



Una cosa es conocer la causa y otra dar con la solución

Casino / Ventura

Diálogo 2

Sobre los estudios observacionales y su tratamiento periodístico

Pablo Francescutti

El segundo diálogo se centró en los estudios observacionales y su tratamiento en los medios de comunicación, pero desbordó este marco inicial para adentrarse en otras cuestiones de interés periodístico. La sesión, también moderada por Gonzalo Casino, contó con la presencia de los dos expertos de la mañana, Erik Cobo y Pablo Alonso, y dos periodistas, Esperanza García Molina y Pablo Francescutti.

Esperanza García Molina inició el debate con una cuestión muy concreta y de gran valor práctico en términos periodísticos: ¿en qué elementos debe fijarse un periodista que tiene entre manos un estudio observacional para saber si merece difusión?

Alonso: Nosotros no podemos entrar a valorar la importancia periodística de un estudio, porque

es algo que os toca a vosotros determinar. Lo que sí podemos es repasar lo que se ha dicho a lo largo de la jornada: diferentes diseños ofrecen diferentes niveles de confianza de entrada, confianza alta los ensayos clínicos y baja los observacionales. Ése es un primer criterio. Otros parámetros a tener en cuenta son que la cuestión que se plantearon los investigadores sea relevante y que el diseño elegido sea adecuado para responder a esa cuestión. Conviene tener presente que en cada estudio observacional, al igual que en los ensayos clínicos, hay una serie de posibles sesgos, sesgos de detección, de realización y otros. Dependiendo del estudio habrá que hacer una serie de preguntas para saber si ha sido bien ejecutado, cómo se reclutó la muestra, si se ajustaron los factores de confusión, y cómo se hizo el seguimiento. Cada estudio tiene unas

características sobre las que hay que hacerse preguntas específicas, y eso requiere un conocimiento básico de cómo leer críticamente un estudio observacional (ya sea, por ejemplo, de casos y controles o de cohortes), de cómo valorar su diseño y ejecución (riesgo de sesgo), y de qué factores aumentan o disminuyen la confianza, por ejemplo, sabiendo que ésta será mayor si se ha establecido una asociación muy importante o existe un gradiente dosis-respuesta (en el caso de los observacionales), y que será menor si los resultados son inconsistentes, si son imprecisos y si no son directamente aplicables a la población que interesa.

García Molina: Mi segunda pregunta tiene que ver con la causalidad. ¿Nunca puede decirse de un estudio observacional que ha establecido una relación causa-efecto?

Cobo: Para que nos entendamos, siempre es delicado hablar de causalidad, incluso en un ensayo clínico. Recordemos que existen dos formas de razonar: una es la deductiva, en la que partimos de unos principios que damos por ciertos, que podemos llamar premisas; y la otra es la inductiva, en la que partiendo de unos pequeños casos intentamos generalizar lo observado a muchos otros. Decía esta mañana que no tenemos que hablar de las leyes de Newton sino del modelo de Newton, porque las piedras no están obligadas a seguirlo o ir a la cárcel si no lo hacen. Por eso resulta una vanidad decir que estamos escribiendo las leyes de la naturaleza; lo que tiene mucho mérito es que somos capaces de elaborar modelos que nos permiten construir edificios que se aguantan, y eso ya es mucho. Un modelo es una esquematización de la realidad en un lenguaje muy abstracto, que tiene la ventaja de que lo interpretan igual un japonés y un americano. Pero estos modelos también tienen sus propias premisas. Todos los ensayos clínicos, que se sitúan en el nivel máximo de evidencia, tienen premisas; por ejemplo, el efecto que observamos en una persona es independiente del efecto que observamos en otra. Esta premisa puede no cumplirse en el caso de una vacuna, porque si un individuo no se contagia, será más difícil que otro lo haga;

por eso algunos ensayos se diseñan para poder estudiar ese contagio entre unidades. Pero por lo habitual se asume que las unidades se comportan independientemente y se hace el análisis sobre esta base. Por lo tanto, la lectura correcta de un ensayo clínico sería la siguiente: asumiendo ciertas todas las premisas, entre las que resalta que el efecto es independiente en las unidades –y habiéndose cumplido todo lo recién dicho por Pablo Alonso– podemos afirmar que el efecto de esta intervención es el siguiente: "...". Por eso, en general debemos ser muy modestos al hablar de los resultados, incluso delante del mejor diseño que tenemos, que es el ensayo clínico. En lo relativo a los estudios observacionales, tenemos que ser muy transparentes con lo que hemos observado y lo que interpretamos. Faltaría más que los autores del estudio Framingham no terminasen postulando que, si se logra bajar la presión arterial de las personas, a largo plazo bajarán los eventos cardiovasculares. Es magnífico que interpretaran así sus resultados, pero tienen que dejar claro que lo que identificaron es una "simple" relación entre dos variables: aquellos con la presión más alta fueron luego quienes sufrieron más eventos cardiovasculares al cabo de unos años. Luego, por supuesto, pueden tocar en la discusión aspectos que no fueron observados en el estudio, y plantear la hipótesis de que si se consigue bajar la presión arterial de las personas, quizás disminuyan a largo plazo los ictus, los infartos, etcétera. Este razonamiento especulativo es lícito, pero hay que dejar claro que lo estás haciendo. Volviendo a tu pregunta: ¿pueden estos autores hablar de causalidad? Sí, pero dejando muy claro que esa relación no se deduce directamente de ningún estudio si no le añade alguna premisa más. Y por supuesto, los observacionales requieren más premisas. Una diferencia del ensayo clínico con el observacional es que cuando le dices a una persona que haga algo, lo hará o no lo hará. Si a alguien le dices que no se vaya de su comunidad durante su tiempo de vida porque quieres hacerle un seguimiento durante 30 años, no tiene por qué hacerte caso. Estas impurezas puedes verlas en un estudio experimental, pero en un estudio observacional no, por lo que es una incertidumbre adicional que conviene destacar.

Casino: Quiero reformular la pregunta de García Molina, que es clave: ¿cuándo puede hacerse una inferencia causal? Pongo un ejemplo: la relación entre el tabaco y el cáncer. Hace unos 60 años se puso en marcha un célebre estudio observacional: el de los médicos británicos, que se prestaron a participar distribuyéndose en dos cohortes, una de fumadores y otra de no fumadores, para ver qué consecuencias tenía para su salud. Recordemos que por aquel entonces incluso algunos médicos aparecían fumando en los anuncios de cigarrillos. Fijaos cómo ha cambiado la situación desde mediados del siglo XX, cuando los facultativos se mostraban fumando, hasta el día de hoy en que se nos dice «Fumar mata» o «Fumar produce cáncer de pulmón». Todo empezó con un estudio observacional, a principios de la década de 1950. En 1964, la Food and Drugs Administration de Estados Unidos ya asociaba el tabaquismo con un mayor riesgo de cáncer de pulmón. Con el tiempo y la acumulación de pruebas científicas, se ha pasado de establecer una asociación a establecer una relación causal.

Cobo: Es magnífico tu ejemplo. Aquí se cumplen de manera impecable y perfecta los criterios fijados por Bradford Hill: si se cumplen, la interpretación más plausible de la relación entre tabaco y cáncer de pulmón es la causalidad. La prueba definitiva la tendríamos el día en que hiciéramos un diseño experimental asignando niños y adolescentes a grupos de fumadores y no fumadores. Pero eso, afortunadamente, no lo podemos hacer; no podemos decir «te ha tocado el grupo fumador de los 13 a los 43 años». Si alguien me pregunta si disponemos de evidencia suficiente para poner en las cajetillas de cigarrillos la leyenda «Fumar mata», tendré que decir que no tenemos el mismo nivel de evidencia que se exige para introducir un fármaco en el mercado, ya que nos basamos en bastantes más presunciones. Dicho esto, añadiré que ha sido triste que hayamos esperado 30 años para colocar dicha leyenda en las cajetillas, porque se tenía que haber hecho caso a Hill mucho antes y, en consecuencia, haber puesto ese anuncio quizás en la década de 1960. La pregunta que debemos formularnos no es «¿qué sabemos sobre el tabaco?» sino

«¿con qué decisión viviremos mejor?». Sanders Greenland, el gran epidemiólogo, zanjó la discusión en los años 1990 diciendo que el problema no radica en lo que no sabemos, sino en decidir qué es lo más sensato hacer. ¿Qué daño causaremos a la humanidad si alertamos de que el tabaquismo mata y luego resulta que la causa del cáncer es, por ejemplo, cierto contaminante común en todas las fábricas? Que algunos dejarán de fumar y unas empresas dejarán de ganar dinero. ¿Y qué daño causaremos si no ponemos el anuncio y más tarde se confirma que el tabaco es realmente la causa? Los epidemiólogos tienen la respuesta: 60.000 muertes al año en España.

Francescutti: Me gustaría profundizar en el asunto planteado por mi colega Esperanza García Molina, esto es, los criterios que debemos seguir los periodistas a la hora de interpretar datos biomédicos. Quiero referirme a la importancia del tamaño de las muestras de los estudios observacionales, y para ello haré un paralelismo con los sondeos electorales. A los periodistas nos han dicho que una muestra de entrevistados inferior a 1000 personas no es representativa de la opinión de toda la sociedad. ¿Podemos trasladar ese parámetro a los estudios observacionales y clínicos? ¿Cuánta gente debe participar para fiarnos de sus resultados? ¿Es el tamaño de la muestra el criterio decisivo, o debemos prestar más atención a la fuerza de la asociación o causalidad identificada?

Cobo: El tema que planteas exige una respuesta técnica. Para seguir en el plano de los sondeos, remitámonos a la opinión de los catalanes sobre el proceso soberanista. Imaginemos que en mi bolsillo derecho llevo la opinión de 5000 amigos míos, y en el izquierdo la de 100 catalanes escogidos al azar por toda Cataluña. ¿Cuál preferís? Obviamente la opinión de los 100 escogidos al azar, porque más importante que la cantidad de los datos es su calidad. La estadística mide el azar; no puede medir cómo son mis amigos, pero sí cuantificar hasta qué punto el azar ha influido en que sean de un color político u otro. La respuesta está en el intervalo de confianza. El mérito de la estadística es el de precisar el intervalo de con-

fianza. Y eso autoriza a decir que las opiniones expresadas por estos 100 individuos nos merecen una confianza del 90%, pongamos por caso. Y si hubieran sido 2000 encuestados, la confianza sería mucho mayor. Ciertamente, en las encuestas hechas al azar intervienen otros factores: que todos los votantes respondan, que no mientan, que no cambien de intención de voto en los días siguientes... Pero si se cumplen esas premisas, la estadística sabe cuantificar el margen de error y el intervalo de confianza. Los encuestadores honestos te dirán que hubo un 15% de entrevistados que no quisieron decir lo que votaban, y entonces lo más honrado sería estimar hasta qué punto esas "no respuestas" afectan sus conclusiones. En el *blog* de la Sociedad Catalana de Estadística se analiza el patinazo generalizado de las encuestas en las últimas elecciones autonómicas, como también lo hubo en 2003. ¿Cómo interpretar estos fallos? Asumiendo que las encuestas no pueden predecir el cambio de voto en las dos o tres semanas previas a las elecciones, un lapso en el que en Cataluña sucedieron muchas cosas que podían cambiar el voto. Intuyo que esas encuestas no son capaces de detectar los cambios finales, pero si tomamos las encuestas a escala global, lo que parece mentira es lo mucho que aciertan preguntando a tan pocos.

Alonso: Aparte del intervalo de confianza, muy útil para determinar si los resultados son precisos o no, existe una regla de la abuela: cuando los datos proceden de ensayos clínicos con menos de 100 eventos, la confianza que nos merecen es baja o muy baja; cuando la muestra cuenta con cientos de eventos, nos inspira una confianza moderada; y cuando disponemos de miles de pacientes o eventos, una confianza alta.

García Molina: He visto publicados estudios neurológicos llevados a cabo con resonancia magnética que sacan conclusiones con apenas 10 pacientes.

Casino: Está claro que cuando nos están hablando de unas pocas decenas de pacientes, hay que extremar las cautelas y tener cuidado con las conclusiones.

Cobo: Un matiz: cuando se describió el sida, ¿en cuántos casos se basaron? En muy pocos. Si se trata de un estudio muy novedoso, unos pocos casos podrán bastar para sacar conclusiones de interés; pero si quieres aportar algo nuevo sobre un tema conocido, debes seguir las reglas que apunta Alonso.

Alonso: Es verdad. Tomemos el descubrimiento del valor terapéutico de la adrenalina en el *shock* anafiláctico. El cambio que produjo en su tratamiento fue tan espectacular que no se necesitaron muchos ensayos para darlo por bueno.

Francescutti: Algo similar ocurrió cuando descubrieron que un tratamiento antibiótico curaba definitivamente la úlcera estomacal.

Alonso: Sí, porque se produjo un efecto muy llamativo. De todos modos, mantengamos la prudencia ante estudios con un intervalo de confianza estrecho, en los que con muy pocos eventos y muy pocos pacientes puede observarse un efecto importante. Hay trabajos que hablan de la fragilidad del valor *p* en tales estudios, ya que cambiando el desenlace (*outcome*) de uno o dos casos tales efectos desaparecen por completo, de modo que extremad la precaución.

Francescutti: Quiero hacer una reflexión acerca de la pirámide de la evidencia que vimos esta mañana. Cuando la explicabais caí en la cuenta de que en periodismo se suele invertir esa pirámide: las noticias basadas en estudios *in vitro* y con ratones, situadas en el nivel más bajo de confianza, compiten por la primera plana con los estudios epidemiológicos, más fiables. Los metaanálisis prácticamente no son noticia, salvo en revistas médicas o reportajes. Me parece lógico que una publicación especializada destaque un estudio con células o ratones, pero me pregunto cuál es el justo lugar que debemos dar a sus resultados cuando escribimos para un medio generalista.

Casino: Voy a devolver la pelota a los científicos. Que los periodistas hablemos de estudios en ratones es a menudo una consecuencia de

la influencia de la maquinaria informativa de la que forman parte los centros de investigación, las revistas de impacto, los comunicados de prensa y los investigadores. No veo el problema sólo en los periodistas que se hacen eco de esos estudios preliminares, sino también en los científicos que se prestan a salir en la radio y otros medios de comunicación para airear, por ejemplo, los efectos beneficiosos del aceite de oliva en animales, sobre todo cuando el oyente escucha la parte final de sus declaraciones sin saber si están hablando de ratones o personas. En muchos casos, la llamada de atención debería dirigirse a los investigadores que quieren salir en los medios para promocionar su carrera profesional y se prestan a informar de temas de salud a partir de estudios en ratones. En cuanto a la pregunta de Francescutti, hoy, con el escaso espacio disponible en los medios, un periodista tiene que pensarlo dos veces antes de cubrir un estudio de laboratorio. Francamente, no veo que estos trabajos justifiquen la mayoría de las veces una noticia destacada en un periódico general, o en un telediario, por más que trate del Alzheimer o de cualquier enfermedad de gran prevalencia, si no es debidamente contextualizada.

Cobo: Lo habéis resumido muy bien: es la paradoja entre el estudio inicial, que es sugestivo, y el final, que es confirmatorio de lo que ya sabíamos. Pero os devuelvo la pregunta: aunque el primer viaje de Cristóbal Colón fuera sólo tentativo, aunque tan sólo lanzase una posibilidad, aunque no estuviera confirmado su descubrimiento, ¿no os hubiera gustado a todos dar la noticia de que Colón regresó de las Indias?

Alonso: Eso no fue un hallazgo con ratones: fue un hallazgo espectacular, aunque Colón no entendiera bien a dónde había llegado. No fue tan preliminar como dices.

Cobo: Coged lo positivo de la analogía con Colón: cuanto más inicial es un descubrimiento, más peligroso resulta dar la noticia. Lanzaos, pero sed muy prudentes y decid enseguida que se trata de un estudio con animales, y que sólo

dentro de bastantes años podría tener algún impacto en la salud humana.

Alonso: Pero al hacerlo estás introduciendo ruido en la población con mensajes que se confunden constantemente. No creo que sea el objetivo de la prensa general desinformar a la población con mensajes contradictorios, por muy atractivos que sean.

Público: A la gente le interesa la noticia que le ofrecemos en la medida en que le afectará en el futuro. Por eso la pregunta habitual al científico que presenta un estudio in vitro es «¿cuándo prevé que tendrá aplicación?». Y te suelen contestar «no te lo puedo decir».

Casino: Quienes hacen investigación básica en el ámbito de la salud son conscientes de que tienen menos posibilidades de obtener repercusión mediática. Y cuando el aparato de comunicación de sus centros envía a los medios un comunicado de prensa sobre su trabajo, los investigadores ya han entrado en la rueda mediática y muy a menudo acaban saliendo en los medios. Pero es preciso que digan a los periodistas «cuidado, esto es una investigación de laboratorio y no estamos hablando de nada que tenga que ver, por ahora, con la salud humana».

Cobo: Hace unos años se produjo una novedad magnífica. Un metaanálisis realizado por la Colaboración Cochrane concluyó que no existían evidencias para aconsejar el reposo en caso de lumbalgia; por el contrario, era aconsejable una actividad moderada. Es de agradecer que alguien hiciera la revisión y que hoy, gracias a ella, la lumbalgia dure menos. ¿Por qué la prensa no se hizo eco de una noticia tan importante? Sospecho que porque preferimos las noticias positivas. Si el metaanálisis, en lugar de decir que no hay que aconsejar el reposo para la lumbalgia, hubiera dicho que el ejercicio moderado puede mejorar la lumbalgia, hubiera ofrecido un resultado positivo y quizás habría tenido más impacto mediático. Por eso, lo que apuntó Valentín Fuster esta mañana me parece muy importante: debe-

mos ser positivos, y tanto más en el entorno en que nos movemos ahora.

Público: Para mí la diferencia entre informar o no de un hallazgo en ciencia básica depende mucho de la revista en que se publique. ¿Tengo que fiarme de la publicación? Porque yo me guío por la publicación: si un artículo sale en *Science* o *Nature* va a misa.

Casino: Si los artículos vinieran con una clasificación de confianza alta, media, baja y muy baja, las cosas serían más sencillas.

Alonso: Esa confianza está implícita en que sean publicados en *Nature* o *Science*.

Casino: Son dos jerarquías diferentes: la de la revista, si es de mayor o menor impacto, y la jerarquía de confianza según la calidad del estudio. Y luego hay una tercera, que a menudo es la que más importa para decidir si se informa o no: el interés periodístico.

Alonso: Mi consejo a un periodista es que si tienes un titular que te quema en las manos, prudencia, porque un gran descubrimiento médico ocurre en contadas ocasiones. Las revistas científicas también viven de la publicidad y necesitan titulares potentes. A veces las cinco grandes de medicina, *The Lancet*, *New England Journal of Medicine*, *British Medical Journal*, *JAMA* y los *Annals*, pecan de lo mismo que algunos periodistas y publican artículos fantásticos sin examinarlos con el rigor necesario. Se han publicado estudios interrumpidos antes de tiempo, sin criterios explicitados a priori para evaluar los resultados, con pocos eventos y pocos pacientes, en revistas de alto impacto, que muchas veces han acabado contradiciéndose, incluida *Nature*. Como periodista tienes que tener criterio propio, no puedes fiarte únicamente de las revistas.

García Molina: A propósito de este punto, Casino señaló un estudio epidemiológico defectuoso publicado en *Nature*, que no tiene gran tradición en esa disciplina.

Casino: Si se ha hecho un estudio de salud muy importante lo normal es que no se publique en *Nature*, sino en alguna de las cinco grandes revistas de medicina. Casi con seguridad, estas revistas habrían rechazado el estudio del que yo hablaba.

Cobo: Cuando tenéis que coger la evidencia existente y determinar hasta qué punto creer los resultados de un trabajo, os puede resultar muy útil la declaración Consort de ensayos clínicos y estudios observaciones. Las grandes revistas médicas piden a los autores de los artículos que digan en qué página o línea se está satisfaciendo cada uno de los ítems de esa declaración. Es una posibilidad que tenéis. Con respecto a lo dicho sobre *Nature*, totalmente de acuerdo; esta revista no sigue esas líneas de comprobación ni esas declaraciones, es de otro estilo y toca otros temas. Otra cosa: atentos al cambio en los estilos de publicación. Existe una iniciativa que traslada el coste de publicación al investigador, ya que él paga por publicar y el lector tiene libre acceso a los artículos. El modelo Plos, por ejemplo. Otro cambio llamativo concierne a las críticas de los revisores y las respuestas de los autores, que ahora se publican en algunas webs. Esto permite a los periodistas conocer las críticas que los revisores han hecho al estudio.

Francescutti: Sin duda, estas enseñanzas presentan el mayor interés, aunque resultan difíciles de aplicar incluso por periodistas especializados. Pero el problema mayor es que ya casi no quedan periodistas especializados en ciencia en las redacciones; en los medios sólo hay “todoterrenos”, periodistas generalistas con menor capacidad de discriminación en estos temas. En contrapartida, prolifera la comunicación institucional de la ciencia, palpable en los *press releases* de los centros, instituciones y universidades que bombardean a los escasos medios que quedan. Y si llevar a la práctica esas enseñanzas en las redacciones parece difícil, no digamos ya aplicarlas en los gabinetes: veo muy difícil que sus miembros vayan a enmendar la plana a los científicos de sus centros y a sus superiores, ambos empeñados en dar la mayor repercusión a su investigación.

Casino: Sí, todo esto es un poco utópico... Así es como deberíamos funcionar los periodistas, en el supuesto de que existiera un escenario adecuado para conducirnos mejor.

Público: ¿Cuántos de los que estamos aquí hacemos periodismo?

Casino: Probablemente muy pocos; la mayoría se dedica a la comunicación.

Público: Quería preguntaros sobre el sesgo que supone la financiación de la mayoría de los estudios médicos por parte de la industria farmacéutica.

Alonso: Hay muchos trabajos que comparan resultados de estudios financiados por la industria con otros no financiados por ella, y detectan un sesgo conocido: en términos estadísticos, los estudios sobre fármacos financiados por la industria son más positivos que los demás. Mi consejo: si un artículo está financiado por laboratorios, examínadlo con más atención, e igual manteneos atentos al sesgo intelectual del investigador, que quiere dar la vuelta a sus datos para que sean noticiables y tengan más impacto.

Cobo: Soy de los pocos estadísticos que pueden presumir de haber participado en el desarro-

llo de dos productos que han llegado al mercado, algo muy difícil para un producto farmacéutico. El dueño de la compañía no tenía ningún interés en publicar los resultados de los ensayos clínicos de sus fármacos. En aquella época no había ningún apremio por publicarlos, pues lo que interesaba era que las agencias reguladoras los aprobasen. Después de esa aprobación ya tiene interés publicar los resultados. Si volvemos al ejemplo de los hipotensores, recordemos que salen al mercado tras demostrar que bajan la tensión. Más tarde, la FDA dice al fabricante «demuéstreme que su hipotensor también reduce el ictus». «De acuerdo –le responde el laboratorio– pero déjeme ponerlo en el mercado», y la FDA le da la autorización. De esta manera se inician los estudios de fase IV, ensayos que se realizan después de la comercialización en los que se estudian muchos aspectos de un fármaco que ya está en el mercado y que el paciente tiene que pagarse, con lo que la investigación le sale prácticamente gratis al laboratorio. Para concluir esta sesión, deciros que si un estudio resulta crucial para conseguir que el fármaco llegue al mercado, tomáoslo en serio; en cambio, si se trata de un estudio de un medicamento que ya está en el mercado o de uso compasivo, miradlo con otros ojos.