

¿Cómo saber si un método sirve?

Antes de examinar en detalle las nociones holísticas, es preciso discutir de manera más general el modo en que funciona la ciencia experimental o empírica. En breve, la pregunta que no sólo los médicos, sino también los pacientes bien informados, debieran formularse es: ¿Cómo puede establecerse con razonable certeza la eficacia de un método de diagnóstico o tratamiento? La pregunta puede asumir un carácter más general, si nos preguntamos *qué cosa distingue a la verdadera ciencia de la falsa ciencia*.

El método inductivo y sus limitaciones

La mayoría de los lectores preferirá –yo al menos eso espero– los procedimientos de la medicina occidental moderna a los de, digamos, un médico brujo del Africa. Sin embargo, tanto el médico occidental como el brujo hacen diagnósticos y prescriben tratamientos sobre la base de ciertas concepciones. ¿Por qué las concepciones del galeno occidental han de preferirse a las del médico brujo? Después de todo, seguramente este último puede presentar *una larga lista de pacientes satisfechos y dispuestos a jurar que el tratamiento les sanó*.

«Si un paciente con fiebre alta llega al doctor y el doctor, luego del examen, diagnostica una infección bacteriana y prescribe penicilina, él está haciendo uso de un cuerpo considerable de conocimiento científico. Si el mismo paciente fuera a un médico brujo las técnicas de examen, el diagnóstico y el remedio prescrito serían muy diferentes. Por ejemplo, el examen podría incluir un estudio de las entrañas de un pollo vivo; el diagnóstico podría ser que la fiebre es el resultado de un hechizo; y el remedio prescrito podría ser alguna suerte de purificación ritual o acción simbólica (como clavar alfileres en una efigie de la persona identificada como la responsable del hechizo)... desde el punto de vista del “círculo mágico” de las ideas de la brujería, estos [procedimientos] no son ni rituales ni simbólicos. Ellos son instrumentales o tecnológicos. En otras palabras, dada la visión del mundo –o sea, las teorías– del médico brujo, sus técnicas de examen y diagnóstico y su remedio prescrito son tan racionales como su contrapartida occidental»

(Briskman, p. 1034).

El profesor Briskman, al que acabamos de citar, plantea el importante problema de la distinción entre ciencia y pseudociencia.

Una de las primeras soluciones modernas a esta cuestión fue propuesta por Francis Bacon (1561-1626). Brevemente, puede exponerse como sigue: La verdadera ciencia se distingue de la falsa por el uso del *método inductivo*. Dicho método consiste en reconocer, a partir de la observación de fenómenos, hechos o casos particulares, las leyes generales que los rigen; va de lo individual a lo universal.

Según Bacon, el científico debe comenzar su estudio de la naturaleza *sin hipótesis preconcebidas*. Él se limita a observar y a reunir los resultados de sus observaciones. Tras haber reunido un *número suficiente* de observaciones, puede formular una teoría que explique *las causas* de los efectos observados, y que permita *generalizar* la aplicación de lo observado en cada caso particular.

La solución de Bacon ha sido muy popular y permanece hasta hoy como la más generalmente aceptada. Sin embargo, su debilidad fue señalada ya en el siglo XVIII por el filósofo David Hume (1711-1766). El problema es que el método inductivo nunca permitiría conocer *nada* si se lo aplicase a rajatabla. Por el contrario, su aplicación *exige* ciertas ideas preconcebidas:

1. que los efectos observados son una guía confiable para determinar sus causas (deducción);
2. que lo observado en casos particulares es una guía confiable para todos los casos similares (inducción) y
3. que en el momento de formularse la teoría existe un número suficiente de observaciones.

Pero precisamente es el propio método inductivo el que rechaza las presuposiciones: la investigación debe iniciarse con la observación desprejuiciada... Sin embargo, el inductivista *no puede, según sus propias reglas, demostrar que la observación desprejuiciada es de veras una guía confiable para conocer cosas que escapan a la observación* (causas a partir de efectos, o lo general desde lo particular).

Además de esto, los hechos demuestran que *todo hombre formula hipótesis acerca de las realidades que le rodean, y los hombres de ciencia no son la excepción*. Por ello la «observación desprejuiciada» de Bacon es una completa ficción.

Una distinción útil

Una solución satisfactoria al problema de distinguir entre ciencia y pseudociencia es la propuesta por el filósofo contemporáneo Karl Popper. Según él, la distinción radicaría en que *las teorías y concepciones de la ciencia genuina están sujetas a la crítica empírica*, o, dicho de otro modo, que son formuladas de manera tal que sea posible demostrar si son falsas.

La verdadera ciencia no busca solamente evidencia que

apoye sus teorías, sino también, y muy especialmente, toda pista o demostración que las ataque. Así, según Bacon la experimentación permite llegar a una teoría, mientras que para Popper la teoría es previa a la experimentación, y ésta sirve para ponerla a prueba.

Por abstractas que las presentes reflexiones puedan parecer, ellas son de la mayor importancia práctica en nuestro estudio, pues nos permiten establecer normas confiables por las cuales evaluar científicamente diversos procedimientos médicos.

En efecto, mientras que cualquier hipótesis puede encontrar evidencia en su favor, no todas pueden sostenerse frente a procedimientos pensados específicamente para probar su falsedad. Si estos fracasan, nuestra teoría se fortalece; si tienen éxito, la teoría debe ser abandonada o modificada.

«En la medida en que la práctica médica occidental explote teorías científicas que, a diferencia de las teorías de la brujería, son susceptibles de pruebas empíricas, han sido severamente puestas a prueba, y han sobrevivido, y en la medida en que la medicina ha sometido a sus propias técnicas independientes a severas críticas empíricas (por ejemplo, en ensayos clínicos) en lugar de meramente buscar “casos favorables”, hemos hecho todo cuanto se puede para eliminar falsas teorías y técnicas inefectivas o incluso dañinas»

(Ibid., p. 1110).

Claro está que esto no significa que todas las teorías y prácticas actuales sean óptimas, ni mucho menos. Precisamente lo que mantiene a la ciencia en general, y a la medicina en particular, en continuo avance es la *actitud crítica* hacia nuestras actuales concepciones y procedimientos.

Ensayos Clínicos

Sobre la base de lo antedicho, intentaremos una descripción, siquiera breve, de la forma en que la ciencia médica convencional del siglo xx evalúa cada nuevo método de prevención, diagnóstico o tratamiento. Todo nuevo método que aspire a ser aceptado por la comunidad médica internacional debe someterse a una serie de pruebas convencionales que permitan determinar su validez. En lo que sigue no se pretende sino ilustrar la manera en que funcionan los métodos de evaluación. Como todo conocimiento humano, los resultados de estas evaluaciones están sujetos a error. Empero, el uso *correcto* de la estadística permite reducir a un mínimo la probabilidad de error, y por lo tanto aumentar la confiabilidad de los resultados.

Supongamos que se propone una nueva prueba diagnóstica para una determinada enfermedad, como la diabetes. Para simplificar, aceptemos que el resultado de la prueba es cualitativo: «positivo» o «negativo». En el caso *ideal*, la prueba debería ser positiva en todos los enfermos y negativa en todos los sanos; pero esto no es lo que ocurre. La mayoría de las pruebas diagnósticas son ocasionalmente positivas en individuos sanos y negativas en sujetos enfermos. Debido a este problema, una prueba diagnóstica debe ser caracterizada por su *sensibilidad* y por su *especificidad*. La sensibilidad de la prueba se expresa como el porcentaje del total de enfermos que dan resultado positivo; una sensibilidad del 92%, por ejemplo, indica que 92 de cada 100 enfermos dan resultado positivo. La especificidad de la prueba se mide como el porcentaje de individuos sanos que dan resultados negativos.¹ La sensibilidad y especificidad de una prueba determinan conjuntamente su *validez* diagnóstica; en el capítulo IX daremos un par de ejemplos vinculados con el diagnóstico iridológico.

Es de desear que la sensibilidad y la especificidad de toda prueba sean ambas de 100%, pero ello no ocurre. Además, para cada prueba la sensibilidad y la especificidad no siempre van en la misma dirección. Así, una prueba puede ser

muy sensible pero poco específica, y en tal caso dará muchos resultados *falsamente positivos* (sujetos normales con resultados anormales). Por el contrario, puede ser muy específica pero poco sensible, y en consecuencia dar muchos resultados *falsamente negativos*, es decir sujetos enfermos con resultados negativos. En general, para toda prueba de aplicación clínica, tanto la sensibilidad como la especificidad deben superar el 95%; de lo contrario la prueba no se considera válida.

Los resultados de la evaluación de una prueba diagnóstica realizada en determinado laboratorio deben ser *reproducibles*. Esto significa que en condiciones similares otro laboratorio que evalúe independientemente el mismo método con iguales criterios, debe obtener resultados comparables a los del primer laboratorio. Además, toda nueva prueba debe *compararse* en cuanto a su sensibilidad, especificidad, inocuidad, sencillez y costo con otras técnicas ya existentes, mediante una rigurosa metodología estadística y el uso de técnicas «a ciegas» cuyo fundamento explicaremos a propósito de la evaluación de tratamientos.

Evaluación de tratamientos

Para que un tratamiento sea reconocido como eficaz por los entes oficiales responsables de la salud, y adoptado masivamente por los profesionales de la salud, es necesario que dicho tratamiento demuestre ser *por lo menos* comparable a los ya existentes en eficacia, inocuidad y costo. Al evaluar un tratamiento existen ciertas consideraciones básicas:

1. Historia natural de la enfermedad a ser tratada

Hay enfermedades que son autolimitadas, y se curan sin necesidad de tratamiento; tal el caso de un resfriado común. Otras enfermedades muestran mejorías y agravamientos a veces estacionales, como ocurre con las alergias, la enfermedad ulcerosa péptica y otras. El curso que sigue una enfermedad sin tratamiento se denomina su *historia natural* y

debe tenerse muy en cuenta a la hora de evaluar un tratamiento. Por ejemplo, si alguien promoviera un medicamento que cura el resfriado común en una semana, no sería tomado en serio, ya que ésa es la duración normal de la enfermedad sin tratamiento.

2. Influencia de las expectativas

Desde hace tiempo la medicina ha reconocido el importante efecto que las expectativas tanto de médicos como de pacientes pueden tener sobre la presunta eficacia de un tratamiento. El llamado efecto *placebo* consiste en que la administración de un «medicamento» inerte (sin acción farmacológica propia) produce a menudo una mejoría subjetiva e incluso objetiva en el paciente. El prestigio y la actitud del médico, la confianza que en él pone el paciente, y la naturaleza de la enfermedad influyen en la magnitud del efecto placebo. Para determinar qué porcentaje de respuestas se debe al efecto placebo y qué porcentaje puede atribuirse a la acción propia del fármaco en cuestión, se requieren estudios controlados.

3. Resultados de estudios controlados

Desde el descubrimiento en el laboratorio de una posible acción terapéutica de una droga o un procedimiento hasta su empleo generalizado hay un intervalo insoslayable durante el cual deben efectuarse una serie de pruebas para establecer su eficacia e inocuidad. La importancia de tales pruebas quedó demostrada de la manera más dramática luego de la tragedia de la talidomida.²

Estudios controlados

El intervalo entre el descubrimiento y el uso corriente en humanos es de unos diez años. Primero deben realizarse numerosas pruebas en animales de laboratorio. Si éstas resultan satisfactorias, puede procederse a los ensayos en

humanos.³ Estos ensayos en voluntarios o pacientes se dividen usualmente en tres *fases*, que deben ser cumplidas sucesivamente.

La fase I comprende el estudio de los efectos del tratamiento en casos individuales o pequeñas series. Allí se determinan las dosis, los efectos terapéuticos e indeseables. Esta fase se denomina preliminar. Si es cumplida satisfactoriamente, y los resultados son alentadores, puede procederse hacia las siguientes fases, de ensayos más rigurosos y en mayor escala, que se denominan *controlados*: la fase II, en grupos relativamente pequeños, y la fase III, en grupos muy grandes de pacientes.

En los ensayos controlados, se compara estadísticamente la eficacia del nuevo producto con la de placebos (sustancias inertes) y con la de otros productos ya aceptados cuya eficacia es conocida. Tales ensayos controlados se realizan según normas internacionalmente aceptadas y con el conocimiento y formal consentimiento de aquellos pacientes involucrados.

Si el paciente, pero no el médico, desconoce qué medicamento está tomando, o si está recibiendo un placebo, el ensayo se denomina «a ciegas». Si tampoco el médico cuál de los tratamientos posibles está recibiendo cada paciente, es un ensayo «doblemente a ciegas». En estos últimos, se toma un grupo de pacientes lo más homogéneo posible, y a cada uno se le adjudica al azar una de entre dos o más modalidades de tratamiento preestablecidas: la droga a ser evaluada, otra droga de eficacia conocida, o un placebo. El tratamiento exacto que cada paciente recibe es conocido sólo por personas que no están en contacto directo con el paciente, pero no por éste, ni por quien lo administra, ni por quien evalúa el resultado del tratamiento en cada caso. Todos los tratamientos lucen iguales, y sus resultados se analizan *antes* de conocerse qué tratamiento recibió cada cual. El objeto de todas estas precauciones es el de *eliminar todo factor ajeno al tratamiento en sí* que pueda influir en el resultado.

«El diseño doblemente a ciegas en los ensayos clínicos corrige dos errores comunes inherentes a los informes anecdóticos: la inclusión de pacientes que hubiesen mejorado sin terapia alguna, como si fuesen éxitos terapéuticos, y la inclusión de pacientes cuya mejoría fue una respuesta al placebo»

(Fields, p. 101).

NOTAS

1. Charles Ralph Buncher y Dan Weiner, *Valores de Referencia*, en L. A. Kaplan y A. J. Pesce (Dir.): *Química Clínica* (Panamericana, Buenos Aires, 1986, p. 349-359).
2. La talidomida es un fármaco que se empleó como sedante, en especial en las embarazadas, a finales de 1950, por su buena tolerancia y ausencia de efectos colaterales. Sin embargo, poseía un efecto teratogénico, o sea causante de malformaciones fetales. En 1960 se vinculó la administración de talidomida con un aumento en la incidencia de una malformación relativamente rara, la focomelia (ausencia congénita de miembros). Desde entonces los organismos de salud de diversos países aumentaron notablemente las exigencias sobre evaluación preclínica de los medicamentos.
3. Por ejemplo, véase el libro de Naranjo y col., citado en la bibliografía general, y Alfred Goodman Gilman y otros, *Goodman and Gilman's The Pharmacologic Basis of Therapeutics* (6ª Ed. Macmillan, New York, 1980, p. 40-55). Hay edición en español de Panamericana.