RADIOLOGIA BASADA EN LA EVIDENCIA: ESTRATEGIA CONCEPTUAL FOCALIZADA PARA LA PRACTICA DE LA IMAGENOLOGIA

Drs. David Busel M^(1,2), Claudio Silva F⁽¹⁾

1. Servicio de Imagenología, Hospital Clínico Universidad de Chile.

2. Servicio de Imagenología, Clínica Alemana de Santiago.

Abstract: According Dodd and Stolberg, Evidence Based Radiology (EBR) fits into a humanistic approach to the practice of Radiology. This is a multiple step process, which includes: conversion of information needs into answerable questions, identification of the best available evidence, critical appraisal of the evidence and clinical expertise to integrate clinical circumstances with available research evidence and patients values and preferences. Main concepts of this article are starting to change the conventional practice. EBR can be implemented in rapidly progressing areas of radiology practice. Clinicians ask radiologists which test to do rather than telling us to "do the test".

Key words: Diagnostic test, Evidence based radiology, Likelihood ratio, Predictive values, Probability, Sensitivity, Specificity.

Resumen: La radiología basada en la evidencia (RBE) para algunos calza con un abordaje humanístico en su práctica que da cuenta de la responsabilidad del radiólogo hacia el paciente y la comunidad. La RBE incentiva utilizar los resultados de la investigación clínica sistemática en la práctica de una disciplina de mayor complejidad en que coexisten numerosas alternativas como herramientas en el diagnóstico. Por otra parte su práctica desincentiva el usar, solamente, la experiencia clínico radiológica no sistemática, la racionalidad de la patología radiológica y la intuición como recursos únicos para tomar decisiones en la selección e interpretación de las diferentes imágenes diagnósticas.

Este abordaje otorga herramientas que nos permiten con un proceso de varios pasos que incluye el convertir la necesidad de información en una pregunta factible de responder, la identificación de la mejor evidencia disponible, la apreciación crítica de la literatura respecto a la evidencia y su posterior integración a las circunstancias clínicas, ir desde la validez de la evidencia anecdótica, analógica o empírica resumidas como evidencias de un experto en una disciplina y desplazarnos hacia el uso de información obtenida por investigación sistemática relevante obtenida por ensayos clínicos y/o revisiones sistemáticas de estudios particulares de sus usos en órganos y sistemas.

Este artículo entrega conceptos básicos esenciales para comenzar a comprender su práctica y focaliza algunos puntos que se pueden explorar en la literatura de referencia.

Palabras Claves: Especificidad, Evidencia, Probabilidad, Radiología, Sensibilidad, Test diagnóstico, Valores predictivos.

Introducción

La radiología basada en la evidencia (RBE), según Stolberg, calza con un abordaje humanístico en su práctica y da cuenta de la responsabilidad del radiólogo hacia el paciente y la comunidad⁽¹⁾. Este abordaje y la necesidad de competencias para practicarla, han motivado este artículo dando cuenta en nuestro medio de una iniciativa que debiera optimizar y dar excelencia a nuestro trabajo.

La RBE incentiva utilizar los resultados de la investigación clínica sistemática para la práctica de nuestra especialidad en donde coexisten numerosas alternativas como herramientas de diagnóstico. De igual forma desincentiva el uso, de la experiencia clínico radiológica no sistemática, la racionalidad de la patología radiológica y la intuición; como recursos únicos para tomar decisiones en la selección e interpretación de las diferentes imágenes diagnósticas.

Busel D, Silva C. Radiología basada en la evidencia: Estrategia conceptual focalizada para la práctica de la imagenología. Rev Chil Radiol 2004; 10:109-117.

Correspondencia: dbusel@hotmail.com

El abordaje basado en la evidencia otorga herramientas que nos permiten, desde la validez de; la evidencia anecdótica dada por el reporte de casos, la evidencia analógica dada por una serie de casos o la empírica dada por la observación aislada de hechos resumidas como evidencias de un experto, desplazarnos hacia el uso de información obtenida por investigación sistemática relevante obtenida por ensayos clínicos y/o revisiones sistemáticas o metanálisis.

El practicar la radiología según estos preceptos nos llevara a usar en nuestros pacientes exámenes efectivos, cuando una amplia gama de test de imágenes se ofrecen para similares situaciones clínicas, y también evitará exponerlos a exámenes supuestamente eficaces, pero inútiles y que pueden causar daño por radiación⁽²⁾. Además podríamos estar generando gastos de recursos innecesarios para la sociedad y las personas. Por último se puede enfocar los resultados de la atención en efectos clínicos que sean deseados para una especial y singular condición.

RBE conceptos generales

La definición de RBE integra la mejor evidencia científica disponible con la experiencia clínica, en este caso la radiológica, considerando además valores y preferencias individuales en la toma de decisiones en el uso e interpretación de las imágenes médicas⁽³⁾. Requiere adquirir nuevas habilidades, que incluye un mayor rigor científico asociado a una eficiente búsqueda de información, apreciación crítica de la literatura (ACL), conocimientos básicos en bioestadística y manejo de las reglas de uso de la evidencia científica. El incorporarlas en las competencias y destrezas de los programas de radiología conlleva un desafío tanto para docentes, como residentes.

El entrenamiento sería de interés relevante focalizando objetivos educacionales, haciendo de los programas de radiología actuales más modernos e innovadores. El método de RBE se resume en cuatro pasos: 1) Asociación de la necesidad de información a la generación de una pregunta con respuesta factible. 2) Selección de la mejor evidencia para contestar esta pregunta a través de publicaciones pertinentes, usando buscadores, en bases de datos biomédicos. 3) Evaluación en forma crítica de la evidencia; como lo son su diseño, material y métodos, análisis de resultados; usando métodos de apreciación crítica de literatura (ACL). 4) Integración de la información crítica con nuestra especialidad, con los datos de nuestros pacientes y sus circunstancias exclusivas.

Si bien estas acciones serían las esenciales también son recomendables para adquirir destrezas dos más, la aplicación de la evidencia en la

selección de métodos diagnósticos e interpretación de sus resultados, y por último evaluar la efectividad en el uso de los pasos previos y mejorar ambas en forma continua.

La incorporación a través de la *WEB* de búsquedas electrónicas de artículos relevantes en revistas y bases de datos, la aplicación de las reglas de apreciación crítica de la literatura (ACL) creadas por Sackett (1985) con resúmenes estructurados y niveles de confiabilidad, las recomendaciones para la práctica definidas por Bandolier (1995) y Guyatt (2000), que han sido revisadas y ampliadas y donde se definen niveles y recomendaciones estructuradas por estos académicos de las Universidades de McMaster y Oxford; aplicadas al caso de los métodos diagnósticos (Tabla I), han colaborado en el desarrollo de esta práctica.

Una parte importante de la información de la validez de los resultados de los test de diagnósticos esta expresada en términos de Sensibilidad y Especificidad, valores predictivos positivos (VPP) y negativos (VPN), *odds* asociados a razones de probabilidades pre y post test, *Likelihood ratios* (LR) (Tabla II) y representaciones gráficas de probabilidad condicional como las curvas receptor operador (ROC) (Gráfico 1).

De acuerdo a la probabilidad de enfermedad y a las características del cuadro clínico, si los test se usan para descartar enfermedad o como cribado o tamizaje se seleccionan por su mayor sensibilidad o especificidad. Si se efectúan con miras de tratamiento y su uso permite confirmar enfermedad para determinar acción (Gráfico 2), los valores predictivos y probabilidades pre y post test, *odds* y LR, pasan a ser los relevantes.

Para conocer sus propiedades se recomienda conocer la frecuencia de la enfermedad (incidencia o prevalencia) en la población en que se obtiene la información de validez del test diagnóstico.

La información adicional de la confiabilidad, consistencia o precisión de un test es aceptada con mediciones de la reproducibilidad de sus resultados en estudios de variabilidad intra e interobservador y condiciones especiales de uso. Información de su seguridad, disponibilidad y costos son considerados hoy en día en su selección. Por otra parte muchos exámenes se utilizan solos o como parte de un grupo de exámenes de imágenes interconectados para cambiar la probabilidad de enfermedad y corresponde al uso de imágenes en serie o en paralelo (Gráfico 3).

En la búsqueda de la información, la evaluación de la validez de los resultados recae en la sección material y métodos de la cual cuatro preguntas que se deben plantear son resumidas en la Tabla III.

Dodd, MacEneaney y Malone⁽³⁾, en su artículo sobre RBE y cómo evaluar rápidamente la validez y fortaleza de las publicaciones de la literatura en

Tabla I. Niveles de evidencia.

NIVEL ESTUDIO DE TEST DIAGNOSTICO

- 1a Revisión sistemática (con homogeneidad) de estudios de test diagnósticos nivel 1; Reglas de decisión clínica con estudios nivel 1b de distintos centros clínicos
- 1b Estudio de cohorte de validación con buenos estándares de referencia o reglas de decisión clínica probada en un centro clínico
- 1c *SpPins* y *SnNouts* absolutos*
- 2a Revisión sistemática (con homogeneidad) de estudios de test diagnóstico nivel 2 o superior
- 2b Estudio de cohorte exploratorio con buenos estándares de referencia; regla de decisión clínica luego de derivación, o validado solo en muestras particionadas o base de datos
- 3a Revisión sistemática (con homogeneidad) de estudios nivel 3b o superior
- 3b Estudio no consecutivo o sin estándares de referencia de aplicación sistemática
- 4 Estudio caso control pobre o con estándar de referencia no independiente
- 5 Opinión de experto sin análisis crítico explícito o basado en fisiología, investigación de laboratorio o "principios básicos"

Traducido del Oxford Centre for Evidence-based Medicine Levels of evidence (Bob Phillips, Chris Ball, Dave Sackett, Dough Badenoch, Sharon Strauss, Brian Haynes, Martin Dawes)

**SpPins*: Test de alta especificidad, *SnNouts*: Test de alta sensibilidad.

Gráfico 1. Curva receptor operador (ROC).

radiología consideran otros puntos adicionales, que se resumen en la tabla IV. En este artículo, plantean una pregunta en relación al uso del test Colangiopancreatografía (CPRM), en la evaluación inicial de pacientes con obstrucción de la vía biliar, analizan la fortaleza de dos estudios para contestar esta pregunta. Con similar objetivo, reproduciremos algunos de estos conceptos y evaluaremos la validez de un ensayo clínico

Tabla II. Análisis estadístico de resultados

		ENFERMEDAD	
		Presente	Ausente
TEST	Positivo	a	b
	Negativo	c	d

VP= a; VN= d;

Sensibilidad= $a/a+c$; Especificidad= $d/b+d$;

VPP= $a/a+b$; VPN= $d/c+d$

LR (+)= $a/a+c$; $b/b+d$; LR (-)= $c/a+c$; $d/b+d$;

Prevalencia= $a+c/a+b+c+d$

VP: Verdaderos positivos; VN: Verdaderos negativos; VPP= Valor predictivo positivo; VPN= Valor predictivo negativo; LR: Likelihood ratios

Interpretación de la LR: 0: excluye enfermedad; infinito: excluye normalidad; >10 o <0.1 : gran cambio en la LR; 5-10 o 0.1-0.2: cambio moderado; 2-5 o 0.2-0.1: cambio pequeño; 1: no hay cambio.

realizado en nuestro medio para contestar esta misma pregunta, en la que CPRM se compara con ultrasonido biliar tradicional⁽⁴⁾.

Evaluando la fortaleza de la sección resultados

A.- Sensibilidad y Especificidad: La Sensibilidad (Sens) se puede definir como la probabilidad de que el test sea positivo si esta presente la enfermedad y la Especificidad (Spec) como la probabilidad de que

Grafico 2

Tabla III. Preguntas para la sección material y métodos:

- ¿Existió una comparación ciega independiente con el estándar de referencia del diagnóstico?
- ¿Fue el test de diagnóstico evaluado con un espectro apropiado de pacientes?
- ¿Fue la referencia estándar aplicada sin relación con el resultado del test de diagnóstico?
- ¿Fue el test o grupo de tests validado en una segunda oportunidad en otro grupo de pacientes?

Umbral de Exclusión

Umbral de Acción

0% 25% 50% 75% 100%

0%: Muy poco probable que tenga la Enf.; 25%: Probablemente no tiene la Enf.; 50%: No se sabe; 75%: Probablemente tiene la Enf.; 100%: Muy probable que tenga la Enf.
Enf.: Enfermedad

Grafico 3

Estrategia Secuencia de eventos Consecuencias

TEST EN SERIAL

Test A y test B y test C son positivos

Disminuye Sensibilidad; Aumenta Especificidad

TEST EN PARALELO

Test A o test B o test C es positivo

Aumenta Sensibilidad; Disminuye Especificidad

Tabla 4: Preguntas adicionales a considerar por el radiólogo:

Como radiólogos, sugerimos contestar las siguientes preguntas:

¿Ha sido descrito el método de imágenes con suficiente detalle para ser reproducido en su departamento?

¿Ha sido el test de imágenes evaluado y el test *gold* estándar desarrollado con el mismo estándar de excelencia?

¿Han sido generaciones de tecnología de desarrollo con la misma modalidad (CT convencional vs CT espiral vs CT multicorte) consideradas adecuadamente en el diseño del estudio en discusión?

¿Ha sido considerada la exposición a la radiación? El concepto de justificación y optimización ha asumido primera importancia en la protección de la radiación de los pacientes.

¿Han sido las imágenes MRI/CT revisadas en un monitor o en una placa? (Por ejemplo en CPRM una pequeña litiasis puede ser muy difícil de ver en una placa. Esta puede llegar a ser visible por cuidadosa manipulación del ancho y el nivel de ventana.

el test sea negativo si la enfermedad esta ausente (Tabla II). Ambos son críticos en la expresión de la exactitud de un test de diagnóstico y deben ser informados como pareja; si alguno falta o no se puede estimar de los resultados, se debe desconfiar de su validez. Si bien los puntos de corte para normalidad o definición de enfermedad pueden hacer variar su estimación, esto pasa a ser menos importante cuando alcanzan valores de un test fuerte o con buen desempeño (>95% de Sens/Spec). Así un test con alta Sens (Sn) con un test negativo (N) descartan efectivamente (*rule out*: descartar) un diagnóstico y se denomina *SnNout*, de igual manera un test con alta Spec (Sp) con un resultado positivo (P) efectivamente confirman (*rule in*: confirmar) un diagnóstico y se denomina un *SpPin*. Un ejemplo de *SnNout* se aprecia en el ensayo clínico de nuestro ejemplo⁽⁴⁾ con una Sens de 95.45% y una Spec de 87%, lo que hace a CPRM muy fuerte para descartar obstrucción pero en menor grado para confirmarla en este estudio.

En el artículo de Dodd⁽³⁾ se demuestra, como la Sensibilidad varia sustancialmente al cambiar el espesor de corte, comparando dos estudios para una misma técnica diagnóstica (CPRM). De manera similar la selección de pacientes o muestra y las características de los sujetos incluidos en el estudio pueden ser relevantes en el resultado final de estos parámetros; como por ejemplo determinación de casos leves o severos de la enfermedad, litiasis de mayor o menor tamaño o tumores pequeños o grandes etc. Todos ello debe estar descrito en la sección material/métodos, ya que la muestra puede hacer variar mucho estas estimaciones⁽⁵⁾.

Como vemos, Sensibilidad y Especificidad son estimaciones puntuales que varían de acuerdo al lugar, método, tecnología y pacientes, necesitamos por lo tanto otra medida de cuan cerca estas estimaciones están de la verdad y estos son los intervalos de confianza (CIs). Así se usa el 95% de los CIs para evaluar el rango de dispersión de estas estimaciones en la literatura médica. Mientras menor el rango más exacta es la estimación y mayor o más fuerte es la validez de los resultados. Este rango delimita de alguna manera no solo la validez interna o exactitud de los hallazgos para los pacientes estudiados, sino la validez externa que se refiere a la generalización de los resultados del estudio a otros pacientes. En nuestro ensayo clínico del ejemplo previo de CPRM *SnSpin* de 95.83%, el 95% de CIs es de 0.798-0.993 con un estrecho margen de rango de variabilidad reflejando sólidos resultados⁽⁶⁾.

La variable de más influencia en los CIs es el tamaño de la muestra, ya que a mayor tamaño más estrecho son los CIs y más significativos son los resultados del estudio. En general se deben

aplicar cuando la muestra es al menos razonable y esto estadísticamente corresponde a muestras no menor a 30 pacientes. El calcular un tamaño de muestra, asegura el adecuado poder del estudio para detectar diferencias significativas reales, no debidas solo al azar, al realizar estudios comparativos de desempeño.

B.- Información de la frecuencia de la enfermedad y probabilidad Pre/Post-test:

En general los radiólogos estimamos la probabilidad pre-test como baja, intermedia o alta asociando la probabilidad de un diagnóstico de enfermedad generalmente a la medida de frecuencia de la enfermedad en una determinada población (prevalencia o incidencia). En el estudio analizado se estimó en la población una frecuencia de enfermedad o prevalencia de obstrucción de un 37.5%. La probabilidad (pre-test) en cambio, es definida como una opinión acerca de la chance de un evento (*odds*) antes de recibir nueva información de un test de diagnóstico a ser considerada.

Los valores predictivos positivo (VVP) y negativo (VPN) expresan la estimación próxima de un valor de la probabilidad post-test dado el resultado del test para detectar la enfermedad que en realidad es una probabilidad condicional. El valor predictivo positivo corresponde a la probabilidad de que con un resultado positivo del test, este efectivamente sea correcto, es decir, el paciente tiene la enfermedad. El valor predictivo negativo, se refiere al grado de exactitud de la certeza con que con un test negativo el paciente efectivamente esté libre de enfermedad⁽⁷⁾.

En nuestro ejemplo⁽⁴⁾, la evaluación de nuestro test CPRM posee un VPP de 82% y VPN de 97% para obstrucción. La crítica a estas estimaciones es la gran expresión de variabilidad de la prevalencia en diferentes estudios y no poder tampoco estratificar de acuerdo al riesgo clínico en un mismo estudio. Los valores predictivos si bien parecen similares a la probabilidad posterior al test en realidad no lo son, ya que son obtenidos en una población definida de un estudio, son definidos como una proporción de pacientes con un resultado de un test que posee o no enfermedad, se calculan con una tabla de 2x2 y sus valores dependen del cambio de prevalencia de la enfermedad en la población en estudio.

Si bien esta aproximación, permite inferir elementos de precisión del test analizado, su aplicación es poco práctica, ya que son dependientes de la prevalencia de enfermedad en la población estudiada. Al bajar la proporción de pacientes portadores de la enfermedad en el grupo estudiado, el valor predictivo positivo va a ser menor, y el valor predictivo negativo, tenderá a aumentar. Es más fácil de entender aún, si consideramos la prevalencia como la probabilidad de una persona de presentar dicha

enfermedad por el solo hecho de entrar a un hospital, obviamente, si esta probabilidad es baja, la posibilidad que el test que se le aplique sea negativo es mayor solamente por azar. Dado que habitualmente la prevalencia de enfermedad de la población general es baja, podemos tener un test con buena Sensibilidad y Especificidad pero con muy bajos valores predictivos. Además, no permiten una comparación válida de estudios realizados en distintos centros con distintas prevalencias de cualquier patología en estudio.

La probabilidad post-test en cambio es definida como la probabilidad de enfermedad después de que nueva información del test es tomada en cuenta, (post-test odds), se estima con el teorema de Bayes y depende de la probabilidad pre-test de enfermedad ya definida. La probabilidad de un evento (odds) es definida como $P/1-P$, donde P es la probabilidad del evento en este caso de enfermedad y $1-P$ la probabilidad de no enfermedad (Gráfico 4).

Gráfico 4.

El uso de los *Likelihood ratios* (LRs) tiene tres poderosas razones diagnósticas, la primera tiene relación con su manera vertical de cálculo, en que en forma similar a la Sensibilidad y Especificidad no necesitan variar significativamente con los cambios de la prevalencia de la enfermedad a diferencia de los valores predictivos. Incluso tienen mejor estabilidad que la Sensibilidad y Especificidad con los cambios de prevalencia. Segundo pueden ser estimados para diferentes niveles de signos, síntomas o resultados del test y no solo en dos niveles fijos como se obtienen en una tabla de contingencia de 2x2, incrementando así significativamente la información diagnóstica obtenida. De esta manera los LRs subdividen de alguna manera a los pacientes sometidos al test en diferentes categorías y distingue entre ellos el comportamiento del test muy bien.

La última de sus propiedades es que pueden ser usados para acotar nuestra lista de hipótesis diagnósticas porque:

La probabilidad (odds) Pre-test de la enfermedad multiplicado por Likelihood ratio para el resultado del test (LR) = Probabilidad (odds) Post-test

Donde LR es una probabilidad condicional definida como:

$$LR = \frac{\text{Frecuencia de resultado de un test en enfermos}}{\text{Frecuencia de resultado de test en no enfermos}}$$

Para una fácil interpretación del resultado del test existe una representación gráfica, el Nomograma de Fagan, usado para interpretar los LRs donde trazando una línea entre la probabilidad pre-test que generalmente se asocia a la prevalencia de la enfermedad y el LR positivo o negativo para enfermedad se obtiene en el otro extremo la probabilidad post-test. El LR de un test negativo LR (-) cae entre 0 y 1, y el LR de un test positivo LR (+) cae entre 1 e infinito. Mientras más cerca de 0 cae un test negativo, más fuerte o exacto es el test e inversamente mientras más alto el valor del LR de un test positivo, posee mayor validez. Un test positivo con un LR de 1 es completamente inútil.

En nuestro ejemplo⁽⁴⁾ de CPRM podemos evaluar los resultados para pacientes con una probabilidad de obstrucción biliar pre-test de 37.5%, así estimando LR(+) de 7.7, la probabilidad post-test para CPRM positiva se incrementa a casi un 80% y para un test negativo con LR(-) de 0.05 baja a un 3%. En la estratificación por riesgo vemos que en pacientes de alto riesgo clínico con probabilidad pre-test en ese grupo de 48%, la probabilidad post-test de un test positivo se eleva a un 85% y con test negativo baja a 0.5%. Ahora para pacientes con baja probabilidad pre-test (bajo y moderado riesgo clínico) con una probabilidad pre-test o prevalencia de 7.8% con un test negativo, esta desciende aproximadamente a un 0.05% (Gráfico 5a). Estos resultados son suficientemente fuertes para excluir enfermedad si el test es negativo o elevar esta a un 87% de acuerdo a su resultado positivo (ver puntos de corte o *thresholds* en Gráfico 1) para determinar acción o tratamiento o para determinar exclusión de enfermedad^(3,7,8).

En este mismo estudio el otro test de imágenes en evaluación, ultrasonido (US), se desempeña con menor exactitud en pacientes con probabilidad pre-test de 37.5%. Con un LR(+) de 3.3, la probabilidad post-test sube a menos de 70% si el test es positivo (no determina acción) se debe solicitar más información y con LR(-) de 0.59 disminuye a solo casi un 30% si el test es negativo, valor insuficiente para excluir enfermedad.

Gráfico 5a.

Ahora si discriminamos y tomamos pacientes de probabilidad pre-test intermedia (alto riesgo) en este caso con prevalencia de obstrucción de un 48% en nuestro estudio con un test positivo la probabilidad post-test sube a 75% suficiente para determinar acción. Por otra parte si el test es negativo en este grupo, la probabilidad pos-test es de aproximadamente un 4% que podría ser suficiente también para excluir enfermedad (Gráfico 5b).

Gráfico 5b.

Así entonces, en los pacientes con baja probabilidad de enfermedad (moderado y bajo riesgo) y con una probabilidad pre-test baja de 7.8% la probabilidad post-test baja a 0.5% si el test es negativo siendo aún mas sólido para descartarla. Por otra parte si es positivo la probabilidad post-test sube solo a 20%, entonces desconfiando de sus resultados por ser insuficientes para descartar, así solo sirven para determinar cambio y otro estudio de imágenes debe ser indicado en este caso el

recomendado sería CPRM de mejor desempeño o exactitud diagnóstica.

Por otra parte si consideramos nuestra definición, un uso juicioso de los recursos y del tiempo para obtener información usando test en serie, la norma general es usar primero el test más específico de modo de no clasificar a pacientes sanos como enfermos. Ahora en nuestro caso ejemplo, dada la estrecha variación en el comportamiento entre las especificidades y los VPN de ambos test, un uso racional de US inicial, para descartar enfermedad en pacientes solo con moderado y bajo riesgo clínico con una baja probabilidad pre-test en nuestro estudio (7.8%), podría ser costo/efectivo y luego CPRM de acuerdo a su resultado. Por otra parte si el paciente posee alto riesgo clínico con probabilidad intermedia de 48%, el uso de CPRM sería de mayor ventaja por su mejor desempeño como test de mayor validez y exactitud para confirmar o descartar el diagnóstico de obstrucción.

Discusión

Como vemos en el ejemplo usado si bien CPRM es el test de elección en el estudio por su exactitud para confirmar la presencia de obstrucción biliar, la exactitud de ultrasonido negativo en pacientes con baja probabilidad pre-test inicial de obstrucción sería recomendable en nuestro medio. Esta última recomendación sería aún más válida si el paciente no ha sido colecistectomizado. En este caso en particular la evidencia es de tipo 1b y 1c para CPRM y de tipo 1b para ultrasonido y obstrucción, con grado A de recomendación para CPRM y B para ultrasonido (Tabla I).

El cambio de paradigma científico en la práctica, se apoya en el análisis de decisiones que están también basados en los riesgos, beneficios y efectividad, disponibilidad y costos de cada técnica de diagnóstico por imágenes en diferentes situaciones clínicas (ej: juzgar la severidad o estadio de la enfermedad, establecer su curso clínico y pronóstico, probable respuesta a la terapia actual o futura), para evaluar la mejor alternativa. En este rol el radiólogo pasa a ser una especie de interconsultor de apoyo con elementos de juicio consciente y de aporte de la evidencia necesaria en el estudio de imágenes de sus pacientes. En esta determinación, el radiólogo debe estar atento al evaluar evidencia de diferentes estudios de los potenciales sesgos involucrados en la selección, medición, análisis y publicación de resultados que pueden o no ser relevantes en un cambio de práctica o selección de un test⁽⁵⁾.

Como vemos la RBE, se refiere a usar la investigación existente y tiene como objetivo mejorar la calidad de atención y elevar los estándares de su entrega, considera el uso adecuado de los recursos y se focaliza en rango de efectos, insistiendo

explícitamente en puntos o efectos clínicos deseados, además es claramente una herramienta para entregar educación.

Un ejemplo de esta focalización en efectos clínicos deseados es la evaluación de ultrasonido con comprensión graduada en el estudio de pacientes con sospecha de apendicitis aguda realizado en la Universidad de Newcastle, Australia⁽⁹⁾. En este estudio la Sensibilidad y Especificidad de este examen en el diagnóstico de apendicitis es de 94.7% (casi *SnNout*) y 88,9% respectivamente. Este ensayo clínico destaca el valor significativo del uso del ultrasonido en el acortamiento del período previo a cirugía y por otra parte entrega la posibilidad de otros diagnósticos relevantes. En este estudio, vemos igualmente la nula o escasa evidencia para otros efectos de su uso previamente aludidos para indicarlo, tales como una mayor tasa de laparotomías no terapéuticas, perforación o cambios del período de hospitalización. En este sentido este tipo de estudios demuestran exactamente la utilidad del test (si los pacientes mejoran más o no de acuerdo a si se usa o no el test y por que variables) y no solo dan cuenta de su grado de exactitud diagnóstica. El adecuado manejo de información de las ventajas y desventajas, así como el de las indicaciones y contraindicaciones de un test son relevantes a la hora de recomendarlos.

Una importante iniciativa en como la información publicada en la evaluación de la exactitud y selección de test diagnósticos, corresponde al grupo o iniciativa *STARD* (Standards for Reporting of Diagnostic Accuracy), que corresponde a estándares para informes de exactitud diagnóstica liderados por el departamento de Epidemiología Clínica de la Universidad de Amsterdam, quienes proponen un diagrama de flujo y una lista limitada de requisitos para la publicación de literatura de validez y exactitud de test de diagnósticos⁽⁹⁾. Otra iniciativa ejemplo de la evaluación de efectos objetivos deseados en la exactitud de las imágenes esta en radiología oncológica con el uso de los llamados criterios *RECIST* para unificar criterios de respuesta de tratamientos oncológicos de tumores sólidos y que facilitarán nuestra práctica disminuyendo el número de controles de imágenes. El uso inadecuado de numerosos exámenes con radiaciones médicas en la búsqueda, seguimiento y control de las enfermedades constituyen ya un problema de preocupación médica, por su asociación con cáncer, y con tratamientos, que pudieran ser disminuidos al optimizar su uso⁽¹⁰⁻¹²⁾.

La RBE, también puede ser vista entonces como una forma de racionalizar y detener el uso de test de diagnósticos y tratamientos de eficacia no probada, aun cuando la experiencia clínica, sugiera

que pueden beneficiar a un paciente. En efecto, intenta mejorar la calidad de atención probando que intervenciones efectivas son poco usadas y mostrando que procedimientos inefectivos son muy usados.

Algunas objeciones se han publicado en relación con su práctica, como por ejemplo, que exige un compromiso de tiempo y recursos, distraídos de los disponibles para la atención de salud, el que no todos los médicos poseen destreza en localizar la información y usar computadores y que se requieren recursos para comprar bases de datos y mantenerlas. Stolberg y Malone en St. Vincent Hospital en Irlanda, han demostrado el uso de software de dominio público que se bajan de Internet y se usa en PDA(s), Notebooks, etc., ellos ayudan en las estimaciones de estos parámetros y facilitan su uso y práctica. Estos autores muestran ejemplos prácticos y concretos para enseñar esta técnica lo que se traduce en un mejor uso, sabido es que la práctica hace al artesano (<http://www.evidencebasedradiology.net>).

No existe duda que nuevas metodologías y objetivos educacionales en los programas de residencia de radiología darán cuenta de este abordaje, existiendo precedentes en muchos lugares tales como en las Universidades de McMaster, Oxford, Rochester y St Vincent y en que ya se aplican en numerosas preguntas para radiología diagnóstica e intervencional⁽¹³⁻¹⁶⁾.

Conclusión

La radiología no está exenta, al igual que otras disciplinas médicas, de evolucionar no solo de acuerdo al avance tecnológico, sino también en el uso racional, juicioso y humanitario de estos avances. El conocimiento científico en este contexto y el abordaje de la RBE en particular colaborarán en identificar usos e indicaciones, ventajas y desventajas entre estas. Por otra parte el informe radiológico dará cuenta de las fortalezas y debilidades así como de las posibles limitaciones de estos métodos en circunstancias particulares y del grado de evidencia para apoyar o refutar hipótesis de diagnóstico de diferentes alternativas en el contexto del paciente. Así el incentivar el uso responsable en las nuevas generaciones de radiólogos de estos avances, indudablemente llevará a una mayor excelencia del ejercicio de nuestra disciplina en favor de la sociedad. Los médicos clínicos optarán en un futuro cercano, por solicitar ayuda en decidir que examen de imágenes deben pedir en vez de solicitar el realizar un examen. Esta visión posiciona al radiólogo como interconsultor y voz de una segunda opinión versus un ejecutor técnico, en la práctica médica.

Bibliografía

1. Stolberg H. Focus on Concepts: Evidence-Based Radiology (EBR) SY 3.5.1 H.O. McMaster University, Hamilton, On, Canada.
2. Berrington de Gonzalez A, Darby S. Risk of cancer from diagnostic X-Rays: estimates for the UK and 14 other countries. *Lancet* 2004; 363: 345-51.
3. Dodd JD, MacEneaney PM, Malone DE. Evidence-based Radiology: how to quickly assess the validity and strength of publications in the diagnostic radiology literature. *Eur Radiol* 2004; 14: 915-922.
4. Busel D, Pérez L, Arroyo A. y col. Colangiografía (CPRM) vs ultrasonido (US) focalizado en pacientes con ictericia o sospecha de obstrucción de la vía biliar. Resultados preliminares. *Rev Chil Radiol* 2003; 9: 173-181.
5. Begg CB. Biases in the Assessment of Diagnostic Tests. *Stat Med* 1987; 6: 411-423.
6. Hanley JA. The Place of Statistical Methods in Radiology (and in the Bigger Picture). *Invest Radiol* 1989; 24: 10-16.
7. Phillips WC. Jr., Scott JA. Medical Decision Making: Practical Points for Practicing Radiologists. *Am J Roentgenol* 1990; 154:1149-1155.
8. Sox HC Jr., Blatt MA, Higgins MC, Marton K Ed. Butterworths. Medical Decision Making. Capítulos 3 y 4, Pag. 27, 101.
9. Douglas CD, Macpherson NE, Davidson PM, Gani JS. Randomised controlled trial of ultrasonography in diagnosis of acute appendicitis, incorporating the Alvarado score. *BMJ* 2000; 321: 919-22.
10. Bossuyt PM, Reitsma JB, Bruns DE. Et al. Towards complete and Accurate Reporting of Studies of Diagnostic Accuracy; The STARD Initiative. *AJR Am J Roentgenol* 2003; 181: 51-55.
11. Therasse P, Arbuck SG, Eisenhauer EA. Et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000; 92: 205-16
12. Borriell Z, Houghton C, Sullivan PJ, Sestini P. Retrospective analysis of evidence base for tests used in diagnosis and monitoring of disease in respiratory medicine. *BMJ* 2003; 327:1136-38.
13. Sackett, Straus, Scott, Rosenberg, Haynes. Medicina basada en la Evidencia. Cómo practicar y enseñar MBE. Harcourt Segunda Edición 2000, Churchill Livingstone.
14. Panzer RJ, Griner PF. Teaching quantities approaches to the use of diagnostic tests and procedure. *Int. J. of technology Assessment in Health Care* 1987; 3: 27-38.
15. Dodd J, Stolberg H, Malone D, Staunton M. EBR in Practice: Practical Examples. (SY3.5.3), St.Vincent's University Hospital, Dublin, Ireland; McMaster University, Hamilton, Ont, Canada.
16. The Evidence Based Radiology Working Group. Evidence Based Radiology: A New Approach to the Practice of Radiology. *Radiology* 2001; 220:566-575. RSNA 2001.