

Estudio psicométrico de la prueba MIR de 2022

José CURBELO, Alberto GARCÍA-GUERRERO, Fernando SÁNCHEZ-LASHERAS, Carmen PEÑALVER-SAN CRISTÓBAL, Tomás VILLACAMPA, Paula JIMÉNEZ-FONSECA

Introducción. En España, para ejercer como especialista de cualquier especialidad médica es necesario tener la titulación correspondiente. Para poder acceder a la formación como especialista es obligatorio superar la prueba MIR. En este artículo se ha realizado un análisis psicométrico de la prueba MIR de 2022, celebrada en enero de 2023, utilizando tanto la teoría clásica de los test como la teoría de respuesta al ítem.

Material y métodos. Se hizo uso de un subconjunto formado por 3.229 médicos que se presentaron a la prueba MIR de 2022 y que introdujeron su plantilla de respuestas en una aplicación informática puesta a su disposición para tal efecto. Los datos se analizaron mediante la teoría clásica de los test y la teoría de la respuesta al ítem. Concretamente, se calcularon el alfa de Cronbach, la fórmula número 21 de Kuder y Richardson, el índice de dificultad, el índice de dificultad con corrección de los efectos del azar, el índice de discriminación, el índice de correlación biserial puntual, así como los índices de dificultad y discriminación según la teoría de respuesta al ítem.

Resultados. La prueba MIR de 2022, al igual que las de años anteriores, presenta una alta fiabilidad. Se analizan tanto los resultados globales de la prueba como detallados por asignaturas y bloques de asignaturas.

Conclusión. Los resultados obtenidos en el presente estudio permiten afirmar que el examen MIR es un examen objetivo, de dificultad y discriminación adecuadas, así como relativamente estable de una convocatoria a otra.

Palabras clave. Estadísticas. Estudiantes de medicina. Medicina. Mediciones educativas. Psicometría. Reproducibilidad de resultados.

Psychometric analysis of the 2022 MIR test

Introduction. In Spain, to practice as a specialist in any medical specialty, it is necessary to have the corresponding qualification. In order to access training as a specialist, it is mandatory to pass the MIR test. In this article, a psychometric analysis of the 2022 MIR test, held in January 2023, has been carried out using both classical test theory and item response theory.

Materials and methods. A subset made up of 3,229 doctors who took the MIR test in 2022 and who entered their response template in a computer application made available for this purpose. Data were analyzed using classical test theory and item response theory. Specifically, Cronbach's alpha, Kuder and Richardson's formula number 21, the difficulty index, the difficulty index corrected for chance effects, the discrimination index, the biserial-point correlation index, as well as the difficulty and discrimination indices according to item response theory were calculated.

Results. The 2022 MIR test, like those of previous years, shows high reliability. Both the overall results of the test and those detailed by subjects and subject blocks are analyzed.

Conclusion. The results obtained in this study allow us to affirm that the MIR exam is an objective exam, of adequate difficulty and discrimination, as well as relatively stable from one call to another.

Key words. Educational measurements. Medical students. Medicine. Psychometrics. Reproducibility of results. Statistics.

Introducción

En España, para el acceso a una plaza de formación como médico especialista es necesario que el aspirante supere la prueba MIR. Dado que el número de plazas es limitado, esta prueba sirve para orde-

nar a los opositores con el fin de escoger entre las distintas especialidades.

La prueba de acceso a la especialización, conocida como prueba MIR, es convocada anualmente desde 1978 por los Ministerios de Sanidad, y Educación y Formación Profesional. La normativa aplica-

Cursos Intensivos MIR Asturias (J. Curbelo, A. García-Guerrero, C. Peñalver-San Cristóbal). Departamento de Matemáticas. Facultad de Ciencias. Universidad de Oviedo (F. Sánchez-Lasheras). Servicio de Oncología Médica. Hospital Universitario Central de Asturias. Oviedo (P. Jiménez-Fonseca). Facultad de Medicina. Universidad Francisco de Vitoria. Pozuelo de Alarcón (J. Curbelo). Clínica Oftalmológica Villacampa. Avilés, España (T. Villacampa).

Correspondencia:

Dr. Fernando Sánchez Lasheras. Departamento de Matemáticas. Facultad de Ciencias. Universidad de Oviedo. C/ Federico García Lorca, 18. E-33007 Oviedo.

E-mail:

sanchezfernando@uniovi.es

Recibido:

21.09.23.

Aceptado:

22.09.23.

Conflicto de intereses:

No declarado.

Competing interests:

None declared.

Cómo citar este artículo:

Curbelo J, García-Guerrero A, Sánchez-Lasheras F, Peñalver-San Cristóbal C, Villacampa T, Jiménez-Fonseca P. Estudio psicométrico de la prueba MIR de 2022. FEM 2023; 26: 219-26. doi: 10.33588/fem.2605.1303.

© 2023 FEM



Artículo open access bajo la licencia CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

ISSN: 2014-9832

ISSN (ed. digital): 2014-9840

ble a la prueba se recoge en el Real Decreto 589/2022 [1]. Este real decreto desarrolla lo dispuesto en el artículo 25 de la Ley 44/2003, de 21 de noviembre, de Ordenación de las Profesiones Sanitarias [2].

Según algunos autores, la aplicación de pruebas de tipo test para la selección de médicos con el fin de acceder a una especialidad [3-5] no es una metodología suficientemente completa para garantizar la calidad del proceso de selección. Efectivamente, aunque sería conveniente que una prueba de esta índole se acompañase de otros criterios de evaluación que tuvieran en cuenta la experiencia previa, la motivación, la empatía, la capacidad para trabajar en equipo y otras variables relacionadas con el perfil profesional del candidato, no es menos cierto que las pruebas de tipo test se emplean también en otros países para el acceso tanto al ejercicio de la profesión médica como de la especialización [6,7]. Además, también cabe destacar la existencia de trabajos previos [8-13] en los que se argumenta que, desde el punto de vista psicométrico y del contenido de las preguntas, la prueba MIR presenta un buen rendimiento.

En la convocatoria de 2022 (celebrada el 21 de enero de 2023), el número total de preguntas fue de 200, junto con otras 10 preguntas de reserva. Una vez realizada la prueba MIR, se calcula la puntuación total que determinará el orden de prelación a la hora de escoger una especialidad. Para el cálculo de la puntuación total se pondera el resultado de la prueba MIR (90%) con el baremo académico obtenido en el grado en Medicina (10%), siguiendo las bases publicadas en la Orden SND/840/2022, de 26 de agosto, por la que se aprueba la oferta de plazas y la convocatoria de pruebas selectivas de 2022 para el acceso en el año 2023 a plazas de formación sanitaria especializada. Además, se establece una nota de corte que se corresponde con el 25% del promedio de la puntuación del 10% de los mejores exámenes de la convocatoria. Así, los candidatos que obtienen una puntuación total igual o superior a la nota de corte se ordenan de mayor a menor en función de su puntuación total, y quedan excluidos del proceso de selección los candidatos con una puntuación total inferior a la nota de corte.

En esta prueba, los candidatos podían optar a un total de 8.550 plazas de formación especializada, lo que suponía un incremento de 545 plazas con respecto a la convocatoria anterior. A esta prueba se presentaron 11.578 personas, de las que 10.793 superaron la nota de corte. Del total de las plazas ofertadas, 599 estaban reservadas a personas con discapacidad. La especialidad con el mayor número de plazas fue medicina familiar y comunitaria, con un total de 2.455.

El objetivo de este trabajo es analizar desde el punto de vista psicométrico los resultados de las pruebas MIR de 2022. Para tal fin, la metodología de trabajo se basará en el análisis de los datos recopilados de las respuestas de un subconjunto de los opositores que se examinaron en esta prueba y de los que se dispone de sus respuestas al test pregunta a pregunta. Una vez realizado el análisis, se presentarán los resultados obtenidos y se discutirán comparándolos con los resultados de convocatorias previas ya analizados en otros estudios [10-14]. Se concluirá con una valoración global de la prueba.

Material y métodos

Bases de datos

Se ha hecho uso de la información disponible a través de una aplicación *ad hoc* desarrollada por el Curso Intensivo MIR Asturias. Este sistema ya se había empleado en estudios anteriores [10-15]. Así, se dispone de las respuestas de 3.229 sujetos que se presentaron a la prueba MIR de 2022. Seguidamente se definen las métricas utilizadas en este trabajo y que ya se habían empleado en investigaciones anteriores [14].

Alfa de Cronbach y fórmula 21 de Kuder y Richardson

La fiabilidad del test se estudió por medio de las fórmulas del alfa de Cronbach y la fórmula número 21 de Kuder y Richardson [16].

La fórmula del alfa de Cronbach viene definida por la ecuación:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_x^2} \right) \quad (1)$$

Donde:

n representa el número de ítems del test.

σ_j^2 es la varianza del ítem j -ésimo

σ_x^2 es el valor de la varianza de las puntuaciones obtenidas por todos los opositores en el test.

La fórmula número 21 de Kuder y Richardson se expresa como:

$$KR_{21} = \frac{n}{n-1} \left(1 - \frac{n \cdot p \cdot (1-p)}{\sigma_x^2} \right) \quad (2)$$

Nótese que (2) n y σ_x^2 tienen el mismo significado que en (1) y que p representa la probabilidad promedio de acierto de todas las preguntas del test.

Índice de dificultad

El índice de dificultad es una métrica adecuada para determinar cuán difícil es una pregunta en un examen. Esto se logra mediante la evaluación de la eficiencia de los estudiantes a la hora de responder a la pregunta en cuestión. Este índice se define por medio de la fórmula [16,17]:

$$DF = \frac{F+D}{N_1+N_2} \quad (3)$$

En (3), F representa el número de personas que aciertan la pregunta de entre las que pertenecen al grupo fuerte, el cual se define como el que está formado por el 27% de los opositores que obtuvieron las puntuaciones más altas en la prueba [10], mientras que D representa el número de respuestas correctas entre los candidatos pertenecientes al grupo débil, siendo dicho grupo el formado por el 27% de los estudiantes que obtuvieron las peores puntuaciones en el test [10]. Esta fórmula también hace uso de las variables N_1 y N_2 que representan el número de estudiantes que respondieron a la pregunta en el grupo fuerte y débil, respectivamente. Valores altos del índice de dificultad significan que la pregunta ha sido respondida de forma correcta por un número elevado de opositores.

Las preguntas se clasifican atendiendo a su índice de dificultad en [10,16]:

- Fácil: $>0,8$.
- Aceptable: $>0,6-0,8$.
- Óptimo: $>0,5-0,6$.
- Aceptable: $>0,3-0,5$.
- Difícil: $0-0,3$.

Índice de dificultad con corrección de los efectos del azar

Otro indicador de la dificultad de una pregunta es el conocido como índice de dificultad con corrección de los efectos del azar [10,17]:

$$ID = \frac{A - \frac{E}{K-1}}{N} \quad (4)$$

En esta fórmula, A representa el número de médicos que responden a la pregunta correctamente, mientras que la suma de los que la fallan o la dejan

en blanco viene dada por E . Además, K representa el número de alternativas de respuesta, que en el caso de la convocatoria MIR objeto de análisis es 4, y N es el número total de opositores en la muestra.

El resultado obtenido de este índice se puede clasificar en las siguientes categorías [10,17]:

- Muy fácil: $>0,8$.
- Fácil: $>0,66-0,8$.
- Óptimo: $>0,33-0,66$.
- Difícil: $>0-0,33$.
- Muy difícil: $-0,33-0$.

Índice de discriminación

Se define el índice de discriminación como [17]:

$$DS = 2 \cdot \frac{F-D}{N_1+N_2} \quad (5)$$

En esta fórmula, F , D , N_1 y N_2 tienen el mismo significado que en la fórmula (1). El índice de discriminación se clasifica en las siguientes categorías [10,16,18,19]:

- Excelente: $>0,34$.
- Bueno: $>0,24-0,34$.
- Revisable: $>0,14-0,24$.
- Malo: $0-0,14$.
- Muy malo: <0 .

Índice de correlación biserial puntual

El índice de correlación biserial puntual ayuda a conocer la capacidad discriminativa de un test. Su fórmula es [16]:

$$\rho_{bis} = \frac{\mu_p - \mu_q}{\sigma_x} \sqrt{\frac{ID}{1-ID}} \quad (6)$$

La variable μ_p equivale a la puntuación media en el test de los sujetos que aciertan el ítem, mientras que μ_q se corresponde con la media de los que lo fallan. Además, la fórmula también utiliza la desviación típica de las puntuaciones totales de todos los opositores, representada por σ_x y el índice de dificultad con corrección de los efectos del azar definido anteriormente en la ecuación (4).

Los resultados obtenidos a través de este índice se clasifican en [10,12]:

- Excelente: $>0,39$.
- Buena: $>0,3-0,39$.
- Regular: $>0,2-0,3$.
- Pobre: $0-0,2$.
- Pésima: <0 .

Índices de dificultad y discriminación según el modelo de dos parámetros de la teoría de respuesta al ítem

Si bien todos los índices definidos con anterioridad corresponden al marco de la teoría clásica de los test, en el presente apartado se definen dos índices, uno de dificultad y otro de discriminación, desarrollados en el contexto de la teoría de respuesta al ítem [11]. Trabajos previos [11,13] han puesto de manifiesto que el modelo que mejor refleja el comportamiento de los opositores que se presentan a la prueba MIR es el modelo de dos parámetros [19]. Dicho modelo expresa la probabilidad que cada opositor tiene de acertar la pregunta i -ésima por medio de la fórmula [20]:

$$P(\mu_j | a_j, b_j) = \frac{\exp[-1,7 \cdot a_j \cdot (\theta_i - b_j)]}{1 + \exp[-1,7 \cdot a_j \cdot (\theta_i - b_j)]} \quad (7)$$

En esta ecuación, θ_i se refiere al nivel de conocimiento del opositor i -ésimo, mientras que a_j y b_j representan, respectivamente, los valores de discriminación y dificultad de la pregunta i -ésima. A la vista de esta fórmula, es posible afirmar que la probabilidad que tiene un individuo de acertar una pregunta viene dada tanto por el nivel de conocimientos que éste tiene como por las características de la pregunta, entendiendo como tales su dificultad y discriminación.

Resultados

El subconjunto de la población total de examinados analizado está formado por un total de 3.229 individuos con un baremo promedio de 7,67 y una mediana de 7,65 sobre un máximo de 10 puntos. Si estos valores de baremo se comparan con los de años anteriores, se observa que tanto el valor promedio como la mediana son muy similares a los de los subconjuntos estudiados en las convocatorias de 2020 y 2021 [14].

Se obtuvo un valor según la fórmula 21 de Kuder y Richardson de 0,9639 y un alfa de Cronbach de 0,9535. Por tanto, esta prueba presenta una alta fiabilidad. Nótese que en esta prueba se anulaban cuatro preguntas y que en la misma había preguntas correspondientes a 34 asignaturas.

De las 206 preguntas no anuladas en la prueba, las asignaturas con un mayor número de preguntas fueron Endocrinología, con 14 preguntas, seguida de Cardiología, Neurología y Preventiva, con 12

preguntas cada una de ellas, y Digestivo, Pediatría y Reumatología, con 10 preguntas cada una.

En lo relativo al reparto de preguntas por bloques de asignaturas, el bloque con mayor número de preguntas fue el de 'Otras especialidades', con el 47,57% de las preguntas, seguido del de 'Aparatos' con el 42,72% y, finalmente, el de 'Básicas' con el 9,71%. Estos resultados son similares a los de las pruebas de 2020 y 2021 [14].

Análisis de las preguntas por asignaturas

La tabla I muestra el número de preguntas por asignatura junto con los valores medios y desviaciones estándar de sus índices de dificultad, índice de dificultad con corrección de los efectos del azar, índice de discriminación, índice de correlación biserial puntual, índice de dificultad del modelo de dos parámetros de la teoría de respuesta al ítem e índice de discriminación del mismo modelo correspondiente a las pruebas MIR de 2022.

En lo que se refiere al índice de dificultad de todas las preguntas de la prueba de 2022, el valor promedio obtenido fue de 0,6791, con una desviación estándar de 0,185, valor que se diferencia en menos de dos centésimas del de las dos convocatorias anteriores [14]. Este resultado permite afirmar que el promedio de dificultad es 'aceptable', pues el valor de la dificultad promedio de las asignaturas de Cardiología, Digestivo, Endocrinología, Neurología, Reumatología, Pediatría y Preventiva está comprendido entre 0,6 y 0,8. Es decir, el índice de dificultad de la prueba es aceptable y no hay grandes diferencias entre los niveles de dificultad de las preguntas de esta prueba en las asignaturas con 10 o más preguntas y las propuestas en las pruebas de 2020 y 2021.

El índice de dificultad con corrección de los efectos del azar alcanzó en el examen MIR de 2022 un valor promedio de 0,5674, con una desviación estándar de 0,2614. Este valor es muy similar al que se obtuvo en las convocatorias de 2020 y 2021, que era 0,5644 y 0,5566, respectivamente [14]. Este resultado supone que, en promedio, la dificultad de las preguntas se colocaría en la categoría 'óptima'. Así, en el caso de la prueba MIR de 2022, de las siete asignaturas con 10 o más preguntas, en seis de ellas el promedio del índice de dificultad con corrección de los efectos del azar se clasifica en la categoría de óptimo, y sólo en una de ellas, Pediatría, en la de fácil. Si se comparan con los de las pruebas de 2020 y 2021 [14], estos resultados indican que, a pesar de algunas variaciones en los valores de dificultad de las preguntas, las pruebas, año tras año, mantienen una dificultad uniforme en sus preguntas.

Tabla I. Prueba MIR de 2022. Número de preguntas por asignatura junto con su valor medio y desviación estándar de sus índices de dificultad, índice de dificultad con corrección de los efectos del azar, índice de discriminación, índice de correlación biserial puntual, índice de dificultad del modelo de dos parámetros de la teoría de respuesta al ítem (TRI) e índice de discriminación del modelo de dos parámetros de la teoría de respuesta al ítem.

Bloque	Asignaturas	Número de preguntas	Índice de dificultad	índice de dificultad con corrección de los efectos del azar	Índice de discriminación	Correlación biserial puntual	Dificultad TRI	Discriminación según la TRI
Aparatos	Cardiología	12	0,6516 (0,1179)	0,5577 (0,1785)	0,3332 (0,1533)	0,3615 (0,1442)	-0,6474 (0,792)	0,9568 (0,5567)
	Digestivo	10	0,6655 (0,1807)	0,5449 (0,2488)	0,2924 (0,1017)	0,3303 (0,1069)	-0,5332 (1,326)	0,8361 (0,3469)
	Endocrinología	14	0,6114 (0,1696)	0,4743 (0,2463)	0,3775 (0,1111)	0,3664 (0,1075)	-0,0509 (1,466)	0,9706 (0,3865)
	Enfermedades infecciosas	8	0,7211 (0,1469)	0,6231 (0,1922)	0,3049 (0,1484)	0,3614 (0,1356)	-0,9239 (0,848)	1,0103 (0,6271)
	Hematología	5	0,6621 (0,2397)	0,5538 (0,3493)	0,3068 (0,0696)	0,3912 (0,1368)	-0,22 (1,67)	1,1466 (0,548)
	Nefrología	9	0,6571 (0,1371)	0,546 (0,2082)	0,3179 (0,1373)	0,3507 (0,1316)	-0,6033 (0,927)	0,9253 (0,4789)
	Neumología	8	0,6718 (0,1866)	0,5459 (0,2755)	0,3442 (0,1493)	0,3531 (0,1122)	-0,3259 (1,75)	0,9402 (0,348)
	Neurología	12	0,6347 (0,1468)	0,4991 (0,2236)	0,2963 (0,1559)	0,3212 (0,1384)	-0,4946 (1,1)	0,8359 (0,481)
Reumatología	10	0,7223 (0,1724)	0,6146 (0,2559)	0,2342 (0,1592)	0,3532 (0,1996)	-0,848 (1,372)	1,2027 (0,9604)	
Básicas	Anatomía patológica	3	0,8664 (0,0926)	0,7923 (0,1381)	0,1628 (0,1013)	0,3209 (0,083)	-2,0492 (0,815)	0,9053 (0,3598)
	Bioquímica	4	0,5442 (0,1812)	0,3424 (0,2156)	0,226 (0,1707)	0,2448 (0,1214)	1,3386 (2,693)	0,5796 (0,3003)
	Farmacología	5	0,5669 (0,1708)	0,4096 (0,2369)	0,3486 (0,1174)	0,3374 (0,131)	0,2523 (1,133)	0,8504 (0,4012)
	Genética	3	0,7789 (0,1704)	0,7303 (0,2464)	0,3776 (0,2335)	0,5155 (0,0191)	-1,0146 (0,849)	1,7416 (0,2649)
	Inmunología	5	0,7027 (0,0764)	0,5902 (0,1142)	0,4637 (0,1054)	0,4632 (0,0178)	-0,5383 (0,411)	1,2521 (0,0486)
Otras	Alergología	2	0,746 (0,1279)	0,5571 (0,0058)	0,3298 (0,0945)	0,3375 (0,0554)	-0,5055 (0,154)	0,8162 (0,137)
	Anestesiología	3	0,565 (0,1081)	0,3851 (0,1428)	0,1784 (0,1152)	0,1994 (0,0982)	0,2939 (1,07)	0,4313 (0,2446)
	Bioética	4	0,7673 (0,2187)	0,6847 (0,2957)	0,1075 (0,1041)	0,2848 (0,1335)	-1,3161 (1,858)	0,8144 (0,5825)
	Cirugía maxilofacial	2	0,9058 (0,0201)	0,89 (0,0272)	0,1566 (0,0141)	0,462 (0,0301)	-1,8588 (0,021)	1,545 (0,3112)
	Cirugía plástica	2	0,4519 (0,5176)	0,2621 (0,6923)	0,0825 (0,1341)	0,1485 (0,2114)	48,4651 (71,467)	0,3563 (0,4781)
	Cirugía vascular	2	0,5665 (0,0025)	0,3774 (0,0673)	0,3605 (0,0275)	0,3007 (0,0226)	0,1097 (0,473)	0,7007 (0,1204)
	Dermatología	3	0,5452 (0,1957)	0,3397 (0,2544)	0,1743 (0,0916)	0,1783 (0,1014)	1,4278 (3,216)	0,3927 (0,2179)
	Fisiología	4	0,4799 (0,2495)	0,2862 (0,3264)	0,2479 (0,3048)	0,2142 (0,2391)	-1,9598 (3,895)	0,5504 (0,6383)
	Geriatría	6	0,6199 (0,2142)	0,4988 (0,302)	0,3325 (0,1444)	0,3815 (0,1102)	-0,2888 (1,256)	1,132 (0,7089)
	Ginecología y obstetricia	9	0,6996 (0,2069)	0,6065 (0,3014)	0,2469 (0,0764)	0,366 (0,1349)	-0,6319 (1,442)	1,076 (0,6046)
Medicina legal	2	0,1973 (0,0211)	0,074 (0,031)	0,1184 (0,0562)	0,1211 (0,0487)	5,2566 (2,448)	0,3318 (0,1332)	
Oftalmología	3	0,7301 (0,1884)	0,6397 (0,2761)	0,2338 (0,1674)	0,3522 (0,1919)	-0,7282 (1,163)	0,9966 (0,6434)	

Tabla I. Prueba MIR de 2022. Número de preguntas por asignatura junto con su valor medio y desviación estándar de sus índices de dificultad, índice de dificultad con corrección de los efectos del azar, índice de discriminación, índice de correlación biserial puntual, índice de dificultad del modelo de dos parámetros de la teoría de respuesta al ítem (TRI) e índice de discriminación del modelo de dos parámetros de la teoría de respuesta al ítem (*cont.*).

Bloque	Asignaturas	Número de preguntas	Índice de dificultad	Índice de dificultad con corrección de los efectos del azar	Índice de discriminación	Correlación biserial puntual	Dificultad TRI	Discriminación según la TRI
Otras	Oncología	6	0,7765 (0,2118)	0,704 (0,3172)	0,2738 (0,0739)	0,44 (0,0873)	-1,0288 (1,233)	1,4095 (0,4351)
	Otorrino-laringología	4	0,8399 (0,0866)	0,7904 (0,136)	0,1997 (0,0891)	0,385 (0,1353)	-1,806 (0,391)	1,1702 (0,5828)
	Paliativos	2	0,6756 (0,0203)	0,5758 (0,0327)	0,3268 (0,0944)	0,3345 (0,0431)	-0,8546 (0,318)	0,7794 (0,1411)
	Pediatría	10	0,7753 (0,1466)	0,7239 (0,2057)	0,2816 (0,0996)	0,4203 (0,1175)	-1,1532 (0,904)	1,2771 (0,5483)
	Preventiva	12	0,7229 (0,2323)	0,622 (0,3141)	0,2778 (0,1661)	0,399 (0,1497)	-0,3271 (2,666)	1,2299 (0,6016)
	Psiquiatría	8	0,7328 (0,1273)	0,6571 (0,1738)	0,3506 (0,1262)	0,4191 (0,0701)	-0,9348 (0,654)	1,1717 (0,3803)
	Traumatología	9	0,6929 (0,1625)	0,5966 (0,2356)	0,2748 (0,1945)	0,3566 (0,1781)	-0,8867 (0,961)	1,0774 (0,7504)
	Urgencias	5	0,8059 (0,1358)	0,7366 (0,1871)	0,2229 (0,2263)	0,3639 (0,1201)	-1,8138 (1,171)	1,0342 (0,4565)
Total		206	0,6791 (0,185)	0,5674 (0,2614)	0,2884 (0,1475)	0,3554 (0,1375)	-0,0957 (7,122)	1,0063 (0,5502)

Al observar los valores del índice de discriminación, se aprecia que el valor promedio de la prueba de 2022 fue de 0,2884, con una desviación estándar de 0,1475; valor que se clasifica en la categoría de bueno, al igual que en las pruebas de las dos convocatorias anteriores [14].

El valor promedio del índice de correlación biserial puntual de las preguntas correspondientes a la prueba de 2022 fue de 0,3554, con una desviación estándar de 0,1375. Este valor, si bien se puede clasificar como bueno, al igual que en las dos convocatorias previas, es unas cinco centésimas superior al de ambas, mientras que la desviación estándar se mantiene en el mismo rango [14].

En lo que se refiere a la dificultad promedio de las preguntas de la prueba MIR de 2022 expresada según el modelo de dos parámetros de la teoría de respuesta al ítem, el valor obtenido fue de -0,0957, con una desviación estándar de 7,122, resultados a menos de una unidad de diferencia de los obtenidos en promedio en las pruebas de las dos convocatorias anteriores [14] y con una desviación estándar también similar.

Por último, en lo relativo a los valores de discriminación según la teoría de respuesta al ítem, el promedio obtenido para la prueba de 2022 fue de 1,0063, con una desviación estándar de 0,5502, resultado con menos de dos décimas de diferencia

con respecto a los promedios de 2020 y 2021 y una desviación estándar también muy similar a la de las dos convocatorias previas [14].

Análisis de las preguntas por bloques

Además del análisis de las preguntas por asignaturas, también se estudiaron por bloques. Concretamente, los bloques en los que se dividen las preguntas son: 'Aparatos', que consta de las asignaturas médicas clásicas y quirúrgicas correspondientes a cada aparato, 'Asignaturas básicas' y 'Otras asignaturas'. La tabla II muestra el número de preguntas en cada bloque, junto con la media y desviación estándar de los índices de dificultad, índice de dificultad con corrección de los efectos del azar, índice de discriminación, índice de correlación biserial puntual, e índice de dificultad y discriminación del modelo de dos parámetros de la teoría de respuesta al ítem.

El índice de dificultad de los tres bloques de preguntas es muy similar, con un valor promedio en la categoría de 'aceptable', que apenas se diferencia en tres centésimas entre bloques. En el caso del índice de dificultad con corrección de los efectos del azar, los promedios de los tres bloques se encuentran en la categoría 'óptima' y la máxima diferencia entre bloques se produce entre 'Aparatos' (el bloque más difícil) y 'Otras especialidades' (el bloque más fácil),

Tabla II. Prueba MIR de 2022. Número de preguntas por bloque de asignaturas junto con su valor medio y desviación estándar de sus índices de dificultad, índice de dificultad con corrección de los efectos del azar, índice de discriminación, índice de correlación biserial puntual, índice de dificultad del modelo de dos parámetros de la teoría de respuesta al ítem (TRI) e índice de discriminación del modelo de dos parámetros de la TRI.

Bloque	Número de preguntas	Índice de dificultad	Índice de dificultad con corrección de los efectos del azar	Índice de discriminación	Correlación biserial puntual	Dificultad TRI	Discriminación según la TRI
Aparatos	88	0,6618 (0,1595)	0,5449 (0,2313)	0,3147 (0,1369)	0,3521 (0,1317)	-0,5086 (1,231)	0,9677 (0,5366)
Básicas	20	0,6730 (0,1757)	0,5468 (0,2437)	0,3293 (0,1688)	0,3746 (0,1278)	-0,2634 (1,679)	1,0386 (0,4649)
Otras	98	0,6958 (0,2072)	0,5919 (0,2893)	0,2564 (0,147)	0,3544 (0,1454)	0,3092 (10,245)	1,0343 (0,5803)
Total	206	0,6791 (0,185)	0,5674 (0,2614)	0,2884 (0,1475)	0,3554 (0,1375)	-0,0957 (7,122)	1,0063 (0,5502)

y dicha diferencia es ligeramente superior a las cinco centésimas.

Los índices de discriminación de los tres bloques y del promedio se pueden clasificar como ‘buenos’; el más alto es el de las asignaturas básicas y el más bajo el de otras asignaturas. En lo referente al coeficiente de correlación biserial puntual, también los índices de todos los bloques y el global se pueden clasificar como ‘buenos’. El valor más alto corresponde a ‘Asignaturas básicas’, y el más bajo, a ‘Aparatos’.

En lo relativo a la dificultad calculada según la teoría de respuesta al ítem, el bloque de preguntas más difícil fue el de ‘Aparatos’, seguido de ‘Básicas’ y, finalmente, el de ‘Otras especialidades’, mientras que en lo referente a la discriminación, el bloque de preguntas más discriminativo resultó ser el de ‘Básicas’, seguido por ‘Otras especialidades’ y ‘Aparatos’.

Además, este estudio se ha desarrollado con la limitación de no contar con datos de los participantes como el número de años desde que obtuvieron su título de grado o licenciatura en Medicina, su edad o el género. Finalmente, como fortaleza, cabe señalar que se presenta un análisis que combina indicadores propios de la teoría clásica de los test con otros de la teoría de respuesta al ítem y que la metodología empleada es comparable con la que se usó para el análisis de convocatorias previas [10-15].

Finalmente, cabe señalar que la prueba MIR se caracteriza por presentar una alta fiabilidad, así como una estabilidad psicométrica entre convocatorias. El examen MIR de la convocatoria de 2022 es el que ha alcanzado un valor promedio de discriminación de sus preguntas más alto desde que los autores comenzaron a medirlo, en la convocatoria del año 2009. Además, el promedio de la dificultad de sus preguntas es similar al medido en los dos años anteriores.

Discusión

Según resultados obtenidos y al igual que en convocatorias anteriores [10,13,14], la prueba MIR de 2022 es altamente fiable. En lo referente a las limitaciones de este estudio, al igual que en otros [14], la más significativa es que no se dispone de información de todos los opositores, sino únicamente de las respuestas de un subconjunto de éstos que, probablemente, tengan en promedio un baremo ligeramente superior al de la población. A pesar de esto, resulta de interés, dado que hasta el momento no se ha publicado artículo alguno relativo a esta convocatoria o alguna otra reciente que haya analizado toda la población. De hecho, el último estudio [8] que analizó la población completa de aspirantes presentados a la prueba MIR fue el correspondiente a las convocatorias de 2005 y 2006.

Bibliografía

1. Real Decreto 589/2022, de 19 de julio, por el que se regulan la formación transversal de las especialidades en Ciencias de la Salud. BOE número 173, de 20 de julio de 2022.
2. Ley 44/2003, de 21 de noviembre, de ordenación de las profesiones sanitarias. BOE número 280, de 22 de noviembre de 2003.
3. Sitges-Serra A. El examen MIR, a examen. FEM 2020; 23 (Supl 1): S23-5.
4. Palés-Argullós J. Contenidos y calidad del examen MIR. FEM 2020; 23 (Supl 1): S11-3.
5. Lobato RD, Lagares A, Villena V, Alen JF, Jiménez-Roldan L, Munárriz PM. El método de selección de los residentes en España. Análisis del examen MIR y propuesta de una nueva metodología. Neurocirugía 2015; 26: 53-63.
6. Ibáñez L. EUNACOM: un esfuerzo señero de las facultades de medicina chilenas. Rev Chil Ped 2013; 84: 365-6.
7. Rivière E, Quinton A, Neau D, Constans J, Vignes JR, Dehail P. Analyse docimologique des premières Épreuves classantes nationales informatisées (ECNi) de 2016: propositions d'amélioration. Rev Méd Interne 2019; 40: 47-51.

8. Bonillo A. Pruebas de acceso a la formación sanitaria especializada para médicos y otros profesionales sanitarios en España: examinando el examen y los examinados. *Gac San* 2012; 26: 231-5.
9. Rodríguez-Díez MC, Alegre M, Díez N, Arbea L, Ferrer M. Technical flaws in multiple-choice questions in the access exam to medical specialties ('examen MIR') in Spain (2009-2013). *BMC Med Educ* 2016; 16: 1-8.
10. Baladrón J, Curbelo J, Sánchez-Lasheras F, Romeo-Ladrero JM, Villacampa T, Fernández-Somoano A. El examen al examen MIR 2015: aproximación a la validez estructural a través de la teoría clásica de los tests. *FEM* 2016; 19: 217-26.
11. Baladrón J, Sánchez-Lasheras F, Villacampa T, Romeo-Ladrero JM, Jiménez-Fonseca P, Curbelo J. El examen MIR 2015 desde el punto de vista de la teoría de respuesta al ítem. *FEM* 2017; 20: 29-38.
12. Baladrón J, Sánchez-Lasheras F, Romeo-Ladrero JM, Curbelo J, Villacampa-Menéndez P, Jiménez-Fonseca P. Evolución de los parámetros dificultad y discriminación en el ejercicio de examen MIR. Análisis de las convocatorias de 2009 a 2017. *FEM* 2018; 21: 181-93.
13. Baladrón J, Sánchez-Lasheras F, Romeo-Ladrero JM, Villacampa T, Curbelo J, Jiménez-Fonseca P. The MIR 2018 exam: psychometric study and comparison with the previous nine years. *Medicina (Kaunas)* 2019; 55: 751.
14. Baladrón J, Sánchez-Lasheras F, Curbelo-García J, García-Guerrero A, Peñalver-San Cristóbal C, Villacampa-Castro T, et al. Estudio psicométrico de las pruebas MIR de 2020 y 2021. *FEM* 2023; 26: 19-28.
15. Baladrón J, Sánchez-Lasheras F, Villacampa T, Romeo-Ladrero JM, Jiménez-Fonseca P, Curbelo J. Propuesta metodológica para la detección de preguntas susceptibles de anulación en la prueba MIR. Aplicación a las convocatorias 2010 a 2015. *FEM* 2017; 20: 161-75.
16. Muñiz J. Teoría clásica de los test. 2 ed. Madrid: Pirámide; 2002.
17. Guilbert JJ. Educational handbook for health personnel 6th ed. Geneva: WHO; 1998.
18. Pruebas selectivas para el acceso a plazas de formación de médicos especialistas (1982-1992). Madrid: Ministerio de Sanidad y Consumo; 1993.
19. Pruebas selectivas para el acceso a plazas de formación de médicos especialistas. Validez estructural, diseño y capacidades exploradas (1988-1992). Madrid: Ministerio de Sanidad y Consumo; 1993.
20. Lord FM. Applications of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum; 1980.