

5. Evaluación de la implementación

Javier Gracia, Carola Orrego, Juan Antonio Blasco

En este capítulo se explica la necesidad de evaluar la implementación de las GPC y se exponen algunas claves para dicha evaluación.

Preguntas que pretende responder el capítulo:

- ¿Ha funcionado la implementación? Evaluación de estructura, proceso y resultado: Utilización de indicadores
- ¿Cuáles son los indicadores clave a medir para evaluar la implementación de una GPC?
- ¿Existen instrumentos para la medición del proceso de implementación?
- ¿Con qué diseños podemos planificar una evaluación de resultados de implementación de GPC?

Presentamos aquí una revisión breve de la evaluación de la implementación de las GPC. Esta evaluación es necesaria para todos los grupos de interés de la GPC (grupo elaborador, implementadores, profesionales y pacientes) porque permite conocer si las estrategias puestas en marcha para implementar la GPC han tenido éxito.

Se debe identificar desde el principio, en el plan de implementación, **qué** es lo que **debe ser evaluado, qué información debe recogerse, las fuentes de información y los instrumentos** para recogerla, así como el **diseño** más apropiado para la realización de esta evaluación.

En la fase de elaboración de la GPC, un elemento importante a considerar es la factibilidad de las recomendaciones y los indicadores clave que nos ayudarán a monitorizarlas para evaluar su implementación. De esta forma se dispone de un punto de partida común en el que el grupo implementador puede basarse para iniciar la evaluación, con posibilidad de evaluar aquello que queremos medir, en lugar de evaluar únicamente lo que podemos medir, “dadas las circunstancias”. Por otro lado, para apoyar este proceso es útil contar con profesionales familiarizados con la metodología de investigación y evaluación que nos ayuden en el diseño.

Las evaluaciones a nivel local son importantes para determinar las próximas fases en los esfuerzos por implementar las GPC, creándose un proceso cíclico. En dichas evaluaciones es también necesario considerar si queremos que los resultados se mantengan a lo largo del tiempo. Esto se relaciona con la fase final de Graham *et al.*⁽¹⁾: “manteniendo el uso del conocimiento en marcha”, y señala la necesidad de crear una cultura organizacional que alimente la adherencia a las GPC⁽²⁾.

5.1. ¿Ha funcionado la implementación? Evaluación de estructura, proceso y resultado. Utilización de indicadores

5.1.1. ¿Qué debe ser evaluado?

La implementación de una GPC puede afectar tanto al proceso asistencial como a los resultados alcanzados con dicho proceso. Cuando se está diseñando la evaluación, debe decidirse qué componentes se quiere medir. Hakkennes y Green⁽³⁾ revisaron la efectividad de intervenciones basadas en GPC, para identificar los tipos de resultados presentados y los métodos utilizados para medir los cambios logrados. Los resultados encontrados fueron agrupados en cinco categorías: dos a nivel de paciente, dos a nivel de los profesionales de salud y una a nivel de la organización o del proceso. Estas categorías se presentan en la tabla 5.1.

Tabla 5.1. Categorización de los resultados a evaluar⁽³⁾

Nivel de paciente

- *Medidas de cambios en la salud de los pacientes:* dolor, depresión, calidad de vida, HbA1c, etc.
- *Medidas subrogadas de cambio en la salud de los pacientes:* actitudes del paciente, tiempo de estancia en el hospital, etc.

Nivel de profesional de la salud

- *Medidas de cambio en la práctica:* confianza en las GPC, cambios en los porcentajes de prescripción, etc.
- *Medidas subrogadas de cambio en la práctica:* conocimiento y actitudes del profesional, etc.

Nivel de organización o de proceso

- *Medidas de cambio en el sistema sanitario* (lista de espera): cambios en políticas, costes y utilidad o extensión de la indicación de determinadas intervenciones.

Investigaciones posteriores en la medición de resultados en relación con la implementación de GPC nos permitirán entender qué tipo de resultados deben ser evaluados en función de las circunstancias particulares. Por ejemplo, aunque el objetivo último de una GPC sea mejorar los resultados en salud (nivel de paciente), cuando una GPC aborda un comportamiento clínico sobre el cual hay fuerte evidencia de que dicho comportamiento se correlaciona con el o los resultados de salud que interesan, puede ser suficiente con observar si el cambio en la práctica diaria del profesional ocurre. En cualquier caso, se están desarrollando cada vez más, en función de áreas específicas de atención sanitaria, medidas de resultado estandarizadas de utilidad global.

Siempre se ha de considerar una revisión de las medidas que puedan ser utilizadas en nuestra evaluación; existen organizaciones internacionales que se han dedicado a la identificación y creación de repositorios de indicadores. Más adelante se comentan algunos de ellos.

5.1.2. ¿Qué información debe recogerse?

Utilización de indicadores

Al ser las GPC parte integral de la mejora de la calidad asistencial, la medición y evaluación de su implementación debe reflejar uno o más de los dominios que definen una *asistencia sanitaria de calidad*, entendiendo como tal una atención “segura, efectiva, centrada en el paciente, a tiempo, eficiente y equitativa”⁽⁴⁾. Los indicadores son elementos que nos van a permitir la medida de estos dominios.

Un indicador es un elemento medible del desarrollo de la práctica para el cual hay evidencia o consenso de que puede ser utilizado para evaluar la calidad de la atención prestada⁽⁵⁾.

Los elementos clave a evaluar para determinar el resultado de la implementación de una GPC tienen que ver con la evaluación de la estructura, el proceso y los resultados de su utilización mediante estos indicadores.

Respecto a la evaluación de la **estructura**, la cuestión clave se refiere a si se han dispuesto los recursos humanos y físicos requeridos o necesarios para implementar las recomendaciones de la GPC. Esto tiene que ver no sólo con los equipos, instrumentos y profesionales requeridos; también con la estructura de la organización, cambios organizativos y cualificación o formación del personal sanitario. Todo ello está relacionado con la forma en que los recursos han sido distribuidos y la manera en que se ha organizado la aplicación de las recomendaciones e intervenciones que se derivan de una GPC.

La evaluación del **proceso** nos permite conocer cómo está funcionando el plan de implementación, y se centra en saber qué se hace y quién lo hace. Se trata de información fundamental para determinar cómo ha resultado el plan de implementación en cuanto a la mejora de la aplicación de las recomendaciones, y cómo puede mejorarse o qué nuevas estrategias de implementación se pueden plantear.

La evaluación del **resultado** está dirigida a identificar los cambios que han ocurrido en los pacientes como resultado de la aplicación de la GPC, comparándolos con los resultados que se esperaban. Idealmente los referidos al efecto de la adherencia a las recomendaciones en términos de resultados en salud medibles, sobre todo a corto plazo.

El objetivo de los indicadores clave en este caso es determinar si los beneficios que se derivan de la implementación han merecido la pena respecto a los costes y recursos utilizados, y en cuanto a mejoras en los resultados en salud⁽⁶⁾.

A la hora de poner en marcha la evaluación de la implementación de la GPC se debe considerar, por un lado, la necesidad y las posibilidades de recogida de datos; por otro, las herramientas que nos puedan facilitar la ejecución de la evaluación, tal y como veremos más adelante. Es evidente que la utilización de datos que se recogen de forma rutinaria dentro de la práctica sanitaria facilitará todo el proceso de evaluación y minimizará los costes.

5.2. Indicadores clave para evaluar la implementación de una GPC⁽⁷⁾

5.2.1. Medición de los indicadores clave de estructura y de proceso

Los resultados en salud son una pieza importante en el proceso de evaluación pero no la única. Hay factores estructurales, como los recursos humanos y materiales, y del proceso, como los conocimientos y las habilidades, que pueden ser importantes. Por ejemplo, en la evaluación de una GPC para el manejo del dolor que depende del control analgésico del paciente, puede requerirse una evaluación del número de bombas de infusión disponibles (estructura) y de la formación de las enfermeras para incrementar su conocimiento acerca del manejo del dolor (proceso). De hecho, los datos recogidos para la evaluación de estructura y proceso permiten explicar resultados que se pueden considerar negativos tras la aplicación de la GPC. Siguiendo con el ejemplo anterior, si no se ha conseguido reducir los niveles de dolor, los datos estructurales pueden indicar que no había suficientes bombas en las unidades, que no se llevaron a cabo los programas educativos necesarios, o que fueron inefectivos a la hora de proveer conocimientos sobre el manejo del dolor para enfermería.

5.2.2. Medición de los indicadores clave de resultado

Los resultados clave suelen ser fáciles de identificar respecto a las recomendaciones establecidas por la GPC. Los indicadores deben ser un número reducido (dos o tres) y deben ir dirigidos a la valoración clara de los beneficios principales derivados de la implementación de la GPC. Es recomendable utilizar criterios de selección de indicadores relacionados con los datos que ya hayan sido utilizados y probados anteriormente o en estudios similares. Preferiblemente, deben ser indicadores estandarizados, relacionados con la intervención y con los resultados en salud mediante los que se pretende evaluar el impacto de la implementación de las recomendaciones de la GPC.

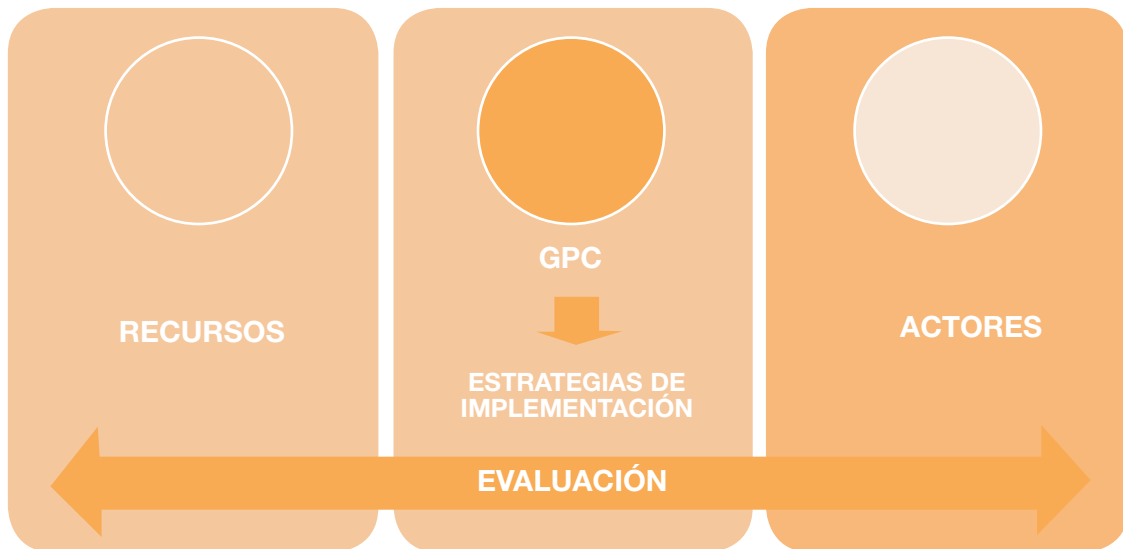
5.2.3. Selección de los indicadores clave

Los indicadores clave deberían surgir de los objetivos principales que se persiguen al implementar la GPC, considerando la población a la que va dirigida así como los recursos requeridos.

Se deben tener en cuenta los temas relacionados con la recolección de datos e información necesaria para el cálculo de los indicadores propuestos para cada elemento, así como las dificultades que puede plantear la obtención de la información. Entre los participantes (actores) que pueden estar implicados en este proceso está la propia administración sanitaria, gestores y grupos de profesionales de la salud.

La figura 5.1 refleja cómo a la hora de seleccionar los indicadores clave se deben considerar las estrategias de implementación propuestas y pensar en la evaluación que se va a llevar a cabo en función de los objetivos de la GPC, sin olvidar los recursos y actores que serán necesarios para la obtención de la información necesaria para la evaluación.

Figura 5.1. Aspectos a tener en cuenta en la selección de indicadores



5.2.4. Categorización de indicadores clave

De forma general podemos decir que hay cuatro **categorías** fundamentales dentro de los potenciales indicadores clave para la evaluación de la implementación de una GPC: **la organización, el proveedor, los pacientes/familiares y los costes.**

La tabla 5.2, que figura a continuación, se puede utilizar como guía para identificar indicadores de estructura, de proceso y de resultado; y contiene algunos ejemplos de los mismos para las categorías citadas.

Tabla 5.2. Guía para la identificación de indicadores clave

Indicadores para evaluación de la implementación de GPC			
Categoría	Estructura (lo que se necesita)	Proceso (cómo hacemos)	Resultado (qué ocurre)
Organización / Unidad	<ul style="list-style-type: none"> Estabilidad organizacional Cultura y soporte para el cambio Mecanismos de aseguramiento calidad Procedimientos Espacios físicos Equipamiento 	<ul style="list-style-type: none"> Desarrollo/modificación de políticas y procedimientos Diagramas 	<ul style="list-style-type: none"> Logro de objetivos para la mejora de resultados en el paciente Logro de objetivos para condiciones específicas
Proveedor	<ul style="list-style-type: none"> Número y cualificación del personal Ratio personal/paciente Roles Programas educativos 	<ul style="list-style-type: none"> Actitudes hacia las GPC Conocimientos/ niveles de habilidades 	<ul style="list-style-type: none"> Asistencia a programas educativos Adherencia a GPC Número de evaluaciones realizadas Número de tratamientos adecuados Satisfacción del profesional
Paciente/ Familiar	<ul style="list-style-type: none"> Características del paciente (demográficas/nivel de riesgo) Implicación en toma de decisiones 	<ul style="list-style-type: none"> Actitudes hacia las GPC Aceptación de la familia/comunidad Conocimiento de la familia/comunidad 	<ul style="list-style-type: none"> Resultados físicos, psicológicos, sociales Salud de la familia Satisfacción con la asistencia recibida Acceso
Costes	<ul style="list-style-type: none"> Costes de personal y recursos adicionales requeridos Nuevo equipamiento 	<ul style="list-style-type: none"> Costes de las estrategias de implementación Educación del personal/pacientes 	<ul style="list-style-type: none"> Costes incrementales de innovación, incluidos productos y fármacos Crecimiento del Servicio Estancia Número de pruebas diagnósticas, intervenciones Visitas a urgencias, % de reingresos

Adaptado de: Registered Nurses Association of Ontario. Toolkit: Implementation of clinical practice guidelines. Toronto, Canada: Registered Nurses Association of Ontario; 2002⁽⁷⁾.

En los anexos 4 y 5 de este manual se presentan dos modelos de ayuda para el diseño e identificación de indicadores.

5.2.5. Algunas fuentes de indicadores clave

Existen indicadores propuestos por la Agencia estadounidense para la investigación en calidad y cuidados de la salud, AHRQ (*Agency for Healthcare Research and Quality*), que resultan muy útiles en la evaluación (www.qualitymeasures.ahrq.gov).

Otro recurso disponible e interesante es el sistema de monitorización *Oryx*, iniciativa de la Comisión conjunta para la acreditación de organizaciones sanitarias (*Joint Commission on Accreditation of Healthcare Organizations*) (<http://www.jointcommission.org/>). Este sistema, basado en datos de la actuación, incluye indicadores básicos de proceso y resultado de la práctica clínica que pueden facilitar la evaluación de la implementación. En la selección de indicadores han colaborado asociaciones profesionales internacionales, proveedores de servicios de salud, asociaciones profesionales, asociaciones de consumidores y expertos en metodología de la calidad.

Una organización destacada en la elaboración de indicadores en guías de práctica clínica es el Instituto estadounidense para la mejora de los sistemas clínicos, ICSI (*Institute for Clinical Systems Improvement*) (<http://www.icsi.org>), que propone una batería de indicadores clave vinculados a unos objetivos priorizados para cada problema clínico asociada a cada guía de práctica clínica.

Otras instituciones en nuestro contexto, como la Fundación Avedis Donabedian (FAD) (<http://www.fadq.org/>), tienen experiencia en el diseño y definición de indicadores clave para la evaluación de la calidad asistencial que nos pueden ser útiles en la evaluación. Uno de los modelos propuestos por la FAD para definir indicadores se presenta en el anexo 5.

5.3. ¿Existen instrumentos para la medición del plan de implementación? Técnicas de recogida de la información

La evaluación de la implementación se basa, como hemos visto, en la evaluación de los resultados obtenidos tras la implementación de las GPC. Otro aspecto que debe tenerse en cuenta son las técnicas de recogida y las fuentes de información que se utilicen para evaluar el impacto de la implementación.

Existe una gran variedad de técnicas de recogida de información para medir los resultados obtenidos con una GPC. Entre ellos, se han descrito los siguientes⁽³⁾:

- Auditoría de registros médicos, revisando historias clínicas, informes o mediante registros electrónicos.
- Cuestionarios o entrevistas a profesionales.
- Cuestionarios o entrevistas a pacientes.
- Cuestionario de recogida de datos ad hoc.

Además, es posible recurrir a las siguientes fuentes de datos:

- Bases de datos (por ejemplo, datos de facturación médica, CMBD).
- Registros de un servicio (por ejemplo, registros de la frecuentación al servicio de urgencias).
- Registros clínicos (por ejemplo, resultados de pruebas de laboratorio, exploraciones clínicas).

Jennet *et al.*⁽⁸⁾ hacen una descripción útil de algunos de estos métodos. Cuando haya que decidir la técnica de recogida o la fuente de datos, se debe verificar si las medidas permitirán obtener la información deseada.

5.3.1. Consideraciones prácticas en la recogida

Se debe tener en cuenta la necesidad de recurrir a un comité de ética en función del tipo de información a recoger, el diseño del estudio, el número de centros participantes y los objetivos de la evaluación.

En general, la información clínica recogida y los datos utilizados para la determinación de los indicadores mediante los que se evalúa la implementación de las GPC deben estar codificados para garantizar la confidencialidad.

Respecto a la recogida de información para la determinación de los indicadores relacionados con medidas de percepción o conocimiento en pacientes o en los profesionales sanitarios, es recomendable disponer de tasas altas de respuesta. Idealmente se debe alcanzar un 75%. Cuanta menor sea la tasa de respuesta, más difícil resultará conocer la representatividad o el sentido de los resultados. Las estrategias para mejorar las tasas de respuesta podrían estar relacionadas con los recursos e incentivos disponibles para la evaluación. En cualquier caso, los cuestionarios o datos a recoger deben ser lo más simples y breves posible para mejorar no sólo la tasa de respuesta, sino la calidad de la información y, por tanto, la utilidad y representatividad del indicador. Durante la recogida y el cálculo de datos para determinar los indicadores es necesario identificar cualquier factor que pueda suponer un sesgo y tratar de tenerlo en cuenta como un factor potencial que puede condicionar los resultados.

5.4. ¿Con qué diseños podemos planear una evaluación de la implementación de GPC?

Para tomar decisiones basadas en la evidencia acerca de la implementación de guías, los responsables de implementarlas deben estar informados sobre la efectividad y la relación coste efectividad de las distintas intervenciones, sobre los posibles modificadores del efecto esperado y sobre los recursos necesarios para realizar la intervención, cuestiones a tener en cuenta también en la evaluación.

Existe poca literatura en relación al diseño, conducción y análisis de la evaluación de intervenciones tan complejas como las que implican la implementación de guías de práctica clínica⁽⁹⁾.

Según los objetivos que se persigan, se pueden utilizar diferentes métodos y diseños para evaluar el efecto de las actividades o intervenciones de implementación. Se debe decidir no sólo qué es factible, sino también qué grado de detalle se requiere. Por ejemplo, una evaluación puede llevarse a cabo para conocer el efecto de una intervención en un proyecto en particular, en un determinado contexto o, a la inversa, podría realizarse para informar sobre cierto conocimiento científico en cuanto a la efectividad de una intervención. Este último abordaje requerirá un estudio con un diseño más complejo⁽¹⁰⁾ que garantice la validez interna del estudio, de tal manera que esa intervención pueda resultar útil en otros entornos.

Por lo tanto, los diferentes diseños de evaluación dependen de si queremos probar el progreso en la implementación de las GPC y cambios en el proceso de atención en un/los centro/s: mejora de la calidad; o si queremos proveer información o conocimiento sobre el valor de una nueva estrategia o intervención para la implementación de una GPC: investigación sobre la efectividad de las intervenciones.

En el primer caso, el principal objetivo es establecer si los cambios deseados han sido alcanzados; en el segundo, lo importante es que los efectos puedan ser atribuidos fiablemente a la intervención. Para el segundo caso podemos decir que para atribuir los efectos de un estudio a una intervención que tiene como objetivo conseguir la aplicación de la evidencia de una GPC, se necesitan evaluaciones que obtengan estimaciones no sesgadas de sus efectos. Estas estimaciones son particularmente necesarias para informar las decisiones políticas mayores y a largo plazo⁽¹¹⁾.

Sin embargo, los recursos para la realización de estas evaluaciones son limitados y en algunas ocasiones puede resultar imposible su realización.

Muchas decisiones relacionadas con estrategias de implementación se refieren a proyectos *mas prácticos*, programas o proyectos pequeños y concretos, tales como programas de mejora de la calidad en un centro específico o en un servicio clínico en concreto⁽⁴⁾, pero también pueden desarrollarse a gran escala, por ejemplo en todos los hospitales o centros de Atención Primaria de una región⁽¹¹⁾.

Existe un amplio rango de diseños de estudio que podrían ser utilizados para la evaluación de la implementación de GPC, y la elección de uno u otro dependerá del objetivo de la evaluación y del grado de control que los investigadores quieran tener sobre la intervención.

Los diseños de estudio utilizados se clasifican usualmente en función de este control sobre la intervención⁽⁶⁾: experimentales (estudios de intervención) y observacionales⁽¹²⁾. A continuación se presentan estos diseños.

5.4.1. Estudios observacionales

Estos estudios, en el contexto de la implementación de GPC, permiten ayudar a esclarecer el proceso de modificación del comportamiento de los profesionales⁽¹³⁾.

Por definición, los estudios observacionales examinan la variación natural entre los participantes de un estudio para explorar los efectos de las intervenciones en los resultados y en los factores asociados⁽¹⁴⁾.

En los estudios observacionales se pueden distinguir dos grupos de estudios, las evaluaciones descriptivas y las evaluaciones comparativas⁽¹⁴⁾:

- **Evaluaciones descriptivas:** El principal objetivo es monitorizar y documentar el cambio en indicadores relevantes para la práctica en entornos específicos. Los estudios descriptivos usualmente incluyen a todos los proveedores sanitarios del proyecto y la medición es relativamente simple. Estos estudios no pueden determinar relaciones causa-efecto y su generalización es a menudo limitada; su valor radica en que permiten comprobar progresos en objetivos planteados para mejorar la adherencia a las recomendaciones clínicas, o los resultados clínicos para un determinado tipo de pacientes, y en que proporcionan información a nivel local de los resultados de programas o proyectos a pequeña escala. Dentro de este tipo de diseño encontramos los estudios de desarrollo y los estudios de casos descriptivos⁽¹⁴⁾.
- **Evaluaciones comparativas:** Permiten monitorizar y documentar el impacto de las estrategias de implementación y utilizar comparaciones entre diferentes sujetos o entornos para identificar potenciales determinantes para el cambio. En muchos casos, una muestra de proveedores sanitarios es suficiente para resultados generalizables. La medición tiende a ser más compleja porque en ella influyen un amplio rango de determinantes. Los estudios de *revisión y monitorización* son los estudios observacionales de utilización más frecuente en la evaluación de implementación de GPC. Son relativamente rápidos y de bajo coste, y pueden ayudar a profesionales sanitarios y gestores en la evaluación del grado de implementación de estrategias con evidencia probada⁽¹⁵⁾. Implican la evaluación o monitorización de la provisión de servicios sanitarios, para garantizar que la calidad de los mismos se corresponda con lo sugerido por la evidencia científica⁽¹⁶⁾. Habitualmente están focalizados en uno o más cambios específicos, que son diseñados para un tema concreto (por ejemplo, asegurar la indicación de aspirina y betabloqueantes al alta de un paciente con infarto). Este tipo de evaluaciones incluye un ciclo de revisiones que permite la monitorización de los cambios. Otros tipos de estudios comparativos, son *estudios de casos comparativos y evaluaciones de impacto*⁽¹⁴⁾.

5.4.2. Estudios de intervención

- **Diseños aleatorizados (o estudios experimentales)**
 - *Ensayos aleatorizados:* Los ensayos aleatorizados son el método *Gold Standard* para la evaluación de intervenciones sanitarias, los más estrictos por su diseño y control de sesgos. Estos ensayos estiman el impacto de una intervención a través de la comparación directa con una asignación aleatoria de un grupo control que no recibe intervención o recibe una intervención alternativa^(11,17). Este tipo de estudio requiere un amplio dominio de esta metodología así como, en general,

numerosos recursos^(11,17). La implementación de una GPC es una intervención compleja que presenta particulares dificultades metodológicas también en su evaluación, ya que suelen abordar múltiples aspectos y son altamente dependientes del contexto social^(11,18,19). Este tipo de estudios de investigación son más frecuentemente llevados a cabo por investigadores con fondos para investigación. El grupo EPOC de la colaboración Cochrane (*Cochrane Effective Practice and Organisation of Care Group*) definió una serie de criterios que tienen que tomarse en cuenta cuando se realizan revisiones sistemáticas orientadas a evaluar la efectividad de las intervenciones para generar el cambio de los profesionales sanitarios⁽²⁰⁾. Estos criterios pueden considerarse como una útil “lista de comprobación” de la calidad de los estudios y por tanto establecen unas condiciones mínimas de calidad para cada tipo de diseño. En el caso de los ensayos aleatorizados, consideran criterios de inclusión los siguientes:

- Diseño en que los individuos (u otras unidades) fueron *definitivamente* asignados de forma prospectiva a una, dos o más formas alternativas de atención sanitaria, utilizando un proceso de asignación aleatoria (por ejemplo tablas de números aleatorios, lanzamiento de una moneda, etc.)⁽²⁰⁾
- *Ensayos controlados*: Se trata de ensayos donde el método de asignación aleatoria de pacientes a cada grupo tiene alguna limitación. Según el grupo EPOC⁽²⁰⁾, puede corresponder a un ensayo en el cual los participantes (u otras unidades) fueron:
 - *definitivamente* asignados de forma prospectiva a una, dos o más formas alternativas de atención sanitaria, utilizando un método *cuasi aleatorio* de asignación (por ejemplo, por fecha de nacimiento o por número de historia).
 - *posiblemente* asignados de forma prospectiva a una o más formas alternativas de atención sanitaria, utilizando un método aleatorio o cuasi aleatorio.

Como ejemplo, se presenta un estudio de Varonen *et al.*⁽²¹⁾, que llevaron a cabo un ensayo multicéntrico controlado y aleatorizado en centros sanitarios, para evaluar el impacto del programa de implementación de una GPC en el manejo de la sinusitis maxilar aguda en el contexto de Atención Primaria. Los centros fueron aleatoriamente seleccionados para implementar la GPC, bien de acuerdo a una metodología de aprendizaje basado en problemas (trabajo en grupos, facilitado por un tutor local, basado en escenarios clínicos), bien por un método de visitas educativas facilitadas por los médicos de familia locales; había también controles externos. Los resultados mostraron que la implementación de la GPC contribuyó a cambios menores en las prácticas recomendadas y que no había diferencias significativas entre los dos métodos de implementación. Menos de la mitad de los centros participantes fueron capaces de abordar el proyecto como se dijo, disminuyendo por tanto la validez interna del estudio. Los autores sugirieron que aproximaciones más centradas en los problemas y las prácticas de cada centro de salud podrían haber mejorado los resultados.

- **Diseños no aleatorizados (o diseños cuasi experimentales)**

Los diseños cuasi experimentales son útiles cuando existen barreras políticas, prácticas o éticas para conducir un ensayo aleatorizado. Bajo estas circunstancias los investiga-

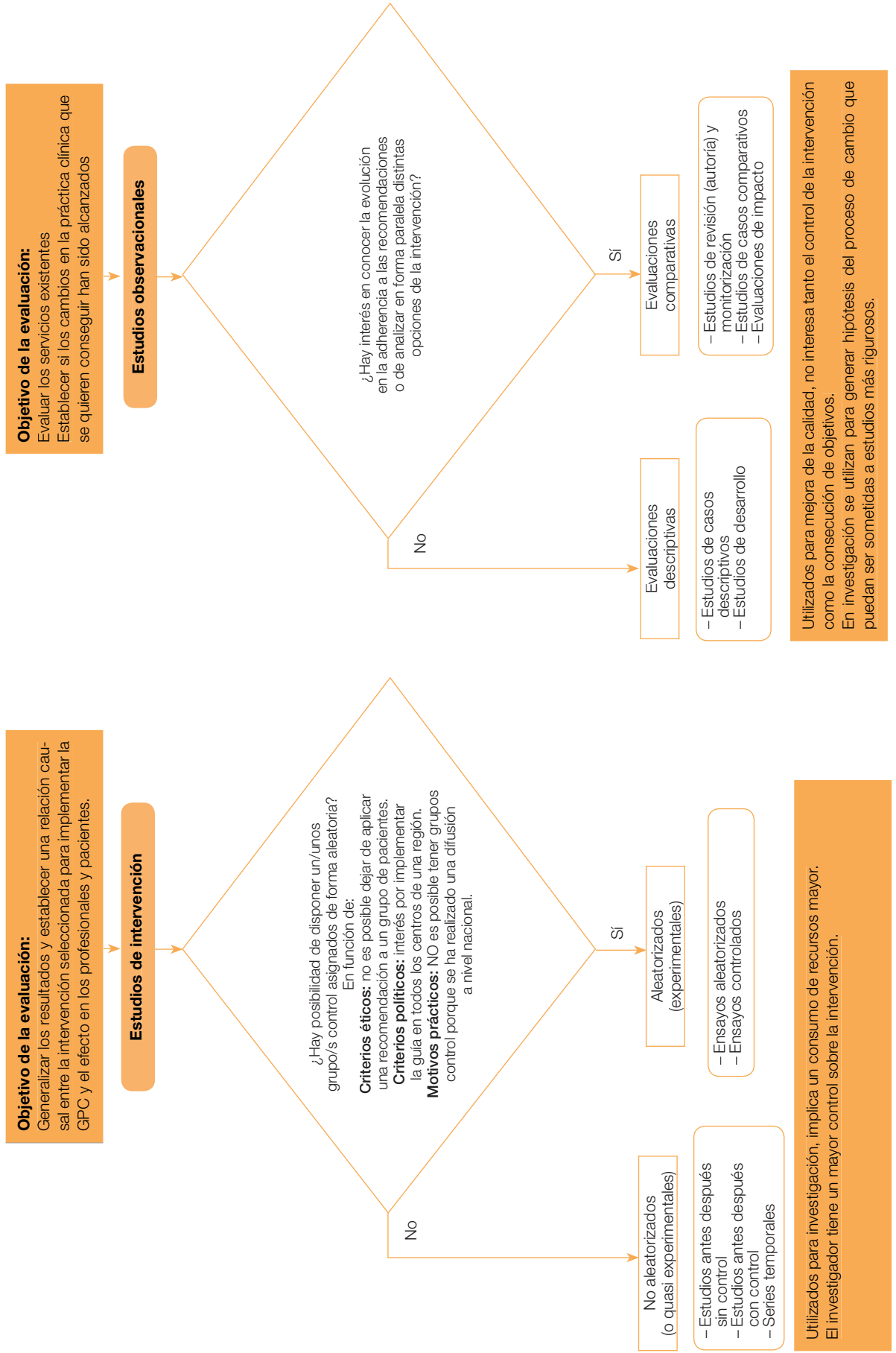
dores tienen menor control sobre la asignación de la intervención y tienen que realizar una evaluación ad hoc para la intervención propuesta⁽¹⁷⁾. Existen diferentes tipos de diseños de estudios cuasiexperimentales^(12,17), pero los más utilizados para la evaluación en estudios de implementación son los siguientes:

- *Estudios no controlados antes y después.* También llamados *preexperimentales*⁽¹²⁾. Estos estudios miden los resultados de la actuación antes y después de la introducción de una intervención⁽¹⁷⁾. Son relativamente simples de conducir y, en términos de causalidad, son superiores a los estudios observacionales, pero tienen debilidades intrínsecas en el diseño, puesto que tendencias temporales o cambios repentinos pueden hacer difícilmente atribuibles los cambios a la intervención. Los cambios en los resultados pueden ser causados por maduración de los sujetos de estudio o por otros factores externos. Este tipo de diseños se excluye de las revisiones sistemáticas coordinadas por el grupo EPOC⁽²²⁾. Se presenta aquí como ejemplo el estudio de Muller *et al.*⁽²³⁾, estudio prospectivo antes y después para determinar la efectividad de un diagrama de flujo para reducir el uso de transfusiones de sangre en pacientes sometidos a cirugía de recambio de cadera y rodilla. El diagrama, desarrollado por médicos y personal de enfermería del hospital, y aceptado por los líderes médicos locales, fue ampliamente distribuido y presentado a médicos y personal de enfermería en sesiones de entrenamiento en grupos pequeños en las que se enfatizaba el sentido de *propiedad* y responsabilidad local del proyecto, y se incluían también flujos de pacientes. Siguiendo esta intervención, el porcentaje de pacientes que recibía sangre después de un recambio total de articulación cayó en más de un 40%, con una reducción concomitante en los costes. Los autores atribuyen la efectividad del diagrama de flujo a su simplicidad, a su amplia distribución, a que no requería grandes cambios, al apoyo de los líderes locales y al sentido de *propiedad*.
- *Estudios controlados antes y después (o diseño de grupos no equivalentes)*⁽¹²⁾. En los estudios controlados antes y después, los investigadores intentan identificar una población control de características similares y realizar la intervención en la población de estudio, recogiendo datos en ambos grupos antes y después de que la intervención se aplique a la población de estudio. Mientras que los estudios bien diseñados antes y después deberían proteger de las tendencias temporales y los cambios repentinos, a veces es difícil identificar grupos control comparables. Incluso, en condiciones aparentes de grupos bien emparejados, la actuación basal a menudo difiere. Otro problema analítico común en este tipo de diseño se da cuando la intervención se realiza a nivel de organización (datos agrupados) y, erróneamente, el investigador los analiza a nivel individual (paciente).
- *Series temporales.* Los diseños de series temporales intentan detectar si una intervención ha tenido un efecto significativamente más grande que las tendencias temporales subyacentes. Son diseños útiles para evaluar los efectos de las intervenciones cuando es difícil aleatorizar o identificar un grupo control apropiado; por ejemplo, después de la diseminación de unas GPC a nivel nacional⁽¹⁷⁾. Los datos se recogen en múltiples observaciones antes y después de la intervención⁽¹²⁾. Los puntos antes de la intervención permiten estimar cualquier tenden-

cia subyacente o cambios cíclicos y los puntos después de la intervención permiten estimar el efecto de la intervención teniendo en cuenta los cambios temporales. Un ejemplo de este diseño lo podemos encontrar en un estudio orientado a evaluar la efectividad de una estrategia de difusión de GPC en médicos de familia, en el que se utilizaron series temporales interrumpidas, observándose pequeñas mejoras en la práctica. De forma paralela, se contrastaron los resultados que se hubieran obtenido utilizando un estudio antes-después, concluyendo que con éste último los resultados se hubiesen sobreestimado⁽²⁴⁾.

A continuación, el algoritmo 5.1, presenta un árbol de decisión que puede facilitar la selección de los diseños presentados, en función de los objetivos requeridos en la evaluación^(12,14,17). Además, en el anexo 2 se presentan ejemplos de algunas experiencias con diseños como los presentados en este capítulo.

Algoritmo 5.1. Alternativas en diseños de estudio para evaluar estrategias de implementación de GPC^(13,14,18)



5.4.3. Consideraciones sobre la evaluación de intervenciones basadas en organizaciones o áreas sanitarias

Tradicionalmente, la evaluación de la atención sanitaria se ha realizado considerando a un sujeto individual como la unidad de intervención y análisis, como por ejemplo un ensayo clínico para evaluar la eficacia de un tratamiento con un medicamento. En las evaluaciones de la implementación de GPC suelen ser unidades de intervención un grupo en un área geográfica o una organización. Los individuos no existen de forma aislada, y los cambios en las políticas sanitarias o en las organizaciones son usualmente implementados dentro de áreas o a nivel de un servicio clínico o centro sanitario. Esto hace que debamos considerar este punto cuando hagamos evaluaciones del efecto de la implementación de GPC⁽¹²⁾.

Además, desde el punto de vista práctico (factibilidad de realizar evaluación de este tipo de intervenciones), existen múltiples razones que hacen que la evaluación basada en grupos o *clusters* puede ser apropiada⁽²⁵⁾:

- Una intervención puede afectar a todos los miembros del área geográfica.
- Incluso cuando una evaluación de individuos es factible, pueden existir problemas éticos para realizar una intervención dentro de un grupo, diferente a otros.
- En algunas circunstancias, puede no resultar conveniente, desde una perspectiva administrativa o desde un punto de vista político, asignar miembros de una misma organización a diferentes grupos de intervención.
- La posibilidad de *contaminación* puede minimizarse a través de la asignación apropiada de los grupos en organizaciones sanitarias. De lo contrario, por ejemplo, los individuos de un grupo podrían transmitir el mensaje al grupo control⁽²⁶⁾.
- Este tipo de estudios puede ser más coste efectivo que la asignación de individuos.

Estas razones justifican que las evaluaciones basadas en áreas u organizaciones puedan requerir la asignación de los individuos a diferentes grupos de intervención. Esto necesariamente implicaría reconocer el *cluster* (organización, área) como la unidad de intervención o asignación, justificar el uso del *cluster* como unidad de intervención o asignación, incluir un número suficiente de *clusters* (evitar estudios con menos de cuatro *clusters* por grupo), aleatorizar los *clusters* siempre que sea posible y utilizar métodos de análisis que tengan en cuenta las consideraciones comentadas previamente^(13,27-31).

Un ejemplo de estudio que considera el *cluster* como una unidad de análisis es el ensayo de Althabe *et al.*⁽³²⁾, donde se pone en marcha una estrategia de implementación para la atención del parto y alumbramiento. Se realiza la aleatorización de 19 hospitales (10 para la intervención y 9 para control), y se reconoce el *cluster* (hospital) como unidad de análisis.

CUADRO RESUMEN DE LOS ASPECTOS CLAVE

- La evaluación de la implementación es necesaria porque nos permite conocer si las estrategias puestas en marcha para implementar la GPC han tenido éxito.
- En el plan de implementación deben identificarse los elementos que se van a evaluar, la información que se va a recopilar, los instrumentos para recopilarla y el diseño de estudios para la evaluación.
- Los resultados a evaluar en la implementación pueden agruparse en el nivel de paciente (cambios en la salud de los pacientes), de profesional (cambios en la práctica) y de la organización (cambio en el sistema sanitario, en políticas, costes y utilidad o extensión de la indicación de determinadas intervenciones).
- Los elementos clave a evaluar para determinar el resultado de la implementación tienen que ver con la evaluación, mediante indicadores, de la estructura, el proceso y los resultados de la utilización de la GPC.
- Se debe contemplar una revisión de medidas de resultado estandarizadas que puedan ser utilizadas globalmente, y definir medidas propias del proceso de evaluación (repositorios de indicadores).
- Para seleccionar los indicadores clave, se deben tener en cuenta las estrategias de implementación propuestas, los objetivos de la GPC, y los recursos y actores necesarios para la obtención de la información, incluidas las técnicas de recogida que se requieran.
- Hay que elegir un diseño adecuado para la evaluación, que dependerá del objetivo: probar el progreso en la implementación de la GPC y cambios en el proceso de atención en un/os centro/s (estudios observacionales), o proveer información generalizable del valor de una nueva estrategia o intervención para la implementación de una GPC (estudios de intervención).
- La evaluación basada en grupos o *clusters* en muchas ocasiones es apropiada, por su factibilidad, para realizar la evaluación de la implementación.

Bibliografía

1. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge: time for a map? *J Contin Educ Health Prof.* 2006;26(1):13–24.
2. Ward MM, Yankey JW, Vaughn TE, et al. Provider adherence to COPD guidelines: relationship to organizational factors. *J Eval Clin Pract.* 2005;11(4):379–87.
3. Hakkennes S, Green S. Measures for assessing practice change in medical practitioners. *Implement Sci.* 2006;1:29.
4. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century [Internet]. Washington: National Academies Press; 2001 [consulta 10 de octubre de 2008]. Disponible en: www.nap.edu/books/0309072808/html.
5. Lawrence M, Olesen F. Indicators of quality in health care. *Eur J Gen Pract.* 1997;3(3):103-8.
6. Porteous NL, Sheldrick BJ, Stewart PJ. Program evaluation toolkit: a blueprint for public health management. Ottawa: Ottawa-Carleton Health Department, Public Health Research, Education and Development Program; 1997.
7. Registered Nurses Association of Ontario. Toolkit: implementation of clinical practice guidelines. Toronto: RNAO; 2002.
8. Jennett PA, Sinclair L, Harrison RV. Methods, tools, and techniques of evaluation. En: Davis D, Barnes BE, Fox R, editores. *The continuing professional development of physicians: from research to practice.* Chicago: American Medical Association; 2003. p. 275-316.

9. Grol R, Baker RH, Moss F. Quality improvement research: understanding the science of change in health care. London: BMJ Books; 2004.
10. Harvey G, Wensing M. Methods for evaluation of small scale quality improvement projects. *Qual Saf Health Care*. 2003;12(3):210-4.
11. Eccles M, Grimshaw J, Campbell M, et al. Experimental evaluations of change and improvement strategies. En: Grol R, Wensing M, Eccles M, editores. *Improving patient care: the implementation of change in clinical practice*. Edinburgh: Elsevier; 2005. p. 235-247.
12. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol Assess*. 1999;3:iii-92.
13. Thorsen T, Mäkelä M. Changing professional practice theory and practice of clinical guidelines implementation. Copenhagen: Danish Institute for Health Services Research and Development; 1999.
14. Grol R, Wensing M, Eccles M. *Improving patient care: the implementation of change in clinical practice*. Edinburgh: Elsevier; 2005.
15. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. *Qual Saf Health Care*. 2002;11(3):270-5.
16. Black N. The relationship between evaluative research and audit. *J Public Health Med*. 1992;14(4):361-6.
17. Eccles M, Grimshaw J, Campbell M, et al. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care*. 2003;12(1):47-52.
18. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ*. 2006;332(7538):413-6.
19. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 2000;321(7262):694-6.
20. Cochrane Effective Practice and Organisation of Care Group (EPOC). Data collection checklist [Internet]. [Consulta 5 de diciembre de 2008] Disponible en: <http://www.epoc.cochrane.org/Files/Website%20files/Documents/Reviewer%20Resources/datacollectionchecklist.pdf>.
21. Varonen H, Rautakorpi UM, Nyberg S, et al. Implementing guidelines on acute maxillary sinusitis in general practice: a randomized controlled trial. *Fam Pract*. 2007;24(2):201-6.
22. Cochrane Effective Practice and Organisation of Care Group (EPOC). Study designs for EPOC reviews [Internet]. [Consulta 10 de diciembre de 2008]. Disponible en: <http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/FAQ%20-%20Included%20Studies%20-%20EPOC%20-%202007-May-01.doc>.
23. Müller U, Exadaktylos A, Roeder C, et al. Effect of a flow chart on use of blood transfusions in primary total hip and knee replacement: prospective before and after study. *BMJ*. 2004;328(7445):934-8.
24. Matowe L, Ramsay CR, Grimshaw JM, Gilbert FJ, Macleod MJ, Needham G. Effects of mailed dissemination of the Royal College of Radiologists' guidelines on general practitioner referrals for radiography: a time series analysis. *Clin Radiol*. 2002;57(7):575-8.

25. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG, Donner A. Methods in health service research: evaluation of health interventions at area and organisation level. *BMJ*. 1999;319(7206):376-9.
26. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ*. 2001;322(7282):355-7.
27. Kerry SM, Bland JM. Trials which randomize practices. II: sample size. *Fam Pract*. 1998;15(1):84-7.
28. Reed JF, III. Adjusted chi-square statistics: application to clustered binary data in primary care. *Ann Fam Med*. 2004;2(3):201-3.
29. Wears RL. Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. *Acad Emerg Med*. 2002;9(4):330-41.
30. Kerry SM, Bland JM. Trials which randomize practices. I: how should they be analysed? *Fam Pract*. 1998;15(1):80-3.
31. Kerry SM, Bland JM. Analysis of a trial randomised in clusters. *BMJ*. 1998;316(7124):54.
32. Althabe F, Buekens P, Bergel E, et al. A behavioral intervention to improve obstetrical care. *N Engl J Med*. 2008;358(18):1929-40.