
El ensayo clínico para un solo paciente. Justificación, metodología y aportaciones bioéticas

Juan B. Cabello López^a, Víctor Abaira^b y Javier Gómez García^{a,c}

^aUnidad de Investigación. Hospital General Universitario de Alicante. ^bUnidad de Bioestadística Clínica. Hospital Ramón y Cajal. Madrid. Centro Coordinador de REUNI. Instituto de Salud Carlos III. Madrid. ^cBecario del Instituto de Salud Carlos III n.º 96/4448.

La actividad clínica constituye un proceso de decisión-acción relativo a un paciente individual y con una finalidad multidimensional. En efecto, la conducción de la anamnesis y la realización del examen y de unas determinadas pruebas diagnósticas (y no otras), así como la selección de un tratamiento, implican continuas decisiones. La ejecución activa de algunas pruebas (p. ej., pruebas hormonales) o la administración del tratamiento son, por su parte, acciones con diferentes niveles de complejidad y diferentes actores. En unos casos la acción y su control son fácilmente realizables, son los casos de acciones de ejecución directa, por ejemplo realizar una angioplastia, una intervención quirúrgica o administrar una inyección intravenosa; tanto la acción como su control (prescripción y ejecución) son realizables por el médico. En otros casos, por ejemplo, la ejecución de un tratamiento con bloqueadores, amitriptilina o teofilina, la realización de la acción y su control ofrece más dificultades por prolongarse en el tiempo y sobre todo por depender también del paciente; así pues, prescripción y observancia son realizadas, respectivamente, por médico y paciente¹.

Por otra parte, los objetivos de un tratamiento pueden ser diversos y afectar a múltiples niveles: mejorar los síntomas de la enfermedad, evitar recidivas de ésta, prevenir deterioros o acontecimientos futuros, remediar las molestias actuales o dejar morir con dignidad al paciente^{2,3}. La definición de objetivos terapéuticos y la ejecución de los mismos son, con las precisiones pertinentes, un proceso común de médico y de paciente.

Este procedimiento de decidir para un individuo está basado en el uso combinado de evidencias y valores. Respecto a las evidencias, se usa fundamentalmente conocimiento operativo

construido a partir de grupos de pacientes y también conocimiento sobre los mecanismos de enfermar y sanar^{4,5}. Como caso particular de este proceso, para la selección de la terapéutica se acepta que el mejor conocimiento sobre tratamientos proviene de los ensayos clínicos controlados y aleatorizados⁶.

Los objetivos del presente artículo son: *a)* plantear los problemas teóricos que existen al aplicar el conocimiento de los ensayos clínicos a casos particulares; *b)* revisar los aspectos más relevantes del diseño, conducción y análisis de los ensayos clínicos para un solo paciente (EC $n = 1$), y *c)* señalar las peculiaridades bioéticas de los EC $n = 1$.

Problemas al aplicar el conocimiento de los ensayos clínicos a casos particulares

El ensayo clínico es el mejor diseño para conocer la relación entre una acción o tipo de tratamiento y un efecto o resultado del ensayo. Dicho de otro modo, es la mejor forma de obtener evidencias sobre la eficacia, efectividad y seguridad de los procedimientos o fármacos para lograr unos objetivos terapéuticos definidos⁶. Pese a la firmeza epistemológica de este conocimiento, existen algunos aspectos de estos estudios que conviene considerar, por cuanto pueden limitar su valor al seleccionar un tratamiento para un paciente concreto.

Dichos ensayos se realizan, usualmente, en un grupo de pacientes definidos por los criterios de elegibilidad (*población de estudio*), y que serán asignados mediante un procedimiento aleatorio a los grupos a comparar. De este modo, el resultado del estudio se refiere, en sentido estricto, sólo a esa población que reúne los criterios de inclusión. El clínico tendrá que decidir en qué medida su paciente es análogo a la población del estudio y, por tanto, en qué medida cabe esperar en él un comportamiento análogo del tratamiento.

Publicado en *Med Clin (Barc)* 1997; 109 (15): 592-598. Este trabajo ha sido financiado por el FIS (proyecto 97/0043-01).

La valoración de la eficacia en un ensayo clínico depende de modo crítico de cuál sea la pregunta principal; en este sentido, los ensayos se diseñan para determinada/s pregunta/s principal/es (habitualmente una), mientras que otras preguntas secundarias influyen menos en el diseño. Por ejemplo, podemos preguntarnos: ¿el fármaco A disminuye la mortalidad o la incidencia de un nuevo accidente coronario?, ambas podrían ser preguntas principales. Las preguntas secundarias podrían ser: ¿el fármaco A mejora el número de crisis de angina, la clase funcional del paciente o la calidad de vida?

Estas preguntas secundarias, aunque suelen estar recogidas en el estudio, no reciben la misma atención en el diseño que las variables principales: por ejemplo, no son sometidas a procesos específicos de mejora de la calidad en la medición (en estudios piloto) y no se evalúan con la misma intensidad que aquéllas. Por otra parte, estas preguntas secundarias no influyen en la estimación del tamaño muestral, y por tanto el diseño del estudio no está condicionado por esas preguntas. Por todo ello, respecto a las preguntas secundarias la calidad de la evidencia no es tan solvente, desde una perspectiva epistemológica, como para las preguntas principales.

En la selección de las preguntas principales influye la percepción de cuál es el objetivo más importante del tratamiento de esa enfermedad, y se eligen a partir de lo que los médicos o investigadores consideran importante para el paciente, por ejemplo prolongar su supervivencia. Sin embargo, el paciente podría tener otros valores y decidir qué otro objetivo es el que a él le importa; esta circunstancia es esencial al elegir el tratamiento y debería considerarse en el diseño de los estudios⁷.

Por otra parte, y salvo que el tratamiento sea extraordinario, en el grupo experimental existirán sujetos en los que el tratamiento estudiado no es efectivo (sujetos no respondedores), ello es así, aun cuando la terapéutica ensayada pueda ser eficaz en el conjunto de pacientes, y la respuesta sea más positiva en este grupo que en el de control. En este sentido, en los ensayos clínicos los indicadores de asociación entre tratamiento y efecto (*odds ratio* [OR], *risk ratio* u otros)⁸ informan sobre la intensidad de una relación, de su independencia del azar y de su posible causalidad. Es decir, afirmar que el tratamiento A es 5 veces mejor que el B y que la asociación no depende del azar (OR = 5; *intervalo de confianza* del 95 %, 1,9-9,2) es sólo una *proposición causal potencial* que informa sobre la asociación, pero nada predice sobre un paciente⁹. La proba-

bilidad de que un paciente responda al tratamiento será la proporción de respuestas en el grupo tratado con A o B, y ésta será una *proposición causal predictiva*⁹.

Con esa evidencia de efecto y esa probabilidad de respuesta, indicaremos un tratamiento de acuerdo con el paciente y sus valores; este tratamiento tendrá frecuentemente un sentido de prueba, y por tanto será preciso controlar sus efectos. Puede ocurrir, es lo más probable, que el tratamiento sea eficaz, pero también que no lo sea e incluso que empeore la situación de nuestro paciente o que produzca efectos adversos.

Para algunos objetivos del tratamiento cuya variable de respuesta es la ocurrencia de un acontecimiento a largo plazo, por ejemplo, mortalidad o nuevo accidente coronario, no podremos saber si nuestro tratamiento está siendo efectivo, y sólo cuando el episodio ocurra podremos pronunciarnos retrospectivamente al respecto. Para otros objetivos, cuyas variables de respuesta son índices de estado (como disnea, capacidad funcional, rigidez matutina, dolor articular, etc.) o episodios contabilizables (número de crisis de angina, de episodios de taquicardia o de crisis epilépticas), sí es posible saber si el tratamiento está siendo efectivo o no, e incluso si está perjudicando. No obstante, en la actividad clínica ese pronunciamiento es frecuentemente difícil.

En resumen, existirán pacientes en los que el resultado del ensayo no es aplicable puesto que reúnen criterios de exclusión y no hubieran participado en el estudio. Otras veces, no se ha realizado el ensayo clínico que necesitaríamos para tomar nuestra decisión, o se ha realizado sobre preguntas diferentes de nuestros objetivos de tratamiento. Finalmente puede ocurrir que, aunque el tratamiento fue correctamente seleccionado, existan dudas sobre si está siendo efectivo, sobre la dosis adecuada y sobre si puede estar perjudicando o produciendo efectos adversos.

Cuando concurren algunas de estas circunstancias (tabla I) y se cumplen algunos requisitos de viabilidad (tabla II), la realización de un EC $n = 1$ es una alternativa a considerar, que suscita cada vez más interés^{10,11}.

Diseño de EC $n = 1$

En esencia, el EC $n = 1$ es una estrategia estructurada de "prueba de tratamientos" para determinar la efectividad de un tratamiento o identificar efectos adversos en un paciente. Estos estudios proceden de áreas de la psicología y la psiquiatría donde se han propuesto diversos diseños cuya descripción se aleja de los objetivos

TABLA I
¿CUÁNDO ESTÁ INDICADO HACER
EC N = 1? (PERTINENCIA)

- I. En relación con la efectividad del tratamiento
El clínico y/o el paciente dudan de la efectividad de la terapéutica administrada
El clínico duda si el tratamiento ha comenzado a actuar en su paciente
El paciente insiste en tomar un tratamiento que el clínico considera que no está siendo efectivo
- II. En relación con la dosis
Existen dudas respecto a la dosis óptima para este paciente
- III. En relación con los efectos adversos
El paciente experimenta síntomas que pudieran ser atribuibles al tratamiento (pero no es seguro)
- IV. El EC n = 1 como test diagnóstico
En aquellos test estímulo-respuesta cuya evaluación es clínica y existen dudas sobre la respuesta al mismo en nuestro paciente (p. ej., hidrocortisona en Addison o pidostigmina en miastenia grave)

EC n = 1: ensayos clínicos para un solo paciente.

de este artículo¹⁰⁻¹⁷. En la actividad clínica el diseño más empleado es el propuesto por Guyatt¹⁰⁻¹², y a él nos referiremos a partir de aquí.

Definición

El EC n = 1 es un estudio (fig. 1) en el cual se prueban en un solo paciente dos tratamientos, un tratamiento en cada subciclo, y se repite la prueba durante varios ciclos. El orden de los tratamientos dentro de cada ciclo es asignado aleatoriamente, y el procedimiento se realiza con doble enmascaramiento de médico y paciente (EC n = 1, ensayo con un solo sujeto, doble enmascaramiento y cruzamiento múltiple).

Desde un punto de vista formal, si se prueba un tratamiento y se observa una mejoría del paciente, ello puede ser debido a la eficacia real o al efecto placebo, o dicho de otro modo, podemos preguntarnos ¿existe una relación causal entre el fármaco y la mejoría, o simplemente se debe al hecho de tratar? Surgirá, pues, la pregunta: ¿si se hubiese administrado otro tratamiento –o placebo–, qué hubiera ocurrido? Este modelo de pregunta es denominada en filosofía de la ciencia *condicional subjuntivo*¹⁸. En los ensayos clínicos usuales, para contestar a esta pregunta se hace necesario disponer de un grupo de

TABLA II
REQUISITOS DE LA ENFERMEDAD Y TRATAMIENTOS

1. Que la enfermedad sea razonablemente estable en el tiempo, de modo que los posibles cambios bruscos en la enfermedad no confundán la evaluación	En este sentido las enfermedades agudas o rápidamente evolutivas no son susceptibles de ser estudiadas con EC n = 1
2. Que la acción de los tratamientos se alcance rápidamente	Fármacos o tratamientos que tarden en alcanzar efecto no son muy adecuados para estos estudios
3. Que al cesar el/los tratamientos la situación pueda considerarse idéntica al estado anterior, de modo que el segundo tratamiento actúe como si el primero no se hubiera administrado	Fármacos con largo período de eliminación (amiodarona) plantean problemas para realizar el EC n = 1. Procedimientos irreversibles (cirugía, angioplastia, etc.) no son adecuados para estos estudios
4. Que la variable respuesta sea de evaluación fácil y de pronta aparición	Para objetivos a medio o largo plazo no es pertinente (o posible) realizar estos estudios
5. Que el tratamiento deba mantenerse durante largo tiempo	En caso contrario, no es eficiente realizar el esfuerzo de un EC n = 1

EC n = 1: ensayos clínicos para un solo paciente.

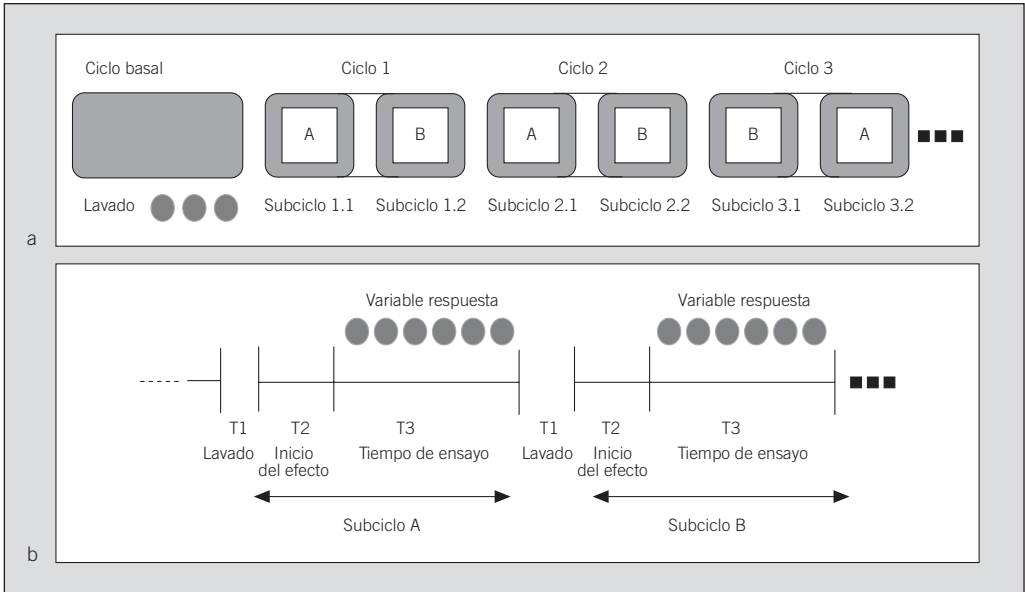


Fig. 1. a) Diseño de un EC $n = 1$. A: tratamiento A; B: tratamiento B. b) Detalle de un ciclo. T1: tiempo de lavado; T2: tiempo para el inicio del efecto; T3: tiempo de ensayo.

control tan parecido como sea posible al grupo de tratamiento, lo cual se consigue mediante la asignación aleatoria.

En nuestro caso, al tratarse de un solo paciente se plantea un problema con la selección del control; no obstante, si la enfermedad es crónica y estable, el tratamiento actúa con rapidez y al suspenderlo deja de actuar también rápidamente, se podrá comparar un paciente consigo mismo en dos momentos diferentes y con diferentes tratamientos. Sin embargo, podría ocurrir que influencias externas o ligeros cambios transitorios en la enfermedad afectaran a la comparación y produjeran que un tratamiento aparezca como mejor que otro por azar. Por ello, es necesaria la repetición de los ciclos de modo que la ventaja de un tratamiento sobre el otro no pueda ser atribuida al azar o a la casualidad.

Los requisitos de la enfermedad y de los tratamientos necesarios para realizar estos estudios se muestran en la tabla II; tales condiciones son comunes a los estudios autocontrolados y cruzados.

Pregunta principal, variable/s respuesta

Existirá un objetivo del estudio que será terapéutico, evaluable a corto plazo y no relacionado con el pronóstico a largo plazo (p. ej., mortalidad

coronaria). La pregunta principal se formulará de modo concreto y con contenido empírico definido. La formulación podrá hacerse como hipótesis de una o de dos colas. La variable resultado deberá ser explícita y evaluable. Por ejemplo, "en nuestro paciente, A es mejor/diferente que B para el tratamiento de su limitación crónica al flujo aéreo. El efecto lo mediremos mediante el cuestionario de Borg¹⁹ para la disnea, el consumo de betaestimulantes y la calidad del sueño, medido con cuestionarios específicos al paciente y al familiar".

La variable de respuesta puede ser un valor analítico (p. ej., potasio o hemoglobina A_{1c}), una imagen ecográfica o radiológica, un registro eléctrico, una prueba funcional, etc. No obstante, hay que señalar que estos ensayos se adaptan de modo especial a la valoración de síntomas y estados funcionales. Los posibles aspectos susceptibles de ser variable de respuesta de estos ensayos son innumerables y sólo limitados por la importancia que el paciente otorga al síntoma o problema concreto y la perspicacia del investigador clínico para la medición del síntoma o para la identificación y medición del constructo adecuado.

Desde un punto de vista teórico, la variable resultado de los EC $n = 1$ podrá ser: a) *variable úni-*

ca: número de crisis epilépticas; b) variables múltiples: disnea, uso de betaestimulantes y alteraciones del sueño; en estos casos podrán ser analizadas de modo separado, jerarquizando la variable principal y las secundarias, y c) un constructo con múltiples componentes (actividades de la vida diaria, calidad de vida y cuestionario general de síntomas) o un índice clinimétrico construido con diversas variables. Esta tercera alternativa tiene interés porque puede ofrecer una perspectiva global del beneficio general/perjuicio de uno u otro tratamiento.

Es recomendable que la/s variable/s resultado sean medidas varias veces en cada subciclo para permitir una mejor evaluación del paciente y para mejorar el análisis estadístico del estudio.

Ciclos y subciclos

En el EC $n = 1$ se denomina ciclo al período de tiempo en el que se administran dos tratamientos a prueba, una vez cada uno, y se observan las variables de respuesta. Un ciclo está formado por dos subciclos, cada uno de los cuales corresponde a uno de los dos fármacos. La duración de cada subciclo de tratamiento y el número de ciclos a realizar constituyen aspectos prácticos "cruciales" en el diseño de los EC $n = 1$; ambas cuestiones desempeñan un papel análogo al tamaño del estudio en los ensayos habituales, y se puede denominar *duración del estudio*.

Período basal, o ciclo basal. Siempre que sea clínicamente posible, es deseable la definición de un período de línea basal durante el que se efectúa el lavado de fármacos previos, se familiariza al paciente con el instrumento de medición, y se realizan las primeras mediciones repetidas. Es posible mantener medicaciones básicas durante todo el estudio y realizar el lavado sólo del tratamiento problema.

La definición de este período permite, así mismo, estabilizar las variaciones en la medición de la variable resultado y mejorar las posibilidades de análisis del estudio. Algunos autores¹⁰ proponen dos subciclos de tratamiento abierto para identificar posibles efectos adversos y de este modo evitar las posibles retiradas de pacientes durante el estudio.

Duración de los subciclos. El tiempo de duración de cada subciclo se calculará sumando los tiempos de lavado (T1), el tiempo de comienzo del efecto (T2) y el tiempo de ensayo (T3).

a) T1. Para evitar el efecto de sobrearrastre es necesario programar entre cada subciclo una

fase de lavado de los fármacos en estudio. Para su cálculo, se usará la información farmacocinética disponible sobre el tiempo que tardan los fármacos en eliminarse y cesar su efecto; en caso de amplia variabilidad farmacocinética se usarán los tiempos máximos, y cuando los diferentes fármacos tengan diferente período de lavado consideraremos el mayor de ellos. Si es necesario se puede realizar previamente, en línea basal, una prueba de eliminación.

La posible contraindicación de esta fase, o del lavado del período basal, debe ser evaluada por el clínico en cada caso.

b) T2. Para el cálculo del mismo se usará, igualmente, el conocimiento de la farmacocinética. También aquí puede ser de utilidad la determinación previa de concentraciones del fármaco en el paciente.

c) T3. Durante el tiempo de ensayo se observarán los efectos y se medirán las variables respuesta. Este período se calculará de diferente modo según el tipo de variable.

Si la variable de respuesta es un índice de estado, o prueba analítica, la duración del componente T3 del subciclo se calculará intuitivamente, de modo que permita la realización de múltiples mediciones. Si la variable respuesta es un episodio ocasional, o periódico, la cuestión es más compleja: si la variable principal $X =$ "número de episodios que ocurren en un intervalo temporal de duración t " se distribuye según un modelo de Poisson (que asume que los eventos son raros, aleatorios en el tiempo y con media constante), la variable $T =$ "tiempo necesario hasta que ocurra el primer episodio" sigue un modelo exponencial²¹.

La función de distribución acumulativa del modelo exponencial es $F(t) = 1 - e^{-\lambda t}$, siendo λ el número medio de episodios que ocurren en la unidad de tiempo; por tanto, para calcular el tiempo para el cual hay una probabilidad prefijada " p " de observar un episodio hay que resolver la ecuación $p = 1 - e^{-\lambda t}$, que puede expresarse como $\lambda t = -\ln(1-p)$, que para $p = 0,95$ es aproximadamente $t = 3/\lambda$. Dicho de otro modo, si en promedio ocurre un episodio cada tiempo t , el tiempo necesario para tener una probabilidad de 0,95 de observar un episodio es $3t$.

Por tanto, desde el punto de vista práctico se tratará de ver si el episodio tiene cierta regularidad, y a continuación calcularemos con qué periodicidad se produce éste; si multiplicamos por tres ese período tendremos una seguridad del 95% de observar al menos un episodio durante el mismo (regla del 3). Ese tiempo será el componente T3 del subciclo en el EC $n = 1$. Si la fre-

cuencia media de aparición del episodio no es constante, no se dan las presunciones para el uso del modelo de Poisson, y además no se dan los requisitos de regularidad en la enfermedad o en su manifestación para plantear un $EC\ n = 1$.

Número de ciclos. El número de ciclos a programar en el ensayo dependerá del modo de evaluar el efecto: si se va a evaluar cada ciclo en términos de éxito/fracaso de un tratamiento, asumiendo independencia entre ciclos se podrá formular una diferencia de éxitos mínima para considerar un fármaco mejor que otro y estimar, a partir de la distribución binomial, el número de repeticiones necesarias. Esta estimación es, sin embargo, la más desfavorable desde el punto de vista de la duración del estudio, que para el mejor de los casos deberá durar 5 ciclos (véase "Análisis de datos").

Si la variable de respuesta es un constructo, índice clínico, o un valor analítico, y puede ser medida muchas veces en cada subciclo, el número de ciclos desempeña un papel menos crítico y se calculará intuitivamente.

En ambos casos es aconsejable fijar *a priori* el número de ciclos, aunque, si la respuesta es evidente, se pueda decidir el final del estudio. En este sentido, la experiencia publicada²⁰ indica que la realización de menos de 3 ciclos difícilmente aporta evidencias clínicas.

Secuencia de tratamientos. El orden de los tratamientos en cada ciclo se asignará de modo aleatorio. Existen varios procedimientos. El más fácil es lanzar una moneda para cada ciclo y el resultado decide el orden de los dos tratamientos; alternativamente se puede elegir una secuencia de números aleatorios y fijar según ella (p. ej., par/impar) el orden de todos los ciclos. Una opción interesante es generar una serie de permutaciones de dos elementos, ordenarlas y asignar a cada una un número aleatorio, y la secuencia se ordenaría mediante esos números; esta opción permite generalizar el procedimiento para más de dos tratamientos. En todo caso, para preservar el enmascaramiento del médico el proceso de aleatorización debe ser ejecutado y conservado por alguien ajeno al estudio (p. ej., unidad de investigación).

Definición de la intervención y conducción del estudio

La intervención debe definirse de modo estricto y explícito (dosis, vía, preparados, etc.), y los placebos o tratamientos de control cuidadosa-

mente preparados; en este sentido, la colaboración de los farmacéuticos del centro es imprescindible. Es posible realizar el ensayo manteniendo en todos los ciclos un tratamiento base orientado a otros objetivos y poner a prueba otros tratamientos, a condición de que no existan interacciones entre tratamiento base y tratamientos en prueba, por ejemplo, se puede mantener tratamiento con anticoagulantes mientras se prueba un fármaco antiarrítmico.

Cuando la evaluación se refiere a síntomas, el enmascaramiento es imprescindible tanto en el clínico como en el paciente: en el clínico, para garantizar la "objetividad" en la evaluación de las variables y para evitar que existan mejores cuidados paralelos en algunos períodos. Ambos son problemas potenciales derivados de preferencias del clínico (conscientes o no) por alguna hipótesis; en el paciente, para evitar el efecto placebo o preferencias *a priori* por alguno de los tratamientos, así como el sesgo de obsequiosidad o tendencia a mostrar comportamientos que agraden al médico. Cuando se trate de variables de más fácil medición (valor analítico o ecográfico) el enmascaramiento es menos relevante, aunque deseable.

Los criterios de conclusión del estudio y las razones de interrupción del ciego deben ser explícitos. Un apartado relevante de la conducción del estudio es la comprobación de la observancia del tratamiento y se realizará mediante las estrategias clásicas de recuento de pastillas y el análisis de las concentraciones plasmáticas del fármaco²²; también es importante el mantenimiento de una relación "estrecha" de animación con el paciente. A este respecto, señalaremos que el $EC\ n = 1$ es un pacto explícito entre médico y paciente en beneficio de este último.

Análisis de datos

El modo de analizar los resultados es, sin duda, el aspecto más complejo y polémico de estos ensayos. Las propuestas van desde el uso de intuitivas soluciones gráficas a procedimientos estadísticos con diferente nivel de complejidad. En este artículo nos limitaremos a plantear el problema del análisis y enunciar algunas soluciones.

Como en cualquier experimento, el análisis de resultados trata de determinar si existen cambios en las variables respuesta en las diferentes condiciones experimentales, o dicho con mayor precisión, la pregunta es si los cambios entre las condiciones son mayores que dentro de cada condición experimental, y en su caso si el resul-

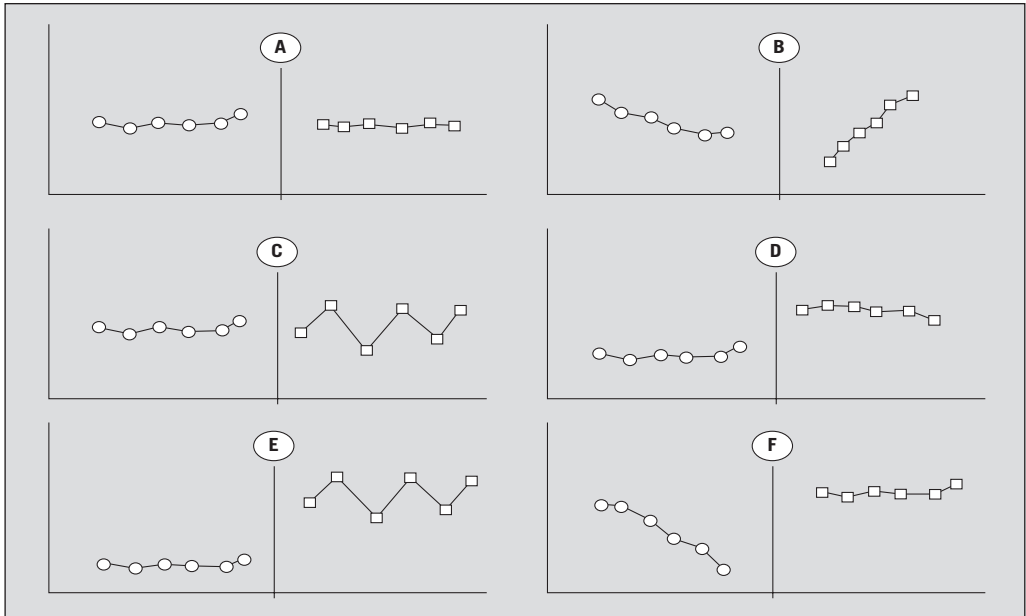


Fig. 2. Esquemas y criterios para el análisis gráfico. A: sin cambio entre 2 subciclos; B: cambio de tendencia; C: cambio de variabilidad; D: cambio de nivel; E: cambio de nivel y variabilidad; F: cambio de nivel y tendencia.

tado es repetible (de aquí la necesidad de varios ciclos).

Este enunciado general del análisis puede, no obstante, verse desde dos perspectivas: si consideramos el EC $n = 1$ como un procedimiento para decidir el mejor tratamiento para un paciente concreto, el resultado del ensayo debe ser claro y comprensible tanto para el médico como para el paciente. Este es el argumento que anima las propuestas de expresar gráficamente los resultados y realizar un análisis visual^{23,24}. El método consiste en presentar en un sistema de ejes cartesianos la variable resultado en función del tiempo, señalando en abscisas los diferentes ciclos y subciclos (fig. 2). Los criterios a considerar para el análisis visual son: a) la variabilidad en la medición general y entre ciclos, y b) cambios entre ciclos en la tendencia, en el nivel o en ambos²⁴.

Además de la sencillez, la ventaja del método es la "insensibilidad" para efectos nimios²⁵. Los inconvenientes residen en que la magnitud del efecto en una gráfica depende de la escala usada en las ordenadas, y que la interpretación presenta poca consistencia entre observadores²³ incluso con entrenamiento adecuado.

Como cuantificación de los cambios en cada ciclo se usa el denominado *split middle method*

(o línea de aceleración)²⁶, consistente en dibujar una recta que pase por los puntos que representan la mediana del resultado en las dos mitades del ciclo basal; a continuación se cuenta el número de puntos por encima y por debajo de esa línea en cada uno de los subciclos de tratamiento. Si no existiera efecto del tratamiento, el número de puntos por encima y por debajo de la recta debería ser idéntico en ambos grupos de subciclos de tratamiento (subciclos A y subciclos B). Esta hipótesis se puede contrastar usando la distribución binomial y asumiendo que, en caso de existir tendencia, ésta es lineal; no obstante, la validez de esta prueba ha sido cuestionada²⁷. Otro método gráfico análogo es la *resistant trend line* (o línea de resistencia de la tendencia)²⁶, que usa tres puntos al calcular la línea de tendencia; este método es más poderoso que el anterior, pero necesita un mínimo de 12 determinaciones basales.

Si, por el contrario, abordamos el análisis desde perspectivas numéricas, los resultados no serán tan subjetivos; sin embargo, existirán dificultades de realización e interpretación de las pruebas estadísticas. Los procedimientos convencionales como el ANOVA o la prueba de la *t* de Student pueden ser inadecuados porque exi-

gen independencia entre las observaciones; en los EC $n = 1$, donde se realizan múltiples determinaciones seriadas de un mismo sujeto, esta condición, en teoría, no se cumple, pues las observaciones estarán correlacionadas entre sí (autocorrelación serial). La importancia de esta autocorrelación es también objeto de debate; algunos autores insisten en ella^{23,24}, mientras que otros rechazan su existencia en la práctica¹¹. Por otra parte, estas pruebas informan sobre los cambios en el valor medio de las variables, cuando en estos estudios puede ser más relevante el cambio de tendencia²⁴.

Existen soluciones matemáticas para eludir esos problemas, la más característica de las cuales es el análisis de series temporales; este método cumple los supuestos estadísticos, pero su interpretación es compleja y precisa de un número de observaciones superior al usual en estos ensayos²⁴. Otra alternativa propuesta es realizar un contraste de hipótesis para encontrar dependencia serial, y en caso de no encontrarla en los datos proceder con pruebas convencionales. Esta última propuesta es problemática puesto que usa a su favor la asimetría de los contrastes de hipótesis en situación de bajo tamaño muestral (por tanto, de baja potencia). De este modo, es probable no poder demostrar dependencia serial en las observaciones aun cuando exista.

Cuando la variable de respuesta es binomial, es decir éxito/fracaso del tratamiento A, la prueba de los signos puede ser adecuada. Con ella se calcula la probabilidad de la serie de respuestas obtenidas en la hipótesis de que la respuesta se deba al azar. En la hipótesis de que los dos tratamientos son iguales, la probabilidad de que un paciente prefiera el tratamiento A en 4 ciclos seguidos será $1/2 \times 1/2 \times 1/2 \times 1/2 = 0,0625$. El principal inconveniente es el número de ciclos necesarios para llegar al nivel de significación convencional, además de que esta prueba también exige independencia.

Una propuesta reciente²⁸, particularmente útil si se usan distintas variables independientes, es definir un "índice beneficio/perjuicio" que englobe todas las variables e integre las dependencias temporales, y sobre ese índice aplicar las pruebas estadísticas o gráficas convencionales.

Hay que hacer notar, sin embargo, que la aparente solidez del análisis estadístico presenta dos puntos críticos: el primero se refiere a la pertinencia de usar números procedentes de escalas ordinales o casi dimensionales como si fueran proyecciones sobre la recta de los números reales, y el segundo se refiere a la escasa consistencia de los resultados según se usen unos u otros

métodos estadísticos, y que ha sido demostrada en los estudios de Nourbakhsh y Ottenbacher²⁹.

En resumen, respecto al análisis de los EC $n = 1$ existen problemas relacionados con la opinión sobre el objetivo del ensayo, y también dificultades de índole metodológica estadística. En todo caso, se ha producido una importante progresión en estos aspectos y ya existe un interesante catálogo de recursos a usar por el investigador. Por ello, la recomendación general es evitar un análisis estándar, diversificando el análisis mediante la combinación de métodos estadísticos con métodos gráficos, y adaptar el mismo a las peculiaridades de cada situación y a la comprensión del clínico y del paciente.

Utilidad de los estudios

En este punto, es preciso plantearse: ¿en qué medida estos estudios son útiles para decidir sobre el tratamiento en pacientes, o en general sobre su manejo clínico?, y también cabe preguntarse ¿cómo pueden contribuir estrategias de este tipo a la mejoría de la evaluación de la efectividad de tratamientos en estudio?

Respecto a la primera cuestión, es obvio que no se trata de un diseño adecuado para todo tipo de pacientes, sino sólo una solución pertinente para pacientes determinados y que además reúnan algunos requisitos respecto a la enfermedad y a los tratamientos a aplicar. La experiencia del grupo de McMaster^{20,30,31}, sobre 73 ensayos $n = 1$, indica que existe un amplio catálogo de enfermedades crónicas en las que han sido usados, aunque confirma que existen muchas situaciones en las que no es posible el uso. Así mismo, señala que existe un número de pacientes y clínicos (7 y 1, respectivamente) que rompen el protocolo del estudio, y en 8 pacientes el ensayo no se pudo comenzar o hubo de suspenderse precozmente por enfermedades o episodios intercurrentes.

Sin embargo, de los 57 ensayos concluidos en el 88 % se obtiene información clínica definitiva, aunque sólo en el 33 % del total se obtuvo significación estadística (el 35 % de los que ofrecen información definitiva), datos que refuerzan los comentarios realizados en el apartado del análisis en relación con el relativo valor de la estadística. Dentro de los ensayos concluidos, en el 37 % el resultado del EC $n = 1$ indujo el cambio del tratamiento previsto inicialmente por el clínico, y en el 43 % el resultado indujo el mantenimiento del mismo. Tras el EC $n = 1$ la confianza en la idoneidad del tratamiento fue alta o muy alta en la mayoría de los clínicos.

Otros estudios^{32,33} indican resultados análogos respecto a la utilidad, y calculan los costes por ensayo³² en 400 dólares de 1991, si bien no ofrecen un cálculo de los costes derivados de la interrupción de las medicaciones, de los gastos de monitorización de tratamientos y del tratamiento de los efectos adversos.

En resumen, cabe concluir que la realización de EC $n = 1$ es un área de frontera que concilia la investigación y la acción en la clínica, que puede mejorar las acciones clínicas, particularmente en casos complejos, y que su realización pudiera ser coste-efectiva. Si el EC $n = 1$ es una estrategia alternativa a la “prueba de tratamientos” habitual de la actividad clínica, ha de responderse mediante la comparación de ambas estrategias con diseños adecuados como los propuestos por Mahon et al³⁴.

Respecto a la segunda pregunta, los EC $n = 1$ incluidos en ensayos paralelos tradicionales pueden mejorar la evaluación de la efectividad del tratamiento en cada individuo de la población de estudio. En este sentido, una estrategia de múltiple cruzamiento y evaluación individual³⁵ ha sido utilizada en ensayos clásicos sobre insuficiencia cardíaca y digital en 1984. Recientemente se ha propuesto²⁸ el uso de un diseño de múltiples EC $n = 1$ y un tratamiento estadístico especial para mejorar la evaluación de la efectividad en ensayos clínicos paralelos orientados hacia las variables de estado (calidad de vida y síntomas). Por tanto, los EC $n = 1$ pueden también ser un instrumento para la progresión en la investigación y evaluación de fármacos o acciones terapéuticas.

Aspectos bioéticos de los EC $n = 1$

Los ensayos clínicos implican la utilización de personas en la construcción de conocimiento y ello plantea un problema desde la visión kantiana de la dignidad del hombre como un fin en sí mismo³⁶. Esta situación, inevitable en la investigación en humanos, plantea cuestiones bioéticas de interés.

El eje de la reflexión bioética asienta en los tres principios señalados por el informe Belmont³⁷: *justicia*, *autonomía* y *beneficencia*. En algunas propuestas posteriores el principio de beneficencia puede contemplarse como subdividido en dos principios con diferente nivel de obligación³⁸: serían los principios de “no maleficencia” y “beneficencia”. La jerarquía de principios podría plantearse en dos niveles de exigencia, un nivel superior donde se ubican los principios de “justicia” y “no maleficencia” (ambos exigibles), y un nivel inferior donde se ubicarían autonomía y beneficencia, principios jerarquizados por ese orden.

En la clínica, los actores de esos principios están claramente definidos; la autonomía sería el principio propio del paciente y la beneficencia el del médico, por su parte sociedad y administración velarían por los principios de “justicia” y de “no maleficencia”³⁸. En el ensayo clínico, el principio de beneficencia se difumina, puesto que el posible beneficio para el paciente o para futuros pacientes sólo es una posibilidad teórica³⁹ (se trata de un tratamiento en prueba) y, por otra parte, existen posibilidades de perjudicarlo. Por ello, en un primer nivel los principios de justicia y no maleficencia deben ser preservados, pero el principio prevalente es el de autonomía, cuya expresión es el *consentimiento informado*. En España, los comités éticos de investigación tienen el encargo legal⁴⁰ de velar por el cumplimiento de estos principios.

En el EC $n = 1$, los beneficios obtenidos de la realización de un ensayo clínico son directamente aplicables a la persona objeto del estudio; por tanto, el principio de beneficencia recupera concreción y se mantiene vigente el principio de autonomía. Como se señaló, el EC $n = 1$ sólo es posible concebido como un pacto explícito entre médico y paciente para conocer, con el mejor método posible, el mejor tratamiento para éste.

En nuestro país no existe derecho positivo en relación con los EC $n = 1$, excepción hecha de un comentario sobre el uso compasivo de fármacos en un solo paciente. En el mundo anglosajón existen pronunciamientos¹¹ de comités éticos de investigación que argumentan que cuando el medicamento ya está probado no es preciso el pronunciamiento explícito de ellos sino de un comité de ética clínica, opinión que es compartida con algunos autores⁴¹.

Finalmente, señalaremos que los EC $n = 1$ suponen un desplazamiento en el centro de gravedad en la investigación terapéutica. En efecto, los síntomas serán el objetivo principal en estos ensayos, frente a los objetivos clásicos de duración de vida, etc. Este desplazamiento es producto de la emergencia del principio de autonomía y plantea, como reto para la investigación clínica, la necesidad de progresar en el desarrollo de técnicas para medición de estados y síntomas (clinimetría)⁴².

La emergencia del principio de autonomía no es un objetivo en sí mismo, pero con ella la relación médico-enfermo tiende a equilibrarse en perjuicio del tradicional paternalismo médico; ese equilibrio es esencial para la consecución de una relación médico-enfermo que supere el puro paternalismo y el puro autonomismo, para fundarse en lo que Laín Entralgo⁴³ denomina *amistad médica*.

Agradecimiento

A los Dres. Jaime Latour, José J. Mira, César Picó y Francisco Pozo por sus valiosos comentarios a versiones previas de este manuscrito.

Así mismo, estamos en deuda con Alicia Picó por su ayuda en la edición del manuscrito.

BIBLIOGRAFÍA

- Durán JA, Figuerola J. Cumplimiento de la medicación: características, factores que lo determinan y recomendaciones para manejarlo. *Med Clin (Barc)* 1988; 90: 338-343.
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Decidir la mejor terapéutica. En: Sackett DL, Haynes RB, Guyatt GH, Tugwell P, editores. *Epidemiología clínica. Ciencia básica para la medicina clínica (2ª ed.)*. Madrid: Editorial Médica Panamericana S.A., 1994; 191-248.
- Taylor DC. The components of sickness: diseases, illnesses and predicaments. *Lancet* 1979; 2: 1.008.
- Patel VL, Evans DA, Groen GJ. Biomedical knowledge and clinical reasoning. En: Evans DA, Patel VL, editores. *Cognitive Science in Medicine. Biomedical modeling*. Cambridge, Massachussets: Massachussets Institute of Technology Press, 1989; 53-85.
- Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *Br Med J* 1995; 310: 1.122-1.126.
- Guyatt GH, Sackett DL, Cook DJ. How to use an article about therapy or prevention. *JAMA* 1993; 270: 2.598-2.601.
- Angell M. Patient's preferences in randomized clinical trials. *N Engl J Med* 1984; 310: 1.385-1.387.
- Álvarez-Dardet C, Bolumar F, Porta Serra M. La medición de frecuencia de enfermedad. *Med Clin (Barc)* 1987; 88: 287-291.
- Kramer MS, Lane DA. Causal propositions in clinical research and practice. *J Clin Epidemiol* 1992; 45: 639-649.
- Guyatt GH, Sackett D, Adachi J, Roberts R, Chong J, Rosebloom D et al. A clinician's guide for conducting randomized trial in individual patients. *Can Med Assoc J* 1988; 139: 497-503.
- Guyatt GH, Sackett D, Taylor W, Chong J, Roberts R, Pugsley S. Determining optimal therapy—randomized trials in individual patients. *N Engl J Med* 1986; 314: 889-892.
- Keller JL, Guyatt GH, Roberts RS, Adachi JD, Rosebloom D. An N of 1 service: applying the scientific method in clinical practice. *Scand J Gastroenterol* 1988; 23 (Supl 147): 22-29.
- Barlow DH, Hersen M. Single-case experimental designs. *Arch Gen Psychiatry* 1973; 29: 319-325.
- Barlow DH, Hersen M. Single-case experimental designs. *Strategies for studying behavior change (2ª ed.)*. Nueva York: Pergamon Press, 1984.
- Kratchwill TR, editor. *Single subject research. Strategies for evaluating change*. Orlando: Academic Press, 1978.
- Kazdin AE. *Single-case research designs. Methods for clinical and applied settings*. Nueva York: Oxford University Press, 1982.
- Spilker B. Single-patient clinical trial. En: Spilker B, editor. *Guide to clinical trials*. Nueva York: Raven Press, 1991; 277-282.
- Wartofsky MW. Introducción a la filosofía de la ciencia. Madrid: Alianza Editorial, 1987; 326-329.
- Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 1970; 23: 92-98.
- Guyatt GH, Keller JL, Jaeschke R, Rosebloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: clinical usefulness. Our three years experience. *Ann Intern Med* 1990; 112: 293-299.
- Abraira V, Pérez de Vargas A. *Bioestadística*. Madrid: Ed. Centro de Estudios Ramón Areces, 1996.
- Privitera MD, Treiman DM, Pledger GW, Sahlroot JT, Handforth A, Linde MS et al. Dezinamide for partial seizures: results of an n-of-1 design trial. *Neurology* 1994; 44: 1.453-1.458.
- Ottenbacher KJ. Reliability and accuracy of visual analyzing graphed data from single-subject designs. *Am J Occup Ther* 1986; 40: 464-469.
- Wolery M, Harris SR. Interpreting results of single-subject research designs. *Phys Ther* 1982; 62: 445-452.
- Baer DM. Perhaps it would be better not to know everything. *J Appl Behav Anal* 1977; 10: 167-172.
- Ottenbacher KJ. Analysis of data in idiographic research. *Am J Phys Med Rehab* 1992; 71: 202-208.
- Crosbie J. The inability of the binomial test to control type I error with single-subject data. *Behav Assess* 1987; 9: 141-150.
- Bagne CA, Lewis RF. Evaluating the effects of drugs on behavior and quality of life: an alternative strategy for clinical trials. *J Consult Clin Psychol* 1992; 60: 225-239.
- Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: a comparative examination. *Phys Ther* 1994; 74: 768-776.
- Patel A, Jaeschke R, Guyatt GH, Keller JL, Newhouse T. Clinical usefulness of n-of-1 randomized trials in patients with nonreversible chronic airflow limitation. *Am Rev Respir Dis* 1991; 144: 962-964.
- Jaeschke R, Adachi J, Guyatt GH, Keller J, Wong B. Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 n-of-1 randomized controlled trials. *J Rheumatol* 1991; 18: 447-451.
- Larson EB, Ellsworth AJ, Oas J. Randomized clinical trials in single patients during a 2 year period. *JAMA* 1993; 270: 2.708-2.712.
- March L, Irwing L, Schwartz J, Simpson J, Chock C, Brooks P. n-of-1 trials comparing a non-steroidal antiinflammatory drug with paracetamol in osteoarthritis. *Br Med J* 1994; 309: 1.041-1.046.
- Mahon JL, Laupacis A, Donner A, Wood T. A randomized trial of n-of-1 trials versus conventional therapy. *Clin Res* 1993; 4: 180A.

35. Lee DC, Johnson RA, Bingham JB, Leahy M, Dinsmore RE, Goroll AH et al. Heart failure in outpatients. A randomized trial of digoxin versus placebo. *N Engl J Med* 1982; 306: 699-705.
36. Rawls J. *A theory of justice*, Cambridge Mass. Belknap Press of Harvard University Press 1971 [traducción española]. Madrid: FCE, 1979.
37. Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. Washington, US Government: OPRR reports, 1979.
38. Gracia D. *Fundamentos de bioética*. Madrid: Eudema Universidad, 1989; 103.
39. Hellman S, Hellman DS. Of mice but no men. Problems of the randomized clinical trial. *N Engl J Med* 1991; 324: 1.585-1.589.
40. Real Decreto 561/1993 de 16 de abril por el que se establecen los requisitos para la realización de Ensayos Clínicos con medicamentos. BOE, 13 de mayo de 1993; 14.346-14.364.
41. Irwing L, Glasziou P, March L. Ethics of n-of-1 trials. *Lancet* 1995; 345: 469.
42. Feinstein AR. *Clinimetrics*. New Haven: Yale University Press, 1987.
43. Laín Entralgo P. *El médico y el enfermo*. Madrid: Alianza Editorial, 1987.