

El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar

María F. Grande Ratti¹, Noelia S. De Masi², José A. Orozco Niño¹ y Sebastián Marciano³

1. Área de Investigación en Medicina Interna, Servicio de Clínica Médica, Hospital Italiano. Buenos Aires, Argentina

2. Departamento Experiencia del Cliente, Hospital Italiano. Buenos Aires, Argentina

3. Secretaría de Investigación Clínica, Universidad Hospital Italiano. Buenos Aires, Argentina

RESUMEN

Introducción: el *benchmarking* se ha consolidado como una herramienta estratégica clave en la gestión sanitaria, alineada con el principio de que aquello que no se mide no puede mejorarse. Aunque es un anglicismo, su uso está plenamente aceptado en la literatura científico-sanitaria. Entendido como un proceso sistemático de medición, comparación y aprendizaje, permite evaluar el desempeño organizacional a partir de indicadores válidos y comparables. En un contexto de creciente demanda de eficiencia, seguridad y resultados centrados en el paciente, el *benchmarking* facilita la mejora continua de la calidad asistencial y el posicionamiento institucional mediante la transparencia y el aprendizaje organizacional.

Estado del arte: la evidencia muestra que el *benchmarking* en salud se apoya en el uso de indicadores de estructura, proceso y resultado, según el modelo de Donabedian. Diversas revisiones sistemáticas han demostrado una asociación positiva entre la implementación de *benchmarking* y mejoras en la calidad asistencial, tanto en procesos como en resultados clínicos. No obstante, su impacto suele potenciarse cuando se combina con intervenciones complementarias como auditorías, retroalimentación, planes de mejora o incentivos, lo que dificulta aislar su efecto específico. Asimismo, persisten desafíos metodológicos relevantes, como la comparabilidad entre instituciones, la necesidad de ajustes por riesgo, la calidad y disponibilidad de los datos, y la sostenibilidad de las mejoras a largo plazo. En este marco, adquiere creciente relevancia la integración de indicadores de experiencia del paciente y datos cualitativos, que amplían la evaluación tradicional del desempeño hacia una perspectiva más integral y centrada en la persona.

Discusión: no debe entenderse como una simple comparación ni como un sustituto de los programas de calidad, sino como un componente esencial de estos. Su implementación efectiva requiere sistemas de información robustos, indicadores bien definidos, equipos interdisciplinarios, liderazgo institucional y una cultura organizacional orientada al aprendizaje y la mejora continua. En conclusión, la ciencia de la implementación resulta clave para cerrar la brecha entre la evidencia y la práctica, y la evaluación comparativa basada en datos de alta calidad se ha vuelto una práctica prácticamente obligatoria en los sistemas de salud modernos.

Palabras clave: benchmarking, gestión de la información en salud, calidad de la atención en salud, indicadores de gestión, mejoramiento de la calidad, satisfacción del paciente.

Autor para correspondencia: maria.grande@hospitalitaliano.org.ar, Grande Ratti MF.

Recibido: 3/02/2026 | Aceptado: 15/09/2026 | Publicado: 28/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1317>

Cómo citar: Grande Ratti MF, De Masi NS, Orozco Niño JA, Marciano S. El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar. *Rev Hosp Ital B. Aires.* 2026;46(3):e0001317

The Role of Benchmarking in Healthcare Management: Beyond Measurement, a Strategy for Learning and Improvement

ABSTRACT

Introduction: benchmarking has become a key strategic tool in healthcare management, aligned with the principle that what is not measured cannot be improved. Although the term is an anglicism, its use is fully accepted in the scientific and healthcare literature. Understood as a systematic process of measurement, comparison, and learning, it allows for the evaluation of organizational performance based on valid and comparable indicators. In a context of increasing demand for efficiency, safety, and patient-centered outcomes, benchmarking facilitates the continuous improvement of healthcare quality and institutional positioning through transparency and organizational learning.

State of the art: evidence shows that benchmarking in healthcare relies on the use of structure, process, and outcome indicators, according to the Donabedian model. Several systematic reviews have demonstrated a positive association between the implementation of benchmarking and improvements in healthcare quality, both in processes and clinical outcomes. However, its impact is usually enhanced when combined with complementary interventions such as audits, feedback, improvement plans, or incentives, which makes it difficult to isolate its specific effect. Furthermore, significant methodological challenges persist, such as comparability between institutions, the need for risk adjustments, data quality and availability, and the long-term sustainability of improvements. Within this framework, the integration of patient experience indicators and qualitative data is gaining increasing importance, broadening the traditional performance evaluation toward a more holistic and person-centered perspective.

Discussion: this should not be understood as a simple comparison or a substitute for quality programs, but rather as an essential component of them. Its effective implementation requires robust information systems, well-defined indicators, interdisciplinary teams, institutional leadership, and an organizational culture oriented toward learning and continuous improvement. In conclusion, implementation science is key to bridging the gap between evidence and practice, and benchmarking based on high-quality data has become virtually mandatory in modern healthcare systems.

Keywords: benchmarking, health information management, quality of health care, management indicators, quality improvement, patient satisfaction.

INTRODUCCIÓN

La presente revisión narrativa tiene como objetivo integrar y analizar los principales componentes conceptuales, metodológicos y organizacionales que condicionan la utilidad del *benchmarking* en la gestión sanitaria. Si bien la medición y comparación sistemática del desempeño constituyen herramientas fundamentales para identificar brechas de calidad y orientar procesos de mejora, la utilidad no depende únicamente de la disponibilidad de indicadores, sino también de su pertinencia, calidad y comparabilidad, así como de la adecuada interpretación de las diferencias observadas en función del contexto en el que se producen.

Desde tal perspectiva, este artículo aborda los fundamentos conceptuales del *benchmarking* y sus principales modalidades de aplicación en salud, examinando el papel de los indicadores clave de desempeño (KPIs por *Key Performance Indicators*), los desafíos asociados a la medición y comparabilidad del desempeño, así como la incorporación de dimensiones vinculadas con la experiencia del paciente.

Asimismo, se analizan las condiciones organizacionales necesarias para que los resultados de la comparación puedan trascender la identificación de diferencias y traducirse en aprendizaje, adopción de mejores prácticas y acciones sostenibles de mejora.

La medición de indicadores permite cuantificar determinados aspectos del desempeño; la monitorización del desempeño implica el seguimiento sistemático de estos indicadores a lo largo del tiempo; los tableros de gestión (*dashboards*) constituyen herramientas para visualizar y comunicar información de desempeño; y los programas de mejora de la calidad comprenden intervenciones destinadas a modificar los procesos y resultados. En contraste, se entiende el *benchmarking* como un proceso sistemático de comparación del desempeño con un referente, orientado a identificar brechas, comprender diferencias y favorecer el aprendizaje y la mejora.

La contribución consiste en articular estas dimensiones desde la gestión sanitaria, destacando que no debe entenderse únicamente como una herramienta

de medición, sino como un mecanismo de aprendizaje organizacional y de mejora continua, que permite identificar referentes, comprender las diferencias observadas y analizar los procesos que podrían explicarlas. El valor no reside en identificar quién obtiene mejores resultados, sino en comprender las diferencias, interpretar los factores que las explican y determinar qué prácticas pueden ser adaptadas al contexto propio. Esta revisión busca aportar un marco conceptual con orientación más pragmática para comprender bajo qué condiciones el *benchmarking* puede constituirse en una herramienta efectiva para mejorar el desempeño y favorecer una atención sanitaria más equitativa, transparente y orientada a la mejora continua.

ESTADO DEL ARTE

“BENCHMARKING” EN SALUD: DEFINICIÓN Y FUNDAMENTOS

La frase “Lo que no se puede medir, no se puede mejorar” suele atribuirse a Peter Drucker, por parafraseos conceptuales similares en sus obras, que hacen referencia a la importancia de medir para gestionar. Sin embargo, fue publicada por primera vez por Ridgway en 1956¹. En salud, el *benchmarking* es un proceso que se basa precisamente en medir, comparar y aprender.

Ante todo, reconocemos que se trata de un anglicismo cuya traducción al español podría ser “evaluación comparativa” o “comparación con estándares de referencia”, aunque no resulte plenamente equivalente ni capte con precisión el uso consolidado del término en la literatura internacional de gestión y evaluación en salud. Dado que el *benchmarking* es un concepto técnico ampliamente aceptado en el ecosistema científico-sanitario, en el presente artículo mantendremos la denominación original en inglés.

En el ámbito sanitario puede entenderse como un proceso sistemático de medición, comparación e interpretación del desempeño respecto de referentes definidos, orientado al aprendizaje y la mejora. Se trata de un campo en expansión y se ha incorporado como una herramienta estratégica de gestión orientada a la mejora continua de la calidad asistencial. Etimológicamente, proviene del concepto de “punto de referencia” (*benchmark*) y su uso se extendió desde el ámbito empresarial, adquiriendo una formulación sistemática como herramienta de gestión a partir de la experiencia de Xerox durante la década de 1970, y posteriormente documentada por Camp en 1989².

En un entorno cada vez más exigente en términos de eficiencia, seguridad y resultados centrados en el paciente³, el *benchmarking* ofrece un marco estructurado para la evaluación y el aprendizaje de una organización. Además de impulsar mejoras internas, permite a las instituciones posicionarse al mostrar resultados transparentes y sostenidos en la calidad, la seguridad y la experiencia.

Recientemente, Byrne y Haut nos recordaron que la atención de los pacientes debe cumplir con un estándar, independientemente de la ubicación geográfica donde se produzca la enfermedad/lesión, o del centro de atención, o del equipo profesional tratante⁴. En otras palabras, siempre se debería garantizar el tratamiento adecuado, en el lugar y el momento apropiados para lograr el mejor

resultado posible⁵. Desafortunadamente, la persistencia de variaciones en el desempeño asistencial pone de manifiesto la necesidad de contar con métodos que permitan medir, comparar e interpretar sistemáticamente esas diferencias. El *benchmarking* fomenta el uso de indicadores válidos, comparables y ajustados, agrupados por factores de riesgo, complejidad o patologías en salud. No es una simple comparación, sino que requiere análisis crítico, interpretación contextual y capacidad de adaptación local. Su uso también abarca la gestión (p. ej., rediseño de procesos y espacios, comunicación, experiencia del paciente), ya que facilita la mejora continua y la toma de decisiones basada en evidencia, orienta la priorización de oportunidades de mejora y promueve el aprendizaje organizacional a partir de experiencias propias y de otras instituciones, contribuyendo así a fortalecer la experiencia de pacientes, acompañantes y equipos de trabajo.

Su eficacia aumenta cuando los equipos participan activamente en la definición e interpretación de los indicadores (instrumentos de medición), y cuando existe liderazgo institucional para traducir los hallazgos en acciones de mejora⁶. Esta metodología se sirve de indicadores correspondientes a las tres dimensiones de calidad (estructura, procesos y resultados) para evaluar y comparar sistemáticamente el desempeño entre organizaciones, servicios o procesos, establecer referentes e identificar prácticas potencialmente transferibles cuya adaptación al contexto local pueda ser evaluada⁷.

TIPOS DE BENCHMARKING

Ya en el año 1994, Benson informaba cuatro tipos diferentes⁸ (Tabla 1).

- Interno, que compara las funciones dentro de una organización.
- Competitivo, ya que compara productos o servicios directamente, que operan en el mismo mercado.
- Funcional, que compara un aspecto o una función específica, como nóminas o compras (p. ej., gestión de turnos), con organizaciones que no necesariamente pertenecen al mismo sector, pero que realizan ese proceso de manera destacada.
- Genérico, permite una comparación más amplia ya que busca identificar y adaptar principios de gestión, enfoques estratégicos o metodologías transversales aplicables a cualquier organización. Se trata de mirar cómo otras empresas hacen las cosas de manera eficiente y pensar en cómo adaptar esos métodos a nuestro contexto. El pensamiento Lean, inspirado en el sistema de producción Toyota, busca maximizar el valor para el usuario mediante la identificación y eliminación de actividades que no agregan valor, la reducción del desperdicio y la mejora continua de los procesos⁹.

INDICADORES Y KPIs: SELECCIÓN, DEFINICIÓN Y ESTANDARIZACIÓN

Según el modelo de Donabedian, los indicadores pueden clasificarse en *estructura*, *procesos* y *resultados*¹⁰. La tabla 2 muestra ejemplos en la atención clínica de personas con diabetes.

Tabla 1. Tipos de *benchmarking* según Benson con ejemplos clínicos y de gestión

Tipo	Ejemplo clínico	Ejemplo de gestión
Interno	Comparación entre distintos servicios de urgencias de una misma institución (Hospital Italiano)	Comparación de la eficiencia en la asignación de turnos entre diferentes unidades administrativas del mismo establecimiento
Competitivo	Tasa de reconsultas a las 72 horas en la Central de Emergencia de Adultos, sede central vs. San Justo Comparación de mortalidad en lista de espera o tiempo de resolución quirúrgica entre hospitales de una misma red	Demora (en días) para la asignación de turno ambulatorio programado entre Ginecología/Obstetricia vs. Hematología
Funcional	Comparación de la gestión del flujo del paciente crítico con centros de emergencias de otro sistema que han optimizado triaje y derivación	Comparación del costo por paciente atendido entre centros de salud de la misma región Comparación de la gestión de turnos o logística de insumos con organizaciones externas
Genérico	Reducir tiempos de espera en consultas externas o mejorar el recorrido del paciente	Un hospital podría inspirarse en la logística de supermercados para organizar la entrega de medicamentos

Fuente: elaboración propia

Tabla 2. Clasificación de indicadores según el modelo de Donabedian

Categoría	Definición	Ejemplo
Estructura	Recursos físicos, humanos y organizacionales disponibles para la atención de salud	Porcentaje de centros ambulatorios que cuentan con al menos un profesional certificado en educación diabetológica Proporción de centros con disponibilidad de HbA1c en laboratorio
Procesos	Actividades y procedimientos mediante los cuales se brinda la atención	Porcentaje de pacientes con HbA1c medida al menos dos veces en los últimos 12 meses Porcentaje de pacientes con evaluación de fondo de ojo anual
Resultados	Cambios en el estado de salud atribuibles a la atención recibida	Proporción de pacientes con HbA1c < 7% (o según metas individualizadas) Porcentaje con LDL < 100 mg/dL (o < 70 mg/dL en alto riesgo)

Fuente: elaboración propia

Su aplicación en salud está relacionada con un trabajo de equipo, que incluya al médico especialista en tema/patología/población que se está evaluando, así como otros miembros con conocimientos técnicos (p. ej., informática, epidemiología, bioestadística, análisis de datos y otras disciplinas afines que permitan la adecuada reutilización de la información derivada de la historia clínica electrónica) y de comunicación –para la medición cuantitativa y su difusión–, y experiencia en gestión de proyectos¹¹.

Los objetivos de las métricas habitualmente más utilizadas en salud (productividad, calidad, tiempos y costos) son fundamentalmente los siguientes:

- Identificar oportunidades de mejora en procesos clínicos (p. ej., tiempo a antibióticos, mortalidad), administrativos o de gestión (p. ej., readmisiones).
- Adoptar mejores prácticas que hayan demostrado ser eficaces en contextos similares (p. ej., resultados en procedimientos complejos como trasplante hepático).
- Reducir la variabilidad injustificada entre unidades, servicios o instituciones.
- Mejorar la toma de decisiones mediante datos comparativos y evidencia.
- Estimular una cultura de calidad centrada en el aprendizaje organizacional (p. ej., establecer planes de mejora tras el análisis de desviaciones).

Dado que los procesos de resultados son complejos y costosos de medir, muchos sistemas sanitarios han recurrido a restringir el uso de mediciones indirectas y más sencillas, como las de estructura y de proceso¹².

Los indicadores clave de desempeño (KPIs por *Key Performance Indicators*) y su construcción deben considerarse entre los pasos principales para estructurar un sistema de monitorización continua que permita evaluar desempeño, identificar brechas y guiar acciones de mejora¹³. Para el diseño de los paneles de gestión hospitalaria, la propia participación de los usuarios principales en la selección y la priorización contribuye a su éxito¹⁴.

Una revisión sistemática de 2024 identificó y clasificó 64 indicadores clave (financieros, gerenciales y clínicos) utilizados para medir el desempeño en la cadena de abastecimiento hospitalaria¹⁵. El estudio remarcó que, para los hospitales, identificar y generar estos indicadores representa un enorme desafío, debido a la frecuente falta de fuentes de datos completas y comparables. El problema mayor radica en que la validez y la fiabilidad dependen de la disponibilidad de datos de alta calidad, en concreto, de los elementos relevantes del conjunto mínimo de datos¹⁶.

Un estudio italiano de 2013 revisó la evidencia y los datos hospitalarios para analizar la relación entre el volumen de actividad (cantidad de casos atendidos) y los resultados sanitarios (especialmente mortalidad intrahospitalaria a 30 días)¹⁷. No obstante, los propios autores advierten que no siempre fue posible fijar umbrales universales y que las diferencias entre condiciones o patologías clínicas obligan a un enfoque contextualizado¹⁷.

Desde la perspectiva de Experiencia del Paciente, el *benchmarking* incorpora una dimensión adicional de interés poblacional al incluir indicadores de percepción

y de participación de pacientes/familias. La integración sistemática de esta información cualitativa en los *dashboards* no solo permite comparar resultados clínicos y de proceso, sino también comprender cómo dichos resultados son vividos y significados por los propios usuarios. En consecuencia, es fundamental integrar datos cualitativos (p. ej., voz del paciente, encuestas de experiencia y relatos narrativos), dentro de los indicadores de resultado¹⁸. Estos insumos aportan dimensiones que no pueden ser captadas exclusivamente por mediciones cuantitativas y permiten una comprensión más profunda¹⁹. No obstante, su interpretación exige una adecuada contextualización según las características de la población atendida, el tipo de servicio y las particularidades de cada proceso. Este enfoque amplía la interpretación tradicional del desempeño institucional, al aportar información contextual, subjetiva y relacional, que resulta esencial para evaluar la calidad del cuidado de manera integral y centrada en la persona.

Las iniciativas orientadas a estandarizar resultados relevantes para las personas, como ICHOM (por International Consortium for Health Outcomes Measurement)²⁰, desplazan el foco desde la comparación exclusiva de actividad o desempeño operativo hacia la comparación de resultados que importan a quienes reciben la atención, ayudando a operacionalizar qué resultados deberían medirse. En este sentido, han desarrollado conjuntos estandarizados de resultados para determinadas patologías (p. ej., fibrilación auricular, enfermedad renal crónica, insuficiencia cardíaca)^{21–23}.

De manera similar, PROMs (por *Patient-Reported Outcome Measures*) y PREMs (por *Patient-Reported Experience Measures*) son fuentes de información para construir medidas de desempeño que permiten incorporar la perspectiva de las personas²⁴. Mientras que las primeras evalúan qué resultado tiene la enfermedad o la atención sobre la salud de la persona, desde su propia perspectiva (p. ej., calidad de vida, capacidad funcional)²⁵, las segundas abordan cómo fue la experiencia¹⁸. Entonces, PROMs/PREMs proporcionan mecanismos para capturar directamente la perspectiva de las personas. Según Jamieson Gilmore y cols.²⁶, los PREMs se utilizan principalmente “a nivel micro” para mejorar la calidad de la atención, mientras que su aplicación “a nivel meso” se vincula con la mejora del desempeño y la financiación basada en calidad, y su uso “a nivel macro” aún continúa siendo limitado.

CALIDAD Y GOBERNANZA DE LOS DATOS

La evidencia destaca la importancia de los datos para tomar decisiones informadas. Por ende, la recolección y la gestión de datos constituyen el pilar. Medir, analizar e interpretar datos permite no solo conocer el estado actual del sistema de salud, sino también anticipar necesidades, corregir desviaciones y promover mejoras continuas.

Sin embargo, la disponibilidad de grandes volúmenes de información no garantiza, por sí misma, que las comparaciones sean válidas. Los datos utilizados para *benchmarking* deben ser suficientemente completos, exactos, consistentes, oportunos, pertinentes y comparables, y contar con definiciones operativas y reglas de cálculo

claramente establecidas (p. ej., criterios homogéneos de inclusión y exclusión, períodos de observación equivalentes y procesos suficientemente estandarizados como para que las diferencias observadas puedan atribuirse al desempeño y no a diferencias en la forma de registrar o brindar la atención). Una revisión sistemática reciente identificó múltiples dimensiones de calidad de los datos en los sistemas de información sanitaria, entre ellas exactitud, completitud, consistencia, oportunidad, fiabilidad, accesibilidad y seguridad, destacando la heterogeneidad existente en su definición y evaluación²⁷.

En este contexto, la gobernanza de datos complementa la evaluación de su calidad. No se limita a aspectos técnicos, sino que comprende definir qué información se recolectará, con qué propósito, mediante qué fuentes, utilizando qué definiciones y bajo qué procedimientos de validación. También requiere establecer mecanismos para detectar errores, datos faltantes, cambios en los procesos de registro y modificaciones en los sistemas de información²⁸.

Entonces, la medición requiere un proceso sistemático y estandarizado que garantice la validez (si los datos realmente miden lo que se supone que deberían, ¿estamos midiendo lo correcto?), fiabilidad (si los datos son consistentes y estables cuando se miden repetidas veces, ¿da el mismo resultado cada vez que se mide?) y comparabilidad de los datos (si los datos se pueden cotejar con otros de diferentes lugares, tiempos o grupos, ¿podemos comparar estos datos con otros y que dicha

comparación tenga sentido?), para lo cual existen 4 pasos fundamentales:

- Definición clara de los indicadores: fuente de datos, método de cálculo, frecuencia de medición y umbrales esperados. Por ejemplo, mortalidad en lista de espera para trasplante hepático en adultos, estimada considerando el trasplante como riesgo competitivo, informada anualmente y comparada con los datos del mismo año de la United Network for Organ Sharing (UNOS). En una aproximación más realista, no basta con hacer un Kaplan-Meier simple, porque eso podría sobreestimar o subestimar el riesgo de muerte al no considerar que algunos se trasplantan antes de morir.

- Sistemas de información integrados: historias clínicas electrónicas, bases de datos administrativas, registros específicos.

- Uso de tecnologías digitales: herramientas de visualización que permitan procesar grandes volúmenes de información en tiempo real, por ejemplo *dashboards*⁵.

- Capacitación del personal: todos los actores implicados deben comprender la importancia de registrar los datos adecuadamente y utilizarlos éticamente²⁹.

El “Marco 5-P” del proceso de selección de indicadores (Fig. 1) consta de cinco dominios³⁰: Propósito (objetivo mediante un proceso operativo o alcance mediante un plan estratégico), Política (gobernanza o comité directivo), Preparación (búsqueda bibliográfica y opinión del usuario final), Procedimiento (método reconocido de construcción de consenso) y Probar (pruebas piloto de calidad, entrenamientos).

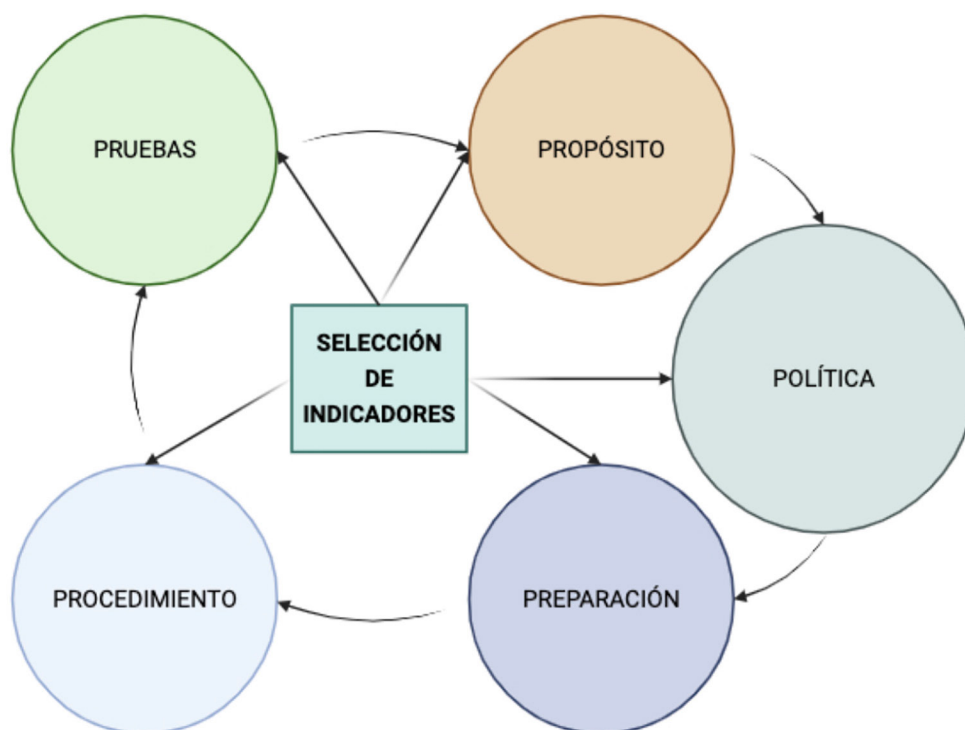


Figura 1. Los cinco dominios ('5-P') del proceso de selección de indicadores.

Fuente: elaboración propia, creada con BioRender.com®.

¿QUÉ SIGNIFICA COMPARAR EN SALUD?

Implica mucho más que disponer de indicadores y establecer diferencias entre dos o más organizaciones, servicios o unidades asistenciales. La comparación permite identificar variaciones en el desempeño y formular preguntas acerca de su magnitud, consistencia y posible relevancia. Sin embargo, una diferencia observada no necesariamente representa una diferencia de calidad. Puede reflejar diferencias en las características de las poblaciones atendidas, complejidad de los casos, disponibilidad de recursos, estructura organizacional, modelos de atención, sistemas de información o contexto socioeconómico. Estas diferencias adquieren especial relevancia en América Latina, donde coexisten sistemas de salud fragmentados y heterogéneos, con variaciones importantes en cobertura, acceso, disponibilidad de recursos y capacidad de los sistemas de información entre países y, en algunos casos, dentro de un mismo país. Por ello, interpretar una comparación requiere considerar tanto las características de las unidades comparadas como las condiciones en las que se generan los resultados.

Este aspecto resulta particularmente relevante en salud, donde la heterogeneidad entre instituciones y poblaciones es considerable. La comparación de tasas de mortalidad, reingreso, complicaciones, tiempos de espera, utilización de recursos o experiencia de los pacientes puede ser informativa, pero su interpretación requiere determinar si las diferencias observadas son esperables, atribuibles al riesgo basal de las poblaciones, consecuencia de diferencias en los procesos de atención o potencialmente evitables. En consecuencia, la identificación de una brecha constituye el punto de partida de la indagación y no necesariamente una conclusión sobre el desempeño. Una comparación orientada al aprendizaje busca avanzar hacia preguntas de mayor profundidad: ¿por qué existen estas diferencias?, ¿qué procesos, prácticas o condiciones podrían explicarlas?, ¿qué podemos aprender de quienes obtienen mejores resultados? y ¿qué elementos podrían adaptarse a nuestro propio contexto? Es en este punto donde la comparación adquiere sentido como componente del *benchmarking*.

En el contexto de la evaluación comparativa es importante reconocer que los indicadores pueden no ser directamente comparables entre instituciones debido a diferencias en sus definiciones operativas, las características de las poblaciones atendidas, los procesos asistenciales y el contexto organizacional, que Byrne y Haut describieron haciendo uso de la expresión *comparing apples to oranges*, que se conoce coloquialmente en nuestro contexto como “peras con manzanas”⁴. Particularmente en América Latina y la Argentina, la heterogeneidad entre instituciones, modelos de atención, disponibilidad de recursos y sistemas de información puede limitar la comparabilidad directa. Incluso los modelos de ajuste de riesgos empleados para generar informes de *benchmarking* podrían no capturar plenamente las diferencias en las características de los pacientes atendidos entre hospitales, lo que puede conducir a comparaciones potencialmente injustas.

La interpretación de las comparaciones en salud requiere integrar consideraciones epidemiológicas, estadísticas y económicas, dado que las poblaciones y los contextos asistenciales rara vez son idénticos. Por ejemplo, si una guardia hospitalaria presentara una demora promedio de dos horas y el *benchmark* establece una hora, la interpretación de esta diferencia requiere considerar el “contexto” y la metodología utilizada para el cálculo del indicador. En este mismo sentido, Van Vliet y cols. compararon tres hospitales oftalmológicos del Reino Unido, Estados Unidos y Países Bajos⁸¹. A pesar de tratarse de importantes centros internacionales de atención terciaria y formación en Oftalmología, los autores se vieron obligados a informar factores ambientales que influyen en los procedimientos de cataratas, viéndose obligados a proporcionar detalles sobre la ubicación, el tipo, la condición de hospital universitario y el volumen operativo de las organizaciones.

De forma similar, Byrne y col. citaron un ejemplo quirúrgico en hospitales de alta complejidad, donde algunos reciben derivaciones de casos de trauma graves (p. ej., heridas de bala), que incrementan su riesgo basal de mortalidad⁴. En tales casos, se sugiere excluir a esta subpoblación para permitir comparaciones más equitativas y justas frente a hospitales que no reciben esas derivaciones.

COMPARABILIDAD E INTERPRETACIÓN DE LAS DIFERENCIAS (Tabla 3)

Ya hacia el año 2006, Arah y cols. hablaban fundamentalmente de un marco metodológico para la medición y comparación del desempeño de sistemas, particularmente de indicadores de calidad de la atención sanitaria entre países pertenecientes a la Organización para la Cooperación y el Desarrollo Económicos (OCDE, es decir, Organization for Economic Cooperation and Development)³². El punto de partida de los autores es muy interesante: antes de seleccionar indicadores hay que preguntarse: ¿El desempeño de qué estamos midiendo, y con qué propósito? Distinguen fundamentalmente entre dos tipos: un marco de desempeño en salud, que adopta una perspectiva más amplia, social o de salud pública, sobre la determinación de la salud; mientras que un marco de desempeño en atención adopta una perspectiva más restringida, principalmente clínica o técnica, sobre la atención médica en relación con las necesidades de salud. Los autores argumentan que el *benchmarking* no debería comenzar simplemente por elegir indicadores y comparar números en un microsistema, sino que debería ir más allá de la medición y favorecer el aprendizaje de procesos.

Kruk y cols. en 2018⁸³ plantearon que históricamente se ha puesto mucho énfasis en cobertura, acceso y utilización, mientras que la calidad de la atención ha recibido relativamente menos atención. Esto genera una situación paradójica: aumentar la cobertura de servicios de baja calidad puede producir poco beneficio e incluso daño. En consecuencia, propusieron pensar la calidad como algo que debe estar presente en todo el sistema sanitario, y no solamente como características de una prestación indivi-

Tabla 3. De la definición de resultados a la comparación del desempeño: conceptos y herramientas clave

Concepto	¿Para qué sirve?
Benchmarking	Permite comparar el desempeño y aprender de las diferencias
VBHC	Ayuda a definir qué resultados importan
ICHOM	Aporta un ejemplo de estandarización sobre qué y cómo medir (para que los resultados sean comparables)
PROMs/PREMs	Contribuyen a medir resultados y experiencia desde la perspectiva del paciente
VBHC: <i>Value-Based Healthcare</i>	
ICHOM: International Consortium for Health Outcomes Measurement	
PROMs: <i>Patient-Reported Outcome Measures</i> . PREMs: <i>Patient-Reported Experience Measures</i>	

dual. Por ende, su aporte es más amplio: proporciona una justificación conceptual para medir y comparar calidad, identificar brechas y utilizar esa información para orientar la mejora. El artículo utiliza datos comparativos y muestra grandes variaciones en indicadores de calidad de atención, incluso entre países con niveles similares de desarrollo o recursos. La comparación permite hacer visible una cuestión fundamental: que existan diferencias de desempeño entre organizaciones abre la posibilidad de preguntarse qué están haciendo diferente aquellos que obtienen mejores resultados, haciendo hincapié en cómo transformar la comparación en aprendizaje organizacional.

Los indicadores permiten detectar esta brecha; el *benchmarking* busca comprenderla y aprender de quienes obtienen mejores resultados. El marco conceptual de atención basada en valor (VBHC, por *Value-Based Healthcare*)³⁴ ayuda a responder preguntas complementarias fundamentales: ¿Qué entendemos por valor? o ¿qué resultados importan? o ¿qué significa realmente “mejor desempeño” en salud?

DEL BENCHMARKING A LA MEJORA CONTINUA

En una revisión sistemática del año 2022 se incluyeron 17 trabajos que evidenciaron una asociación positiva entre la implementación de *benchmarking* y las mejoras en la calidad asistencial: 10 estudios mostraron mejoras en procesos, 13 en resultados y 7 en ambos tipos de indicadores³⁵. Willmington y cols. informaron que la mayoría de los estudios (12 de 17) combinaron el *benchmarking* con al menos una intervención complementaria (p. ej., reuniones entre participantes clave, auditorías y retroalimentación, planes de mejora y/o incentivos financieros), y en 10 estudios se usaron múltiples estrategias de forma simultánea³⁵. Esto último dificulta discernir cuál es el efecto atribuible estrictamente al *benchmarking* y subraya que su eficacia parece reforzarse cuando se acompaña de intervenciones adicionales. Otra limitación mencionada es que muchos estudios tienen períodos de seguimiento cortos (p. ej., entre seis meses y dos años), lo que genera dudas sobre la sostenibilidad del efecto observado, pues las mejoras iniciales podrían

deberse al efecto posimplementación inmediato del proyecto más que a un cambio organizacional y cultural perdurable a largo plazo³⁵. En este sentido, el *benchmarking* debería considerarse parte de los programas de gestión de calidad, y no un sustituto³⁶.

DISCUSIÓN

Los cambios organizacionales que podrían mejorar la salud de los pacientes suelen ser difíciles de implementar en la práctica, incluso cuando están respaldados por evidencia sólida; se requieren equipos interdisciplinarios y una cultura colaborativa³⁷. La “ciencia de la implementación” es el estudio de métodos y estrategias para promover la adopción e integración de tratamientos clínicos, prácticas, intervenciones organizacionales y/o de gestión comprobados en la práctica clínica habitual y, por lo tanto, mejorar la salud³⁸. Es aquella que cierra el salto “de la teoría a la práctica”, estudiando cómo las intervenciones o las prácticas basadas en la evidencia se difunden en el mundo real y comprendiendo las barreras que las limitan³⁸. En particular, existen herramientas que pueden utilizarse antes, durante o después del proceso de comparación para transformar los hallazgos del *benchmarking* en acciones de mejora como: a) Lean, que ayuda a identificar desperdicios, variabilidad y actividades que no agregan valor³⁹; b) Six Sigma, que aporta herramientas para reducir variabilidad y defectos, y mejorar la capacidad de los procesos⁴⁰ y c) ciclos PDSA (de *plan-do-study-act*), un enfoque complementario que proporciona un mecanismo iterativo para probar y ajustar cambios⁴¹.

Es el paso final clave para llevar la atención de las “mejores prácticas” y potencialmente transferibles a un gran número de pacientes⁴². Sin embargo, la gran cantidad y complejidad de los datos del mundo real también exigen el desarrollo de técnicas de procesamiento y análisis de datos más apropiadas, sofisticadas e innovadoras, manteniendo el rigor científico de los hallazgos de investigación y prestando atención a la ética de los datos para aprovechar al máximo su potencial⁴³.

Muchas organizaciones aún no entienden que poner al paciente en el centro es el propósito principal de la

atención sanitaria (p. ej., escepticismo o miedo a alguna aplicación punitiva), lo que implica integrarlo activamente en las mejoras del sistema, algo que sigue generando resistencia entre algunos líderes y profesionales³. Un proceso continuo de evaluación y comunicación puede contribuir a fomentar una cultura de transparencia, confianza y aprendizaje.

El objetivo final de medir la calidad de la atención es proporcionar nuevos conocimientos sobre cómo mejorar el rendimiento: qué cambios se pueden realizar en las estructuras o procesos existentes para mejorar los resultados de los pacientes. La evaluación comparativa del desempeño se ha incorporado ampliamente a las estrategias de evaluación y mejora de la calidad de la atención sanitaria⁴. Sin embargo, su utilidad no depende únicamente de la disponibilidad de indicadores comparables, sino también de la calidad de los datos, la validez de las medidas y la consideración de las diferencias en el riesgo, el contexto de las poblaciones evaluadas, gobernanza clara, transparencia y una cultura de mejora no punitiva.

CONCLUSIONES

El *benchmarking* puede constituir una herramienta de alto valor para la gestión sanitaria al transformar la medición del desempeño en información orientada a la mejora. Su utilidad radica en identificar brechas, comprender las diferencias observadas y favorecer el aprendizaje y la adopción de prácticas basadas en la evidencia, considerando el contexto clínico, institucional y poblacional.

Su capacidad para generar valor depende, sin embargo, de comparaciones metodológicamente válidas, sustentadas en datos de calidad, indicadores comparables y ajustes adecuados. Integrado a estrategias de mejora de la calidad y a organizaciones que favorezcan la transparencia, la confianza y el aprendizaje, el *benchmarking* puede contribuir a traducir los hallazgos en acciones de mejora. Por el contrario, datos incompletos, indicadores no comparables o ajustes insuficientes pueden conducir a interpretaciones erróneas o comparaciones injustas. La incorporación de la experiencia del paciente puede contribuir a orientar estas evaluaciones hacia resultados relevantes para las personas.

Contribuciones de los autores: Conceptualización, Administración del proyecto (MFGR, SM). Curación de datos, Análisis formal, Investigación, Metodología (MFGR, JAON, SM). Recursos, Supervisión (MFGR). Validación, Visualización: MFGR, SM. Redacción–borrador original, Redacción–revisión y edición (MFGR, NSDM, JAON, SM).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Ridgway VF. Dysfunctional consequences of performance measurements. *Adm Sci Q*. 1956;1(2):240-247. <https://doi.org/10.2307/2390989>.
2. Camp RC. Benchmarking: the search for industry best practices that lead to superior performance. New York: CRC Press; 2024. <https://doi.org/10.4324/9781003578871>.
3. Edgman-Levitan S, Schoenbaum SC. Patient-centered care: achieving higher quality by designing care through the patient's eyes. *Isr J Health Policy Res*. 2021;10(1):21. <https://doi.org/10.1186/s13584-021-00459-9>.
4. Byrne JP, Haut ER. Facts and fallacy of benchmark performance indicators. *Adv Surg*. 2022;56(1):89-109. <https://doi.org/10.1016/j.yasu.2022.03.002>.
5. Frutos EL, Iannicelli A, Luna D, et al. Implementation of a real-time dashboard for the emergency department. *Stud Health Technol Inform*. 2025;329:184-188. <https://doi.org/10.3233/SHTI250826>.
6. Nilsen P, Seing I, Ericsson C, et al. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. *BMC Health Serv Res*. 2020;20:147. <https://doi.org/10.1186/s12913-020-4999-8>.
7. Grande Ratti MF, Luna DR. Indicadores en salud: uso en gestión para mejorar la toma de decisiones. *Rev Hosp Ital B Aires*. 2021;41(2):82-83. <https://doi.org/10.51987/revhospitalbaires.v41i2.92>.
8. Benson HR. An introduction to benchmarking in healthcare. *Radiol Manage*. 1994;16(4):35-39.
9. Grande Ratti MF, De Masi NS, Fedele GV, et al. Implementación del método 5S en salud: de la industria automotriz a los traslados intrahospitalarios. *Evid Actual Pract Ambul*. 2026;29(2):e007223. <https://doi.org/10.51987/evidencia.v29i3.7223>.
10. Breyer JZ, Giacomazzi J, Kuhmmer R, et al. Hospital quality indicators: a systematic review. *Int J Health Care Qual Assur*. 2019;32(2):474-487. <https://doi.org/10.1108/IJHCQA-04-2018-0091>.
11. Walsh KE, Razzaghi H, Hartley DM, et al. Testing the use of data drawn from the electronic health record to compare quality. *Pediatr Qual Saf*. 2021;6(4):e432. <https://doi.org/10.1097/pq.0000000000000432>.
12. Salamessy BH, Portrait FRM, van der Hijden E, et al. On the correlation between outcome indicators and the structure and process indicators used to proxy them in public health care reporting. *Eur J Health Econ*. 2021;22(8):1239-1251. <https://doi.org/10.1007/s10198-021-01333-w>.
13. Cossio-Gil Y, Pérez-Sádaba FJ, Ribera J, et al. Identifying potential predictable indicators for the management of tertiary hospitals. *Int J Health Plann Manage*. 2024;39(2):278-292. <https://doi.org/10.1002/hpm.3710>.
14. Nabovati E, Farrahi R, Sadeqi Jabali M, et al. Identifying and prioritizing the key performance indicators for hospital management dashboard at a national level: viewpoint of hospital managers. *Health Informatics J*. 2023;29(4):14604582231221139. <https://doi.org/10.1177/14604582231221139>.
15. Fallahnezhad M, Langarizadeh M, Vahabzadeh A. Key performance indicators of hospital supply chain: a systematic review. *BMC Health Serv Res*. 2024;24:1610. <https://doi.org/10.1186/s12913-024-11954-5>.
16. Gannon B, Jones C, Wakai A, O'Sullivan R. The cost of data collection for performance monitoring in hospitals: protocol for a systematic review. *Syst Rev*. 2014;3:65. <https://doi.org/10.1186/2046-4053-3-65>.
17. Amato L, Colais P, Davoli M, et al. Volumi di attività ed esiti delle cure: prove scientifiche dalla letteratura e dalle valutazioni empiriche in Italia. *Epidemiol Prev*. 2013;37(2-3 Suppl 2):1-100.
18. Grande Ratti MF, De Masi N, Garfi L, et al. A case report of a mixed-methods assessment of patient experiences to inform quality improvement in an emergency department in Argentina. *Int J Health Plann Manage*. 2025;40(3):620-629. <https://doi.org/10.1002/hpm.3906>.
19. Medina-Orozco I. La narrativa médica como herramienta educativa. *Rev Med Inst Mex Seguro Soc*. 2025;63(6):e6840. <https://doi.org/10.5281/zenodo.17342650>.
20. Kelley TA. International Consortium for Health Outcomes Measurement (ICHOM). *Trials*. 2015;16(Suppl 3):O4. <https://doi.org/10.1186/1745-6215-16-S3-O4>.

21. Seligman WH, Das-Gupta Z, Jobi-Odeneye AO, et al. Development of an international standard set of outcome measures for patients with atrial fibrillation: a report of the International Consortium for Health Outcomes Measurement (ICHOM) atrial fibrillation working group. *Eur Heart J*. 2020;41(10):1132-1140. <https://doi.org/10.1093/eurheartj/ehz871>.
22. Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an international standard set of value-based outcome measures for patients with chronic kidney disease: a report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD Working Group. *Am J Kidney Dis*. 2019;73(3):372-384. <https://doi.org/10.1053/j.ajkd.2018.10.007>.
23. Burns DJP, Arora J, Okunade O, et al. International Consortium for Health Outcomes Measurement (ICHOM): standardized patient-centered outcomes measurement set for heart failure patients. *JACC Heart Fail*. 2020;8(3):212-222. <https://doi.org/10.1016/j.jchf.2019.09.007>.
24. Bull C, Teede H, Watson D, et al. Selecting and implementing patient-reported outcome and experience measures to assess health system performance. *JAMA Health Forum*. 2022;3(4):e220326. <https://doi.org/10.1001/jamahealthforum.2022.0326>.
25. Gil Folgar ML, Grande Ratti MF, Perman G, et al. Resultados reportados por pacientes muy mayores tras implante valvular aórtico percutáneo. *Arch Peru Cardiol Cir Cardiovasc*. 2026;7(2):99-107. <https://doi.org/10.47487/apcyccv.7i2.634>.
26. Jamieson Gilmore K, Corazza I, Coletta L, et al. The uses of Patient Reported Experience Measures in health systems: a systematic narrative review. *Health Policy*. 2023;128:1-10. <https://doi.org/10.1016/j.healthpol.2022.07.008>.
27. Ghalavand H, Shirshahi S, Rahimi A, et al. Common data quality elements for health information systems: a systematic review. *BMC Med Inform Decis Mak*. 2024;24:243. <https://doi.org/10.1186/s12911-024-02644-7>.
28. Hooshafza S, Mc Quaid L, Stephens G, et al. Development of a framework to assess the quality of data sources in healthcare settings. *J Am Med Inform Assoc*. 2022;29(5):944-952. <https://doi.org/10.1093/jamia/ocac017>.
29. Abbassi F, Barkun J. Artificial intelligence in surgical outcomes reporting: the next best thing, or just artificially intelligent “garbage in, garbage out”? *Ann Surg*. 2025;282(6):897-899. <https://doi.org/10.1097/SLA.0000000000006789>.
30. Heenan MA, Randall GE, Evans JM. Selecting performance indicators and targets in health care: an international scoping review and standardized process framework. *Risk Manag Healthc Policy*. 2022;15:747-764. <https://doi.org/10.2147/RMHP.S357561>.
31. Van Vliet EJ, Bredenhoff E, Sermeus W, et al. Exploring the relation between process design and efficiency in high-volume cataract pathways from a lean thinking perspective. *Int J Qual Health Care*. 2011;23(1):83-93. <https://doi.org/10.1093/intqhc/mzq071>.
32. Arah OA, Westert GP, Hurst J, et al. A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care*. 2006;18(Suppl 1):5-13. <https://doi.org/10.1093/intqhc/mzl024>.
33. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health*. 2018;6(11):e1196-e1252. [https://doi.org/10.1016/S2214-109X\(18\)30386-3](https://doi.org/10.1016/S2214-109X(18)30386-3). Errata en: *Lancet Glob Health*. 2018;6(11):e1162. [https://doi.org/10.1016/S2214-109X\(18\)30456-X](https://doi.org/10.1016/S2214-109X(18)30456-X); *Lancet Glob Health*. 2021;9(8):e1067. [https://doi.org/10.1016/S2214-109X\(21\)00250-3](https://doi.org/10.1016/S2214-109X(21)00250-3).
34. Teisberg E, Wallace S, O'Hara S. Defining and implementing value-based health care: a strategic framework. *Acad Med*. 2020;95(5):682-685. <https://doi.org/10.1097/ACM.00000000000003122>.
35. Willmington C, Belardi P, Murante AM, et al. The contribution of benchmarking to quality improvement in healthcare: a systematic literature review. *BMC Health Serv Res*. 2022;22:139. <https://doi.org/10.1186/s12913-022-07467-8>.
36. Reponen E, Rundall TG, Shortell SM, et al. Benchmarking outcomes on multiple contextual levels in lean healthcare: a systematic review, development of a conceptual framework, and a research agenda. *BMC Health Serv Res*. 2021;21:161. <https://doi.org/10.1186/s12913-021-06160-6>.
37. Pronovost PJ, Berenholtz SM, Needham DM. Translating evidence into practice: a model for large scale knowledge translation. *BMJ*. 2008;337:a1714. <https://doi.org/10.1136/bmj.a1714>.
38. Neuman HB, Kaji AH, Haut ER. Practical guide to implementation science. *JAMA Surg*. 2020;155(5):434-435. <https://doi.org/10.1001/jamasurg.2019.5149>.
39. Cohen RI. Lean methodology in health care. *Chest*. 2018;154(6):1448-1454. <https://doi.org/10.1016/j.chest.2018.06.005>.
40. Niñerola A, Sánchez-Rebull MV, Hernández-Lara AB. Quality improvement in healthcare: Six Sigma systematic review. *Health Policy*. 2020;124(4):438-445. <https://doi.org/10.1016/j.healthpol.2020.01.002>.
41. Leis JA, Shojania KG. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. *BMJ Qual Saf*. 2017;26(7):572-577. <https://doi.org/10.1136/bmjqs-2016-006245>.
42. Dang A. Real-world evidence: a primer. *Pharmaceut Med*. 2023;37(1):25-36. <https://doi.org/10.1007/s40290-022-00456-6>.
43. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol*. 2022;22:287. <https://doi.org/10.1186/s12874-022-01768-6>.