



EVALUACIÓN DE MITAD DE PERÍODO
ESTRATEGIA NACIONAL DE SALUD 2010 - 2020
REVISIÓN BIBLIOGRÁFICA

Julio 2015

Departamento Estrategia Nacional de Salud

JAVIERA PACHECO
CAMILA SILVA
ODETTE URRUTIA
MIJAIL VILLAGRA
ANDREA ALBAGLI



Contenido

Contenido.....	1
Introducción.....	2
Síntesis de Revisión Bibliográfica.....	6
Definiciones.....	6
Funciones.....	7
Modelos de Evaluación.....	9
Temporalidad.....	9
Marco de Aprendizaje.....	13
Fundamentos Epistemológicos.....	14
Objetivos de Evaluación.....	18
Evaluación de Plan de Salud.....	26
Equipo de Evaluación.....	28
Referencias.....	30

INTRODUCCIÓN

La primera tarea de la Evaluación de Mitad de Período consistió en una revisión bibliográfica para orientar el diseño de la evaluación. Esta revisión se inició con la lectura del trabajo realizado previamente en la evaluación de los Objetivos Sanitarios 2000 – 2010 y otros documentos de la literatura a nivel nacional e internacional. Esto con el fin de reunir fuentes de referencia relevantes sobre evaluación de programas, que aportaran conocimientos sobre metodología y marco conceptual de evaluación, a la vez de contribuir con la identificación de temas centrales dentro del campo de evaluación de programas de salud que permite guiar la búsqueda de literatura académica en la materia.

Siendo que la evaluación de programas es un proceso iterativo de varias fases la revisión bibliográfica también se extenderá a lo largo de todo este trabajo. La síntesis presente en este informe corresponde a la revisión bibliográfica llevada a cabo para la primera fase de la evaluación de mitad de período; el diseño de la metodología de evaluación. En esta fase se revisaron un total de 29 artículos, y otros 23 quedaron seleccionados como material relevante para revisión en futuras fases de la evaluación.

Tabla N°1: Selección Artículos

ARTÍCULOS LEÍDOS EN PRIMERA REVISIÓN PARA ORIENTAR BÚSQUEDA			
Título	Fuente	Fecha	Temas para revisión bibliográfica
Objetivos Sanitarios para la Década 2000 - 2010. Evaluación a mitad del período.	Ministerio de Salud Chile.	2006	Tipos de evaluación: de resultados y proceso
Evaluación Objetivos Sanitarios de la Década 2000 – 2010.	Ministerio de Salud Chile.	2010	
Integración de los Determinantes sociales y equidad en Salud en el rediseño de los programas de Salud : Revisando algunas experiencias (PPT)	Dra. Orielle Solar – FLACSO Chile	2008	Evaluación desde la teoría Modelo de evaluación de equidad de acceso y barreras a servicios de salud
Documentos Técnicos Secretaría de DSS y Equidad N°1.	Ministerio de Salud, Chile.	2008	Evaluación realista
Making Evaluations Matter: a Workshop	Taller MINSAL-UofT	2009	
Monitoring, evaluation and review of	OMS	2011	Indicadores de inequidad

national health strategies Monitoring and evaluation of health systems strengthening	WHO	2009	Measuring service availability and readiness Measuring impact in the millennium development goals
---	-----	------	--

LISTA DE ARTÍCULOS INCLUIDOS EN REVISIÓN

LITERATURA GRIS

Título	Autor (es)	Año
Objetivos Sanitarios para la Década 2000 - 2010. Evaluación a mitad del período.	Ministerio de Salud	2006
Evaluación Objetivos Sanitarios de la Década 2000 – 2010.	Ministerio de Salud	2010
Documento Técnico I Secretaría de DSS y Equidad.	Ministerio de Salud	2008
Integración de los Determinantes sociales y equidad en Salud en el rediseño de los programas de Salud	Orielle Solar – FLACSO	2008
Monitoring, evaluation and review of national health strategies <i>Monitoreo, evaluación y análisis de la estrategias nacional de salud</i>	OMS	2011
Monitoring and evaluation of health systems strengthening <i>Monitoreo y Evaluación del fortalecimiento de los sistemas de salud</i>	OMS	2009
Midterm review of national health plans: an example from the United Republic of Tanzania <i>Evaluación de mitad de período de la estrategia nacional de salud: un ejemplo de la República Unida de Tanzania</i>	WHO	2015
Evaluation Brief describing the Five-Year Evaluation <i>Resumen Ejecutivo de la Evaluación tras cinco años</i>	The Global Fund ¹	2006
Health Inequality Monitoring <i>Monitoreo de Inequidades en Salud</i>	WHO	2013
Improving health service delivery in developing countries: from evidence to action <i>Mejorando los servicios de salud en países en desarrollo: de la evidencia a la acción</i>	The World bank	2009
Writing Terms of Reference for an Evaluation; A how-to guide <i>Una guía para escribir términos de referencia para una evaluación</i>	The World Bank	2011
Planning Evaluability Assessments: A synthesis of the literature with recommendations	Davis, R.	2013

LITERATURA ACADÉMICA

Título	Autor (es)	Año
--------	------------	-----

¹ “The Global Fund to Fight AIDS, Tuberculosis and Malaria mobilizes and invests nearly US\$4 billion a year to support programs run by local experts in countries and communities most in need. As a partnership between governments, civil society, the private sector and people affected by the diseases, the Global Fund is accelerating the end of AIDS, TB and malaria as epidemics”.

The Practice of Health Program Evaluation <i>La práctica de evaluación de programas de salud</i>	David Grembowski	2001
Realist Evaluation <i>Evaluación Realista</i>	Ray Pawson & Nick Tilley	1997
Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations <i>Medición de impacto en la era de los Objetivos del Milenio y más allá: un nuevo enfoque para la evaluación de efectividad a gran escala</i>	Cesar G Victora y cols.	2010
The evaluation of health and social care partnerships: an analysis of approaches and synthesis for the future <i>Evaluación de alianzas de trabajo en salud y proyectos sociales: un análisis de enfoques y sistematización para el futuro</i>	Helen Dickinson	2006
Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research <i>¿Cumple con su promesa la evaluación realista? Una revisión de estudios empíricos publicados sobre investigación de sistemas de salud</i>	Bruno Marchal y cols.	2012
Theories of Change and Realistic Evaluation <i>Teorías del cambio y Evaluación Realista</i>	Avril Blamey & Mhairi Mackenzie	2007
Evaluating complex public health interventions: theory, methods and scope of realist enquiry <i>Evaluando intervenciones complejas de salud pública: teoría, métodos y alcance de la evaluación realista</i>	James B. Connelly	2007
Ten steps to making evaluation matter <i>10 pasos para una evaluación significativa</i>	Sanjeev Sridharan & April Nakaima	2010
Health service coverage and its evaluation <i>Evaluación de la cobertura de servicios de salud</i>	Tanahashi	1978
Monitoring Intervention Coverage in the Context of Universal Health Coverage <i>Monitoreo de la cobertura de intervención en el contexto de cobertura universal de salud</i>	Ties Boerma y cols.	2014
Monitoring Progress towards Universal Health Coverage at Country and Global Levels <i>Monitoreo del progreso hacia la cobertura universal de salud a nivel país y global</i>	Ties Boerma y cols.	2014
Evaluation models and evaluation use <i>Modelos de evaluación y usos de la evaluación</i>	Damien Contandriopoulos & Astrid Brousselle	2012
Evaluation Designs for Adequacy, Plausibility, and Probability of Public Health Programme Performance and Impact <i>Diseños de evaluación sobre la adecuación, posibilidad y probabilidad del rendimiento e impacto de programas de salud pública</i>	Habicht, J.P., C.G. Victora, & J.P. Vaughan.	1999



Equity-Oriented Monitoring in the Context of Universal Health Coverage <i>Monitoreo de equidad en el contexto de cobertura universal de salud</i>	Ahmad Reza Hosseinpoor y cols.	2014
Evaluación de programas de salud	María Naranjo	2006
Evaluación en salud pública; ¿todo vale?	Editorial Sanitaria	Gaceta 2007

SÍNTESIS DE REVISIÓN BIBLIOGRÁFICA

Definiciones

Los programas y políticas públicas de salud se diseñan con la intención de mejorar el estado actual de una condición que se considera perjudicial (OMS, 2013). La evaluación de programas y políticas es un proceso de análisis que busca identificar si los programas y políticas diseñados para subsanar una dificultad observada cumplen o no con su fin.

El monitoreo es un proceso de levantamiento de datos que señala los efectos de los programas implementados, indicando si es necesario realizar cambios en las acciones tomadas para alcanzar los cambios esperados. De este modo, el monitoreo de programas de salud es tanto útil como necesario, ya que tiene la habilidad de trazar los resultados de una intervención. El monitoreo entonces, debe ser un proceso iterativo y cíclico que se ejecute continuamente (OMS, 2013).

“EL MONITOREO ES UN PROCESO DE PLANTEAR REITERADAMENTE UNA PREGUNTA DE INVESTIGACIÓN A LO LARGO DEL TIEMPO. EN EL MUNDO DE LAS POLÍTICAS PÚBLICAS, LA PREGUNTA DE INVESTIGACIÓN USUALMENTE REFLEJA LA MEDICIÓN DE UNA CONDICIÓN QUE LA POLÍTICA BUSCA CAMBIAR”. OMS, 2013¹.

Si bien el monitoreo contribuye a identificar el éxito o estancamiento de las intervenciones, éste por sí solo no puede explicar las causas de las dificultades observadas. El monitoreo se puede entender entonces como un sistema de vigilancia que identifique las áreas de trabajo más débiles, para entonces destinar recursos para investigar en mayor profundidad e identificar las causas de los problemas evidenciados (OMS, 2013).

La evaluación es un proceso de análisis que se construye en base a los datos que provee el monitoreo. En esta instancia se hace una reflexión más profunda que toma en consideración los cambios contextuales y las explicaciones alternativas a los cambios observados, preguntándose así por la posibilidad de atribuir con certeza dichos cambios a la intervención bajo estudio (OMS, 2011).

Según la Organización Mundial de la Salud, la evaluación es un área comúnmente ignorada en el trabajo de E&M (evaluación y monitoreo) (OMS, 2011).

“LA EVALUACIÓN SE DEFINE COMO UN PROCESO DE ANÁLISIS RIGUROSO Y CIENTÍFICO SOBRE LA INFORMACIÓN DE LAS ACTIVIDADES, CARACTERÍSTICAS, RESULTADOS E IMPACTO DE UN PROGRAMA, QUE DETERMINA LOS MÉRITOS DE ÉSTE. ESTO ES DISTINTO AL MONITOREO, QUE SE DEFINE COMO EL SEGUIMIENTO Y REPORTE CONSTANTE SOBRE LA INFORMACIÓN PRIORITARIA DE UN PROGRAMA Y SUS RESULTADOS E IMPACTOS ESPERADOS”¹. OMS, 2011.

El Banco Mundial (2009) hace referencia a la distinción entre monitoreo y evaluación al señalar que saber **cómo** fortalecer servicios de salud es diferente de saber **qué** se debe fortalecer. Junto con lo anterior, el tipo de evidencia científica que se necesita para responder a las preguntas **como** es diferente al tipo de evidencia necesaria para responder las preguntas **qué**. Según el Banco Mundial (2009), ambos tipos de conocimiento son importantes para los tomadores de decisión.

Funciones

La función principal de la evaluación es la mejora constante de programas y políticas públicas. Todo programa puede ser mejorado, independiente de los resultados que se obtengan, incluso si muestra ser exitoso. En este sentido, la evaluación resulta imprescindible como instrumento de mejora y perfeccionamiento y como garante de la calidad de una intervención.

La evaluación también tiene la función de proveer un medio para rendir cuentas sobre los esfuerzos desplegados por la intervención y los resultados obtenidos. Esto adquiere un mayor sentido cuando se refiere a programas y políticas públicas, ya que el rendimiento de cuentas debe formar parte indiscutible del sistema democrático. Este rendimiento de cuentas debe darse a varios niveles, incluyendo de forma especial a la ciudadanía: el ciudadano/a tiene derecho a saber no sólo en qué se están empleando los fondos públicos sino también con qué grado de idoneidad, eficacia y eficiencia se están asignando, gestionando y empleando dichos fondos.

Por último, también cumple la función de informar acciones futuras. Aunque una evaluación se centra esencialmente en un programa concreto en el presente y dentro de su contexto dado, las evaluaciones arrojan una información sistemática que contribuye al "acervo científico" no sólo de ese programa, servicio o política concreta, sino también al abordaje general de determinados problemas públicos. Como dice Carol Weiss, *"el estudio de las políticas y programas en acción es una oportunidad para estudiar los estímulos que conducen al cambio y las condiciones ambientales bajo las cuales es más probable que ocurran cambios"* (Weiss, 1998).

En cuanto a la primera función— su uso para adoptar medidas de mejora a la intervención bajo estudio — Contandriopoulos & Brousselle (2012), plantean que la naturaleza de los sistemas colectivos sugiere que la evidencia científica rara vez resuelve directamente las dificultades organizacionales o a nivel de políticas públicas. Los autores plantean que para ser relevante, útil y significativa, la evidencia debe enmarcarse dentro de las opciones políticas. David Grembowski (2001) plantea algo similar, por lo cual llama a la evaluación de programas *"Una Obra de Tres Actos"*, ilustrando el trabajo de evaluación de programas como una obra de teatro en tanto considera que hay dos procesos dentro de una evaluación; uno técnico y uno político. El proceso técnico es la aplicación de métodos de investigación con el fin de llegar a conclusiones prácticas que informen decisiones racionales. En la práctica, el autor plantea que hay un segundo proceso, donde la evaluación de programas no es sólo una actividad técnica, sino que un proceso influenciado por intereses políticos. Con esto Grembowski hace referencia al cúmulo de valores inherentes a la forma en que los involucrados comprenden el fenómeno en el cual intervienen. Por ejemplo, los actores involucrados en el programa de salud bajo estudio pueden o no priorizar la equidad en salud como un elemento importante, lo cual se reflejará al plantear preguntas de evaluación, incluyendo o excluyendo dicho criterio del proceso.

"PORQUE LA POLÍTICA ES EL MODO EN QUE ADJUNTAMOS VALOR A LOS HECHOS EN NUESTRA SOCIEDAD, LA POLÍTICA Y LOS VALORES SON INSEPARABLES DE LA EVALUACIÓN DE PROGRAMAS DE SALUD". GREMBOWSKI, 2001

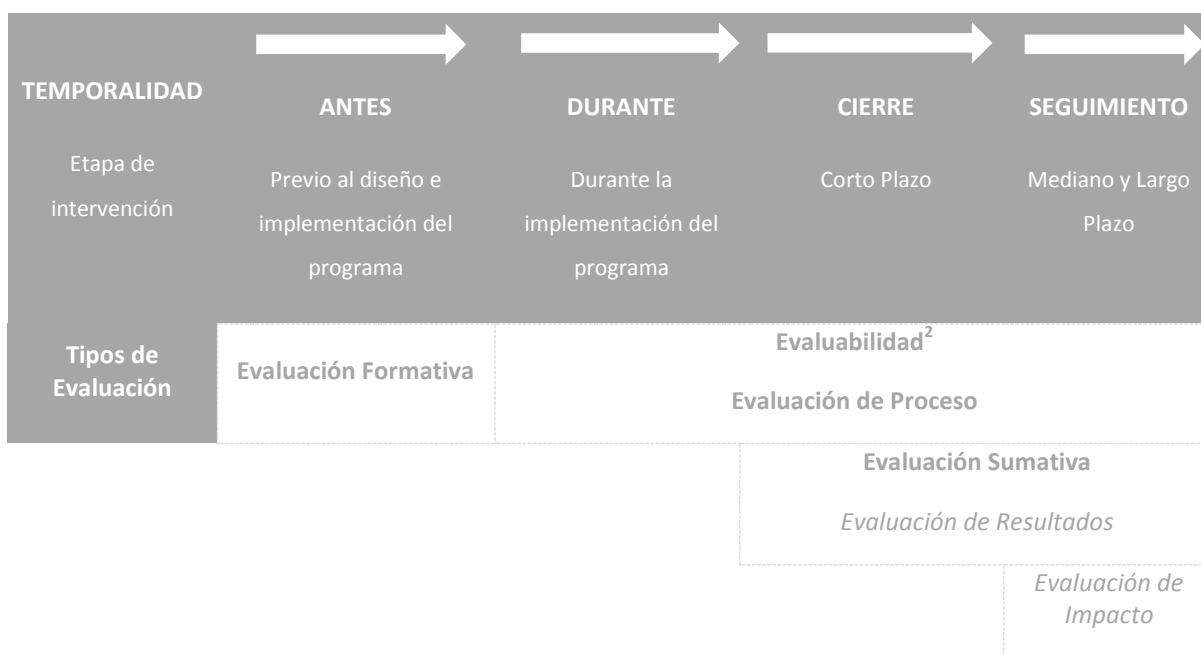
Modelos de Evaluación

Los tipos de evaluación pueden clasificarse según varios criterios, los que **no son mutuamente excluyentes**. A continuación se describen los diferentes enfoques hacia la evaluación en base a una serie de criterios; temporalidad, marco de aprendizaje, fundamentos epistemológicos y objetivos de evaluación. Para cerrar esta sección se exponen las consideraciones del trabajo de evaluación que son específicas a la evaluación de sistemas de salud.

Temporalidad

Una distinción general de los enfoques en evaluación de programas se relaciona al período en que se instala dentro del desarrollo de implementación del mismo; antes del diseño e implementación del programa, durante su implementación o en las fases posteriores a su ejecución completa (seguimiento). A su vez, en cada uno de estos momentos se pueden plantear diferentes objetivos de evaluación.

Figura N°1: Temporalidad en Evaluación de Programas



² Traducción no oficial del término en Inglés: 'Evaluability Assessment'.

Evaluación Formativa: En la fase previa a la implementación de programas no se pueden plantear preguntas sobre el funcionamiento del mismo debido a que no hay experiencia de trabajo para hacerlo. En esta fase inicial, las preguntas de evaluación se orientan a rescatar información relevante que permita perfeccionar el diseño del programa, ya sea ante el diseño de un programa nuevo o la adaptación de un programa que se ha desarrollado en otro contexto. Esta evaluación es la evaluación formativa, y tal como sugiere su nombre su objetivo es informar el diseño e implementación de programas aún por ejecutarse, por medio de revisión de la literatura, y análisis situacional del contexto de implementación por medio de entrevistas con beneficiarios, profesionales y actores relevantes del espacio a intervenir (O'Hara, 2013).

Evaluabilidad (*Evaluability Assessment*): Davies (2013), realizó una revisión sistemática que incluyó 133 documentos sobre la 'evaluabilidad' de programas, cuya definición se ha consensuado entre agencias de desarrollo internacional como *"el nivel en que una actividad o proyecto puede ser evaluado de modo confiable"*. Este proceso de indagación previo a una evaluación implica que no siempre es posible o recomendable evaluar un programa dado. Según Davies (2013), la evaluabilidad de un programa debe definirse en términos teóricos y prácticos. La evaluabilidad teórica explora la naturaleza del programa al visualizar su teoría de causa-efecto³, preguntándose si es posible evaluar la intervención siguiendo la lógica que se ha planteado de esta. La evaluabilidad práctica tiene relación con la disponibilidad de información relevante y el funcionamiento de los sistemas que levantan dichos datos. Junto con los aspectos teóricos y prácticos de la evaluabilidad de un programa, se debe considerar también la utilidad de la evaluación. En resumen, una indagación de evaluabilidad debe considerar los siguientes aspectos:

- ✘ El diseño del programa
- ✘ La disponibilidad de información
- ✘ El contexto institucional

Según el autor (Davies, 2013), idealmente, la evaluabilidad debe desarrollarse por evaluadores externos y sus resultados deben tener consecuencias para el diseño de la evaluación posterior (en

³ La teoría de programas será descrita en mayor profundidad más adelante en este documento.

caso de proseguir a una evaluación del programa), el diseño de un marco de evaluación y administración del programa, y/o el diseño del programa. Si bien no existe un marco estándar que indique los procesos de una evaluabilidad, algunos pasos comunes observados en la literatura son:

- ✧ Identificación del alcance del programa y los resultados esperados
- ✧ Identificación de recursos disponibles
- ✧ Revisión de la documentación disponible
- ✧ Negociación con actores clave
- ✧ Desarrollo de recomendaciones
- ✧ Retroalimentación de hallazgos a los actores relevantes
 - Diseño y teoría del programa
 - Sistemas y capacidad de evaluación y administración del programa
 - Preguntas de evaluación de interés para los actores involucrados
 - Posibles diseños de evaluación

Este tipo de evaluación puede llevarse a cabo durante cualquier fase del programa una vez iniciado, dado que es un paso previo que estima la pertinencia de desarrollar una evaluación y los métodos más adecuados para hacerlo, independiente del tipo de preguntas que se quiera plantear en la evaluación misma.

Evaluación de proceso: Una **evaluación de proceso** indaga en la implementación de tareas y estrategias diseñadas para identificar si en efecto se aplican como fue programado. Examina tareas y procedimientos del programa, la cantidad de actividades realizadas y la audiencia a quien logra llegar las intervenciones. Este tipo de evaluación es fundamental para entender los motivos por los cuales un programa alcanza o no sus objetivos (Grembowski, 2001). La evaluación de proceso puede llevarse a cabo durante la implementación de la intervención o tras a su término. Grembowski (2001) sugiere las siguientes preguntas como parte de una evaluación de proceso:

- ✧ La intervención, ¿se implementó tal como fue planificada?
- ✧ ¿Alcanzó el grupo beneficiario identificado?
- ✧ ¿Qué tipo de servicios recibieron los beneficiarios?
- ✧ ¿Estuvieron satisfechos los beneficiarios con los servicios entregados?
- ✧ ¿Cuál es el costo promedio de la intervención por persona?

Evaluación Sumativa: Grembowski (2001) define a la evaluación sumativa como aquella que busca establecer los efectos alcanzados por una intervención. Este tipo de evaluación se pregunta por el grado en que se alcanzó o no los objetivos propuestos. Adicionalmente, se puede distinguir dentro de este tipo de evaluación entre evaluación de resultados y evaluación de impacto. Esta distinción busca especificar la cadena de resultados, diferenciando los resultados directos que tiene y espera tener un programa y los impactos de largo plazo que pueden o no desprenderse de esos resultados. Por ejemplo, en el área de la salud sexual y reproductiva, los resultados inmediatos de la educación en salud a la población adolescente, disponibilidad de métodos anticonceptivos modernos y el aumento de accesibilidad a los mismos, puede o no generar el impacto de largo plazo deseado de reducir la tasa de embarazos no planificados en la adolescencia.

La evaluación sumativa sólo puede llevarse a cabo una vez que el programa haya finalizado, o, en el caso de intervenciones permanentes, una vez que su implementación lleve el tiempo suficiente para que los cambios esperados se hayan podido manifestar. Las preguntas orientadas a levantar información sobre los resultados de corto plazo pueden plantearse más tempranamente que aquellas dirigidas a los impactos de largo plazo. El tiempo de evaluación de cada uno de estos aspectos dependerá del tema de salud en cuestión según las características del fenómeno. Siguiendo el ejemplo anterior, el acceso a métodos anticonceptivos por la población adolescente podría indagarse tras un año de iniciado el programa, mientras que la tasa de embarazo adolescente es un cambio epidemiológico que tarda más años en fluctuar como consecuencia del aumento de acceso.

Es importante que dentro de una evaluación sumativa se evalúen tanto los resultados esperados como los resultados no esperados. Esto último se revisará con mayor profundidad más adelante en este documento.

Marco de Aprendizaje

Sridharan & Nakaima (2010) señalan que existe una tendencia a asumir que existe sólo un tipo de aprendizaje a desprender de procesos de evaluación. Sin embargo, los autores señalan que una evaluación puede tener distintos objetivos de aprendizaje y es importante que éstos sean explícitos en el trabajo evaluativo. Los objetivos de aprendizaje posibles que Sridharan & Nakaima señalan son:

Aprendizaje sobre la política pública (*policy learning*): Toda intervención de salud consiste en la traducción de una idea de tomadores de decisión a su implementación en terreno, proceso en el cual hay una adaptación de la idea inicial a su ejecución. Este tipo de aprendizaje – sobre la política pública – consiste en el intento de visualizar y entender esta traducción desde el marco conceptual a su aterrizaje operativo. Algunos ejemplos de preguntas que apuntan a este tipo de aprendizaje son:

- ✧ ¿Cuáles eran las aspiraciones de la política pública?
- ✧ ¿Cuál fue la teoría fundacional de esta política?
- ✧ ¿Son consistentes los objetivos de la intervención con las aspiraciones de la política?
- ✧ ¿Es la implementación de la intervención consistente con las aspiraciones de la política?

Aprendizaje organizacional (*organizational learning*): Las evaluaciones son también una oportunidad para aprender sobre las estructuras organizacionales que sustentan el programa bajo estudio. Este marco de aprendizaje permitiría entender mejor qué estructuras operan en el programa dado, cuán adecuadas son para dar respuesta a las necesidades presentes y qué tipo de estructura se necesitaría reforzar o agregar para mejorar el funcionamiento del programa.

Aprendizaje de proceso: Se orienta a levantar información sobre los procesos de diseño e implementación de las intervenciones bajo estudio. A lo anterior Sridharan & Nakaima (2010) agregan que dentro de un análisis de proceso también es de gran valor agregar preguntas sobre la sostenibilidad de los programas.

Aprendizaje sobre barreras: Conocer acabadamente las barreras que experimenta la intervención en su implementación es de gran utilidad para entender los motivos de éxito o de fracaso de ésta misma. En base a este conocimiento se pueden levantar medidas de mejora para la intervención bajo estudio y para futuras intervenciones.

Aprendizaje sobre impacto a nivel individual: La diferencia que una intervención hace sobre la vida de los beneficiarios a quienes se dirigió el trabajo. *¿Mejóro el programa la calidad de vida de los involucrados?* Según los autores, la mayoría de las evaluaciones de programas se centran en esta pregunta, obviando los otros aprendizajes que pueden obtenerse de un proceso de evaluación.

Según Sridharan & Nakaima (2010), definir el marco de aprendizaje que se espera obtener de una evaluación es un paso esencial en un proceso de diseño. Dependiendo de los recursos disponibles y la magnitud del programa, la evaluación puede incluir a sólo uno o una combinación de los marcos de aprendizaje aquí mencionados.

Fundamentos Epistemológicos

La **evaluación realista** es un tipo de evaluación basada en la teoría creada por Pawson & Tilley (1997). Considera que todo programa de intervención son teorías y el rol del evaluador es identificar dichas teorías y su interacción con el contexto, reconociendo así **qué** intervenciones funcionan, para **quiénes**, bajo **qué circunstancias** (Contandriopoulos & Brousselle, 2012). La evaluación realista surgió como respuesta a la limitación de estudios experimentales, que si bien pueden ser adecuados para estudios de efectividad no logran entregar información ante programas implementados en sistemas complejos. Es por esto que la implementación de la teoría a la práctica resulta frecuentemente infructuosa (Marchal y cols., 2012).

Según sus creadores (Pawson & Tilley, 1997) la evaluación realista se basa en las siguientes premisas; los programas son **teorías**, están **insertos** en un contexto, son **activos** y son parte de sistemas **abiertos**. Un programa como teoría significa que en los fundamentos de una intervención existe una serie de explicaciones de causa y efecto y ligan a una situación considerada problemática con sus

orígenes y los modos de incentivar el cambio. El estar insertos en un contexto implica que es a través de la interacción de sus sistemas sociales que se generan cambios, y por lo mismo los efectos de un programa están sujetos a una serie de condiciones. Por ejemplo, el éxito de un programa de educación y capacitación laboral para personas privadas de libertad dependerá de las capacidades individuales de los participantes y los monitores, de la relación interpersonal entre ambos, el balance que la institución haga entre el tiempo dedicado a rehabilitación y a represión de quienes están presos, y, por último, al sistema socio político y estado de bienestar que puede apoyar o limitar la reincorporación de personas privadas de libertad en la sociedad. La condición de programas como activos se refiere a que la gran mayoría de los programas requieren de una población beneficiaria que se relacione activamente con el programa. Son muy pocos los ejemplos de programas ‘pasivos’, tales como la fluorización del agua que tienen efectos en la población independiente de la voluntad de ésta. Esta condición de programas de intervención implica que la interpretación que hacen los usuarios o beneficiarios de un programa acerca de dicho programa se torna en un elemento esencial. Por último, el ser parte de sistemas abiertos significa que los programas no pueden ser aislados ni mantenidos constantes. El sistema social en el que se inserta un programa está sujeto a cambios que no son necesariamente anticipados, tales como cambios políticos, cambios de personal contratado, cambios tecnológicos entre otros. Estos cambios siempre tendrán un impacto sobre el programa y su funcionamiento. Más aún, un programa exitoso puede generar ciertos cambios en el contexto, cambiando las mismas condiciones que lo hicieron funcionar en primera instancia (Pawson & Tilley, 1997). Por todo lo anterior, sólo se puede evaluar un programa adecuadamente indagando en su proceso de cambio, considerando la interacción entre mecanismos y contexto.

Un concepto central de la evaluación realista es el **mecanismo**. Los mecanismos son los medios que generan cambios, que son gatillados por la interacción entre el contexto y la intervención. Por ejemplo, el “club del desayuno” es una medida común para mejorar el rendimiento escolar temprano de los alumnos. La intervención consiste en compartir un desayuno entre padres, profesores y alumnos todas las mañanas antes del comienzo de la jornada escolar. La teoría de esta intervención consiste en los efectos cognitivos positivos que una buena alimentación puede tener a nivel individual. No obstante, ante el éxito del programa existen diferentes interpretaciones, que ejemplifican los diversos mecanismos que pueden ser gatillados por la intervención. 1) El

rendimiento escolar de los estudiantes mejoró debido a que gracias al desayuno previo a la jornada escolar, los alumnos tenían tiempo de socializar con sus compañeros, de modo que al comenzar las clases podían prestar más atención a los contenidos. 2) El rendimiento escolar de los estudiantes mejoró debido a que gracias al desayuno previo a la jornada escolar los estudiantes pudieron compartir socialmente con sus profesores, generando un vínculo más estrecho que facilitó la dinámica en la sala de clases. 3) El tener reuniones diarias entre apoderados y profesores facilitó la comunicación entre ambos, facilitando así una mejor coordinación entre el hogar y la escuela. La intervención del desayuno en conjunto puede gatillar una serie de mecanismos que generan un cambio, lo que a su vez depende del contexto en que se aplica (Pawson & Tilley, 1997).

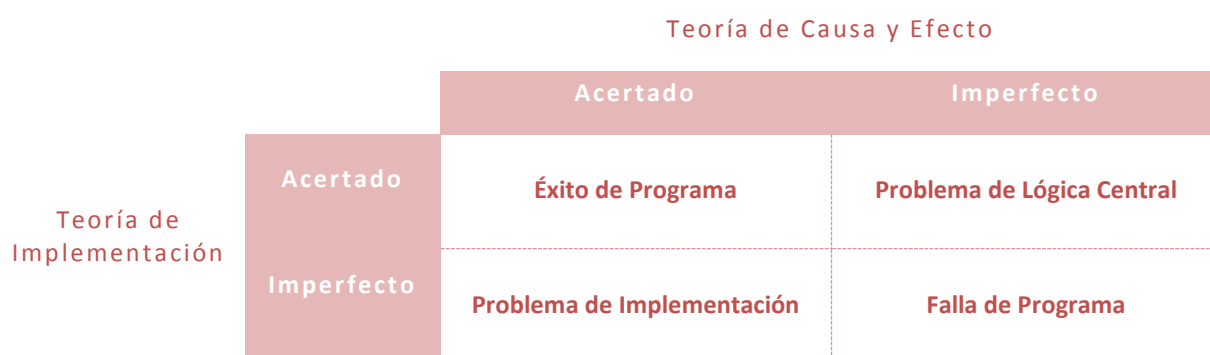
David Grembowski (2001) también acoge un enfoque de evaluación basada en la teoría. Para el autor la evaluación de programas es un acto que consta de tres etapas; 1) planteamiento de la pregunta, 2) aplicación de metodologías para responder la pregunta y 3) uso de respuestas para la toma de decisiones. A primera vista, este esquema ya releva un punto importante dentro de la evaluación, situando el planteamiento de preguntas al mismo nivel de importancia que el trabajo de indagación para responderla mediante métodos de recolección y análisis de datos. En otras palabras, dado que la evaluación es en esencia un proceso de plantear y responder preguntas asociadas a una intervención que busca generar un cambio en la sociedad, el éxito de una evaluación depende en primer lugar, de cuán buenas preguntas se plantean. Una metodología de calidad que logra levantar información confiable y válida no proveerá mayor beneficio si responde a una pregunta que no es relevante o prioritaria.

El autor especifica cuatro pasos para plantear preguntas de evaluación; 1) especificar la teoría del programa, 2) especificar los objetivos del programa, 3) traducir teoría y objetivos del programa en preguntas de evaluación y 4) priorizar las preguntas a responder.

Especificar la teoría de programa: El éxito de un programa depende de; 1) cuán adecuado es su fundamento de base para el cambio esperado (teoría de causa y efecto), 2) cuán apropiada es su estrategia de implementación (teoría de implementación) y 3) si la estrategia escogida se implementó como se pretendía en el diseño.

La teoría de causa y efecto es la lógica subyacente a una intervención que explica la relación de cambio entre una acción y sus efectos esperados, y la teoría de implementación es la elección de estrategias para implementar estas medidas. Examinar la teoría de causa y efecto y la teoría de implementación nos permite entender por qué un programa funciona o no funciona (Grembowski, 2001).

Figura N°2: Teoría de Programa



Entender la teoría de causa y efecto de un programa no sólo implica analizar los fundamentos de base del programa en cuestión, sino que implica hacer una revisión de los **modelos teóricos en la literatura** sobre el fenómeno de salud que el programa busca tratar. Así, la evaluación puede señalar nuevas ideas sobre cómo el programa debiera funcionar. Esta información ayuda a plantear preguntas de evaluación para entender qué fenómenos (quizás no cubiertos por el programa) están teniendo un impacto en los resultados de la intervención.

Especificar los objetivos del programa: El equipo evaluador debe formular todos los objetivos del programa en modos que puedan ser medidos, siendo éstos los criterios de éxito en base a los cuales se obtendrán los juicios sobre la intervención. La mayoría de los programas tienen múltiples objetivos que se relacionan entre sí – objetivos de corto y mediano plazo y objetivos de largo plazo, que también pueden ser divididos en objetivos de implementación y objetivos de resultado, los que conforman una **jerarquía de objetivos** o **cadena de resultados**.

TODO PROGRAMA DE SALUD SE BASA EN UNA TEORÍA DE CAUSA Y EFECTO Y EN UNA TEORÍA DE IMPLEMENTACIÓN,
SEA ESTA EXPLÍCITA O IMPLÍCITA.

Objetivos de Evaluación

Otra forma de categorizar la literatura sobre evaluación es mirar los objetivos de evaluación que se plantean. Es importante reiterar que las categorías expuestas en este documento son sólo para fines de organización del material. Los enfoques expuestos a continuación proponen una pregunta de evaluación específica, un área relevante a indagar en una intervención, lo que es complementario con las alternativas de las categorías expuestas anteriormente. Las preguntas a continuación pueden plantearse independiente de la fase en que se encuentre el programa, las bases epistemológicas con las que se comprende el fenómeno y su estudio y el nivel de aprendizaje que se desee obtener. En esta sección se expondrán las siguientes preguntas de evaluación, los que no son exhaustivos ni excluyentes:

- ✘ Diseños de evaluación para estudios de adecuación, plausibilidad y probabilidad del desempeño e impacto de programas de salud pública – Habicht & cols
- ✘ Evaluación Democrática – House & Howe
- ✘ Evaluación de Empoderamiento – Fetterman & Wandersman
- ✘ Evaluación de Alianzas de Trabajo – Dickinson H.
- ✘ Consecuencias no Intencionadas – Grembowski

Adecuación, plausibilidad y probabilidad en la provisión, utilización, cobertura e impacto:

Habicht & cols (1999), proponen una matriz compuesta por dos ejes para escoger un diseño de evaluación. El primer eje señala los indicadores de interés, siendo éstos de disponibilidad, utilización, cobertura o impacto. El segundo eje expone el grado de seguridad que se desea tener en conclusiones de atribución; inferencias de adecuación, plausibilidad o probabilidad. Esta matriz es

una guía que permite traducir los motivos que hacen necesaria una evaluación en un diseño de evaluación adecuado para su respuesta.

Los autores señalan que en la literatura hay poco debate sobre los motivos por los que se desarrolla una intervención. Según ellos sin embargo, esta pregunta es esencial para escoger un buen modelo de evaluación. El objetivo principal de una evaluación es influenciar las decisiones sobre un programa, y cuán compleja y precisa debe ser depende de quiénes son los tomadores de decisión y qué tipos de decisiones tomarán según los resultados obtenidos. Ante el caso que haya varios tomadores de decisión involucrados con expectativas diversas de uso de los resultados, la evaluación debe adaptarse de tal modo que acoja todas dichas necesidades.

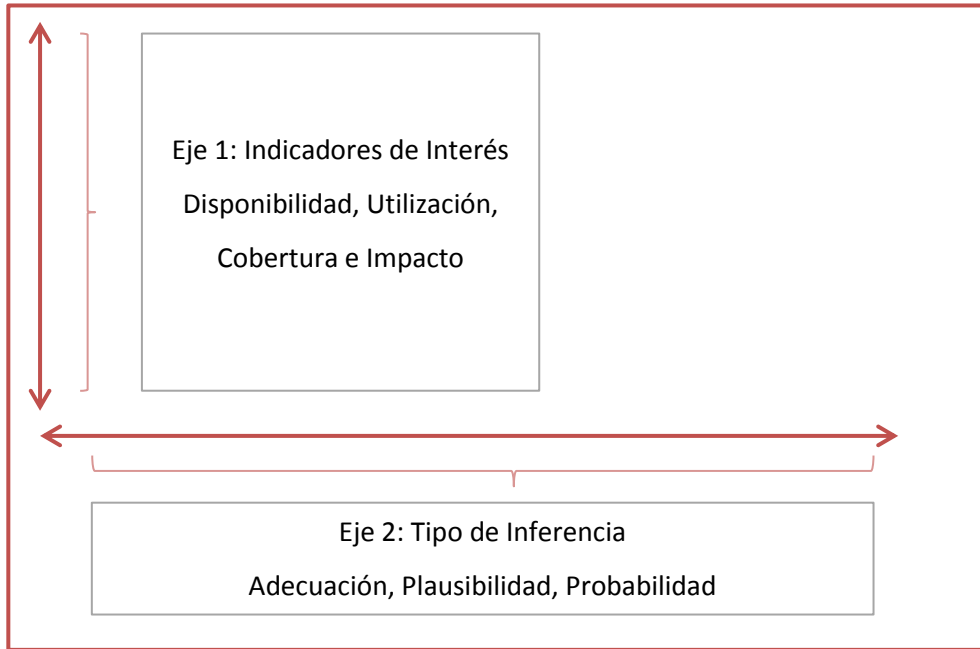
Según Habicht & cols (1999) hay otros factores que también afectan la elección de un modelo de evaluación, tales como los plazos de trabajo y los costos. Los plazos de trabajo son relevantes dado que la evaluación debiera entregar respuestas a los tomadores de decisión de modo oportuno para que puedan incorporarlo en su proceso de toma de decisión. Es inútil proveer información confiable y válida de una evaluación perfecta si su entrega es posterior al plazo de toma de decisión. Es por esto que los evaluadores, además de determinar qué decisiones se van a tomar en base a los resultados de la evaluación, deben identificar el período en que dichas decisiones ocurrirán. El diseño y conducción de una evaluación debe entonces planificarse de modo que cumpla con estas fechas.

“TODOS LOS DISEÑOS DE EVALUACIÓN DEBEN INCLUIR LA OPORTUNIDAD DE LA INFORMACIÓN COMO PARTE DE SUS OBJETIVOS”. HABICHT & COLS (1999)⁴

Dado lo anterior, la primera tarea del equipo evaluador es definir su grupo objetivo de los resultados de la evaluación, dado que sus intereses, experiencia, y plan de trabajo, influirán en el tipo de preguntas a plantearse y el modelo de evaluación a adoptar para su respuesta.

⁴ Traducción no oficial desde texto Habicht & cols, 1999.

Figura N° 3: Ejes de matriz de toma de decisión



Los indicadores de interés – disponibilidad, utilización, cobertura e impacto – se refieren al aspecto de la intervención que la evaluación quiere estudiar; la capacidad de los servicios, el uso de éstos, el uso de éstos en relación a la población que presenta necesidad o el impacto que generan los servicios en el estado de salud resultante. El tipo de inferencia por su parte se refiere al grado de certeza con el cual se quiere hacer un juicio de atribución entre resultados e intervención; es decir, con cuánta seguridad se puede decir que los resultados observados se deben a la intervención bajo análisis y no a otras causas.

La evaluación de **adecuación** indaga en la capacidad de los programas en alcanzar los objetivos propuestos según los criterios de éxito establecidos en un inicio. Este tipo de evaluación mide si los indicadores de comportamiento o de salud han mejorado en los beneficiarios de la intervención bajo estudio, ya sea con mediciones longitudinales o de corte transversal. Para este tipo de estudio se pueden utilizar datos secundarios lo que puede reducir importantemente los costos. Dado a que la inferencia de adecuación se limita a observar los cambios generados por la

intervención, no se puede afirmar con precisión si se deben a dichas iniciativas o a condiciones secundarias. Los cambios pueden obedecer a causas externas tales como la mejora de condiciones socioeconómicas o el trabajo de otros programas en el área de análisis.

Por otra parte, este tipo de evaluación puede subestimar éxitos alcanzados por el programa en caso que su presencia haya mantenido indicadores de salud previos bajo condiciones socio-ambientales empeoradas, como podría ser el caso de sequía, aumento del desempleo o desastres naturales.

A pesar de sus limitaciones en establecer validación causal entre la intervención y sus consecuencias, este enfoque puede ser considerado como una aproximación al estudio de programas y, en caso que los resultados señalen patrones de cambio no esperados se puede profundizar en el estudio con metodologías más sofisticadas y costosas.

Las preguntas de **plausibilidad** suman un elemento extra a la medición de cambios en la población beneficiaria, incorporando en el análisis los factores externos que podrían haber intervenido para así aumentar la confiabilidad en la atribución de causalidad. Las preguntas de plausibilidad requieren de diseños metodológicos de mayor rigurosidad con la inclusión de grupos control. La característica central de una evaluación de plausibilidad es la consideración de factores externos al programa en la cadena de causalidad.

Las evaluaciones de **probabilidad** buscan asegurar que sólo existe una pequeña probabilidad que la diferencia entre grupos control y grupos con intervención se deba a factores de confusión, sesgo o casualidad. Esto requiere de randomización de muestras. Si bien esta modalidad ofrece una mayor confiabilidad en los resultados, hay ciertas consideraciones que lo hacen poco factible para estudios de salud pública. En primer lugar, el equipo evaluador debe estar presente en etapas tempranas de la intervención para la randomización de las muestras, cuando en la mayoría de los casos los programas suelen solicitar evaluaciones tiempo después de su implementación. Por otro lado las presiones políticas pueden limitar las indicaciones del equipo evaluador en cuando a la selección de poblaciones intervenidas y poblaciones control. Una alternativa a consideraciones políticas es el diseño de randomización por etapas, donde una intervención se implementa en un inicio siguiendo

la randomización del estudio y es luego extendida a todas las poblaciones, incluyendo las que en un comienzo fueron de control. Esta alternativa es importante de considerar por motivos tanto políticos como éticos, pero es sólo apropiada para evaluar efectos que no tardan un largo tiempo en revelarse luego del comienzo de la intervención. Las dificultades de implementación de este tipo de estudios puede generar la creación de espacios de intervención que son diferentes a la realidad, por lo cual si bien aumenta la validez interna del estudio, su validez externa disminuye. Por todas estas razones el uso de estas metodologías es limitado en evaluaciones de programas (Habicht & cols, 1999). Esto es también el caso para los estudios de plausibilidad que también requieren del estudio de grupos de control.

Consideraciones de Habicht & cols, 1999 sobre Temporalidad y Evaluación

- ✘ Las inferencias de probabilidad requieren que el equipo evaluador esté integrado a la implementación previo a su comienzo para permitir la randomización de la muestra
- ✘ Todos los métodos longitudinales requieren de una línea base a recopilar antes del inicio del programa, o la inclusión de datos secundarios confiables en caso de no tener una línea base propia del programa
- ✘ Las evaluaciones sobre disponibilidad y utilización de servicios puede llevarse a cabo en fases tempranas de la implementación y con mayor frecuencia
- ✘ Las evaluaciones de cobertura e impacto se efectúan más adelante en el ciclo del programa. Los impactos en el estado de salud toman al menos entre 3 y 5 años en efectuarse, e incluso más en enfermedades con largos períodos de incubación como el VIH-SIDA

Evaluación Democrática

Contandriopoulos & Brousselle (2012) hacen una revisión del modelo de evaluación propuesto por House y Howe (1999); evaluación democrática, cuyo fin es proveer a la sociedad información confiable y relevante para apoyar debates públicos. Este tipo de evaluación debe considerar todos los argumentos presentes dentro de una situación dada, para lo cual se necesita de un evaluador externo e independiente.

Este tipo de evaluación es adecuada para contextos donde hay gran polarización entre los actores clave dentro de un debate. Además, se suma el reconocimiento que la distribución de poder es

desigual entre las partes, por lo cual el rol del evaluador es fomentar y facilitar el debate democrático restituyendo estas dinámicas de poder.

Evaluación de Empoderamiento

Contandriopoulos & Brousselle (2012) también describen el modelo de evaluación de Fetterman y Wandersman (2005); evaluación de empoderamiento. Este modelo busca la mejora del programa bajo estudio por medio del desarrollo de las habilidades y autonomía de sus participantes. Este tipo de evaluación se desarrolla con la colaboración cercana entre miembros del personal ejecutor del programa y actores clave, quienes participarán del planteamiento de preguntas, la recolección y análisis de datos y planificación de pasos a seguir. Este tipo de evaluación requiere de la construcción de consensos entre todos los participantes, quienes se han de apropiar del proceso de evaluación.

Evaluación de Alianzas de Trabajo

Dickinson (2006) reconoce el carácter intersectorial del trabajo de sectores públicos, por lo cual sugiere un modelo de evaluación que se oriente a indagar sobre este aspecto en particular; las alianzas de trabajo (*health and social care partnerships*). El artículo señala que a pesar de la importancia reiterada que se le da a este trabajo colaborativo, aún no se ha logrado probar su aporte a la mejora de servicios para sus usuarios. Según el autor, la evaluación basada en la teoría ha enfocado esta pregunta con la intención de visibilizar los procesos de cambio que se permiten con el trabajo de alianzas estratégicas.

Consecuencias no Intencionadas

Toda intervención puede tener consecuencias no previstas inicialmente, tanto favorables como perjudiciales para el bienestar de los beneficiarios. Para David Grembowski (2001), es un requisito incluir en toda evaluación de programas un análisis de estas consecuencias no intencionadas. Al ser no intencionadas y por lo mismo, no estar previstas, es muy difícil llevar a cabo este análisis. Para esto Grembowski (2001) sugiere una serie de tareas que facilita su inclusión. Un evaluador/a debe:

- ✘ Comprender en profundidad el programa evaluado por medio de la revisión de documentos, conversaciones con directores, coordinadores, staff, participantes.
- ✘ Conversar con personas que creen que el programa va a fallar
- ✘ Conversar con personas que están familiarizadas con el programa pero no son responsables de su funcionamiento
- ✘ Hablar con otros evaluadores/as de programas
- ✘ Examinar los resultados de otros programas similares ya implementados

Meta-Evaluación

La meta evaluación es un proceso crucial de todo proceso de evaluación de programas, que consiste en el estudio de los mecanismos de evaluación y monitoreo del programa mismo. Así como la evaluación es una oportunidad de aprendizaje para la ejecución de programas de salud, también debe ser un proceso de aprendizaje para los medios por los cuales se estudian los procesos de cambio que ocurren en consecuencia. La OMS (2011) subraya la necesidad de este ejercicio de revisión para así constatar que los indicadores están midiendo lo que pretenden medir, que los datos son recolectados de acuerdo a estándares de calidad adecuados y que el análisis de datos y difusión de resultados entrega oportunamente la información requerida por tomadores de decisión.

A continuación se resumen los puntos centrales que deben estar presentes en un sistema de monitoreo y evaluación según la literatura revisada. A partir de estas consideraciones se pueden levantar preguntas de evaluación para la evaluación de Mitad de Período y otros planes y/o programas. Este resumen se ordena en una tabla para facilitar la revisión del material.

Tabla N°2: Meta-evaluación

Ítem	Descripción	Fuente
Pertinencia de Indicadores (<i>fit for purpose</i>)	<ul style="list-style-type: none"> ✘ Relevantes para las necesidades de diferentes usuarios de la información ✘ Sensibles al cambio en las variables bajo estudio 	<p>OMS, 2011</p> <p><i>Sugiere que este análisis se haga</i></p>

Calidad de información	<ul style="list-style-type: none"> ✘ Validez de indicadores (que midan lo que esperan medir) ✘ Confianza (consistencia en mediciones) ✘ Transparencia 	<i>cada dos años y que sus resultados sean de acceso público</i>
Identificación de necesidades de capacitación	<ul style="list-style-type: none"> ✘ Recolección y administración de información de registros clínicos (susceptibles a sesgos ante incentivos financieros) ✘ Recolección y almacenamiento de datos de diversas fuentes relevantes. Dar acceso público a esta información es clave para la transparencia. ✘ Monitoreo de calidad de datos del sector público por parte de instituciones académicas independientes. ✘ Síntesis de datos de múltiples fuentes para la elaboración de informes, planificación, evaluación y difusión. ✘ Estimación y modelos estadísticos: Aplicación de estándares internacionales de calidad, herramientas de corrección de sesgos y datos perdidos, cálculo de estimaciones y proyecciones. ✘ Presentación y difusión de datos para diferentes audiencias: Una buena difusión de resultados es esencial para que los resultados de procesos de evaluación puedan informar toma de decisiones futuras. ✘ Inclusión en estrategia de metas nacionales e internacionales. 	OMS, 2011
Estandarización de procesos	Debiera identificar con claridad indicadores y sus fuentes, los procesos de análisis y difusión de datos	OMS, 2009
Criterios de selección de indicadores	<ul style="list-style-type: none"> ✘ Reflejan todos los aspectos del desempeño de sistemas de salud y cubre todos los dominios de la cadena de resultados; recursos, procesos, productos, resultados esperados, impacto ✘ Se basan en listas de indicadores internacionales ✘ Son científicamente robustos, útiles, accesibles y comprensibles ✘ Cumplen con los requisitos de indicadores SMART; específicos (<i>specific</i>), medibles (<i>measurable</i>), alcanzable (<i>achievable</i>), relevante (<i>relevant</i>) y con indicación de plazos (<i>time-bound</i>) 	OMS, 2009

Evaluación de Plan de Salud

Existe una diferencia de alcance en las evaluaciones de programas y planes de salud. Dado que esta evaluación en particular – Evaluación de Mitad de Período de la Estrategia Nacional de Salud – es una evaluación a nivel de plan de salud, y no así de programas en particular, a continuación se señalan los apuntes de la literatura específicos sobre este nivel de estudio.

Tabla N°3: Componentes Esenciales en el Monitoreo y Evaluación de Estrategias Nacionales de Salud

Ítem	Descripción	Fuente
Tipos de informes en evaluación y monitores de Estrategias Nacionales de Salud	<ul style="list-style-type: none"> ✘ Evaluación Anual: Focalizada en indicadores específicos de la planificación anual operativa. Consiste en indicadores de recursos, procesos y productos. De estar disponibles también se pueden incluir datos sobre cobertura. Estos informes deben contribuir a la evaluación de modo sistemático. Cada país puede tener necesidades específicas de profundizar en algún área particular en sus informes anuales. ✘ Evaluación de mitad de período: Cubre todas las metas mencionadas en la estrategia, incluyendo metas de resultados esperados e impacto, tomando en consideración cambios contextuales. Esta evaluación debiera coincidir con el informe anual recién mencionado. ✘ Evaluación de final de período: Incluye un análisis integral de proceso y desempeño para el período completo de la estrategia nacional de salud. Se construye en base a los informes anuales, la evaluación de mitad de período, incluyendo también resultados de evaluación prospectiva (formativa) para construir la estrategia siguiente. 	OMS, 2011
Metodología de investigación	El análisis del desempeño de servicios de salud y el progreso de indicadores de salud debe complementar dimensiones cuantitativas y cualitativas.	OMS, 2011

Elementos a incluir en un sistema de monitoreo y evaluación	<ul style="list-style-type: none"> ✘ Progreso hacia las metas de la estrategia nacional de salud ✘ Equidad ✘ Eficiencia ✘ Análisis cualitativo de cambios contextuales ✘ Comparación (<i>Benchmarking</i>) 	OMS, 2011
Local y global	El monitoreo y evaluación de una estrategia nacional de salud debe estar enfocado principalmente en prioridades locales y a la vez entregar una base para monitoreo global	OMS, 2009
Diversidad de audiencias	El sistema de monitoreo y evaluación debe acoger las necesidades de una diversidad de actores para múltiples propósitos, incluyendo el monitoreo de recursos de programa, procesos y resultados	OMS, 2009
Áreas a evaluar	<ul style="list-style-type: none"> ✘ Indicadores de desempeño en términos de sus sistemas de salud – disponibilidad, acceso, calidad, eficiencia ✘ Indicadores de salud poblacional – estado de salud, satisfacción usuaria, protección financiera. <p>Estos son elementos a capturar por el sistema global de monitoreo y evaluación, no por cada uno de los tipos de evaluación (anual, mitad de período, cierre de período).</p>	OMS, 2009
Elementos a incluir en un sistema de monitoreo y evaluación	<ul style="list-style-type: none"> ✘ Variables contextuales ✘ Indicadores de; Recursos y productos del sistema de salud; Cobertura; Factores de riesgo; Estado de salud ✘ Evaluación periódica de la calidad de sus datos ✘ Información cualitativa para responder preguntas de evaluación 	Victoria & cols., 2010

<p>Características de una evaluación de mitad de período</p>	<ul style="list-style-type: none"> ✘ De mayor amplitud en el alcance de sus mediciones y su proceso en comparación con revisiones anuales. Pueden reemplazar o complementar revisiones anuales. ✘ Integran la evidencia relevante de años anteriores para evaluar el progreso y desempeño del sector salud en general. ✘ Debiera analizar la calidad de los datos, patrones de largo plazo, cambios contextuales y consideraciones de equidad. ✘ Debiera comparar resultados de nivel nacional con otros países. ✘ Debiera ser también una base para el análisis situacional del próximo plan nacional. ✘ Debiera ser ejecutada por instituciones locales de salud pública e instituciones académicas en conjunto con aliados internacionales. 	<p>OMS, 2015 <i>“Evaluación de mitad de período de la estrategia nacional de salud: un ejemplo de la República Unida de Tanzania”</i></p>
<p>Resultados intermedios a evaluar en sistemas de salud</p>	<ul style="list-style-type: none"> ✘ Efectividad ✘ Eficiencia ✘ Equidad 	<p>Grembowski, 2001</p>

Equipo de Evaluación

La Organización Mundial de la Salud (2011), señala que dentro de un proceso de evaluación es importante incluir el aporte de instituciones que son independientes de la implementación de programas, de modo de aumentar su objetividad. Instituciones académicas y de salud pública estarían en condiciones de asumir este rol. Otros agentes a involucrar son funcionarios del Ministerio de Salud representantes de departamentos relevantes, de equipos regionales y otros ministerios que tienen injerencia en temas de salud. En el 2015, la Organización Mundial de la Salud por medio de su experiencia de evaluación de mitad de período de la estrategia nacional de Tanzania, reitera la importancia de involucrar a los diversos actores interesados en la evaluación, para lo cual se destinaron meses de trabajo para elaborar un diseño que contemplaba las diferentes perspectivas de los involucrados. Los integrantes del equipo, reitera la OMS (2015), deben ser miembros del



ministerio de salud, instituciones académicas independientes y agencias internacionales. En esta experiencia, el período de diseño de evaluación tardó más que el período de recolección de datos, para lo cual los miembros del equipo de evaluación distribuyeron tareas y fuentes de información.

Contandriopoulos & Brousselle (2012) coinciden con lo anterior, señalando que para formar un equipo de evaluación es necesario identificar a quienes harán uso de la información que proveerá el estudio, para entonces incluirlos en el equipo de consulta para incluir adecuadamente sus intereses. Habicht & cols (1999), también refuerzan la necesidad de procesos participativos que involucren a todos los actores relevantes de los programas bajo evaluación, de modo que sus resultados sean relevantes para ser utilizados en la mejora de los mismos.

REFERENCIAS

- Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation peas in a pod or apples and oranges?. *Evaluation*, 13(4), 439-455.
- Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77.
- Davies, R. (2013) Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations. 40 Working Paper, DFID.https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf
- Dickinson, H. (2006). The evaluation of health and social care partnerships: an analysis of approaches and synthesis for the future. *Health & social care in the community*, 14(5), 375-383.
- Grembowski, D. (2001). *The practice of health program evaluation*. Sage Publications.
- Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International journal of epidemiology*, 28(1), 10-18.
- Marchal, B., van Belle, S., van Olmen, J., Hoérée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), 192-212.
- Organización Mundial de la Salud (2011). *Monitoring, evaluation and review of national health strategies*.
- Organización Mundial de la Salud (2013). *Handbook on health inequality monitoring: with a special focus on low- and middle-income countries*.
- Organización Mundial de la Salud (2015). *Midterm review of national health plans: an example from the United Republic of Tanzania*.
- Pawson, R., and Tilley, N. 1997. *Realistic Evaluation*. London: Sage
- Peters, David H.; El-Saharty, Sameh; Siadat, Banafsheh; Janovsky, Katja; Vujicic, Marko. (2009). *Improving Health Service Delivery in Developing Countries : From Evidence to Action*. Washington, DC: World Bank. © World Bank. <https://openknowledge.worldbank.org/handle/10986/12335> License: CC BY 3.0 IGO.



Sridharan, S., & Nakaima, A. (2011). Ten steps to making evaluation matter. *Evaluation and program planning*, 34(2), 135-146.

Tanahashi, T. (1978). Health service coverage and its evaluation. *Bulletin of the World Health Organization*, 56(2), 295.

Weiss., C. (1998). *Evaluation*. New Jersey: Prentice Hall.