

REVISIÓN DE TEMA**EVALUACIÓN DE PROGRAMAS DE SALUD****Alfonso Tenorio Gnecco *, Regina Victoria Plaza******RESUMEN**

Está plenamente aceptado que la implementación de cualquier intervención, programa, iniciativa o proyecto en salud debe contar desde su diseño con una metodología de evaluación que permita determinar el cumplimiento de sus objetivos, así como generar elementos que faciliten la toma de decisiones. Sin embargo, en términos generales esto no se cumple y en ocasiones los programas o proyectos no son evaluados en ninguna de sus fases de implementación, o se evalúan tardíamente sin contar con los insumos necesarios para este propósito.

No obstante la importancia de este tema, llama la atención la escasez de bibliografía internacional sobre la evaluación de programas de salud pública. En este artículo se ofrece una amplia revisión del tema, considerando no solamente aspectos relacionados con la definición, tipos de evaluación, métodos, sino también con la metodología de descripción de programas, el costeo y el análisis. En este artículo se propone una definición de evaluación que fusiona conceptos de diferentes autores y pretende ser más completa que las actualmente conocidas.

Palabras clave: *evaluación, programas de salud, metodología, salud pública.*

ABSTRACT

He fully accepted that the implementation of any intervention programme, an initiative or project on health should run from their design with an evaluation methodology for determining compliance with its objectives and generate elements that will facilitate decision-making. However, overall this is not fulfilled and sometimes programmes or projects are not evaluated at any stage of implementation, or are assessed late without having the necessary inputs for this purpose. Notwithstanding the importance of this issue, striking shortage of international literature on the assessment of public health programs. This article provides a comprehensive review of the topic, considering not only issues related to the definition, types of evaluation methods, but also with the methodology description of programs, costing and analysis. This article proposes a definition of assessment that blends concepts of different authors and pretend to be more complete than those currently known.

Key Words: *assessment, health programs, methodology, public health*

Recibido para evaluación: enero 24 de 2008. **Aprobado para publicación:** junio 15 de 2008.

* Médico Internista, docente del Departamento de Medicina Interna, Facultad de Ciencias de la Salud, Universidad del Cauca
** Médica y cirujana, docente del Departamento de Medicina Interna, Facultad de Ciencias de la Salud, Universidad del Cauca.

Correspondencia: Dr. Alfonso Tenorio., Carrera 11 No. 18N-66, Popayán, Cauca, Colombia. E mail: alftenorio@gmail.com

INTRODUCCIÓN

Es importante reconocer que la evaluación de los programas de salud permite medir los logros alcanzados, la utilidad de los mismos y en general, permite mejorar el desempeño e incluso el impacto del programa. Es imposible conocer la evolución de un programa o intervención en salud si no tenemos métodos de evaluación que nos permitan ajustar las estrategias, realizar cambios, optimizar los recursos y facilitar la toma de decisiones basadas en resultados.

Podemos evaluar un programa desde dos tipos de abordajes, la evaluación formativa que permite modificar y afinar un programa en ejecución, y la sumativa que evalúa programas finalizados. Esta última permite que el tomador de decisiones defina cambios importantes para el nuevo periodo, continúe o modifique el programa, de acuerdo a la correlación entre los objetivos planeados, la forma de ejecución y los resultados obtenidos.

Reconociendo que el componente de evaluación de cualquier programa o estrategia en salud es fundamental, es necesario que los funcionarios que están involucrados en el diseño e implementación de ellos, planifiquen desde un inicio la metodología de la evaluación que se llevará a cabo en el transcurso de la implementación o a su finalización.

DEFINICIÓN DE EVALUACIÓN

Rossi y Freeman consideran que la evaluación es la aplicación sistemática de procedimientos de investigación para determinar la conceptualización, el diseño, la implementación y la utilidad de los programas, mientras que Patton la define como la sistemática recopilación de información relacionada con las características y resultados de programas implementados en poblaciones específicas, y así mejorar la efectividad del programa y facilitar la toma de decisiones. Cronbach define la evaluación como la compilación y uso de la información que facilite la toma de decisiones en un programa.

Cuando se pretende evaluar los sistemas de salud, es necesario diferenciar los tres niveles en los cuales se pueden describir lo evaluado: programa, proyecto e intervención. Usando un ejemplo muy claro, el de la tuberculosis, existen programas de prevención y control en diferentes niveles, nacional, departamental y municipal, pero a su vez, los programas pueden adoptar diversos proyectos o estrategias de intervención. La Organización Mundial de la Salud (OMS) recomendó que todos los programas de tuberculosis, a todos los niveles, implementaran la estrategia de Tratamiento Acortado Supervisado (DOTS) y así lo hicieron

la mayoría de países desde 1995 a 2005. Actualmente se recomienda la implementación de la Estrategia Mundial "Alto a la Tuberculosis", la cual incluye los elementos de la estrategia DOTS como fundamentales. A su vez, cada estrategia o proyecto se compone de diversos elementos, por ejemplo el suministro supervisado de medicamentos es sólo un elemento de los 6 que componen la estrategia Alto a la TB. Aunque es claro que la evaluación se puede hacer en cualquiera de los niveles (programa, estrategia o intervención), para efectos de esta revisión, se usará el término 'programa' para definir las intervenciones múltiples orientadas a lograr una meta general.

Es necesario distinguir la evaluación de programas, de la investigación evaluativa y de la investigación básica. La investigación evaluativa es la aplicación de métodos de investigación a la evaluación de programas para llegar a conclusiones sobre la eficacia de un programa. Weiss considera que el propósito de la investigación evaluativa es medir los efectos de un programa al compararlos con las metas que se propuso lograr y así contribuir a la toma de decisiones del programa y a la optimización futura de los recursos. La evaluación de programas es el proceso por el cual se genera información sobre las operaciones y el impacto de la aplicación de programas o políticas. La investigación básica es un requisito para muchos procesos de investigación evaluativa en el sentido de proporcionar el sustrato sobre el cual se diseñan los programas.

La definición más ajustada al proceso de evaluación de programas de salud en Colombia es el producto de la fusión de definiciones de Rossi, Freeman y Cronbach, así: evaluación es la aplicación sistemática de procedimientos de investigación para determinar la conceptualización, el diseño, la implementación y la utilidad de los programas, que facilite la toma de decisiones.

¿POR QUÉ EVALUAR UN PROGRAMA?

El objetivo principal de una evaluación es ayudar en el proceso de toma de decisiones. Los tomadores de decisiones están cada vez más convencidos de la importancia de las evaluaciones como un instrumento útil que va a permitir cambios o ajustes en las políticas de implementación de los programas.

Existen varios objetivos en los estudios de evaluación, entre ellos: determinar si las metas son alcanzadas y medir tanto el nivel del logro, así como la forma como se ejecutó; comparar la eficiencia y adecuación de varios programas y métodos que tienen iguales objetivos o en otras palabras

emplear controles para la comparación; proporcionar controles de calidad; examinar la efectividad de diferentes estructuras organizativas y sistemas de prestación de servicios; proporcionar conocimiento científico y proponer hipótesis para estudios futuros; y desarrollar nuevos abordajes y estrategias para futuros programas. Una de las maneras como se logra hacer salud pública basada en la evidencia, es a través de la evaluación de programas.

Pero de manera más simple, la evaluación de un programa sirve para determinar su continuidad, así como mejorar su desempeño, añadir o suprimir algunas actividades, asignar recursos, generar conocimiento y, muy importante, replicar el programa en otros lugares y otros contextos.

TIPOS DE EVALUACIÓN

En cuanto al momento en el que se hace la evaluación, existen dos abordajes grandes, la sumativa relacionada con programas finalizados y la formativa que permite evaluar un programa en la ejecución del proceso para afinar las acciones de implementación. La evaluación sumativa permite que los responsables de los programas de salud pública tomen decisiones relacionadas con continuar, cambiar, expandir o finalizar un proyecto, un programa o una intervención. Los resultados de las evaluaciones serán usados para tomar decisiones sobre el programa evaluado.

Potvin divide las evaluaciones en otros dos tipos: las de proceso y las de resultado. Aunque la literatura convencional las pone en polos opuestos, las evaluaciones de proceso, también consideradas formativas por otros, se relacionan con el control, aseguran la calidad de la práctica y son usadas para evaluar cómo están implementados los procesos (responsables, acciones que se ofrecen, población objeto y bajo qué condiciones). La evaluación con base en resultados se entiende como un abordaje de comparación enfocado tanto en los logros como en los procesos, que se desprenden de la misión, objetivos y actividades del programa.

Habicht et al. proponen un modelo de evaluación basado en dos ejes, el primero de ellos relacionado con los indicadores de interés como los de provisión, utilización de servicios, cobertura o impacto, y un segundo eje relacionado con el grado de inferencia: adecuación, plausibilidad y probabilidad.

El eje de inferencia tiene dos componentes, uno categórico y otro relacionado con la fuerza de la inferencia. El categórico se relaciona con los tres diseños, el de adecuación, el de probabilidad y el de plausibilidad, mientras el eje de la fuerza inferencial depende de la existencia o no de un grupo

de control. Este poder inferencial aumenta desde estudios sin grupo de control, pasando por estudios con grupo de control probablemente sesgados, hasta estudios con grupos de control aleatorizados.

Las consideraciones sobre la inferencia se relacionan con la confianza con la cual se puede afirmar que los resultados observados son o no debidos al programa o la intervención. La evaluación por adecuación se limita a determinar si los cambios esperados como consecuencia del programa ocurrieron o no. La inferencia relacionada con la adecuación de un programa depende de la comparación del desempeño o impacto de un proyecto con criterios de adecuación previamente establecidos. En caso que no existan criterios estandarizados con los cuales se pueda comparar la adecuación de un programa, el desempeño o impacto puede ser evaluado midiendo las tendencias generales en el tiempo, como por ejemplo la disminución de las tasas de incidencia o incremento de coberturas.

Los diseños de evaluación por adecuación no requieren grupos de control cuando sus resultados son comparados con criterios previamente establecidos. Mientras que para evaluar los cambios de adecuación en el tiempo se requiere por lo menos de dos mediciones, lo cual aumenta la complejidad y los costos del diseño. Sin embargo, está claramente definido que los diseños de plausibilidad y probabilidad son mucho más costosos.

El tipo de evaluación de adecuación por desempeño incluye indicadores como provisión, utilización y cobertura. Este diseño valora cómo las actividades del programa cumplieron con los objetivos esperados. Las evaluaciones de adecuación pueden ser transversales o longitudinales. Las primeras incluyen una medición única durante o al final del programa, mientras las longitudinales requieren una línea de base o hacer mediciones repetidas en el tiempo para determinar las tendencias.

El otro diseño de evaluación de adecuación incluye indicadores de impacto, el cual valora si los indicadores de salud han mejorado entre los beneficiarios de los programas o entre las poblaciones objeto. De igual manera que en los diseños de adecuación por desempeño, los diseños de impacto pueden hacerse de manera transversal o longitudinal. Aunque los diseños de evaluación por adecuación no permiten hacer asociaciones de causalidad entre los cambios observados y las actividades realizadas, sí permiten definir claramente si se cumplieron los objetivos preestablecidos.

El diseño de evaluación por plausibilidad responde a la pregunta: ¿parece que el programa ha tenido un efecto más allá

de las influencias externas? La valoración por plausibilidad va más allá de la lograda por adecuación, pues permite descartar las influencias de factores externos sobre los resultados. El control de factores de confusión se logra a través de grupos de control que se escogen antes de iniciar la evaluación o durante el análisis de los datos. Los grupos de control que se usan en los diseños de evaluación por plausibilidad incluyen grupos de control históricos, los cuales toman la misma institución o población y comparan los resultados antes y después de la implementación del programa. Otros diseños incluyen grupos de control interno, los que consideran subgrupos de población quienes han recibido la intervención en diferentes grados de intensidad. Los diseños que utilizan grupos de control externo comparan los resultados de la intervención en áreas o poblaciones que no han recibido la intervención con aquellas que si la recibieron.

Los diseños de evaluación por plausibilidad toman en cuenta también indicadores de desempeño (provisión, utilización y cobertura) y de impacto. El enfoque puede ser a través de modelos transversales, longitudinales y longitudinales con control para el caso de los indicadores de desempeño, y transversales, longitudinales, longitudinales con control, y casos y controles cuando los indicadores son de impacto.

Los diseños de evaluación por probabilidad pretenden asegurar que los resultados obtenidos son explicados por el programa y no debidos a factores de confusión, sesgos o al azar. Una de las grandes limitaciones para usar este diseño es que requiere que el evaluador esté presente antes del inicio del programa para aleatorizar los grupos de acuerdo a la intervención. El rigor de este diseño de evaluación, a menudo lleva a que las condiciones en las que se desarrolla sean artificiosas o diferentes a las condiciones reales, lo que dificulta la extrapolación de los resultados, afectando la validez externa. Los grandes esfuerzos para aumentar la validez interna de este diseño, pueden menoscabar la validez externa y su utilidad en el momento de tomar decisiones basadas en sus resultados.

Dado que los diseños de probabilidad y plausibilidad pueden afirmar con seguridad que los efectos se debieron al programa, sin entrar a detallar si fueron adecuados o no, se recomienda que estos dos tipos de diseño incluyan idealmente objetivos de adecuación, que incluso son de muy bajo costo. Por el contrario, no es lógico adicionar objetivos de plausibilidad a los diseños de evaluación por probabilidad, ya que ambos pretenden inferir que los resultados son consecuencia de la intervención. El diseño de evaluación por probabilidad lo hace intentando excluir otras explicaciones para los hallazgos y el diseño de evaluación de plausibilidad, a través de análisis estadísticos.

Otros autores utilizan los términos de evaluación experimental, cuasi-experimental y no experimental de manera muy similar a los expuestos en adecuación, plausibilidad y probabilidad. Los diseños experimentales comparan los resultados de dos o más grupos aleatorios equivalentes, mientras que los diseños cuasi-experimentales usan grupos de comparación pero no asignados aleatoriamente.

MÉTODOS DE EVALUACIÓN

Cada uno de los diseños descritos en la literatura para la evaluación de programas propone un método de evaluación específico con algunas variaciones entre ellos, sin embargo es Rottman y colaboradores quien resume en términos generales los aspectos más importantes que deben ser considerados en el momento de evaluar un programa, independiente del diseño empleado. Las fases metodológicas sugeridas por este autor se sintetizan en lo siguiente:

- Descripción del programa a estudiar.
- Identificación de los temas o preguntas a tener en cuenta en la evaluación.
- Definición del proceso de obtención de la información requerida.
- Recolección de la información según lo definido.
- Análisis de la información recolectada.
- Recomendaciones a los tomadores de decisión.
- Divulgación de los hallazgos a los tomadores de decisiones y otras audiencias.

DESCRIPCIÓN DE LOS PROGRAMAS

En cuanto a la descripción de un programa, Potvin afirma que los programas de prevención y control de las enfermedades no deben ser considerados como dispositivos que se distribuyen y se usan por los profesionales de la salud con mínimas referencias al entorno en el cual se implementan. Al contrario, cada programa de prevención y control constituye una organización social y debe ser caracterizado como un sistema abierto y dinámico.

Potvin propone que los programas de salud pueden ser estudiados como sistemas dinámicos a través de tres dimensiones: dimensión contextual, componentes del programa y la evolución del programa. La dimensión contextual comprende la interacción entre el programa y el entorno, de tal manera que el programa procura modificar algunos aspectos del entorno (metas de cambio) y a su vez, el programa es influenciado por elementos contextuales. Las características internas de los programas son los objetivos,

los componentes estructurales (recursos y servicios provistos) y la transformación de procesos. La evolución del programa hace referencia a los cambios que ocurren en el transcurso de su implementación.

Interacciones entre el programa y el entorno

Por entorno se entiende todos los aspectos del sistema social en los cuales un programa es implementado, tales como actores y recursos sociales. Además de los aspectos sociales, también incluye los económicos, culturales y políticos en donde se enmarca el programa. Los sistemas sociales pueden ser comunidades, provincias, municipalidades o ciudades. En el entorno se identifican dos factores principales: las condiciones del entorno y la población objetivo en la que se espera el cambio, las cuales están divididas en condiciones iniciales y condiciones resultantes.

Un modelo actualmente vigente es la participación de la comunidad en las diferentes fases de la implementación de los programas, lo que es sin duda una de las más importantes evidencias de la interacción que existe entre los programas y la comunidad.

Componentes del programa

Las categorías en las cuales se incluyen los componentes de los programas son objetivos, recursos y actividades/servicios. Los objetivos o resultados esperados son la forma que da dirección al programa y son formulados en términos de cambios particulares en el entorno. Los recursos son la materia prima, o sea el conocimiento, dinero, equipos e infraestructura física necesarios para la operativización del programa. Las actividades y servicios son el medio por el cual el programa persigue sus objetivos y comprenden las estrategias de intervención.

Evolución del programa

En los programas se identifican tres fases de evolución en el tiempo, la fase de diseño, de implementación y la de terminación. La fase de diseño incluye aspectos relacionados con las metas, identificación de la población objeto, estudios de factibilidad y de aceptabilidad de las estrategias de intervención, determinar los objetivos, establecer un cronograma, establecer un presupuesto, las fuentes de financiamiento y la manera como se van a entregar los servicios. Los aspectos externos al programa que se relacionan con esta fase son las fuentes de financiamiento, las políticas en salud o normativas a las cuales el programa debe ajustarse y la situación epidemiológica previa a la implementación del programa.

En la segunda fase o fase de implementación se canalizan todos los recursos hacia la población objetivo a través de las actividades y servicios prestados. Esta fase es la más larga y en ella ocurren las mayores interacciones entre el programa y el entorno. El programa entra en su fase final de terminación, ya sea porque cumplió sus objetivos y las fuentes se agotaron, o porque entra en otro ciclo con actividades y objetivos diferentes.

Preguntas de evaluación

Las preguntas de evaluación se pueden enfocar tanto a los componentes del programa como a las interacciones entre el programa y el entorno, y deben ser planteadas en fases específicas del ciclo del programa. Así, preguntas relacionadas con la coherencia y pertinencia son apropiadas para la fase de diseño, mientras que las preguntas relacionadas con los logros se aplican principalmente en la fase de implementación, sin embargo, las preguntas relacionadas con los resultados no están ligadas a ninguna fase específica del programa.

Los responsables de diseñar la evaluación deben trabajar en conjunto con aquellos que tienen la capacidad para tomar las decisiones respecto a los programas evaluados. La manera útil para abordar las evaluaciones de salud es preguntar "qué va a ser evaluado", lo cual permite definir "qué es lo que se medirá". Los diseños de evaluación no están obligados a considerar todos los aspectos de manera simultánea, y más bien pueden ser enfocados a puntos específicos de interés. Las preguntas de evaluación se pueden categorizar según el uso que se le va dar a los resultados o el tipo de indicadores de desempeño que se van a usar. Las primeras pueden ser de enfoque formativo o sumativo, y las segundas son evaluaciones de proceso o de resultados.

Momento de evaluación y cronogramas

Algunos diseños de evaluación de programas requieren que los evaluadores estén presentes desde antes del inicio del programa, este es el caso de los diseños de probabilidad. Sin embargo, en los otros diseños con enfoque longitudinal se requiere información de línea de base que permitirá hacer las comparaciones. Por lo general, estudios con indicadores de provisión y utilización son elaborados en menos tiempo y permiten tomar decisiones más rápidamente. Por otro lado, los estudios de impacto y cobertura son más demorados, requieren que la intervención haya acabado por lo menos un ciclo. Se afirma que por lo menos 3 a 5 años son necesarios para que una intervención muestre un impacto, e incluso varios años o décadas se necesitarán para demostrar un impacto en intervenciones relacionadas con enfermedades con largos periodos de incubación, enfermedades crónicas

o efectos generacionales de intervenciones nutricionales.

Es una regla general que las intervenciones en salud pública tengan un plan de evaluación desde el principio, sin embargo diseños de evaluación por adecuación y plausibilidad pueden desarrollarse cuando la intervención ya se ha hecho. Los evaluadores deben tener en cuenta el tiempo en el cual los resultados se necesitan, de tal manera que puedan ser conocidos por los tomadores de decisión de forma oportuna. Se recomienda, por tanto, que el diseño de evaluación incluya un cronograma que garantice que los resultados serán entregados oportunamente.

Muestreo

A diferencia de los estándares aceptados sobre error alfa y beta en la epidemiología convencional, el nivel de error aceptado para estos estudios depende del grado de certeza que desee el tomador de decisiones a quien se le presentarán los resultados. Sin duda, a menor nivel de error alfa y beta, mayor será la muestra, lo cual tiene implicaciones de costo.

Es aceptado que en diseños de evaluación de programas de salud pública se establezcan niveles de error más altos que los que se usan en ensayos de eficacia, siempre y cuando los resultados no generen inferencias del orden clínico, en donde se requieren bajos niveles de error alfa.

Costo

Es suficientemente claro que el costo es uno de los aspectos que más influyen en la escogencia del diseño de evaluación, especialmente porque en muchas ocasiones las actividades relacionadas con la evaluación no están consideradas en el presupuesto del programa que se va a evaluar. El tamaño de la muestra y la frecuencia con que se haga el muestreo es uno de los factores más importantes para determinar el costo del estudio. Los evaluadores deben discutir con el tomador de decisiones algunos aspectos relacionados con la complejidad de los análisis que se harán, si es completamente necesario hacer una evaluación sumativa total, si el diseño incluirá grupos de control o simplemente comparación de un antes y un después, y que tan infrecuente es el evento que se va a medir.

ESCOGIENDO UN DISEÑO DE EVALUACIÓN

Uno de los aspectos más importantes en el momento de decidir un diseño de evaluación es conocer claramente el tipo de

personas quienes van a recibir los resultados de evaluación y que tienen la responsabilidad de tomar decisiones.

El grado de complejidad del diseño de evaluación escogido dependerá de lo que se desee evaluar. Si se desea evaluar la disponibilidad de un servicio, bastará con usar indicadores de provisión en un diseño de adecuación, pero si el propósito es evaluar el grado de utilización de este servicio, se deberían usar indicadores de utilización también en un diseño de adecuación. Por otro lado, si lo que se desea es establecer la cobertura alcanzada con una intervención o medir su impacto, la evaluación se fundamentará en indicadores de cobertura o de impacto, lo que se puede lograr a través de diseños de adecuación y plausibilidad. Este enfoque escalonado para escoger el diseño de evaluación permite racionalizar los recursos disponibles, optimizar el tiempo y considerar aspectos administrativos y políticos de mucho interés para el tomador de decisiones.

Además de los indicadores que se deseen medir y el grado de certeza que se desee, existen otros factores que deben ser tomados en cuenta en el momento de elegir el tipo de diseño de evaluación. Estos incluyen la eficacia de intervenciones a gran escala, el área de conocimiento en el que se está trabajando y el tiempo disponible para la evaluación. Si la eficacia de una intervención es conocida, diseños para determinar el impacto o la causalidad entre la intervención y el resultado serían innecesarios. Cuando se conoce la eficacia de la intervención, pero queremos evaluar las actividades relacionadas con su implementación, bastará con aplicar un diseño de adecuación que nos permita determinar la provisión, utilización y cobertura de la intervención.

ANÁLISIS

Análisis de la coherencia

Determinar si un programa es coherente se logra analizando, primero, la relación existente entre los objetivos del programa, las actividades y servicios planeados y los recursos, y generalmente se determina en la fase de diseño, pero también analizando las interacciones entre el programa y el entorno.

Análisis de la pertinencia del programa

Las preguntas de pertinencia son formuladas durante la fase de diseño y están dirigidas a determinar la relación entre el programa y sus metas. Este enfoque evalúa la conveniencia de los objetivos con relación a los problemas de salud y las condiciones de vida de la población objeto. En la literatura

se conoce también como evaluación estratégica. Usualmente los diseñadores de programas acuden a la literatura científica para determinar la pertinencia de las metas, pero también consultan a la población objeto sobre las hipótesis de la implementación de la intervención.

Análisis de la receptividad del programa

Las preguntas de receptividad son formuladas durante la fase de implementación. Estas evaluaciones pretenden definir cómo las condiciones de contexto modifican varios componentes del programa. En la literatura se evidencia cada vez más la tendencia de estrechar el enfoque de este tipo de evaluación, concentrándose en cómo el programa fue implementado, qué tan bien las actividades se ajustan al diseño original, quiénes recibieron los servicios prestados, la magnitud de la cobertura de la intervención y los factores externos que compiten con los efectos del programa.

Análisis de los logros del programa

Las preguntas relacionadas con los logros se relacionan con el programa en sí mismo y se formulan durante la fase de implementación. Estas preguntas procuran dar respuesta a si el programa cumple el plan de acción. Dos temas pueden ser abordados a través de las preguntas de logros, uno relacionado con el tema de la calidad y otro con el tema de la productividad. El tema de la calidad se relaciona con las características organizacionales y estructurales de la prestación del servicio, tales como la disponibilidad, accesibilidad y calidad técnica, o también se puede relacionar con la cobertura en la población objetivo y la satisfacción del cliente. El tema de la productividad tiene que ver con el manejo eficiente de los recursos, o en otras palabras por el recurso necesario para producir una actividad dada. La evaluación de logros adopta diferentes nombres según el diseño, así como también se le conoce como monitoreo del programa, evaluación de procesos o evaluación formativa.

Análisis de los resultados del programa

Análisis de resultados es una evaluación que examina los cambios en el entorno asociados con el programa. Los resultados que se evalúan pueden incluir tanto aquellos esperados como los no previstos. Se han identificado dos tipos de dificultades asociados con esta modalidad de evaluación, una es la capacidad para atribuir los cambios documentados a la existencia del programa y la otra tiene que ver con la selección de los indicadores para evaluar los efectos.

Algunos autores consideran que los modelos de evaluación por resultados se oponen a los modelos de evaluación de

impacto o proceso, y consideran que solo la evaluación de resultados refleja el verdadero efecto del programa, descuidando el valor de los procesos intermedios, de los cuales finalmente depende el resultado final.

Ante estas dificultades relacionadas con la conceptualización de la evaluación, otros autores proponen que la evaluación con base en resultados se fundamenta en la comparación enfocada tanto en los procesos como en los logros obtenidos a la luz de la misión estratégica del programa.

El análisis de resultados puede incluir indicadores de efectividad, impacto y relación costo-beneficio con el propósito de identificar si el programa logra sus objetivos y metas, si el programa o servicio genera una diferencia significativa con otra intervención o si el programa permite recuperar la inversión.

El enfoque tradicional de la evaluación de programas, diferencia la evaluación de procesos y resultados del monitoreo administrativo, el cual se fundamenta en verificar que los recursos del programa sean transformados en servicios y actividades, tratando de documentar y justificar los gastos. Las evaluaciones de resultados procuran documentar si un programa produjo los cambios planeados. El desafío del evaluador es tener las herramientas válidas para afirmar la asociación causal entre el programa y los cambios observados en el entorno. Tratar de determinar la causalidad en un modelo de evaluación de resultados requiere usualmente de métodos epidemiológicos y estadísticos complejos.

Desde un punto de vista tradicional, las evaluaciones de proceso analizan las actividades del programa tal como fueron implementadas, y de este modo sirven para tres propósitos generales. El primero es mejorar la capacidad de los componentes del programa para lograr la cobertura en la población objeto; segundo, dar explicación a las discrepancias encontradas entre lo que se planeó y lo que se implementó; y tercero, proveer información necesaria para interpretar los resultados.

En conclusión, la evaluación de los programas es un elemento fundamental que permite determinar si los recursos, actividades, estrategias e intervenciones empleadas tienen un real impacto y permiten alcanzar los objetivos propuestos en la población escogida. Ejecutar un programa de salud en beneficio de una población es bueno, ejecutarlo y documentar sus resultados es mejor, pero ejecutarlo, documentar sus resultados y compararlos con lo que se pretendía alcanzar, es óptimo.

BIBLIOGRAFÍA RECOMENDADA

1. **de Salazar L, Díaz C.** La evaluación-sistematización: una propuesta metodológica para la evaluación en promoción de la salud. Un estudio de caso en Cali, Colombia. *Ciência & Saúde Coletiva*, 9(3):545-555, 2004
2. **Habicht JP, Victora CG, Vaughan JP.** Evaluation designs for adequacy, plausibility and probability for public health programme performance and impact. *IJE*, 1999; 28(1):10-18.
3. **Ligia de Salazar.** Evaluación de Efectividad en Promoción de la Salud. Guía de Evaluación Rápida. Universidad del Valle, CEDETES, Cali, 2004
4. **Gustavo Bergonzoli.** Propuesta: La epidemiología y la planificación local: Medidas para la evaluación del impacto potencial. *Colomb Med* 2005; 36: 44-49
5. **Eiliana Montero Rojas.** Marco conceptual para la evaluación de programas de salud. *Centro Centroamericano de Población*, 2004
6. **Rossi PH, Freeman HE.** Evaluation. Newbury Park: Sage Publications. 1993
7. **Habicht JP, Mason JB, Tabatabai H.** Basic concepts for the design of evaluation during programme implementation. In: Sahn DE, Lockwood R. *Methods for the evaluation of impact of Food and Nutrition Programmes*. Tokyo: United Nations University, 1984, 1-25.
8. **Rothman KJ.** Modern Epidemiology. Boston: Little, Brown and Company, 1986.
9. **Rothman KJ, Greenland S.** Modern Epidemiology, 2edn. Philadelphia: Lippincott Raven, 1998.
10. **Rootman, Irving, Ed.; Goodstadt, Michael, Ed.; Hyndman, Brian, Ed.; McQueen, David V., Ed.; Potvin, Louise, Ed.; Springett, Jane, Ed.; Ziglio, Erio, Ed.** Evaluation in Health Promotion: Principles and Perspectives. *WHO Regional Publications, European Series*, 2001
11. Organización Mundial de la Salud. Cuadragésima cuarta Asamblea Mundial de la Salud. WHA44/1991/REC/1.1991.
12. Guía para la evaluación de un programa nacional de tuberculosis. Organización Mundial de la Salud (OMS). WHO/TB/98.240. Ginebra, 1998.
13. **Rodrigo T, Cayla JA, Galdos Tanguis, Jansa JM, Brugal T, Garcia de Olalla P** Evaluación de los Programas de Tuberculosis de las Comunidades Autónomas de España. *Med Clin (Barc)* 1999; 113: 604 – 607
14. **Rodrigo T, Cayla JA, Galdos Tanguis, Jansa JM, Garcia de Olalla P, Brugal T.** Proposing indicators for evaluation of a tuberculosis control programme in large cities based on the experiencia. *Int J Tuberc Lung Dis* 2001; 5 : 1- 9
15. **Rodrigo T , Cayla JA.** por el Grupo de Trabajo para Evaluar Programas de Control de Tuberculosis. Efectividad de los programas de control de tuberculosis en España. *Med Clin (Barc)*, 2003; 121:375-7.
16. **Frenk J.** El concepto y la medición de accesibilidad. Investigaciones sobre servicios de salud. Una antología. Pag 929-43
17. **Cruz LF, Guerrero R. y Alzate A.** La salud como empresa social. Aproximaciones, reflexiones y propuestas. Fundación Carvajal. Universidad del Valle. 2000.
18. **Donabedian, A.** La calidad de la atención médica. Definición y métodos de evaluación. La Prensa Médica Mexicana, México, 1984.
19. **Hennigan KM, Flay BR, Haag RA.** Esclarecimiento de conceptos y términos usuales en la investigación evaluativa. En **Klein R, Read M, Riecken HW.** Evaluación del impacto de los programas de nutrición y de salud. Organización Panamericana de la Salud (OPS/OMS). Washington DC, 1982: 200-218.