

Framework per la valutazione dei programmi nella sanità pubblica

Prefazione

Ciò che i professionisti della sanità pubblica si sforzano di ottenere è il miglioramento della salute. Per raggiungere questo obiettivo dobbiamo dedicare le nostre capacità – e la nostra volontà – alla stima degli effetti delle azioni di sanità pubblica. Siccome gli obiettivi delle azioni di sanità pubblica si sono estesi oltre le malattie infettive per includere anche malattie croniche, violenza, nuovi patogeni, minacce di bioterrorismo ed i contesti sociali che influenzano la disuguaglianza nella salute e nell'accesso ai servizi sanitari, anche l'obiettivo della valutazione è diventato più complesso. I CDC hanno sviluppato una cornice concettuale per la valutazione dei programmi al fine di assicurare che, nel corso di una fase di complessa transizione nella sanità pubblica, l'azione sarà incentrata nell'ottenere outcomes di salute quantificabili.

Integrando i principi di questo framework in tutti i programmi operativi dei CDC, risulterà stimolata la ricerca di strategie innovative per il raggiungimento di migliori outcomes, e ci si troverà in una migliore posizione per riconoscere l'impatto dei programmi. Una valutazione più efficiente e tempestiva di questi effetti accrescerà la nostra abilità nel tradurre i risultati della ricerca per incorporarli nella pratica.

Anche l'approccio di base alla pianificazione dei programmi si evolverà, guidato dall'organizzazione in fasi e dagli standard del framework. Le scoperte della ricerca, nel campo della prevenzione, condurranno alla produzione di programmi più chiari e logici; una maggiore collaborazione permetterà agli operatori di concentrare gli sforzi per il raggiungimento di obiettivi comuni; l'integrazione dei sistemi informativi supporterà valutazioni sistematiche; e le lezioni apprese dalla valutazione saranno usate più efficacemente per modificare le strategie della sanità pubblica.

La pubblicazione di questo framework dà importanza al costante impegno dei CDC nel migliorare ovunque le condizioni di salute della comunità. Siccome le strategie di prevenzione non possono avere successo con interventi isolati (sporadici o limitati), i professionisti della sanità pubblica che lavorano su aree omogenee di programma, devono cooperare per poter stimare la loro influenza combinata sulla salute della comunità. Solo allora saremo capaci di realizzare e dimostrare il successo della nostra visione: persone in salute in un mondo sano, con la prevenzione.

Jeffrey P. Koplan, M.D., M.P.H.

Director, Centers for Disease Control and Prevention

Administrator, Agency for Toxic Substances and Disease Registry

Framework per la valutazione dei programmi nella sanità pubblica

Sommario

Un'efficace valutazione dei programmi rappresenta un modo sistematico per migliorare e valutare le azioni di sanità pubblica attraverso l'adozione di procedure utili, realizzabili, etiche ed accurate. Il framework guida i professionisti della sanità pubblica nella valutazione dei programmi. Si tratta di uno strumento pratico non prescrittivo, ideato per riassumere ed organizzare gli elementi essenziali della valutazione dei programmi. Il framework comprende sia le fasi pratiche della valutazione che gli standard per un'efficace valutazione. Seguendo le fasi e gli standard del framework sarà possibile comprendere bene il contesto di ogni programma e migliorerà sia la percezione che la conduzione della valutazione dei programmi. Inoltre, il framework incoraggia un approccio alla valutazione integrato con le operazioni routinarie degli interventi. L'enfasi è posta su pratiche e strategie progressive di valutazione che includono non soltanto gli esperti della valutazione ma anche tutti i destinatari. La comprensione e l'applicazione degli elementi di questo framework possono rappresentare un impulso a pianificare strategie di sanità pubblica efficaci, migliorare i programmi esistenti e dimostrare i risultati degli investimenti in termini di risorse.

Introduzione

La valutazione dei programmi è pratica organizzativa fondamentale in sanità pubblica (1); ciononostante, la valutazione dei programmi non è eseguita in modo consistente nelle diverse aree in cui si effettuano programmi di sanità pubblica, né è sufficientemente integrata nella gestione - giorno per giorno - della maggior parte dei programmi. La valutazione dei programmi è anche necessaria per aderire pienamente ai principi operativi che i CDC si sono dati per guidare le attività di sanità pubblica, che includono:

- a) usare le conoscenze scientifiche come strumento per decidere e agire nella sanità pubblica;
- b) estendere la ricerca dell'equità sociale attraverso l'intervento di sanità pubblica;
- c) funzionare in modo efficace come un'agenzia di servizi;
- d) indirizzare gli sforzi ai risultati;
- e) assumere responsabilità.

Questi principi operativi suggeriscono varie alternative per migliorare il modo in cui le attività sanitarie sono pianificate e gestite. Sottolineano che per avere buoni programmi, è necessario sviluppare piani espliciti e chiari, incluso la dichiarazione delle *partnership*, e sistemi di feedback che permettono l'apprendimento dall'esperienza e il miglioramento progressivo. Un modo per assicurare il rispetto di questi principi, sia per i nuovi programmi che per quelli già esistenti, è effettuare, di routine, valutazioni pratiche che forniscono informazioni per la corretta gestione dei programmi e ne migliorano l'efficacia.

Questo report presenta un framework per comprendere come si valutano i programmi e facilitare l'integrazione della valutazione nel sistema sanitario. I propositi di questo report sono:

- riassumere gli elementi essenziali della valutazione dei programmi;
- fornire un framework per condurre valutazioni efficaci dei programmi;
- rendere chiare le fasi della valutazione;
- riesaminare gli standard per una valutazione efficace;
- correggere informazioni erranee riguardanti i fini ed i metodi della valutazione.

BACKGROUND

La valutazione è stata definita un'indagine sistematica nel merito, valore o importanza di un oggetto (3,4). Durante gli ultimi tre decenni, la pratica della valutazione si è sviluppata come disciplina con proprie definizioni, metodi, approcci e applicazioni a diverse materie e situazioni (4-7). Nonostante questi miglioramenti, un framework di base organizzativo non era ancora stato sviluppato. Nel maggio 1997, il direttore e lo staff esecutivo dei CDC hanno riconosciuto la necessità di un tale framework e la necessità di combinare la valutazione con la gestione dei programmi. Inoltre è stata posta in rilievo l'esigenza di studi valutativi che dimostrino la relazione tra attività di programma ed efficacia, in termini di prevenzione. I CDC convocarono un gruppo di lavoro incaricato di sviluppare un framework che riassume ed organizza gli elementi di base della valutazione dei programmi.

Procedure per sviluppare il framework

Il gruppo di lavoro per la valutazione, con rappresentanti dai CDC e in collaborazione con operatori della sanità a livello statale e locale, ha cercato input da otto diversi gruppi di referenti, durante la lunga fase di raccolta delle informazioni. Sono stati richiesti contributi a:

- esperti di valutazione,
- manager e staff impegnati nella gestione di programmi di sanità pubblica,
- operatori di sanità pubblica di livello locale e statale,
- direttori di programmi di sanità pubblica non federali,
- rappresentanti e insegnanti di organizzazioni di sanità pubblica,
- ricercatori impegnati nello studio di problemi di salute della comunità,
- rappresentanti del Public Health Service, e staff dei CDC.

Nel febbraio 1998, il gruppo di lavoro ha realizzato un workshop per lo Sviluppo di un Framework per la Valutazione nella Pratica di Sanità Pubblica, cui hanno partecipato, approssimativamente, 90 rappresentanti. Inoltre, il gruppo di lavoro ha condotto interviste a circa 250 persone, ha riesaminato report sulla valutazione pubblicati e non, ha consultato i destinatari di vari programmi circa l'applicazione del framework, e ha mantenuto un sito Web per diffondere documenti e ricevere osservazioni. Nell'Ottobre 1998 è stato condotto un corso nazionale di apprendimento a distanza per la caratterizzazione del framework, attraverso il *Public Health Training Network* dei CDC (8). Il pubblico ha incluso approssimativamente 10.000 professionisti. Queste strategie di condivisione delle informazioni hanno fornito al gruppo di lavoro numerose opportunità per esaminare e rifinire il framework con esperti della sanità pubblica.

Definizione dei concetti chiave

In questo rapporto il termine *programma* è usato per descrivere l'oggetto della valutazione, il quale coincide con ogni azione organizzata di sanità pubblica. Questa definizione è deliberatamente ampia poiché il framework può essere applicato a quasi ogni attività organizzata di sanità pubblica, inclusi interventi che consistono di servizi direttamente gestiti, sforzi di mobilitazione della comunità, attività di ricerca, sistemi di sorveglianza, azioni di sviluppo di politiche sanitarie, investigazione di focolai epidemici, diagnostica di laboratorio, campagne di comunicazione, progetti di costruzione di infrastrutture, servizi di training, aggiornamento ed educativi, e sistemi amministrativi. Il termine *destinatari* indica i committenti della valutazione. La committenza può essere un

vero e proprio mandato cui corrisponde una adeguata remunerazione, oppure il termine di destinatari deve esser inteso in senso più astratto come tutti coloro che sono interessati ai risultati della valutazione, perché a seguito dei risultati possono prendere decisioni o intraprendere azioni adeguate. In altri casi questo termine è stato tradotto come: fruitori o utenti.

Ulteriori termini definiti in questo rapporto sono stati scelti al fine di stabilire un vocabolario comune per la valutazione, ad uso dei professionisti della sanità pubblica.

Integrare la valutazione nella pratica di routine dei programmi

Quando si pone l'accento sulla concretezza e sulla non sporadicità della valutazione, in cui vengono coinvolti, non solo gli esperti della valutazione, ma anche il personale e gli operatori del programma, la valutazione può essere congiunta all'operatività routinaria dei programmi. La valutazione fa allora da complemento alla gestione dei programmi, raccogliendo le informazioni necessarie per migliorare e quantificare l'efficacia pratica degli stessi. I professionisti della sanità pubblica hanno utilizzato di routine, nella pratica, processi di valutazione rispondendo a domande fatte da persone interessate, consultando i collaboratori, ottenendo giudizi basati sul feedback e rifinando gli aspetti operativi del programma (9). Questi processi di valutazione, sebbene informali, sono adeguati per un verificare di continuo il funzionamento dei programmi, al fine di decidere piccoli aggiustamenti nelle funzioni e negli obiettivi. Tuttavia, allorché le possibilità di decisioni potenziali o di modifiche nei e dei programmi aumentano (per esempio al momento di decidere quali servizi offrire in un programma nazionale di promozione della salute), l'impiego di procedure di valutazione esplicite, formali e giustificabili diventa importante.

Assegnare un Valore alle Attività del Programma

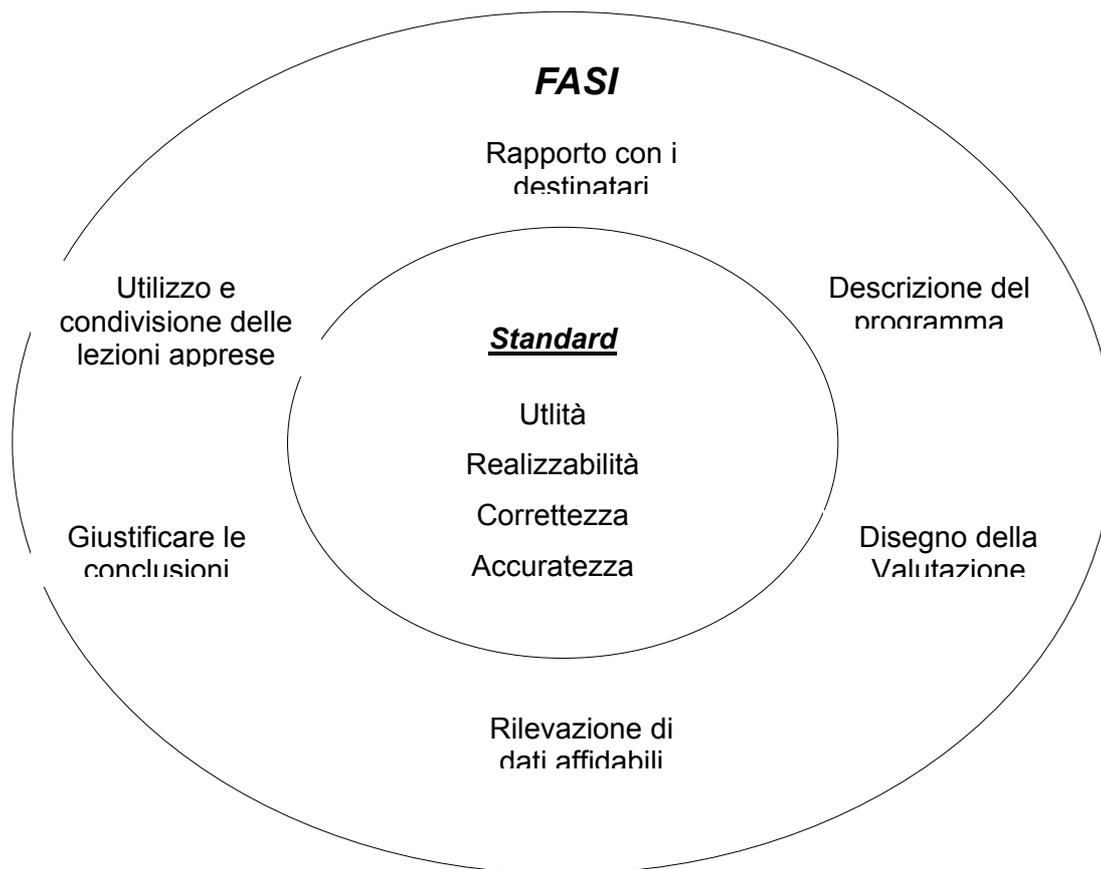
Le domande che riguardano i valori, in contrasto con quelle che riguardano i fatti, includono generalmente tre problemi interrelati: merito (cioè, qualità), valore (cioè, costo-efficacia), significatività (cioè, importanza) (3). Se un programma è giudicato degno di merito, ci si può domandare se, oltre al merito, il programma valga il proprio costo. Ancora, ci si può domandare se programmi di valore contribuiscano a determinare cambiamenti rilevanti. Per assegnare valore ed effettuare giudizi riguardanti un programma sulla base dell'evidenza, bisogna dare risposta alle seguenti domande (3,4,11):

- Cosa sarà valutato? (cioè, in cosa consiste il programma e in quali contesti funziona?)
- Quali aspetti del programma saranno considerati al momento di giudicare le sue prestazioni?
- Quali standard (cioè, tipo o livello di prestazione) devono essere raggiunti affinché il programma sia giudicato riuscito?
- Quale prova (evidenza) sarà usata per indicare come ha funzionato il programma?
- Quali conclusioni, riguardo alle prestazioni dei programmi, sono giustificate comparando l'evidenza disponibile con gli standard selezionati?
- In che modo saranno usate le lezioni apprese dalla valutazione per migliorare l'efficacia pratica della sanità pubblica?

Queste domande andrebbero poste all'inizio di un programma e ripetute durante la messa in pratica. Il framework, descritto in questo rapporto, fornisce un approccio sistematico per rispondere a queste domande.

Framework per i programmi di valutazione nella sanità pubblica

Una valutazione efficace dei programmi è uno strumento sistematico per migliorare le azioni di sanità pubblica utilizzando procedure utili, realizzabili, etiche e accurate. Il framework raccomandato è stato sviluppato per guidare professionisti della sanità nell'usare la valutazione dei programmi. È uno strumento pratico e non prescrittivo, ideato per riassumere e organizzare gli elementi essenziali della valutazione dei programmi. Il framework comprende fasi di pratica della valutazione e standard per una valutazione efficace (figura 1).



Il framework è composto di sei fasi che devono essere utilizzate in ogni valutazione. Sono punti di inizio per indirizzare la valutazione verso un particolare intervento della sanità pubblica, in un determinato arco di tempo. Siccome le fasi sono interdipendenti, potrebbero, in realtà, essere eseguite in una sequenza non lineare; comunque, un ordine esiste – le prime tappe forniscono le fondamenta per il progresso successivo. Per questo, le decisioni che riguardano il modo in cui eseguire una tappa sono iterative e non dovrebbero essere messe in pratica, finché le fasi precedenti non sono state completamente applicate.

Le fasi sono le seguenti:

Fase 1: Coinvolgere i destinatari.

Fase 2: Descrivere il programma.

Fase 3: Mettere a fuoco il disegno della valutazione.

Fase 4: Raccogliere prove credibili.

Fase 5: Giustificare le conclusioni.

Fase 6: Usare i risultati e condividere le lezioni apprese.

Se si seguono queste sei fasi, si facilita la comprensione del contesto del programma (ad esempio, la storia, la messa in pratica e l'organizzazione del programma) e migliora il modo in cui la maggior parte delle valutazioni viene disegnata e condotta.

Il secondo elemento del framework è un insieme di 30 standard, per valutare la qualità della valutazione, organizzati nei seguenti quattro gruppi:

Standard 1: utilità,

Standard 2: realizzabilità

Standard 3: correttezza, e

Standard 4: accuratezza

Questi standard, adottati dal Joint Committee for Educative Evaluation Standards (12), rispondono alla domanda: "questa valutazione sarà efficace?", e sono raccomandati come criteri per giudicare la qualità degli studi di valutazione dei programmi di sanità pubblica. Il resto di questo report discute ogni fase, i suoi sottopunti, e gli standard che governano un'efficace valutazione dei programmi (*Box 1*).

Box 1. Le fasi della valutazione in pratica e standard per una valutazione efficace.

Fasi nella Pratica di Valutazione

- 1) Impegnare i destinatari: Le persone coinvolte o influenzate dal programma e gli utenti primari della valutazione (i decision makers).
- 2) Descrivere il programma: Necessità, effetti auspicati, attività, risorse, stadio di sviluppo, contesto, modello logico.
- 3) Mettere a fuoco il progetto di valutazione: Scopo, utenti, usi, quesiti, metodi, accordi.
- 4) Raccogliere prove credibili: Indicatori, fonti dei dati, qualità, quantità, logistica.
- 5) Giustificare le conclusioni: Standard, analisi/sintesi, interpretazione, giudizio, raccomandazioni.
- 6) Assicurare l'uso e condividere gli insegnamenti: Progetto, preparazione, feedback, follow-up, divulgazione.

Standard per la Valutazione Efficace

- 1) Utilità : Aderisce alle esigenze informative degli utenti prestabiliti.
- 2) Attuabilità: Essere realisti, prudenti, diplomatici, ed economici.
- 3) Correttezza: Comportarsi correttamente sotto il profilo giuridico, etico, e con riguardo verso il benessere di coloro che sono coinvolti e di coloro che ne sono influenzati.
- 4) Accuratezza: Rivelare e rendere note informazioni tecnicamente accurate.

Fasi della valutazione dei programmi

Step 1: i destinatari – per chi (si lavora) si avvia il processo di valutazione

Il ciclo della valutazione comincia con il rapporto con i destinatari (cioè, le persone o le organizzazioni che dovranno investire in ciò che sarà appreso da una valutazione e ciò che sarà eseguito con la conoscenza acquisita). Il lavoro di sanità pubblica include la collaborazione; perciò, ogni valutazione di un programma di sanità pubblica richiede la considerazione dei sistemi di valori dei partner, siano essi strutture, istituzioni, organizzazioni.

I destinatari devono essere coinvolti nell'indagine per assicurare che le loro prospettive siano comprese e tenute nel giusto conto. Quando i destinatari non sono coinvolti, una valutazione potrebbe non trattare elementi importanti degli obiettivi, operazioni, e outcomes del programma. Quindi, i risultati della valutazione potrebbero essere ignorati,

criticati, o, il loro utilizzo ostacolato, perché la valutazione non tiene conto degli interessi e valori dei destinatari (12). Dopo essere stati coinvolti, i destinatari contribuiscono all'esecuzione delle altre fasi. E' fondamentale identificare e stabilire un rapporto impegnativo con i seguenti tre gruppi principali di destinatari:

- a) coloro che sono coinvolti nelle operazioni dei programmi (vale a dire sponsor, collaboratori, partner cooperativi, ispettori di fondi d'investimento, amministratori, manager, e staff);
- b) coloro che sono utilizzati o interessati dal programma (ad esempio clienti, famiglie, organizzazioni di quartiere o vicinato, istituzioni accademiche, autorità, gruppi di difesa dei consumatori o degli ammalati ecc., associazioni professionali, scettici, oppositori, e staff delle organizzazioni collegate o competenti); e
- c) gli utenti primari della valutazione.

Persone Coinvolte nelle Operazioni del Programma.

Le persone o le organizzazioni coinvolte operativamente nel programma sono interessate su come saranno condotte le attività di valutazione, poiché il programma potrebbe dover essere modificato da quanto appreso a seguito della valutazione. Sebbene lo staff, gli ispettori delle agenzie che hanno finanziato l'intervento ed i collaboratori lavorino insieme, non necessariamente rappresentano un unico gruppo d'interessi. I sottogruppi potrebbero avere prospettive differenti e seguire ordini di priorità alternativi; inoltre, siccome costoro detengono un ruolo professionale nel programma, potrebbero sospettare che la valutazione dei programmi sia eseguita con l'intento di giudicarli personalmente. La valutazione dei programmi è collegata alla valutazione del personale, ma deve essere tenuta distinta da quest'ultima, che opera secondo standard differenti (13).

Coloro che sono serviti o influenzati dal programma.

Le persone o le organizzazioni che usufruiscono del programma o direttamente (per esempio, ricevendo servizi) o indirettamente (per esempio, beneficiando di maggiore sostegno da parte della comunità), dovrebbero essere identificati e coinvolti nella valutazione, nella misura del possibile. Sebbene sia naturale impegnare i sostenitori di un programma, può essere meno ovvio che anche coloro che sono apertamente scettici e antagonisti verso il programma potrebbero rappresentare importanti fruitori da impegnare. L'opposizione, infatti, potrebbe dipendere da differenti considerazioni riguardanti gli obiettivi dell'intervento (ovvero quale cambiamento è necessario) o il modo di conseguirli. Aprire una valutazione a prospettive antitetiche e ottenere l'appoggio degli oppositori del programma durante l'inchiesta, potrebbe costituire una scelta prudente, perché con questi sforzi la credibilità della valutazione può essere migliorata.

Utenti Primari della Valutazione

Gli utenti primari della valutazione sono le persone che, per la posizione che occupano, possono eseguire o decidere qualcosa che riguardi il programma. In pratica, gli utenti primari saranno un sottoinsieme di tutti i destinatari identificati. Una valutazione che abbia successo deve individuare, ben presto durante il proprio sviluppo, gli utenti primari, e mantenere frequenti interazioni con loro, cosicché la valutazione rispecchi i loro interessi e soddisfi le loro peculiari esigenze informative (7). La portata e il livello del coinvolgimento dei fruitori varierà per ogni valutazione. Diverse attività riflettono la necessità di impegnare i fruitori (Box 2) (14). Per esempio, i committenti potrebbero essere direttamente coinvolti nel pianificare e condurre la valutazione. Inoltre, si potrebbe

mantenerli informati circa il progresso della valutazione attraverso incontri periodici, rapporti, e altri mezzi di comunicazione. La condivisione delle decisioni e la soluzione dei conflitti aiuta a evitare che sia attribuita eccessiva importanza alle esigenze di uno specifico fruitore (15). Occasionalmente, i fruitori potrebbero approfittare del coinvolgimento nell'attività di valutazione per sabotare, distorcere o screditare il programma. La fiducia tra i fruitori è essenziale; quindi, è necessaria cautela per prevenire il cattivo uso del processo di valutazione.

BOX 2. Impegnare i destinatari

Definizione

Incoraggiare input, partecipazione, e condivisione dei poteri tra coloro che hanno investito nella conduzione della valutazione e dei risultati; è soprattutto importante impegnare gli utenti primari della valutazione, intesi come coloro che hanno responsabilità nel programma e possono prendere decisioni in merito ad esso.

Funzione

Aiuta ad aumentare la probabilità che la valutazione avrà utili ricadute pratiche; può aumentare la credibilità di una valutazione, chiarire ruoli e responsabilità, accrescere la competenza culturale, aiutare a proteggere gli esseri umani, ed evitare reali o presunti conflitti d'interesse.

Esempi di Attività

- ❑ Consultare gli addetti (ad esempio leader, staff, clienti, e fonti di finanziamento del programma) ed esterni (ad esempio, scettici);
- ❑ Fare sforzi particolari per promuovere l'inclusione di gruppi o individui meno potenti;
- ❑ Coordinare l'input dei destinatari in tutto il processo della valutazione: dal progetto, alla messa in opera, all'uso; e
- ❑ Evitare un'identificazione eccessiva, che potrebbe ostacolare la continuazione della valutazione.

Adattato dal Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

Fase 2: Descrivere il Programma

La descrizione comprende la missione e gli obiettivi del programma che viene valutato. La descrizione dovrebbe essere sufficientemente dettagliata, così da assicurare la comprensione degli obiettivi generali e delle strategie del programma. Bisognerebbe prendere in esame la capacità del programma di conseguire i cambiamenti desiderati, lo stadio di sviluppo, e il modo in cui va ad inserirsi nella più ampia organizzazione sociale e nella comunità. Le descrizioni dei programmi definiscono la struttura di riferimento per tutte le decisioni che si dovranno, successivamente, prendere nel processo di valutazione. La descrizione rende possibili paragoni con programmi simili e facilita i tentativi di collegare le componenti del programma ai loro effetti (12). Inoltre, i destinatari della valutazione potrebbero avere idee differenti riguardo agli obiettivi e propositi dei programmi. Valutazioni eseguite senza un accordo comune sulla definizione del programma avranno probabilmente una utilità limitata. A volte, negoziare con i destinatari per giungere ad una descrizione chiara e logica, apporterà benefici prima ancora che i dati necessari a valutare l'efficacia del programma siano disponibili (7).

Aspetti da includere in una descrizione sono: il bisogno (perché?), gli effetti previsti (per ottenere cosa?), le attività, le risorse, lo stadio di sviluppo, il contesto e il modello logico.

BOX 3. Descrivere il programma

Definizione

Scrutinare le caratteristiche del programma che si sta per valutare, incluso il suo scopo e il posto che occupa in un contesto più ampio. La descrizione include le informazioni circa il modo in cui si riteneva dovesse funzionare il programma ed il modo in cui è stato attuato nella realtà. Essa inoltre include le caratteristiche del contesto che potrebbero influenzare le conclusioni sul programma.

Funzione

Aumenta l'imparzialità e l'accuratezza della valutazione; permette una valutazione equilibrata dei punti di forza e di debolezza ed aiuta i destinatari a capire come le caratteristiche del programma si adattano e si collegano ad un contesto più ampio.

Esempi di Attività

- ❑ Caratterizzare la necessità (o insieme di necessità) a cui si rivolge il programma;
- ❑ Elencare le aspettative specifiche: scopi, obiettivi e criteri per la buona riuscita;
- ❑ Chiarire perché le attività di programma sono considerate valide per raggiungere i cambiamenti attesi;
- ❑ Valutare la maturità del programma o lo stadio di sviluppo;
- ❑ Analizzare il contesto all'interno del quale opera il programma;
- ❑ Considerare in che modo il programma è collegato ad altri sforzi in corso; e
- ❑ Evitare la creazione di una descrizione eccessivamente precisa per un programma è in via di sviluppo.

Adattato dal Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs - Thousand Oaks, CA: Sage Publications, 1994

Bisogno.

Un'asserzione sul bisogno serve a descrivere il problema su cui è focalizzato il programma e il modo in cui il programma risponde ad esso. Elementi importanti per la descrizione del bisogno da cui nasce il programma, includono: (a) la natura e l'importanza del problema, (b) quali popolazioni sono interessate, (c) se i bisogni, per cui nasce il programma, stanno cambiando, e (d) in che modo stanno cambiando.

Effetti Previsti.

Le descrizioni dei risultati attesi rendono l'idea di ciò che il programma deve raggiungere perché il suo esito sia considerato positivo (cioè, gli effetti del programma). Gli effetti della maggior parte dei programmi, si rivelano col tempo; perciò le descrizioni delle aspettative dovrebbero essere strutturate nel tempo, procedendo da effetti più specifici (immediati) ad effetti più generali (a lungo termine). Missione, scopo e obiettivi di un programma, rappresentano, a vari livelli di specificità, le aspettative del programma. Bisognerebbe anche fornire previsioni su potenziali conseguenze indesiderate del programma.

Attività.

La descrizione delle attività del programma (quello che il programma compie per ottenere il cambiamento), consente di disporre in sequenza logica le specifiche fasi, strategie o azioni. Ciò dimostra in che modo ogni attività è correlata alle altre e chiarisce il modo di funzionare del programma o la teoria del cambiamento su cui è costruito (16, 17).

Inoltre, le descrizioni dell'attività del programma dovrebbero distinguere le attività di cui il programma è direttamente responsabile da quelle condotte da programmi o da partner ad esso collegati (18).

Dovrebbero essere rilevati anche fattori esterni che potrebbero influire sulla riuscita del programma (ad esempio i trend secolari).

Risorse.

Le risorse comprendono il tempo, le competenze e le risorse umane, la tecnologia, le attrezzature, le informazioni, il denaro e altri beni a disposizione della conduzione delle attività di programma. Le descrizioni delle risorse del programma dovrebbero rendere l'idea del valore e della portata dei servizi del programma e consentire di apprezzare se c'è o no corrispondenza tra le attività che si intende avviare e le risorse disponibili per eseguire tali attività. Inoltre, per le valutazioni economiche è necessario tenere in conto tutte le risorse assegnate e dei costi diretti ed indiretti del programma (19-21).

Stadio di Sviluppo.

I programmi della Sanità Pubblica maturano e cambiano nel tempo; perciò, lo stadio di sviluppo di un programma riflette la sua maturità. I programmi che hanno ricevuto di recente l'autorizzazione e il finanziamento iniziali saranno diversi da quelli che operano di continuo da un decennio. Nel processo di valutazione si dovrebbe considerare il diverso grado di maturità nella pratica del programma (22). Devono essere individuati un minimo di tre stadi di sviluppo: pianificazione, esecuzione, ed effetti. Durante la pianificazione non si possono esaminare le attività di programma, e lo scopo della valutazione è di mettere a fuoco i piani.

Durante l'esecuzione, sono esaminate sul campo e modificate le attività di programma; lo scopo della valutazione è la caratterizzazione delle reali (in quanto opposte a ideali) attività di programma e di migliorare le operazioni, magari rivedendo i piani.

Durante l'ultimo stadio, è trascorso un tempo sufficiente a far emergere gli effetti del programma; lo scopo della valutazione è di identificare e rendere conto sia degli effetti intenzionali sia di quelli non desiderati.

Contesto.

Le descrizioni del contesto del programma dovrebbero includere la collocazione spaziale e le influenze ambientali (ad esempio, storia, geografia, politica, condizioni economiche e sociali, e sforzi delle organizzazioni collegate o competenti) al cui interno opera il programma (6). La comprensione di tale influenza dell'ambiente è necessaria per elaborare una valutazione che tenga conto del contesto, e aiutare i fruitori ad interpretare i risultati in modo accurato e a valutare la generalizzabilità delle scoperte.

Modello Logico.

Un modello logico descrive la sequenza degli eventi necessari per ottenere il cambiamento, sintetizzando i principali elementi del programma in un'immagine di come si pensa debba funzionare il programma (23-35).

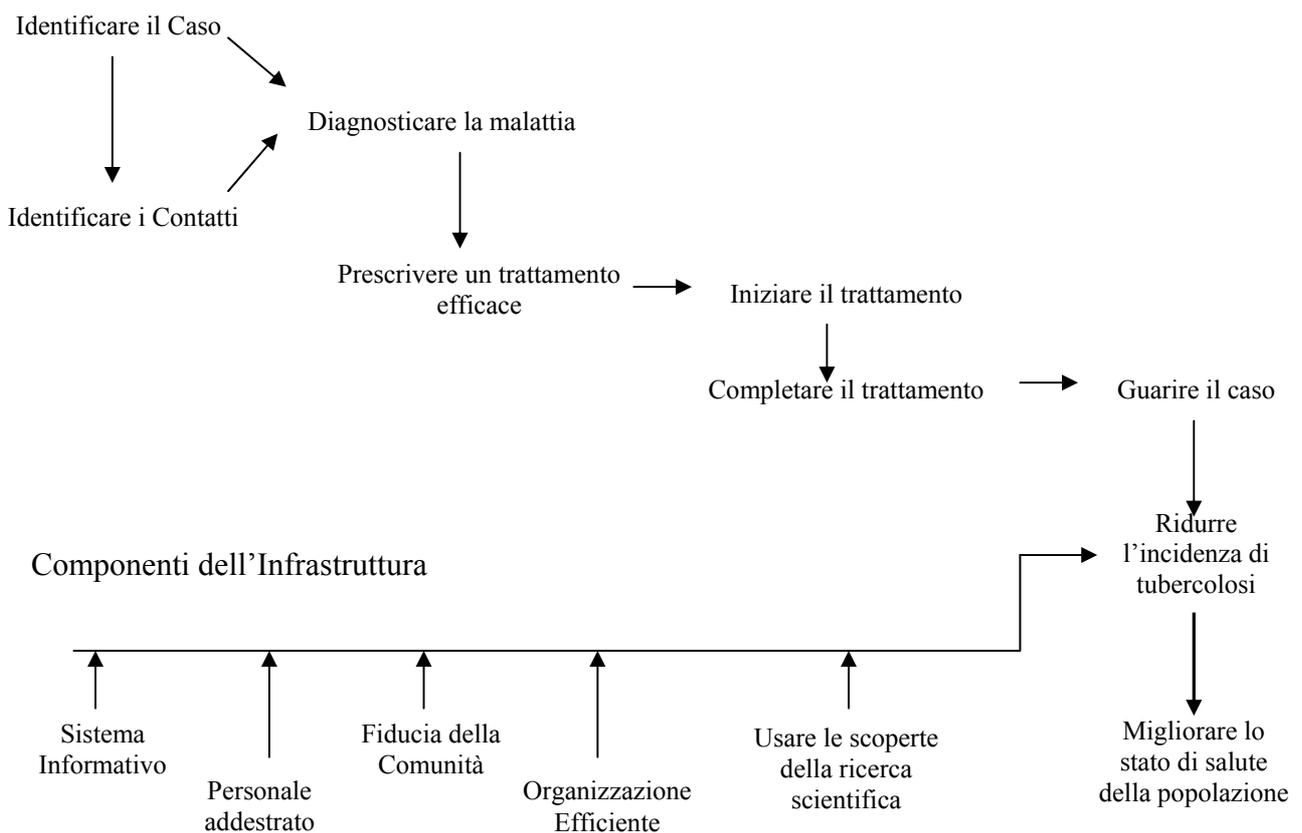
Spesso, tale modello è rappresentato da un diagramma di flusso, una mappa, o una tabella che ritrae la sequenza delle fasi che conducono ai risultati del programma (Fig. 2).

Una delle virtù di un modello logico, è la sua capacità di riassumere il meccanismo generale su cui si basa il programma collegando tra loro i processi (ad esempio la diagnosi di laboratorio delle malattie) ad eventuali effetti (ad esempio la riduzione dell'incidenza della tubercolosi).

Il modello logico può anche mostrare l'infrastruttura di cui si necessita per sostenere le operazioni del programma. Gli elementi correlati in un modello logico possono variare, ma generalmente includono Input (ad esempio staff esperto), Attività (ad esempio l'identificazione dei casi), Output (ad esempio persone che completano il trattamento), e Risultati: dagli effetti immediati (ad esempio, la cura di persone malate), a quelli intermedi (ad esempio, la riduzione del tasso di tubercolosi), a quelli a lungo termine (ad esempio, il miglioramento delle condizioni di salute della popolazione). La creazione di un modello logico permette ai committenti di aver chiare le strategie del programma; perciò, il modello logico migliora ed inquadra meglio il percorso e la direzione del programma. Esso, tra l'altro, mostra le assunzioni riguardo le condizioni in cui programma funziona e fornisce uno schema di riferimento per una o più valutazioni del programma. Un modello logico dettagliato può anche rinforzare l'esigenza di causalità e rappresentare una base per stimare gli effetti del programma ad end point, non misurati direttamente, ma legati in una catena causale messa in luce dalla ricerca precedente (35). Famiglie di modelli logici possono essere create per mostrare un programma a diversi livelli di specificità, o per diversi destinatari.

Fig.2 Logica del modello di un programma di controllo della tubercolosi

Componenti del Servizio



Le descrizioni del programma variano per ogni valutazione, e varie attività riflettono l'esigenza di descrivere il programma (ad esempio, l'uso di fonti multiple d'informazioni per costruire una descrizione appropriata) (Box* 3). L'accuratezza di una descrizione di programma può essere confermata attraverso la consultazione di vari destinatari della

valutazione, e la descrizione dell'attività può essere controllata alla luce dell'osservazione diretta delle attività sul campo. Una descrizione stringata può essere migliorata focalizzando alcuni fattori quali il turnover dello staff, l' inadeguatezza delle risorse, le pressioni politiche, oppure la forte partecipazione della comunità che potrebbero influenzarne la performance.

Fase 3: Mettere a fuoco il disegno della Valutazione

Per usare il tempo e le risorse nel modo più efficiente possibile, la valutazione deve essere concentrata sui problemi di grande interesse per i committenti ed i destinatari (7, 36, 37). Non tutti i possibili progetti sono ugualmente adeguati per rispondere alle necessità di informazione dei committenti.

A raccolta dei dati iniziata, un cambiamento di procedure potrebbe essere difficile o impossibile, anche se è ovvio che bisogna sempre sforzarsi di migliorare i metodi. Un piano accurato è in grado di prevedere gli usi prestabiliti e crea una strategia di valutazione che ha la maggiore probabilità di mostrarsi utile, realizzabile, etica ed accurata. Tra i temi da considerare, allorché si progetta una valutazione, vi sono il proposito, gli utenti, gli usi, i quesiti, i metodi, e i consensi.

Il proposito.

Definendo il proposito di una valutazione (cioè che cosa si intende fare), si evita di prendere decisioni affrettate sul modo di condurre la valutazione.

Le caratteristiche del programma, soprattutto lo stadio di sviluppo ed il contesto, influenzeranno il proposito della valutazione. Le valutazioni di sanità pubblica hanno quattro propositi generali (Box* 4).

Il primo è di aumentare l'*insight*, il che avviene, per esempio, quando si valuta la realizzabilità di un approccio innovativo alla pratica. La conoscenza acquisita, tramite la valutazione, sulla praticità del nuovo approccio, porta poi a decidere se esso può essere utilizzato per includerlo in un programma di cui poi si verificherà l'efficacia. Per un programma in via di sviluppo, informazioni da precedenti valutazioni possono fornire l'insight necessaria a chiarire il modo in cui le sue attività debbano essere progettate per apportare i cambiamenti desiderati.

Un secondo proposito della valutazione dei programmi è di cambiare la pratica, il che è appropriato nello stadio dell'esecuzione quando si cerca di descrivere cosa ha fatto un definito programma, e in quale misura. Tale informazione può essere utilizzata per descrivere meglio le procedure, per migliorare il modo in cui il programma opera, e per mettere a punto bene la strategia generale del programma. Le valutazioni fatte con questo proposito comprendono gli sforzi per migliorare la qualità, l'efficacia o efficienza delle attività del programma.

Un terzo proposito della valutazione è di valutare gli effetti. Le valutazioni fatte con questo proposito esaminano il rapporto tra le attività del programma e le conseguenze osservate. Questo tipo di valutazione si addice ai programmi maturi per cui è possibile definire quali interventi sono stati effettuati e quale proporzione della popolazione bersaglio è stata coinvolta. Conoscere gli effetti potenziali del programma e sapere metterli in evidenza può assicurare che le conseguenze significative non siano trascurate. Un insieme di effetti può provenire da diretto rapporto causa-effetto con il programma. Se questo è il caso, si potrebbero ricercare le prove per attribuire tali effetti esclusivamente al programma. In altri casi, gli effetti osservati potrebbero essere solo in parte attribuiti al programma, e potrebbero essere frutto del contributo di diversi fattori. Ad esempio, se le attività di un programma sono parallele a quelle di altri interventi che hanno operato nello stesso ambiente, alcuni effetti (ad esempio, la creazione di nuove politiche o leggi) non possono essere attribuibili esclusivamente ad un programma o ad un altro. In questo tipo di

situazioni, lo scopo della valutazione è di raccogliere prove credibili che descrivano il contributo di ciascun programma nel comune sforzo al cambiamento. Solo se ci è capaci di condurre valutazioni che misurino entrambi i contributi relativi dei programmi, si potrà asserire se e quanto quel programma è responsabile di quei risultati.

Un quarto proposito, possibile ad ogni stadio di sviluppo del programma, è l'uso del processo di valutazione per interessare coloro che partecipano nella ricerca. La riflessione logica e sistematica necessaria alla ricerca, può funzionare come un catalizzatore di un cambiamento autogestito per i partecipanti al programma (e al processo di valutazione). Si può iniziare una valutazione con l'intento di influenzare positivamente i fruitori. Tali influenze possono essere: integrare l'intervento del programma (ad esempio, utilizzando un questionario di follow up per rafforzare i messaggi del programma); rafforzare i partecipanti al programma (ad esempio, aumentando un certo senso di controllo del cliente sulla direzione del programma); promuovere la crescita culturale dello staff (ad esempio, insegnando come raccogliere, analizzare, e interpretare le evidenze); contribuire alla crescita organizzativa (ad esempio, chiarendo il modo in cui il programma si relaziona alla missione dell'organizzazione); o facilitare la trasformazione sociale (ad esempio, far avanzare la lotta della comunità per l'auto-determinazione) (7, 38-42).

BOX 4. Usi della valutazione nella pratica della sanità pubblica, per tipo di finalità

1. Ricavare una intuizione dei problemi (insight)

- Valutare i bisogni, le aspettative ed i desideri dei membri della comunità.
- Identificare gli ostacoli e le agevolazioni per l'uso del servizio.
- Imparare a descrivere e a misurare le attività di programma e gli effetti.

2. Cambiare la pratica

- Mettere a punto dei piani per introdurre un nuovo servizio.
- Caratterizzare il grado di esecuzione dei piani d'intervento.
- Migliorare il contenuto dei materiali educativi.
- Accrescere la competenza culturale del programma.
- Verificare che i diritti dei partecipanti siano tutelati.
- Stabilire le priorità per l'addestramento del personale.
- Fare modifiche intermedie per migliorare il flusso paziente/cliente.
- Migliorare la chiarezza dei messaggi di comunicazione sulla salute.
- Mobilitare il sostegno della comunità al programma.
- Valutare gli effetti
- Valutare lo sviluppo delle abilità dei partecipanti al programma.
- Confrontare i cambiamenti nel comportamento dei fornitori nel tempo.
- Comparare i costi con i benefici.
- Individuare quali partecipanti agiscono bene nel programma.
- Decidere dove collocare nuove risorse.
- Documentare il livello di riuscita nel raggiungimento degli obiettivi.
- Dimostrare che le richieste, sul piano delle responsabilità, sono adempiute.
- Aggregare informazioni da diverse valutazioni per stimare gli effetti, in termini di outcome, per programmi simili.
- Raccogliere storie riuscite.

3. Influenzare i partecipanti

- Rafforzare i messaggi d'intervento.
- Stimolare al dialogo e suscitare la consapevolezza riguardo ai problemi di salute.
- Allargare l'area del consenso tra i membri della coalizione circa gli scopi del programma.
- Insegnare le abilità di valutazione allo staff e ad altri destinatari.
- Sostenere il cambiamento e lo sviluppo organizzativo.

Utenti

Gli utenti sono le specifiche persone che riceveranno i risultati della valutazione.

Trattandosi di coloro che poi fanno esperienza diretta dei risultati del progetto, essi dovrebbero prendere parte alla scelta del punto centrale della valutazione (7).

La partecipazione dell'utente è necessaria per mettere in chiaro quale uso si vuole fare dei risultati (a che serve), stabilire un ordine di priorità dei problemi e dei metodi, e impedire che la valutazione sia fuorviata o diventi irrilevante.

Usi

Gli usi sono i modi specifici in cui saranno applicate le informazioni generate dalla valutazione. Esistono diversi possibili usi della valutazione di programma (Box 4). Parlare degli usi in termini vaghi potrebbe compiacere molti fruitori, ma aumenta le possibilità che la valutazione non centri i bisogni di alcuno. Gli usi dovrebbero essere pianificati e ordinati per priorità con il contributo degli utenti e con riguardo per lo stadio di sviluppo del programma ed il contesto in cui si trova. Tutti gli usi devono essere collegati ad uno o più utenti specifici.

Quesiti

Sono le domande che i fruitori si pongono sul funzionamento di un programma. I quesiti stabiliscono i confini della valutazione, definendo quali aspetti del programma saranno oggetto della valutazione (5-7). Mettendo a punto i quesiti si spronano i committenti (i fruitori) a rivelare a cosa, secondo loro, la valutazione debba rispondere. Inoltre, negoziando e discutendo l'ordine di importanza dei quesiti dei committenti è possibile migliorare la comprensione di quale sia il punto centrale della valutazione. La fase dello sviluppo dei quesiti potrebbe anche far venire alla luce possibili opinioni divergenti dei committenti su quale sia la migliore unità di analisi. Alcuni committenti potrebbero voler studiare in che modo diversi programmi operano sinergicamente, come un sistema integrato di interventi, per ottenere il cambiamento all'interno di una comunità. Altri committenti potrebbero avere quesiti riguardanti la performance di un singolo programma o di un progetto locale all'interno di un programma. Altri ancora potrebbero volersi concentrare su specifici sottoelementi o procedure di un progetto. Decisioni chiare circa i quesiti e le corrispondenti unità d'analisi prese all'inizio, guidano la scelta del metodo e la raccolta delle prove, nelle fasi successive della valutazione.

Metodi

I metodi per la valutazione sono forniti dalla metodologia della ricerca scientifica, in particolar modo dall'insieme dei metodi elaborati nelle scienze sociali, comportamentali, e sanitarie (5-7,43-48). Una classificazione dei tipi di disegno include progetti sperimentali, quasi-sperimentali, e osservazionali (43-48).

In nessun caso un disegno è migliore di un altro. I metodi di valutazione dovrebbero essere selezionati per fornire le informazioni appropriate per rispondere ai quesiti dei fruitori (vale a dire che i metodi dovrebbero essere appaiati agli utenti primari, agli usi che bisogna fare dei risultati ed ai quesiti). I progetti sperimentali usano l'assegnazione casuale per confrontare gli effetti di un intervento (in un gruppo) con gruppi altrimenti equivalenti (49). I metodi quasi-sperimentali comparano i gruppi non equivalenti (ad esempio, quello formato dai partecipanti al programma contro il gruppo delle persone in lista d'attesa) oppure utilizzano serie di dati riguardanti lo stesso gruppo, osservato nel tempo, in periodi in cui l'intervento è effettuato e periodi senza l'intervento, per stabilire

un paragone (un esempio è quello delle serie temporali interrotte) (50-51). I metodi osservazionali utilizzano confronti all'interno di un gruppo per spiegare le caratteristiche specifiche dei suoi membri (ad esempio, studi comparativi di casi o studi trasversali) (45, 52-54). La scelta del disegno ha conseguenze su cosa emergerà come evidenza, il modo in cui quella evidenza sarà raccolta, e che genere di conclusione si può trarre (incluso la validità interna ed esterna delle conclusioni) (55).

Inoltre, le decisioni metodologiche chiariscono il modo in cui opererà la valutazione (ad esempio, in quale misura i partecipanti al programma saranno coinvolti; il modo in cui le fonti d'informazioni saranno selezionate; quali strumenti per la raccolta dati saranno utilizzati; chi raccoglierà i dati; quali sistemi di gestione dei dati si renderanno necessari; e quali sono i metodi di analisi, sintesi, interpretazione, e presentazione più appropriati). Poiché ogni metodo ha le sue limitazioni e comporta specifici tipi di bias, le valutazioni che utilizzano diversi metodi sono generalmente più efficaci (44,56-58). Nel corso di una valutazione, potrebbe essere necessario rivedere o modificare i metodi. Inoltre, possono mutare le circostanze che rendono credibile e utile un particolare tipo di approccio. Per esempio, l'uso che si prevede di fare della valutazione può slittare dal migliorare l'operatività corrente del programma allo stabilire se è il caso di estendere i servizi del programma ad un nuovo gruppo di popolazione. In tal modo, il cambiamento delle condizioni potrebbe richiedere la modifica o la riprogettazione iterativa dei metodi, per tenere sul giusto binario la valutazione (22).

Accordi

Gli accordi riassumono le procedure e chiariscono i ruoli e le responsabilità tra coloro che eseguiranno il piano di valutazione (6,12).

Gli accordi descrivono come il piano di valutazione sarà eseguito utilizzando le risorse disponibili (ad esempio, denaro, personale, tempo, ed informazioni) (36-37). Gli accordi inoltre stabiliscono quali salvaguardie sono predisposte per proteggere i soggetti umani e, quando appropriato, quali approvazioni etiche (ad esempio, comitato etico dell'istituzione) o amministrative (ad esempio, la redazione del rapporto) devono essere ottenute (59-60). Gli elementi di un accordo includono dichiarazioni concernenti il proposito, gli utenti, gli usi, i quesiti, e i metodi, assieme alla consegna, periodo di tempo, e budget. L'accordo può includere tutti i committenti impegnati, ma come minimo, deve coinvolgere gli utenti primari, i fornitori di materiali o risorse finanziarie, e quelle persone che condurranno la valutazione e faciliteranno il suo uso e la divulgazione. Il grado di formalizzazione di un accordo può variare secondo le relazioni esistenti tra i committenti. Un accordo può essere un contratto legale, un protocollo dettagliato, o un memorandum di intesa. La creazione di un accordo esplicito costituisce una verifica dell'intesa reciproca, necessaria per la riuscita di una valutazione. L'accordo scritto è anche una base per modificare o rinegoziare le procedure se necessario. Varie attività riflettono la necessità di mettere bene a fuoco il progetto di valutazione (Box 5). Sia i sostenitori sia gli scettici verso il programma possono essere consultati per garantire che i quesiti di valutazione proposti siano politicamente consistenti (cioè rispondenti alle varie posizioni dei diversi gruppi di interesse).

È possibile far circolare tra i fruitori un elenco degli usi potenziali della valutazione, appropriati allo stadio di sviluppo del programma e al contesto, per stabilire qual è il più interessante. Si possono intervistare specifici utenti predefiniti, per comprendere meglio i loro bisogni informativi e l'intervallo temporale per l'azione. La quantità di risorse necessarie potrebbe ridursi se gli utenti vogliono impiegare metodi di valutazione più rapidi ma meno precisi.

BOX 5. Centrare il disegno della valutazione

Definizione

Pianificare in anticipo verso quali obiettivi dirigere la valutazione e quali fasi saranno intraprese; il processo è iterativo (cioè, continua fin quando non viene individuato un preciso approccio che risponda ai quesiti della valutazione, con metodi approvati dai committenti, in quanto ritenuti utili, realizzabili, etici ed accurati); i quesiti ed il metodo della valutazione potrebbero essere modificati per ottenere un adattamento ottimale che faciliti l'uso da parte degli utenti primari.

Funzione

Fornisce l'investimento in qualità; aumenta le possibilità che la valutazione abbia una buona riuscita, identificando procedure che siano pratiche, che siano politicamente fruttuose, ed efficaci rispetto al costo; il fallimento nella pianificazione può essere autolesionista in quanto può portare ad una valutazione che potrebbe essere impraticabile o inutile; quando i committenti sono d'accordo su un centro d'interesse del progetto, esso è utilizzato per tutto il processo di valutazione per fare in modo che il progetto continui per la giusta strada.

Esempi di attività:

- ❑ Incontrarsi con i destinatari per chiarire l'intento o il proposito della valutazione;
- ❑ imparare quali persone sono nella posizione di usare effettivamente i risultati, per orientare il piano facendo in modo che incontri le loro necessità;
- ❑ Capire come verranno usati i risultati della valutazione;
- ❑ Stilare espliciti quesiti di valutazione a cui dare delle risposte;
- ❑ Descrivere metodi pratici per campionamento, raccolta dati, analisi dei dati, interpretazione, e giudizio;
- ❑ Preparare un protocollo scritto o un accordo che riassume le procedure di valutazione, con ruoli chiari e responsabilità per tutti i committenti; e
- ❑ Rivedere tutto o parte del piano di valutazione al cambiare delle circostanze cruciali.

Adattato dal Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and b) U.S. General Accounting Office. Designing evaluations. Washington, DC: U.S. General Accounting Office, 1991; publication no. GAO/PEMD-10.1.4.

Fase 4: Raccogliere Prove Credibili.

Una valutazione dovrebbe sforzarsi di raccogliere le informazioni che possono rendere un'immagine completa del programma, cosicché l'informazione risulti credibile agli occhi degli utenti primari della valutazione. I fruitori dovrebbero poter percepire che l'informazione (vale a dire le evidenze, le prove) acquisita è affidabile e rilevante per rispondere ai loro quesiti. Se sarà l'informazione sarà ritenuta credibile o no dipende dai quesiti che sono stati posti e dai motivi che hanno portato a tali quesiti. Per alcuni quesiti, il criterio di credibilità del fruitore potrebbe richiedere l'acquisizione di risultati frutto di un esperimento controllato; mentre per un altro quesito, un insieme di osservazioni sistematiche (ad esempio, interazioni tra un osservatore esterno e gli abitanti della comunità) costituirebbe l'evidenza dotata del maggior grado di credibilità. In alcune situazioni, potrebbe rendersi necessaria la consultazione di specialisti in metodologia della valutazione. Per esempio quando l'interesse per la qualità dei dati è alto o vi sono delle serie conseguenze associate con il commettere errori di conclusione (cioè, concludere che gli effetti di programma esistono quando non ve n'è alcuno, oppure concludere che il programma non ha alcun effetto quando invece vi è l'effetto c'è, o attribuire degli effetti ad un programma che non è stato adeguatamente eseguito) (61,62).

Evidenze credibili rafforzano i giudizi conclusivi della valutazione e le raccomandazioni che ne seguono. Sebbene qualsiasi tipo di dati abbia delle limitazioni, si può accrescere la credibilità generale di una valutazione utilizzando procedure multiple per la raccolta, l'analisi, e l'interpretazione dei dati. Incoraggiare la partecipazione dei committenti può anche aumentare in costoro la fiducia e la credibilità. Quando i committenti sono impegnati nella definizione e nella raccolta di dati che essi stessi ritengono credibili, saranno più propensi ad accettare le conclusioni di una valutazione e ad agire sulla scorta delle sue raccomandazioni (7,38).

Gli aspetti della raccolta delle evidenze che tipicamente influenzano la percezione di credibilità includono gli indicatori, le fonti, la qualità, la quantità, e la logistica.

Indicatori

Gli indicatori definiscono gli attributi di un programma che riguardano il punto centrale ed i quesiti della valutazione (63-66). Siccome gli indicatori traducono in misure specifiche, dotate di significato, concetti che riguardano il programma, il suo contesto e i suoi effetti previsti, essi forniscono una base per raccogliere evidenze, che possano essere ritenute valide e riproducibili, per gli usi a cui la valutazione è rivolta. Gli indicatori si riferiscono ai criteri che saranno usati nel giudicare il programma (66-70). Esempi di indicatori includono misure delle attività del programma (per esempio, la capacità del programma di offrire servizi; il tasso di partecipazione; livelli di soddisfazione del cliente; l'efficienza dell'uso delle risorse; e l'ammontare dell'esposizione all'intervento) e misure degli effetti del programma (per esempio: cambiamenti del comportamento dei partecipanti, delle norme, politiche e pratiche della comunità, dello stato di salute, qualità della vita oppure le condizioni o l'ambiente che circondano il programma). Definire troppi indicatori può distrarre dagli obiettivi della valutazione; comunque, diversi indicatori sono necessari per seguire l'esecuzione e gli effetti di un programma. Un approccio allo sviluppo di indicatori multipli è basato sul modello logico del programma (sviluppato nella seconda fase della valutazione). Il modello logico può essere utilizzato per sagomare uno spettro di indicatori, che dalle attività di programma conducano agli effetti attesi (23,29-35). Per ogni fase del modello si possono sviluppare indicatori quali-quantitativi adattati al concetto in questione, all'informazione disponibile, ed alle utilizzazioni previste dei dati (ricadute pratiche). Collegare gli indicatori al modello logico consente di rintracciare piccoli cambiamenti del funzionamento più rapidamente che se, al contrario, fosse adottata un'unica misura di outcome.

Attraverso questo approccio, sono inoltre precisate le linee di responsabilità, in quanto per ogni fase strategica del programma c'è un indicatore. In più, questo approccio offre un'ampia base di misure che rivelano come gli outcome sanitari sono conseguenza di effetti intermedi del programma.

Anche fattori intangibili (ad es. la qualità del servizio, le capacità della comunità [71], o le relazioni inter-organizzazioni), che pure influiscono sul programma, possono essere misurati, registrando sistematicamente markers di cosa vien detto e cosa vien fatto, quando il concetto è espresso (72,73).

Durante una valutazione, potrebbe essere necessaria la modifica di precedenti indicatori o l'adozione di nuovi. Misurare le performance del programma tramite indicatori è soltanto una parte della valutazione e non deve essere confuso come l'unica base per prendere decisioni. È ben documentato che si creano problemi quando gli indicatori di performance sostituiscono un processo di valutazione completo e la ricerca di conclusioni pienamente giustificate (66, 67,74). Un indicatore (ad esempio, un tasso

di malattia in aumento) potrebbe essere interpretato come il riflesso del fallimento di un programma, quando, in realtà l'indicatore potrebbe essere stato influenzato dal cambiamento di condizioni esterne, non controllabili del programma.

Fonti

Le fonti di prove in una valutazione sono le persone, i documenti, o le osservazioni che forniscono informazioni per la ricerca (box 6). Per ogni indicatore da misurare, può essere utilizzata più di una fonte per raccogliere prove. Selezionare fonti multiple, fornisce l'opportunità di includere prospettive differenti riguardo al programma, accrescendo l'affidabilità della valutazione.

BOX 6. Fonti delle informazioni in una valutazione

1. Persone:

- Clienti, partecipanti al programma, non partecipanti;
- Staff, manager del programma, amministratori;
- Utenza generale;
- Informatori chiave;
- Finanziatori;
- Critici/scettici;
- Staff di altre agenzie;
- Rappresentanti dei gruppi di difesa dei cittadini;
- legislatori, progettatori delle linee guida;
- ufficiali della sanità locale e statale.

2. Documenti

- Proposte di finanziamento, newsletter, servizi stampa;
- Brevi riunioni, schede amministrative, moduli o schede di registrazione/iscrizione;
- Materiali pubblicitari, rapporti trimestrali;
- Pubblicazioni, articoli scientifici, poster;
- Precedenti rapporti di valutazione;
- Stime delle risorse e delle esisgenze;
- Rapporti dalla sorveglianza;
- Database esistenti;
- Registrazioni eseguite da addetti delle agenzie finanziatrici o collaboratori;
- Pagine Internet; e
- Grafici, mappe, diagrammi, fotografie, registrazioni video.
- Osservazioni
- Riunioni, eventi speciali /attività, incontri di servizio.

Adattato da Taylor-Powell E, Rossing B, Geran J. Evaluating collaboratives: reaching the potential. Madison, WI: University of Wisconsin Cooperative Extension, 1998.

Una prospettiva "dall'interno" può essere compresa dai documenti interni e dai commenti dello staff o del manager del programma, mentre clienti, osservatori neutrali, o quelli che non sostengono il programma potrebbero fornire una prospettiva diversa, ma ugualmente rilevante. Contemperare queste ed altre prospettive, dà una visione più ampia del programma.

I criteri usati per la selezione delle fonti dovrebbero essere dichiarati apertamente così che gli utenti ed altri fruitori siano messi in grado di interpretare accuratamente le

prove e valutare se esse siano distorte (45,75-77). Inoltre, alcune fonti hanno forma narrativa, altre numerica. L'integrazione di informazioni qualitative e quantitative può aumentare le possibilità che le prove siano ben equilibrate, e perciò soddisfacenti i bisogni e le aspettative dei diversi utenti (43, 45, 56, 57, 78-80). Infine, in alcuni casi, si possono utilizzare più valutazioni separate, come fonti, per condurre a una sintesi più ampia (58, 81, 82).

Qualità

La qualità si riferisce all'appropriatezza e all'integrità delle informazioni utilizzate in una valutazione. Dati di alta qualità sono riproducibili, validi, ed informativi per quanto riguarda il loro utilizzo previsto. Una buona definizione degli indicatori facilita la raccolta di dati di qualità. Altri fattori che influenzano la qualità includono il progetto dello strumento di misurazione (ad es. il questionario), le procedure per la raccolta dati, l'addestramento degli addetti alla raccolta dati, la scelta delle fonti, codifica, gestione dati, e controllo di routine degli errori. Ottenere dati di qualità obbligherà a fare delle scelte (per esempio, numerosità del campione contro quantità di informazioni, ovvero più rappresentatività contro più profondità) che dovrebbero essere negoziate con e tra i fruitori. Visto che tutti i dati hanno limiti, l'intento di una valutazione pratica è di sforzarsi per ottenere un livello di qualità che si adatti al tipo di decisioni che si devono prendere ed alla soglia di credibilità dei fruitori.

Quantità

La quantità si riferisce all'ammontare di prove che bisogna raccogliere nel corso della valutazione. La somma di informazioni richieste dovrebbe essere stabilita in anticipo oppure, laddove siano utilizzati processi in evoluzione, dovrebbero essere stabiliti dei criteri per decidere quando fermare la raccolta dei dati. La quantità influenza il potenziale livello di fiducia o la precisione statistica delle conclusioni della valutazione. Inoltre, essa determina in parte se l'evidenza avrà una potenza statistica sufficiente a rivelare gli effetti (83). Si dovrebbe sapere in anticipo, il più chiaramente possibile, quale sarà l'utilizzo pratico di ogni prova raccolta. D'altra parte, i rispondenti dovrebbero avere il minimo fastidio possibile.

Logistica

La logistica comprende i metodi, la tempistica, e le infrastrutture fisiche per la raccolta e il trattamento delle prove. Ogni tecnica selezionata per la raccolta delle prove (Box 7) deve essere appropriata alla/e fonte/i, al piano d'analisi, e alla strategia di comunicazione dei risultati.

Box 7. Tecniche per raccogliere le informazioni

- ❑ Survey con questionari (es. distribuiti a mano, per telefono, fax, posta, posta elettronica, o Internet);
- ❑ Intervista personale (es., individui o gruppi; strutturata semi strutturata, o conversazionale);
- ❑ Osservazione;
- ❑ Analisi dei documenti;
- ❑ Case Study;
- ❑ Discussioni di gruppo (es. brainstorming o gruppi nominali, una riunione strutturata condotta per elicitar ed ordinare le priorità, gli obiettivi o per identificare i problemi);
- ❑ Gioco di ruolo, drammatizzazione;
- ❑ Peer Review o Expert Review (verifica tra pari, o verifica con un esperto);

- ❑ Testimonials;
- ❑ Tecniche dei Differenziali semantici, confronti appaiati , test di similarità o dissimilarità;
- ❑ Costruzione di Scenari ipotetici;
- ❑ Narrazione di storie;
- ❑ Cartografia geografica;
- ❑ Cartografia concettuale;
- ❑ Pile sorting *letteralmente* “mettere in ordine il mucchio” (cioè, una tecnica che permette ai rispondenti di classificare liberamente gli item, rivelando il modo in cui essi percepiscono la struttura di un argomento);
- ❑ Free Listing (una tecnica per elicitare una lista completa di tutti gli argomenti in un area culturale);
- ❑ Diagramma della rete sociale
- ❑ Simulazione, costruzione di modelli;
- ❑ Contabilità;
- ❑ Fotografia, disegno, arte videografia;
- ❑ Diari o riviste; e
- ❑ Registri, moduli di attività.

Adattato da: (a) Taylor-Powell E, Rossing B, Geran J. *Evaluating collaboratives: reaching the potential*. Madison, WI: University of Wisconsin Cooperative Extension, 1998; (b) Phillips JJ. *Handbook of training evaluation and measurement methods*. 3rd ed. Houston, TX: Gulf Publishing Company, 1997; c) Weller SC. *Systematic data collection*. Thousand Oaks, CA: Sage Publications, Inc. 1988; and (d) Trochim WMK. *Introduction to concept mapping for planning and evaluation*. Available at <http://trochim.human.corne.edu/research/epp1/epp1.htm>. Accessed July 1999.

Le persone e le organizzazioni hanno preferenze culturali che dettano quali modi di porre le domande e di raccogliere informazioni sono accettabili, incluso la persona che può essere considerata appropriata a porre i quesiti. Per esempio, alcuni partecipanti potrebbero essere più disponibili a discutere il loro comportamento sanitario con un estraneo, laddove altri sono più a loro agio con qualcuno che conoscono. Le procedure per raccogliere prove in una valutazione (Box 8) devono essere adattate alle condizioni culturali dell’ambiente e studiate per assicurare che la privacy e la confidenzialità delle informazioni e delle fonti siano protette (59, 60, 84).

BOX 8. Raccogliere evidenze credibili

Definizione

Elencare le informazioni che i destinatari percepiscono come affidabili e rilevanti per trovare soluzioni ai loro problemi.

Tale evidenza può essere di tipo sperimentale, oppure osservazionale, qualitativa o quantitativa, oppure può includere un mix di metodi. Dati adeguati potrebbero essere già disponibili e facilmente accessibili, oppure si dovrebbe definire quale è il bisogno informativo e raccogliere nuovi dati. Un insieme di prove è più o meno affidabile, per i destinatari, a seconda di come erano stati poste le domande principali cui si intendeva dare una risposta), a seconda delle fonti di informazione, delle condizioni in cui sono stati raccolti i dati, della riproducibilità delle misure, della validità delle interpretazioni, e delle procedure di controllo di qualità.

Funzione

Aumenta l’utilità e l’accuratezza della valutazione; guida la scelta delle informazioni da raccogliere e da la priorità alle fonti di informazioni più difendibili; promuove la raccolta sistematica di informazioni valide e riproducibili, e ciò è a fondamento di ogni valutazione efficace.

Esempi di Attività

- ❑ Scegliere indicatori dotati di significato per i problemi della valutazione;
- ❑ Descrivere in modo completo gli attributi delle fonti di informazione ed il rationale che sta alla base della loro scelta;
- ❑ Stabilire procedure chiare ed addestrare lo staff a rilevare informazione di elevata qualità;
- ❑ Monitorare periodicamente la qualità dell'informazione raccolta e fare sforzi per migliorare la qualità;
- ❑ Stimare in anticipo l'ammontare di informazione necessaria oppure stabilire dei criteri per decidere quando mettere fine alla rilevazione, in situazioni in cui viene messo in atto un processo iterativo o progressivo; e
- ❑ Salvaguardare la confidentialità dell'informazione e delle fonti di informazioni.

Adapted from Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

Fase 5: Giustificare le Conclusioni.

Le conclusioni della valutazione sono giustificate quando sono collegate alle prove raccolte e giudicate tenendo conto di valori o standard stabiliti dai fruitori e/o committenti. E' necessario che i committenti concordino sulla giustezza della valutazione prima di poter utilizzare – sentendosi sicuri - i risultati della valutazione. Per giustificare le conclusioni sulla scorta delle prove, bisogna prendere in esame gli standard, l'analisi e la sintesi, l'interpretazione, il giudizio, e le raccomandazioni.

Standard

Gli standard rispecchiano i valori dei committenti, valori che forniscono la base per formulare giudizi riguardo l'esecuzione del programma. L'uso di standard espliciti distingue la valutazione da altri approcci al management strategico in cui le priorità sono stabilite senza riferimento a valori espliciti. In pratica, quando i committenti articolano e negoziano i loro valori, questi diventano gli standard per giudicare se una certa esecuzione di un programma sarà, ad esempio, considerata mal riuscita, ben riuscita o adeguata. Un insieme di sistemi di valore che potrebbero servire come fonti di standard sono riportati nel Box 9. Quando messi in opera, questi standard consentono confronti con cui il programma può essere giudicato (3, 7, 12).

BOX 9. Fonti di standard per giudicare la performance di un programma	
<ul style="list-style-type: none">❑ Esigenze dei partecipanti;❑ Valori, aspettative e norme della comunità;❑ Grado di partecipazione;❑ Obiettivi del Programma;❑ Protocolli e procedure del Programma;❑ Performance attesa, previsioni, stime;❑ Realizzabilità;❑ Assenza di rischi;❑ Target o criteri di performance prefissati;❑ Modifica della performance nel tempo;❑ Efficienza nell'uso delle risorse;❑ Standard Professionali;	<ul style="list-style-type: none">❑ Performance mediante il confronto con programmi precedenti o simili;❑ Performance mediante un gruppo di controllo o di confronto;❑ Mandati, politiche, statuti, regolamenti, leggi;❑ Giudizi da parte di gruppi di riferimento (es: partecipanti, staff, esperti, e ispettori delle agenzie che hanno finanziato il programma);❑ Finalità Istituzionali;❑ Ideologia Politica;❑ Equità Sociale;❑ Volontà Politica; eDiritti umani.

Adattato da a) Patton MQ. Utilization-focused evaluation: the new century text. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997; b) Scriven M. Minimalist theory of evaluation: the least theory that practice requires. American Journal of Evaluation 1998;19(1):57-70; c) McKenzie JF. Planning, implementing, and evaluating health promotion programs: a primer. New York, NY: Macmillan Publishing Company, 1993; d) Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and e) Gostin L, Mann JM. Towards the development of a human rights impact assessment for the formulation and evaluation of public health policies. Health and Human Rights 1994;1:59-80.

Analisi e Sintesi

Attraverso l'analisi e la sintesi dei risultati di una valutazione, è possibile individuare i pattern dell'evidenza, o enucleando alcuni specifici risultati importanti (analisi) o combinando le fonti di informazioni per raggiungere una comprensione più ampia (sintesi). Metodi di valutazione misti consistono nell'analisi separata di ogni elemento di prova e nella sintesi delle varie fonti, per esaminare pattern di accordo, convergenza o complessità. Decifrare i fatti da un insieme di prove implica la decisione del modo in cui organizzare, classificare, collegare, comparare, e mostrare le informazioni (7, 85-87). Queste decisioni vengono prese in base a diversi fattori: i quesiti posti, il tipo di dati disponibili, l'input dei committenti e dei fruitori primari.

Interpretazione

Con l'interpretazione si afferma che cosa significano i risultati. Essa è una parte del tentativo di comprensione dell'evidenza complessiva raccolta in uno studio di valutazione (88). Mettere in luce fatti riguardanti l'esecuzione di un programma, non basta per trarre le conclusioni della valutazione. I risultati devono essere interpretati per derivare le conseguenze pratiche di quanto che si è appreso. Questo è tanto più realizzabile, quanto più la valutazione ha tenuto conto delle esigenze dei committenti e quanto più questi partecipano ed interagiscono nella valutazione.

Giudizi

I giudizi sono affermazioni riguardo al merito, valore, o importanza del programma. Essi sono formulati confrontando risultati ed interpretazioni sul programma con uno o più standard selezionati. Dal momento che gli standard con cui confrontare la performance di un dato programma possono essere molteplici, i fruitori potrebbero arrivare a giudizi diversi o addirittura conflittuali. Per esempio, un programma che aumenta la produttività del 10% rispetto all'anno precedente dovrebbe essere considerata positivamente dai manager che usano come standard l'incremento nel tempo. D'altra parte, i membri della comunità potrebbero ritenere che, nonostante i miglioramenti, non è stata raggiunta la soglia minima di accessibilità ai servizi. Quindi, utilizzando lo standard dell'equità sociale il loro giudizio riguardo l'esecuzione del programma, sarebbe negativo. Richieste conflittuali riguardo la qualità del programma, valore, o importanza spesso indicano che i committenti stanno usando standard di giudizio diversi. Nel contesto di un processo di valutazione questi disaccordi possono funzionare come dei catalizzatori per chiarire quali sono i valori rilevanti e per discutere e negoziare quali siano le basi adeguate su cui il programma dovrebbe essere giudicato.

Raccomandazioni

Le raccomandazioni sono le azioni che bisognerebbe intraprendere in considerazione dei risultati della valutazione. Formulare raccomandazioni è una parte distinta del processo di valutazione, per cui sono necessarie informazioni che vanno al di là di quelle che erano state assunte per i giudizi sull'esecuzione del programma (3). Sapere che un intervento riduce il rischio di una malattia non si traduce necessariamente in una raccomandazione a continuare lo sforzo, specialmente quando esistono differenti priorità o interventi alternativi.

Così, le raccomandazioni a continuare, espandere, riprogettare, o mettere fine a un programma costituiscono un frutto del processo di valutazione, diverso dal giudizio sull'efficacia del programma. La formulazione di raccomandazioni richiede informazioni sul contesto, in particolare sul contesto organizzativo, in cui saranno prese le decisioni programmatiche (89). Raccomandazioni che non poggiano su una sufficiente mole di evidenze o quelle non coerenti con i valori dei committenti/fruitori, possono condurre a perdita di credibilità della valutazione. Invece, l'impatto di una valutazione può essere aumentato da raccomandazioni in linea con le sensibilità politiche dei fruitori e che li illuminano su punti fondamentali che essi hanno il potere di controllare o influenzare (7). Condividere definite raccomandazioni, sollecitare le reazioni da molteplici committenti, e presentare varie opzioni, anziché consigli diretti, aumenta la probabilità che le raccomandazioni saranno rilevanti e ben accolte. Varie attività soddisfano la necessità di giustificare le conclusioni in una valutazione (Box 10). Le conclusioni potrebbero essere rafforzate da:

- a) elencare e riassumere i meccanismi di cambiamento plausibili;
- b) delineare la sequenza temporale tra attività ed effetti;
- c) illustrare spiegazioni alternative e mostrare perché queste non sono sostenute da prove;
- d) mostrare che gli effetti possono essere ripetuti.

BOX 10. Giustificare le conclusioni

Definizione

Fare affermazioni riguardanti il programma giustificate sulla base di dati che sono stati valutati alla luce di idee di merito, valore o importanza (ovverossia alla luce di standard di valore) pertinenti e difendibili; le conclusioni sono giustificate quando sono legate all'evidenza raccolta e coerenti con valori condivisi o con gli standard dei destinatari.

Funzione

Rinforza le conclusioni e ciò è cruciale per rendere la valutazione utile ed accurata; implica la esplicitazione dei valori, l'analisi dei dati qualitativi e quantitativi e la sintesi, l'interpretazione sistematica, e il confronto appropriato con standard rilevanti ai fini del giudizio.

Esempi di Attività

- ❑ Usare metodi di analisi e sintesi appropriati per riassumere i risultati;
- ❑ Interpretare il significato dei risultati per decidere che cosa essi vogliono dire;
- ❑ Giudicare in accordo a valori chiaramente definiti che consentono di classificare un risultato, ad esempio come positivo o negativo, alto o basso;
- ❑ Considerare modi alternativi per confrontare i risultati (es: confrontati con gli obiettivi prefissati, con un gruppo di controllo, livelli stabiliti da norme statali, performance del passato, esigenze);
- ❑ Generare spiegazioni alternative per i risultati ed indicare perché tali spiegazioni dovrebbero o meno essere tenute in conto;
- ❑ Raccomandare azioni o decisioni coerenti con le conclusioni; e

- Limitare le conclusioni a situazioni, orizzonti temporali, persone, contesti, e finalità cui i risultati sono applicabili.

Adattato da: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

In presenza di diverse conclusioni, tutte ben sostenute, ognuna potrebbe essere presentata con un sommario dei punti di forza e di debolezza.

Ci si potrebbe servire di tecniche creative (per esempio, la tecnica Delphi) per stabilire consenso tra i committenti nell'assegnare giudizi di valore (90). Si dovrebbe ricercare l'accordo sulle tecniche per l'analisi, la sintesi, e l'interpretazione dei risultati, prima che la raccolta dati cominci, per assicurare che saranno disponibili tutti i dati necessari.

Fase 6: Assicurare l'Uso e la Condivisione delle Lezioni Apprese

Le lezioni apprese nel corso di una valutazione non si traducono automaticamente in decisioni informate e razionali, nè in azioni adeguate. Per garantire che i processi di valutazione e i risultati ottenuti siano utilizzati in modo appropriato e disseminati accuratamente c'è bisogno di un impegno deliberato. La valutazione conduce all'azione quando, già nella primissima fase di rapporto con i committenti/fruitori e poi in tutto il processo di valutazione, la preparazione è basata su un modo di pensare finalizzato alla pratica. Gli elementi critici per assicurare l'utilizzo pratico di una valutazione sono cinque: disegno dello studio, preparazione, feedback, follow-up e disseminazione.

Disegno

Il disegno si riferisce al modo in cui vengono sviluppati i metodi, i problemi e tutto il processo della valutazione. Come è stato discusso nella terza fase di questo framework, il disegno dovrebbe essere organizzato, fin dall'inizio, per ottenere risultati utilizzabili dagli utenti primari. Un disegno chiaro, finalizzato all'utilizzo pratico, aiuta chi eseguirà la valutazione a sapere precisamente "chi farà cosa" a seguito dei risultati e chi trarrà beneficio dal partecipare alla valutazione. Inoltre, se si richiede ed utilizza il contributo dei fruitori e committenti, durante la messa a punto del disegno dello studio, aumenta la rilevanza, la credibilità ed utilità complessiva della valutazione.

Preparazione

La preparazione si riferisce alle fasi in cui sono enumerati i possibili utilizzi pratici dei risultati della valutazione. L'abilità a tradurre la nuova conoscenza in un'azione appropriata è una capacità che può essere rafforzata attraverso la pratica. Costruire questa capacità può rappresentare di per se un beneficio per la valutazione (38, 39, 91). Elencare i modi in cui i potenziali risultati (particolarmente quelli negativi) potrebbero influire sulle decisioni, preparerà i committenti agli eventuali utilizzi delle evidenze raccolte (92). Si potrebbero fornire ai committenti e ad altri fruitori un insieme di risultati ipotetici, chiedendogli di descrivere quali azioni o decisioni, essi ritengono adeguate, sulla base di quanto scoperto. Se essi suggeriscono che le prove presentate sono incomplete e che non si dovrebbe intraprendere nessuna azione, è segno che la valutazione progettata deve essere modificata. Preparare all'utilizzo lascia ai committenti il tempo necessario per esplorare implicazioni positive e negative dei risultati potenziali e per individuare opzioni per il miglioramento del programma.

Feedback

Il Feedback è la comunicazione che avviene tra tutti i partecipanti alla valutazione. Dare e ricevere feedback crea un'atmosfera di fiducia tra i fruitori; mantenendo informati coloro che collaborano su come sta procedendo la valutazione, il feedback la tiene sui giusti binari. Gli utenti primari e altri committenti hanno il diritto di prendere parte a decisioni che potrebbero condizionare la probabilità di ottenere informazioni utili. Il feedback da parte dei committenti è una parte integrante della valutazione, che serve soprattutto ad assicurarne l'uso. Il conseguimento del feedback può essere raggiunto tenendo discussioni periodiche durante ogni fase della valutazione e attraverso la comunicazione di routine dei risultati ottenuti nel frattempo, delle interpretazioni provvisorie e dei primi abbozzi di rapporti sull'argomento.

Follow-Up

Il Follow-up si riferisce al supporto tecnico ed emozionale di cui i fruitori necessitano durante la valutazione e dopo che hanno ricevuto i risultati. A causa dell'impegno profuso, il raggiungimento di conclusioni valide può sembrare di per sé la fine di una valutazione; ma, un follow-up attivo potrebbe servire per ricordare agli utenti l'uso, che essi stessi avevano pianificato di fare, dei risultati della valutazione. Il follow-up potrebbe anche essere necessario per impedire che le lezioni apprese vadano perdute o ignorate, quando bisogna prendere decisioni complesse o che abbiano una rilevanza politica. Perché ciò non avvenga, chi è impegnato nella valutazione dovrebbe trasformarsi in una sorta di difensore delle scoperte della valutazione, durante la fase decisionale. Questo tipo di difesa ("advocacy") migliora l'apprezzamento dei risultati e di quali siano le azioni coerenti con i risultati della valutazione. Semplificare l'uso dei risultati della valutazione comporta anche la responsabilità di prevenire il maluso di essi (7, 12, 74, 93, 94). I risultati di una valutazione sono sempre legati al contesto in cui la valutazione è stata eseguita. Ciononostante, a volte i committenti potrebbero essere tentati di estrapolare i risultati dal loro contesto o utilizzarli per fini diversi da quelli stabiliti. Ad esempio, la generalizzazione inappropriata dei risultati ottenuti da uno studio locale per prendere decisioni che interessano molte altre aree della nazione, rappresenta un utilizzo errato della valutazione effettuata in ambito locale. Analogamente, committenti che cercassero di screditare un programma di interventi, potrebbero utilizzare in modo errato i risultati enfatizzando eccessivamente i risultati negativi, senza valorizzare gli attributi positivi del programma. Un follow-up attivo può prevenire queste ed altre forme di cattivo utilizzo, assicurando che le prove non siano interpretate in modo errato né siano applicate a problemi diversi da quelli messi a fuoco dalla valutazione.

Disseminazione

La disseminazione è il processo tramite il quale le procedure o le lezioni apprese da una valutazione sono comunicate a un pubblico interessato, in modo tempestivo, corretto e coerente. Sebbene sia necessario che esista la documentazione di una valutazione, non sempre un rapporto formale si rivela il miglior prodotto. Come altri elementi di una valutazione, la strategia per la documentazione dovrebbe essere discussa in anticipo con gli utenti e gli altri committenti. Questa consultazione preventiva garantisce che i bisogni di informazioni dei destinatari più importanti saranno soddisfatti.

Per programmare una comunicazione efficace c'è bisogno di considerare i tempi, lo stile, il tono, la fonte del messaggio, il mezzo, e il formato dei prodotti informativi. Una checklist dei punti da considerare allorché si prepara il rapporto della valutazione, include: l'adeguamento dei contenuti del rapporto ai destinatari, la spiegazione del

focus della valutazione e dei suoi limiti, e l'elencazione dei punti di forza così come dei punti di debolezza della valutazione (Box 11) (6).

BOX 11. Checklist per assicurare report efficaci

- ❑ Fornire agli utenti finali intermedi e finali in tempo per essere usati.
- ❑ Adattare contenuto, formato, e stile all'audience coinvolgendo membri dell'audience.
- ❑ Includere un sommario.
- ❑ Riassumere la descrizione dei destinatari e come questi erano stati impegnati.
- ❑ Descrivere i caratteri essenziali del programma (incluso il modello logico).
- ❑ Evidenziare il focus della valutazione e le sue limitazioni.
- ❑ Includere un sommario adeguato del piano e delle procedure della valutazione.
- ❑ Fornire tutta l'informazione tecnica necessaria (casomai in appendice).
- ❑ Specificare gli standard e i criteri in base ai quali sono stati emessi i giudizi valutativi.
- ❑ Spiegare i giudizi valutativi e quanto essi sono sostenuti dall'evidenza.
- ❑ Elencare sia i punti di forza che i punti di debolezza della valutazione.
- ❑ Discutere le raccomandazioni per l'azione con i loro vantaggi, svantaggi, e implicazioni in termini di risorse.
- ❑ Assicurare protezione per i clienti ed altri destinatari del programma.
- ❑ Prevedere come la gente o le organizzazioni potrebbero essere influenzate dai risultati.
- ❑ Presentare le opinioni delle minoranze o le controrepliche laddove necessario.
- ❑ Verificare che il report sia accurato e senza distorsioni.
- ❑ Organizzare il report logicamente, includendo i dettagli appropriati.
- ❑ Eliminare il gergo tecnico
- ❑ Usare esempi, figure, grafici, e storie.

Adattato da: Worthen BR, Sanders JR, Fitzpatrick JL. Program evaluation:

alternative approaches and practical guidelines. 2nd ed. New York, NY: Logman, Inc. 1996.

Usi Ulteriori

Usi ulteriori della valutazione possono emergere nel corso dello studio; questi usi hanno un loro valore e dovrebbero essere incoraggiati in quanto complementari agli usi predefiniti (Box 12) (7,38,93,94).

Le persone che prendono parte ad una valutazione possono subire profondi cambiamenti nel modo di pensare e nel comportamento.

In particolare, quando persone alle prime esperienze nella valutazione iniziano a pensare come valutatori, si possono verificare importanti modifiche dei loro punti di vista. La valutazione spinge i membri dello staff a chiarificare quanto essi hanno capito degli scopi del programma. Questa maggiore chiarezza permette allo staff di funzionare in modo coeso, come una squadra concentrata sul raggiungimento di un fine comune (95). Immergersi nella logica, nel tipo di ragionamento, e nei valori della valutazione può avere un impatto duraturo (come l'imparare a basare le decisioni sull'analisi sistematica dell'evidenza invece che su assunti che possono essere infondati) (7).

Tra gli usi ulteriori, derivati dal solo fatto di partecipare alla valutazione, vi è la definizione di indicatori per comprendere cosa è davvero importante per coloro che prendono decisioni e definire gli outcomes importanti, modificando l'incentivazione connessa con il conseguimento dell'outcome (ad esempio, pagando dividendi basati sull'outcome a quei programmi che, attraverso la prevenzione, generano risparmi) (96). I benefici che derivano da questi ed altri usi forniscono ulteriori motivazioni per cominciare le attività di valutazione fin dall'inizio di un programma.

BOX 12. Assicurare l'uso e condividere le lezioni apprese

Definizione

Assicurare che a) i destinatari siano consapevoli delle procedure e dei risultati della valutazione; b) i risultati vengano presi in considerazione quando bisogna prendere decisioni o intraprendere azioni che influenzano il programma; e c) coloro che hanno partecipato alla valutazione abbiano fatto un'esperienza fruttuosa (utilità di processo).

Funzione

Assicura che la valutazione raggiunge il suo principale scopo – essere utile; tuttavia, diversi fattori potrebbero influire sul grado di utilizzo, incluso la credibilità del valutatore, la chiarezza del report, la tempestività del report e la diffusione e divulgazione dei risultati, l'imparzialità del reporting, e i cambiamenti nel programma o nel contesto organizzativo.

Esempi di Attività

- Disegnare la valutazione per raggiungere l'uso voluto dagli utenti prescelti;
- Preparare i destinatari per l'eventuale utilizzo dei possibili risultati costruendo ipotesi su come varie conclusioni porterebbero a differenti decisioni influenzando sull'operatività del programma;
- Fornire continuamente feedback ai destinatari su risultati intermedi, interpretazioni provvisorie, e decisioni che potrebbero influire sulla probabilità che lo studio produca risultati utilizzabili;
- Mettere in agenda incontri di follow-up con i destinatari per facilitare il trasferimento delle conclusioni della valutazione in azioni e decisioni appropriate; e
- Comunicare le procedure usate e le lezioni apprese dalla valutazione ai destinatari, usando strategie di comunicazione adeguate capaci di soddisfare le loro specifiche esigenze.

Adattato da: (a) Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and (b) Patton MQ. Utilization-focused evaluation: the new century text. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997.

Standard per la Valutazione Efficace

I professionisti della sanità pubblica riconosceranno che le fasi basilari del framework per la valutazione dei programmi sono parte del loro lavoro abituale. Nella pratica corrente della sanità pubblica, si consultano i committenti; si definiscono gli scopi del programma; si avanzano quesiti guida; sono raccolti, analizzati ed interpretati i dati; si formulano dei giudizi; e si condivide ciò che si è appreso. Sebbene nella pratica routinaria la valutazione viene fatta in modo informale, esistono degli standard per giudicare se un insieme di attività valutative è ben progettato e potenzialmente funzionale. A questo scopo, il *Joint Committee on Standards for Educational Evaluation* ha definito degli standard di valutazione di programma (12). Questi standard, progettati per valutare programmi educativi, sono importanti anche per i programmi di sanità pubblica.

Gli standard per la valutazione di programma fanno in modo che la pratica valutativa sia chiara e corretta. Gli standard forniscono linee-guida da seguire allorché si deve decidere tra diverse opzioni della valutazione.

Gli standard contribuiscono ad evitare di condurre valutazioni non equilibrate (ad esempio, studi accurati e realizzabili ma inutili oppure studi potenzialmente utili e accurati ma irrealizzabili). Inoltre, gli standard possono essere applicati quando bisogna pianificare una valutazione e durante l'esecuzione. Il Joint Committee è ben chiaro su questo punto: "gli standard sono principi guida, non regole meccaniche. . . . Infine, è sempre possibile giudicare se un dato standard è stato messo in pratica adeguatamente in una situazione particolare" (12). Nel rapporto del Joint Committee gli standard sono raggruppati nelle seguenti 4 categorie ed includono un totale di 30 standard specifici (Box 13-16). Così

come è descritto nel resoconto, ogni categoria ha un relativo elenco di linee-guida ed errori più comuni, illustrati con esempi di casi applicati: utilità, attuabilità, appropriatezza, e accuratezza.

Standard 1: Utilità

Gli standard di utilità garantiscono che siano soddisfatti i bisogni informativi dei destinatari della valutazione. Sette standard di utilità (Box 13) definiscono altrettante tematiche come, ad esempio, l'identificazione di coloro per cui la valutazione avrà una ricaduta pratica, l'ammontare ed il tipo di informazioni raccolte, i valori utilizzati nell'interpretazione dei risultati della valutazione, e la chiarezza e la tempestività dei rapporti della valutazione.

BOX 13. Standard di utilità

I seguenti standard di utilità assicurano che una valutazione andrà incontro alle esigenze di coloro che devono utilizzare i risultati, cioè gli utenti o destinatari:

- Identificazione dei destinatari. Le persone coinvolte in- o influenzate dalla valutazione dovrebbero essere identificate, le loro esigenze definite e la valutazione rivolta ad esse.
- Credibilità del valutatore. Le persone che conducono la valutazione dovrebbero essere affidabili e competenti nel campo per raggiungere il massimo di credibilità ed accettazione.
- Scelta dell'informazione. L'informazione da raccogliere dovrebbe essere in grado di dare risposta a questioni pertinenti riguardo il programma ed costituire una risposta alle esigenze e interessi degli utenti o di altri destinatari specificati.
- Identificazione dei Valori. La prospettiva, le procedure, ed il rationale impiegato per interpretare i risultati dovrebbero essere descritti con cura, sicché la base per i giudizi di valori appaia chiaramente.
- Chiarezza del report. I reports dovrebbero descrivere chiaramente il programma da valutare, incluso il contesto e le finalità, le procedure, ed i risultati della valutazione, in modo che l'informazione essenziale possa essere facilmente compresa.
- Tempestività e distribuzione del report. Risultati intermedi importanti e report della valutazione dovrebbero essere distribuiti ai destinatari in in tempo utile da consentire a costoro di farne uso.
- Valutazione d'impatto. La pianificazione e la conduzione delle valutazioni e la produzione dei report dovrebbero avvenire in modo che tutto il processo possa essere seguito dai destinatari per aumentare la probabilità che i risultati potranno essere utilizzati.

Fonte: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994

Standard 2: Attuabilità

Gli standard di attuabilità garantiscono che la valutazione sia possibile e concreta. I tre standard di attuabilità (Box 14) mettono in evidenza che: la valutazione dovrebbe impiegare procedure pratiche e non dispersive; che gli interessi politici divergenti di coloro che sono coinvolti dovrebbero essere analizzati in anticipo e resi noti; e che l'uso delle risorse nella esecuzione della valutazione dovrebbe essere prudente e tendere a produrre risultati rilevanti.

BOX 14. Standard di Realizzabilità

I seguenti standard di fattibilità assicurano che la valutazione sarà realistica, prudente, diplomatica, ed economica:

- Procedure pratiche. Le procedure valutative dovrebbero essere pratiche e l'informazione necessaria dovrebbe essere ottenuta col minimo possibile di complicazioni organizzative.
- Fattibilità politica. Durante la pianificazione e la conduzione della valutazione, è necessario considerare attentamente le diverse posizioni dei vari gruppi di interesse in modo da ottenere la loro cooperazione e da evitare o attenuare possibili tentativi da parte di qualche gruppo di screditare la valutazione o distorcere o applicare male i risultati.
- Costo-efficacia. La valutazione dovrebbe essere efficiente, produrre informazioni importanti per giustificare le risorse impegnate in essa.

Fonte: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

Standard 3: Correttezza

Gli standard di correttezza garantiscono che la valutazione sia etica (cioè, condotta tenendo conto dei diritti e degli interessi di coloro che sono coinvolti e in qualche modo implicati). Gli otto standard di correttezza (Box 15) riguardano temi come la necessità di approntare protocolli concordati per la valutazione; la protezione del benessere dei soggetti umani; ponderare e rivelare i risultati in modo completo ed equilibrato; e rivolgersi a qualsiasi conflitto d'interesse in modo aperto e conveniente.

BOX 15. Standard di Correttezza

I seguenti standard di correttezza assicurano che la conduzione della valutazione sia legalmente ed eticamente corretta riguardo al benessere di coloro che sono coinvolti nella valutazione e di coloro che saranno coinvolti a seguito dei risultati della valutazione stessa:

- Orientamento al servizio. La valutazione dovrebbe essere progettata in modo da assistere le organizzazioni a rispondere più efficacemente ai bisogni degli utenti.
- Accordi Formali. Tutti gli attori principali coinvolti in una valutazione dovrebbero trovarsi d'accordo nel sottoscrivere i loro obblighi (che cosa bisogna fare, come, da chi, e quando), così che ciascuno possa rispettare le condizioni dell'accordo o rinegoziare l'accordo.
- Diritti dei soggetti umani. La valutazione dovrebbe essere progettata e condotta in modo da rispettare e proteggere i diritti ed il benessere dei soggetti umani.
- Rapporti umani. I valutatori dovrebbero interagire rispettando le altre persone che partecipano alla valutazione, in modo che i partecipanti non siano danneggiati.
- Valutazione completa ed imparziale. La valutazione dovrebbe essere completa ed imparziale nell'esaminare e registrare i punti di forza ed i punti di debolezza del programma così che le parti ben funzionanti possano essere ulteriormente rafforzate e i problemi individuati e capiti.
- Divulgazione dei risultati. I principali partecipanti ad una valutazione dovrebbero garantire che tutte le persone su cui, in qualche modo, la valutazione influirà (e ogni altro che abbia un diritto legalmente riconosciuto) possano effettuare un completo esame dei risultati, compresi i loro limiti.
- Conflitto di interessi. Il conflitto di interessi dovrebbe essere trattato in modo aperto e leale sì da non compromettere il processo ed i risultati della valutazione.

Fonte: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

Standard 4: Accuratezza

Gli standard di accuratezza garantiscono che la valutazione produca risultati che siano considerati corretti. Dodici standard di accuratezza (Box 16) includono temi come la descrizione del programma e del suo contesto; l'articolazione nei dettagli dello scopo e

dei metodi della valutazione; l'impiego di procedure sistematiche per raccogliere informazioni valide e credibili; l'applicazione di metodi qualitativi o quantitativi durante l'analisi e la sintesi, e la produzione di risultati imparziali che contengano conclusioni che siano giustificate. Le fasi e gli standard sono utilizzati insieme nel corso del processo di valutazione. Per ogni fase, si dovrebbe considerare un sottoinsieme di standard rilevanti (Box 17).

BOX 16. Standard di accuratezza

I seguenti standard di accuratezza assicurano che la valutazione fornirà informazione tecnicamente adeguata riguardo le caratteristiche del programma che determinano il suo valore.

- Documentazione. Il programma da valutare dovrebbe essere chiaramente ed accuratamente documentato.
- Analisi del contesto. Il contesto in cui è attivo il programma dovrebbe essere analizzato sufficientemente a fondo, sì da identificare le probabili influenze sul programma.
- Descrizione di scopi e procedure. Scopi e procedure dovrebbero essere descritte dettagliatamente in modo che il lettore possa farsi un proprio giudizio.
- Fonti di informazioni difendibili. Le fonti informative dovrebbero essere riportate dettagliatamente, in modo da valutarne l'attendibilità.
- Informazione valida. Le procedure di rilevazione dell'informazione dovrebbero essere messe a punto in modo da consentire una valida interpretazione per gli usi previsti.
- Informazione riproducibile. Le procedure di rilevazione dovrebbero essere messe a punto in modo da assicurare che l'informazione raccolta sarà sufficientemente riproducibile per gli usi previsti.
- Informazione sistematica. L'informazione raccolta, elaborata e pubblicata dovrebbe essere verificata in modo sistematico e ogni errore dovrebbe essere corretto.
- Analisi di dati quantitativi. I dati quantitativi dovrebbero essere analizzati in modo appropriato e sistematico in modo da avere risposte efficaci ai quesiti della valutazione.
- Analisi di dati qualitativi. I dati qualitativi dovrebbero essere analizzati in modo appropriato e sistematico in modo da avere risposte efficaci ai quesiti della valutazione.
- Giustificazione delle conclusioni. Le conclusioni dovrebbero essere giustificate in modo esplicito per consentire ai destinatari di farsi un proprio giudizio.
- Reporting imparziale. Bisognerebbe stare in guardia, nel reporting, da distorsioni causate da sentimenti personali o punti di vista pregiudiziali di persone coinvolte nella valutazione, in modo da riflettere fedelmente i risultati.
- Metavalutazione. Bisognerebbe che la valutazione a sua volta giudicata, in base a questi o altri standard pertinenti in modo da portarla avanti in modo appropriato e, alla fine, consentire ai destinatari di rendersi conto dei punti di forza e dei punti di debolezza della valutazione stessa.

Fonte: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

BOX 17. Standard rilevanti fase per fase

Fase della valutazione	Standard Rilevante	Gruppo/Box
Rapporto con i Destinatari	Identificazione dei Destinatari	Utilità/13A
	Credibilità del valutatore	Utilità/13B
	Accordi formali	Correttezza/15B
	Diritti dei soggetti umani	Correttezza/15C
	Interazioni personali	Correttezza/15D
	Conflitti di interesse	Correttezza/15G
	Metavalutazione	Correttezza/15L
Descrizione del Programma	Valutazione completa e imparziale	Correttezza/15C
	Documentazione sul programma	Accuratezza/16A
	Analisi del contesto	Accuratezza/16B
	Metavalutazione	Accuratezza/16L
Mettere a fuoco il disegno della valutazione	Valutazione dell'impatto	Utilità/13G
	Procedure pratiche	Realizzabilità/14A
	Fattibilità politica	Realizzabilità/14B
	Costo efficacia	Realizzabilità/14C
	Orientamento al servizio	Correttezza/15A
	Valutazione completa e imparziale	Correttezza/15E
	Responsabilità fiscale	Correttezza/15H
Rilevare prove credibili	Descrizione degli scopi e delle procedure	Accuratezza/16C
	Metavalutazione	Accuratezza/16L
	Scelta dell'informazione	Utilità/13C
	Fonti di informazione affidabili e difendibili	Accuratezza/16D
	Informazioni valide	Accuratezza/16E
	Informazioni riproducibili	Accuratezza/16F
Giustificazione delle conclusioni	Informazioni sistematiche	Accuratezza/16G
	Metavalutazione	Accuratezza/16L
	Identificazione dei valori	Utilità/13D
	Analisi dell'informazione quantitativa	Accuratezza/16H
	Analisi dell'informazione qualitativa	Accuratezza/16G
Garantire la messa in pratica e la condivisione delle lezioni apprese	Conclusioni giustificate	Accuratezza/16J
	Metavalutazione	Accuratezza/16L
	Credibilità del valutatore	Utilità/13B
	Chiarezza del report	Utilità/13E
	Tempestività e buona distribuzione del report	Utilità/13F
	Impatto della valutazione	Utilità/13G
	Divulgazione dei risultati	Correttezza/15F
Reporting imparziale	Accuratezza/16K	
Metavalutazione	Accuratezza/16L	

APPLICAZIONE DEL FRAMEWORK

Condurre Valutazioni Ottimali

I professionisti della sanità pubblica non possono più chiedersi se valutare i loro programmi; le domande appropriate sono, invece,

- Qual è il modo migliore di valutare?
- Cosa si sta per apprendere dalla valutazione? E,
- in che modo le informazioni apprese dalle valutazioni saranno usate per rendere gli sforzi nella sanità pubblica più efficaci e influenti?

Il framework per la valutazione dei programmi aiuta a rispondere a queste domande guidando i propri utenti nella selezione di strategie di valutazione che possano rivelarsi utili, attuabili, etiche, e accurate.

Per usare il framework consigliato in uno specifico contesto di programma c'è bisogno della pratica, la quale sviluppa le abilità sia nella scienza sia nell'arte della valutazione dei programmi. Quando si applica il framework, la cosa difficile è escogitare un'ottimale - in quanto opposta ad ideale - strategia.

La strategia ottimale esegue ogni fase descritta nel framework adattandola al contesto del programma, soddisfacendo tutti gli standard rilevanti o facendo addirittura meglio. Le valutazioni dei CDC dei programmi di prevenzione del virus dell'immunodeficienza, inclusi i programmi nelle scuole, forniscono esempi di strategie ottimali per la valutazione a livello nazionale, statale, e locale (97,98).

Formare il team della Valutazione

Mettere insieme e concentrare gli sforzi di un gruppo cooperativo è un approccio per condurre una valutazione ottimale (24, 25). Un approccio di squadra si realizza quando un piccolo gruppo di persone, accuratamente selezionato, decide cosa si deve ottenere dalla valutazione e mette in comune le risorse per eseguire il piano. I destinatari potrebbero avere variabili livelli di coinvolgimento nella squadra a secondo dei loro particolari punti di vista, abilità operative, e interessi. La designazione di un leader è necessaria per il coordinamento della squadra ed il mantenimento della continuità per tutto il processo; dopodiché, le fasi pratiche della valutazione guidano la selezione dei membri della squadra. Per esempio:

- Coloro che sono diplomatici e sono in contatto con altre reti di relazioni professionali possono "agganciare" altri destinatari e mantenere la continuità.
- Quando si descrive il programma, è necessario che lo facciano persone che hanno compreso la storia del programma, lo scopo, e l'operatività.
- Altri ancora possono essere dotati di abilità come facilitatori del lavoro di gruppo; a costoro si potrebbe chiedere di aiutare a rendere esplicite aspettative riguardanti il programma passate sotto silenzio. Tali facilitatori possono inoltre aiutare i destinatari a creare modelli logici che descrivano il programma e chiariscano il tipo di relazioni tra i mezzi ed i fini, presente nel programma.
- Coloro che prendono decisioni e gli altri che dirigono il programma possono contribuire a incentrare il progetto di valutazione sui problemi che interessano specifici utenti e portino a specifici usi. Essi possono inoltre stabilire la logistica: il campo d'azione, la tempistica, e la data di consegna dei risultati della valutazione.
- Gli scienziati, in particolar modo gli scienziati sociali, possono apportare expertise allo sviluppo dei problemi, metodi, e strategie per la rilevazione delle evidenze. Essi possono inoltre aiutare ad analizzare il modo in cui un programma opera nel suo contesto di organizzazione o di comunità.

- Persone fidate ("super partes") che non hanno un diretto interesse nella valutazione, possono fungere da garanti che i valori dei partecipanti siano trattati adeguatamente allorquando si applicano gli standard, si interpretano i fatti, e si raggiungono conclusioni giustificate.

- Gli avvocati, comunicatori, creativi, e membri della struttura di potere possono aiutare a garantire che gli insegnamenti appresi dalla valutazione influenzeranno in futuro il processo decisionale circa la strategia del programma.

Tutte le organizzazioni, anche quelle che sarebbero in grado di formare la squadra di valutazione al proprio interno, dovrebbero ricercare nella comunità i partner con cui collaborare, avvantaggiandosi delle risorse della comunità, quando assemblano il team per la valutazione.

Questa strategia aumenta le risorse disponibili ed eleva la credibilità della valutazione. Per di più, una squadra formata da differenti destinatari ha una più alta probabilità di condurre una valutazione attenta alle differenze culturali (vale a dire, comprensiva e sensibile verso le differenti persone, condizioni, e contesti associati al programma) (99,100).

Sebbene costituisca una vera sfida per i coordinatori ed i partecipanti, l'approccio collaborativo è pratico per i benefici che apporta (ad esempio, riduce sospetti e timori, aumenta la consapevolezza e l'impegno, aumenta le possibilità che i risultati siano utilizzati, e permette di considerare diverse prospettive) (8,24).

Rivolgersi ad Interessi Comuni

Utilizzando il framework è possibile chiarire gli interessi comuni riguardo la valutazione di un programma. Potrebbe succedere che le valutazioni non vengano intraprese perché ritenute, a torto, costose. In ogni caso, la spesa di una valutazione è relativa; il costo dipende dai quesiti posti e dal livello di precisione che si desidera dalle risposte. Aniché scoraggiare le valutazioni perché tolgono tempo e perché sono tangenziali rispetto alle operazioni di programma (ad esempio, lasciate nell'ultimo periodo nel progetto del programma), il framework suggerisce di condurre la valutazione, fin dall'inizio, scadenzata strategicamente, per fornire il feedback necessario per guidare l'azione. Ciò fa in modo che la valutazione sia integrata con l'attività pratica del programma. Un'altra preoccupazione è incentrata sulle presunte difficoltà tecniche per progettare e condurre una valutazione. Sebbene vi siano circostanze in cui sono necessari ambienti controllati e tecniche analitiche elaborate, la maggior parte dei programmi di valutazione nella sanità pubblica non richiede questi metodi. Invece, l'approccio pratico adottato da questo framework, focalizza l'attenzione su quesiti che miglioreranno il programma, tramite l'uso di metodi sensibili al contesto e tecniche analitiche che riassumano accuratamente il significato delle informazioni qualitative e quantitative. Infine, la prospettiva della valutazione può turbare i membri dello staff del programma perché essi percepiscono i metodi di valutazione come punitivi, escludenti, o contrari. Il framework incoraggia un approccio di valutazione che sia progettato per essere utile, ed impegna tutti i destinatari interessati in un processo che ben accoglie la loro partecipazione. Le sanzioni che si applicano, qualora ve ne siano, non dovrebbero derivare dal fatto che ci si accorga di risultati negativi, ma dal non riuscire a sfruttare l'insegnamento ricavato per modifiche volte ad una maggiore efficacia (10).

TENDENZE DELLA VALUTAZIONE

L'interesse nel miglioramento e nella responsabilità di un programma continua a crescere nel governo, e nei settori del privato e del non-profit. *The Government Performance and Results Act* richiede alle agenzie federali di stabilire gli scopi dell'esecuzione e di misurare i risultati annualmente. Le organizzazioni donatrici non-profit (ad esempio, United Way) hanno integrato la valutazione nei loro programmi di attività ed ora esigono che chi riceve finanziamenti misuri gli outcomes del programma (30).

Anche le fondazioni orientate sulla sanità pubblica (ad esempio, W.K. Kellogg Foundation) hanno iniziato a dare importanza al ruolo della valutazione nella loro programmazione (24). Inoltre, sono emersi approcci innovativi alla formazione dello staff per la valutazione dei programmi. Per esempio, *American Cancer Society (ACS) Collaborative Evaluation Fellows Project*, per valutare i programmi di prevenzione locali sul cancro, collega studenti e facoltà in 17 scuole di sanità pubblica con gli uffici nazionali e regionali dell'ACS (101). Questo tipo di attività nei settori pubblici e privati è il riflesso di un investimento collettivo per creare una capacità di valutazione che migliori l'esecuzione e che sia responsabile dei risultati ottenuti nella sanità pubblica.

Vengono effettuati investimenti nella capacità di valutazione al fine di migliorare la qualità e l'efficacia del programma. Uno dei migliori esempi di effetti benefici di condurre valutazioni è il *the Malcolm Baldrige National Quality Award Program* (102).***

L'evidenza dimostra che i processi valutativi richiesti per vincere il premio Balbridge, hanno aiutato le aziende americane a sopravanzare i loro competitori (103). Adesso questi stessi effetti sulla qualità e sull'esecuzione vengono tradotti nel settore della sanità e dei servizi sociali. Di recente, per giudicare il grado di perfezione delle organizzazioni sanitarie sono stati adattati i criteri del premio Balbridge (104). Questa estensione all'industria sanitaria illustra il ruolo critico della valutazione nel raggiungimento di obiettivi nella sanità e nei servizi sociali. Allo stesso modo, il framework per la valutazione di programma è stato sviluppato per contribuire all'integrazione della valutazione nella cultura organizzativa della sanità pubblica e per aderire ai principi operativi dei CDC nella pratica di sanità pubblica.

Costruire la capacità di valutazione per tutti gli operatori della sanità pubblica è un obiettivo condiviso anche dal *Public Health Functions Steering Committee*. Insieme a *U.S. Surgeon General*, questo comitato ha identificato le competenze nella valutazione come essenziali per gli operatori di sanità pubblica nel ventunesimo secolo (105).

Essendo la valutazione focalizzata sul rendere accessibile la valutazione a tutto il personale ed ai committenti, il framework aiuta a promuovere il grado di conoscenza della valutazione e la competenza tra tutti i professionisti di sanità pubblica.

SOMMARIO

La valutazione è l'unico modo per separare programmi che promuovono la salute e prevengono gli incidenti, malattie, e disabilità da quelli che non funzionano; essa costituisce un punto di forza ed una guida per la pianificazione di strategie di sanità pubblica efficaci, il miglioramento di programmi già esistenti, e la dimostrazione dei risultati derivanti dagli investimenti delle risorse. La valutazione focalizza anche l'attenzione sulle finalità condivise dei programmi di sanità pubblica e pone la domanda se la dimensione dell'investimento è adeguata ad i compiti che devono essere svolti (95). Il framework raccomandato rappresenta una sintesi di pratiche di valutazione già esistenti e

uno standard per miglioramenti ulteriori. Esso sostiene un approccio pratico alla valutazione che è basato su fasi e standard applicabili nel contesto della sanità pubblica. Poiché il framework, di proposito, ha un alto grado di generalizzazione, esso fornisce una guida per la progettazione e la condotta di progetti specifici di valutazione in numerose e diverse aree di programma. Inoltre, il framework può essere utilizzato come modello per creare altre nuove linee-guida di valutazione, adattate a specifici programmi, che rendano le fasi e gli standard meglio applicabili, in pratica, per ciascun programma (20,96,106-112).

In definitiva, il framework raccomandato è uno dei diversi strumenti che i CDC , assieme ai loro partner, possono utilizzare per migliorare e assumere maggiori responsabilità nel lavoro di promozione della salute e di prevenzione delle malattie e degli incidenti.

ULTERIORI INFORMAZIONI

Fonti di ulteriori informazioni sono disponibili per coloro che desiderano iniziare ad applicare il framework presentato in questa relazione o che desiderino aumentare la loro comprensione della valutazione. In particolare, si raccomandano le seguenti risorse:

is an Internet resource for health promotion and community development that contains information regarding how to conduct public health work and social change on a community level. Because they consider program evaluation to be a critical part of successful community-based health promotion, the CTB team used the framework for program evaluation to create a unique gateway to evaluation-related ideas and tools. This gateway can be accessed at <<http://ctb.lsi.ukans.edu/ctb/c30/ProgEval.html>>. The CDC Evaluation Working Group has compiled a list of additional resources for program evaluation. These resources address such topics as a) ethics, principles, and standards for program evaluation; b) evaluation-related organizations, societies, foundations, and associations; c) journals and on-line publications; d) step-by-step evaluation manuals; e) resources for developing logic models; f) planning- and performance-improvement tools; and g) evaluation-related publications. This list of resources can be obtained through the Working Group's website at <<http://www.cdc.gov/eval/index.htm>> or by sending an electronic message to <eval@cdc.gov>.

"Practical Evaluation of Public Health Programs" (course no. VC0017) è un corso d'insegnamento a distanza della durata di 5 ore che usa anche il framework presentato in questa relazione. Sviluppato attraverso CDC's Public Health Training Network (PHTN) (8), il corso consiste di due videocassette e un quaderno, che possono essere utilizzati da singoli per studio autodidatta o da piccoli gruppi con l'aggiunta di altre attività facoltative. Per questo corso è disponibile il credito per la continuazione dell'istruzione. Ulteriori informazioni sono disponibili al sito Web PHTN:

<http://www.cdc.gov/PHTN> o chiamando, gratuitamente, 800-41-Train (800-418-7246). Inoltre, materiale per il corso può essere acquistato chiamando gratuitamente la Public Health Foundation 877-252-1200 oppure usando la forma di ordinazione on-line a:

<http://bookstore.phf.org./prod41.htm>.

Per scopi informativi, il quaderno può essere visionato su Internet al sito:

<http://www.cdc.gov/eval/workbook.pdf>.

The Community Toolbox (CTB) è una risorsa Internet per la promozione della salute che contiene informazioni sul modo di condurre il lavoro di sanità pubblica e il cambiamento sociale al livello della comunità locale. Si può accedere a tale strumento al sito:

<<http://ctb.lsi.ukans.edu/ctb/c30/ProgEval.html>>.

Il Gruppo di Lavoro per la Valutazione dei CDC ha compilato una lista di risorse ulteriori per la valutazione di programma. Queste risorse si rivolgono ad argomenti quali a) etica, principi, e standard per la valutazione di programma; b) organizzazioni correlate alla valutazione, società, fondazioni, ed associazioni; c) riviste e pubblicazioni on-line; d) manuali di valutazione fase per fase; e) risorse per lo sviluppo di modelli logici ; f) miglioramento della pianificazione e dell'esecuzione; e g) pubblicazioni correlate alla valutazione. Questa lista di risorse può essere ottenuta attraverso il sito Web

<<http://www.cdc.gov/eval/index.htm>>

o inviando un messaggio elettronico a

<eval@cdc.gov>

Riferimenti bibliografici

1. Dyal WW. Ten organizational practices of public health: a historical perspective. *Am J Prev Med* 1995;11(6)Suppl 2:6-8.
2. Koplan JP. CDC sets millennium priorities. *US Medicine* 1999;4-7.
3. Scriven M. Minimalist theory of evaluation: the least theory that practice requires. *American Journal of Evaluation* 1998;19:57-70.
4. Shadish WR, Cook TD, Leviton LC. *Foundations of program evaluation: theories of practice*. Newbury Park, CA: Sage Publications, 1991.
5. Weiss CH. *Evaluation: methods for studying programs and policies*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1998.
6. Worthen BR, Sanders JR, Fitzpatrick, JL. *Program evaluation: alternative approaches and practical guidelines*. 2nd ed. New York, NY: Longman, 1996.
7. Patton MQ. *Utilization-focused evaluation: the new century text*. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997.
8. CDC. *Practical evaluation of public health programs*. Atlanta, GA: US Department of Health and Human Services, CDC, Public Health Training Network, 1998; PHTN course no. VC-0017.
9. Love A. *Internal evaluation: building organizations from within*. *Applied social research*, vol 24. Newbury Park, CA: Sage Publications, 1991.
10. Sanders JR. Uses of evaluation as a means toward organizational effectiveness. In: Gray ST, ed. *Leadership IS: a vision of evaluation; a report of learnings from Independent Sector's work on evaluation*. Washington, DC: Independent Sector, 1993.
11. Shadish WR. Evaluation theory is who we are. *American Journal of Evaluation* 1998;19(1):1-19.
12. Joint Committee on Standards for Educational Evaluation. *Program evaluation standards: how to assess evaluations of educational programs*. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
13. Joint Committee on Standards for Educational Evaluation. *Personnel evaluation standards: how to assess systems for evaluating educators*. Newbury Park, CA: Sage Publications, 1988.
14. CDC. *Principles of community engagement*. Atlanta, GA: CDC, Public Health Practice Program Office, 1997.
15. Mertens DM. Inclusive evaluation: implications of transformative theory for evaluation. *American Journal of Evaluation* 1999;20(1):1-14.
16. Connell JP, Kubisch AC. Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. In: Fulbright-Anderson K, Kubisch AC, Connell JP, eds. *New approaches to evaluating community initiatives: theory, measurement, and analysis*. Washington, DC: Aspen Institute, 1998.
17. Chen HT. *Theory driven evaluations*. Newbury Park, CA: Sage Publications, 1990.
18. U.S. General Accounting Office. *Managing for results: measuring program results that are under limited federal control*. Washington, DC: U.S. General Accounting Office 1998; publication no. GAO/GGD-99-16.
19. Haddix AC, Teutsch SM, Shaffer PA, Duñet DO, eds. *Prevention effectiveness: a guide to decision analysis and economic evaluation*. New York, NY: Oxford University Press, 1996.
20. CDC. Framework for assessing the effectiveness of disease and injury prevention. *MMWR* 1992;41(No. RR-3):1-13.
21. Gold ME, Russell LB, Siegel JE, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
22. Eoyang GH, Berkas T. Evaluation in a complex adaptive system. In: Lissack M, Gunz H, eds. *Managing complexity in organizations*. Westport, CT: Quorum Books, 1999.
23. Rush B, Ogbourne A. Program logic models: expanding their role and structure for program planning and evaluation. *Canadian Journal of Program Evaluation* 1991;6(2):95-106.
24. W.K. Kellogg Foundation. *W.K. Foundation Evaluation Handbook*. Battle Creek, MI: W.K. Kellogg Foundation, 1998.
25. Taylor-Powell E, Rossing B, Geran J. *Evaluating collaboratives: reaching the potential*. Madison, WI: University of Wisconsin--Cooperative Extension, 1998.
26. Poreteous NL, Sheldrick BJ, Stewart PJ. *Program evaluation tool kit: a blueprint for public health management*. Ottawa, Ontario, Canada: Ottawa-Carleton Health Department, Public Health Research, Education and Development Program, 1997.
27. Weiss CH. Nothing as practical as a good theory: exploring theory-based evaluation for comprehensive community initiatives for families and children. In: Connell JP, Kubisch AC, Schorr LB, Weiss CH, eds. *New approaches to evaluating community initiatives: concepts, methods, and contexts*. Washington, DC: Aspen Institute, 1995.
28. Weiss CH. How can theory-based evaluation make greater headway? *Evaluation Review* 1997;21:501-24.

29. Wong-Reiger D, David L. Using program logic models to plan and evaluate education and prevention programs. In: Love AJ, ed. *Evaluation methods sourcebook II*. Ottawa, Ontario, Canada: Canadian Evaluation Society, 1995.
30. United Way of America. *Measuring program outcomes: a practical approach*. Alexandria, VA: United Way of America, 1996.
31. Moyer A, Verhovsek H, Wilson VL. Facilitating the shift to population-based public health programs: innovation through the use of framework and logic model tools. *Can J Public Health* 1997;88(2):95-8.
32. McLaughlin JA, Jordan GB. Logic models: a tool for telling your program's performance story. *Evaluation and Program Planning* 1999;22(1):65-72.
33. McEwan KL, Bigelow DA. Using a logic model to focus health services on population health goals. *Canadian Journal of Program Evaluation* 1997;12(1):167-74.
34. Julian D. Utilization of the logic model as a system level planning and evaluation device. *Evaluation and Program Planning* 1997;20(3):251-7.
35. Lipsey MW. Theory as method: small theories of treatments. *New Directions for Program Evaluation* 1993;57:5-38.
36. U.S. General Accounting Office. *Designing evaluations*. Washington, DC: U.S. General Accounting Office, 1991; publication no. GAO/PEMD-10.1.4.
37. Taylor-Powell E, Steele S, Douglah M. *Planning a program evaluation*. Madison, WI: University of Wisconsin Cooperative Extension, 1996.
38. Fetterman DM, Kaftarian SJ, Wandersman A, eds. *Empowerment evaluation: knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage Publications, 1996.
39. Preskill HS, Torres RT. *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage Publications, 1999.
40. Patton MQ. Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice* 1997;18(2):147-63.
41. Wandersman A, Morrissey E, Davino K, et al. Comprehensive quality programming and accountability: eight essential strategies for implementing successful prevention programs. *Journal of Primary Prevention* 1998;19(1):3-30.
42. Cousins JB, Whitmore E. Framing participatory evaluation. In: Whitmore E, ed. *Understanding and practicing participatory evaluation*, vol 80. San Francisco, CA: Jossey-Bass, 1998:5-24.
43. Bickman L, Rog DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998.
44. Cook TD, Reichardt CS, eds. *Qualitative and quantitative methods in evaluation research*. Sage research progress series in evaluation, vol 1. Beverly Hills, CA: Sage Publications, 1979.
45. Patton MQ. *Qualitative evaluation and research methods*. 2nd ed. Newbury Park, CA: Sage Publications, 1990.
46. Posavac EJ, Carey RG. *Program evaluation: methods and case studies*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
47. Rossi PH, Freeman HE, Lipsey MW. *Evaluation: a systematic approach*. 6th ed. Thousand Oaks, CA: Sage Publications, 1999.
48. Trochim WMK. *Research methods knowledge base [on-line textbook]*. 2nd ed. 1999. Available at <<http://trochim.human.cornell.edu>>. Accessed June 1999.
49. Boruch RF. Randomized controlled experiments for evaluation and planning. In: Bickman L, Rog DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:161-92.
50. Cook TD, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Boston, MA: Houghton Mifflin, 1979.
51. Reichardt CS, Mark MM. Quasi-experimentation. In: Bickman L, Rog DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:193-228.
52. Yin RK. *Case study research: design and methods*. 2nd ed. Applied Social Research Methods Series, vol 5. Thousand Oaks, CA: Sage Publications, 1994.
53. Yin RK. Abridged version of case study research: design and method. In: Bickman L, Rog DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:229-60.
54. U.S. General Accounting Office. *Case study evaluations*. Washington, DC: U.S. General Accounting Office, 1990; publication no. GAO/PEMD-91-10.1.9.
55. McQueen DV, Anderson LM. What counts as evidence? Issues and debates on evidence relevant to the evaluation of community health promotion programs. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. *Evaluation in health promotion: principles and perspectives*. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
56. Green JC, Caracelli V, eds. *Advances in mixed-method evaluation: the challenges and benefits of integrating diverse paradigms*. *New Directions for Program Evaluation*, vol 74; San Francisco, CA: Jossey-Bass 1997.
57. Frechtling J, Sharp L. *User-friendly handbook for mixed method evaluations*. Arlington, VA: National Science Foundation, 1997.

58. U.S. General Accounting Office. Evaluation synthesis. Washington, DC: U.S. General Accounting Office, 1992. publication no. GAO/PEMD-10.1.2.
59. Newman DL, Brown RD. Applied ethics for program evaluation. Thousand Oaks, CA: Sage Publications, 1996.
60. Fitzpatrick JL, Morris M, eds. Current and emerging ethical challenges in evaluation. *New Directions for Program Evaluation*, vol 82; San Francisco, CA: Jossey-Bass 1999.
61. Newcomer K. Using statistics appropriately. In: Wholey J, Hatry H, Newcomer K, eds. *Handbook of practical program evaluation*. San Francisco, CA: Jossey-Bass, 1994.
62. Basch CE, Silepcevich EM, Gold RS, Duncan DF, Kolbe LJ. Avoiding type III errors in health education program evaluation: a case study. *Health Education Quarterly* 1985;12(4):315-31.
63. Perrin EB, Koshel JJ, eds. *Assessment of performance measures for public health, substance abuse, and mental health*. Washington, DC: National Academy Press, 1997.
64. Innes JE. *Knowledge and public policy: the search for meaningful indicators*. 2nd expanded ed. New Brunswick, NJ: Transaction Publishers, 1990.
65. McRae D Jr. *Policy indicators: links between social science and public debate*. Chapel Hill, NC: University of North Carolina Press, 1985.
66. Institute of Medicine. *Improving health in the community: a role for performance monitoring*. Durch, JS, Bailey LA, Stoto MA, eds. Washington, DC: National Academy Press, 1997.
67. Eddy DM. Performance measurement: problems and solutions. *Health Aff* 1998;17(4):7-25.
68. Harvard Family Research Project. Performance measurement. *Evaluation Exchange* 1998;4(1):1-15.
69. Rugg D. New activities and use of indicators in the evaluation of HIV prevention efforts at CDC. *National Alliance of State and Territorial AIDS Directors (NASTAD) HIV Prevention Community Planning Bulletin*;1997:2-3.
70. Center for Substance Abuse Prevention. *Measurements in prevention: a manual on selecting and using instruments to evaluate prevention programs*. CSAP technical report no. 8 (SMA)93-2041. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, 1993.
71. Goodman RM, Speers MA, McLeroy K, et al. Identifying and defining the dimensions of community capacity to provide a basis for measurement. *Health Educ Behav* 1998;25(3):258-78.
72. Phillips JJ. *Handbook of training evaluation and measurement methods*. 3rd ed. Houston, TX: Gulf Publishing Company, 1997.
73. Knauff EB. What Independent Sector learned from an evaluation of its own hard-to-measure programs. In: Gray ST, ed. *Leadership IS: a vision of evaluation; a report of learnings from Independent Sector's work on evaluation*. Washington, DC: Independent Sector, 1993.
74. Perrin B. Effective use and misuse of performance measurement. *American Journal of Evaluation* 1998;19(3):367-79.
75. U.S. General Accounting Office. Using statistical sampling. Washington, DC: U.S. General Accounting Office, 1992; publication no. GAO/PEMD-10.1.6.
76. Henry GT. Practical sampling. In: Bickman L, Rog DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:101-26.
77. Johnson JC. Selecting ethnographic informants. *Qualitative Research Methods Series*, vol 22; Thousand Oaks, CA: Sage Publications, 1991,
78. Steckler A, McLeroy KR, Goodman RM, Bird ST, McCormick L. Toward integrating qualitative and quantitative methods: an introduction. *Health Education Quarterly* 1992;19(1):191-8.
79. Greene JC. Qualitative program evaluation: practice and promise. In: Denzin NK, Lincoln YS, eds. *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications, 1994.
80. de Vries H, Weijts W, Dijkstra M, Kok G. Utilization of qualitative and quantitative data for health education program planning, implementation, and evaluation: a spiral approach. *Health Education Quarterly* 1992;19:101-15.
81. U.S. General Accounting Office. *Prospective evaluation methods: the prospective evaluation synthesis*. Washington, DC: U.S. General Accounting Office, 1990; publication no. GAO/PEMD-10.1.10.
82. Lipsey MW. What can you build with thousands of bricks? Musings on the cumulation of knowledge in program evaluation. *New Directions for Evaluation* 1997;76:7-23.
83. Lipy MW. Design sensitivity: statistical power for applied experimental research. In: Bickman L, Rob DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:39-68.
84. Sieber JE. Planning ethically responsible research. In: Bickman L, Rob DJ, eds. *Handbook of applied social research methods*. Thousand Oaks, CA: Sage Publications, 1998:127-56.
85. Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook*. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
86. Hennessy M. Evaluation. In: Stroup D, Teutsch S., eds. *Statistics in Public Health*. New York, NY: Oxford University Press, 1998:193-219.

87. Henry GT. Graphing data. In: Bickman L, Rog DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:527-56.
88. Weick KE. Sensemaking in organizations. Thousand Oaks, CA: Sage Publications, 1995.
89. Rogers PJ, Hough G. Improving the effectiveness of evaluations: making the link to organizational theory. *Evaluation and Program Planning* 1995;18(4):321-32.
90. Adler M, Ziglio E. Gazing into the oracle: the Delphi method and its application to social policy and public health. Bristol, PA: Jessica Kingsley Publishers, 1996.
91. Torres RT, Preskill HS, Piontek ME. Evaluation strategies for communicating and reporting: enhancing learning in organizations. Thousand Oaks, CA: Sage Publications, 1996.
92. Zaltman G, Barabba VP. Hearing the voice of the market: competitive advantage through creative use of market information. Boston, MA: Harvard Business School Press, 1990.
93. Shulha LM, Cousins JB. Evaluation use: theory, research, and practice since 1986. *Evaluation Practice* 1997;18(3):195-208.
94. Weiss CH. Have we learned anything new about the use of evaluation? *American Journal of Evaluation* 1998;19(1):21-33.
95. Schorr LB. Common purpose: strengthening families and neighborhoods to rebuild America. New York, NY: Doubleday, Anchor Books, 1997.
96. Fawcett SB, Paine-Andrews A, Francisco VT, et al. Evaluating community initiatives for health and development. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. *Evaluation in health promotion: principles and perspectives*. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
97. Collins J, Rugg D, Kann L, Banspach S, and Kolbe L. Evaluating a national program of school-based HIV prevention. *Evaluation and Program Planning* 1996;19(3):209-18.
98. Rugg D, Buehler J, Renaud M, et al. Evaluating HIV prevention: a framework for national, state, and local levels. *American Journal of Evaluation* 1999;20(1):35-56.
99. Greene J, Lincoln Y, Mathison S, Mertens DM, Ryan K. Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation* 1998;19(1):101-22.
100. Smith NL. The context of investigations in cross-cultural evaluations. *Studies in Educational Evaluation* 1991;17:3-21.
101. Bonnet, DG. An evaluation of year 1 of the American Cancer Society's collaborative evaluation fellows project. Indianapolis, IN: D. Bonnet Associates, 1999.
102. National Institute of Standards and Technology. National Quality Program. National Institute of Standards and Technology, 1999. Available at <<http://www.quality.nist.gov>>. Accessed June 1999.
103. National Institute of Standards and Technology. Baldrige index outperforms S&P 500 for fifth year. National Institute of Standards and Technology, National Quality Program, 1999. Available at <http://www.nist.gov/public_affairs/releases/n99-02.htm>. Accessed June 1999.
104. National Institute of Standards and Technology. Health care criteria for performance excellence. Gaithersburg, MD: National Institute of Standards and Technology, Baldrige National Quality Program, 1999. Also available at <<http://www.quality.nist.gov>>. Accessed June 1999.
105. Public Health Service. Public health workforce: an agenda for the 21st century: a report of the public health functions project. Washington, DC: US Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 1997.
106. CDC. Handbook for evaluating HIV education. Atlanta, GA: US Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, Division of Adolescent and School Health, 1993.
107. Thompson, NJ, McClintock HO. Demonstrating your program's worth: a primer on evaluation for programs to prevent unintentional injury. Atlanta: CDC, National Center for Injury Prevention and Control, 1998.
108. CDC. Guidelines for evaluating surveillance systems. *MMWR* 1988;37(No. SS-5):1-18.
109. Fawcett SB, Sterling TD, Paine-Andrews A, et al. Evaluating community efforts to prevent cardiovascular diseases. Atlanta, GA: U.S. Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, 1995.
110. Rootman I, Goodstadt M, Hyndman B, et al., eds. *Evaluation in health promotion: principles and perspectives*. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
111. CDCynergy (Version 1.0) [CD ROM]. Atlanta, GA: U.S. Department of Health and Human Services, CDC, Office of Communication, 1998.
112. Linney JA, Wandersman A. Prevention plus III: assessing alcohol and other drug prevention programs at the school and community level: a four-step guide to useful program assessment. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, Office for Substance Abuse Prevention, 1991. publication no. (ADM)91-1817.

* The program evaluation standards are an approved standard by the American National Standards Institute (ANSI) and have been endorsed by the American Evaluation Association and 14 other professional organizations (ANSI Standard No. JSEE-PR 1994, Approved March 15, 1994).

** Developed by the Rand Corporation, the Delphi process is an iterative method for arriving at a consensus concerning an issue or problem by circulating questions and responses to a panel of qualified reviewers whose identities are usually not revealed to one another. The questions and responses are progressively refined with each round until a viable option or solution is reached.

*** The Malcolm Baldrige National Quality Improvement Act of 1987 (Public Law 100-107) established a public-private partnership focused on encouraging American business and other organizations to practice effective quality management. The annual award process, which involves external review as well as self-assessment against Criteria for Performance Excellence, provides a proven course for organizations to improve significantly the quality of their goods and services.