

## Marco conceptual para la evaluación de programas de salud

*Eiliana Montero Rojas*<sup>1</sup>

### RESUMEN

La evaluación como disciplina es un área relativamente nueva de las Ciencias Sociales; consecuentemente, su incorporación explícita en los programas y sistemas de salud ha llegado algo tardíamente. Este artículo presenta elementos de un marco conceptual para concebir y utilizar la evaluación en programas de salud. Se define la evaluación como una herramienta para la toma de decisiones, se discuten tendencias y temáticas de actualidad en el campo y también conceptos clave tales como teoría del programa y valoración de necesidades. Finalmente, se esboza una primera aproximación a dos constructos de gran interés en la evaluación de programas de salud: calidad y equidad.

### ABSTRACT

Evaluation, as a discipline, is an relatively new area of the Social Sciences, thus, its explicit incorporation in health programs and systems has appeared somehow late. This article presents elements of a conceptual framework to conceive and use the evaluation in health programs. Evaluation is defined as a tool for decision making, current tendencies and themes in the field are discussed, as well as key concepts such as program theory and needs assessment. Finally, a first approach to two key concepts in evaluation of health programs, quality and equity, is presented.

### ¿QUÉ ES EVALUACIÓN?

El concepto de evaluación es quizá una de las nociones más sujetas a equívocos y controversias. Para muchos, la sola palabra “evaluación” trae a la memoria experiencias juveniles no muy placenteras asociadas a pruebas escolares y asignación de notas. En contextos institucionales es común la creencia de que una evaluación se debe realizar cuando se sospecha que las cosas no andan bien en la ejecución de un programa o proyecto. La sola mención de la palabra puede poner nerviosas a muchas personas, especialmente a los gestores y ejecutores de los programas.

Es cierto que la evaluación puede a veces ser empleada equivocadamente como un “instrumento de castigo”, cuando, por ejemplo, el profesor amenaza a sus alumnos con la administración de una prueba, o cuando se usa como una forma de auditoría. Sin embargo, no son éstos sus usos recomendados y quizá es precisamente la mala práctica de la evaluación la que ha contribuido a darle a este concepto una imagen punitiva y poco favorable.

Por el contrario, la evaluación debe ser parte integral del diseño y desarrollo de cualquier programa o intervención que pretenda aliviar o solucionar problemas asociados a la calidad de

---

<sup>1</sup> PhD. Evaluación Educativa, Florida State University, Investigadora Instituto de Investigaciones Psicológicas, Universidad de Costa Rica, emontero@cariari.ucr.ac.cr

vida de las personas. La OECD (1998) indica que “es una valoración tan sistemática y objetiva como sea posible de un proyecto, programa o política que se está desarrollando o se ha completado, su diseño, implementación y resultados. La evaluación debería brindar información que es creíble y útil, posibilitando la incorporación de las lecciones aprendidas en el proceso de toma de decisiones”. Por su parte, Rossi et al (1999) definen evaluación como “el uso de procedimientos científicos para sistemáticamente investigar la efectividad de los programas. La evaluación es un instrumento para estudiar, comprender y ayudar a mejorar los programas en todos sus aspectos importantes, incluyendo el diagnóstico de los problemas a los que se dirigen, su conceptualización y diseño, su implementación y administración, sus resultados y su eficiencia”.

A partir de estas definiciones resulta claro que la evaluación es una herramienta fundamental para iluminar la toma de decisiones en las diferentes fases de ejecución de los programas e intervenciones, desde su concepción y diseño hasta su posterior valoración en términos de logros e impactos.

Evaluar implica establecer un juicio sobre el valor o mérito de algo, una conclusión valorativa. Para ello se requiere siempre de la comparación con algún estándar o punto de referencia, que ha sido establecido explícita o implícitamente. Determinar, por ejemplo, que un cierto programa de inmunizaciones tiene un 82% de cobertura no es una evaluación, sino solo una medición. La evaluación se realiza cuando se compara este resultado contra el nivel planeado, digamos 90% y se establece un juicio valorativo con base en esta comparación. El establecimiento de significados sustantivos para la interpretación a partir de la comparación entre resultados obtenidos y planeados es, en ocasiones, un proceso de cierta complejidad que requiere del concurso experto (Mohr, 1995). Se da con cierta frecuencia que ni los mismos ejecutores de programas han determinado con precisión cuales son esos niveles esperados de desempeño del programa, es entonces tarea de los evaluadores determinar estos valores de manera más exacta.

Para garantizar la validez de sus interpretaciones la evaluación debe regirse por criterios de rigurosidad científica. De esta forma los evaluadores utilizan teorías, métodos y técnicas de la ciencia, especialmente de las ciencias sociales. En evaluación, al igual que en una investigación científica, se debe recolectar, analizar e interpretar evidencia empírica para fundamentar las conclusiones. Entre más sólida sea la evidencia más credibilidad tendrá el juicio evaluativo.

Más específicamente la evaluación de programas permite contestar preguntas como las siguientes (Rossi et al, 1999):

- 1- Sobre la necesidad por los servicios (valoración de necesidades): ¿Hay necesidad de este programa? ¿Cuáles son las condiciones sociales y de salud a las que debe dirigirse el programa?
- 2- Sobre la conceptualización del programa y su diseño: ¿Está el programa apropiadamente diseñado?
- 3- Sobre las operaciones del programa y la prestación de los servicios: ¿Se está implementando el programa adecuadamente?
- 4- Sobre los resultados e impactos: ¿Realmente tiene el programa los resultados deseados? ¿Hay efectos colaterales no esperados?
- 5- Sobre el costo y la eficiencia: ¿Cuál es el costo del programa? ¿Es costo-efectivo? ¿Es sostenible?

No necesariamente en un solo proceso evaluativo se debe pretender contestar cabalmente a estas cinco dimensiones de interrogantes, más bien dependiendo del grado de desarrollo que presenten

los programas y de los objetivos de la evaluación, el proceso puede concentrarse en responder específicamente solo a algunas de estas interrogantes.

Debe siempre tenerse presente que la evaluación debería orientar a los tomadores de decisiones para establecer las posibles razones que explican el comportamiento del programa en cada una de esas dimensiones de interrogantes y con esto ayudar a determinar líneas de acción que mejoren su efectividad. Sin evaluación es imposible saber si los recursos están siendo aprovechados apropiadamente, si las acciones emprendidas se están realizando según lo planeado y si esas acciones están provocando los cambios deseados en las dimensiones de calidad de vida que constituyen las metas de impacto finales para cualquier proyecto o programa de salud.

Entonces a partir de la discusión anterior es claro que la evaluación es una herramienta poderosa, pues informa la toma de decisiones, ayuda a identificar y corregir errores en la implementación de los programas y permite hacer las intervenciones más eficientes y efectivas. No debe nunca ser usada como un instrumento de castigo, ni tampoco como una excusa para realizar auditorías o contralorías. El enfoque que inspira la evaluación, a diferencia de una auditoría, se basa en recolectar información útil que permita juzgar el valor o mérito del programa, para así orientar la toma de decisiones, pero no bajo la óptica de revisar, confrontar, recoger “pruebas”, o identificar responsables que sí se aplica en el caso de la auditoría.

En general, se pueden distinguir dos tipos de evaluación: formativa y sumativa. Formativa es aquella que se realiza regularmente para guiar y posiblemente corregir el desempeño de un programa. Su objetivo fundamental es el mejoramiento. La sumativa, por el contrario pretende establecer en qué grado el programa, una vez implementado, ha generado en la población meta los cambios esperados en las variables de resultado o de impacto.

### **¿QUIÉNES DEBEN HACER EVALUACIONES?**

Una de las discusiones que ha entretenido por años a los estudiosos en evaluación ha sido el tema de quienes son los más indicados para llevar a cabo este tipo de actividad. De esta forma, tradicionalmente se ha distinguido entre evaluación interna y evaluación externa. En el caso de la evaluación interna son los mismos ejecutores y responsables de los programas (y a veces hasta los mismos beneficiarios) los que llevan a cabo la evaluación. Por esta razón también recibe el nombre de auto-evaluación. Por el contrario, en la evaluación externa o independiente es un profesional o grupo de profesionales traído expresamente para ese efecto el que está a cargo del diseño y desarrollo del proceso evaluativo.

Comparando las bondades y limitaciones de ambos tipos de evaluación, interna y externa, se puede decir que las debilidades de una son las fortalezas de la otra. Por ejemplo, una de las ventajas de la evaluación interna es, que al ser los mismos evaluadores miembros del equipo de ejecución, tienen un conocimiento profundo del funcionamiento del programa y del contexto que lo rodea, sin embargo precisamente por su condición de ejecutores pueden también ser más proclives a incorporar percepciones subjetivas demasiado optimistas en su interpretación evaluativa y sobreestimar las bondades del programa. En realidad parece que la opción más indicada es una adecuada combinación de ambos enfoques, el interno y el externo (Worthen & Sanders, 1987).

Independientemente de cual sea la mejor combinación entre ambos tipos de evaluación externa e interna, la experiencia y la práctica ha enseñado a los evaluadores que, en la medida en que los ejecutores y responsables de los programas se sientan involucrados con el proceso evaluativo y

más aún, en la medida en que todas las audiencias interesadas se puedan “apropiar” de ese proceso, en esa medida habrá más posibilidades de que la evaluación se use efectivamente para iluminar la toma de decisiones. (Cook & Reichardt, 1995).

## **TENDENCIAS Y TEMÁTICAS DE ACTUALIDAD EN LA EVALUACIÓN DE PROGRAMAS Y SISTEMAS DE SALUD**

La incorporación explícita de procesos de evaluación en los sistemas y programas de salud ha llegado algo tardíamente, si los comparamos con otros tipos de programas de desarrollo. Ya desde los años 50 se realizaban evaluaciones en los Estados Unidos para buscar evidencias de efectividad en programas de capacitación laboral. Muchos programas educativos fueron pioneros en cuanto a ser valorados con modelos de evaluación. En el campo de la salud pública desde hace varias décadas se realizan evaluaciones de intervenciones específicas, especialmente aquellas que involucran un fuerte componente psicosocial, tales como proyectos para evitar el abuso en el consumo de ciertas sustancias y drogas (Rossi et al, 1999).

Por otro lado, la creación, después de la Segunda Guerra Mundial, de organizaciones internacionales de asistencia para el desarrollo de las naciones menos industrializadas llevó consigo la incorporación de procesos evaluativos para muchas de esas intervenciones. En el campo de la salud se pueden mencionar especialmente las evaluaciones de los programas nacionales de planificación familiar y salud reproductiva (Bertrand et al, 1996). Sin embargo, el interés por evaluar de manera global todo el sistema de salud pública de una nación es relativamente reciente. (Aday et al, 1998)

Una indicación de este relativo atraso en el reconocimiento de la evaluación como disciplina entre los profesionales de la salud es el hecho de que según Rossi et al (1999) en 1993 menos de un 3% de los miembros de la AEA (American Evaluation Association) declararon salud como su disciplina principal, siendo las frecuencias mayores para educación (22%) y psicología (18%).

No se trata, sin embargo, de que la evaluación haya sido completamente ignorada en los sistemas de salud. Hay muy importantes elementos y actividades evaluativas que llevan años desarrollándose y aplicándose, aún en países como Costa Rica. Desde esta perspectiva, se pueden mencionar los esfuerzos realizados en cuanto a medición de la calidad de la atención en los servicios de salud, los estudios de satisfacción de los usuarios, los análisis de situación, las auditorías médicas y de establecimientos, los estudios epidemiológicos y la investigación clínica. Todas estas actividades representan una clara evidencia de que los procesos evaluativos han estado presentes en los sistemas de salud. No obstante, estos esfuerzos se han realizado en su mayoría sin utilizar un enfoque sistémico y no corresponden con una concepción integral de la evaluación.

Otro dato interesante es que uno de los pocos textos que trata específicamente el tema de evaluación de sistemas de salud, “Evaluating the Healthcare System: Effectiveness, Efficiency and Equity” de Aday et al (1998) no presenta una definición formal del concepto de evaluación, ni discute aspectos asociados al marco conceptual de esta disciplina, aunque sí pone especial énfasis en la definición y discusión en torno a los conceptos de efectividad, eficiencia y equidad.

Actualmente, sin embargo, parece que se está generando un genuino interés por la evaluación en el campo de la salud. Ya existe en la American Evaluation Society un grupo de interés en el tema de evaluación de salud (health evaluation) y la prestigiosa revista “Evaluation: The International Journal of Theory, Research, and Practice” dedicó casi completamente su número 4 (octubre del

2001) a evaluaciones relacionadas con programas de salud. Por otra parte, existe una revista de evaluación especializada para los profesionales de la salud: “Evaluation and the Health Professions”.

Una búsqueda bibliográfica en el portal de internet de la Organización Panamericana de la Salud probó ser infructuosa al tratar de localizar documentos bajo los descriptores “evaluación”, “evaluación de programas” y “evaluación de sistemas de salud”, tanto en inglés como en español. Esto no significa, como se apuntó antes, que la evaluación esté completamente ausente de los estudios y documentos que genera esta institución, estamos seguros que no es así, pero sí llama la atención y es significativo que se use tan poco el término “evaluación de programas”.

Por su parte, en el portal de la Organización Mundial de la Salud se localizaron 571 publicaciones usando el descriptor “program evaluation” pero casi en su totalidad se referían a evaluaciones de intervenciones específicas y no de sistemas de salud.

Uno de los documentos localizados en esta búsqueda se refiere a la V Conferencia Global sobre Promoción de la Salud que tuvo lugar en la ciudad de México del 5 al 9 de junio del año 2000. En ésta se creó un grupo de trabajo Ad Hoc sobre evaluación que, entre otras llegó a las siguientes conclusiones:

1. Se debe distinguir los diferentes propósitos de la evaluación y los tipos de evidencia que se recolecta para cada una. Uno de los propósitos es el beneficio del programa y sus audiencias involucradas (stakeholders). La evidencia recolectada se usa para guiar el programa y mejorarlo. Un segundo propósito es probar a los donantes, agencias financiadoras y hacedores de política que el programa tiene un impacto, un efecto, un valor.
2. Todavía falta por oír muchas voces en el debate sobre lo que constituye evidencia, especialmente se deben incluir a representantes de las naciones en desarrollo y a la gente y los grupos que quedan automáticamente excluidos al no mantener cargos públicos que puedan traerlos a la mesa de discusión.
3. Cualquiera que sea el contexto para la evaluación de la promoción de la salud, tales evaluaciones deben ser conducidas en colaboración con los “stakeholders” (audiencias interesadas), incluyendo las comunidades en las cuales los programas se están ejecutando. Esto requiere que las medidas e indicadores de evaluación seleccionados tengan significado para esos diversos grupos y comunidades.

Por otra parte, una de las principales recomendaciones de este grupo Ad Hoc apuntó a que la OMS debería establecer un grupo de trabajo permanente que sería el responsable de crear un plan para el desarrollo de la evaluación globalmente. También se recomendó que los modelos de evaluación reconozcan la importancia de la equidad al conducir evaluaciones locales y enfatizen el uso de enfoques participativos y el involucramiento multisectorial.

En lo que toca a esfuerzos internacionales que se están realizando actualmente en los países en desarrollo para incorporar la evaluación y el monitoreo como parte integral de la ejecución de los programas, se puede mencionar al Programa MEASURE *Evaluation*. Es financiado por USAID y ejecutado por el CPC (Carolina Population Center) de la Universidad de Carolina del Norte en Chapel Hill. Tiene como objetivos principales generar enfoques coordinados, costo-efectivos y eficientes en la recolección de datos, el monitoreo y la evaluación de servicios en las áreas de población, salud y nutrición a escala mundial. En Costa Rica el Centro Centroamericano de Población de la Universidad de Costa Rica, conjuntamente con MEASURE *Evaluation*, ha organizado durante cuatro años consecutivos un taller de tres semanas de duración, dirigido al

área de América Latina, sobre métodos avanzados para la evaluación de impacto de programas de población, salud y nutrición.

Por su parte, el Banco Mundial cuenta actualmente con un grupo de trabajo en evaluación que pertenece a su Departamento de Evaluación de Operaciones. Este grupo está desarrollando iniciativas de evaluación de programas desde un marco de referencia comprensivo del desarrollo (Comprehensive Development Framework), idea que fue articulada por el presidente del Banco Mundial en enero de 1999. Este marco de referencia, según sus creadores, contiene un conjunto de principios que un número de naciones y agencias de asistencia para el desarrollo han tratado de poner en práctica concienzudamente durante los últimos años. Las estrategias se concentran en los programas de reducción de la pobreza y se pone especial énfasis en la necesidad de incorporar procesos participativos, dentro de un marco de desarrollo más amplio que el tradicionalmente usado por las instituciones financieras. Los criterios que se proponen para la evaluación de los programas y proyectos financiados por el Banco Mundial son: pertinencia, eficiencia, eficacia, impacto y sostenibilidad. Este esfuerzo ha producido ya importantes resultados que se reportan en un conjunto de publicaciones e informes que, entre otras cosas, describen experiencias piloto en países en donde se ha tratado de implementar el enfoque. Las actividades ejecutadas también incluyen un importante componente de capacitación por medio de talleres y cursos para sus funcionarios.

Finalmente, aunque no se trata de un esfuerzo internacional, es necesario mencionar la gran importancia que ha adquirido la evaluación para el “Center for Disease Control” (CDC) de los Estados Unidos. El CDC también cuenta con un grupo de trabajo para, en sus palabras, “promover la evaluación de programas en salud pública”. Como parte de su página en internet se incluyen dos sitios sobre evaluación. Uno de ellos presenta proyectos e investigaciones en este campo en que el CDC participa o ha participado. El otro sitio web se titula “recursos de evaluación” (evaluation resources) y muestra diferentes publicaciones, asociaciones profesionales y centros de investigación dedicados al tema de la evaluación. Una búsqueda en línea bajo “program evaluation” en el portal del CDC arrojó los títulos de 686 documentos. Sin embargo, al hacer la búsqueda bajo “health systems evaluation” o “evaluation of health systems” solo se encontró un documento.

En relación con temas y asuntos de actualidad que se discuten en el campo de la evaluación en general y específicamente en programas de salud se pueden mencionar los siguientes:

- a. **Discusiones en torno al concepto de valoración de necesidades:** En el enfoque dominante de la evaluación como disciplina científica el concepto de valoración de necesidades (needs assessment) es relativamente nuevo, pues los acercamientos más antiguos establecían que la evaluación consistía en comparar los resultados del programa con los objetivos planteados inicialmente. Sin embargo, más recientemente se ha planteado también la importancia de evaluar en sí mismos los objetivos de los programas en cuanto a su pertinencia para satisfacer las necesidades de la población meta (Scriven, 1991). En el caso de los programas de salud los llamados análisis de situación pueden servir como punto de partida para esta valoración de necesidades, aunque sería necesario establecer si es suficiente con la información suministrada por ellos para establecer si efectivamente se ha hecho una adecuada valoración de necesidades desde el punto de vista de la población meta.
- b. **Teoría del Programa (program theory):** Este es otro concepto fundamental que se ha incorporado recientemente de manera explícita a la teoría y práctica de la evaluación, aunque según Rossi et al (1999) ya desde los años 70 varios autores líderes en el campo reconocieron y discutieron la importancia de nociones similares. De esta forma, se han

utilizado diferentes nombres para identificar este concepto y se han dado diferentes descripciones para definirlo. Así, se habla de “model logic” (modelo lógico o lógica del modelo), “program logic” (lógica del programa), “causal logic” (lógica causal), línea de resultados, mapa causal y también de “theory driven evaluation” (evaluación dirigida por la teoría) (H.T. Chen, 1990). Michael Scriven es un autor que ha dedicado gran parte de su obra a discutir y elaborar sobre la lógica causal en los programas y en cómo interpretar la evidencia a la luz de esta lógica.

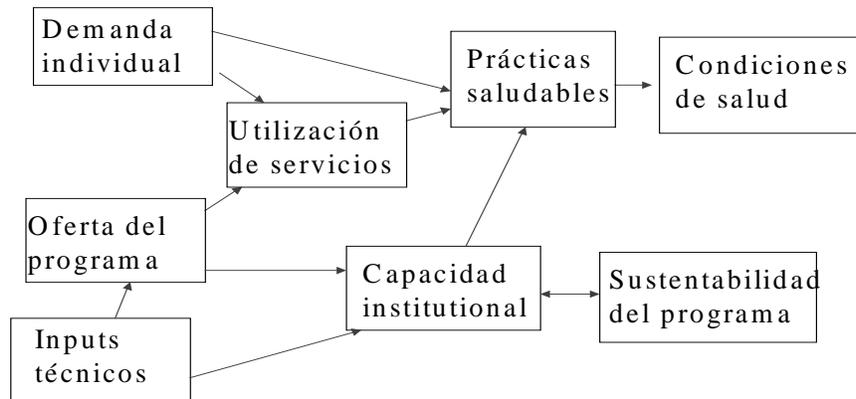
En general, podemos decir que la teoría del programa explica de qué forma las acciones del programa producirán los resultados deseados en la población meta. Rossi et al. (1999) la definen más específicamente como el conjunto de supuestos sobre la forma en que el programa (o proyecto) provoca los beneficios que se espera que produzca, incluyendo las estrategias y tácticas que ha adoptado para lograr sus metas y objetivos.

Muchas veces la teoría del programa no ha sido articulada de antemano por los diseñadores o ejecutores del programa, no ha sido formalmente establecida. Entonces es tarea de los evaluadores hacerla explícita. Precisamente se llama “evaluación de caja negra” a aquella evaluación que se realiza sin haber articulado formalmente la teoría del programa. Sin embargo, la recomendación más reciente de la mayoría de los miembros de la comunidad académica en evaluación apunta a que, en general, en toda evaluación se deben por lo menos esbozar los elementos más relevantes de la teoría del programa (Rossi et al, 1999; Mohr, 1995). En los enfoques más tradicionales, cuando se concebía la evaluación como un simple contraste entre objetivos y resultados, es cierto que no hacía falta la teoría del programa. No obstante, la disciplina evaluativa ha evolucionado para plantear actualmente la necesidad de evaluar también la pertinencia de los mismos objetivos del programa y para ello la teoría del programa es una herramienta esencial.

Por lo expuesto anteriormente es que hemos creído conveniente elaborar un poco más sobre este concepto. Distinguímos entonces dos elementos básicos en la teoría del programa:

- ❑ La teoría del proceso, describe los planes: organizativo y el de utilización de servicios del programa (aquí se hace monitoreo y evaluación de proceso).
- ❑ La teoría del impacto, relata la naturaleza del cambio en las condiciones de vida provocadas por la acción del programa (aquí se hace evaluación de resultados y de impacto).

En el siguiente esquema se presentan de manera muy sucinta los elementos fundamentales para la creación de la teoría del programa en el caso de programas de salud.

**Figura 1. Esquema para elaborar la Teoría del Programa en intervenciones de salud**

Fuente: Programa MEASURE *Evaluation*. Carolina Population Center. Universidad de Carolina del Norte en Chapel Hill.

- c. **Evaluación de impacto:** Es este un concepto que se ha “puesto de moda” entre los consumidores y clientes de evaluaciones, aunque de manera poco precisa. Aún cuando este término se ha vuelto casi un lugar común, es empleado muchas veces con poca rigurosidad. Valga decir aquí que realizar una evaluación de impacto implica establecer evidencia de una relación causal entre el programa y los resultados de salud que se observan en la población. Como es bien sabido por los investigadores en cualquier campo, el establecimiento de relaciones causales es uno de los retos más grandes en la investigación científica. Entonces al realizar una evaluación de impacto es indispensable cumplir con los necesarios criterios de rigor científico. Para ello se requiere diseñar estudios que permitan controlar el efecto de las “terceras variables” o “variables confusoras”, que son aquellos otros factores que pueden afectar los resultados de salud y que no tienen que ver con las acciones del programa.

El impacto se refiere a resultados del programa que pueden ser asociados a dimensiones de calidad de vida o desarrollo humano. Los cambios esperados pueden ser cualitativos o cuantitativos pero siempre deben ser medibles.

Así, la evaluación de impacto puede definirse como un conjunto de actividades de investigación tendientes a establecer si el comportamiento observado en las variables de interés (variables de impacto) puede ser atribuido al programa. La meta es entonces establecer evidencias de una relación causal entre los resultados observados y las actividades del programa de salud.

Podemos decir que la evaluación de impacto en salud está incluida dentro de un enfoque más general que se denomina investigación de resultados (outcomes research) y medición de resultados (outcomes measurement).

Una vez que se han definido los indicadores apropiados el problema mayor que se enfrenta en una evaluación de impacto es precisamente determinar el grado en que los resultados observados pueden ser atribuidos al programa. Esto porque normalmente los estados de salud que se miden pueden estar determinados o influidos por variables o factores exógenos al programa.

A raíz de lo anterior, para realizar evaluaciones de impacto se requiere contar con diseños metodológicos que permitan aproximarse a establecer evidencias del grado de relación causal entre los resultados y el programa. Tradicionalmente los experimentos y los cuasi-experimentos eran los diseños recomendados en estas situaciones. Sin embargo, debido a las limitaciones prácticas y éticas en su aplicación con seres humanos, más recientemente han aparecido alternativas metodológicas que permiten por medio de modelos estadísticos estimar esa relación causal en contextos de investigación más “naturales”, tales como una encuesta (Angeles & Mroz, 2001).

- d. **Costo-efectividad:** Los métodos tradicionales, costo-beneficio, para la valoración económica de los programas de desarrollo, incluyendo salud y educación, han sido criticados, pues muchas veces es difícil o imposible asignar valores monetarios a resultados que tienen que ver con condiciones de salud y calidad de vida. Los métodos de costo-efectividad miden los resultados en términos no monetarios pero tienen la desventaja que solo se pueden usar en programas que tengan los mismos objetivos en cuanto a resultados de salud, por ejemplo disminuir la tasa de mortalidad infantil. En todo caso la valoración económica de los programas y sistemas de salud sigue siendo tema de intenso debate entre los economistas de la salud.
- e. **Uso complementario de métodos cualitativos y cuantitativos:** En el paradigma dominante (mainstream) de la evaluación de programas parece haberse llegado actualmente a un consenso en cuanto a la necesidad de utilizar conjuntamente los métodos y técnicas cuantitativas y cualitativas. Las evaluaciones suelen ser más comprensivas y más útiles cuando se utiliza un enfoque complementario. Este enfoque ecléctico ha ganado más terreno quizá porque, tal y como lo indican Worthen y Sanders (1987), el evaluador a diferencia del investigador, no puede darse el lujo de escoger un método de investigación sobre otro. Se trata, evidentemente, de conocer las bondades y limitaciones de cada uno y usar aquel o aquellos que más se ajusten al tipo de pregunta específica que se quiera contestar. Los métodos cuantitativos, por ejemplo, ofrecen herramientas para generalizar las conclusiones hacia poblaciones mayores y para medir productos y resultados. Los métodos cualitativos, por su parte, dado su acercamiento más intimista y profundo, son más indicados para explicar y documentar los procesos y para elaborar hipótesis acerca de por qué funcionan o no funcionan los programas.
- f. **Sistemas de información:** Un aspecto que parece indispensable para hacer de la evaluación una adecuada herramienta en la toma de decisiones es la incorporación de sistemas de información en lo que toca al diseño, organización, acceso y manipulación de datos, incluyendo la creación de bases de datos relacionales, acceso en línea o vía redes y sistemas de información geográficos. Parece claro que la incorporación de estas nuevas tecnologías ayudará a lograr que la evaluación se vuelva un componente regular de la ejecución de los programas y que más audiencias tengan acceso a información generada con propósitos evaluativos. Especialmente la comunicación en línea permite que todo el proceso de recolección y análisis se pueda realizar de manera ágil y oportuna.

## **UNA PRIMERA APROXIMACIÓN A DOS CONSTRUCTOS DE INTERÉS EN LA EVALUACIÓN DE PROGRAMAS DE SALUD: CALIDAD Y EQUIDAD**

### **Calidad**

Según Carey & Lloyd (2001) virtualmente se han escrito toneladas de textos alrededor del concepto de calidad. Aunque en el campo de la salud la discusión y producción ha sido muy fértil, según estos autores los profesionales de la salud entraron tardíamente a esta arena y no es hasta los años 90 que se comienza a generar una “explosión” de reportes, discusiones y estudios dirigidos específicamente al tema de calidad de la atención. Ellos afirman que en el campo de la salud, históricamente, el concepto de calidad había sido definido por los proveedores de los servicios. Sin embargo, a partir de los años 80 y sobre todo en los 90 ha habido por parte de los clientes o beneficiarios un movimiento para establecer ellos mismos su definición de calidad.

Algunos de los primeros debates con relación al término “calidad de la atención” se dieron en programas de planificación familiar y salud reproductiva. Ciertas audiencias involucradas en la ejecución y evaluación de estos programas realizaron un trabajo pionero en cuanto a establecer la necesidad de incluir la voz de la paciente en el proceso de valoración de la calidad en la prestación de servicios. La teoría de género y los enfoques participativos han producido un marco de referencia que cuestiona la relación tradicional médico-paciente y que tiene consecuencias en la definición y componentes de la dimensión calidad de la atención (Bruce, 1990).

Al respecto se puede mencionar específicamente que en Costa Rica el Hospital de las Mujeres está implementando de manera piloto un programa de atención a la mujer con enfoque de género y participativo. Esta experiencia implica cambiar paradigmas dominantes y modificar actitudes y valores tanto en los prestatarios como en las clientes o beneficiarias de los servicios. El enfoque cuestiona tanto los marcos conceptuales como la práctica y procedimientos que se siguen en la atención. Este modelo ciertamente promete una aportación importante en términos de su posible generalización a todo el sistema de salud.

Para dar tan solo un ejemplo de como la perspectiva de género puede ayudar en una mejora de la calidad de los servicios de atención en todo el sistema de salud basta mencionar el caso ampliamente conocido de las mujeres que asisten a consulta externa en los hospitales y clínicas de la Caja Costarricense del Seguro Social, por razones que podríamos llamar “psicológicas”. Muchos profesionales de la salud piensan que estas mujeres asisten a la consulta no por razones estrictamente médicas sino de apoyo emocional o para tener la oportunidad de ser atendidas y escuchadas. Por años se ha reportado esta tendencia y algunos la han etiquetado como un problema, argumentando que este grupo está “sobre-utilizando” los servicios que otros sectores de población necesitan con mayor urgencia y arriesgando la capacidad de respuesta del sistema, ya de por sí muy maltrecha. Pero desde una perspectiva más amplia del concepto de salud y utilizando también la teoría de género, una situación como ésta debería llevarnos a explorar la naturaleza y las causas de este fenómeno, en otras palabras hacer una valoración de necesidades para poder generar el tratamiento apropiado. Este acercamiento también es consistente con la perspectiva poblacional en la conceptualización y prestación de los servicios de salud. Así, se podrían diseñar intervenciones de calidad que se dirijan a aliviar este problema.

Volviendo ahora más específicamente a la definición del constructo, según Aday et al. (1998) “calidad es un atributo del proceso de salud que tiene que ver con estar haciendo lo apropiado, lo correcto y estarlo haciendo bien”. La Agency for Healthcare Research and Quality de los Estados Unidos (2001) presenta una definición muy similar de calidad de la atención pues afirma que

significa hacer lo correcto, en el momento correcto, en la forma correcta, por el personal correcto y teniendo los mejores resultados posibles. Según esta agencia hay dos tipos básicos de medidas de calidad: las que incluyen la valoración de los consumidores (clientes o pacientes) y las que involucran indicadores del desempeño clínico, también llamados medidas de calidad técnica. Ambos tipos de medidas están basadas en lo que se denomina “outcomes research” (investigación de resultados).

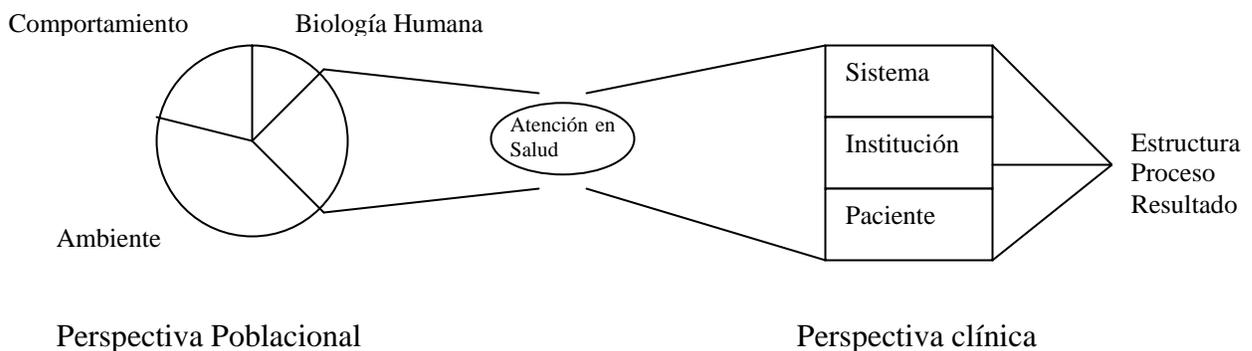
Desde nuestra perspectiva, las medidas de resultado deben incluir indicadores de satisfacción y calidad de vida en sus diferentes aspectos, es decir no solo deben considerarse los indicadores clásicos de mortalidad y morbilidad. Para el desempeño clínico se pueden mencionar las dimensiones de diagnóstico, tratamiento, seguimiento y medicación apropiados, así como la capacidad de prevenir antes que curar.

Refiriéndonos ahora al grado de propiedad en la prestación de los servicios, es relevante mencionar que según un reporte del Departamento de Salud y Servicios Humanos de los Estados Unidos (1998) muchas veces se lastima a los pacientes durante el tratamiento y existen porcentajes no tolerables de errores en los exámenes de laboratorio. Por ejemplo, alrededor de 30% de las pruebas de Papanicolau fueron clasificadas incorrectamente como normales. En nuestro país también resulta de extrema importancia la realización de este tipo de investigaciones con propósitos evaluativos. En el caso de las inmunizaciones, para poner solo una ilustración, se sabe que pueden ser totalmente inefectivas si no se cumple con las normas de temperatura de almacenaje.

Para Aday et al. (1998) es importante considerar tanto la perspectiva clínica como la poblacional al definir y determinar las dimensiones de calidad que se evaluarán en un sistema de salud. La perspectiva poblacional argumenta que se requieren tanto inversiones médicas como no médicas para el mejoramiento de la salud de individuos y comunidades, mientras que la perspectiva clínica se orienta a optimizar la precisión de la atención. Un esquema que ilustra ambas perspectivas se presenta a continuación:

**Figura 2. Las perspectivas clínica y poblacional en el sistema de salud**

**Fuente: Aday et al, 1998.**



Estos mismos autores (op.cit) establecen con base en este marco de referencia que las siguientes dimensiones deberían incluirse en la valoración de la calidad del sistema de salud.

**En Estructura:**

- 1- Políticas de salud (nacional y local)
- 2- El sistema de atención y prestación de servicios (disponibilidad, organización y financiamiento)
- 3- Poblaciones en riesgo (factores predisponentes, necesidad)
- 4- Ambiente físico, social y económico de las comunidades

**En Proceso:**

- 1- Acceso real (utilización y satisfacción)
- 2- Los riesgos de salud (del ambiente y de comportamiento)

**En resultados intermedios:**

- 1- Eficiencia
- 2- Efectividad clínica y poblacional
- 3- Equidad

**En resultados de impacto (finales):**

Salud de los individuos y de las comunidades

Así, desde esta perspectiva los indicadores de calidad deben referirse a ambas dimensiones, la clínica y la poblacional.

Por su parte, el programa nacional de calidad Baldrige de los Estados Unidos (2002) ha determinado criterios de calidad del desempeño de la atención en salud. Establecen que los indicadores de calidad deben referirse a las siguientes dimensiones, entre otras:

1. Liderazgo visionario
2. Excelencia enfocada en el paciente
3. Aprendizaje continuo en el ámbito organizacional y personal
4. Valoración positiva del personal y colaboradores
5. Agilidad (la capacidad para el cambio rápido y flexibilidad)
6. Gerencia para la innovación
7. Responsabilidad pública y salud comunitaria (entendiendo esto como las expectativas básicas del sistema de salud en relación con las prácticas éticas y la protección de la salud pública, seguridad y el ambiente)
8. Énfasis en resultados y creación de valor
9. Uso de un enfoque de sistemas

Birch et al (2000), siguiendo el marco de referencia para el estudio de la calidad de los servicios de salud establecido por Donabedian (1980), establecen que deben considerarse tres grandes dimensiones de calidad: la calidad médica o profesional que es vista como la atención óptima en salud dados los conocimientos biomédicos actuales, la calidad desde el punto de vista del cliente/usuario en la que se valora tanto la forma en que el servicio es brindado como los resultados de esa atención y, la calidad gerencial que incorpora factores como el uso más apropiado de los recursos tanto humanos como no humanos.

Hablando específicamente de los factores que deben evaluarse en el proceso de atención, estos autores citan a Maxwell (1984) para establecer los siguientes aspectos:

**Efectividad:** ¿Es el servicio o tratamiento que se brinda el mejor disponible en el sentido técnico?  
¿Cuál es esa evidencia?

Aceptabilidad: ¿Cuán humanamente y consideradamente es ofrecido ese tratamiento? ¿Qué piensan los pacientes de él? ¿Qué pensaría un observador calificado? ¿Cuál es el contexto en que se brinda el servicio? ¿Se resguarda la confidencialidad y la privacidad?

Eficiencia: ¿Se maximiza el producto para un insumo dado o de manera inversa, se minimiza el insumo para un nivel dado de producto?

Acceso: ¿Puede la gente obtener este tratamiento o servicio cuando lo necesita? ¿Hay barreras identificables en los servicios (distancia, falta de capacidad de pago, listas de espera, tiempos de espera)?, o de plano, ¿hay una escasez de la oferta de servicios?

Equidad: ¿Es este paciente o grupo de pacientes tratado de manera justa en relación con otros? ¿Hay fallas identificables en equidad (por grupo socioeconómico, grupo étnico, sexo)?

Debe aclararse aquí que en este enfoque de Maxwell “equidad” se ve como con componente o factor del constructo mayor “calidad”. En otras aproximaciones, sin embargo, calidad y equidad se analizan como constructos diferentes. En este trabajo se ha tomado este último camino y es por eso que en la sección siguiente se discute el constructo equidad.

Pertinencia: ¿Es el patrón general de la atención y el balance de los servicios el mejor que pudo haber sido logrado, tomando en cuenta las necesidades y los deseos de la población como un todo?

Por último, no queremos finalizar este apartado sin mencionar la necesidad de que cualquier evaluación de la calidad del sistema de salud considere el concepto de promoción de la salud y diseñe indicadores para su estudio y medición. El concepto de promoción de la salud se gesta como parte de lo que Restrepo y Málaga (2001) llaman “un cambio trascendental en el concepto de salud”. Según estos autores hay una diferencia entre la promoción de la salud y la prevención de la enfermedad, siendo la primera dirigida principalmente a la población y la segunda al individuo. De esta forma podemos afirmar que el enfoque de promoción de la salud es acorde con la perspectiva poblacional y el enfoque de prevención de la enfermedad sigue más bien el paradigma clínico o biomédico tradicional. Es pertinente mencionar aquí que el sistema de salud costarricense, durante los años 70, ejecutó en algunas zonas geográficas del país un programa modelo para la promoción de la salud que se llamó “hospital sin paredes”. Esta fue una experiencia aparentemente muy exitosa que puede considerarse como pionera a nivel mundial.

## **Equidad**

El concepto de equidad, al igual que el de evaluación, despierta controversia. Aday et al (1998) afirman que la equidad trata sobre las disparidades en salud y la justicia y efectividad de los procedimientos para enfrentarlos. Según Bambas y Casas (2001) equidad no es igualdad en la distribución, sino más bien justicia en la distribución.

Sin embargo, para Aday et al (1998) la última prueba de la equidad en las políticas de salud se da en la medida en la cual las disparidades o desigualdades en salud persistan entre diferentes subgrupos de la población. Se refleja equidad sustantiva cuando se minimizan las disparidades de salud entre los subgrupos. Los estratos de población que se comparan en este análisis de equidad involucran usualmente variables como nivel de educación, nivel de ingreso, sexo, condición geográfica y grupo étnico. Según estos autores una variedad de enfoques en los Estados Unidos se centran en barreras reales o potenciales para el acceso. Otro tipo de comparaciones, por ejemplo,

por grupos ocupacionales o aún por orientación sexual del cliente son menos usuales, pero no por ello menos relevantes.

Parece existir, entre los estudiosos del concepto de equidad un debate paradigmático que tiene que ver con diferentes concepciones de justicia. Este debate involucra, en un extremo, nociones como justicia distributiva y el marco de los derechos individuales, en el otro extremo se ubican conceptos como la justicia social y el bienestar público. Desde nuestro punto de vista un marco de referencia más comprensivo debe incorporar elementos de ambos paradigmas.

También hay una relación entre estos paradigmas para definir la equidad y la perspectiva poblacional y la clínica en los sistemas de atención en salud, siendo el paradigma de la justicia distributiva más acorde con la perspectiva clínica y el paradigma de la justicia social más asociado a la perspectiva poblacional.

Se puede entonces decir que el paradigma de la justicia distributiva se preocupa más por la igualdad en las condiciones de disponibilidad, acceso y tratamiento en la atención, mientras que el paradigma de la justicia social va más allá e identifica falta de equidad cuando los resultados o estados de salud son diferentes en distintos grupos de población.

Si usamos una perspectiva integrada de los enfoques poblacional y clínico parece claro que las medidas de equidad deben referirse no solo al acceso y prestación de servicios sino a las condiciones del ambiente y comportamiento de los individuos que hacen que ciertos grupos presenten mayores riesgos de salud que otros.

Bambas y Casas (2001), por su parte, distinguen dos tipos de equidad. La equidad horizontal se refiere a la distribución de recursos iguales o equivalentes para personas o grupos con las mismas necesidades. La equidad vertical examina si se da la mejor distribución de los recursos de acuerdo con los diferentes niveles de necesidad en la población. Parece claro, a partir de estas definiciones, que la medición de la equidad horizontal es más sencilla que la medición de la equidad vertical. Estos autores también afirman que para establecer que en una situación se presenta falta de equidad, las diferencias en la distribución de los recursos de salud deben satisfacer los siguientes tres criterios:

- las diferencias en la distribución deben ser evitables
- no deben reflejar una opción voluntaria
- el reclamo debe ligar la distribución con un agente responsable

De nuevo los enfoques poblacionales y la perspectiva de género nos ayudan a iluminar la discusión en términos de equidad. Los estudiosos de la teoría de género han argumentado, de manera muy válida, que equidad no es lo mismo que igual tratamiento y que la equidad muchas veces no puede medirse únicamente por medio de la ausencia de diferencias. Más aún, muchos argumentan que equidad implica tratamiento diferenciado. Favorecen entonces el concepto de equidad vertical. Un ejemplo muy sencillo de este concepto se puede establecer cuando pensamos que por razones netamente fisiológicas las mujeres van más frecuentemente al baño que los hombres. Sin embargo, en la mayoría de los sitios públicos y oficinas el número de baños para mujeres y hombres es el mismo. Esta igualdad representa entonces falta de equidad.

Finalmente Aday et al (1998), mencionan entre otras, las siguientes dimensiones como punto de partida para generar indicadores de equidad:

Participación: La extensión en la cual aquellos que son afectados participan en la toma de decisiones. Una forma en que se puede expresar la participación es la naturaleza y calidad de la comunicación entre los prestadores del servicio y los clientes o beneficiarios. También los grados de involucramiento o representación comunal.

Libertad para escoger: Enfatiza la importancia de la autonomía personal y tiene que ver con la distribución y disponibilidad de los recursos del sistema de salud a los beneficiarios.

Tratamiento similar versus tratamiento diferencial de individuos en diferentes contextos: Aquí se usa la comunidad como unidad de análisis. Lo que buscan las intervenciones es alterar las raíces de los problemas de salud atacando los correlatos socio-estructurales de las desigualdades en salud y atención y que están inmersos en los ambientes físicos y económicos en que los individuos viven y trabajan.

Necesidad: Se debe asegurar que hay una igualdad justa en las oportunidades para vivir una vida normal y proveer los servicios de acuerdo con las necesidades de los individuos y comunidades. Aquí los indicadores de equidad intentan valorar la magnitud de los riesgos y las disparidades en salud de una población. Indicadores de mortalidad, morbilidad, años de vida perdidos o años de calidad de vida ganados ayudan a establecer si las necesidades de los diversos subgrupos de población han sido satisfechas.

## CONCLUSIÓN

La intención de este artículo ha sido, fundamentalmente, presentar elementos clave de un marco de referencia para realizar evaluación de programas y sistemas de salud. Se discuten conceptos derivados de la evaluación como disciplina científica (como valoración de necesidades, teoría del programa) y se argumenta sobre la necesidad de conocer y aplicar estos conceptos cuando estudiamos el valor o mérito de intervenciones en salud.

El paradigma que inspira esta propuesta establece que la evaluación es una herramienta esencial para la toma de decisiones y para el mejoramiento de los programas y proyectos. Además, quienes hacen evaluaciones estarán mejor capacitados para llevarlas a buen término en la medida en que estén familiarizados con este marco de referencia y hayan recibido formación académica en evaluación.

## REFERENCIAS

- Aday, L.A. et al. (1998). Evaluating the Healthcare System: Effectiveness, Efficiency and Equity. Chicago, Illinois: Health Administration Press.
- Agency for Healthcare Research and Quality, Estados Unidos (2001). A Quick Look at Quality. Your Guide to Choosing Quality Health Care. AHCPR Publication No. 99-0012. Rockville, MD. <http://www.ahrq.gov/consumer/qntascii/qntqlook.htm>
- Angeles, G. & Mroz T.A. (2001). A Simple Guide to Using Multilevel Models for the Evaluation of Program Impacts. Universidad de Carolina del Norte en Chapel Hill, The Carolina Population Center: *MEASURE-Evaluation*.

- Baldrige National Quality Program, Estados Unidos (2002). Health Care Criteria for Performance Excellence. Documento disponible en internet en [www.quality.nist.gov](http://www.quality.nist.gov)
- Bambas, A. y Casas, J.A. (2001). Assessing equity in health: Conceptual criteria. En *Equity and Health: Views from the Pan American Sanitary Bureau*. Documento de internet disponible en <http://www.paho.org/English/DBI/OP08.htm>
- Bertrand, J. T. et al. (1996). Evaluación de Programas de Planificación Familiar con Adaptaciones para Salud Reproductiva. Universidad de Carolina del Norte en Chapel Hill, The Carolina Population Center: MEASURE-Evaluation.
- Birch, K.; Field, S. & Scrivens, E. (2000). Quality in General Practice. Abingdon, Oxon, Inglaterra: Radcliffe Medical Press.
- Bruce, J. (1990). Fundamental elements of the quality of care: A simple framework. *Studies in Family Planning*, vol 21, no. 2.
- Carey, R.G. y Lloyd, R.C. (2001). Measuring Quality Improvement in Healthcare: A guide to Statistical Process Control Applications. Milwaukee, Wisconsin: American Society for Quality.
- Cook, T.D. y Reichardt, Ch.S. (1995). Métodos cualitativos y cuantitativos en investigación evaluativa. Madrid: Ediciones Morata.
- Chen, H.T. (1990). Theory Driven Evaluations. Newbury Park, CA: Sage
- Departamento de Salud y Servicios Humanos, Estados Unidos (1998). "The Challenge and Potential for Assuring Quality Health Care for the 21st Century". Informe.
- Donabedian, A. (1980). The definition of quality: A conceptual exploration. En *Explorations in Quality Assessment and Monitoring*, vol. 1. Ann Arbor: Health Administration Press.
- Mohr, L. (1995). Impact Analysis for Program Evaluation. Sage Publications.
- OECD (1998). Review of the DAC Principles. <http://www.oecd.org/dac/evaluation>
- Restrepo, H.E. & Málaga, H. (2001). Promoción de la Salud: Cómo Construir Vida Saludable. Bogotá: Editorial Médica Panamericana.
- Rossi, P.H.; Freeman, H.E. & Lipsey, M.W. (1999). Evaluation: A Systematic Approach. Sage Publications.
- Scriven, M. (1991). Evaluation Thesaurus. Cuarta edición. Newbury Park, CA: Sage.
- Worthen, B.R. y Sanders, J.R. (1987). Educational Evaluation: Alternative Approaches and Practical Guidelines. New York: Longman.