

Lo strumento di raccolta dei dati e la procedura di revisione sistematica

Stephanie Zaza, MD, MPH, Linda K. Wright-De Agüero, PhD, MPH, Peter A. Briss, MD, Benedict I. Truman, MD, MPH, David P. Hopkins, MD, MPH, Michael H. Hennessy, PhD, MPH, Daniel M. Sosin, MD, MPH, Laurie Anderson, PhD, Vilma G. Carande – Kulis, PhD, Steven M. Teutsch, MD, MPH, Marguerite Pappaioanou, DVM, PhD, Task Force on Community Preventive Services

Division of Prevention Research and Analytic Methods, Epidemiology Program Office (Zaza, Wright – De Agüero, Briss, Truman, Hopkins, Anderson, Carande – Kulis), Centers for Disease Control and Prevention (CDC); Division of STD Prevention, National Center for HIV, STD, and TB Prevention (Hennessy); National Center for Injury Prevention and Control (Sosin); Merck & Co., Inc. (Teutsch), West Point, Pennsylvania. Marguerite Pappaioanou che operava presso la Division of Prevention Research and Analytic Methods, CDC, Atlanta, Georgia, nel periodo in cui è stato portato a termine il presente lavoro, è attualmente presso l' Office of Global Health, CDC, Atlanta, Georgia.

Corrispondenza e altre richieste possono essere indirizzate a: Stephanie Zaza, MD, MPH, Community Preventive Services Guide Development Activity, Epidemiology Program Office, MS – K-73, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341.

Introduzione

Al fine di assicurare la coerenza, ridurre le distorsioni e migliorare la validità e attendibilità della *Guide to Community Preventive Services: Systematic Reviews and Evidence – Based Recommendations* (la Guida) sono stati elaborati una procedura e un modulo standardizzato di raccolta dei dati.

Lo strumento di raccolta dei dati

Il contenuto del modulo standardizzato di raccolta dei dati è stato stabilito sulla base di: metodologie adottate in altre revisioni sistematiche; standard di pubblicazione fissati dalle principali riviste specializzate nel settore della sanità e delle scienze sociali; la letteratura statistica, meta-analitica e di valutazione; revisioni e pareri di esperti; i risultati della verifica pilota. Il modulo viene utilizzato per classificare e descrivere le principali caratteristiche dell'intervento e della valutazione (26 domande) e verificare la qualità di esecuzione dello studio (23 domande). Vengono prese in esame le procedure e i risultati degli studi, valutando gli specifici limiti che possono inficiare la validità degli studi stessi in relazione a sei categorie di domande (relative a: descrizioni degli interventi e degli studi, campionamento, misurazione, analisi, interpretazione dei risultati e altri aspetti dell'esecuzione).

Procedure di raccolta dei dati

Ciascuno studio viene esaminato e riassunto da due revisori indipendenti e le opinioni vengono armonizzate dal gruppo di lavoro incaricato della stesura di ciascun capitolo. I revisori ricevono specifica formazione e feedback.

Discussione

La selezione e le modalità di sintesi dei dati sono scelte discrezionali che influiscono sulle conclusioni tratte in merito alla qualità di esecuzione dello studio e alla sua efficacia. Il modulo proposto mira a coniugare la flessibilità necessaria nella valutazione di studi alquanto eterogenei sia dal punto di vista del disegno che degli interventi oggetto d'indagine, con l'esigenza di formulare domande specifiche al fine di massimizzare la validità e l'attendibilità. Esso fornisce un'impostazione strutturata che può essere utilizzata dai ricercatori e da altri lettori della Guida per valutare i contenuti e la qualità degli studi, condurre revisioni sistematiche o redigere lavori scientifici. Un approccio sistematico all'elaborazione e valutazione dei documenti contribuirà a promuovere il generale miglioramento della letteratura scientifica.

Parole chiave: data abstraction, evaluation, study design, study quality
Am J Prev Med 2000; 18 (1S): 44–74 © 2000 American Journal of Preventive Medicine

Introduzione

Nel presente documento vengono descritti lo strumento e la procedura adottati per raccogliere e valutare i dati tratti dai singoli studi di verifica dell'efficacia degli interventi, uno degli aspetti chiave della metodologia utilizzata per l'elaborazione della Guida. Il modulo esemplifica l'approccio adottato dalla *Task Force* nel classificare le informazioni relative al disegno di studio, ai contenuti e alla qualità della letteratura scientifica. Tale approccio risulterà utile ad altri nell'esame della letteratura scientifica, nella stesura di articoli di carattere scientifico, nella pro-

gettazione di studi di valutazione nonché nell'insegnamento dell'epidemiologia e delle metodologie di valutazione.

Lo strumento di raccolta dei dati

Nel mettere a punto lo strumento di raccolta dei dati sono stati perseguiti sei scopi principali:

- Registrazione del processo di revisione degli articoli nelle sue diverse fasi. Il modulo raccoglie le informazioni necessarie per monitorare lo stato di screening, revisione e sintesi di ciascun articolo da parte di più revisori.

TABELLA 1. Categorie di domande finalizzate a valutare i limiti che possono potenzialmente inficiare la validità di ciascuno studio, tratte dallo strumento di raccolta dei dati della *Guide to Community Preventive Services*.

Categorie	Limiti che possono potenzialmente inficiare la validità
Descrizioni Esempio: l'intervento è descritto bene?	Distorsioni introdotte dalla mancata conservazione dell'integrità dell'intervento
Campionamento Esempio: gli autori hanno specificato i criteri di screening per la ammissibilità degli studi	Distorsione da selezione
Misurazione Esempio: le misure relative a esposizione ed esiti erano valide e attendibili	Distorsioni da misurazione osservatore/intervistatore dati auto-riportati dati ricordati (recall) altri Distorsione da errata classificazione esposizione esiti
Analisi Esempio: gli autori hanno condotto un'analisi appropriata effettuando test statistici, controllando misure ripetute, ecc.	Distorsioni analitiche misure ripetute esposizione differenziale effetti dovuti al disegno distorsione cross-level altri
Interpretazione dei risultati Esempio: gli autori hanno corretto per i fattori confondenti controllabili?	Distorsione da esclusione Fattori confondenti Trend secolari Tutti gli altri

- Elaborazione delle tabelle sinottiche delle prove di efficacia. Il modulo registra dati descrittivi dettagliati relativi all'intervento e alla valutazione; tali dati vengono utilizzati per elaborare tabelle sinottiche delle evidenze per ciascun intervento.^{1,2}
 - Classificazione di altre caratteristiche chiave dell'intervento e della valutazione. Vengono raccolti dati descrittivi supplementari per costituire una banca dati che potrà essere consultata quale risorsa a disposizione dei ricercatori e dei responsabili della programmazione degli interventi.
 - Valutazione della qualità di esecuzione dello studio. I revisori individuano e registrano i limiti potenziali che possono inficiare la validità di ciascuno studio, dovuti a difetti di esecuzione o misurazione. Tali informazioni vengono utilizzate quale criterio per l'inclusione dello studio nel corpus di evidenze relativo a un intervento.¹
 - Individuazione di altre informazioni pertinenti. Il modulo registra informazioni relative all'applicabilità dell'intervento in contesti e popolazioni diverse da quelle studiate dai ricercatori, dati economici relativi all'intervento e altri effetti positivi o negativi dell'intervento.
 - Individuazione di studi supplementari che dovrebbero essere oggetto di revisione. Al fine di assicurare che non venga omissso nessuno studio rilevante, i revisori consultano le bibliografie contenute in ciascuno studio preso in esame ed elencano gli articoli attinenti potenzialmente idonei ad essere inseriti nel processo di revisione.
- Il contenuto del modulo è stato stabilito sulla base di: metodologie di revisione tratte da altre revisioni sistematiche (quali quel-

le utilizzate da Cochrane Collaboration); standard di pubblicazione fissati dalle principali riviste specializzate nel settore della sanità e delle scienze sociali; la letteratura statistica, meta-analitica e di valutazione; parere di esperti e la revisione delle bozze preliminari del modulo.³⁻¹³ Sulla base di tale revisione della letteratura e delle specifiche esigenze del processo di valutazione della Guida si è stabilito quali elementi includere nel modulo. Nella fase preliminare di elaborazione del modulo, i membri dei gruppi di lavoro incaricati della stesura dei capitoli e altri esperti hanno effettuato una verifica pilota della chiarezza e attendibilità delle risposte fra i vari revisori. Il modulo è stato rivisto e utilizzato per l'elaborazione di

numerosi capitoli. Il modulo di raccolta dei dati è stato oggetto di un'ulteriore revisione nella quale si è tenuto conto di questi primi riscontri, dell'avvio dell'elaborazione dei capitoli sui vari argomenti, dei suggerimenti forniti dai revisori, delle valutazioni dei coordinatori scientifici con formazione pluridisciplinare e di interviste con i coordinatori scientifici finalizzate a rilevare le incoerenze nell'interpretazione delle domande. La revisione mirava a chiarire le domande ambigue o confondenti, a estendere le istruzioni e gli esempi per guidare i revisori, a migliorare il formato del modulo, specificare i rinvii fra le domande, nonché elevare la qualità delle categorie di esecuzione. Il modulo di raccolta dei dati è costituito da un opuscolo che riporta nelle pagine a sinistra le istruzioni corrispondenti alle domande indicate nelle pagine di destra. Il modulo ha una lunghezza di 26 pagine, compresi la copertina con dati identificativi, le istruzioni e le pagine delle risposte. Esso contiene 26 domande relative ai contenuti dello studio e 23 domande relative alla qualità di esecuzione dello stesso. Per leggere ciascun documento ed estrarre i dati utilizzando il modulo sono necessarie due-tre ore. (Il modulo completo può essere richiesto all'indirizzo indicato nella nota in apertura del capitolo). Le domande sono suddivise in tre sezioni: informazioni relative alla classificazione, informazioni descrittive e informazioni relative alla qualità di esecuzione. Le informazioni relative alla classificazione vengono compilate dal gruppo di lavoro incaricato della stesura di ciascun capitolo al fine di indirizzare la valutazione dello studio da parte del revisore. Questa prima sezione comprende uno schema di classificazione dei disegni de-

gli studi che consente di assegnare ciascuno studio ad una sola categoria. L'adeguatezza del disegno di studio ai fini della valutazione dell'efficacia di un intervento viene valutata e qualificata separatamente (ossia, non utilizzando il modulo di raccolta dei dati).¹ Inoltre le componenti degli interventi vengono classificate in quattro raggruppamenti principali a fini di registrazione e ricerca. Infine, dal gruppo incaricato della stesura dei capitoli vengono stabiliti gli esiti rilevanti da registrare in base al quadro teorico relativo al capitolo. Le risposte vengono verificate e corrette, se necessario, dal revisore.

La seconda sezione del modulo permette al revisore di documentare i metodi e i risultati dello studio. In primo luogo, ai revisori viene richiesto di descrivere l'intervento (che cosa, come, dove e a chi); l'impianto teorico sul quale è basato l'intervento; il tipo di organizzazione che ha attuato l'intervento; qualsiasi intervento fornito ai gruppi di controllo. In secondo luogo, i revisori vengono guidati attraverso una serie di domande a precisare le caratteristiche dello studio di valutazione stesso:

- il luogo di valutazione, ivi compresi ubicazione, densità di popolazione e contesto (se diverso da quello di attuazione dell'intervento);
- la popolazione oggetto di studio (ossia dimensioni del campione e metodo di selezione, valutazione dell'esposizione all'intervento, aspetti demografici e informazione sui fattori di rischio);
- qualsiasi altra popolazione descritta dagli autori che potrebbe essere toccata dall'intervento;
- misurazione degli esiti e di altre caratteristiche della popolazione di studio.

In terzo luogo, i revisori seguono un'impostazione strutturata per riportare i risultati degli studi, ivi compresi: le misure degli effetti, i dati relativi a ciascun intervento e la potenza comparativa (comparison arm) registrati nello studio, il software utilizzato, i metodi analitici, la verifica delle ipotesi e la potenza dello studio. In quarto luogo, i revisori registrano le informazioni relative ad altre questioni chiave affrontate nello studio che possono risultare utili in futuro per il gruppo di lavoro incaricato della stesura dei capitoli nonché i riferimenti bibliografici che eventualmente conformi ai criteri di inclusione per la revisione. La terza sezione del modulo di raccolta dei dati registra il giudizio del revisore sulla qualità di esecuzione dello studio. I limiti specifici che possono inficiare la validità dello studio (Tabella 1) vengono valutati in relazione a sei categorie di distorsione (descrizione, campionamento, misurazione, analisi, interpretazione dei risultati e altri). Data la difficoltà di fornire un'interpretazione coerente di tali interrogativi, ai quali si tende a fornire una risposta soggettiva, le istruzioni forniscono norme esplicite di decisione (ossia che cosa considerare, che cosa non considerare) ed esempi specifici di come rispondere alla domanda in varie circostanze. Ai revisori è stato chiesto di fornire una spiegazione delle risposte per ciascuna domanda.

La qualità di valutazione viene definita sulla base di dati descrittivi raccolti dalla relazione. Per assistere i revisori, le domande

di questa sezione del modulo rimandano alle domande attinenti contenute nelle prime due sezioni. Per esempio, per stabilire l'adeguatezza del campione di popolazione oggetto di studio, si rinvia il revisore alle domande nella seconda sezione del modulo in cui si chiede di specificare la descrizione della popolazione presa in esame, il metodo di campionamento, nonché le dimensioni del campione.

Ciascun tipo di disegno di studio comprende particolari aspetti che possono influire sul punteggio assegnato alla qualità di esecuzione. Per verificare la qualità di studi che utilizzano disegni diversi, sono state formulate domande che forniscono una valutazione dell'impostazione generale, mentre nelle istruzioni sono riportati esempi specifici al fine di agevolare il revisore. Per esempio, sono state inserite due domande generali relative alla validità e attendibilità per valutare i problemi potenziali concernenti la misurazione degli esiti (Tabella 2). Per un trial randomizzato la mancata applicazione del metodo in cieco a osservatori o intervistatori implica una limitazione della validità di misurazione degli esiti. Per uno studio con un disegno basato sulle serie temporali, invece, tale aspetto non verrebbe preso in considerazione ai fini della valutazione della validità della misurazione degli esiti, mentre ne verrebbero esaminati altri attinenti alla validità. Per tutte le domande della sezione relativa alla qualità di esecuzione, se l'aspetto della qualità avente specifica attinenza con il disegno di studio non viene affrontato esplicitamente nella relazione oggetto di revisione, allo studio viene attribuito un limite per la domanda in questione.

Procedure di raccolta dei dati

I dati vengono raccolti da ciascuno studio da due revisori indipendenti. Se i revisori rispondono alle domande fornendo informazioni diverse, il gruppo incaricato della stesura del capitolo armonizza le due diverse revisioni. L'incarico di effettuare le revisioni è affidato a membri dei gruppi incaricati dell'elaborazione dei capitoli, laureati ai corsi di Masters of Public Health, candidati a dottorati e medici che frequentano corsi di formazione di medicina preventiva. I revisori vengono selezionati in base alla loro esperienza nelle aree tematiche considerate e nell'esecuzione di revisioni basate su evidenze, nonché in base alle loro competenze specifiche in materia di metodologia e disegno sperimentale.

La formazione si articola in tre fasi. In primo luogo vengono fornite informazioni preliminari sul processo di elaborazione della Guida e sulla metodologia adottata, utilizzando fra l'altro uno studio campione corredato di un modulo compilato e di una tavola sinottica delle evidenze. In secondo luogo viene chiesto a ogni candidato di valutare uno studio che è stato precedentemente rivisto dal gruppo di lavoro incaricato della stesura dei capitoli. Questa prima revisione viene successivamente discussa approfonditamente con il candidato e il coordinatore scientifico fornisce ulteriori istruzioni per interpretare le domande. Nella terza fase i candidati selezionati prendono in esame gruppi di studi su interventi correlati ricevendo un feed-

Tabella 2. Domande tratte dal modulo di raccolta dei dati della Guide to Community Preventive Services

La variabile dipendente ed altre variabili indipendenti (o predittive) erano misure valide per gli esiti di interesse? Gli autori dovrebbero aver fornito indicazioni su uno o più dei seguenti punti:				
• Chiara definizione della variabile dipendente (outcome variable)				
• Misurazione degli esiti in vari modi. Esempio: analisi correlazionale fra gli esiti misurati per dimostrare la validità convergente (ossia 2 o più misure riflettono lo stesso processo sottostante) o la validità divergente (ossia 2 o più misure riflettono dimensioni diverse). Un esempio di validità convergente è dato da 5 item sull'auto-efficacia che presentano un alto grado di reciproca correlazione; un esempio di validità divergente è il fatto che le misure di auto-efficacia non presentano un alto grado di correlazione con le misure di atteggiamento.				
• Vengono citate o discusse le ragioni per cui il ricorso a tali misure è valido. Esempio: si veda sopra.				
• Altro. Esempio: nel caso in cui lo studio non venga eseguito in cieco – se applicabile – la risposta a questa domanda dovrebbe essere «no».				
La variabile dipendente ed altre variabili indipendenti (o predittive) erano misure attendibili (coerenti e riproducibili) per gli esiti di interesse? Gli autori dovrebbero aver fornito indicazioni su uno o più dei seguenti punti:				
• Misure di coerenza interna. Esempio: cfr. 3B				
• Misure degli esiti ottenute in vari modi. Esempio: cfr. 3B e 3C (sopra)				
• Si è tenuto conto della coerenza del punteggio codificato o della categorizzazione fra osservatori (verifiche di attendibilità fra valutatori) o fra diverse misure degli esiti. Esempio: percentuale di concordanza, Kappa.				
• Si è tenuto conto di come il contesto e il campionamento della popolazione di studio possa influire sull'attendibilità.				
• Vengono citate o discusse le ragioni per cui il ricorso a queste misure va considerato attendibile. Esempio: cfr. 3B.				
• Altro				
Opzioni di risposta	Si	No	N/R	Domande collegate
La variabile dipendente e altre variabili indipendenti (o predittive) erano:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	I/10
Valide?				
attendibili (coerenti e riproducibili?)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	II, 8, 9, 910, 18, 20

back costante da parte del gruppo di lavoro incaricato dell'elaborazione dei capitoli.

Discussione

Le revisioni sistematiche su quattro argomenti (fumo, attività fisica, lesioni a carico di occupanti di veicoli a motore e malattie prevenibili mediante vaccinazione) trattati nella Guida sono state condotte utilizzando le procedure e lo strumento di raccolta dei dati summenzionati. Sono stati passati in rassegna oltre 400 studi basati su disegni di vario tipo – che andavano da sperimentazioni randomizzate controllate a studi di serie temporali e a studi trasversali – da più di 40 revisori, ognuno dei quali ha fornito feedback e suggerimenti per migliorare le istruzioni e l'impostazione del modulo.

L'elaborazione di un modulo standardizzato di raccolta dei dati per una serie di revisioni sistematiche di carattere così vario come quelle contenute nella Guida ha posto due difficoltà fondamentali. Ci si è proposti di mettere a punto un modulo che consentisse, da un lato, la flessibilità necessaria per valutare studi con tipi di disegni e interventi alquanto diversi. Ci si è preoccupati che domande e istruzioni eccessivamente specifiche e dirette potessero limitare la flessibilità necessaria per affrontare questioni rilevanti in relazione ad aree tematiche e disegni di studio diversi. D'altra parte, occorreva coniugare l'esigenza di flessibilità con il rischio di una perdita di specificità nella modalità di formulazione delle domande, pregiudicando così potenzialmente l'attendibilità fra i vari valutatori e fra i vari capitoli. Si è dovuto inoltre conciliare il criterio della semplicità e concisione dal punto di vista formale con l'esigenza di poter

contare su informazioni approfondite.

Si è tentato di raggiungere un equilibrio fra questi vari aspetti attraverso un processo interattivo di elaborazione delle domande e delle istruzioni, utilizzando le domande e istruzioni per rivedere studi di aree tematiche diverse e con disegni diversi, valutando le risposte fornite nella revisione dei lavori, sollecitando il feedback da parte dei revisori e apportando modifiche al testo delle domande e delle istruzioni. Inoltre l'uso effettivo dei dati tratti da ciascuna domanda del modulo è stato posto a confronto con l'uso proposto; sono state eliminate dal modulo le domande che non fornivano i dati inseriti nelle tabelle delle evidenze, le informazioni necessarie per valutare la qualità di esecuzione degli studi, le informazioni complementari da utilizzare per l'elaborazione dei capitoli o il materiale da inserire nella banca dati della Guida.

Il modulo di raccolta dei dati pone i presupposti per trarre le conclusioni sui singoli studi. In qualsiasi processo di revisione sistematica, i singoli studi forniscono gli elementi d'informazione per rispondere ai quesiti della ricerca. La revisione e valutazione di ogni studio costituisce un'analisi qualitativa in sé e per sé. La decisione di quali dati considerare e di come sintetizzarli costituisce una scelta di tipo analitico che influisce sulle conclusioni tratte in merito all'esecuzione dello studio e alla sua efficacia.¹⁴ La validità e attendibilità del modulo di raccolta dei dati sono di cruciale importanza ai fini della fiducia nei risultati delle revisioni.

La validità formale e di contenuto del modulo viene rafforzata dal metodo adottato per la sua elaborazione: il modulo è stato predisposto sulla base di documenti simili precedenti ed è sta-

to rivisto e modificato nel corso di un processo interattivo al quale hanno concorso le revisioni di esperti, le verifiche della coerenza di interpretazione e codifica, nonché l'esame dei prodotti derivati dalle revisioni (ossia tabelle delle evidenze e raccomandazioni). Inoltre, il contenuto del modulo è stato posto a confronto con quello di strumenti analoghi utilizzati in altre occasioni e con i criteri qualitativi adottati da varie riviste per valutare gli studi da destinare alla pubblicazione. Questo processo di convalida mira fondamentalmente ad assicurare la chiarezza, completezza e rilevanza delle domande contenute nel modulo in relazione ai concetti chiave affrontati in ciascuna sezione. Le future versioni del modulo si baseranno sulla costante revisione della validità dello stesso. Per esempio, un quesito di ricerca ancora aperto è in che modo il rating di ciascuna categoria relativa alla qualità (cfr. Parte III del modulo) influisca sulla valutazione di un corpus di prove di efficacia dell'intervento e sulle raccomandazioni che ne derivano.

In sede di verifica delle prime versioni del modulo si è proceduto a valutare, oltre che la validità del modulo stesso, anche l'attendibilità delle risposte fornite dai vari revisori.^a In particolare è stata rivolta particolare attenzione alle domande che presentavano un minor grado di attendibilità nel corso della revisione del modulo. Benché tale analisi non sia stata ripetuta per la versione finale, i miglioramenti apportati alle istruzioni e alla formattazione del modulo dovrebbero aumentare la coerenza fra i vari valutatori.

Oltre ad aumentare la validità e attendibilità delle revisioni destinate alla Guida, questo modulo standardizzato di raccolta dei dati fornisce un'impostazione strutturata che i ricercatori e altri lettori della letteratura scientifica possono utilizzare per passare in rassegna il contenuto e la qualità degli studi, condurre altre revisioni basate sulle evidenze o elaborare lavori scientifici da sottoporre a riviste che si avvalgono di revisioni di esperti. Il miglioramento della letteratura e della metodologia nel settore della sanità pubblica costituisce uno degli obiettivi della *Task Force*. Al fine di promuovere un tale miglioramento, il modulo è stato elaborato con una procedura di dominio pubblico e può essere copiato e utilizzato liberamente.

Ringraziamenti: David Atkins, MD, MPH (Agency for Health Care Policy Research); Donna Higgins, PhD (CDC); David McQueen, PhD (CDC); Mike Stoto, PhD (George Washington University); e Steven Woolf, MD, MPH (Medical College of Virginia) per i loro utili commenti, valutazioni e suggerimenti volti a migliorare lo strumento di raccolta dei dati.

Note:

^a È stata valutata l'attendibilità fra revisori incaricati di esaminare lo stesso studio per singole domande relative a due capitoli (lesioni a carico degli occupanti di veicoli a motore e malattie prevenibili mediante vaccinazione). La percentuale di concordanza fra i revisori per le domande descrittive è risultata compresa fra il 34,5% e il 96,7% (mediana = 75,0%). La percentuale di concordanza fra revisori per le domande relative alla qualità di esecuzione è risultata compresa fra il 65,2% e l'85,6% (mediana = 79,5%).

Bibliografia

1. Briss PA, Zaza S, Pappaioanou M, et al. and the Task Force on Community Preventive Services. Developing an evidence – based Guide to Community Preventive Services – methods. *Am J Prev Med* 2000; 18(suppl 1): 35– 43.
2. Shefer A, Briss PA, Rodewald L, et al. Improving immunization coverage rates: an evidence – based review of the literature. *Epidemiol Rev* 1999; 21: 96–142.
3. Stool SE, Berg AO, Berman S, et al. Otitis media with effusion in children. Guideline Technical Report Number 12, *AHCPR* Publication No. 95– 0621. Rockville, MD: Agency for Health Care Policy Research, Public Health Service, U. S. Dept. of Health and Human Services, July 1994.
4. Canadian Task Force on the Periodic Health Examination. *Canadian guide to clinical preventive health care*. Ottawa: Canada Communication Group, 1994.
5. Gyorkos TW, Tannenbaum TN, Abrahamawicz M, Oxman AD, 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.
6. Hennessy M. Evaluation. In: Stroup DF, Teutsch SM, eds. *Statistics and public health: quantitative approaches to public health problems*. New York: Oxford University Press, 1998: 193–219.
7. Kafadar K, Andrews JS. *Investigating health effects and hazards in the community*. In: Stroup DF, Teutsch SM, eds. *Statistics and public health: quantitative approaches to public health problems*. New York: Oxford University Press, 1998: 93–122.
8. The Joint Committee on Standards for Educational Evaluation. *The program evaluation standards: how to assess evaluations of educational programs* 2 ed. Thousand Oaks, CA: Sage Publications, 1994.
9. Cooper H, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
10. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996; 276: 637–9.
11. US Preventive Services Task Force. *Guide to Clinical Preventive Services*. 2 ed. Alexandria, VA: *International Medical Publishing*, 1996.
12. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence – based medicine: how to practice and teach EBM*. New York: Churchill Livingstone, 1997.
13. *HIV Prevention Among Drug Users: A Resource Book for Community Planners and Program Managers*. Washington, DC: Academy for Educational Development, 1997.
14. Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, CA: Sage Publications, 1994.