

# Interpretación y consideraciones metodológicas y estadísticas en los estudios de rescate

Federico Pulido<sup>a</sup> y Asunción Hernando<sup>b</sup>

<sup>a</sup>Unidad VIH. Hospital Universitario 12 de Octubre. Madrid. España.

<sup>b</sup>Departamento de Especialidades Médicas. Universidad Europea de Madrid. Villaviciosa de Odón. Madrid. España.

**La metodología utilizada en los estudios de tratamiento de rescate en ocasiones es compleja. Esto se debe a la heterogeneidad de objetivos y opciones. En primer lugar la definición de fracaso está sujeta a múltiples interpretaciones y matices. En segundo lugar, el objetivo a lograr en el tratamiento de estos pacientes ha variado en función de las opciones de tratamiento existentes en cada caso y en cada momento. Por último, la metodología utilizada para desarrollar los estudios de tratamiento de rescate ha ido variando con el tiempo, a la par que han cambiado los objetivos y las opciones. En estos momentos es esperable un nuevo cambio para ajustarnos a la situación actual, ya que el número de opciones terapéuticas para el tratamiento de rescate se ha incrementado de forma considerable en el último año. En la presente revisión se comenta cómo ha ido cambiando el diseño de estos estudios, los principales aspectos metodológicos a considerar y las recomendaciones de las agencias al respecto.**

**Palabras clave:** Metodología. Tratamiento de rescate. Tratamiento antirretroviral.

Interpreting methodological and statistical considerations in studies of rescue therapy

**The methodology used in studies of rescue therapy is sometimes complex. This is because of the heterogeneity of objectives and options. Firstly, the definition of failure has multiple interpretations and subtle distinctions. Secondly, the aim of treatment in these patients has varied according to the available treatment options in each case and at each moment of time. Lastly, the methodology used to develop these studies of rescue therapy has varied over time in line with changes in their aims and options. Currently, a new change can be expected to adjust to the current situation, since the number of therapeutic options for rescue therapy has substantially increased in the last year.**

Correspondencia: Dr. F. Pulido.  
Unidad VIH. Hospital 12 de Octubre.  
Avda. de Córdoba, s/n. 28041 Madrid. España.  
Correo electrónico: pulidof@gmail.com

**The present review discusses changes in the design of these studies, the main methodological issues to be taken into account and the recommendations on this subject.**

**Key words:** Methodology. Rescue therapy. Antiretroviral therapy.

## Introducción

A diferencia de los estudios realizados en pacientes sin tratamiento previo, donde la homogeneidad de los pacientes es relativamente fácil de conseguir, en los estudios dirigidos a comparar fármacos, combinaciones de fármacos o estrategias de uso del tratamiento antirretroviral en personas que han presentado fracaso con un tratamiento previo, el principal problema es la heterogeneidad de estas situaciones.

En primer lugar la definición de fracaso está sujeta a múltiples interpretaciones y matices. En segundo lugar, el objetivo a lograr en el tratamiento de estos pacientes ha variado en función de las opciones de tratamiento existentes en cada caso y en cada momento. Por último, la metodología utilizada para desarrollar los estudios de tratamiento de rescate ha ido variando con el tiempo, a la par que han cambiado los objetivos y las opciones. En estos momentos es esperable un nuevo cambio para ajustarnos a la situación actual, ya que el número de opciones terapéuticas para el tratamiento de rescate se ha incrementado de forma considerable en el último año.

En esta revisión intentaremos detallar la evolución que han tenido estos conceptos y repasar los aspectos metodológicos fundamentales en el diseño y la interpretación de este tipo de estudios.

## Definición de tratamiento de rescate

Se denomina tratamiento de rescate, en un sentido amplio, a todo el que se aplica tras haber suspendido un tratamiento previo por cualquier causa y en cualquier momento de la historia terapéutica del paciente. El tratamiento puede ser cambiado por 4 motivos fundamentales: por intolerancia, toxicidad, interacciones farmacológicas o fracaso. A su vez, el fracaso puede ser virológico (incapacidad para conseguir o mantener recuentos de ARN viral en plasma por debajo del límite de cuantificación de la prueba), inmunológico (incapacidad para recuperar la inmunodeficiencia, restaurando las cifras de linfocitos CD4+, aunque no existe acuerdo de cuáles deben de ser estas cifras ni el plazo razonable para alcan-

zarlas) o clínico (aunque en este caso, la posible aparición de cuadros clínicos de reconstitución inmunitaria en el seno de tratamientos efectivos dificulta también que haya una definición consensuada). En cualquier caso, se asume que el fracaso virológico acompaña y precede a los fracasos inmunológico o clínico, y dado su carácter cuantificable es, prácticamente siempre, de elección como criterio principal de fracaso en los ensayos clínicos con antirretrovirales<sup>1</sup>.

En los casos de cambio por intolerancia, toxicidad o interacciones farmacológicas, sólo es necesario cambiar el fármaco causal, siempre que haya un control adecuado de la replicación viral. En el cambio por fracaso, sin embargo, se debe dar una pauta con el mayor número posible de fármacos nuevos, con perfiles de resistencia no coincidentes con los de los fármacos sustituidos. En esta revisión el término tratamiento de rescate será sinónimo del cambio de tratamiento por fracaso virológico.

En la práctica clínica, para poder evaluar correctamente un episodio de fracaso terapéutico, es imprescindible tener en cuenta varios aspectos: la historia de tratamientos antirretrovirales previos, la presencia de signos y síntomas de progresión clínica, la adherencia, la tolerabilidad y las posibles interacciones que pudieran estar en el origen del fracaso, y los resultados de las pruebas de resistencia actuales y previas, si las hubiese. Con todo ello se pueden identificar las mejores opciones de tratamiento de rescate de entre las combinaciones disponibles (tratamiento optimizado). En un ensayo clínico que evalúa tratamientos de rescate, se ha de confirmar que todos estos aspectos han sido valorados y tenidos en cuenta a la hora de elegir la pauta de tratamiento, ya sea como base sobre la que se construye el régimen experimental (tratamiento optimizado de base [TOB]), ya sea como garantía de que los pacientes asignados a un tratamiento prediseñado tengan posibilidades reales de respuesta con él.

## Bases de la interpretación metodológica

Para interpretar correctamente los resultados de un ensayo clínico, es imprescindible conocer la metodología empleada en el estudio<sup>2,3</sup>. Pequeños cambios metodológicos en las definiciones, los procedimientos o el análisis pueden tener como consecuencia grandes cambios en los resultados. Desconocer, por tanto, la metodología utilizada sólo puede dar lugar a extraer conclusiones equivocadas.

De los muchos aspectos relativos a la metodología, y que por tanto deben estar descritos en el apartado de «métodos» en la publicación de un ensayo clínico, hay 2 preguntas imprescindibles que se debe responder antes de pasar al análisis o la interpretación de los resultados. La primera es conocer el objetivo primario del estudio. Aunque un ensayo clínico pueda dar respuesta a múltiples preguntas (objetivos secundarios), el estudio se diseña para responder a una única pregunta (objetivo primario). El tamaño del estudio se calcula (o se debería calcular) para que pueda dar respuesta certera a esa (y sólo a esa) pregunta. Cualquier otro análisis que se haga del estudio debe de interpretarse con precaución, poniéndolo en el contexto de la respuesta al objetivo primario y conociendo que el nivel de evidencia aportado es necesariamente menor. Sería un error (desgraciadamente frecuente) que, ante la ausen-

cia de la respuesta esperada al objetivo primario, las conclusiones del estudio se basen sólo en los resultados de los objetivos secundarios<sup>4</sup>.

El objetivo primario en los estudios de tratamiento antirretroviral de rescate ha cambiado con el tiempo. Mientras el tratamiento de rescate consistía en el reciclaje de fármacos pertenecientes a las 3 familias iniciales de antirretrovirales (inhibidores de la transcriptasa inversa análogos o no análogos de nucleósidos, e inhibidores de la proteasa), la proporción de pacientes con fracasos previos múltiples en los que se conseguía el control de la replicación viral hasta alcanzar cifras de carga viral indetectable (< 50 copias/ml) era muy pequeña<sup>5-7</sup>. Debido a esto, las agencias reguladoras recomendaban como objetivo primario el descenso de carga viral ponderado por el tiempo de seguimiento, considerando como clínicamente relevantes unas diferencias de bajada de carga viral mayores a 0,5 log. Ejemplos de este análisis se pueden encontrar en el estudio Context (que comparaba fosamprenavir/ritonavir con lopinavir/ritonavir como tratamiento de rescate)<sup>8</sup> y en el estudio O45 (que comparaba atazanavir/ritonavir con lopinavir/ritonavir en ese mismo contexto)<sup>9</sup>. La introducción de nuevas familias de fármacos, sobre todo cuando se usan en combinación, ha permitido obtener tasas de eficacia cercanas a las esperables en pacientes *naïve*, por lo que, desde el punto de vista tanto clínico<sup>1</sup> como metodológico, se ha establecido como objetivo principal en el tratamiento de rescate la consecución de cargas virales inferiores a 50 copias/ml.

El segundo aspecto fundamental a aclarar en la metodología de un ensayo clínico que compara tratamientos es si el estudio se ha diseñado para probar la superioridad (la diferencia de eficacia es difícilmente explicable por el azar; la diferencia es estadísticamente significativa; el intervalo de confianza del 95% de la diferencia no incluye el valor 0), la equivalencia (ambos extremos del intervalo de confianza del 95% de la diferencia no sobrepasan el valor de diferencia que se considera clínicamente relevante y que ha sido definido a priori) o la no inferioridad (la máxima diferencia posible a favor del tratamiento control, marcada por el extremo correspondiente del intervalo de confianza del 95% de la diferencia, no sobrepasa el valor de diferencia que se considera clínicamente relevante y que ha sido definido a priori)<sup>10</sup>. Cada uno de estos planteamientos conlleva la elección de hipótesis diferentes, diseño diferente, análisis diferentes y distinta interpretación de los resultados<sup>3</sup>. Uno de los errores más frecuentes y que más confusión generan al interpretar los resultados es asumir la «igualdad» de 2 tratamientos simplemente porque ninguno de ellos demuestra ser superior al otro (la diferencia no es significativa)<sup>11</sup>. Es necesario recordar e insistir en que la ausencia de diferencias significativas (una  $p > 0,05$ ) no indica que la diferencia no exista<sup>12</sup>. Otro posible error es el del cambio a posteriori del modelo de análisis, intentando establecer la no inferioridad en un estudio diseñado para probar superioridad o viceversa. Este cambio rara vez será posible, pues se requiere que se cumplan una serie de requisitos metodológicos difícilmente alcanzables si no se tuvo en cuenta esta posibilidad durante la fase de diseño del estudio. Estos requisitos han sido descritos por la Agencia Europea del Medicamento (EMA) en un documento específico<sup>13</sup>.

Al igual que ha sucedido en los estudios de tratamiento de inicio, durante años los estudios se han diseñado para

demostrar la superioridad frente a regímenes de eficacia limitada. Sin embargo, las elevadas tasas de respuesta que se describen con los tratamientos de rescate que incluyen al menos 2 fármacos plenamente activos (algo que garantiza el hecho de disponer de nuevas familias) va a hacer difícil que se mantenga este tipo de análisis; es de prever que también en el campo del rescate empecemos a ver estudios de no inferioridad.

## Metodología de los estudios de rescate

En los estudios de rescate podemos distinguir 3 grandes «épocas», que a efectos descriptivos vamos a denominar como: *a)* búsqueda de tratamientos de rescate eficaces; *b)* optimización de pautas de rescate, y *c)* nuevos fármacos en el rescate de pacientes con mult fracasos.

La primera época se corresponde con los intentos para encontrar pautas de rescate eficaces tras el fracaso de una terapia triple que contenía un inhibidor de proteasa no potenciado. Desde el punto de vista metodológico, destaca la heterogeneidad de los diseños y objetivos, generalmente en estudios con un reducido número de pacientes<sup>5</sup>. Como consecuencia de estos estudios, se establece el papel fundamental de los inhibidores de proteasa potenciados (con o sin un nucleósido asociado) en el rescate. Sin embargo, las tasas de respuesta virológica, aun utilizando un objetivo menos ambicioso (carga viral inferior a 500 copias/ml), eran sistemáticamente bajas: el 38% a las 24 semanas, en un metaanálisis de este tipo de estudios<sup>5</sup>.

Más interesante, desde el punto de vista metodológico, es la segunda época, en la que se buscaba comparar la eficacia de 2 inhibidores de proteasa potenciados: lopinavir/ritonavir, considerado el fármaco estándar en ese momento, con fosamprenavir/ritonavir (Estudio Context)<sup>8</sup> o con atazanavir/ritonavir (Estudio 045)<sup>9</sup>.

En las recomendaciones de la FDA para realizar ensayos clínicos con antirretrovirales, publicada en el 2002<sup>14</sup>, se hacía la siguiente consideración relativa al objetivo primario a establecer en los estudios de rescate: «En los estudios en pacientes con pocas posibilidades de conseguir la máxima supresión viral se pueden considerar otras opciones diferentes de la proporción de pacientes con carga viral indetectable, como la proporción de pacientes con descenso mantenido de la carga viral mayor de 1 logaritmo decimal. Asimismo, utilizar los cambios medios de la carga viral respecto a la basal, incluyendo los cambios ponderados por el tiempo, puede ser útil en pacientes pretratados en los que no es frecuente alcanzar una carga viral indetectable. Reducciones mantenidas de al menos 0,5 logaritmos se correlacionan con reducciones en el riesgo de progresión. La utilización de los cambios medios de la carga viral respecto al basal no se considera adecuada, sin embargo, en poblaciones de pacientes que pueden alcanzar reducciones de la carga viral por debajo del límite de detección. Los promotores deberían discutir previamente con la agencia qué objetivo es el adecuado para un protocolo concreto».

Ambos estudios, siguiendo las directrices de la FDA, establecen como objetivo primario el descenso de la carga viral ponderado por el tiempo de seguimiento, y dejan como objetivos secundarios la valoración del porcentaje de pacientes con carga viral indetectable (< 400 copias/ml o

< 50 copias/ml). También ambos estudios están diseñados como estudios de no inferioridad (del nuevo inhibidor de proteasa potenciado con respecto a lopinavir/ritonavir) y establecen como margen de no inferioridad una diferencia de 0,5 logaritmos a favor de lopinavir/ritonavir (tratamiento control). El margen de no inferioridad es la máxima diferencia que se consideraría clínicamente aceptable. Este valor tiene que ser definido a priori (indicado en el protocolo) y justificarse clínicamente. En este caso, el valor seleccionado fue el indicado por la FDA. Se debe recordar que este margen no indica la máxima diferencia que se puede dar en el estudio (estimación puntual de la diferencia), sino la máxima diferencia (a favor del fármaco control) que marca el intervalo de confianza bilateral del 95% de la diferencia encontrada en el estudio. Por lo tanto, la diferencia hallada en el estudio tiene que ser necesariamente inferior a la prefijada, para que su intervalo de confianza no la supere.

La diferencia en el descenso de carga viral ponderado por tiempo en el estudio 045 fue de 0,13 logaritmos (estimación puntual), con un intervalo de confianza que indicaba una diferencia máxima a favor de lopinavir/ritonavir de 0,39 logaritmos. Al ser la diferencia máxima a favor del tratamiento control menor que la preestablecida como clínicamente relevante (0,5 logaritmos), el estudio confirmó la no inferioridad de atazanavir/ritonavir en ese escenario.

En el estudio Context, sin embargo, la diferencia encontrada (estimación puntual) fue de 0,24 logaritmos más de descenso con lopinavir/ritonavir, con una diferencia posible a favor de lopinavir/ritonavir (según el límite del intervalo de confianza) de 0,536 logaritmos. Puesto que la diferencia a favor del tratamiento control puede superar la definida como relevante (0,5), el estudio no puede confirmar la no inferioridad de fosamprenavir/ritonavir respecto de lopinavir/ritonavir. Cabe destacar que el hecho de no demostrarse la no inferioridad de fosamprenavir no implica poder asumir la superioridad de lopinavir. Al tratarse de un estudio de no inferioridad, no puede concluirse algo no buscado, como sería la superioridad. En cualquier caso, la diferencia a favor de lopinavir no alcanza significación estadística ( $p > 0,05$ ; el intervalo de confianza cruza el 0), por lo que lopinavir no es superior a fosamprenavir, que sin embargo no ha demostrado ser no-inferior al primero. Este aparente galimatías esconde, desde un punto de vista estrictamente metodológico, una realidad desoladora: el estudio ha sido incapaz de responder a la pregunta planteada y no se ha resuelto la duda o, lo que es lo mismo, el estudio no vale para nada. Quizá esta ausencia de conclusión tenga algo que ver con el hecho de que este estudio nunca se haya llegado a publicar.

La tercera época la inauguran los estudios TORO<sup>6,7</sup>. Al tratar de demostrar la eficacia de un fármaco de una nueva familia, el diseño cambia. En estos estudios la comparación se realiza entre recibir o no recibir enfuvirtida añadido al mejor tratamiento disponible en ese momento (TOB), para tratar a pacientes con resistencia a las 3 familias clásicas de antirretrovirales. Obviamente, el diseño se establece para buscar la superioridad (no tendría sentido querer demostrar la no inferioridad frente a un tratamiento igual en todo menos en el uso del fármaco experimental). El diseño de este estudio servirá posteriormente de modelo para todos los estudios de tratamiento de rescate en pacientes con fracaso múltiple que han probado la eficacia de los nue-

vos antirretrovirales pertenecientes a una nueva familia: raltegravir (estudios BENCHMRK-1 y 2)<sup>15,16</sup> y maraviroc (estudios MOTIVATE 1 y 2)<sup>17,18</sup> o fármacos de familias preexistentes que mantienen actividad contra virus resistentes: tipranavir (RESIST 1 y 2)<sup>19,20</sup>, darunavir (POWER 1 y 2)<sup>21,22</sup> y etravirina (DUET 1 y 2)<sup>23,24</sup>. Otra característica común de los estudios TORO con los estudios que siguen su estela es el hecho de que haya 2 estudios simultáneos prácticamente idénticos (1 y 2) con un análisis primario establecido a las 24 semanas. Esto se debe a la exigencia por parte de la FDA de que la aprobación acelerada de un nuevo fármaco debe ir avalada por 2 ensayos clínicos controlados con 24 semanas seguimiento<sup>14</sup>.

El objetivo principal de los estudios TORO vuelve a ser el descenso de carga viral a las 24 semanas de tratamiento.

Los siguientes estudios, cronológicamente hablando, fueron los RESIST 1 y 2<sup>19,20</sup>, que estudian la eficacia de tipranavir/ritonavir en el tratamiento de rescate. Desde el punto de vista metodológico, es interesante destacar que estos estudios se diseñan como una comparación de tipranavir/ritonavir con los otros inhibidores de proteasa potenciados disponibles, por lo que se marcan un objetivo de no inferioridad frente al tratamiento estándar, pero en el protocolo ya recogen la intención de realizar un segundo análisis de superioridad si se conseguía (como fue el caso) demostrar la no inferioridad. Esto les ha permitido concluir que en el rescate de pacientes multitratados la eficacia de tipranavir/ritonavir es superior a la de los inhibidores de proteasa potenciados disponibles previamente.

De estos primeros estudios se aprendió que las tasas de respuesta con los nuevos fármacos eran superiores a las que se obtenían previamente y que el principal factor que incrementaba la eficacia era el número de fármacos plenamente activos que se incluían en el tratamiento (al menos hasta un máximo de 3). Por ello, la metodología de los estudios más recientes (BENCHMRK<sup>15,16</sup> y DUET<sup>23</sup>) presta especial atención a este aspecto y permite la utilización en el TOB de otros de los fármacos en investigación o desarrollo reciente (como darunavir y enfuvirtida). La mejor respuesta esperada (y confirmada) en estos estudios aconsejó utilizar de nuevo un análisis de superioridad (nuevo fármaco + TOB frente a placebo + TOB). También permitió modificar el objetivo principal, que deja de ser el descenso de la carga viral para convertirse (al igual que en estudios en pacientes *naïve*) en la proporción de pacientes con respuesta virológica.

Merece la pena hacer una última consideración metodológica respecto a la valoración de respuesta en estos estudios. Se trata de la importancia que dan al análisis «por intención de tratar», que de hecho, en ellos, es el tipo de análisis primario. Este análisis, que incluye a todos los pacientes aleatorizados y considera como fracasos de tratamiento a los pacientes perdidos, es fundamental y debe de ser aportado siempre, pero con mayor motivo en los estudios, como los que comentamos, cuyo objetivo es demostrar la superioridad. Su importancia radica en que aporta la única manera de mantener intacta hasta el final la asignación aleatoria de los sujetos, disminuyendo sesgos, y además refleja mejor las expectativas que se pueden tener de un tratamiento en el momento de tomar una decisión para utilizarlo. De todos modos no debe ser éste el único análisis que se presente, sino que las agencias y las revistas de calidad exigen siempre la realización de otros tipos

de análisis (denominados «análisis de sensibilidad») que permitan tener una visión más completa de los resultados del estudio. Entre estos análisis se encuentra el de «datos observados», que sólo incluye a los pacientes que permanecen en seguimiento y con el tratamiento asignado. También son convenientes otros análisis por intención de tratar (que, recordamos, incluye a todos los pacientes) pero cambiando los criterios de respuesta, como los que no consideran fracaso el cambio de tratamiento si la eficacia se mantiene, los que sólo tienen en cuenta el cambio si es por determinadas causas y no por otras, o considerar como fracaso (o no) los repuntes transitorios de carga viral que se controlan sin modificar el tratamiento, entre otras múltiples variantes de los análisis. Para intentar homogeneizar al menos el análisis fundamental, la FDA ha establecido un tipo de análisis primario en estudios con antirretrovirales en pacientes *naïve*, que ya se está utilizando también en los estudios de rescate, como se hace en los estudios DUET<sup>23,24</sup>. Se trata del análisis según el tiempo hasta la pérdida de respuesta virológica, más conocido por su acrónimo en inglés como TLOVR<sup>14</sup>.

Dado su creciente interés, pasamos a detallar las características del análisis TLOVR tal y como las enumera la propia FDA:

1. Para este análisis no se toman en consideración las visitas en las que no hay datos de carga viral. Sí se tendrán en cuenta, sin embargo, todas las visitas con determinación de carga viral, incluidas las que no se ajustan al protocolo y las realizadas después de la semana 48. No se deben interpolar datos para rellenar las visitas sin información.

2. Los sujetos que nunca consiguen una carga viral indetectable confirmada (en 2 visitas consecutivas), antes de presentar cualquiera de los eventos detallados a continuación, se considerará como fracasos en el tiempo 0.

a. Muerte.

b. Introducción de un nuevo antirretroviral al régimen. Se puede hacer excepciones para algunos cambios en los tratamientos de base (no en los comparativos) preespecificados si se deben a toxicidad o intolerancia de los fármacos acompañantes y no del fármaco en estudio.

c. Última visita disponible.

3. Para los sujetos que tienen una carga viral indetectable confirmada (2 visitas consecutivas con carga viral indetectable), el tiempo hasta el fracaso es el más corto de los siguientes (con la excepción detallada en el punto 4):

a. Tiempo hasta el evento descrito en 2b.

b. Tiempo hasta la pérdida de seguimiento.

c. Tiempo hasta presentar una carga viral detectable confirmada. Se define confirmada como 2 cargas virales detectables consecutivas o sólo una si tras ella se pierde el seguimiento.

d. Tiempo hasta el fallecimiento.

4. Si el tiempo hasta el fracaso virológico definido en el punto anterior está precedido por una o varias visitas programadas perdidas, se considerará que el tiempo hasta el fracaso virológico lo marca el tiempo hasta la primera de esas visitas perdidas.

## Directrices del la EMEA para el diseño de estudios de rescate<sup>25</sup>

Tras las 3 etapas detalladas en el apartado anterior, es previsible que en la actualidad estemos entrando en una nueva etapa, caracterizada por la alta proporción de pacientes que alcanzan una carga viral indetectable por debajo de 50 copias/ml, y por la existencia de opciones terapéuticas aceptables para la mayoría de los pacientes. En estas nuevas circunstancias es previsible que el diseño de los nuevos estudios adopte el objetivo de la no inferioridad, a la búsqueda de beneficios adicionales que añadir a la eficacia (simplicidad, tolerancia, coste).

Para conducir esta nueva fase, la EMEA ha actualizado su documento de recomendaciones para el diseño de estudios con antirretrovirales, aún en fase de discusión, y que por su importancia en la regulación de estos aspectos resumimos a continuación<sup>25</sup>. El documento diferencia entre 2 tipos fundamentales de tratamiento de rescate: el de los pacientes en fracaso que aún mantienen opciones suficientemente eficaces de tratamiento, y el de los pacientes (cada vez menos frecuentes) que tras el fracaso no disponen de opciones aceptables sólo con los fármacos ya comercializados.

### Pacientes que mantienen opciones válidas de tratamiento tras el fracaso

La decisión de cuándo y cómo cambiar un régimen que aparentemente está fracasando no siempre es fácil y se recomienda que los criterios de inclusión en un ensayo clínico para este tipo de pacientes se definan de acuerdo con las recomendaciones actualizadas de tratamiento antirretroviral.

Estos criterios de inclusión deberían tener en cuenta los antecedentes de tratamiento previo, junto con la información derivada de las pruebas de resistencias (actual y previas si las hubiese).

Hay varios diseños posibles pero, en cualquier caso, el régimen comparador debería ser adecuado, conforme a las recomendaciones de tratamiento vigentes, para todos los pacientes incluidos. Si el fármaco nuevo pertenece a una familia de fármacos ya disponible, el diseño más simple es el de seleccionar a pacientes sin experiencia previa con fármacos de esa familia y comparar de forma aleatorizada el nuevo fármaco con otro de su misma clase, asociados ambos con el mejor tratamiento de base disponible (tratamiento optimizado de base; TOB) o con un régimen predefinido cuya selección se pueda justificar de forma adecuada. Este diseño se puede aplicar también en el caso de que el fármaco experimental pertenezca a una nueva clase de antirretrovirales si se compara con un fármaco de otra familia que no ha sido utilizada con anterioridad por el paciente. Cuando se utilice como acompañante un TOB, es importante utilizar un sistema de valoración del tratamiento de base que permita objetivar que está compuesto por al menos 2 fármacos activos. Para este propósito se suele utilizar una escala de recuento de fármacos activos de acuerdo con los resultados del estudio de resistencias genotípico (GSS) o fenotípico (PSS), teniendo siempre en cuenta, además, la información aportada por los antecedentes terapéuticos del paciente. La administración de un nuevo fármaco sin modificar el tratamiento acompañante (monoterapia funcional) puede ser aceptable durante un breve período antes de optimizar el tratamiento de

base si la información existente sobre la potencia antiviral del nuevo fármaco es insuficiente.

El objetivo terapéutico en la práctica clínica es conseguir una carga viral inferior al límite de detección (actualmente es < 50 copias/ml), por lo que la proporción de pacientes que alcanzan este grado de supresión debe especificarse siempre. En la mayoría de los casos, además, esta proporción (pacientes con carga viral menor de 50 copias/ml) referida a un tiempo suficiente de seguimiento (por lo general, 48 semanas) debería ser el criterio primario de valoración (*endpoint*). Se deberían definir también los criterios primarios y secundarios de «fracaso virológico», teniendo en cuenta la eficacia esperable del régimen comparador y los criterios definidos en las guías de tratamiento antirretroviral vigentes. Si el objetivo del ensayo es demostrar la superioridad del nuevo fármaco, el análisis primario de eficacia se puede establecer a las 24 semanas, pero si lo que se pretende es establecer la no inferioridad, la duración debería ser de al menos 48 semanas. En el caso de que la demostración de la no inferioridad sea un objetivo razonable, el protocolo debería establecer de forma inequívoca el margen de no inferioridad establecido a priori (correspondiente a la máxima diferencia considerada clínicamente aceptable o irrelevante) y justificar científicamente los motivos de la elección de dicho margen. En este tipo de estudios es fundamental tener una baja tasa de pérdidas de seguimiento y realizar «análisis de sensibilidad» (análisis alternativos incorporando posibles modificaciones en los criterios de imputación de fracaso, como, por ejemplo, análisis según datos observados, análisis por intención de tratar con datos perdidos considerados como fracaso y cambios de tratamiento considerados o no como fracasos, etc.).

### Pacientes con escasas o nulas opciones terapéuticas tras el fracaso si sólo se utilizaran los fármacos disponibles comercialmente

Este tipo de pacientes es aquel que, conforme a su historia terapéutica y los estudios de resistencia a los antirretrovirales comercializados, no tiene más de 2 fármacos que se podrían considerar activos. En ellos, la posibilidad de utilizar nuevos fármacos en investigación podría constituir la única posibilidad de diseñar regímenes suficientemente eficaces. En esta situación, y para evitar comprometer opciones futuras de tratamiento en el paciente, se debería evitar la utilización de monoterapias funcionales (cuando el nuevo fármaco es el único fármaco activo) por períodos prolongados. Igualmente se debería minimizar la duración de regímenes que incluyan solo 2 fármacos activos. Teniendo esto en cuenta, los posibles diseños de estudios incluyen:

1. Si hay datos convincentes de la magnitud del efecto del tratamiento y la durabilidad de la respuesta, procedentes de estudios comparativos realizados en pacientes menos experimentados, podrían servir como base ante las agencias reguladoras para justificar el tratamiento con el fármaco nuevo. Ese tipo de datos permitiría presuponer la eficacia potencial y la seguridad a medio-largo plazo en condiciones controladas. No obstante, si se trata de un nuevo fármaco de una familia ya existente, habría que diseñar estudios de monoterapia funcional durante un corto período con objeto de conocer cómo afectan a la actividad del fármaco las mutaciones previas frente a esa familia de antirretrovirales. Si, por el con-

trario, se trata de un fármaco de una nueva familia, un corto ensayo de monoterapia funcional podría reafirmar la actividad antiviral del compuesto en la población con múltiples fracasos previos. Antes de iniciar un período de monoterapia funcional se requiere un mínimo de 8 semanas de tratamiento antirretroviral estable para que los resultados se puedan interpretar correctamente. La duración adecuada de la monoterapia se debería decidir en cada caso en función de la información disponible sobre el fármaco en estudio y su familia. En todo caso, la prueba con monoterapia funcional debería seguirse de un cambio temprano en el tratamiento acompañante que incluya al menos un nuevo fármaco activo.

2. Para los pacientes en que se dispone de 2 fármacos activos comercializados para elaborar un TOB, el diseño que incluya una comparación con placebo podría ser aceptable. El criterio primario de valoración de respuesta en este tipo de estudios podría ser el tiempo hasta la respuesta virológica (definida por lo general como una carga viral < 50 copias/ml) o la tasa de respuesta sostenida en tiempo predefinido. Para los pacientes que tras el procedimiento de selección se consideren no elegibles para el estudio porque no se disponga de 2 fármacos activos que utilizar en el TOB, se puede diseñar un estudio paralelo, no comparativo, en el que se les proporcione a todos el fármaco nuevo junto al TOB. En este caso se podría permitir la utilización concomitante en el TOB de algún otro fármaco en investigación. El seguimiento de estos pacientes del estudio paralelo tendría que ser igual que el de los pacientes del estudio comparativo, con objeto de conseguir datos de seguridad adicionales. La EMEA considera preferible esta forma de valoración de nuevos compuestos sobre el uso exclusivo de programas de acceso expandido. En el caso de que coincida en el tiempo el desarrollo de 2 nuevos fármacos que pudieran ser utilizados en este tipo de pacientes, las agencias consideran aceptable la realización de estudios con diseño factorial para estudiar la eficacia y la seguridad de ambos compuestos en un mismo estudio. Sin embargo, si se prevé que por las características de los fármacos pudiera haber interacciones entre ellos, se debería realizar primero un estudio para aclarar estas interacciones. Dada la especial situación de los pacientes con fracaso terapéutico múltiple, y la necesidad de nuevos fármacos en este escenario, las agencias pueden autorizar la comercialización de los nuevos fármacos para esta indicación de uso si se disponen de datos de eficacia a 24 semanas y no hay datos que indiquen que hay toxicidad.

### Declaración de conflicto de intereses

Los autores han declarado no tener ningún conflicto de intereses.

### Bibliografía

- Panel de expertos de Gesida/PNS. Recomendaciones de GESIDA/Plan Nacional sobre el Sida respecto al tratamiento antirretroviral en adultos infectados por el virus de la inmunodeficiencia humana (actualización enero de 2007). *Enferm Infecc Microbiol Clin.* 2007;25:32-53.
- Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet.* 2001;357:1191-4.
- ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Stat Med.* 1999;18:1905-42.
- Messerli FH. ALLHAT, or the soft science of the secondary end point. *Ann Intern Med.* 2003;139:777-80.
- Losina E, Islam R, Pollock AC, Sax PE, Freedberg KA, Walensky RP. Effectiveness of antiretroviral therapy after protease inhibitor failure: an analytic overview. *Clin Infect Dis.* 2004;38:1613-22.
- Lalezari JP, Henry K, O'Hearn M, Montaner JS, Piliero PJ, Trottier B, et al. Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America. *N Engl J Med.* 2003;348:2175-85.
- Lazzarin A, Clotet B, Cooper D, Reynes J, Arasteh K, Nelson M, et al. Efficacy of enfuvirtide in patients infected with drug-resistant HIV-1 in Europe and Australia. *N Engl J Med.* 2003;348:2186-95.
- Elston R, Yates P, Tisdale M, Richards N, White S, DeJesus E. GW433908/ritonavir: 48 weeks results in PI-experienced subjects: a retrospective analysis of virological response based on baseline genotype and phenotype. En: Program and Abstracts of the XIV International AIDS Conference, Bangkok 2004. Abstract MoOrB1055.
- Johnson M, Grinsztejn B, Rodriguez C, Coco J, DeJesus E, Lazzarin A, et al. Atazanavir plus ritonavir or saquinavir, and lopinavir/ritonavir in patients experiencing multiple virological failures. *AIDS.* 2005;19:685-94.
- Kaul S, Diamond GA. Good enough: a primer on the analysis and interpretation of noninferiority trials. *Ann Intern Med.* 2006;145:62-9.
- Greene WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? *Ann Intern Med.* 2000;132:715-22.
- Alderson P. Absence of evidence is not evidence of absence. *BMJ.* 2004;328:476-7.
- EMEA. The European Agency for the Evaluation of Medical Products. Committee for proprietary medicinal products. Points to consider on switching between superiority and non-inferiority. CPMP/EWP/482/99 2000.
- FDA. U.S. Department of Health and Human Services. Center for Drug Evaluation and Research. Guidance for Industry. Antiretroviral Drugs Using Plasma HIV RNA Measurements — Clinical Considerations for Accelerated and Traditional Approval. Octubre 2002.
- Cooper D, Gatell J, Rockstroh J, Katlama C, Yeni P, Lazzarin A, et al. Resultados a las 48 semanas de BENCHMRK-1, un estudio en fase III de raltegravir (RAL) en pacientes con fracaso del tratamiento antirretroviral (TAR) y VIH-1 resistente a las tres clases de medicamentos. En: Program and Abstracts of the 15th Conference on Retroviruses and Opportunistic Infections. Boston 2008. Abstract 788.
- Steigbigel R, Kumar P, Eron J, Schechter M, Markowitz M, Loutfy M, et al. Resultados a las 48 semanas de BENCHMRK-2, un estudio en fase III de raltegravir (RAL) en pacientes con fracaso del tratamiento antirretroviral (TAR) y VIH-1 resistente a las tres clases de medicamentos. En: Program and Abstracts of the 15th Conference on Retroviruses and Opportunistic Infections. Boston 2008. Abstract 789.
- Lalezari J, Goodrich J, DeJesus E, Lampiris H, Gulick R, Saag M, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy In Viremic, ART-Experienced Patients Infected With CCR5-Tropic HIV-1: 24-Week Motivate 1 Study. En: Program and Abstracts of the 14th Conference on Retroviruses and Opportunistic Infections. Los Angeles. 2007. Abstract 104bLB.
- Nelson M, Fätkenheuer G, Konourina I, Lazzarin A, Clumeck N, Horban A, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy In Viremic, ART-Experienced Patients Infected With CCR5-Tropic HIV-1: 24-Week Motivate 2 Study. En: Program and Abstracts of the 14th Conference on Retroviruses and Opportunistic Infections. Los Angeles. 2007. Abstract 104aLB.
- Gathe J, Cooper DA, Farthing C, Jayaweera D, Norris D, Pierone G Jr, et al. Efficacy of the protease inhibitors tipranavir plus ritonavir in treatment-experienced patients: 24-week analysis from the RESIST-1 trial. *Clin Infect Dis.* 2006;43:1337-46.
- Cahn P, Villacian J, Lazzarin A, Katlama C, Grinsztejn B, Arasteh K, et al. Ritonavir-boosted tipranavir demonstrates superior efficacy to ritonavir-boosted protease inhibitors in treatment-experienced HIV-infected patients: 24-week results of the RESIST-2 trial. *Clin Infect Dis.* 2006;43:1347-56.
- Katlama C, Esposito R, Gatell JM, Goffard JC, Grinsztejn B, Pozniak A, et al. Efficacy and safety of TMC114/ritonavir in treatment-experienced HIV patients: 24-week results of POWER 1. *AIDS.* 2007;21:395-402.
- Clotet B, Bellos N, Molina JM, Cooper D, Goffard JC, Lazzarin A, et al. Efficacy and safety of darunavir-ritonavir at week 48 in treatment-experienced patients with HIV-1 infection in POWER 1 and 2: a pooled subgroup analysis of data from two randomised trials. *Lancet.* 2007;369:1169-78.
- Madruca JV, Cahn P, Grinsztejn B, Haubrich R, Lalezari J, Mills A, et al. Efficacy and safety of TMC125 (etravirine) in treatment-experienced HIV-1-infected patients in DUE1-1: 24-week results from a randomised, double-blind, placebo-controlled trial. *Lancet.* 2007;370:29-38.
- Lazzarin A, Campbell T, Clotet B, Johnson M, Katlama C, Moll A, et al. Efficacy and safety of TMC125 (etravirine) in treatment-experienced HIV-1-infected patients in DUE1-2: 24-week results from a randomised, double-blind, placebo-controlled trial. *Lancet.* 2007;370:39-48.
- EMEA. The European Agency for the Evaluation of Medical Products. Committee for medicinal products for human use. Guideline on the clinical development of medicinal products for the treatment of HIV infection. Draft. EMEA/CPMP/EWP/633/02 Rev. 2 October 2007.