

## Aspetti metodologici

**Peter A. Briss, MD, Stephanie Zaza, MD, MPH, Marguerite Pappaioanou, DVM, PhD, Jonathan Fielding, MD, MPH, MBA, Linda Wright-De Agüero, PhD, MPH, Benedict I. Truman, MD, MPH, David P. Hopkins, MD, MPH, Patricia Dolan Mullen, DrPH, Robert S. Thompson, MD, Steven H. Woolf, MD, MPH, Vilma G. Carande – Kulis, MS, PhD, Laurie Anderson, PhD, MPH, Alan R. Hinman, MD, MPH, David V. McQueen, ScD, Steven M. Teutsch, MD, MPH, Jeffrey R. Harris, MD, MPH, The Task Force on Community Preventive Services**

*Epidemiology Program Office (Briss, Zaza, Wright – De Agüero, Truman, Hopkins, Carande – Kulis, Anderson, Harris), Centers for Disease Control and Prevention (CDC), Atlanta, Georgia; Office of Global Health (Pappaioanou), CDC, Atlanta, Georgia; Los Angeles Department of Health Services (Fielding), Los Angeles, California; University of Texas-Houston School of Public Health (Dolan Mullen), Houston, Texas; Department of Preventive Care, Group Health Cooperative of Puget Sound (Thompson), Seattle, Washington; Medical College of Virginia (Woolf), Fairfax, Virginia; Task Force for Child Survival and Development (Hinman), Atlanta, Georgia; National Center for Chronic Disease Prevention and Health Promotion (McQueen), CDC, Atlanta, Georgia; Merck & Co., Inc. (Teutsch), West Point, Pennsylvania*

*Corrispondenza e altre richieste possono esser indirizzate a: Peter A. Briss, MD, Senior Scientist and Development Coordinator, Community Preventive Services Guide Development Activity, Epidemiology Program Office, MS K73, Centers for Disease Control and Prevention, 4770 Buford Highway MS-K73, Atlanta, GA 30341; E-mail: pxb5@cdc.gov.*

*Parte del presente contributo è stato pubblicato precedentemente in: Shefer A, Briss P, Rodewald L, et al. Improving immunization coverage rates: An evidence-based review of the literature. *Epidemiologic Reviews* 1999; 20: 96–142.*

### Premessa

Le revisioni sistematiche e le raccomandazioni basate su evidenze scientifiche stanno assumendo un'importanza crescente nel processo decisionale in ambito medico e sanitario. Negli ultimi 20 anni la mole di informazioni sui metodi scientifici di sintesi dei risultati sperimentali è andata aumentando in misura esponenziale. Alcuni metodi di revisione sistematica delle prove di efficacia dei servizi clinici di prevenzione e assistenza sanitaria possono tuttavia risultare inadeguati per valutare interventi a livello di popolazione. Inoltre i metodi per collegare le evidenze scientifiche alle raccomandazioni sono meno perfezionati rispetto a quelli utilizzati per la sintesi delle evidenze.

La *Guide to Community Preventive Services: Systematic Reviews and Evidence-Based Recommendations* (la Guida) si propone di valutare e formulare raccomandazioni relative a interventi di popolazione e nel settore della sanità pubblica. Il presente documento fornisce una panoramica sul procedimento adottato per effettuare revisioni sistematiche delle evidenze scientifiche e tradurle in raccomandazioni.

Scopo della Guida è quello di valutare le prove di efficacia, la loro applicabilità (ossia in che misura i dati disponibili relativi all'efficacia si possono considerare applicabili ad altre popolazioni e contesti), altri effetti prodotti dagli interventi (effetti collaterali degni di nota), l'impatto economico e gli ostacoli all'attuazione degli interventi.

Al fine di raccogliere e valutare le evidenze di supporto alle raccomandazioni è necessario: 1) formare gruppi di lavoro interdisciplinari incaricati di elaborare i singoli capitoli; 2) sviluppare un protocollo per l'organizzazione, classificazione, selezione e valutazione degli interventi esaminati in ciascun capitolo; 3) selezionare gli interventi oggetto di valutazione; 4) procedere alla ricerca sistematica e catalogazione delle prove; 5) valutare la qualità e sintetizzare il corpus di prove di efficacia; 6) tradurre il corpus di prove di efficacia in raccomandazioni; 7) prendere in esame le informazioni riguardanti evidenze diverse dalle prove di efficacia; 8) individuare e sintetizzare eventuali lacune nella ricerca.

La revisione sistematica e la formulazione di raccomandazioni suffragate da dati scientifici su interventi sanitari di popolazione rappresentano un compito estremamente arduo e un ambito metodologico in continua evoluzione. Il ricorso ad un approccio basato su evidenze per individuare e raccomandare interventi efficaci finalizzati a raggiungere specifici obiettivi di sanità pubblica può tuttavia consentire di limitare gli errori nelle modalità di raccolta e interpretazione dei dati, di mettere in luce lacune importanti nello stato attuale delle conoscenze, indirizzando così le ricerche successive, e aumentare la capacità da parte dei lettori della Guida di valutare se, dal proprio punto di vista, le raccomandazioni sono valide e prudenti. Tutti i vantaggi summenzionati potrebbero via via contribuire a raggiungere un maggiore consenso sulle strategie sanitarie più adeguate da adottare a livello di comunità e ad estenderne l'applicazione.

**Parole chiave:** community health services; decision making; evidencebased medicine; systematic reviews; methods; population-based interventions; practice guidelines; preventive health services; public health practice; Task Force

*Am J Prev Med* 2000; 18(1S): 35–43 © 2000 American Journal of Preventive Medicine

### Introduzione

Le revisioni sistematiche e le raccomandazioni basate su evidenze scientifiche stanno assumendo un'importanza crescente nel processo decisionale in ambito medico e sanitario. Negli ultimi 20 anni la mole di informazioni relative ai metodi scientifici adottati per la sintesi dei risultati sperimentali è andata aumentando in misura esponenziale.<sup>1,2</sup> Gran parte dei lavori disponibili relativi a sintesi di studi su problematiche sanitarie, tuttavia, si incentra sui servizi clinici di prevenzione (ossia sulle prassi di prevenzione applicate a condizioni bersaglio presso individui asintomatici)<sup>3</sup> e sull'assistenza medica.<sup>2,4,5</sup> Inoltre, benché anche nell'ambito del collegamento delle evidenze scientifiche a raccomandazioni pratiche sia stata acquisita una certa esperienza,<sup>3,6</sup> essa risulta senz'altro più limitata. Pertanto, la *Task Force* si sta adoperando al fine di perfezionare i metodi di conduzione delle revisioni sistematiche e di formulazione di raccomandazioni basate su evidenze scientifiche relative a interventi di popolazione nel settore della sanità pubblica. Il presente documento intende tracciare una panoramica del procedimento attualmente adottato per valutare le evidenze e per tradurle in raccomandazioni contenute nella Guida. Un esempio di revisione e delle raccomandazioni che ne sono state tratte utilizzando questi metodi viene illustrato altrove in questo numero della rivista.<sup>7,8</sup>

### Metodi di elaborazione delle revisioni e delle raccomandazioni

La *Task Force* ha stabilito che le raccomandazioni contenute nella Guida devono essere basate su revisioni sistematiche delle evidenze finalizzate a mostrare il nesso fra l'intervento e determinati esiti e su un procedimento esplicito di traduzione delle evidenze in raccomandazioni. Nella Guida il termine «evidenze» viene utilizzato per indicare:

- informazioni considerate adeguate per rispondere ai quesiti relativi all'efficacia degli interventi;
- applicabilità delle prove di efficacia (ossia in che misura si ritiene che i dati relativi all'efficacia disponibili possano essere applicabili anche ad altre popolazioni e contesti);
- altri effetti prodotti dagli interventi (effetti collaterali, ivi compresi esiti rilevanti degli interventi non oggetto della valutazione di efficacia, sia che si tratti di effetti dannosi o benefici, previsti o impreveduti, di esiti sanitari o non sanitari);
- impatto economico;
- ostacoli riscontrati nell'attuazione degli interventi.

Le raccomandazioni contenute nella Guida sono basate fondamentalmente sulle prove di efficacia.

Le evidenze considerate in base alle finalità della Guida sono ottenute generalmente con metodi osservativi o sperimentali. I metodi di raccolta e valutazione delle evidenze giudicati accettabili possono variare in funzione dell'argomento preso in esame. Per esempio, nella procedura adottata nella Guida, per rispondere a quesiti in merito all'efficacia o meno degli interventi vengono utilizzati dati tratti da studi comparativi, in cui

vengono posti a confronto gli esiti registrati in un gruppo sottoposto all'intervento con gli esiti ottenuti in un gruppo concomitante o storico non esposto o esposto in misura minore. Studi non comparativi possono tuttavia essere utilizzati per evidenziare gli ostacoli all'attuazione e per raccogliere dati economici. Per l'ulteriore esame di altri metodi di valutazione di vari tipi di evidenze si rinvia a *From the Division of Prevention Research and Analytic Methods*, a una trattazione successiva più dettagliata di seguito nel presente capitolo e ad altri contributi pubblicati in questo supplemento.

Al fine di raccogliere e valutare le evidenze e di tradurre tali evidenze in raccomandazioni la *Task Force* ha ritenuto necessario:

- formare gruppi di lavoro interdisciplinari incaricati di elaborare i singoli capitoli;
- sviluppare un protocollo per l'organizzazione, classificazione e selezione degli interventi valutati in ciascun capitolo;
- selezionare gli interventi oggetto di valutazione;
- procedere alla ricerca sistematica e catalogazione delle prove;
- valutare la qualità e sintetizzare il corpus di prove di efficacia;
- tradurre il corpus di prove di efficacia in raccomandazioni;
- prendere in esame evidenze diverse dalle prove di efficacia;
- individuare e sintetizzare eventuali lacune nella ricerca.

### Composizione dei gruppi di lavoro incaricati della stesura dei capitoli

Data l'ampia portata e il carattere multidisciplinare di numerose problematiche relative alla sanità pubblica, la *Task Force* si avvale di gruppi di lavoro rappresentativi di diverse prospettive per l'elaborazione dei singoli capitoli. La supervisione della stesura di ciascun capitolo è affidata a circa 4-10 esperti dotati di competenze specifiche di tipo metodologico o attinenti allo specifico argomento affrontato. Altri 15-20 specialisti, fra i quali operatori sanitari, forniscono la propria consulenza per l'elaborazione dei capitoli. La vasta esperienza dei membri dei gruppi di lavoro è un presupposto di fondamentale importanza per varie ragioni:

- per garantire la validità e completezza dell'impostazione teorica adottata nel capitolo;
- per assicurare un grado adeguato di conoscenza ed esperienza nell'esaminare numerosi tipi di interventi e aumentare l'utilità delle revisioni degli interventi selezionati in via definitiva;
- per ridurre la probabilità che importanti informazioni vengano ignorate;
- per limitare le possibilità di errori o bias nell'interpretazione dei dati acquisiti.

### Elaborazione di un protocollo per l'organizzazione, la classificazione, la selezione e la valutazione degli interventi esaminati in ciascun capitolo

Data l'ampiezza di ciascun capitolo della Guida, è necessario che il gruppo di lavoro incaricato della stesura di ogni singolo capitolo individui alcuni ambiti chiave sui quali concentrare l'attenzione. Un quadro logico è uno schema in cui viene trac-

ciata una catena di ipotetici rapporti causali fra fattori determinanti, esiti intermedi ed esiti sanitari. Il quadro logico viene utilizzato per stabilire i nessi fra: determinanti sociali, ambientali e biologici e i relativi esiti; punti strategici ai fini dell'azione; interventi che possono influire su tali punti. L'aspetto forse più importante è che i quadri logici forniscono al gruppo di lavoro incaricato della stesura di ciascun capitolo una struttura per descrivere gli interventi utilizzabili per conseguire specifici obiettivi nell'ambito della sanità pubblica e consente al gruppo stesso di stabilire quali delle opzioni disponibili saranno oggetto di revisione nel capitolo. Alcuni esempi di quadri logici sono stati forniti in altra sede.<sup>7,9,10</sup> Una volta selezionati gli interventi, per ciascuno viene sviluppato un quadro analitico dettagliato che traccia gli ipotetici nessi fra l'intervento e gli effetti sanitari e di altro tipo. I quadri analitici sono fondamentalmente piani analitici dettagliati che costituiscono porzioni di quadri logici più vasti. I quadri analitici delineano il piano di valutazione di ciascun intervento, orientando così la ricerca delle evidenze da parte del gruppo di lavoro. Quadri analoghi sono stati utilizzati anche per guidare altre revisioni sistematiche.<sup>11</sup>

Per poter essere considerate prove di efficacia secondo gli scopi perseguiti dalla Guida, le evidenze empiriche devono dimostrare che un intervento è in grado di migliorare gli esiti sanitari. Tale dimostrazione può essere diretta, come nel caso dell'effetto della fluorizzazione dell'acqua nel ridurre l'insorgenza della carie dentaria. Più spesso, però, la dimostrazione è indiretta: l'aumento del prezzo dei prodotti del tabacco, ad esempio, potrebbe determinare una riduzione dell'uso del tabacco che, a sua volta, può ridurre la morbilità e la mortalità. Nei casi in cui l'esistenza di un nesso fra esiti intermedi ed esiti sanitari sia stata ben documentata in altra sede (come ad esempio il nesso fra fumo e danni alla salute o fra vaccinazione e riduzione della malattia) si può rinviare a tali evidenze, mentre la ricerca di prove da parte della Guida si concentrerà esclusivamente sul rapporto fra gli interventi e gli esiti intermedi. L'adozione di un quadro logico consente di esplicitare scelte di questo tipo.

### Selezione degli interventi oggetto di valutazione

Un intervento è caratterizzato dai seguenti elementi: l'azione svolta, le modalità con cui è stata condotta, i destinatari e il luogo di esecuzione. Gli interventi possono inoltre essere costituiti da un'unica componente – ossia prevedere una sola attività – o da più componenti, ossia utilizzare più attività collegate fra loro. Data l'eterogeneità che caratterizza generalmente gli interventi attuati a livello di popolazione, il gruppo incaricato dell'elaborazione di ogni capitolo deve stabilire esplicitamente in che misura gli interventi verranno considerati parte dello stesso corpus di evidenze. Tali valutazioni dipendono dalle caratteristiche dell'intervento in questione, dal grado di approfondimento della letteratura disponibile, da considerazioni teoriche e di altro tipo.

Nel selezionare i tipi di intervento oggetto di valutazione nei vari capitoli, i gruppi di lavoro tengono conto dei seguenti aspetti:

- la capacità potenziale di ridurre l'entità del problema, ovvero malattie e lesioni;
- la capacità potenziale di aumentare comportamenti giovevoli per la salute e di ridurre comportamenti che nuocciono alla salute;
- la capacità potenziale di estendere l'attuazione di interventi efficaci non diffusi su larga scala;
- la capacità potenziale di favorire l'abbandono progressivo di interventi meno efficaci ma attuati su larga scala a favore di opzioni più efficaci o economicamente più vantaggiose;
- l'attuale livello di interesse presso i responsabili decisionali e gli erogatori di servizi sanitari.

La valutazione della quantità di letteratura disponibile non costituisce un criterio di selezione degli interventi. Gli interventi che soddisfano uno o più dei criteri summenzionati ma non sono stati oggetto di studi approfonditi dovrebbero venire valutati sistematicamente al fine di documentare importanti lacune nello stato attuale della ricerca. La procedura di selezione degli interventi, pur essendo sistematica, implica la formulazione di giudizi: un diverso gruppo di partecipanti potrebbe scegliere di prendere in esame un insieme diverso di interventi.

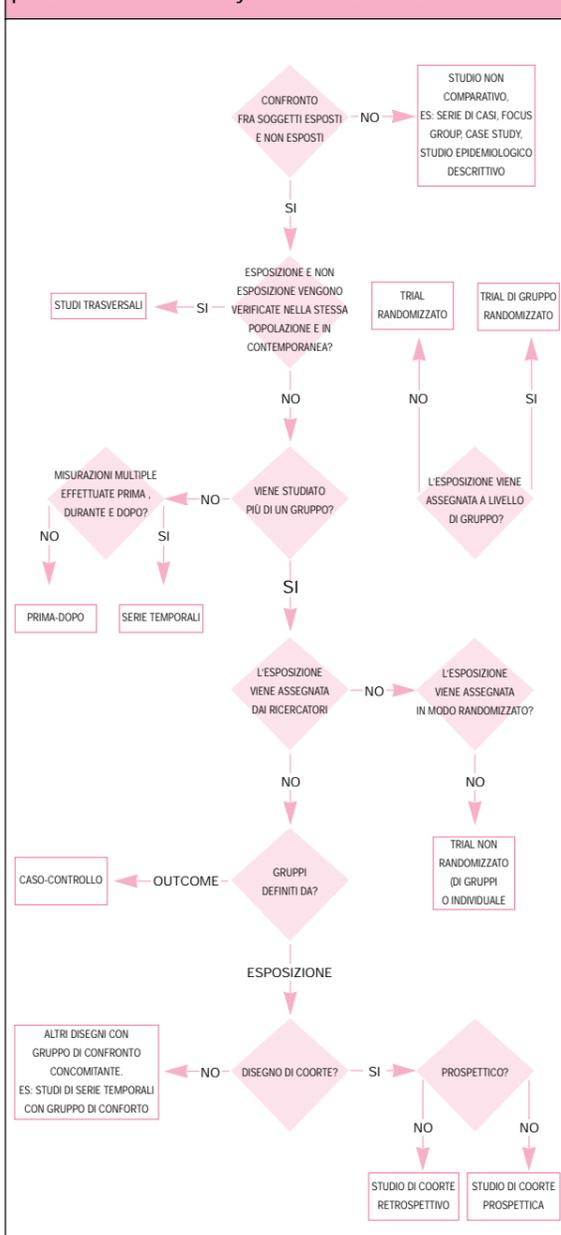
### Ricerca sistematica e catalogazione delle evidenze

I quadri analitici stabiliscono alcuni dei criteri di inclusione per la determinazione delle evidenze, specificando l'intervento o gli interventi considerati e l'esito o gli esiti relativi. Vengono indicati inoltre altri criteri di inclusione, quali i paesi e gli anni nei quali è stato condotto lo studio e le lingue nelle quali è stato pubblicato. L'esame della letteratura viene condotto su libri e riviste conformi ai criteri di inclusione e prevede ricerche su numerose banche dati computerizzate, revisioni bibliografiche e consultazioni di esperti. La necessità di reperire dati non pubblicati o tratti da altre fonti e i metodi utilizzati a tal fine vengono valutati per ogni singolo intervento da parte dei gruppi di lavoro incaricati dell'elaborazione dei capitoli. Vengono condotte ricerche esaustive al fine di ridurre le possibilità che venga privilegiata unilateralmente la ricerca di dati a sostegno di una determinata conclusione, ignorando altre informazioni.

TABELLA 1. Adeguatezza del disegno dello studio per valutare l'efficacia nella Community Guide

Adeguatezza	Caratteristiche
Elevata	Gruppi di controllo concomitanti e misure successive di esposizione e outcome
Moderata	Tutti i disegni retrospettivi o i disegni con misurazioni multiple pre o post ma senza gruppo di controllo concomitante.
Bassa	Misurazioni singole pre e post senza gruppo di controllo concomitante o esposizione e outcome misurati su un singolo gruppo nello stesso momento

ILLUSTRAZIONE 1. Classificazione dei disegni di studio per la Guide to Community Preventive Services



### Valutazione della qualità e sintesi delle prove di efficacia

Dopo aver selezionato i singoli studi che formano il corpus delle prove di efficacia relative ad un intervento, si procede alla loro valutazione, all'estrazione dei risultati, alla sintesi del corpus generale di evidenze e alla valutazione della forza delle prove di efficacia (ossia degli elementi che attestano che i cambiamenti degli esiti sono attribuibili agli interventi considerati).

Ciascuno studio conforme ai criteri espliciti di inclusione viene letto da due revisori che utilizzano un modulo di sintesi standardizzato<sup>12</sup> per registrare le informazioni relative ai seguenti aspetti:

- intervento preso in esame;
- contesto nel quale è stato condotto lo studio (popolazione, setting);
- disegno di valutazione;
- qualità dello studio;
- risultati.

Qualsiasi divergenza di opinione fra i due revisori viene conciliata raggiungendo il consenso all'interno del gruppo di lavoro incaricato della stesura del capitolo nel corso del processo di sintesi dei risultati nelle tabelle sinottiche delle evidenze.

Ciascuno studio viene qualificato considerando sia l'adeguatezza del disegno utilizzato per valutare l'efficacia che la qualità di esecuzione dello studio stesso. I disegni vengono classificati utilizzando uno schema standard (Illustrazione 1). L'adeguatezza del disegno sperimentale (Tabella 1) viene verificata alla luce di numerose caratteristiche che contribuiscono a evitare una serie di potenziali limiti che possono inficiarne la validità.

In base alla procedura adottata per l'elaborazione della Guida, affinché lo studio possa essere utilizzato per valutare l'efficacia è necessario che il disegno sperimentale preveda un confronto simultaneo oppure un confronto «prima - dopo». E' impossibile, infatti, determinare in che misura l'intervento sia stato efficace nel gruppo sottoposto all'intervento in mancanza di una valutazione della misura in cui gli esiti auspicati si sono verificati fra persone non esposte all'intervento. Anche altre caratteristiche del disegno contribuiscono ad aumentare l'adeguatezza dello studio nel valutare l'efficacia degli interventi. Un disegno che prevede un gruppo di controllo concomitante consente di evitare di interpretare erroneamente cambiamenti degli esiti non attribuibili all'intervento considerato. Sia pure in misura minore, anche studi con misurazioni multiple degli esiti compiute a distanza di tempo possono tutelare da eventuali cambiamenti concomitanti non attribuibili all'intervento. I disegni nei quali la valutazione dell'esposizione precede la valutazione degli esiti consentono di evitare errori sistematici nel rilevamento delle esposizioni.

I revisori valutano la qualità di esecuzione degli studi tenendo conto di sei categorie di limiti che possono inficiare la validità e che possono riguardare i seguenti aspetti: descrizioni delle popolazioni e degli interventi, campionamento, misurazione dell'esposizione e degli esiti, analisi dei dati, interpretazione dei risultati (ivi compresi follow-up, bias e confondimenti) e altro. Ciascuna categoria comprende numerosi quesiti.<sup>12</sup> Sono ammessi complessivamente 9 limiti. La qualità di esecuzione di ciascuno studio viene definita come buona, discreta o insufficiente sulla base del numero di limiti ri-

levati. La qualità degli studi nei quali sono stati riscontrati 0-1, 2-4, e 5 o più limiti di esecuzione viene classificata rispettivamente come buona, discreta e insufficiente. Gli studi la cui qualità di esecuzione viene giudicata insufficiente non vengono inseriti nei corpus di prove di supporto delle raccomandazioni. In generale, i dati relativi alla qualità di esecuzione degli studi si basano esclusivamente sulle informazioni tratte da relazioni pubblicate. Ciò in quanto, da un lato, la scarsa disponibilità o la qualità variabile di informazioni supplementari fornite dagli autori potrebbero dare adito a errori di valutazione e, dall'altro, può risultare impossibile ottenere ulteriori informazioni dagli autori.

I risultati tratti da un gruppo di studi collegati vengono sintetizzati dal punto di vista qualitativo e, ogni qual volta è possibile, raggruppati utilizzando statistiche descrittive quali la mediana e il campo di variazione o l'intervallo interquartile delle entità degli effetti. A seconda dell'adeguatezza e applicabilità di una sintesi quantitativa e della disponibilità di misurazioni statistiche della variabilità o dei dati a partire dai quali calcolarle, possono essere adottate anche procedure formali di raggruppamento statistico per descrivere una misura riassuntiva degli effetti.

Inoltre il corpus delle prove di efficacia viene qualificato come forte, sufficiente o insufficiente in base al numero di studi disponibili, alla qualità del disegno e dell'esecuzione, nonché all'entità e coerenza degli effetti registrati (Tabella 2). Prove classificabili come sufficienti o forti possono essere ottenute con numerosi procedimenti diversi, in grado di coniugare rigore scientifico, fattibilità e adeguatezza della valutazione dell'ampia gamma di interventi utilizzati nelle strategie di popolazione finalizzate a migliorare le condizioni di salute. La *Task Force* si riserva inoltre l'opzione di ricorrere al parere di esperti nelle rare circostanze in cui non siano disponibili altre evidenze e si ritenga che l'intervento sia sufficientemente significativo e/o utilizzato su vasta scala da richiedere una raccomandazione al riguardo.

Nel procedere alla classificazione dei corpus di prove di efficacia come forti, sufficienti o insufficienti si è fatto riferimento a una serie di principi. Le prove forti o sufficienti possono essere basate su un numero ridotto di studi che presentano una maggiore validità di esecuzione e un disegno più adeguato o da un numero più elevato di studi con disegno meno adeguato o maggiori punti deboli nell'esecuzione (Tabella 2). Per tutte le prove classificate come forti o sufficienti, i risultati degli studi devono generalmente raggiungere un grado sufficiente di coerenza dal punto di vista della direzione e dell'entità. Benché un singolo studio possa presentare evidenze sufficienti o forti, posto che tutte le altre caratteristiche risultino confrontabili, un maggior numero di studi fornisce un corpus di evidenze più forte in quanto un più elevato numero di repliche può contribuire a ridurre la possibilità che i risultati dei singoli studi siano stati ottenuti casualmente o siano dovuti a errori sistematici, qualora i potenziali difetti dei singoli studi non siano identici. In generale si

ritiene che effetti di maggiore entità (ossia rischi assoluti o relativi) forniscano prove di efficacia più forti rispetto a effetti di minore entità. La valutazione dell'entità degli effetti viene effettuata dalla *Task Force* caso per caso.

### Traduzione delle prove di efficacia in raccomandazioni

#### Efficacia

In generale la forza delle prove di efficacia (Tabella 2) è correlata direttamente con la forza delle raccomandazioni (Tabella 3). Evidenze giudicate incoerenti dal punto di vista della direzione o dell'entità degli effetti in base a caratteristiche definibili della popolazione, del setting o dell'intervento dovrebbero indurre a formulare raccomandazioni distinte per situazioni specifiche. Per esempio, alcuni interventi potrebbero essere raccomandati per popolazioni urbane ma non per popolazioni rurali. Altri interventi potrebbero essere indicati per strutture cliniche di dipartimenti sanitari ma non per organizzazioni di managed care. Inoltre potrebbero venire raccomandati interventi più o meno intensivi. La presenza di evidenze insufficienti o contraddittorie, in mancanza di prove che attestino l'efficacia dell'intervento in uno o più contesti definibili, dovrebbe indurre a concludere che le evidenze sono insufficienti per valutare l'efficacia dell'intervento presso qualsiasi popolazione. L'inefficacia documentata presso talune popolazioni, in mancanza di prove che attestino l'efficacia di un intervento in situazioni definibili, deve indurre a sconsigliare l'attuazione dell'intervento stesso in tutte le popolazioni. Talvolta nelle raccomandazioni della *Task Force* si fa riferimento a evidenze diverse dalle prove di efficacia ed esse vengono sintetizzate di routine ad uso dei lettori.

#### Applicabilità

Per aiutare gli utenti della Guida a stabilire con che grado di probabilità le informazioni fornite possano applicarsi o meno alle proprie realtà locali, i gruppi di lavoro incaricati della stesura dei capitoli hanno:

- definito le popolazioni bersaglio e i contesti per i quali l'intervento può essere preso in considerazione;
  - verificato se gli studi disponibili abbiano valutato l'intervento in tali popolazioni e contesti;
  - valutato in che misura le popolazioni o i contesti presi in esame in tali studi possano essere considerati rappresentativi delle popolazioni bersaglio e dei contesti di interesse;
  - formulato alcuni giudizi sulla possibilità che l'intervento in questione possa funzionare meglio o peggio in talune popolazioni e in determinati contesti rispetto ad altri.
- Sulla base di tali informazioni la *Task Force* esprimerà il proprio parere sulle possibilità di applicazione delle relative raccomandazioni e individuerà spunti per ulteriori ricerche.

#### Altri effetti

Gli interventi sanitari finalizzati a influire su uno o più esiti sanitari possono talvolta produrre effetti di altro tipo (ossia effet-

**TABELLA 2. Valutazione della forza di un corpus di prove di efficacia relativo a interventi di popolazione nella Guide to Community Preventive Services**

Prove di efficacia <sup>a</sup>	Esecuzione:buona o discreta <sup>b</sup>	Adeguatezza del disegno: massima, media o minima	Numero di studi	Coerenza <sup>c</sup>	Entità degli effetti <sup>d</sup>	Parere di esperti <sup>e</sup>
<b>Forte</b>	Buona	Massima	Almeno 2	Si	Sufficiente	Non utilizzato
	Buona	Massima o media	Almeno 5	Si	Sufficiente	Non utilizzato
	Buona o discreta	Massima	Almeno 5	Si	Sufficiente	Non utilizzato
	Vengono soddisfatti i criteri relativi a disegno, esecuzione e coerenza per evidenze sufficienti ma non forti				Elevata	Non utilizzato
<b>Sufficiente</b>	Buona	Massima	1	Non applicabile	Sufficiente	Non utilizzato
	Buona o discreta	Massima o media	Almeno 3	Si	Sufficiente	Non utilizzato
	Buona o discreta	Massima, media o minima	Almeno 5	Si	Sufficiente	Non utilizzato
<b>Parere di esperti</b>	Variabile	Variabile	Variabile	Variabile	Sufficiente	A sostegno di una raccomandazione
<b>Insufficiente<sup>f</sup></b>	A. Disegni o esecuzioni insufficienti		B. Numero di studi insufficiente	C. Incoerente	D. Insufficiente	E. Non utilizzato

<sup>a</sup> Le categorie non si escludono reciprocamente; un corpus di evidenze che soddisfa i criteri relativi a più categorie dovrebbe essere classificato nella categoria più elevata possibile.

<sup>b</sup> Gli studi con limiti di esecuzione non vengono utilizzati per la valutazione dell'efficacia.

<sup>c</sup> Coerenza generale nella direzione ed entità.

<sup>d</sup> Entità degli effetti sufficienti e elevate vengono definite caso per caso sulla base del parere della Task Force.

<sup>e</sup> Le ragioni per qualificare le evidenze come insufficienti vengono così riassunte: A. Disegno o esecuzione insufficiente. B. Numero di studi insufficiente. C. Incoerenza D. Entità degli effetti insufficiente E. Il parere degli esperti non viene utilizzato. Queste categorie non si escludono a vicenda e un corpus di evidenze che non soddisfa i criteri relativi a evidenze forti o sufficienti può rientrare in una o più categorie.

ti collaterali, comprendenti esiti degni di nota dell'intervento che non rientrano nella valutazione dell'efficacia, siano essi dannosi o benefici, previsti o imprevisi e connessi a esiti sanitari o non sanitari). Ad esempio, un intervento finalizzato a migliorare la qualità delle acque dei canali navigabili può avere effetti positivi dal punto di vista ricreativo, oltre che sanitario. Le revisioni contenute nella Guida possono individuare altri effetti importanti, effettuare ricerche sistematiche e valutare la forza delle evidenze a sostegno di tali effetti utilizzando lo stesso procedimento utilizzato per le prove di efficacia. Nel caso in cui vengano dimostrati altri effetti, essi verranno discussi nel relativo capitolo. Qualora vengano documentati effetti dannosi che superano i benefici il ricorso agli interventi in questione verrà sconsigliato. Gli interventi per i quali si pongono importanti considerazioni distributive (ossia interventi che risultano efficaci presso alcune popolazioni ma dannosi presso una o più «popolazioni collaterali») potrebbero indurre a formulare raccomandazioni più mirate o di tenore meno positivo. Qualsiasi informazione disponibile su altri effetti documentati verrà segnalata; va comunque tenuto presente che, in generale, la constatazione di altri effetti di ordine non sanitario (in mancanza di evidenze sufficienti o forti che attestino il miglioramento degli esiti sanitari) non verrà utilizzata per giustificare raccomandazioni positive.

#### Valutazioni economiche

I gruppi incaricati della stesura dei capitoli effettuano ricerche sistematiche delle valutazioni economiche disponibili relative a

interventi raccomandati o fortemente raccomandati. Una volta che tali studi sono stati individuati, essi vengono valutati e sintetizzati utilizzando un modulo standard di valutazione<sup>13</sup> e, nella misura del possibile, standardizzati al caso di riferimento raccomandato dal *Panel on Cost-Effectiveness in Health and Medicine*<sup>14,15</sup>. La standardizzazione dei risultati è utile, data la grande variabilità dei disegni e delle modalità di esposizione delle valutazioni economiche che limita notevolmente la confrontabilità dei risultati fra i vari studi.

Le risorse destinate a interventi volti a migliorare le condizioni di salute sono sempre limitate. Le analisi economiche possono fornire ai responsabili decisionali utili informazioni in merito alle risorse necessarie per l'attuazione di vari interventi finalizzati a migliorare gli esiti sanitari e aiutandoli a stanziare tali risorse in modo da massimizzare i benefici ottenuti. Quale ausilio per il processo decisionale, nella Guida i dati economici disponibili verranno riportati in sintesi. Di norma, tuttavia, le informazioni economiche non influiranno sul tenore delle raccomandazioni contenute nella Guida, considerando i limiti riscontrati nella disponibilità e qualità dei dati e il fatto che i vari utenti possono valutare diversamente se e come utilizzare le informazioni economiche in sede decisionale.

#### Sintesi degli ostacoli all'attuazione degli interventi

Per ciascun intervento valutato nelle revisioni contenute nella Guida verranno prese in esame le informazioni applicabili sugli ostacoli riscontrati in fase di attuazione degli interventi. Que-

ste informazioni vengono inserite principalmente affinché i responsabili decisionali possano tenerne conto nel selezionare gli interventi. Benché la Guida non sia stata concepita quale manuale di attuazione degli interventi, questa sezione può risultare utile anche agli operatori sanitari chiamati ad attuare gli interventi. In linea generale i dati relativi agli ostacoli di attuazione non influiscono sulle raccomandazioni.

#### Sintesi delle lacune nella ricerca

Le revisioni sistematiche condotte nella Guida sono volte a individuare le informazioni esistenti sulla base delle quali formulare raccomandazioni nel settore della sanità pubblica. Un importante vantaggio supplementare di tali revisioni consiste nell'individuazione di ambiti nei quali le informazioni mancano o sono di basso livello qualitativo.

#### Discussione

La Guida fa riferimento alle precedenti esperienze, che sono cospicue nell'ambito delle revisioni sistematiche<sup>1, 5</sup> e più limitate in quello del collegamento delle evidenze alle raccomandazioni<sup>3, 16-18</sup>. Degli sforzi precedenti la Guida condivide l'impegno ad adottare una procedura sistematica e a esplicitare i presupposti.

Gli interventi di popolazione si differenziano dagli interventi individuali, sui quali sono incentrate molte delle precedenti revisioni sistematiche in ambito sanitario, da vari punti di vista: per la scala di attuazione degli interventi e di misurazione degli esiti; per il maggior numero, la variabilità e la complessità di possibili opzioni di intervento; per la complessità e la variabilità dei contesti in cui vengono effettuati gli interventi. Queste differenze fra gli interventi individuali e di popolazione hanno indotto la *Task Force* ad adeguare i metodi esistenti di valutazione delle evidenze e di collegamento delle evidenze alle raccomandazioni.

La randomizzazione rappresenta un efficace strumento di controllo dei confondimenti nella ricerca e pertanto nella Guida i trial clinici randomizzati e i trial randomizzati di gruppo vengono classificati come due dei disegni sperimentali più idonei per la valutazione dell'efficacia. Tuttavia la randomizzazione non è sempre praticabile o eticamente accettabile nell'ambito della ricerca di popolazione. Inoltre i trial randomizzati di gruppo spesso non sono in grado di randomizzare in modo plausibile un numero sufficiente di unità, assicurando in tal modo la distribuzione omogenea di potenziali effetti confondenti fra i gruppi<sup>19</sup>; possono dunque venire meno alcuni vantaggi rispetto a strategie alternative di controllo dei confondimenti presenti nei trial clinici randomizzati degli studi finalizzati a verificare l'efficacia ed efficienza clinica.

I criteri di valutazione dell'adeguatezza dei disegni di studio adottati nella Guida consentono di prendere in esame studi che raccolgono sia dati individuali che ecologici. Questa scelta è stata compiuta in quanto numerosi interventi di comunità vengono applicati a varie popolazioni e potrebbe risultare difficile, im-

**Tabella 3. Corrispondenza fra forza delle prove di efficacia e forza delle raccomandazioni**

Forza delle prove di efficacia	Raccomandazione
Forti	Fortemente raccomandato
Sufficienti	Raccomandato
Dati empirici insufficienti integrati dal parere di esperti	Raccomandato sulla base del parere di esperti
Insufficienti	Gli studi disponibili non forniscono evidenze sufficienti ai fini della valutazione
Prove forti o sufficienti attestano l'inefficacia o il danno	Sconsigliato

possibile o inopportuno utilizzare dati a livello individuale. Inoltre i dati a livello individuale e quelli a livello di popolazione hanno punti di forza e limiti diversi e possono fornire informazioni complementari. Gli studi che si avvalgono di dati ecologici talvolta sono limitati da effetti attenuati o difficoltà di controllo dei confondimenti<sup>20</sup> e, se del caso, si tiene conto di tali limiti nel valutare la qualità di esecuzione degli studi.

La Guida valuta approfonditamente e tiene conto tanto della qualità di esecuzione degli studi che del disegno. Ciò consente, per esempio, di attribuire un maggior peso a studi caso-controllo o a studi prospettivi di coorte ben eseguiti rispetto a trial randomizzati che presentano una bassa qualità di esecuzione. Esistono numerosi metodi diversi di misurazione della qualità di esecuzione degli studi.<sup>21-23</sup> I metodi di valutazione della qualità possono variare tanto dal punto di vista delle caratteristiche considerate (ovvero misurazioni di validità interna, esterna, statistica e di costrutto) che dei metodi di misurazione di tali caratteristiche. L'impatto dei vari approcci alla qualità sui risultati registrati negli studi<sup>22</sup> e sui risultati delle metanalisi sono attualmente oggetto di notevole interesse nell'ambito della ricerca e, per il momento, non esiste nessuno «gold standard». I metodi di valutazione della qualità di studi non randomizzati (probabilmente la parte più cospicua degli studi presi in esame nella Guida) sono forse ancor meno sviluppati rispetto a quelli adottati per gli studi randomizzati. Per la Guida è stato messo a punto un approccio all'esecuzione degli studi che si richiama a: le metodologie applicate ad altre revisioni sistematiche; standard di pubblicazione fissati dalle principali riviste specializzate nel settore sanitario e delle scienze sociali; la letteratura statistica, metanalitica e di valutazione; la revisione e l'opinione degli esperti. La validità di contenuto dell'approccio è rafforzata ulteriormente dalla metodologia adottata per la sua elaborazione. Tale approccio resta comunque uno fra i molti accettabili e ci auguriamo che i sistemi di valutazione dell'esecuzione degli studi continuino ad evolversi.

Nella scelta della metodologia da utilizzare la *Task Force* ha tenuto conto di numerosi obiettivi: 1) ottenere e utilizzare le migliori evidenze empiriche disponibili a supporto del processo decisionale;

- fissare standard in grado di migliorare la disponibilità e qualità delle evidenze nel corso del tempo;
- formulare raccomandazioni che non richiedono dati di qualità inaccessibile;
- coniugare la necessità di adottare un approccio coerente per l'intera Guida con l'esigenza di un metodo di valutazione che sia appropriato e praticabile per tutti gli argomenti;
- far fronte a limitazioni di tempo e risorse.

Per mettere a punto una procedura sistematica e coerente sono state compiute alcune scelte metodologiche. E' stato stabilito, per esempio:

- quali problematiche prendere in esame nelle revisioni sistematiche;
- quali evidenze considerare adeguate ai fini della trattazione di tali problematiche (ossia quali i tipi di disegno ed esecuzione degli studi fossero migliori o peggiori, per valutare l'efficacia, per raccogliere i dati economici o per descrivere gli ostacoli all'attuazione degli interventi);
- come valutare la qualità degli studi;
- come distinguere fra evidenze insufficienti, sufficienti e forti;
- come collegare le evidenze alle raccomandazioni (ossia che peso attribuire all'efficacia rispetto agli effetti indesiderati, i costi o gli ostacoli).

La *Task Force* ritiene che le motivazioni alla base delle proprie scelte siano valide e che la procedura adottata consenta di trarre conclusioni ragionevoli e giustificate. Tuttavia, l'attuazione e valutazione degli interventi di popolazione finalizzati a migliorare le condizioni di salute e a prevenire le malattie costituiscono un ambito in rapida evoluzione. I metodi applicati nella Guida per valutare in modo esplicito e attendibile gli interventi di popolazione si sono via via perfezionati nel corso dell'elaborazione della Guida e continueranno a progredire per mantenere il passo con gli sviluppi nel settore.

Le ricerche approfondite di informazioni rappresentano un compito arduo e che richiede molto tempo. Attualmente non sono disponibili fonti di informazione facilmente accessibili, consultabili e ben indicizzate per reperire dati sugli interventi di popolazione. Le revisioni e raccomandazioni esistenti sono generalmente buone fonti bibliografiche ma per il resto possono risultare difficili da integrare nel processo di elaborazione della Guida in quanto numerosi metodi di revisione non vengono formulati esplicitamente e a causa delle differenze nei criteri di inclusione, degli esiti presi in esame o degli scopi perseguiti. Idealmente un corpus di evidenze complete dovrebbe comprendere informazioni pubblicate su libri e riviste nonché informazioni tratte da altre fonti (quali, relazioni governative, atti di congressi, tesi di laurea e dissertazioni). Vista la maggiore probabilità che i dati pubblicati nelle riviste possano attestare effetti positivi<sup>24</sup> rispetto ad altre fonti di informazione, limitare una revisione ai dati pubblicati potrebbe indurre a sovrastimare gli effetti (bias di pubblicazione). Tuttavia, la capacità di identificare e comprendere altre fonti d'informazione deve essere valutata rispetto ai limiti pratici nel reperire ed sintetizzare i dati for-

niti nonché alla possibilità che taluni tipi di dati siano di qualità inferiore rispetto a quelli pubblicati nelle riviste. Infine, nel processo di elaborazione della Guida si è fatto riferimento principalmente alla letteratura di lingua inglese e ciò potrebbe comportare l'omissione di alcune informazioni applicabili.

La definizione e classificazione degli interventi ai fini della valutazione solleva una serie di difficili problemi. La classificazione e sintesi fra interventi eterogenei inevitabilmente determina la perdita di alcuni dettagli relativi alle caratteristiche degli interventi e del contesto nel quale vengono effettuati. Le difficoltà si moltiplicano nel caso degli interventi con componenti multiple. Proprio per questo la Guida non si propone quale manuale di attuazione degli interventi, benché le sintesi attente delle evidenze relative a interventi simili consentiranno la descrizione dei tipi di attività intrapresi e di esprimere considerazioni informate sui relativi vantaggi e svantaggi. Risulterà inoltre possibile la maggiore comprensione della misura in cui gli interventi raggiungono effettivamente gli esiti auspicati. La capacità di mettere a confronto e paragonare le esperienze dovrebbe determinare un miglioramento del processo decisionale relativo alla scelta degli interventi da attuare.

La Guida non dovrebbe essere considerata come l'unica fonte alla quale far riferimento in un processo decisionale informato nell'ambito della sanità pubblica, data l'importanza che assumono anche le informazioni locali contestuali. La scelta degli interventi da attuare sarà influenzata da molti fattori non affrontati nella Guida (quali la disponibilità di risorse, la giustizia sociale, la partecipazione della collettività, l'adeguatezza culturale, l'incidenza e prevalenza a livello locale di malattie e fattori di rischio nonché considerazioni politiche). La Guida, tuttavia, fornirà informazioni approfondite e raccolte sistematicamente su numerose problematiche rilevanti per gli operatori sanitari e i responsabili decisionali della sanità pubblica, informazioni che è difficile o improduttivo elaborare a livello locale. Le revisioni e raccomandazioni contenute nella Guida risulteranno della massima utilità se associate ad un processo partecipativo di pianificazione a livello di comunità che chiarisca le esigenze e gli obiettivi e che tenga conto delle revisioni delle prove e delle raccomandazioni formulate nella Guida insieme ad ulteriori informazioni applicabili specificatamente connesse alla situazione della comunità.

Le raccomandazioni basate sulle evidenze non sono prive di potenziali svantaggi. L'adozione di approcci basati sulle evidenze potrebbe comportare difficoltà nella formulazione delle raccomandazioni dovute a mancanza di informazioni, a problemi connessi alla qualità delle informazioni disponibili, a necessità di valutare esiti importanti difficili da misurare o previsti in un lontano futuro o ad approcci promettenti che non sono ancora stati adeguatamente verificati. In alternativa, approcci basati sulle evidenze potrebbero dare adito a raccomandazioni di portata troppo generale o basati su informazioni insufficienti. Nonostante l'interesse suscitato dai metodi basati sulle evidenze, resta tuttora aperto il quesito se e in che misura essi possano de-

terminare cambiamenti nella prassi professionale o esiti sanitari migliori rispetto a linee guide fondate su altri presupposti. La *Task Force* è giunta alla conclusione che i vantaggi dei metodi basati sulle evidenze superino i potenziali svantaggi. I processi sistematici e partecipativi ridurranno gli errori commessi nelle modalità di raccolta e interpretazione dei dati, riducendo la probabilità che le raccomandazioni riflettano esclusivamente informazioni selezionate o un punto di vista unilaterale. Facendo riferimento a chiari principi analitici di base nel raccomandare determinati interventi gli utenti della Guida avranno maggiori possibilità di valutare se le raccomandazioni siano valide e prudenti dal loro punto di vista, se esse risultino applicabili ai rispettivi contesti locali e se consentano di raggiungere gli obiettivi ai quali essi attribuiscono importanza. Infine, le revisioni basate sulle evidenze rilevano esplicitamente tutti i limiti e le incertezze presenti nei dati disponibili, offrendo così l'opportunità di migliorare la qualità della ricerca e di stimolarla a colmare importanti lacune. Tutti questi vantaggi possono progressivamente aumentare il consenso su strategie sanitarie di comunità adeguate e contribuire a estendere la loro attuazione.

#### Bibliografia

1. Cooper H, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
2. Mulrow CD, Oxman AD, eds. *Cochrane collaboration handbook*. In: *The Cochrane library* [CDROM, updated September 1997]. The Cochrane Collaboration. Oxford: Update Software, 1997: Issue 4.
3. US Preventive Services Task Force. *Guide to clinical preventive services*. 2nd ed. Alexandria, VA: *International Medical Publishing*, 1996.
4. The Cochrane Collaboration. *The Cochrane database of systematic reviews* 1999 vol 2 [online database]. Available at <http://www.updateusa.com/clubip/club.htm>. Accessed July 22, 1999.
5. Agency for Health Care Policy and Research, Rockville MD. *Evidence-based practice centers: overview*, December 1998. Available at: <http://www.ahrp.gov/clinic/epc/>. Accessed July 22, 1999.
6. CDC Guidelines: Improving the Quality. 1996. Centers for Disease Control and Prevention, Atlanta, GA.
7. Briss PA, Rodewald L, Hinman A, et al. and the Task Force on Community Preventive Services. Reviews of evidence for interventions to improve vaccination coverage in children, adolescents and adults. *Am J Prev Med* 2000; 18(suppl 1): 97-140.
8. Task Force on Community Preventive Services. Recommendations on interventions to improve vaccination coverage in children, adolescents and adults. *Am J Prev Med* 2000; 18(suppl 1): 92-96.
9. Shefer A, Briss P, Rodewald L, Bernier R, Strikas R, Yusuf H, et al. Improving immunization coverage rates: an evidence-based review of the literature. *Epidemiol Rev* 1999; 20: 96-142.
10. The Task Force on Community Preventive Services. *The guide to community preventive services*. Available at: <http://web.health.gov/communityguide>. Accessed July 22, 1999.
11. Woolf SH. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale. In: McCormick KA, Moore SR, Siegel RA, eds. *Clinical practice guideline development: methodology perspectives*. Washington, DC: US Dept. of Health and Human Services, Agency for Health Care Policy and Research, 1994: 105-13.
12. Zaza S, Wright-De Agüero LK, Briss PA, et al. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. *Am J Prev Med* 2000;(suppl 1): 44-74.
13. Carande- Kulis VG, Maciosek MV, Briss PA, et al. and the Task Force on Community Preventive Services. Methods for systematic review of economic evaluations for the Guide to Community Preventive Services. *Am J Prev Med* 2000; 18(suppl 1): 75-91.
14. Gould MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-effective-ness in health and medicine*. Oxford: Oxford University Press, 1996.
15. Haddix AC, Teutsch SM, Shaffer PA, Dunet DO. *Decision analysis and economic evaluation*. Oxford: Oxford University Press, 1996.
16. Novick LF. Public health practice guidelines: a case study. *J Public Health Management and Practice* 1997; 3: 59-64.
17. Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman MD, Scott EAF, Millson ME, et al. An approach to the development of practice guidelines for community health interventions. *Can J Public Health* 1994; 85(suppl 1): S8-S13.
18. Canadian Task Force on the Periodic Health Examination. *Canadian guide to clinical, preventive health care*. Ottawa, Canada: Canada Communication Group, 1994.
19. Murray DM. Design and analysis of group-randomized trials: a review of recent developments. *Ann Epidemiol* 1999; 7(suppl 7): S69-S77.
20. Rothman KJ. *Modern epidemiology*. Boston: Little, Brown, and Company, 1986.
21. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials* 1995; 16: 621-13.
22. Moher D, Pham B, Jones A, et al. Does quality of reports of randomized trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; 352: 609-13.
23. Wortman PM. Judging research quality. In: Cooper H, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994: 97-109.
24. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992; 267: 374-8. *Am J Prev Med* 2000; 18(1S) 43 Reprinted by permission of Elsevier Science from: Developing an Evidence-Based Guide to Community Preventive Services - Methods. Peter A. Briss, Stephanie Zaza, Marguerite Pappaioanou et al., *American Journal of Preventive Medicine*, Vol 18 No 1S, pp 35-43, Copyright 2000 by *American Journal of Preventive Medicine*.