this information is not reported can it be filled in from elsewhere?
- policy or practice should change as a result

Procedure per migliorare la validità interna ed esterna dei *cluster randomised controlled trial* (Eldridge 2008)

Internal validity

Accounting for clustering in sample size calculation

Accounting for clustering in analysis

Protected against recruitment/identification bias when identifying/recruiting patients

Blinding of individual participants to allocation status

Assessment of primary outcome blind to allocation status

External validity

Full information on number of clusters approached, recruited, and analysed

Comparison of characteristics of clusters recruited and those not recruited

Discussion of cluster generalisability

Discussion of how the clusters analysed might differ from other clusters

Some information about acceptability and/or feasibility

Allegato 3. Strumento per la valutazione critica delle *interrupted time series design* (ITS)

Quality criteria for ITS designs (Ramsay 2003)

1. Intervention occurred independently of other changes over time

DONE The intervention occurred independently of other changes over time

NOT CLEAR Not specified (will be treated as NOT DONE if information cannot be obtained from the authors)

NOT DONE Reported that intervention was not independent of other changes in time

2. Intervention was unlikely to affect data collection

DONE Reported that intervention itself was unlikely to affect data collection (for example, sources and methods of data collection were the same before and after the intervention)

NOT CLEAR Not specified (treated as NOT DONE if information cannot be obtained from the authors)

NOT DONE Intervention itself was likely to affect data collection (for example, any change in source or method of data collection reported)

3. The primary outcome was assessed blindly or was measured objectively

DONE Stated explicitly the primary outcome variables were assessed blindly or outcome variables are objective e.g., length of hospital stay, drug levels assessed by a standardized test

NOT CLEAR Not specified (treated as NOT DONE if information cannot be obtained from the authors)

NOT DONE Outcomes were not assessed blindly

4. The primary outcome was reliable or was measured objectively

DONE Two or more raters with agreement $\geq 90\%$ or kappa ≤ 0.8 or outcome assessment is objective, e.g., length of hospital stay, drug levels assessed by a standardized test

NOT CLEAR Reliability not reported for outcome measures obtained by chart extraction or collected by an individual (will be treated as NOT DONE if information cannot be obtained from the authors)

NOT DONE Two or more raters with agreement $< 90\%$ or kappa < 0.8

5. The composition of the data set at each time point covered at least 80% of the total number of participants in the study

DONE Data set covers 80-100% of total number of participants or episodes of care in the study

NOT CLEAR Not specified (will be treated as NOT DONE if information cannot be obtained from the authors)

NOT DONE Data set covers less than 80% of the total number of participants or episodes of care in the study

6. The shape of the intervention effect was prespecified

DONE A rational explanation for the shape of intervention effect was given by the author(s)

NOT CLEAR Not specified

NOT DONE Any of the conditions above are not met

7. A rationale for the number and spacing of data points was described

DONE Rationale for the number of points stated (e.g., monthly data for 12 months postintervention was used because the anticipated effect was expected to decay) or sample size calculation performed

NOT CLEAR Not specified

NOT DONE Any of the conditions above are not met

8. The study was analyzed appropriately using time series techniques

DONE ARIMA models were used or time series regression models were used to analyze the data and serial correlation was adjusted/tested for

NOT CLEAR Not specified

NOT DONE Any of the conditions above are not met

ITS: interrupted time series; ARIMA: autoregressive integrated moving average

Altri criteri di qualità delle ITS

Caratteristiche principali dell'intervento

- Definizione esatta dell'intervento.
- Tempo puntuale di applicazione dell'intervento.

Identificazione di trend secolari

- L'intervento è indipendente da altri cambiamenti.
- Misurazioni prima e dopo l'intervento (almeno tre + tre).
- Applicazione di test per i trend.

Protezione dal detection bias

- L'intervento non ha potuto modificare la raccolta delle informazioni.
- La misura dell'outcome primario è in cieco o è costituita da dati oggettivi.
- Se la misura dell'outcome primario non è oggettiva, deve essere rilevata da due misuratori che raggiungano una concordanza pari a $k \geq 0.8$ o di almeno il 90%.
- Completezza del data set (deve coprire almeno l'80% dei partecipanti).

Bibliografia

- Agency for Healthcare Research and Quality (AHRQ). Treating Tobacco Use and Dependence: 2008. <http://www.ahrq.gov/path/tobacco.htm> (ultimo accesso: 10 maggio 2010).
- Associazione Italiana di Epidemiologia (AIE). Trattamento dei rifiuti e salute. Posizione dell'Associazione Italiana di Epidemiologia. *Epidemiol Prev* 2008; 32: 183-7.
- Atkins D, Best D, Briss PA et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328: 1490-7.
- Bero L, Grilli R, Grimshaw JM, et al. The Cochrane Effective Practice and Organization of Care Group (EPOC) Module. In: *The Cochrane Library*. Oxford, Update Software, 2002, Issue 1.
- Bo S, Valpreda S, Scaglione L, Boscolo D, Piobbici M, Bo M, Ciccone G. Implementing hospital guidelines improves warfarin use in non-valvular atrial fibrillation: a before-after study. *BMC Public Health* 2007; 7: 203.
- Centro nazionale per la prevenzione e il controllo delle malattie (Ccm). Diffusione della "Ebp" nei dipartimenti di Prevenzione del SSN. Ultimo aggiornamento: 2008. <http://www.ccm-network.it/>.
- Centro nazionale per la prevenzione e il controllo delle malattie (Ccm). Guadagnare salute. Ultimo aggiornamento: 28 Gennaio 2010. <http://www.ccm-network.it/>.
- Cook DT, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Chicago, Rand McNally, 1979.
- D'Argenio P. Critiche allo studio «Trattamento dei rifiuti in Campania: impatto sulla salute umana». *Epidemiol Prev* 2008; 32: 189-92.
- Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F et al. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003; 7 (27): iii-x, 1-173.
- Devereaux PJ, Choi PTL, Lacchetti C, Weaver B, Schünemann HJ, Haines T et al. A systematic review and meta-analysis of studies comparing mortality rates of private for profit and private for non profit hospitals. *CMAJ* 2002; 166(11): 1399-406.
- Eccles M, Grimshaw J, Campbell M et al. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003; 12: 47-52.
- Egger M, Smith GD. Potentials and Promise. *BMJ* 1997; 315: 1371-4.
- Eldridge S, Ashby D, Bennett C, Wakelin M, Feder G. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008; 336(7649): 876-80.
- Evidence-Based Medicine (EBM) Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA* 1992; 268(17): 2420-5.
- Faggiano F, Gelormino E, Mathis F et al. Cessazione del fumo di tabacco. Linee Guida clinico-organizzative per la Regione Piemonte. Commissione Regionale Anti-tabacco. Quaderno n. 3, maggio 2007.
- Faggiano F, Gelormino E. Gli studi per la valutazione di impatto per la salute. In: Faggiano F, Donato F, Barbone F (eds). *Applicazioni di Epidemiologia per la Sanità Pubblica*. Centro Scientifico Editore, Torino, 2006: 231-76.
- Fineberg HV. Screening for HIV infection and public health policy. *Law Med Health Care* 1990; 18: 29-32.
- Grol R, Dalhuijsen J, Thomas J, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998; 317: 858-61.
- Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336: 924-6. (a)
- Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ et al. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008; 336: 995-8. (b)
- Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann et al. Going from evidence to recommendations. *BMJ* 2008; 336: 1049-51. (c)

- Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993; 270(21): 2598-601
- Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994; 271(1): 59-63.
- Habicht JP, Victora CG and Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol* 1999; 28: 10-8.
- Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. www.cochrane-handbook.org.
- Kingdon JW. *Agendas, alternatives and public policy*. Boston, Mass Little, Brown, 1984.
- Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J. Knowledge Transfer Study Group. How can research organizations more effectively transfer research knowledge to decision makers? *Milbank Q* 2003; 81(2): 221-48, 171-2.
- Lopez AD, Mathers CD, Ezzati M, Jamison DT, and Murray CJL. *Global Burden of Disease and Risk Factors*. New York, Oxford University Press, 2006.
- Main C, Mather L, Ogilvie D, Petticrew M, Sowden A, Thomas S et al. The inverse evidence law in tobacco control: findings from a systematic review of reviews. *J Epidemiol Community Health* 2005; 59(Suppl 1): A24.
- Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Clin Epidemiol* 2009; 62(10): 1006-12.
- Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, Marteau T. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998; 2(3): i-iv, 1-88.
- Murray DM, Varnell SP, Blitstein JL. Design and analysis of Group-randomized trials: a review of recent methodological developments. *Am J Public Health* 2004; 94: 423-32.
- National Institute for Health and Clinical Excellence (NICE). *Methods for development of NICE public health guidance - Second edition*. Nice, NHS, 2009 (a).
- National Institute for Health and Clinical Excellence (NICE). *The Guidelines Manual*. NICE, NHS, 2009 (b).
- Ogilvie D, Hamilton V, Egan M, Petticrew M. Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *J Epidemiol Community Health* 2005; 59: 804-8.
- Oxman AD, Cook DJ, Guyatt GH. Users' guide to the medical literature. VI. How to use an overview. *JAMA* 1994; 272(17): 1367-71.
- Piano Nazionale Linee Guida (PNLG). *Come produrre, diffondere e aggiornare raccomandazioni per la pratica clinica - Manuale metodologico*. PNLG, Istituto Superiore di Sanità e Agenzia per i Servizi Sanitari Regionali, Roma, 2002.
- Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. Interrupted Time series designs in Health Technology Assessment: lessons from two systematic reviews of behavior change strategies. *Int J Technol Assess Health Care* 2003; 19(4): 613-23.
- Regonini G. *Capire le politiche pubbliche*. Bologna, Il Mulino, 2001.
- Rychetnik L, Hawe P, Waters E, Barratt A, Frommer M. A glossary for evidence based public health. *J Epidemiol Community Health* 2004; 58(7): 538-45.
- Sackett DL, Straus SE, Richardson WS et al. *Evidence Based Medicine. How to practice and teach EBM*. London, Churchill Livingstone 2000.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007; 7: 10.
- Swedish National Institute of Public Health

(SNIPH). Knowledge based Public Health Work, 2001.

The ADAPTE Collaboration (2009). The ADAPTE Process: Resource Toolkit for Guideline Adaptation. Version 2.0. www.g-i-n.net

The AGREE Collaboration. Development and validation of an International Appraisal Instrument for assessing the quality of clinical practice guidelines: the AGREE Project. *Qual Saf Health Care* 2003; 12: 18-23.

Torun P, Heller RF, Verma A. Potential population impact of changes in heroin treatment and smoking prevalence rates: using Population Impact Measures. *Eur J Public Health* 2009; 19(1): 28-31.

van der Maas PJ. How summary measures of population health are affecting health agendas. *Bull World Health Organ* 2003; 81(5): 314.

Victora CG, Habicht J-P, Bryce J. Evidence-Based Public Health: Moving Beyond Randomized Trials. *Am J Public Health* 2004; 94: 400-5.

Weightman A, Ellis S, Cullum A et al. Grading evidence and recommendations for public health interventions: developing and piloting a framework. NHS Health Development Agency, 2005.

West SG, Duan N, Pequegnat W, Gaist P, Des Jarlais DC, Holtgrave D et al. Alternative to the randomized controlled trial. *Am J Public Health* 2008; 98(8): 1359-66.

Questo documento è consultabile anche nel sito internet <http://www.snlg-iss.it>

Il Sistema nazionale per le linee guida (SNLG)

In Italia, l'elaborazione di linee guida e di altri strumenti di indirizzo finalizzati al miglioramento della qualità dell'assistenza avviene all'interno del Sistema nazionale per le linee guida (SNLG).

La legislazione vigente propone l'adozione di linee guida come richiamo all'utilizzo efficiente ed efficace delle risorse disponibili e come miglioramento dell'appropriatezza delle prescrizioni.

Queste sono le finalità del SNLG con i compiti specifici di:

- produrre informazioni utili a indirizzare le decisioni degli operatori, clinici e non, verso una maggiore efficacia e appropriatezza, oltre che verso una maggiore efficienza nell'uso delle risorse;
- renderle facilmente accessibili;
- seguirne l'adozione esaminando le condizioni ottimali per la loro introduzione nella pratica;
- valutarne l'impatto organizzativo e di risultato.

Gli strumenti utilizzati per perseguire questi fini sono appunto linee guida clinico-organizzative, documenti derivanti da consensus conference, revisioni rapide di procedure e interventi, documenti implementativi e per la valutazione dei servizi.

Il Network italiano di Evidence-based Prevention (NIEbP)

Il Network Italiano di EBP si è sviluppato nel corso dell'ultimo decennio per affrontare anche in Italia il tema dell'efficacia degli interventi in sanità pubblica, ancora oggi troppo ancorati a pratiche obsolete e a visioni tramontate. Sulla base di queste istanze di rinnovamento il Centro nazionale per la prevenzione e il controllo delle malattie (Ccm) del Ministero della Salute ha promosso un primo progetto nel 2005, che poi fu ampliato.

È proprio nell'ambito di tale ampliamento che si colloca l'iniziativa di questo Manuale per produrre, diffondere e aggiornare linee guida per la salute pubblica, che sarà la base delle nuove linee guida di prevenzione che verranno elaborate in accordo con il SNLG.