

## El baremo académico en el acceso a la formación médica especializada en España

C. Delia Dávila-Quintana, Beatriz G. López-Valcárcel, Patricia Barber, Vicente Ortún

**Introducción.** La selección en la formación sanitaria especializada de los médicos (MIR) se basa en la equidad interpersonal que premia a los candidatos según su esfuerzo. El rendimiento académico, a través del baremo, refleja el 'esfuerzo de fondo', mientras que el examen MIR mide el 'esfuerzo *sprint*'. Este artículo discute los pros y los contras de mantener el baremo académico, que actualmente pesa un 10%. Por un lado, mide capacidades y habilidades más allá del mero conocimiento que aporta el examen; pero, por otro, puede tener problemas de validez, no modificar elecciones debido a su bajo peso o no ser coste-efectivo.

**Materiales y métodos.** Registro de adjudicatarios de plazas MIR 2012-2013 (Ministerio de Sanidad, Servicios Sociales e Igualdad) y encuesta propia a los MIR-1 2012-2013. Modelo multinivel para estimar el efecto universidad en el baremo académico de los candidatos.

**Resultados.** Hay diferencias sistemáticas en el baremo de los candidatos atribuibles a la universidad de procedencia cuando se comparan todas las universidades, españolas y extranjeras (correlación intragrupo: 47%), pero no hay diferencias significativas entre universidades españolas (correlación intragrupo: 7%). El 62% de los candidatos sube o baja menos de cien puestos en el *ranking*. Al menos el 1% de los candidatos cambiaría su especialidad elegida si se suprimiera el baremo académico, lo que supone un coste máximo de 576 euros por cambio.

**Conclusiones.** El baremo académico es válido para las universidades españolas, pero no para las extranjeras. Su supresión modificaría marginalmente el orden de elección y las asignaciones MIR.

**Palabras clave.** Baremo académico. Efecto universidad. Modelos multinivel. Pruebas MIR.

### The academic grade in access to specialist medical training in Spain

**Introduction.** The selection on specialized health training of doctors (MIR) is based on interpersonal fairness that rewards candidates by effort. Academic achievement, through the average grade, reflects the long-distance race, while the MIR test measures the sprint effort. This article discusses the pros and cons of maintaining the academic grade, that currently weighs 10%. On the one hand, it measures skills and abilities beyond the mere knowledge that provides the test, but on the other, it may have problems of validity, it could not change elections, given their low weight, and/or it could be not cost-effective.

**Materials and methods.** Database of MIR 2012-2013 (Ministry of Health, Social Services and Equality) and own survey of MIR-1 2012-2013. We used a multilevel model to estimate the effect of the university on the candidate's academic grade.

**Results.** There are systematic differences in the grades of the candidates attributable to the university when all universities, Spanish and foreign are compared (intragroup correlation: 47%). There are no significant differences between Spanish universities (intragroup correlation: 7%). 62% of the candidates move less than a hundred places in the ranking if the academic grade would be suppressed, and at least 1% of the candidates would change their chosen specialty. It represents a maximum cost of 576 euros per change.

**Conclusions.** The academic grade is valid for the Spanish universities but not for foreign universities. Its suppression would marginally alter the order of choice and MIR assignments.

**Key words.** MIR tests. Multilevel models. Scales. University bias.

Departamento de Métodos Cuantitativos; Universidad de Las Palmas de Gran Canaria; Las Palmas, Gran Canaria (C.D. Dávila-Quintana, B.G. López-Valcárcel, P. Barber). Departamento de Economía y Empresa; Universitat Pompeu Fabra; Barcelona, España (V. Ortún).

#### Correspondencia:

Profa. Patricia Barber Pérez. Departamento de Métodos Cuantitativos D-4-10. Universidad de Las Palmas de Gran Canaria. Campus de Tafira. E-35017 Las Palmas de Gran Canaria.

#### Fax:

+34 928 458 225.

#### E-mail:

patricia.barber@ulpgc.es

#### Financiación:

Proyecto de investigación ECO2010-21558 del Plan Nacional de I+D+i y Subdirección General de Ordenación Profesional del Ministerio de Sanidad, Servicios Sociales e Igualdad.

#### Conflicto de intereses:

No declarado.

#### Competing interests:

None declared.

© 2015 FEM

## Introducción

En España, la asignación de candidatos a plazas MIR de formación sanitaria especializada se basa en la elección ordenada [1] según los resultados de un examen nacional tipo test (90%) y de la media del baremo académico durante la licenciatura/grado en medicina (10%). Hasta 2009, el baremo académico ponderaba un 25% [2]. De la suma de ambos resulta la posición en el *ranking* final, que otorga la prioridad en la elección de plaza de formación MIR. La plaza se define por la especialidad (entre las 47 reconocidas actualmente) y el centro. Aunque ha habido algunas reformas, como la eliminación del baremo de los estudios de doctorado en 2000 [3] o la disminución del peso de la tesis doctoral en 2008 [4], las reglas del juego han permanecido relativamente estables. Hay una convocatoria anual: 6.389 plazas MIR en 2012-2013 y 6.149 plazas en 2013-2014.

El examen MIR es un cuestionario tipo test. Los médicos extracomunitarios están afectados por un cupo máximo de plazas, que en primera vuelta fue del 8% en 2012-2013 (4% en 2013-2014). Por su parte, el baremo académico pondera las calificaciones de los expedientes académicos y la tesis doctoral con una puntuación entre 1 y 5. Mientras el baremo académico refleja el 'esfuerzo de fondo', el examen MIR mide el 'esfuerzo *sprint*'.

Conceptualmente, la asignación de candidatos a plazas MIR debería cumplir dos criterios que podrían entrar en conflicto: la eficiencia social y la equidad interpersonal [5]. La eficiencia social es la capacidad de resolver problemas de salud –conseguir AVACS (años de vida ajustados por calidad)– en un horizonte temporal de largo plazo, del conjunto de especialistas formados en una cohorte determinada. Tiene que ver, pues, con la productividad o efectividad de su trabajo como especialistas. La equidad interpersonal se define por el premio diferencial a los candidatos según su esfuerzo o mérito. Este es el único criterio que se emplea en la práctica en España [6].

En los últimos años ha surgido cierto debate acerca de si es pertinente o no el uso del baremo académico en la obtención del *ranking* de elección MIR. En contra se aduce que, dada la variedad y heterogeneidad de facultades de procedencia de los aspirantes, no existe garantía de que las calificaciones sean homogéneas, que reflejen méritos iguales. Otro argumento es que, debido a su bajo peso en la puntuación final, el baremo prácticamente 'no modifica conducta', por lo que su beneficio no compensaría el coste de la baremación.

Los argumentos a favor ganan peso tras la implantación de los nuevos títulos de grado posreforma

de Bolonia. Las habilidades de comunicación, de trabajo en equipo, etc. que se aplican a lo largo del proceso de evaluación del grado se reflejan, aunque indirectamente, en el baremo académico.

El objetivo de este trabajo es discutir y valorar los argumentos a favor y en contra de mantener el baremo académico como criterio de ordenación de los candidatos MIR, en base a un marco conceptual definido y a una base de datos, en parte inéditos, procedentes de encuestas a los nuevos médicos residentes en su primera semana de incorporación a la plaza, en 2012 y 2013.

## Materiales y métodos

La base de datos de elecciones individuales de plazas MIR en la convocatoria 2012-2013 se ha enlazado con los datos de una encuesta realizada por el Ministerio de Sanidad, Servicios Sociales e Igualdad (Subdirección General de Ordenación Profesional) con el Grupo de Economía de la Salud de la Universidad de Las Palmas de Gran Canaria. Dicha encuesta tiene dos oleadas, para las convocatorias 2011-2012 y 2012-2013 (Tabla I). La población son los MIR-1 recién incorporados a su plaza. En la encuesta, que se cumplimenta *online*, se pregunta al adjudicatario sobre la plaza elegida y sobre la preferida, se pide su valoración de una serie de atributos asociados a esa elección, y se pregunta por sus preferencias, movilidad, reespecialización y recirculación. En 2012-2013 se han recogido 4.221 respuestas (67% de la población), aunque la muestra efectiva es de 3.397 adjudicatarios MIR porque ha sido imposible enlazar ambas bases (encuesta y registro del ministerio) cuando no aportaban una dirección enlazable de correo electrónico.

En primer lugar, presentamos una descripción numérica y gráfica de las puntuaciones medias de baremo y de *ranking* por convocatorias y universidades, incluyendo la correlación entre las notas del expediente académico y del examen MIR para la convocatoria 2012-2013, distinguiendo entre universidades españolas y extranjeras.

En segundo lugar, valoramos la validez del baremo académico como indicador de esfuerzo de fondo de los candidatos. Para ello se parte de la hipótesis nula de homogeneidad entre universidades: las puntuaciones del baremo académico de las distintas universidades miden lo mismo y son comparables. Un sobresaliente en la universidad A equivale en conocimientos a un sobresaliente en la universidad B. Para contrastar esta hipótesis estimamos el modelo siguiente:

$$\text{Bar}_{ij} = \beta Ex_{ij} + \gamma_j + \eta_i + \varepsilon_{ij} \text{ [ec. 1]},$$

donde la nota media del expediente académico (*Bar*) del candidato *i*, que ha estudiado en la universidad *j* y se presenta en la convocatoria 2012-2013, dada su nota del examen MIR (*Ex*), difiere entre universidades. Un candidato que tiene mejor formación de base también rendirá más en el examen MIR, de ahí la correlación parcial entre ambas que asume el modelo. Pero después de controlar por dicha correlación, todavía hay un efecto sistemático de la universidad ( $\gamma_j$ ), que queremos estimar, y del candidato ( $\eta_i$ ).  $\varepsilon_{ij}$  es el error aleatorio del modelo, que se supone normal. Si los baremos de las distintas universidades son comparables, el término  $\gamma_j$  debería ser no significativo. En el modelo,  $\gamma_j$  se supone aleatorio y su varianza es uno de los parámetros a estimar. Una vez estimadas ambas varianzas, la de  $\gamma_j$  y la del error individual  $\varepsilon_{ij}$ , calculamos la correlación intragrupo, que mide el porcentaje de la varianza total en el baremo que es atribuible a la universidad. El efecto de cada universidad se estima con el predictor *a posteriori* bayesiano [7].

En relación al argumento de que el baremo no cambia conductas, hacemos dos análisis: cuantificamos los cambios en el orden de elección de plaza MIR que se producirían en caso de suprimir el 10% del baremo académico para la convocatoria 2012-13, y estimamos para esa misma convocatoria los cambios de elección de especialidad que se producirían, a partir de la información suministrada por cada elector en la encuesta MIR, sobre la plaza que habría elegido si no hubiese tenido restricciones de orden.

Por último, en relación al argumento del coste, cuantificamos el coste máximo de un cambio en la elección de especialidad.

## Resultados

En la convocatoria 2012-2013 el 22% de los aspirantes quedó excluido por no alcanzar la nota mínima requerida en el examen MIR. Nueve de las diez primeras universidades en nota media del examen MIR son españolas, pero siete de las diez universidades en las que más mejoran su posición los adjudicatarios gracias al baremo son extranjeras.

Las correlaciones entre la nota del baremo y la del examen para los alumnos de universidades españolas es positiva y bastante alta (0,57), pero es mucho menor para los de universidades extranjeras (0,16). En la convocatoria 2012-2013 (incluyendo sólo aquellas con más de diez alumnos con plaza MIR asignada en 2012-2013), ocho de las diez universidades con

**Tabla I.** Encuesta a médicos adjudicatarios de plaza MIR (convocatoria 2012-2013).

1. ¿Qué especialidad has elegido? (lista desplegable de especialidades)
2. ¿Qué centro has elegido? (lista desplegable de centros según especialidad) ¿Por qué? Ordena los siguientes atributos según la importancia que ha tenido cada uno en tu elección:
Retribución económica
Reconocimiento de la labor profesional por parte de los pacientes
Horario, vacaciones y posibilidad de compatibilizar con la vida familiar
Cercanía a tu domicilio (residencia habitual)
Posibilidades de promoción y desarrollo profesional futuro dentro de la especialidad (nuevos campos, nuevas técnicas, avances científicos...)
Prestigio, reconocimiento por compañeros y reconocimiento social
Posibilidad de investigar
Seguridad/estabilidad en el empleo
Posibilidad de ejercer la práctica privada
Probabilidad de encontrar trabajo con facilidad
3. La plaza que has elegido, ¿a cuántos kilómetros aproximadamente está de tu casa (residencia habitual)?
4. ¿Has tenido/tendrás que cambiar tu residencia habitual para ocupar la plaza MIR que has elegido?
5. Si hubieras podido elegir cualquier especialidad en esta convocatoria sin restricción de nota, ¿cuál habrías elegido como primera opción? Siguiendo con la pregunta anterior (si hubieras podido elegir tu especialidad preferida sin restricción), ¿qué centro habrías elegido para hacerla?
6. ¿Qué especialidad habrías elegido como segunda opción (recuerda, si no hubieras tenido restricción de nota)? ¿Qué centro habrías elegido para esa segunda opción?
7. ¿Es la primera vez que te examinas del MIR? ¿Has empezado alguna otra especialidad MIR antes? ¿Cuál? ¿La has terminado? ¿En que año?
8. Sexo
9. Edad
10. Nacionalidad
11. ¿Cuál es tu provincia de residencia habitual (generalmente, la que figura en tu DNI/pasaporte)?
12. Código postal de tu residencia habitual

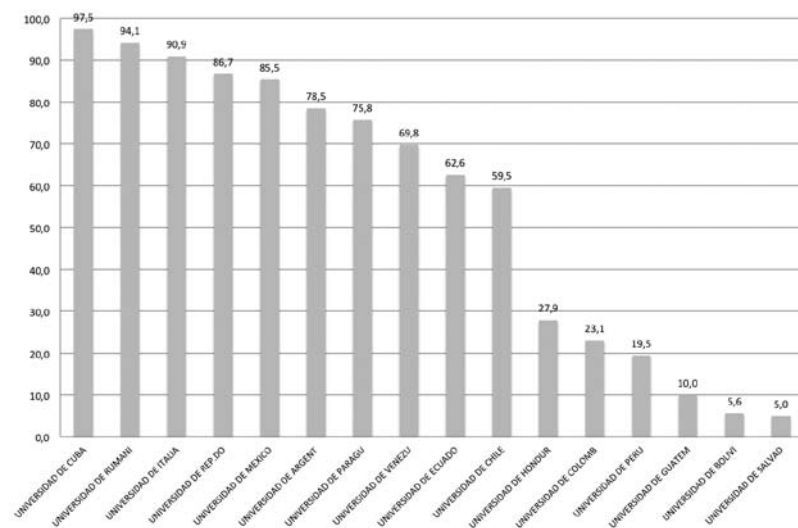
mejor nota media del expediente académico son extranjeras, pero cinco de ellas están entre las de peor resultado promedio en el examen MIR.

La estimación del modelo multinivel [ec. 1] confirma efectos significativos de las universidades en el baremo académico ( $p = 0,000$ ) después de ajustar por sexo y por la nota del examen MIR. Los detalles del modelo se encuentran en la tabla II. La correlación intragrupo (las variaciones en la nota del expediente atribuibles a las diferencias sistemáticas entre universidades) es del 43%. Por tanto, los datos

**Tabla II.** Modelo multinivel del baremo académico individual.

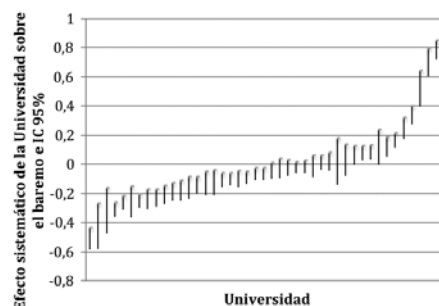
Modelo de efectos mixtos	Test Wald $\chi^2(3) = 2430,77$				
Numero de observaciones: 6.347	Log likelihood = -2761,1008				
Variable agrupación: universidad	Prob > $\chi^2 = 0,0000$				
Numero de grupos: 73					
Baremo acad.	Coefficiente	Error estándar	z	P >  z	Int. confianza 95%
Ptosexam	0,0029047	0,0000589	49,29	0,000	0,0027892-0,0030202
mujer	0,025789	0,0100874	2,56	0,011	0,0060181-0,0455598
_cons	0,8212789	0,0472583	17,38	0,000	0,7286543-0,9139034
Parámetros efectos aleatorios	Estimación	Error estándar	Int. confianza 95%		
var(_cons)	0,1035023	0,020878	0,0697027-0,1536917		
var(Residual)	0,1348839	0,0024086	0,1302449-0,1396882		
Test LR frente a regresión lineal: $\chi^2(13) = 1601,64$ . Prob $\geq \chi^2 = 0,0000$					

**Figura 2.** Porcentaje de aspirantes que mejoran su posición final gracias al baremo (universidades extranjeras con más de diez aspirantes).



sugieren que hay diferencias sistemáticas entre universidades en la forma de calificar a los estudiantes de medicina. La figura 1 representa las estimaciones (calculadas mediante el predictor bayesiano) de esos efectos, y los correspondientes intervalos de confianza al 95%, para las 44 universidades con más de diez estudiantes (las extranjeras no aparecen individualizadas, sino por países, como por ejemplo, ‘universidades de Rumanía’). Salvo una, las universida-

**Figura 1.** Estimación del efecto universidad en el baremo académico estimado por el modelo multinivel (predictor bayesiano *a posteriori*).



des de ambos extremos (es decir, las que sistemáticamente dan calificaciones muy bajas o muy altas) son extranjeras. En el extremo negativo (bajas calificaciones) están las de Bolivia, El Salvador, Guatemala, Perú, Honduras y Colombia; en el positivo (calificaciones infladas), Rumanía, Cuba, México, Italia, República Dominicana, Argentina, Venezuela, Ecuador y Paraguay (Fig. 1).

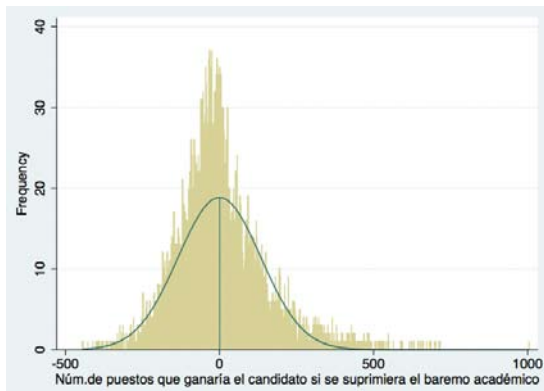
Contestando a la primera pregunta del modelo se concluye que no se percibe heterogeneidad sistemática en la valoración de los estudiantes de grado entre universidades españolas, pero sí respecto a las extranjeras. Cuando el modelo se estima para los candidatos de universidades españolas, la correlación intragrupo baja al 7%.

La figura 2 representa el porcentaje de estudiantes de las universidades extranjeras (con más de 10 candidatos adjudicatarios) que mejoran su posición en el *ranking* gracias al baremo académico. La universidad española con mayor porcentaje de titulados que mejoran su posición gracias al baremo tiene un 63,4%, 13 puntos por encima del 50% que se esperaría bajo hipótesis de azar.

La supresión del baremo modificaría el *ranking* de elección, aunque el 62% de los candidatos subirían o bajarían menos de 100 puestos en el *ranking* final (Fig. 3). En la convocatoria 2012-2013, el 53% de los candidatos mejoraría su lugar en el *ranking* si se eliminara el baremo académico. En general, los más beneficiados de ese *sprint* final, los que mejorarían su posición si no existiese el baremo, serían los más jóvenes, los hombres (55% frente al 51% de mujeres) y los de nacionalidad española (56% frente al 45% de extranjeros).

Para trasladar los cambios en el *ranking* a cambios en las elecciones hemos cruzado los datos del

**Figura 3.** Cambios de puesto en el *ranking* final si se suprime el baremo académico.



registro MIR con los de la encuesta. Ésta concluye que el 57% de los electores en 2012-2013 han podido elegir su especialidad preferida (62% para los españoles y 38% para los extranjeros). Naturalmente, el porcentaje es menor a medida que se desciende en el *ranking* (Fig. 4). Entre los mil primeros (96% de vocaciones satisfechas) y los mil últimos (20%) hay una gran diferencia. Naturalmente, el éxito se reparte desigualmente entre las especialidades. Así, mientras que todos los MIR-1 de cirugía pediátrica o torácica seguirían eligiendo esa especialidad, tres de cada cuatro que han elegido medicina familiar y comunitaria no lo hubiese hecho de haber tenido todas las opciones disponibles.

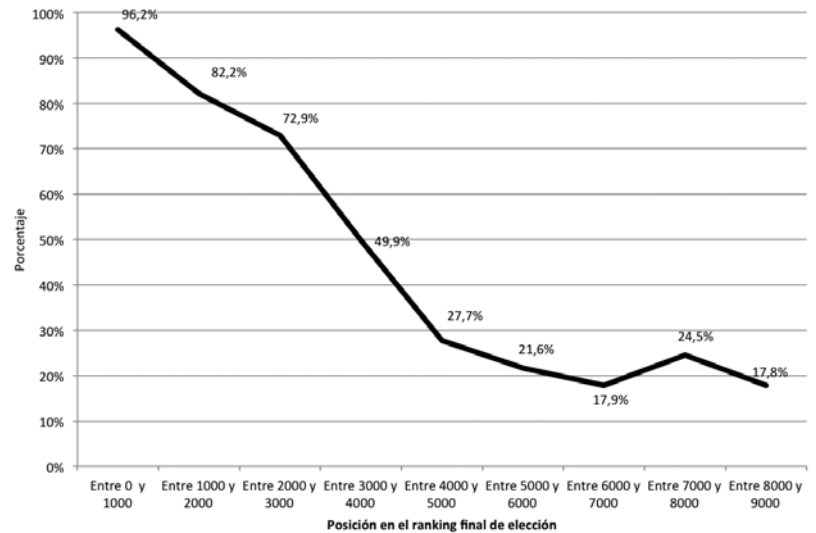
Los cambios en el *ranking* que se producirían de suprimir el baremo académico harían que el 0,82% de los candidatos que han elegido su especialidad preferida no podrían hacerlo, por bajar en el *ranking*, y el 1,35% de los que no tenían acceso a su especialidad preferida lo conseguirían. Por tanto, contestando a la segunda pregunta, mantenemos que suprimir el baremo académico modificaría elecciones, pero sólo marginalmente.

Dado que el coste de la baremación es aproximadamente de 3,30 euros por candidato admitido y que hemos calculado una cota inferior para el porcentaje de cambios, el coste de cambiar una elección es como máximo de 576 euros.

## Discusión

Sería de esperar que los estudiantes con buena trayectoria académica obtuvieran también buenos re-

**Figura 4.** Porcentaje de éxito del sistema (plaza elegida igual a plaza preferida).



sultados en el examen de acceso MIR, lo que quedaría reflejado en correlaciones altas entre esos dos componentes del *ranking* final para cada uno de los adjudicatarios. Hemos demostrado que esas correlaciones son relativamente altas y significativas para las universidades españolas, pero no para las extranjeras. Hay signos de heterogeneidad en la forma en que las distintas universidades extranjeras califican a sus estudiantes, por lo que el baremo no es un indicador válido del esfuerzo de fondo, al menos cuando se incluye a los candidatos formados como médicos en el extranjero.

Algunos estudiantes consiguen mejorar sustancialmente su posición de elección mediante un buen examen; otros, sin embargo, ven cómo su posición de ventaja, fruto del trabajo de fondo reflejado en un buen expediente académico, empeora notablemente sus posibilidades de elección tras la prueba de *sprint* (examen nacional). El análisis que hemos presentado concluye que el efecto de la inclusión del baremo es muy desigual, al menos entre universidades españolas y extranjeras. A las diferencias sistemáticas y persistentes entre universidades en el rendimiento en el examen MIR en España [8,9] hay que añadir una clara evidencia de sobreacreditación académica para los estudiantes procedentes de determinadas universidades extranjeras.

Este hecho, que si bien por sí mismo podría no justificar suficientemente la eliminación del baremo (dada la poca capacidad de modificar conducta en su



formulación actual), podría ser el principal argumento frente a los que defienden incrementar la ponderación actual del expediente académico, al menos hasta que se introdujera algún ajuste en el proceso de baremación que tuviera en cuenta las diferencias sistemáticas no justificadas en los baremos.

Además, la inequidad entre universidades españolas y extranjeras tras el proceso de evaluación del expediente académico se agrava en cuanto que el principal argumento para defender la inclusión de la baremación es que, en el marco del Plan de Bolonia, se evalúan competencias, habilidades, capacidades, etc., pero este plan afecta sólo a las universidades europeas.

Si el único criterio decisor fuera el coste del proceso, habría que decidir hasta qué punto la baremación justifica esta diferencia de 'éxito'. La baremación de los expedientes tiene un coste unitario en torno a 3,30 euros. Prácticamente todos los admitidos a examen, que en cada una de las tres últimas convocatorias han sobrepasado los 13.000, requieren la evaluación de su expediente, aunque finalmente el número de presentados es del 85% sobre admitidos y las plazas disponibles, prácticamente la mitad. Por otra parte, y aunque no se conoce su coste con exactitud, la gestión completa del proceso MIR (admisión, gestión, coordinación, corrección...) requiere un importante despliegue de recursos económicos, tanto puntuales como estructurales. Por tanto, en términos relativos, el coste de la baremación es pequeño y los argumentos de coste no justificarían la supresión del baremo.

Una posible vía de avance podría ser utilizar el *Situational Judgement Test* (SJT), como en Reino Unido, Bélgica y otros países. Se trata de una herramienta de selección útil y coste-efectiva [10,11] que evalúa también constructos no académicos, como

las habilidades interpersonales y la capacidad de trabajar en equipo, con el fin de predecir el futuro rendimiento en el trabajo y los resultados de la formación, tan importantes en esta profesión.

#### Bibliografía

1. Orden de 27 de junio de 1989 por la que se establecen las normas reguladoras de las pruebas selectivas para el acceso a plazas de formación sanitaria especializada. Boletín Oficial del Estado n.º 153, de 28 de junio de 1989.
2. Orden SAS/2158/2010 de 28 de julio de 2010 por la que se establecen las normas reguladoras de las pruebas selectivas para el acceso a plazas de formación sanitaria especializada. Boletín Oficial del Estado n.º 191, de 7 de agosto de 2010.
3. Orden 13.281 de 11 de julio de 2000 por la que se modifica el baremo de los méritos académicos de las pruebas selectivas para el acceso a plazas de formación sanitaria especializada. Boletín Oficial del Estado n.º 167, de 13 de julio de 2000.
4. Real Decreto 183/2008, de 8 de febrero, por el que se determinan y clasifican las especialidades en ciencias de la salud y se desarrollan determinados aspectos del sistema de formación sanitaria especializada. Boletín Oficial del Estado n.º 45, de 21 de febrero de 2008.
5. Harris JE, López-Valcárcel BG, Barber P, Ortún V. Efficiency versus equity in the allocation of medical specialty training positions in Spain: a health policy simulation based on a discrete choice model. NBER Working Paper; febrero de 2014.
6. Lorusso N, González López-Valcárcel B. Variabilidad de los procesos selectivos para el acceso a las especialidades médicas en Europa: entre convergencia y divergencia. FEM 2013; 16: 215-8.
7. Box GEP, Tiao GC. Bayesian inference in statistical analysis. New York: Addison-Wesley; 1973.
8. López-Valcárcel BG, Ortún V, Barber P, Harris JE, García B. Ranking Spain's medical schools by their performance in the national residency examination. Rev Clin Esp 2013; 213: 428-34.
9. González López-Valcárcel B, Ortún V, Barber P, Harris JE. Important differences between faculties of medicine. Implications for family and community medicine. Aten Primaria 2014; 46: 140-6.
10. Patterson F, Ashworth V, Zibarras L, Coan P, Kerrin M, O'Neill P. Evaluations of situational judgement tests to assess non-academic attributes in selection. Med Educ 2012; 46: 850-68.
11. Patterson F, Carr V, Zibarras L, Burr B, Berkin L, Plint S. New machine-marked tests for selection into core medical training: evidence from two validation studies. Clin Med 2009; 5: 417-20.