

Acuerdo entre evaluadores. Metodología para unificar criterios diagnósticos

M. Martínez-González. *Departamento de Fisioterapia. Universidad de Murcia*

A. Gómez Conesa. *Departamento de Fisioterapia. Universidad de Murcia*

M.^a D. Hidalgo Montesinos. *Departamento de Psicología Básica y Metodología. Universidad de Murcia*

RESUMEN

El presente estudio valora un método para mejorar el acuerdo entre los fisioterapeutas en la realización e interpretación de pruebas diagnósticas.

En situaciones simuladas y durante tres sesiones, dos fisioterapeutas inexpertos fueron entrenados por expertos en procedimientos de evaluación de ejecuciones motoras. Posteriormente, en situación real y de forma simultánea e independiente, los inexpertos observaron, evaluaron y registraron la ejecución de los ítems de cuatro pruebas motoras (coordinación, flexibilidad y dos pruebas de equilibrio) efectuadas por 11 individuos.

El acuerdo entre los inexpertos en la evaluación efectuada en situación real se analizó mediante coeficientes de correlación. Las puntuaciones resultantes (coordinación, $r = 1$; flexibilidad, $r = 0,975$ y dos pruebas de equilibrio, $r = 0,921$ y $r = 0,832$) muestran que el grado de acuerdo entre los evaluadores ha sido elevado en todas las pruebas realizadas, por lo que esta intervención ha sido efectiva.

Palabras clave: Diagnóstico, evaluación, Fisioterapia, interobservadores.

ABSTRACT

The present study tests a method to improve the ratio of agreement between physiotherapists in the execution and interpretation of diagnostic tests.

During three simulated sessions, two inexperienced physiotherapists were trained by experts in procedures of the evaluation of motor-skills. Later, under natural conditions the inexperienced physiotherapists simultaneously and independently observed, evaluated and registered the execution of four motor-skills tests (co-ordination, flexibility and two balance tests), undertaken by 11 subjects.

The ratio of agreement between the inexperienced physiotherapists in the evaluation made under the natural conditions was analyzed using correlation coefficients. The resulting scores (co-ordination, $r = 1$, flexibility, $r = 0,975$, two balance tests, $r = 0,921$ and $r = 0,832$) show that the ratio of agreement between the appraisers was high in all the tests carried out. We consider that our intervention has been effective for this reason.

Key words: Diagnosis, evaluation, physiotherapy, inter-observer.

INTRODUCCIÓN

El avance de la Fisioterapia ha ocasionado que en la actualidad, tanto en los estudios experimentales como en la actividad clínica, los fisioterapeutas utilicen diferentes escalas y métodos de valoración o diagnóstico para conocer el estado de los pacientes. La cuantificación de aspectos biológicos, sociales y psicológicos, no sólo antes y después de los tratamientos experimentales y clínicos, sino también durante la aplicación de éstos, permite detectar los cambios ocurridos como consecuencia de la intervención fisioterapéutica.

En su actividad, el fisioterapeuta lleva a cabo tanto una valoración inicial, para el tratamiento, como una valoración posterior, del tratamiento. La valoración de diagnóstico para el tratamiento fisioterapéutico se centra en la recogida de información, con la selección y aplicación de los procedimientos (escalas, pruebas, protocolos, etc.) más relevantes para medir el problema de salud, así como en el establecimiento del plan de tratamiento que se debe seguir [1]. Este plan debe ser siempre acordado con el enfermo/usuario tras proporcionarle una información adecuada [2,3].

Por otro lado, mediante la valoración del tratamiento, el fisioterapeuta formula las hipótesis contrastables, planifica el tratamiento eligiendo el diseño más adecuado, aplica las técnicas seleccionadas valorando su eficacia de manera continuada y, finalmente, valora el éxito de la actuación [1].

A lo largo del proceso de actuación profesional, tanto en la selección de las pruebas de exploración y evaluación como en el desarrollo del plan de tratamiento, es necesario seguir una metodología rigurosa que permita obtener resultados y conclusiones comparables, y gracias a su posterior difusión, facilite el trabajo de otros fisioterapeutas. El rigor

metodológico, que debe primar en cualquier actuación fisioterapéutica, garantiza la actividad profesional y promueve el avance científico de la Fisioterapia.

Lamentablemente, en ocasiones se observa que, aunque diferentes estudios han utilizado la misma prueba de evaluación, cuando se valoran aspectos relacionados con la observación directa y la interpretación de ejecuciones motrices (como normalidad, deficiencia o alteraciones), los resultados no pueden ser comparables debido a problemas de fiabilidad del evaluador o a su inexperiencia como tal [4]. La inexperiencia puede convertirse en una fuente de error del examinador [5], ya que aunque éste la introduce de forma involuntaria en el procedimiento de aplicación de una determinada prueba, resta rigor a los resultados de ésta. Un ejemplo de escasez en el rigor metodológico lo constituye la variabilidad existente entre dos o más observadores al cuantificar la misma conducta o gesto motor durante la evaluación o diagnóstico fisioterapéutico, ya que como señalan Koch y cols. [6], el acuerdo entre evaluadores permite determinar la fiabilidad.

Asimismo, la ausencia o escasez en la unificación de criterios previos (como la observación, recogida de información, registro de datos, etc.) para la interpretación y valoración de conductas motoras, dificulta el contraste y generalización de los resultados de la prueba.

OBJETIVOS

Con la finalidad de mejorar la evaluación diagnóstica en Fisioterapia, este estudio se ha dirigido a aumentar el acuerdo entre los fisioterapeutas, tanto en la observación de gestos y actividades motrices, como en la cuantificación y registro de los resultados de dichas ejecuciones.

El estudio se desarrolló en dos fases diferentes:

Fase de adquisición de habilidades de observación (en situaciones simuladas)

Incluyó la programación de actividades de entrenamiento de evaluadores.

Objetivos

- Unificar criterios de valoración.
- Desarrollar la capacidad de observación.
- Favorecer el procesamiento cognoscitivo de la evaluación.
- Capacitar de recursos de atención y comunicación.
- Familiarizar a los individuos con las pruebas.

Fase de ejecución (en situaciones reales)

Aplicación de las pruebas y establecimiento del diagnóstico de los individuos evaluados según los resultados obtenidos.

Objetivos

- Entrenar en la observación de conductas motoras.
- Integrar aprendizajes previos.
- Aplicar las pruebas en contextos reales.

MATERIAL Y MÉTODOS

Individuos

Tras obtener el consentimiento informado de los individuos implicados en el estudio (participantes, dirección y responsables del

centro de mayores), dos fisioterapeutas inexpertos en procedimientos diagnósticos fueron entrenados por dos fisioterapeutas expertos y especializados en técnicas de evaluación.

Asimismo, en la fase de ejecución, 11 individuos (5 hombres y 6 mujeres) mayores de 70 años, con deambulación y semejante estado de salud, institucionalizados en una residencia de ancianos de la ciudad de Murcia, participaron en la evaluación.

Material

En ambas fases se emplearon cuatro pruebas de evaluación cuyos ítems miden actividades o gestos psicomotores; en concreto:

- Prueba de flexibilidad.
- Coordinación óculo-manual [7].
- Prueba de equilibrio estático y dinámico.
- Prueba de equilibrio de Tinetti [8].

La prueba para valorar la flexibilidad contiene 5 ítems, la de la coordinación óculo-manual 1, la prueba de Tinetti valora el equilibrio mediante 16 ítems, y la segunda prueba de valoración del equilibrio estático y dinámico contiene 8 ítems. Excepto en la prueba de Tinetti, que valora cada uno de los ítems con una puntuación de 0 o 1, el resto de las pruebas presenta ítems que se valoran de 0 a 3 puntos en función de la bondad en la ejecución.

Procedimiento

Fase de adquisición de habilidades

Se desarrolló en condiciones simuladas durante 3 sesiones, con una duración variable entre 2 y 3 horas por sesión y un intervalo de 3 a 4 días entre sesiones.

Sesión 1 (2 horas): Los expertos entregaron a cada evaluador inexperto una copia con las cuatro pruebas de evaluación que se emplearían en el experimento. Asimismo, se les informó de los objetivos del estudio y del procedimiento y metodología que se debía seguir.

Los inexpertos efectuaron una lectura comprensiva de las pruebas y, posteriormente, plantearon dudas que se resolvieron en una discusión en grupo. Seguidamente, se concertó la segunda sesión para tres días después y se programaron tareas para casa: estudio y análisis de cada uno de los ítems de las cuatro pruebas de evaluación.

Sesión 2 (3 horas): La sesión se inició con el planteamiento y resolución de dudas derivadas del trabajo individual efectuado en casa. A continuación, los expertos explicaron aspectos relativos a la evaluación en situaciones reales, como por ejemplo: problemas de comunicación que pueden presentarse, importancia de la empatía, estrategias para mejorar las habilidades sociales y de comunicación con los evaluados, etc. Se revisaron estrategias para guiar la evaluación con disciplina y evitar interrupciones no deseadas durante su aplicación, ya que además de hacer perder la concentración en la prueba, ocasionan un aumento no deseado en el tiempo de aplicación. Estos aspectos se discutieron en grupo.

La segunda parte de la sesión se dedicó a la práctica con *role-playing*, en la que uno de los expertos simuló ser un individuo evaluado, mientras los dos evaluadores noveles y el otro experto lo valoraban. Posteriormente, se contrastaron las evaluaciones, tanto entre los tres evaluadores como con el simulador, respecto a lo que éste había representado. Los expertos analizaron razonadamente las diferencias aparecidas en esta práctica evaluativa, se discutió y se unificaron de nuevo los criterios con el acuerdo de los cuatro componentes (expertos e inexpertos).

Se concertó para cuatro días más tarde la siguiente sesión y se programaron tareas para casa: análisis y reflexión del trabajo realizado durante la segunda sesión y repaso de la prueba.

Sesión 3 (3 horas): Se inició con la discusión en grupo de las dudas surgidas, para proseguir con la realización de evaluaciones simuladas. De forma individual, tanto expertos como inexpertos, alternativamente, simulaban las ejecuciones. Al final de cada evaluación, las discrepancias se contabilizaron y se compararon los resultados de la valoración, para continuar con la discusión en grupo siguiendo el modelo iniciado en la sesión anterior. Este procedimiento se repitió hasta alcanzar la unificación de criterios entre los cuatro evaluadores y en todos los ítems de las cuatro pruebas. Posteriormente, se repitieron algunas pruebas y se comprobó que el acuerdo se mantenía.

Los evaluadores fueron citados para el día siguiente con la finalidad de efectuar la evaluación en situación real. La tabla 1 presenta un esquema con los contenidos de la primera fase del estudio.

Fase de ejecución en situación real

Se llevó a cabo mediante la evaluación de 11 individuos mayores de 70 años institucionalizados, efectuada por los dos fisioterapeutas inexpertos.

Ambos fisioterapeutas observaron, evaluaron y registraron de forma simultánea e independiente, la ejecución de cada uno de los ítems que componen las pruebas de flexibilidad, equilibrio y coordinación óculo-manual; es decir, las cuatro pruebas de valoración para las que habían sido entrenados.

Durante el desarrollo de la evaluación, uno de los expertos se encargó del control del

TABLA 1. Fase de adquisición de habilidades. Contenido de las sesiones

<i>Sesión 1</i>	(2 horas) Presentación: objetivos y metodología del estudio Entrega de las pruebas de evaluación Lectura y análisis de los ítems de cada una de las pruebas Discusión en grupo y resolución de dudas Tareas para casa: estudio y análisis de los ítems de las cuatro pruebas
<i>Sesión 2</i>	(3 horas) Resolución de dudas Explicación y discusión de aspectos relativos a la evaluación en situaciones reales <i>Role-playing</i> . Con evaluación de la simulación, discusión y unificación de criterios Tareas para casa: análisis y reflexión sobre la sesión y repaso de las pruebas
<i>Sesión 3</i>	(3 horas) Resolución de dudas Actividades de <i>Role-playing</i> . Con evaluación de las simulaciones, discusión y unificación de criterios

proceso: atención a los evaluados, apoyo para los evaluadores, a la vez que comprobaba la independencia entre los evaluadores en el registro de datos.

RESULTADOS

Análisis estadístico

Para comprobar el acuerdo entre los observadores y, por lo tanto, verificar la unificación de criterios en la evaluación, en cada una de las cuatro pruebas se analizaron los resultados de ambos evaluadores por medio del coeficiente de correlación de Pearson (r), prueba que se obtiene de dividir la covarianza por el producto de las desviaciones típicas y cuyos valores sólo pueden oscilar entre -1 y $+1$. El 1 indica la máxima relación entre

las variables y 0 , la ausencia de relación. Asimismo, los signos $+$ (positivo) y $-$ (negativo), indican la dirección de la relación: directa $(+)$ e inversa $(-)$.

El análisis de los datos se efectuó con el paquete estadístico Systat (Versión 8.0) [9,10].

Resultados de los análisis

Los coeficientes de correlación obtenidos para cada uno de las cuatro pruebas fueron los siguientes:

- Prueba de coordinación óculomanual, $r = 1$.
- Prueba de flexibilidad, $r = 0,975$.
- Prueba de equilibrio estático y dinámico, $r = 0,921$.
- Prueba de Tinetti, $r = 0,832$.

La figura 1 muestra un gráfico con los resultados.

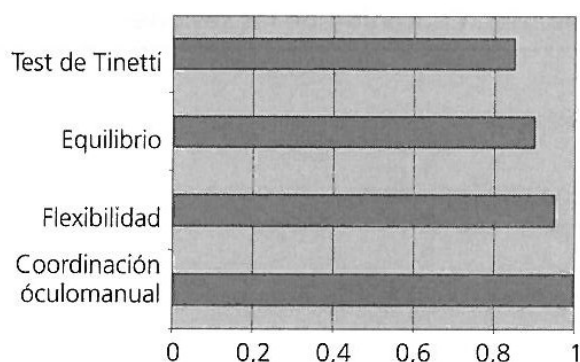


FIG. 1. Resultados de los coeficientes de correlación de Pearson. Muestra el grado de acuerdo entre los evaluadores en cada una de las pruebas efectuadas.

DISCUSIÓN

Como se desprende de los datos y la representación gráfica de la figura 1, los coeficientes de correlación resultaron elevados para cada una de las pruebas de evaluación utilizadas en este estudio.

En las pruebas de valoración de equilibrio estático y dinámico, de flexibilidad y prueba de Tinetti, se obtuvieron coeficientes de correlación elevados, lo cual indica un importante acuerdo entre los observadores a la hora de medir las ejecuciones motoras.

Respecto a la valoración de la prueba de coordinación óculomanual, se obtuvo un efecto «techo» en el acuerdo interobservadores.

Por otro lado, al examinar dos individuos simultáneamente las ejecuciones, éstas no estaban afectadas por la aplicación de una prueba repetida, lo que ha permitido obtener dos medidas de cada individuo para cada una de las ejecuciones de las cuatro pruebas.

CONCLUSIONES

Los índices de correlación resultantes del análisis comparativo de los datos recogidos

por ambos evaluadores muestran un importante grado de acuerdo interobservadores al valorar los ítems de las ejecuciones motrices de las pruebas de equilibrio, flexibilidad y coordinación óculomanual empleadas.

Los resultados demuestran que la intervención ha sido efectiva para lograr el objetivo principal del experimento: unificar los criterios de valoración diagnóstica entre los dos evaluadores. En la misma línea, otros autores [11] también han encontrado exactitud interevaluadora al analizar las medidas obtenidas por distintos examinadores entrenados mediante procedimientos de instrucción, prácticas y evaluación práctica.

En esta investigación, los fisioterapeutas entrenados fueron capaces de llegar a un acuerdo en los resultados de los exámenes y obtener un excelente grado de exactitud. En opinión de los autores, el entrenamiento de los fisioterapeutas en procedimientos de observación e interpretación de conductas motrices y en la valoración de las ejecuciones, permite mejorar el acuerdo a la hora de efectuar una evaluación funcional y diagnosticar un problema.

Finalmente, estos hallazgos muestran el efecto positivo de las sesiones de entrenamiento realizadas, ya que proporcionan educación y experiencia clínica y permiten mejorar la fiabilidad de los fisioterapeutas tanto en la observación directa como en la interpretación de las ejecuciones.

BIBLIOGRAFÍA

1. Gómez, A., Méndez, F.X., Olivares, J.: Proceso de actuación fisioterapéutica. *Revista Iberoamericana de Fisioterapia y Kinesiología*; 2(1): 31-43, 1999.
2. Carles, R., Gómez, A., Pérez-Cárceles, M.D., Osuna, E.: Consentimiento informado en

- Fisioterapia. *Fisioterapia y Calidad de Vida*; 1(1): 40-42, 1998.
3. Gómez, A., Carles, R., Pérez-Cárceles, M.D., Osuna, E.: Principios éticos y deontológicos en fisioterapia. Su comparación con otras profesiones sanitarias. *Cuestiones de Fisioterapia*; 11: 1-7, 1999.
 4. Brooks, D., Thomas, J.: Interrater reliability of auscultation of breath sounds among physical therapist. *Physical Therapy*; 75 (12): 1.082-1.088, 1995.
 5. Chesworth, B.M., MacDermid, J.C., Roth, J.H., Patterson, S.D.: Movement diagram and «end-feel» reliability when measuring passive lateral rotation of the shoulder in patients with shoulder pathology. *Physical Therapy*; 78 (6): 593-595, 1998.
 6. Koch, M., Gottschalk, M., Baker, D.I., Palumbo, S., Tinetti, M.E.: An impairment and disability assessment and treatment protocol for community-living elderly persons. *Physical Therapy*; 74 (4): 286-298, 1994.
 7. Durante, P., Hernando, A.L.: Actividad física con los muy viejos. Metodología y resultados. *Revista Española de Geriatria y Gerontología*; 29 (4): 197-203, 1994.
 8. Tinetti, M.E., Speechley, M., Ginter, S.F.: Risk factors for falls among elderly persons living in the community. *The New England Journal of Medicine*; 319: 1.701-1.707, 1988.
 9. Ato, M., López, J.J.: *Fundamentos de estadística con SYSTAT*. Madrid: RA-MA, 1994.
 10. Martínez, R.: *Psicometría*. Madrid: Síntesis, 1995.
 11. Van Dillen, L.R., Sahrmann, S.A., Norton, B.J., Caldwell, Fleming D.A., McDonnell, M.K., Woolsey, N.B.: Reliability of physical examination items used for classification of patients with low back pain. *Physical Therapy*; 78(9): 979-988, 1998.