



ORIGINAL

Evaluación de resultados clínicos individuales: a propósito de un caso tratado con terapia electroconvulsiva



Ioseba Iraurgi^{a,*}, Susana Gorbeña^a, Miren-Itxaso Martínez-Cubillos^b,
Margarita Escribano^c y Pablo Gómez-de-Maintenant^a

^a Departamento de Personalidad, Evaluación y Tratamientos Psicológicos, Universidad de Deusto, Bilbao, Vizcaya, España

^b Fundación Eragintza, Bilbao, Vizcaya, España

^c Unidad de Hospitalización, Avances Médicos (AMSA), Bilbao, Vizcaya, España

Recibido el 1 de julio de 2014; aceptado el 9 de julio de 2014

Disponible en Internet el 3 de octubre de 2014

PALABRAS CLAVE

Evaluación clínica;
Índice de cambio
fiable;
Efectividad;
Estudio de caso

Resumen La valoración de resultados terapéuticos y la evaluación de la eficacia y efectividad de los tratamientos es un área de interés para clínicos e investigadores. Habitualmente se utiliza la metodología basada en diseños controlados aleatorizados y la comparación intergrupos con un mínimo de participantes en cada brazo de tratamiento. No obstante, estos procedimientos raramente pueden ser aplicados en la práctica clínica habitual, si bien es importante para el clínico evaluar si los resultados del tratamiento propuesto para una persona específica cumplen con los objetivos terapéuticos, tanto por su significación estadística como por su relevancia clínica. La metodología basada en el índice de cambio fiable (Jacobson y Truax)¹ permite estimar esta doble consideración. El objetivo del artículo es proponer un procedimiento de aplicación a través del estudio de un caso con diagnóstico de depresión mayor y tratado con terapia electroconvulsiva.

© 2014 SEP y SEPB. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

KEYWORDS

Clinical assessment;
Reliable change
index;
Effectiveness;
Case study

Assessment of individual clinical outcomes: Regarding an electroconvulsive therapy case

Abstract Evaluation of therapeutic results and of the efficacy and effectiveness of treatments is an area of interest both for clinicians and researchers. In general, randomized controlled trial designs have been used as the methodology of choice in which intergroup comparisons are made having a minimum of participants in each arm of treatment. However, these procedures are seldom used in daily clinical practice. Despite this fact, the evaluation of treatment results for a specific patient is important for the clinician in order to address if therapeutic goals have

* Autor para correspondencia.

Correo electrónico: ioseba.iraurgi@deusto.es (I. Iraurgi).

been accomplished both in terms of statistical significance and clinical meaningfulness. The methodology based on the reliable change index (Jacobson y Truax)¹ provides an estimate of these two criteria. The goal of this article is to propose a procedure to apply the methodology with a single case study of a woman diagnosed with major depression and treated with electroconvulsive therapy.

© 2014 SEP y SEPB. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

La evaluación es un elemento clave en cualquier proceso clínico para el diagnóstico, la planificación del tratamiento y la valoración de sus efectos. Si los resultados son los esperados, se considera que la intervención ha sido exitosa y se concluye con el alta terapéutica, pero si no lo son es necesario tomar nuevas decisiones, realizar cambios, eliminar el tratamiento prescrito o adoptar nuevas recomendaciones terapéuticas. En cualquier caso, la información sobre los resultados clínicos de la intervención es de suma importancia para los profesionales sanitarios².

La evaluación de la eficacia terapéutica basada en pruebas científicas utiliza una comparativa de grupos expuestos a diferentes intervenciones: uno de ellos es el objeto de la evaluación, el otro cuenta con un valor terapéutico demostrado. El diseño básico se conoce como ensayo clínico aleatorizado, un procedimiento en el que se define un protocolo para evitar los efectos del sesgo e identificar la causalidad de la intervención en el efecto terapéutico. El control de la variabilidad grupal y el empleo de la estadística permiten la toma de decisiones respecto a la causa (azar o intervención) de los cambios observados³.

Sin embargo, en la práctica clínica diaria, los profesionales de la salud rara vez cuentan con grupos numerosos que permitan comparaciones, y muy a menudo sus decisiones se basan en casos concretos. El procedimiento habitual consiste en considerar el cambio experimentado por el individuo mediante el análisis de las variaciones observadas en el estado anímico o rasgo de interés, de acuerdo con índices biológicos, sintomáticos o órcticos. Por ejemplo, en materia de salud mental, la terapia tiene como objetivo obtener cambios en el estado anímico (por ejemplo, en los signos y síntomas depresivos) o en los rasgos de personalidad (por ejemplo, en el neuroticismo). Para lograrlo se evalúan los estados o rasgos con instrumentos estandarizados que permitan calificar el caso respecto a un grupo normativo. Por lo tanto, si se evalúa a una persona con el mismo instrumento antes y después del tratamiento, la diferencia en las puntuaciones puede ofrecer una puntuación determinada de los cambios mediante incrementos, descensos o igualdad en el valor. Además, los resultados se pueden analizar con referencia a una muestra normativa, lo que permite situar a la persona dentro de un intervalo funcional o disfuncional. Para este fin, en los instrumentos de este tipo se suelen fijar unos puntos de corte (PC) que permiten la toma de decisiones clínicas.

Sin embargo, en comparación con la metodología empleada en los estudios clínicos aleatorizados para la determinación de la eficacia del tratamiento, no disponemos

de herramientas que establezcan la significación clínica y estadística del cambio observado en los casos individuales. Necesitamos saber si el cambio posee una magnitud suficiente como para ser considerado un resultado clínico importante y descartar el azar. El objetivo de este trabajo, basado en un caso clínico tratado con terapia electroconvulsiva (TEC), es ofrecer una respuesta a esta interrogante mediante la aplicación del cálculo del índice de cambio fiable (RCI)^{1,4} para la toma de decisiones clínica y estadísticamente significativas en casos individuales.

Caso clínico

Se trata de una paciente de 45 años de edad, casada y madre de 2 hijos. Es profesora y se encuentra de baja por enfermedad. Fue ingresada en la unidad de psiquiatría de un hospital general y se trata de su quinto ingreso. Fue derivada a un centro psiquiátrico debido a síntomas depresivos graves durante las últimas semanas en el contexto de un trastorno bipolar.

La paciente presentaba síntomas depresivos, especialmente incapacidad para disfrutar de las actividades diarias exacerbada por la mañana, astenia intensa, lentitud psicomotora, dolores de cabeza tensionales, deterioro atencional y de concentración así como pensamientos no delirantes de inadecuación, abandono, discapacidad y culpa.

Su funcionamiento diario se caracteriza por ser autoexigente, y se presenta como persona muy respetuosa, cortés y bien educada, hasta el punto de ser rígida por un firme sentido del deber. Sus valores mostraron un fuerte compromiso con sus responsabilidades en torno a su papel como buena madre, hija, esposa y trabajadora. Muestra rasgos obsesivos (orden y control) y una negación notable de cualquier manifestación de hostilidad. Declara poseer una rutina diaria muy organizada que le mantiene estable.

Su historia clínica indica hipotiroidismo tratado con hormonas. Con referencia a los antecedentes psiquiátricos, relata varios problemas afectivos desde joven, inicialmente un signo depresivo que remitió con benzodiazepinas, e incluso de forma espontánea. Hace 3 años presentó un episodio maniaco con síntomas psicóticos y estados de ánimo congruentes e incongruentes (paranoide). Refiere que no existen antecedentes familiares psiquiátricos.

La paciente señala que su primer episodio depresivo ocurrió hace 6 años, tras el fallecimiento de su padre. Fue tratada con antidepresivos, que continúa tomando en la actualidad. Inicialmente respondió de forma rápida y eficaz a los antidepresivos IRSA con una recuperación completa. Desde entonces, se han producido al menos 5 episodios

depresivos con una respuesta al tratamiento más deficiente, una recuperación más lenta, sin alcanzar un verdadero estado eutímico, y el mantenimiento de síntomas depresivos subclínicos que interfieren con el funcionamiento diario y la calidad de vida. Se han realizado varios cambios en el tratamiento: prescripción de fármacos duales (venlafaxina) y posteriormente, una combinación de antidepresivos a dosis máximas (venlafaxina y mirtazapina) y la incorporación de otros fármacos estabilizadores del ánimo (sal de litio o ácido valproico). La intensidad y la discapacidad funcional causadas por los síntomas requirieron tratamiento hospitalario. También recibe psicoterapia cognitiva una vez a la semana y muestra buena adherencia en los últimos 5 años.

Tras varias intervenciones farmacológicas sin buenos resultados terapéuticos (la última, clorimipramina 225 mg/día, que se suspendió debido a la falta de respuesta y a efectos secundarios graves) se le recomienda TEC. Se programan 12 sesiones bilaterales, 3 por semana en días alternos. Se observó una respuesta muy favorable. Durante las entrevistas realizadas tras la tercera sesión presenta estado eutímico, con una mejora significativa en todos los síntomas. Dada la respuesta positiva, se administraron solo 10 sesiones y se recomendaron sesiones de mantenimiento tras el alta siguiendo las recomendaciones de Martínez-Amoros et al.⁵.

Cuando la paciente fue ingresada para el tratamiento con TEC, se realizó una entrevista clínica a fin de evaluar la frecuencia y la intensidad de los síntomas depresivos y los rasgos de personalidad. Se utilizó la versión española adaptada de 3 instrumentos estandarizados: la escala de depresión de Hamilton (HAM-D)⁶, el inventario de depresión de Beck (DBI)⁷ y el NEO-PI-R⁸, un test de personalidad ampliamente utilizado que mide 5 rasgos: Neuroticismo, Extraversión, Apertura a la experiencia, Afabilidad, Responsabilidad. Se utilizaron el mismo protocolo e instrumentos al final de la segunda semana. La [tabla 1](#) muestra todas las puntuaciones en las 3 escalas antes y después del tratamiento.

Se obtuvo el consentimiento de la paciente tanto para la evaluación y como para la TEC.

Metodología de la evaluación: índice de cambio fiable

El RCI constituye un conjunto de procedimientos estadísticos utilizados en muchos campos de la evaluación psicológica que sirve para determinar si las puntuaciones de un individuo en una prueba psicométrica han cambiado respecto a la evaluación anterior con el mismo instrumento⁴. Jacobson y Truax¹ describieron por primera vez las técnicas de cambio fiable diseñadas para determinar cambios terapéuticos significativos en pacientes que habían experimentado un trastorno o una intervención clínica y aplicarlas en la toma de decisiones.

Primero propusieron una definición concreta de lo que consideraban cambio clínicamente significativo en terapia: «cuando el cliente se mueve del rango disfuncional al funcional durante el curso de la terapia»¹ y para ello se precisa de instrumentos estandarizados que cuenten con un sistema baremado de puntuaciones y la fijación de puntos de corte que delimiten las áreas funcionales y disfuncionales. En nuestro caso, se utilizaron los datos normativos españoles

de las siguientes escalas: HAM-D⁶ con un alfa de Cronbach de 0,82; media de 5,8 y desviación estándar de 2,2; DBI⁷ con un alfa de Cronbach 0,83; media de 5,4, y desviación estándar de 5,4 y el NEO PI-R⁸. Los PC para los 3 instrumentos se presentan en la [tabla 1](#).

Adicionalmente estos autores, y otros posteriores (ver Iraurgi⁴ para una revisión) proponen que la forma más efectiva de determinar hasta qué punto una puntuación individual ha cambiado de forma fiable y significativa consiste en expresar la magnitud del efecto alcanzado, para lo cual ponen en relación la magnitud del cambio (diferencia entre el postest y el pretest) respecto a un error de estimación de medida que proviene, por una parte, de la precisión del instrumento de evaluación utilizado y, por otra, de la variabilidad de puntuaciones dentro del grupo normativo. Esta estimación constituye el RCI y los cálculos necesarios son los siguientes:

Paso 1. Cálculo del error estándar de la media (S_E): $S_E = S_x (1-r_{xx})^{1/2}$, o $S_E = S_x \sqrt{(1-r_{xx})}$

Paso 2. Cálculo del error estándar de la diferencia (SE_{dif}): $SE_{dif} = (2S_E^2)^{1/2}$, o $SE_{dif} = S_E \sqrt{2}$

Paso 3. Cálculo de la puntuación de RCI: $RCI = (x_2 - x_1) / SE_{dif}$

Donde S_x es la desviación típica de las puntuaciones de la prueba en el grupo normativo; r_{xx} es el coeficiente de fiabilidad de la prueba; x_2 es la puntuación tras la prueba y x_1 es la puntuación obtenida antes de la prueba. El RCI resultante se expresa utilizando las puntuaciones estandarizadas de z : valores $\leq -1,96$ o $\geq 1,96$ se consideran estadísticamente significativos con una probabilidad inferior al 5% ($p \leq 0,05$). La [tabla 1](#) muestra el RCI para las escalas y subescalas de los 3 instrumentos utilizados con esta paciente, obtenido con el algoritmo descrito y utilizando los índices de fiabilidad y la desviación estándar de las escalas. Se diseñó una aplicación con Excel para realizar los cálculos.

Finalmente, Jacobson y Truax¹ definen un cambio clínico significativo como aquel que, por una parte, alcanza el PC en términos de nivel de funcionamiento y, por otra, presenta una magnitud estadísticamente significativa. 'No cambio' (cuando no se cumple el criterio de magnitud de cambio clínico ni estadístico), 'Cambio No Significativo' (cuando se sobrepasa un punto de corte clínico pero no hay significación estadística), 'Cambio Significativo' (cuando se produce un resultado estadísticamente significativo pero no sobrepasa el punto de corte) y 'Cambio Clínicamente Significativo' (cuando además de producirse un resultado estadísticamente significativo se sobrepasa el punto de corte hacia la funcionalidad). La [tabla 1](#) muestra la categoría alcanzada en este estudio de caso en las escalas y subescalas de los instrumentos utilizados.

Para ilustrar los resultados utilizaremos como ejemplo la HAM-D (última columna/fila en [tabla 1](#)). La paciente obtuvo antes de someterse a la TEC una puntuación de 28, lo que le sitúa muy por encima del punto de corte que señala la clasificación de 'depresión grave' (PC = 19). Por tanto, la paciente se encontraba con alta intensidad y frecuencia de sintomatología que la localiza por encima del percentil 99 en el baremo normativo. Tras la intervención, la puntuación en el HAM-D es de 12, lo que le sitúa entre los puntos de corte de depresión leve (PC = 8) y moderada (PC = 14), si bien le

Tabla 1 Cambio clínica y estadísticamente significativo en los factores NEO-PI-R, el DBI y la HAM-D

	Pretratamiento	Postratamiento	Puntos de corte clínicos	Cambio clínico	Diferencia cambio	RCI	Cambio estadísticamente significativo*	Interpretación
<i>NEO-PI-R</i>								
Neuroticismo	63	58	35/65	No	-5	-0,913	No	Sin cambio (SC)
Extraversión	37	37	35/65	No	0	0,000	No	SC
Apertura a la experiencia	37	37	35/65	No	0	0,000	No	SC
Afabilidad	58	57	35/65	No	-1	-0,183	No	SC
Responsabilidad	53	47	35/65	No	-6	-1,096	No	SC
Ansiedad	63	58	35/65	No	-5	-0,913	No	SC
Hostilidad	43	56	35/65	No	+13	2,373	Sí	Cambio significativo
Depresión	68	51	35/65	Sí	-17	-3,104	Sí	Cambio clínico significativo
Ansiedad social	68	60	35/65	Sí	-8	-1,461	No	Cambio no significativo
Impulsividad	49	53	35/65	No	+4	0,730	No	SC
Vulnerabilidad al estrés	73	66	35/65	No	-7	-1,278	No	SC
Cordialidad	37	43	35/65	No	+6	1,096	No	SC
Gregarismo	56	42	35/65	No	-14	-2,556	Sí	Cambio significativo
Asertividad	29	37	35/65	Sí	+8	1,461	No	Cambio no significativo
Actividad	42	47	35/65	No	+7	1,278	No	SC
Búsqueda de emociones	43	50	35/65	No	+7	1,278	No	SC
Emociones positivas	37	37	35/65	No	0	0,000	No	SC
Fantasía	34	37	35/65	Sí	+3	0,548	No	Cambio no significativo
Estética	47	47	35/65	No	0	0,000	No	SC
Sentimientos	43	46	35/65	No	+3	0,548	No	SC

Tabla 1 (continuación)

	Pretratamiento	Postratamiento	Puntos de corte clínicos	Cambio clínico	Diferencia cambio	RCI	Cambio estadísticamente significativo*	Interpretación
Acciones	37	40	35/65	No	+3	0,548	No	SC
Ideas	42	40	35/65	No	-2	-0,365	No	SC
Valores	43	40	35/65	No	-3	-0,548	No	SC
Confianza	57	57	35/65	No	0	0,000	No	SC
Franqueza	73	60	35/65	Sí	-13	-2,373	Sí	Cambio clínico significativo
Altruismo	58	47	35/65	No	-11	-2,008	Sí	Cambio significativo
Actitud conciliadora	63	57	35/65	No	-6	-1,096	No	SC
Modestia	50	57	35/65	No	+7	1,278	No	SC
Sensibilidad a los demás	31	44	35/65	Sí	+13	2,373	Sí	Cambio clínico significativo
Competencia	57	44	35/65	No	-13	-2,373	Sí	Cambio significativo
Orden	53	49	35/65	No	-4	-0,730	No	SC
Sentido del deber	57	43	35/65	No	-14	-2,556	Sí	Cambio significativo
Necesidad de logro	53	50	35/65	No	-3	-0,548	No	SC
Autodisciplina	34	47	35/65	Sí	+13	2,373	Sí	Cambio clínico significativo
Deliberación	60	53	35/65	No	-7	-1,278	No	SC
Deseabilidad Social	63	46	35/65	No	-17	-3,104	Sí	Cambio significativo
<i>Escalas sintomáticas de depresión</i>			#					
			L / M / G					
BDI	30	8	10 / 19 / 30	Sí	-22	-6,987	Sí	Cambio clínico significativo
HAM-D	28	12	8 / 14 / 19	Sí	-16	-12,531	Sí	Cambio clínico significativo

Puntos de corte para DBI y HAM-D: L = depresión Leve; M = Moderada; G = Grave.

* Valores de RCI > 1,96 o > -1,96 resultan estadísticamente significativos para $p < 0,05$. A medida que se incrementa el valor de la RCI menor es la probabilidad de que el cambio observado sea debido al azar.

situaría todavía por encima del percentil 95 respecto a la distribución normativa. La magnitud del cambio, por tanto, es de -16 puntos, que lleva asociado un RCI de -12,5 indicativo de haberse producido un cambio estadísticamente muy significativo (el valor de p asociado a un RCI de esta magnitud es inferior a 0,0001). Si bien, en este caso, no se ha alcanzado la puntuación de corte que indicaría la funcionalidad ($PC < 8$), sí se ha logrado reducir dos niveles de clasificación sintomática, lo que indicaría una clara mejoría de la intensidad y frecuencia de síntomas depresivos que además no se han producido por azar. Es decir, nos encontramos ante un resultado clínicamente relevante y estadísticamente significativo.

Discusión

Del total de escalas y subescalas consideradas para valorar el cambio experimentado por la paciente tras ser sometida a una TEC, un total de 38 aspectos, seis han resultado ser estadística y clínicamente relevantes, siendo tres de ellas claramente específicas de la sintomatología que promueve la intervención (el HAM-D, el DBI y la subescala de depresión del NEO PI-R).

No hubo cambios en los 5 factores del NEO-PI-R, un hallazgo esperado dados los supuestos subyacentes en cuanto a su naturaleza como inventario general de los rasgos de personalidad y estabilidad. Las medidas de rasgo son sensibles a los cambios a largo plazo. En la misma línea, no se identificaron cambios en las 18 facetas del NEO-PI-R, pero es importante mencionar que algunos cambios observados fueron estadísticamente significativos y otros no. Por ejemplo, en la faceta de ansiedad social, la paciente pasó de una puntuación de 68 antes de TEC a 60 después de la terapia, un cambio que no es significativo. Es decir, de una puntuación disfuncional ($68 > PC = 65$) pasa a localizarse dentro de la funcionalidad y esto sería interpretado como una mejoría. No obstante, esta mejoría podría deberse al azar ya que la magnitud de la diferencia del cambio está dentro del error de medida esperado. Por tanto, contar con el RCI nos permite cuestionar este cambio y ser prudentes en su interpretación.

Por otro lado, también encontramos resultados que pueden ser significativos desde un punto de vista clínico, aunque los criterios basados estrictamente en los valores normativos nos podrían llevar a descartarlos. Es el caso de los resultados en la faceta Hostilidad (diferencia del cambio = 13, $RCI = 2,37$) o gregarismo (diferencia del cambio = 14, $RCI = 2,55$). En ambos casos, los valores de RCI indicaban un cambio estadísticamente significativo con una $p < 0,008$, es decir, en 8 de cada 1.000 ocasiones nos equivocaríamos al considerar que la magnitud del cambio observado no ha sido debido al azar. No obstante, en ambos casos las puntuaciones en el pretest y en el postest se sitúan dentro del área de funcionalidad y no resultarían interpretadas. Ahora bien, un cambio de 13 puntos de incremento en hostilidad y una reducción de 14 puntos en gregarismo podrían ser susceptibles de interpretación clínica y tendidas en cuenta como cambios atribuibles al efecto de la intervención.

Por último, existe otra puntuación que puede tener implicaciones clínicas: la vulnerabilidad. Una puntuación de 73 antes de la TEC pasa a 66 tras la intervención en esta paciente. En ambos casos las puntuaciones se hallan

dentro de los criterios de disfuncionalidad y el valor de RCI (-1,28) no nos permite concluir que el cambio observado sea estadísticamente significativo, lo que nos lleva a la conclusión de no cambio. Pero consideremos que la puntuación pretest fuera de 76 y, por tanto, se alcanzase un cambio de -10 puntos. El cálculo del RCI bajo esta nueva condición sería de -2,01, lo que llevaría a la conclusión de que se ha producido un cambio significativo, aunque no pueda concluirse que clínicamente sea relevante al no situarse la puntuación postest dentro del área de funcionalidad. No obstante, este resultado sería interesante para el clínico ya que le informaría que el cambio obtenido está en la dirección de los objetivos de reducción sintomática que se persiguen y, por tanto, serían un criterio más que permitiría mantener la prescripción, por ejemplo.

Conclusión

Los datos presentados en este caso clínico muestran que el RCI puede ser de interés para los profesionales de la salud a la hora de guiar la toma de decisiones de acuerdo con los resultados observados en casos individuales. Se trata de una herramienta sólida que va más allá de los juicios subjetivos y ofrece datos que sirven para respaldar las decisiones, controlar el efecto del azar en los resultados esperados y los cambios. Además, el sistema de juicios basados en los dos niveles de decisión, clínico y estadístico, permite considerar otros tipos de resultados que clínicamente pueden tener implicaciones sustantivas. Específicamente, en un caso de TEC podría ser una herramienta útil para evaluar la eficacia de una estrategia de continuidad y mantenimiento, como la recomendada por Martínez-Amorós et al⁵.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychology*. 1991;59:12-9.
2. Fernández-Ballesteros R. Evaluación psicológica. Conceptos, métodos y estudio de casos. Madrid: Pirámide; 2011.
3. Sederer LI, Dickey B. Outcomes assessment in clinical practice. Baltimore: Williams & Wilkins; 1995.
4. Iraurgi I. Evaluación de resultados clínicos III: Estimación del cambio clínicamente significativo utilizando los índices de cambio fiable. *Norte de Salud Mental*. 2010;36:105-22.
5. Martínez-Amorós E, Cardoner N, Gálvez V, Urretavizcaya M. Eficacia y patrón de uso de la terapia electroconvulsiva de continuación y mantenimiento en el trastorno depresivo mayor. *Rev Psiquiatr Salud Ment*. 2012;5:241-53.
6. López-Pina JA, Sánchez-Meca J, Rosa-Alcázar AI. The Hamilton Rating Scale for Depression: A meta-analytic reliability generalization study. *Int J Clin Health Psychology*. 2009;9:143-59.
7. Sanz J, Vázquez C. Fiabilidad, validez y datos normativos del inventario de depresión de Beck. *Psicothema*. 1998;10:303-18.
8. Costa PT, McCrae RR. NEO PI-R, Inventario de Personalidad NEO revisado. Manual Profesional. Madrid: TEA Ediciones; 2008.