

A pesar de los continuos avances en Medicina, la toma de decisiones clínicas cualificadas sigue siendo la pieza clave de la práctica médica. El objetivo de la toma de decisiones médicas es optimizar la salud y la asistencia sanitaria del paciente. Esta optimización precisa la integración y aplicación de habilidades cognitivas, saber médico y juicio con sensibilidad social, cultural e interpersonal. La toma de decisiones médicas es un proceso con múltiples pasos, que comprende identificación del problema, selección y evaluación de información diagnóstica y elección de intervenciones. La toma de decisiones médicas supone un reto. Las decisiones médicas conllevan riesgos elevados y se caracterizan por información imperfecta y problemas complicados con diferente presentación y respuesta al tratamiento. La información médica es abundante, de distribución generalizada, en cambio constante, a menudo contradictoria y de validez y fiabilidad inciertas. A menudo no se dispone de información importante. Toda la información clínica tiene valor diagnóstico, aunque imperfecto y sujeto a error, con una validez y fiabilidad imperfectas. Por tanto, todas las decisiones médicas están sujetas a un cierto grado de incertidumbre. Paradójicamente, el tremendo progreso de la ciencia médica, aunque ha aumentado nuestra capacidad para diagnosticar y tratar enfermedades y discapacidades, ha provocado que el proceso de toma de decisiones sea más difícil y complicado.

La toma de decisiones médicas es todavía más compleja por la diversidad de la enfermedad entre sistemas biológicos y por las preferencias y valores entre individuos, por la incertidumbre de la información médica y por la escasez de recursos, que obliga a una asignación justa y a valorar los equilibrios entre beneficios, riesgos y costes. La tarea de tomar decisiones se torna más complicada por factores sociales, organizativos, económicos y ambientales, como la necesidad de incorporar las perspectivas, preferencias, valores y necesidades de los pacientes, familias y proveedores, por unos sistemas de suministro y financiación muy fragmentados, por unos sistemas de información clínica inadecuados y por unos incentivos equivocados.

El objetivo de este capítulo es proporcionar un marco básico para abordar, evaluar y organizar la información médica. El capítulo expone los aspectos psicológicos cognitivos, conductuales y ambientales de la toma de decisiones, la evaluación y el uso de las pruebas diagnósticas y la evaluación y selección de las prácticas e intervenciones médicas. En el capítulo 2 se expone la toma de decisiones relacionadas con las limitaciones de los recursos y su coste-efectividad, un componente de relevancia creciente.

## **Influencias cognitivas, conductuales y ambientales en la toma de decisiones**

### **Modelos de toma de decisiones (editar)**

#### **Razonamiento heurístico**

La heurística emplea “atajos” para ayudar a afrontar las limitaciones cognitivas y psicológicas y reducir la complejidad de la información a un nivel controlable. Al aumentar la experiencia, la obtención e interpretación de los datos y la acción se hacen más selectivas y automáticas. Aunque generalmente útil, puede provocar errores previsibles en la toma de decisiones.

La información se interpreta a la luz de la experiencia previa y del contexto en que se produce. Las percepciones están influenciadas de forma selectiva e intensa por las creencias, experiencias y expectativas. El conocimiento y la experiencia pueden acentuar la repercusión de nociones preconcebidas sobre las observaciones (p. ej., los mismos sonidos cardíacos pueden interpretarse de forma diferente según lo que espera oír el

médico). La formulación e interpretación del problema están influenciadas por la forma de elaborar la información (es decir, la ordenación o formateado, uso de frases cargadas de valor, episodios presentes como ganancias o pérdidas). Por ejemplo, la elaboración de una alternativa como intento de “ayudar al paciente” frente a “no hacer nada” sesga la decisión a favor de la intervención.

La toma de decisiones está influenciada por factores y limitaciones psicológicas y cognitivas inherentes a la condición humana. La capacidad de memoria a corto plazo es limitada (la media de una persona es de siete alternativas a la vez). La memoria es dependiente del contexto y puede fallar. Más que representar un recuerdo preciso de experiencias o episodios pasados, las memorias se reconstruyen a partir de piezas de información almacenadas por separado, olvidando detalles obtenidos por inferencia lógica, lo que da lugar a una nueva memoria combinada. Las decisiones médicas pueden estar distorsionadas por “disonancia cognitiva”, alineando las creencias con la conducta (bien antes o después del hecho) para reducir o eliminar las incoherencias y contradicciones psicológicas. Por tanto, antes de embarcarse en decisiones médicas importantes, los médicos deben considerar su grado de motivación por creencias o expectativas previas para interpretar los resultados y cómo los mismos resultados pueden interpretarse de otro modo en ausencia de dichas motivaciones y expectativas.

Los médicos intentan con frecuencia clasificar las circunstancias de los pacientes en base a su analogía con las circunstancias de otros pacientes (representatividad). Aunque a menudo resulta útil para formular diagnósticos o estimar la respuesta al tratamiento, la representatividad heurística puede, mediante una concentración inadecuada en componentes engañosos o no informativos de las manifestaciones clínicas, conducir a una estimación inadecuada de las probabilidades (p. ej., la subestimación frecuente de la probabilidad de isquemia aguda en mujeres con dolor torácico), a no tener en cuenta la prevalencia de posibilidades alternativas, insensibilidad al tamaño de la muestra (otorgando la misma importancia a estudios pequeños o amplios, aunque las observaciones basadas en muestras extensas deben engendrar más confianza que las derivadas de muestras de pequeño tamaño), concepciones erróneas de la probabilidad (p. ej., asumir que si se produce un acontecimiento adverso infrecuente no volverá a aparecer en un futuro próximo), regresión a la media (p. ej., falta de reconocimiento de que los pacientes tienden a acudir cuando los síntomas empeoran y atribuir así de forma equivocada un beneficio terapéutico a la evolución normal de los síntomas) e infracorrección de la correlación (p. ej., sobrestimación del valor de las pruebas diagnósticas que miden información con correlación estrecha).

De forma similar, existe una tendencia a estimar la probabilidad de un episodio por la facilidad con que aparece (disponibilidad), lo que sobrestima la probabilidad de episodios que resultan familiares, destacados, fácilmente imaginables o recientes (p. ej., la tendencia a evaluar la probabilidad de que un paciente responda a un antihipertensivo concreto basándose en la experiencia propia de las últimas prescripciones de dicho medicamento más que en los resultados de estudios clínicos rigurosos y extensos) o intentar predecir episodios que son imprevisibles de forma inherente (correlación ilusoria).

La información diagnóstica se obtiene en una secuencia temporal, con revisión de las estimaciones de probabilidad iniciales conforme se obtiene más información. Sin embargo, las estimaciones de probabilidad iniciales (incluso cuando están basadas en información incompleta) se ajustan habitualmente de forma insuficiente a la información subsiguiente (ajuste y anclaje). Por ejemplo, la información inicial apunta a un dolor torácico no isquémico (mujer premenopáusica) seguida de información más sospechosa de cardiopatías isquémica (dolor torácico subesternal con el ejercicio) puede

llevar a una subestimación de la probabilidad de isquemia coronaria, que no se produciría si el orden de presentación fuera el inverso. La tendencia a observar los episodios que ya han pasado como relativamente inevitables y obvios (sesgo retrospectivo) distorsiona también la estimación de la probabilidad al asignar una mayor probabilidad previa al episodio después del hecho que cuando se estima antes del episodio.

### **Reducción de errores psicológicos y heurísticos**

Se producen errores cognitivos en todas las fases del razonamiento diagnóstico: generación de hipótesis, recogida y procesado de la información, diagnóstico de los resultados y tratamiento. La mejor manera de evitar o reducir estos errores es estar alerta sobre su existencia y tomar medidas conscientes para subsanarlo. Así, los médicos deben evitar la tendencia a desarrollar conjuntos de investigación limitados mediante una consideración consciente de perspectivas alternativas y ampliando la consideración al intervalo de posibles alternativas diagnósticas y terapéuticas, reducción de la incertidumbre de las predicciones aumentando los intervalos de confianza, manteniendo más tiempo las hipótesis alternativas, sustituyendo las hipótesis rechazadas por nuevas hipótesis (en lugar de reducir las opciones conforme se obtiene información adicional) y buscando y reconociendo información relacionada, buscando atentamente pruebas que apoyen los diagnósticos apropiados o que descarten las hipótesis de trabajo, sobre todo en circunstancias infrecuentes pero catastróficas, mantenimiento de registros precisos y considerando de forma explícita por qué o cómo podrían ser diferentes los resultados. Los errores en la toma de decisiones pueden reducirse también mediante la identificación de errores estadísticos frecuentes, como los menores intervalos de confianza presentes en los estudios más extensos y, cuando hay más experiencia, utilizando la regresión a la media y distinguiendo entre relación y causalidad, así como entre significación clínica y estadística (como puede suceder en los estudios muy extensos, en los que pequeñas diferencias pueden ser estadísticamente significativas pero carecer de relevancia clínica).

### **Influencias conductuales y ambientales**

La toma de decisiones médicas suele presentarse como una secuencia lógica y racional, dirigida mediante una evaluación desapasionada del conocimiento científico. No obstante, las decisiones médicas están influenciadas por distintos factores conductuales, sociales, organizativos y ambientales.

Las personas son organismos sociales de forma inherente. Obtenemos pistas conductuales y estamos preocupados de cómo somos vistos por los demás. Cada vez más, las decisiones médicas se toman en conjunto con otras personas. Sin embargo, la valoración y la toma de decisiones en grupo son propensas a los sesgos, como una consideración insuficiente de la información sorprendente o contraria, lo que provoca un exceso de confianza y juicios y conducta de autorrefuerzo (atribución del error al grupo) y una polarización de la opinión. De forma característica, los grupos rinden más que la media de sus miembros, pero no tan bien como el mejor miembro. Los errores atribuibles al juicio y la toma de decisiones en grupo pueden disminuirse mediante una “tormenta de ideas” de manera independiente más que en grupo y, dentro de los grupos, alentando la disidencia y las críticas, evitando alegar las preferencias personales desde el principio y actuando como “abogado del diablo”. En las decisiones especialmente importantes, la toma de decisiones puede mejorarse mediante la creación de grupos independientes para considerar el mismo problema e invitando a expertos externos y a

colegas cualificados a desafiar el acuerdo del grupo.

Las características del médico (edad, conocimientos, formación, actitud, experiencia, práctica y organización, localización geográfica, integración social, rasgos de la personalidad y estilo de decisión general como actitud ante el riesgo) influyen en la decisión, al igual que la forma de comunicar la información (características del mensaje, el mensajero y el medio y contexto en el que se comunica). Aunque la información médica se transmite por diferentes canales (revistas y libros, cursos de formación continuada, sociedades profesionales, anuncios, medios de comunicación), los líderes de opinión son los más influyentes. También influyen en las decisiones las propias características de la información (p. ej., coste, riesgo, capacidad de aplicación, facilidad de reversibilidad).

La formación (incluidas las directrices de práctica clínica) tienen una gran influencia en la toma de decisiones. No obstante, la aceptación y credibilidad de la formación dependen del patrocinador, el formato y el contexto, así como de la formación y el entrenamiento previo. Los esfuerzos de formación pueden verse distorsionados o deteriorados también si entran en conflicto con otros incentivos, estilo de práctica, reglas o normas y valores locales y profesionales.

El ambiente en el que se practica la medicina también influye mucho en la toma de decisiones médicas. La estructura, proceso y regulación administrativa pueden influir en los médicos en la toma de decisiones mediante la creación o destrucción de barreras y una provisión adecuada en el tiempo de pistas, recordatorios, retroalimentación y apoyo. Sin embargo, a menos que sean especialmente potentes o incómodos, estos factores se superan fácilmente y eliminan con rapidez. La retroalimentación de resultados creíbles y relevantes, sobre todo si proceden de fuentes fiables y se producen de forma secuencial (p. ej., recordatorios computarizados generados por expertos en el momento del servicio), puede influir mucho en las decisiones médicas. Los incentivos económicos influyen de forma notoria en la toma de decisiones y la práctica, pero son difíciles de calibrar, por lo que deben ser coherentes con el conocimiento, los valores y las normas profesionales, sociales, culturales y organizativas. Por tanto, un amplio abanico de factores ejercen una influencia considerable en las decisiones médicas, sobre todo cuando están en línea con otros y se refuerzan con el tiempo.

### **Toma de decisiones diagnósticas**

Las pruebas diagnósticas definen problemas médicos con precisión suficiente para guiar el tratamiento de los pacientes. La utilidad de las pruebas diagnósticas depende de la aportación de nueva información que afecta de forma beneficiosa al proceso diagnóstico, el tratamiento clínico y el resultado para la salud. La información nueva tiene poco valor si no cambia la asistencia o el estado de salud de los pacientes.

La interpretación y el uso de pruebas diagnósticas están determinados por un objetivo clínico, la capacidad de dicha prueba para discriminar entre la presencia y ausencia de enfermedad y la seguridad, eficacia, efectividad y coste-efectividad de la prueba y del tratamiento.

Para cada circunstancia clínica existe una probabilidad de enfermedad por encima de la cual está indicada la intervención, “el umbral de tratamiento de la prueba” y una probabilidad por debajo de la cual puede descartarse la enfermedad con suficiente confianza y deben considerarse otras posibilidades diagnósticas, “el umbral de no tratamiento de la prueba”. Cuando la probabilidad de enfermedad se encuentra entre estos dos umbrales, las pruebas diagnósticas, en conjunto o por separado, son útiles

según cómo se comporta la probabilidad por encima del umbral de tratamiento o por debajo del umbral de no tratamiento de la prueba. Los umbrales de prueba y de tratamiento se determinan mediante las características de rendimiento de la prueba diagnóstica (sensibilidad y especificidad) y los riesgos, costes y beneficios (mortalidad, morbilidad y calidad de vida) de las opciones terapéuticas.

Los tratamientos menos eficaces, con mayor riesgo o más caros aumentan tanto el umbral de tratamiento como el de no tratamiento de la prueba (es necesaria una mayor certeza diagnóstica para comenzar el tratamiento y una menor certeza para descartar la enfermedad). A la inversa, cuanto más eficaz, segura y barata sea la prueba, es necesaria una menor certeza diagnóstica antes de comenzar el tratamiento o la prueba (menor umbral de tratamiento y menor umbral de no tratamiento de la prueba). Un menor riesgo y coste de la prueba y una mayor sensibilidad y especificidad reducen el intervalo de probabilidad en el que es apropiada (mayor umbral de tratamiento y menor umbral de no tratamiento de la prueba), mientras que un mayor coste y riesgo de la prueba y una menor sensibilidad y especificidad reducen el intervalo de la prueba (menor umbral de tratamiento y mayor umbral de no tratamiento de la prueba). Cuando más grave es la enfermedad (mayor morbimortalidad), menor es el umbral de la prueba y del tratamiento.

### **Rendimiento de la prueba diagnóstica**

Aunque las pruebas diagnósticas deben ser precisas y fiables (ofrecer un resultado coherente) y válidas y exactas (ofrecer un resultado correcto), el indicador principal del rendimiento de una prueba diagnóstica es la capacidad para distinguir entre la presencia y la ausencia de una enfermedad. Una prueba diagnóstica ideal discrimina perfectamente las personas con enfermedad (positivos verdaderos PV) y sin enfermedad (negativos verdaderos NV).

Sin embargo, las pruebas diagnósticas detectan marcadores correlacionados de forma imperfecta con la presencia o ausencia de enfermedad. Por esta razón, en su forma dicotómica más simple (interpretada como positiva o negativa) se producen dos tipos de errores: personas sin enfermedad se consideran de forma incorrecta como enfermas (positivos falsos PF) y personas con la enfermedad se consideran de forma incorrecta como sanos (negativos falsos NF).

### **Sesgos en la evaluación del rendimiento de una prueba diagnóstica**

La evaluación de una prueba diagnóstica se ve alterada con frecuencia por sesgos que tienden a sobrestimar su rendimiento.

Cuanto mayor sea la discrepancia entre las condiciones del estudio y el contexto en que se emplea la prueba, menos aplicables serán los resultados del estudio a la práctica clínica.

**PROBLEMAS DE ESTÁNDAR DE REFERENCIA.** El rendimiento de una prueba puede evaluarse en relación con un estándar de referencia independiente que establece la presencia o ausencia de una enfermedad. Sin embargo, a menudo resulta difícil determinar el estado verdadero de enfermedad por limitaciones de seguridad, coste, éticas y científicas.

Un estándar de referencia imperfecto produce errores de clasificación, lo que distorsiona las estimaciones de sensibilidad y especificidad de la prueba. Por ejemplo, aunque la coronariografía es un estándar de referencia para establecer el diagnóstico de oclusión de una arteria coronaria, no permite identificar a los pacientes con un riesgo aumentado

de episodio cardíaco con lesiones coronarias no oclusivas (p. ej., placa coronaria de pared fina químicamente inestable).

#### SESGO DE ESPECTRO.

Las pruebas diagnósticas deben evaluarse sobre una representación amplia y suficiente de personas con y sin enfermedad, para estimar el rendimiento de la prueba en subgrupos clínicamente relevantes, incluidas personas sanas sin enfermedad, así como pacientes sin enfermedad pero con trastornos comórbidos frecuentes, comienzo reciente de una enfermedad asintomática, enfermedad asintomática establecida, enfermedad avanzada, enfermedad terminal y otras enfermedades que afectan a los mismos órganos anatómicos.

La evaluación de las pruebas diagnósticas en una muestra no representativa provoca una sobrestimación de la sensibilidad y especificidad de la prueba. Por ejemplo, la sensibilidad de la prueba de esfuerzo varía con la gravedad de la enfermedad (mayor en la enfermedad de tres vasos, intermedia en otros casos de enfermedad de varios vasos, menor en la enfermedad de un solo vaso que no sea el tronco coronario izquierdo), localización de la lesión (mayor en las lesiones del tronco coronario izquierdo y progresivamente menor en las lesiones de la descendente anterior, coronaria derecha y circunfleja izquierda, respectivamente) y presentación de los síntomas (mayor en pacientes con el clásico dolor subesternal durante el ejercicio).

#### SESGO DE REFERENCIA, VERIFICACIÓN E INVESTIGACIÓN

El sesgo de referencia se produce cuando el resultado de la prueba evaluada se emplea para determinar que personas precisan una evaluación adicional mediante el estándar de referencia (p. ej. , cuanto mayor riesgo, coste o incómodo sea el estándar de referencia, mayor es el sesgo). La especificidad de la coronariografía con radioisótopos tras ejercicio es menor en la práctica que la señalada en los primeros estudios publicados, porque los médicos utilizaban esta prueba para seleccionar a los pacientes candidatos a cateterismo cardíaco. El sesgo de investigación sobrestima la sensibilidad pero subestima la especificidad de la prueba evaluada porque las personas con resultados positivos en esta prueba tienen más probabilidad de someterse a la prueba de referencia.

#### INTERPRETACIÓN CIEGA

Los resultados de la prueba evaluada deberían interpretarse sin conocer el resultado de la prueba estándar de referencia ni el verdadero diagnóstico. Cuando no se emplea esta interpretación ciega independiente de la prueba, es posible sobrestimar su sensibilidad y especificidad.

#### RESULTADOS NO INTERPRETABLES

Los resultados no interpretables o no diagnósticos son frecuentes. Es posible que un paciente no sea capaz de alcanzar el grado de ejercicio necesario para una prueba de esfuerzo, y la capacidad de los técnicos para obtener ecocardiogramas de alta calidad en pacientes obesos o de evitar la atenuación de la pared inferior en la tomografía computarizada por emisión de fotón único es variable. Estos resultados no interpretables de la prueba no se incluyen habitualmente en los cálculos del rendimiento de una prueba diagnóstica, lo que provoca una sobrestimación de la sensibilidad de la misma y son especialmente relevantes al comparar pruebas diagnósticas alternativas cuando la frecuencia de resultados no interpretables defiere entre las pruebas.

## UNIDAD PRUEBA-INTERPRETADOR

El rendimiento de una prueba diagnóstica depende tanto de las características de la prueba como de la experiencia de la persona que la dirige. Esta dependencia es especialmente importante en muchas pruebas cardíacas en las que algunos fármacos para imagen, protocolos, experiencia y habilidad del técnico y del interpretador afectan su rendimiento. Cuanto más relevantes sean estos factores, mayor será la dificultad para comparar o generalizar el rendimiento diagnóstico entre pruebas, interpretadores y centros.

### **Selección y uso de pruebas diagnósticas**

Los resultados positivos verdaderos permiten seleccionar el tratamiento apropiado, reducir la morbimortalidad y mejorar la función. Los resultados negativos verdaderos aportan confianza y evitan el riesgo, incomodidad y costes innecesarios. Por otro lado, los resultados positivos falsos conducen a realizar pruebas y tratamientos innecesarios y pueden aumentar la ansiedad y los resultados negativos falsos retrasan o pasan por alto el diagnóstico y tratamiento, con resultados adversos, pruebas innecesarias y aumento de la incertidumbre y ansiedad.

Las pruebas con alta sensibilidad son preferibles cuando los costes de un diagnóstico pasado por alto son elevados, el tratamiento es eficaz, seguro y barato, los resultados positivos falsos no provocan un riesgo serio y el objetivo es minimizar los resultados negativos falsos (como en el cribado o cuando se intenta descartar una enfermedad). Son preferibles las pruebas con alta especificidad cuando el objetivo es minimizar los resultados positivos falsos: cuanto menor es la efectividad y mayor el coste y el riesgo del tratamiento (p. ej., cirugía), los resultados positivos falsos pueden ocasionar un riesgo serio.

La selección de las pruebas diagnósticas varía según la fase del proceso diagnóstico. Por ejemplo, aunque un antecedente de dolor torácico tiene una sensibilidad modesta para detectar una enfermedad coronaria isquémica aguda, no es específico (es decir, tiene una tasa alta de positivos falsos). Por esta razón, los médicos evalúan a estos pacientes con una prueba de imagen (p. ej., prueba de esfuerzo, ecocardiografía de sobrecarga, gammagrafía de sobrecarga) que, a pesar de ser más cara y difícil de realizar, tiene mayor sensibilidad. Sin embargo, son necesarias pruebas con una especificidad muy alta (p. ej., coronariografía) antes de comenzar intervenciones caras o con riesgo.

El clínico puede optimizar la sensibilidad y especificidad bien mediante una selección cuidadosa entre pruebas alternativas con diferentes características de rendimiento o alterando el límite de una prueba para destacar bien la sensibilidad o la especificidad (pero no ambas simultáneamente).

### **Cribado**

Las pruebas diagnósticas se emplean habitualmente como cribado (detección de una enfermedad o riesgo elevado de enfermedad en personas aparentemente sanas antes del comienzo de los síntomas). Debido a la escasa prevalencia de enfermedad en las poblaciones sometidas a cribado, las pruebas siempre tienen una tasa elevada de positivos falsos (VP+ bajo). A menos que el cribado se limite a una población con un riesgo muy alto (50%), incluso las mejores pruebas ofrecen más resultados positivos falsos que positivos verdaderos. Además, la gran mayoría de las personas sometidas a cribado no se benefician de éste. Por este motivo, es mejor limitar el cribado a enfermedades graves para las que existe un tratamiento seguro y eficaz y en las que el diagnóstico y tratamiento temprano mejoran de forma significativa el resultado. Una

prueba de cribado debe tener una alta sensibilidad (para reducir al mínimo los resultados negativos falsos) y especificidad y debe ser segura, barata, cómoda y aceptable para la población evaluada. Además, debido a que el rendimiento de la prueba de cribado produce numerosos resultados positivos falsos, debemos disponer de pruebas de seguimiento apropiadas con un rendimiento y seguridad diagnóstica suficientemente altos y una coste razonable para identificar y descartar a las personas sin la enfermedad y para confirmar la presencia de enfermedad en las personas afectadas. Por esta razón, por lo general es mejor limitar incluso las mejores pruebas de cribado a las poblaciones y pacientes con un mayor riesgo de enfermedad según determinan los factores de riesgo conocidos.

Con frecuencia se sobrestima la efectividad y los beneficios de las pruebas y los programas de cribado. Al decidir si se realiza el cribado de una enfermedad, hay que determinar si el diagnóstico temprano mejora el resultado. Un error frecuente en la evaluación de los programas de cribado es medir el tiempo entre el diagnóstico de la enfermedad y el fallecimiento. El tiempo guía, el intervalo entre la detección de la enfermedad mediante cribado y el momento habitual de diagnóstico sintomático, es una función de la velocidad de progresión biológica de una enfermedad y de la sensibilidad de la prueba de cribado. Con frecuencia se sobrestima la efectividad del cribado, como resultado de una falsa prolongación de la supervivencia debido a una detección precoz de la enfermedad en ausencia de una mayor efectividad de una intervención más temprana (sesgo de tiempo guía), una mayor participación de las personas con más probabilidades de cumplir el tratamiento recomendado (sesgo de adherencia o cumplimiento) y un impacto no representativo de la detección de casos prevalentes en ciclos de cribado precoces, con un número desproporcionado de casos progresivamente lentos respecto a una mayor proporción de casos nuevos con progresión más rápida en los siguientes ciclos de cribado (sesgo de prevalencia). Por tanto, es esencial una evaluación rigurosa de pruebas y programas de cribado prometedores mediante estudios clínicos aleatorizados durante múltiples ciclos de cribado antes de su adopción generalizada.

### **Contenido de la información**

El valor de una prueba diagnóstica está en función de la cantidad de nueva información aportada. Incluso la mejor prueba diagnóstica no aporta nueva información diagnóstica cuando se está seguro de que la enfermedad está presente o ausente. Por otro lado, las pruebas diagnósticas aportan la mejor información diagnóstica cuanto mayor es la incertidumbre ante la prueba (probabilidad antes de la prueba o prevalencia intermedia). Las pruebas diagnósticas se interpretan habitualmente de forma dicotómica como positivas o negativas. El contenido de la información de una prueba diagnóstica es mayor cuando se considera también el grado de positividad, de forma que una prueba ligeramente anormal tiene más probabilidad de ser un resultado positivo falso que una prueba notablemente anormal, que tiene más probabilidad de ser un positivo verdadero. Por ejemplo, una prueba de esfuerzo con un descenso del segmento ST de 3 mm. En una fase temprana del ejercicio tiene más probabilidades de ser un resultado positivo verdadero que si se produce un descenso de 1 mm. En una fase más avanzada de ejercicio. De forma similar, una prueba muy negativa tiene más probabilidad de ser un negativo verdadero que una prueba ligeramente negativa, que tiene más probabilidades de ser un resultado negativo falso. Los resultados limítrofes tienen habitualmente un valor diagnóstico limitado. Por tanto, si tenemos en cuenta el grado de positividad de la prueba de esfuerzo conseguiremos aproximadamente un tercio más de información diagnóstica que si tan sólo interpretamos el resultado como positivo o negativo.



### Uso de cocientes de posibilidades en la práctica clínica

Resulta difícil utilizar clínicamente la sensibilidad y especificidad de una prueba por la complejidad de computación de las probabilidades revisadas. La relación de probabilidad (y los cocientes de posibilidades, que expresan el rendimiento de la prueba en forma de proporción) son equiparables a las probabilidades, pueden multiplicarse entre sí y son más fáciles de usar para revisar la probabilidad de enfermedad.

El cociente de posibilidades de una prueba positiva (CP+) representa la proporción PV/PF para un intervalo de resultados de la prueba y se obtiene calculando la probabilidad de un resultado positivo en las personas con la enfermedad dividida por el mismo resultado positivo en las personas sin la enfermedad.

Hay estos cocientes:

$$CP+ = \frac{\text{Probabilidad (resultado x)/pacientes con enfermedad}}{\text{Probabilidad (resultado x)/pacientes sin enfermedad}}$$

=

$$CP+ = \frac{\text{Sensibilidad}}{1 - \text{especificidad}}$$

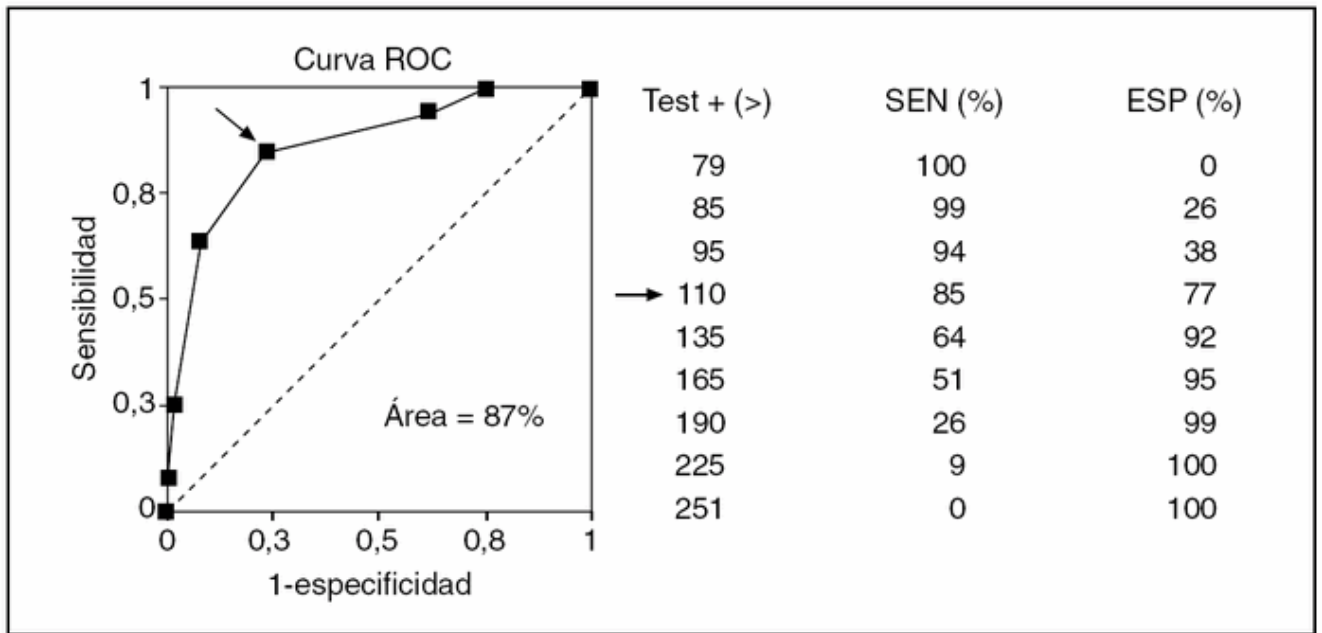
=

$$CP+ = \frac{\text{Probabilidad de PV/ pacientes con enfermedad}}{\text{Probabilidad de PV/pacientes sin enfermedad}}$$

Una CP+ mayor de 1 indica que la probabilidad de enfermedad posterior a la prueba es mayor que la probabilidad de enfermedad anterior a la prueba, mientras que una CP+ menor de 1 indica que el resultado de la prueba reduce la probabilidad de enfermedad (es decir, la probabilidad de enfermedad posterior a la prueba es menor que la probabilidad de enfermedad anterior a la prueba). Cuanto mayor sea el cociente de posibilidades de una prueba, mayor será la probabilidad de que el resultado de la prueba se produzca en una persona con la enfermedad a diferencia de una persona sin la enfermedad, mayor la probabilidad de que sea un positivo verdadero y mayor la probabilidad de enfermedad posterior a la prueba. Al aportar información sobre los rangos de los resultados de la prueba, los cocientes de posibilidades maximizan la información proporcionada por las pruebas diagnósticas. Puede utilizarse un nomograma simple para convertir la probabilidad de enfermedad estimada subjetivamente anterior a la prueba respecto a la probabilidad de enfermedad posterior a la prueba para un CP+ determinado, lo que facilita su uso clínico. Diamond y cols. Realizaron una estimación basada en los datos publicados sobre el rendimiento diagnóstico de la prueba de esfuerzo para al insuficiencia cardíaca en pacientes con síndromes de dolor torácico para diferentes magnitudes de descenso del segmento ST horizontal o en pendiente inducido por el ejercicio, empleando como estándar de referencia un estrechamiento del diámetro del 50% en al menos una arteria coronaria. Con un criterio de resultado dicotómico de descenso del segmento ST de 1 mm. Como resultado positivo, la sensibilidad de esta prueba fue 0,65 y la especificidad 0,89. La consideración del grado de positividad de la prueba mejora la discriminación entre aquellos con y sin la enfermedad.

### Curvas ROC

La sensibilidad y especificidad de una prueba dependen del límite elegido entre un resultado de la prueba normal y anormal. La selección de un límite que hace que no se considere de forma incorrecta que personas sin la enfermedad la padecen (especificidad = 1) produce numerosos resultados negativos falsos y reduce la sensibilidad de la prueba. Por otro lado, un límite que produce una identificación correcta de todas las personas con la enfermedad (sensibilidad = 1) provoca errores de clasificación de personas sin ella como enfermas (aumento de PF), lo que reduce la especificidad.



Una curva ROC es un trazado de la sensibilidad de la prueba frente a 1 - especificidad conforme varía la definición de resultado positivo a lo largo de todo el intervalo de valores clínicamente relevantes.

Debido al desfase inherente entre sensibilidad y especificidad conforme el criterio de positividad de un aprueba varía a lo largo de un intervalo de límites, no resulta apropiado o informativo hablar de sensibilidad o especificidad sin tener en cuenta a la otra, porque la sensibilidad de la prueba siempre puede aumentarse adoptando un criterio de positividad menos estricto, pero sólo a costa de una disminución de la especificidad o viceversa.

El límite óptimo de una prueba diagnóstica depende de la prevalencia y probabilidad de enfermedad estimada antes de la prueba, sensibilidad y especificidad y los riesgos, costes y beneficios de los diagnósticos correctos e incorrectos (estos mismos dependen de la gravedad de la enfermedad y de la seguridad, efectividad y coste de las intervenciones terapéuticas).

Sensibilidad

1-especificidad

=

Coste neto NV-coste neto PF

Coste neto NV-coste neto NF

x

probabilidad de no enfermedad posterior a la prueba

probabilidad de enfermedad posterior a la prueba

Los límites de una prueba diagnóstica publicados están determinados por lo general de forma más subjetiva, por lo que es posible que no sean los valores de uso clínico más apropiados.

### COMPARACIÓN DEL RENDIMIENTO DE SISTEMAS DE PRUEBAS DIAGNÓSTICAS ALTERNATIVAS

Es posible comparar el rendimiento de pruebas diagnósticas competitivas para un intervalo de límites, lo que requiere una curva Roc. Cuando se emplean curvas ROC para comparar el rendimiento de pruebas diagnósticas competitivas, la tecnología diagnóstica superior está más arriba y más a la izquierda en la curva ROC (es decir, la sensibilidad es más alta para un nivel de especificidad determinado y viceversa) y mayor es la superficie bajo la curva ROC (o mayor es la superficie bajo la curva ROC dentro de un intervalo clínicamente relevante de límites de la prueba) mejor es la prueba diagnóstica.

#### **Uso de pruebas diagnósticas en combinación**

En la mayoría de los casos se emplean varias pruebas diagnósticas combinadas para obtener la información diagnóstica necesaria para guiar el tratamiento del paciente. La realización de varias pruebas ofrece las ventajas de una información más abundante, mejor diagnóstico, velocidad de diagnóstico, comodidad y, en algunas circunstancias, reducción del coste. Sin embargo, las múltiples pruebas exponen al paciente a un mayor riesgo y coste y, en ocasiones, a un retraso diagnóstico y terapéutico. Además, el número de pruebas diagnósticas aumenta la probabilidad de discrepancias entre los resultados de las pruebas de forma considerable, lo que incrementa la complejidad de la interpretación de las pruebas.

Cuando se realizan varias pruebas, hay que adoptar criterios para interpretar combinaciones discrepantes. Una combinación de pruebas puede interpretarse como positiva si todas son positivas (criterio de positividad conjuntiva). Por otro lado, una combinación de pruebas puede interpretarse como positiva si alguna de las realizadas es positiva (criterios de positividad disyuntivos). Los criterios disyuntivos para interpretar combinaciones de pruebas aumentan la sensibilidad y disminuyen la especificidad respecto a las pruebas individuales, mientras que los criterios conjuntivos reducen la sensibilidad pero elevan la especificidad.

Por tanto, los criterios conjuntivos son preferibles cuando quieren evitarse los resultados positivos falsos y potenciarse al máximo la especificidad. Las pruebas disyuntivas se emplean cuando queremos evitar los resultados negativos falsos y aumentarse al máximo la sensibilidad.

#### **CORRELACION DE LA PRUEBA Y DEPENDENCIA CONDICIONAL.**

Las pruebas utilizadas en combinación están con frecuencia correlacionadas entre sí. Cuanto mayor sea la correlación entre las pruebas, menor información adicional se obtiene al usar dichas pruebas en combinación. Resulta más probable que las pruebas

que miden aspectos diferentes de una circunstancia patológica sospechada o que emplean diferentes métodos diagnósticos (p. ej., pruebas serológicas y de imagen) aporten información independiente respecto a las combinaciones de pruebas que emplean métodos similares.

### **PRUEBAS SECUENCIALES FRENTE A SIMULTÁNEAS.**

Cuando se emplean las pruebas en combinación, pueden realizarse bien de forma simultánea o secuencial, de modo que el rendimiento de las siguientes pruebas depende de los resultados de las previas. La estrategia simultánea acelera el diagnóstico pero precisa la realización de más pruebas que la secuencial. Por esta razón, la estrategia secuencial es en general más apropiada cuando la evolución del problema es lenta y cuando el coste de un diagnóstico más lento es menor (p.ej., pacientes ambulatorios), el riesgo de la prueba es elevado y los costes del tratamiento diferido son bajos. La estrategia simultánea es preferible cuando la evolución del problema es rápida y los costes de un retraso elevados (p. ej., pacientes hospitalizados), el riesgo de las pruebas es bajo y los costes de un tratamiento diferido son altos.

### **Toma de decisiones terapéuticas**

El objetivo del tratamiento es mejorar el resultado y la salud del paciente. Las intervenciones médicas se valoran en términos de seguridad (efectos secundarios aceptables), eficacia (beneficio neto en condiciones ideales), efectividad (beneficio neto en condiciones habituales) y coste efectividad (aumento de los beneficios respecto al aumento del coste). Como en la práctica todas las intervenciones tienen beneficios y riesgos, el reto es determinar los equilibrios entre beneficio y riesgo en pacientes concretos (es decir, cuál es el beneficio neto, en que pacientes y en que condiciones). La toma de decisiones terapéuticas requiere 1) identificar y definir la posible evolución, 2 calcular los posibles resultados y su probabilidad de presentación, 3 evaluar el valor de cada resultado según los desfases entre riesgo y beneficio (especialmente difícil cuando se hace una selección entre opciones alternativas competitivas) y 4) conocer cómo los pacientes valoran los resultados alternativos. Aunque la mayoría de las decisiones clínicas se toman de forma subjetiva con una combinación de métodos y principios que ya se han comentado, los métodos de apoyo de una decisión cuantitativa pueden ayudar al médico es esta tarea compleja y desafiante.

Los modelos predictivos emplean diferentes tipos de regresión estadística y matemática, así como técnicas relacionadas (p. ej., redes neurales) para identificar un conjunto frugal de variables clínicas que predicen un estado de enfermedad o el resultado de un tratamiento. Por ejemplo, los criterios de Jones aportan un modelo predictivo para la fiebre reumática. Las ecuaciones de riesgo de Framingham calculan el riesgo de presentar insuficiencia cardíaca y se han creado modelos para predecir el riesgo de injerto de derivación arterial coronaria y angioplastia coronaria. Los modelos predictivos son simplemente pruebas diagnósticas matemáticas, por lo que su uso e interpretación están sujetos a los principios descritos con anterioridad para las pruebas diagnósticas.

Los modelos de apoyo a la decisión se emplean para un elenco amplio de decisiones clínicas, desde problemas complejos caracterizados por una incertidumbre, riesgo o coste elevado a la normalización del tratamiento para problemas rutinarios frecuentes. Los modelos de decisión (entre los cuales el análisis de decisión es el más frecuente y familiar) son ayudas preceptivas a la decisión (es decir, buscan mejorar la toma de decisiones médicas al proponer como deben tomarse las decisiones más que a describir como se toman en realidad). Además de ayudar a guiar las decisiones clínicas, estos

modelos son especialmente útiles para informar de la toma de decisiones clínicas: organizando los aspectos destacados de la toma de decisiones clínicas tradicional, identificando (mediante análisis de sensibilidad multifactoriales o unifactoriales, en el que los cálculos puntuales de los parámetros del modelo cambian a lo largo de un intervalo de valores verosímiles) los elementos críticos que determinan las decisiones y que por tanto deben definirse de forma más precisa y explícitamente clarificar los desfases entre seguridad, efectividad y coste.

Los modelos de apoyo a la decisión estructuran las decisiones clínicas, definiendo con precisión la cuestión analizada, identificando acciones alternativas exhaustivas, relevantes y mutuamente excluyentes que pueden iniciarse y sus posibles resultados y calculando la probabilidad de que se produzcan episodios adversos, así como sus costes y valores asociados (es decir, “utilidades”). Se calcula e integran los valores del parámetro a partir de los mejores estudios disponibles experimentales (estudios aleatorizados), cuasi-experimentales (estudios observacionales, de cohortes y de casos controles) y no experimentales (metaanálisis, modelos, opinión de expertos). Se calcula el valor esperado de cada vía de decisión y se agrega a cada nodo de decisión, con la decisión que proporciona la mayor versatilidad esperada para la opción terapéutica preferida. Los análisis de sensibilidad identifican la repercusión de cálculos alternativos en el resultado y en las decisiones y aportan una estimación de la solidez de los resultados.

La ventaja principal del modelo de decisión es su naturaleza explícita y estructurada, que fuerza un análisis sistemático del problema y la asignación de valores explícitos, evita los errores durante el procesamiento de la información y centra la atención en parámetros que dirigen las decisiones clínicas. Las desventajas principales son la complejidad de desarrollo y explicación, así como la dificultad para mantener y actualizar los modelos. Por estas razones, incluso cuando no son aplicables directamente al lugar de asistencia de pacientes, estos modelos son extremadamente útiles para obtener información sobre los métodos clínicos generales y para dirigir el desarrollo de directrices clínicas.

En teoría, al tomar decisiones el médico debe actuar como representante del paciente, eligiendo la opción terapéutica que éste preferiría si pudiera manejar la misma formación que el médico. Para muchos problemas de salud existe más de una opción terapéutica razonable y la elección está determinada con frecuencia por los valores y preferencias del paciente por resultados beneficiosos alternativos y por la tolerabilidad de acontecimientos adversos y efectos de las secuelas de la enfermedad. Debido a que el conocimiento por parte del médico de las preferencias y valores del paciente es por lo general incompleto, hay que identificarlos en términos de riesgo, así como identificar e incorporar los posibles beneficios a las decisiones clínicas.

Las intervenciones y sistemas de apoyo a la decisión son métodos cada vez más atractivos para complementar el saber de los médicos, ya que aportan recordatorios e indicios mediante retroalimentación en el momento adecuado. Sin embargo, el desarrollo de sistemas de apoyo a la decisión resulta complejo y también es difícil su actualización y puesta en práctica, aunque al aumentar la computarización y sobre todo con la llegada de sistemas de registro médico informatizados es más sencillo obtener este apoyo en un tiempo adecuado.

## **Referencias**

Sanford Schwartz J. Toma de decisiones clínicas en cardiología. En: Zipes DP, Libby P, Bonow RO y Braunwald E, coordinadores. Tratado de Cardiología. 7ª ed. Madrid: Elsevier; 2006. p. 27-34.

2009