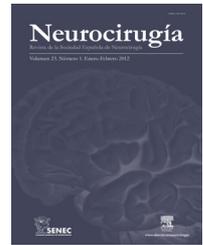


NEUROCIRUGÍA

www.elsevier.es/neurocirugia



Artículo de opinión

La selección de los graduados en Medicina para los puestos de residencia. Estudio comparativo de los métodos usados en diferentes países



Ramiro D. Lobato^{a,b,*}, Alfonso Lagares^{a,b}, Victoria Villena^{a,c}, Jorge García Seoane^a, Luis Jiménez-Roldán^{a,b}, Pablo M. Munarriz^{a,b}, Ana M. Castaño-Leon^{a,b} y José F. Alén^{a,b}

^a Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España

^b Servicio de Neurocirugía, Hospital Universitario 12 de Octubre, Madrid, España

^c Servicio de Neumología, Hospital Universitario 12 de Octubre, Madrid, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 4 de diciembre de 2014

Aceptado el 15 de diciembre de 2014

On-line el 16 de enero de 2015

Palabras clave:

Grado de medicina

Selección de residentes

Especialización de posgrado

Residencia

Test de conocimientos

Examen MIR

Competencia clínica

RESUMEN

Introducción: El diseño de una metodología apropiada para la selección de los graduados que acceden a los puestos de residencia tiene gran importancia porque, aparte de su eficiencia en la elección de los candidatos, condiciona decisivamente el funcionamiento curricular en las escuelas de medicina, incluyendo la manera de estudiar y aprender del estudiante y la actividad docente del profesor y su motivación para perseguir de la innovación curricular. En la actualidad existe una gran dispersión en los métodos de selección de residentes en diferentes países, evidenciando que no hay ninguno perfecto. El uso aislado de criterios tales como el peso del expediente académico, los test de conocimiento teórico, determinadas pruebas prácticas de la competencia clínica, o las entrevistas tienen un valor limitado para predecir qué candidatos van a ser más competentes durante la residencia o la práctica autónoma.

Objetivos: Describir las metodologías utilizadas para la selección de los residentes en los países de nuestro entorno prestando especial atención a las empleadas en el Reino Unido y los EE. UU. donde se analiza de manera sistemática el funcionamiento de los currícula en las escuelas de medicina. Se comentan las ventajas y desventajas de los exámenes de ámbito nacional para otorgar el grado y/o construir el listado de acceso a la residencia, y su utilidad para favorecer el proceso de convergencia y armonización de los grados y los programas de residencia a nivel transnacional europeo o internacional.

El presente análisis se utiliza para establecer una comparación con la metodología de selección aplicada actualmente en España en un intento de diseñar un nuevo método multicriterio más eficiente que se describirá en otro artículo publicado en el próximo número de esta revista

Véase contenido relacionado en DOI: <http://dx.doi.org/10.1016/j.neucir.2014.12.005>.

* Autor para correspondencia.

Correo electrónico: ramirodiezlobato@gmail.com (R.D. Lobato).

<http://dx.doi.org/10.1016/j.neucir.2014.12.003>

1130-1473/© 2014 Sociedad Española de Neurocirugía. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Conclusiones: Los métodos multicriterio para la selección de los residentes en el Reino Unido y los EE. UU. parecen los más consistentes por lo que serán utilizados para construir el nuevo método que podría ser aplicado en España. Aunque muchos expertos en docencia médica rechazan los exámenes de ámbito nacional para otorgar el grado y/o establecer el orden de acceso a los puestos de residencia porque limitan la diversidad y la innovación curriculares, cuando están bien diseñados sirven para verificar el alcance del nivel mínimo aceptable de competencia clínica del graduado sin distorsionar el buen funcionamiento y la mejora progresiva del currículum.

© 2014 Sociedad Española de Neurocirugía. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Selection of medical graduates for residency posts. A comparative study of the methodologies used in different countries

A B S T R A C T

Keywords:

Medical grade
Selection for residency
Postgraduate specialization
Residency
Clinical competence
MIR examination
Knowledge testing

Introduction: The design of an appropriate method for the selection of medical graduates for residency posts is extremely important, not only for the efficiency of the method itself (accurate identification of most competent candidates), but also for its influence on the study and teaching methodologies operating in medical schools. Currently, there is a great variation in the criteria used in different countries and there is no definitively appropriate method. The use of isolated or combined criteria, such as the marks obtained by students in medical schools, their performance in tests of theoretical knowledge and evaluations of clinical competence, or personal interviews, have a limited value for identifying those candidates who will perform better during the residency and later on during independent practice.

Objectives: To analyse the variability in the methodologies used for the selection of residents employed in different countries, in particular those used in the United Kingdom and USA, where external agencies and medical schools make systematic analyses of curriculum development. The advantages and disadvantages of national or transnational licensing examinations on the process of convergence and harmonization of medical degrees and residency programmes through Europe are discussed.

The present analysis is used to design a new and more efficient multi-criteria methodology for resident selection in Spain, which will be published in the next issue of this journal.

Conclusions: Since the multi-criteria methods used in UK and USA appear to be most consistent, these have been employed for designing the new methodology that could be applied in Spain. Although many experts in medical education reject national examinations for awarding medical degrees or ranking candidates for residency posts, it seems that, when appropriately designed, they can be used to verify the level of competence of graduating students without necessarily distorting curriculum implementation or improvement.

© 2014 Sociedad Española de Neurocirugía. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

Aún no se ha diseñado una metodología definitivamente satisfactoria para seleccionar a los graduados que compiten por los puestos de residencia. Los criterios utilizados en diferentes países son muy variables, lo que revela que no existe ningún método definitivamente bueno, y que estamos ante un problema por resolver. Cualquier método inadecuado perjudicará a unos candidatos frente a otros y condicionará negativamente el aprendizaje del estudiante y otros aspectos del funcionamiento curricular. Por ello, muchos expertos, agencias reguladoras e instituciones relacionadas con la docencia

médica enfatizan la necesidad de perfeccionar los criterios de selección y aunarlos en lo posible para favorecer la convergencia de programas y títulos a nivel del pre-posgrado entre diferentes países.

En España la selección de los residentes utiliza como criterio fundamental el rendimiento en un examen de conocimiento (MIR) que tiene una validez limitada para ordenar a los candidatos y ejerce una influencia negativa sobre el funcionamiento curricular¹⁻⁶. Con el objetivo de diseñar una metodología más eficiente hemos realizado un análisis descriptivo y comparativo de los métodos utilizados en los países europeos y Norteamérica que se presenta en este trabajo. El análisis de las limitaciones del método actualmente empleado

en España y la descripción de la nueva metodología que proponemos para sustituirlo se incluyen en otro estudio que será publicado en el próximo número de esta revista.

Análisis de los métodos de selección de residentes en los países europeos

La selección de los residentes en los países de la Unión Europea

En Europa se distinguen globalmente dos metodologías de selección⁶. El llamado *modelo mediterráneo*, aplicado en Francia, España, Italia, Portugal, Rumania y Grecia, utiliza una prueba escrita administrada a nivel nacional o regional que consiste en un test teórico con preguntas, usualmente de opción múltiple y respuesta única. En algunos países se añade una simulación escrita de caso clínico, o se valora el expediente académico, pero de manera diferente y con tendencia al escaso peso relativo. Este modelo, que tiene la ventaja de exponer a los candidatos a la misma prueba garantizando la llamada equidad, no evalúa, sin embargo, el nivel de competencia clínica adquirido en las escuelas o facultades, ni entrevista al candidato. El *modelo anglosajón*, que se aplica en el Reino Unido, Suecia, Alemania, Austria, Bélgica, Irlanda, Países Bajos, Polonia y República Checa, no utiliza exámenes de conocimiento de ámbito nacional y recurre a la solicitud individualizada del candidato y una entrevista personal. Este modelo tampoco determina el nivel de competencia clínica adquirido en las escuelas, y al ser de aplicación localista no garantiza la equidad de derechos del modelo mediterráneo. A continuación se describen con más detalle los criterios utilizados en los principales países europeos.

En Francia no se valora el expediente académico del grado y se recurre a un test teórico de ámbito nacional que está en fase de cambio; el nuevo examen, llamado @Examen Clasant National, a introducir en el 2015, se pasará con iPad y constará de una serie de casos clínicos con 20 preguntas concatenadas de cada uno de diferentes contenidos y formatos. En Portugal, tras finalizar el grado, se concede de manera inmediata a un año de internado, y seis meses después de finalizado el grado se pasa un examen teórico cuyo resultado ordena a los candidatos para el acceso a la especialidad; el test consta de 100 preguntas de elección múltiple con 5 opciones relacionadas con los bloques de Neurología, Digestivo, Neumología, Cardiología y Nefrología.

En Italia, el Ministerio de Investigación-Universidades convoca un examen de ámbito nacional en el que la puntuación máxima obtenible (135 puntos) se establece con: 1) la nota del expediente académico (máximo 15 puntos, que incluye la nota de graduación y la media aritmética de las asignaturas, además de tesis experimental y títulos de investigación si los hubiere); y 2) la puntuación en un test de conocimiento que tiene una primera parte común, con 70 preguntas sobre temas generales (70 puntos), y otra segunda con 50 preguntas de temas específicos (50 puntos); las preguntas son de elección múltiple con cinco opciones y respuesta única.

En Alemania cada servicio establece su propio procedimiento de selección; el jefe de servicio realiza una entrevista que es determinante, y no siempre se valora el expediente

académico del grado; tampoco se valora la nota del Staatsexam de suficiencia (tipo MIR que se realiza en 5.º curso), o del Praktisches Jahr que desde el 2013 incorpora un examen de competencia clínica tipo examen clínico objetivo estructurado (ECO); estos exámenes son usados para otorgar el grado, pero no para ordenar el acceso a los puestos de residencia; se valoran otros méritos como el trabajo de Doctorado cuando se hace durante la carrera, publicaciones con revisión por pares, o la participación en ONG. En Suecia se exige un periodo de preregistro (AT) de 18 meses y una vez conseguida la condición de preregistrado, el aspirante se dirige a una de las 6 universidades existentes para solicitar una plaza de formación en una de las especialidades reconocidas en el país; además, se realiza una entrevista. En los Países Bajos, donde la formación está vinculada también a las universidades, el candidato presenta una solicitud a la Universidad, y aporta una carta de presentación y su expediente académico, pero tampoco hay examen de entrada.

En el Reino Unido el graduado accede al Foundation Programme (2 años de inmersión clínica con el «role» de médico entre el grado de 5 años y la residencia) mediante un proceso discretamente competitivo, que se describe más adelante, y que incluye la valoración del expediente del grado y un test de juicio situacional que mide atributos profesionales para la práctica clínica⁷⁻¹⁰. El acceso a la residencia se basa en un proceso complejo y descentralizado en el que las plazas se convocan en el BMJ o el Lancet y el aspirante envía su expediente y una solicitud al servicio elegido donde se le convoca para una entrevista, y en algunos casos para la realización de un examen teórico. Sin embargo, hay que aclarar que la posición ocupada en el Foundation Programme condiciona el acceso a la residencia, por lo que consideramos el proceso de selección para entrar en este programa equivalente al que podría aplicarse para acceder a la residencia (ver más adelante).

A continuación se analizan más en detalle los criterios usados para otorgar el grado y/o clasificar a los candidatos para el acceso a la residencia en el Reino Unido y EE. UU. de selección que podría ser aplicado en nuestro país en sustitución del actual examen MIR. La escasa trazabilidad de publicaciones sobre el tema posiblemente existentes en algunos países punteros en docencia médica impide establecer otras comparaciones que pudieran resultar de utilidad para este propósito.

El método de selección en el Reino Unido

Los criterios para ordenar al graduado que va a acceder al Foundation Programme y elegir la «Foundation School» de su preferencia (desde la cual pasará a la residencia) son: 1) la Educational Performance Measure (expediente académico) que otorga entre 34 (10% percentil bajo) y 43 (10% percentil alto) puntos según las notas de la carrera, a los que se pueden añadir 7 puntos más (máximo) por méritos diversos tales como grados (doctor, máster) presentaciones, premios internacionales, o participación en ONG (máxima puntuación con este criterio de 50); así, la diferencia máxima entre el estudiante de mayor excelencia y otro con el más bajo rendimiento en la carrera es de 17 puntos; y, 2) el Situational Judgement Test

(SJT), que estima el juicio del candidato en diferentes situaciones durante la práctica clínica y otorga una puntuación entre 0 y 50 puntos⁹. Analicemos el valor de estos dos criterios de selección y las limitaciones de su uso.

La valoración y el peso del expediente académico en el Reino Unido. En este país existen dos sistemas de inspección y encuesta de los métodos de evaluación aplicados en las escuelas; la Quality Assurance Agency for Higher Education, que funciona como agencia externa, y la Quality Assurance of Basic Undergraduate Education, que pertenece al General Medical Council (cuerpo regulador de la profesión médica). En el año 2009¹⁰ se pasó una encuesta a las 29 escuelas del país sobre los exámenes que aplicaban a lo largo del grado requiriendo datos acerca de: a) la naturaleza, el formato y los contenidos de los exámenes teóricos y pruebas prácticas empleados, así como el uso de portafolios, «audits», etc; b) los momentos de su aplicación a lo largo del grado (anuales, semestrales, de año final, etc); c) la capacitación de los examinadores y su preparación específica para esta tarea; y, d) los criterios de puntuación, y la fijación del estándar de corte. Al tiempo, se requirió la opinión sobre la pertinencia de compartir exámenes entre escuelas y la de aplicar uno o más exámenes de ámbito nacional al final del grado, bien para otorgar el título, o bien para establecer además un ranking de candidatos para acceder al Foundation Programme.

La respuesta a la encuesta fue del 100%. En relación con los exámenes teóricos 23 escuelas usaban test con preguntas del tipo «extended matching question» (de apareamiento múltiple entre un listado de opciones y otro de ítems que son preferentemente viñetas clínicas), que sustituyen cada vez más a las de elección múltiple y respuesta única (MCQ) para evaluar razonamiento y juicio clínico; 16 usaban preguntas del tipo MCQ, 14 preguntas de respuesta corta o del tipo «ensayo», y solo 5 preguntas verdadero-falso. Las variaciones en el número de exámenes por año fue de 1 a 6, y su duración osciló de 1 a 3 horas. Veintidós de las 29 escuelas especificaron el método de fijación del estándar (Angoff o Hofstee).

En cuanto a las pruebas prácticas (exámenes de competencia clínica), 25 de las 29 escuelas usan exámenes clínicos objetivos estructurados (ECOE), con un número de estaciones entre 4 y 44 de una duración entre 4 y 30 minutos cada una; se reportaron variaciones en los dominios de competencias evaluados, en los criterios de paso, en el uso de listados de ítems vs. escalas globales, y en los tipos de observador de la actuación del alumno (el propio paciente vs. el profesor). Ocho escuelas usan pruebas de evaluación objetiva estructurada con enfermo real tipo mini-CEX (mini Clinical Evaluation Exercise), 8 usan el OSLER (Objective Structured Long Examination Record), y otras 8 la observación directa de procedimiento práctico. La mayoría de las escuelas usan portafolio con enfoques sumativos o formativos, y grados variables de «diario de reflexión» sobre la experiencia clínica. Quince escuelas usaron «logbook» de los pacientes examinados en los encuentros clínicos. Un dato a destacar es que todas las escuelas disponen de expertos en educación y más específicamente en evaluación y psicometría.

Las conclusiones del estudio son que a pesar de emplearse un currículum nuclear («core competences»), cada escuela tiene su propio sistema de instrucción y evaluación, y si bien estos son parecidos, no se puede garantizar la equivalencia

en el nivel de competencia exigido en cada una de ellas. Al no existir un registro nacional de los resultados de los exámenes, no se pueden establecer comparaciones cuantitativas de la equivalencia de las notas otorgadas en diferentes escuelas. Sin embargo, existe evidencia de que todas garantizan un nivel aceptable-bueno de competencia, lo que revalida la inclusión del expediente para ordenar el acceso al Foundation Programme.

A pesar de que la mayoría de los responsables docentes defienden la conveniencia de homogeneizar las evaluaciones a nivel nacional persiguiendo una mayor estandarización, y de aplicar exámenes de cualificación final de grado comparables entre escuelas, la opinión generalizada en el Reino Unido es que estos no deben usarse para ordenar el acceso a la especialidad, ni para clasificar la calidad de las diferentes escuelas. Hay unanimidad en mantener la libertad de los currículos porque homogeneizarlos rígidamente comprometería la competitividad entre escuelas y la creatividad e innovación curriculares (ver el Comentario final). Se advierte además que las evaluaciones tipo «disparo único» (como el MIR) provocan desafección del profesorado hacia el proceso de evaluación y empujan a los estudiantes a preparar este examen abandonando otros tipos de aprendizaje y el objetivo de llegar a ser un doctor competente («well-rounded doctor»)¹⁰⁻¹³.

La valoración del Situational Judgement Test. Este test mide una variedad de atributos no académicos especialmente relevantes en la práctica clínica, tales como profesionalismo, capacidad para afrontar la presión, comunicación efectiva con el paciente y el personal docente y sanitario, relación con el paciente, trabajo efectivo en equipo, capacidad de organización, utilización del tiempo, resolución de problemas, toma de decisiones, autocrítica y cuidado personal^{7,9}. No mide «valores éticos» propiamente dichos, sino la capacidad de juicio del estudiante en situaciones del trabajo real diario y la conciencia de hasta qué punto se es competente en el contexto clínico y en las relaciones interpersonales; así, se usa en combinación con tests que miden conocimiento y competencia clínica para obtener una impresión global de la aptitud del estudiante para el trabajo clínico. Utiliza 60 preguntas de las que una es de elección múltiple con 8 opciones y 3 respuestas correctas, y los otros dos tercios son de ordenamiento de opciones, que son construidas por los clínicos de los hospitales donde trabajan los residentes del Foundation Programme asegurando así que los escenarios contemplados en las preguntas son relevantes, realistas y claros.

Diferentes estudios demuestran que el SJT tiene una fiabilidad interna alta (alfa = 0,79-0,85) y una validez de contenido, criterio y constructo buena-alta para estimar atributos profesionales como los mencionados anteriormente y otros como empatía, integridad-honestidad, fiabilidad, entereza-resistencia, tenacidad e iniciativa personal. Estudios de validez a largo plazo han demostrado que el test predice el desempeño durante el aprendizaje de la práctica clínica evaluado por el supervisor docente⁸. Por otra parte, su aceptabilidad («face validity»), y aplicabilidad son también buenas; se puede pasar en papel o en video y corregir por computador, por lo que resulta coste-eficiente (se usa típicamente para grandes grupos de 7000 graduados). Así, este test, que predice aspectos relevantes de la eficiencia en el desempeño clínico, se considera válido para decidir sobre la empleabilidad

de los graduados que acceden al Foundation Programme. Al medir competencias relacionadas con el trabajo clínico, el SJT no puede ser superado por graduados que no hayan estado expuestos a la práctica clínica de manera sostenida. Además, si se diseña bien, la preparación en «academias» es de escasa utilidad.

Tras sus primeras aplicaciones el SJT ha sufrido críticas por parte de los estudiantes que, si bien admiten que estima competencias y características personales no medibles de otro modo, censuran su limitado poder discriminante y alegan que es injusto que pese tanto como el expediente académico; y es que con el actual método de selección un estudiante «malo» (con 34 puntos en el expediente y un buen resultado en el SJT) iguala a otro con el mejor expediente y un SJT medio¹⁴.

Conclusión sobre el método de selección en el Reino Unido. A pesar de existir un «core curriculum» con «competencias nucleares», la diversidad en las metodologías de instrucción (coexisten abordajes curriculares convencionales con otros basados en la resolución de problemas, o en objetivos y competencias) y de evaluación entre escuelas impide establecer una equivalencia estrecha entre las calificaciones otorgadas en unas y otras. No obstante, al expediente, que se estima refleja la formación recibida y el nivel de competencia clínica alcanzado, se le otorga la mitad del peso total, adjudicando la otra mitad al resultado del SJT que mide atributos, habilidades y actitudes profesionales que solo se pueden adquirir con el ejercicio clínico. Este sistema de selección para el acceso al Foundation Programme es poco competitivo en un país donde se cuida la adecuación entre el número de egresados de las escuelas y el de puestos disponibles en este programa, pero permite elegir el lugar para trabajar en la Foundation School que a su vez determina la accesibilidad al área del país de mayor atractivo para hacer la residencia (p.ej. Londres frente a una región periférica) a la que, como ya se apuntó, se accede mediante solicitud y entrevista en el servicio solicitado.

El método de selección de residentes en EE. UU.

La acreditación de las escuelas y la obtención del grado de medicina está regulada por el Liaison Committee for Medical Education, esponsorizado por la Association of American Medical Colleges, y la American Medical Association¹⁵. En el año 2010, y de la mano de la Association of American Medical Colleges, se pasó una encuesta a las 128 escuelas de

EE. UU. y Canadá¹⁶ que, al proporcionar datos sobre el perfil y la operatividad del currículum a nivel nacional, permite conocer de manera indirecta la consistencia de la metodología utilizada para hacer la selección de los residentes en esos países. La encuesta requirió información sobre: 1) el funcionamiento de la Oficina Docente, una estructura transdepartamental que gobierna, sostiene y desarrolla el currículum y verifica el efecto de los cambios curriculares sobre la enseñanza y la formación del profesorado; 2) la promoción y titulación del profesorado; 3) el nivel de implantación de la enseñanza basada en competencias; 4) la implantación de nuevas corrientes y paradigmas pedagógicos (reducción de clases, clases interactivas, instrucción basada en casos o problemas, «work-based» y «on line learning», simulación, uso de enfermos estandarizados, trabajo en equipo, etc); 5) mejoras en evaluación del tipo ECOE, práctica reflexiva y portafolios; y

6) mejoras en la experiencia clínica (inmersión clínica precoz, expansión hacia hospitales y comunidad). La aportación de un dossier detallado de la organización curricular en relación con los anteriores ítems permitió comprobar que todas las escuelas habían renovado ya, o lo estaban haciendo, sus métodos de instrucción y evaluación, lo que ha resultado en una corriente general de renovación curricular cara al siglo XXI similar a la producida por la revolución flexneriana del año 1910.

Para seleccionar a los residentes en los EE. UU. se utiliza el llamado «National Resident Matching Program» («Electronic Residency Application Service» con oficina central en San Francisco) que para las peticiones que los solicitantes envían a los servicios de su preferencia con los criterios de aceptación internos de los propios servicios que seleccionan a algunos solicitantes para ser entrevistados. Los criterios usados para hacer la selección son: a) el expediente académico; b) el rendimiento en los dos test de conocimientos en ciencias básicas y clínicas que son los «steps» 1 y 2 de conocimiento clínico del United States Medical Licensing Examination (USMLE)¹⁷; y c) una entrevista semiestructurada. Sobre la base de estos criterios, el comité de selección del servicio elabora una lista que ordena a los candidatos («Rank Order List»). Veamos la influencia y las limitaciones de cada uno de los criterios empleados para la selección.

La valoración y el peso del expediente académico. Los directores de programas y los comités locales de selección requieren información sobre el desempeño del candidato durante el grado y valoran dos documentos relacionados que son la transcripción de las notas (el expediente convencional) y el llamado «Medical Student Performance Evaluation» modificado en el 2002 por la Association of American Medical Colleges, que detalla el rendimiento comparativo del estudiante a lo largo del grado, tanto en conocimientos (ciencias básicas y clínicas), como en su desempeño en las rotaciones clínicas (Medicina, Cirugía, Pediatría, Psiquiatría, Medicina de Familia), y sus atributos profesionales (profesionalismo, trato empático y compasivo al paciente, honestidad e integridad, respeto por los demás, y anteponer las necesidades del paciente a las propias). Además, se exigen tres cartas de recomendación, un «personal statement» y el currículum vitae, del que pueden valorarse experiencia en investigación, publicaciones y póster, premios, congresos y otros, becas obtenidas, participación en ONG, y otros.

El peso relativo concedido a los componentes del expediente académico varía entre servicios, y es que a pesar de existir un «core curriculum» y una notable homogeneidad en los métodos de instrucción a nivel nacional, hay una considerable dispersión en las calificaciones otorgadas en las diferentes escuelas, que además pueden estar «infladas» por el sesgo de los profesores a exaltar las virtudes de sus estudiantes y el atractivo de su escuela. En una encuesta específica pasada a 119 escuelas de EE. UU.-Canadá en el año 2009¹⁸ se recabó información sobre los procedimientos de evaluación del desempeño clínico de los estudiantes en las rotaciones clínicas, observándose una notable dispersión en las evaluaciones tanto inter- como intraescuela; así, solo 6 escuelas usaron un sistema de dos grados o escalones (pasa-no pasa); 15 usaron 3 grados (pasa-no pasa y «honors»), 63 cuatro grados, y 35 usaron cinco o más grados; el porcentaje de estudiantes que obtuvieron el grado más alto en diferentes escuelas osciló

entre 10 al 77% en las rotaciones por Medicina Interna, 7 al 67% en las de Cirugía, 8 al 60% en Pediatría, 11 al 81% en Psiquiatría, 2 al 68% en Ginecología y del 10 al 87% en Medicina de Familia.

Esta variabilidad ha llevado a proponer la reducción de las opciones a pasa-no pasa y poco más (añadiendo quizás uno o dos grados de «honors»), y homogeneizar los criterios a nivel nacional en lo posible. Sin embargo, muchos expertos rehúsan reducir a tan solo pasa/no pasa las calificaciones en las rotaciones clínicas nucleares porque ello impediría hacer comparaciones entre estudiantes. Así, una opción es usar «pasa-no pasa» otorgando la calificación de «bueno-adequado» si se alcanzan los estándares, y añadir «atributos únicos» o escalas de evaluación secundarias para reconocer algo que sea especial en el desempeño y cualificar a los mejores para una «honors evaluation».

El peso y la utilidad del rendimiento en el examen USMLE. El examen USMLE, que según regula el Liaison Committee for Medical Education ha de ser superado para obtener el grado en EE. UU. tiene cuatro componentes: 1) los steps 1 y 2 de conocimiento clínico («clinical knowledge»), ya mencionados, que emplean test teóricos con preguntas construidas para medir contextualización clínica de los conocimientos en ciencias básicas y clínicas, respectivamente; 2) el step 2 de habilidades clínicas («clinical skills») diseñado para medir competencia clínica (diagnóstico y manejo clínico básicos y habilidades de comunicación), que consiste en un ECOE multiestación, con casos balanceados en todos los parámetros (agudeza-cronicidad, edades, sexo, «mix» de patologías: problemas comunes en cardiovascular, digestivo, genitourinario, respiratorio, neurológico, psiquiátrico, ginecológico); y 3) el step 3, que evalúa manejo inicial y a largo plazo de los problemas clínicos más comunes en pacientes ambulatorios, e incluye simulaciones de pacientes que se pasan por ordenador^{17,19}.

Globalmente el USMLE utiliza 1200 preguntas, encuentros clínicos con 12 enfermos estandarizados, exige la resolución de 20 casos clínicos completos, y tiene la duración de cinco días (8 horas por jornada). La validez, reproducibilidad y transparencia de este examen son buenas-altas y hay que enfatizar que incluso las preguntas del step 1, relacionadas con las enseñanzas de las ciencias básicas, requieren contextualización clínica de los conocimientos evaluados lo cual se corresponde con el currículum utilizado en los EE. UU. que, en comparación con el nuestro, está mejor integrado verticalmente y contempla aspectos diagnósticos y patológicos ya en los dos primeros cursos del grado. También hay que destacar la excelente calidad en la construcción de las preguntas que contrasta con las del examen MIR.

La gran mayoría de los servicios incluyen el resultado de los steps 1 y 2 de conocimiento clínico en la Rank Order List, si bien en ocasiones el resultado del step 2 (que se pasa al final del grado) no está disponible en el momento de la valoración del candidato. Un estudio reciente encontró una buena correlación entre los resultados de los steps 1-2 de conocimiento clínico y el rendimiento de los estudiantes, tanto en los dos cursos básicos del grado (medido con la nota media obtenida en las asignaturas de 1.º y 2.º cursos), como en el desempeño durante las rotaciones en los cursos

clínicos (medido con el instrumento llamado «subject examination performance» que emplea el National Board of Medical Examiners)²⁰; la conclusión del estudio es que los steps 1-2 de conocimiento clínico sirven para medir la capacidad para contextualizar el conocimiento sobre la realidad clínica y otros aspectos de la competencia clínica. Ya otro análisis anterior había mostrado igualmente una buena correlación entre el examen de licenciatura y el desempeño como médico de Atención Primaria²¹.

Sin embargo, en otro estudio los resultados de los steps 1-2 de conocimiento clínico no se correlacionaron ni con los resultados obtenidos por los estudiantes en el step 2 de habilidades clínicas, ni con su desempeño posterior en la etapa de residencia o durante el ejercicio como profesional autónomo²², concluyendo que los steps 1-2 de conocimiento clínico no deben ser utilizados para ordenar a los residentes, ni para clasificar la calidad de las escuelas de medicina. A pesar de estos resultados contradictorios sobre su validez para medir el nivel de competencia clínica, y de que no fueron diseñados para «ordenar» el acceso a la residencia, sino para otorgar el grado, los steps 1-2 de conocimiento clínico se usan como criterios en 2.º y 5.º lugar para hacer esta selección en la mayoría de los servicios del país.

Sería de interés correlacionar los resultados de los steps 2 de habilidades clínicas y el step 3 con el desempeño competencial durante la residencia y la práctica autónoma, o dicho de otro modo, verificar si el poder predictivo de estos dos exámenes de competencia clínica supera al de los test de conocimiento (steps 1 y 2 de conocimiento clínico). Diferentes estudios han encontrado correlación entre el nivel de competencia clínica registrado en las rotaciones clínicas del grado y el desempeño clínico durante la residencia y la actividad autónoma^{23,24}. El problema para incluir el resultado en el step 2 de habilidades clínicas como criterio en el listado de acceso a la residencia es que se pasa en diferentes ciudades y momentos por lo que el examen no es exactamente el mismo para todos los candidatos; si bien las presentaciones de la prueba están estandarizadas al máximo (misma planilla de datos y mismo número de candidatos clínicos), las tasas de varianza no atribuibles al candidato, como son las relacionadas con los contenidos (distintas estacionamientos) y los observadores (distintos desahogos; es el enfermo quien rellena los listados de ítems) desaconsejan su uso como criterio para hacer la selección a los puestos de residencia.

Independientemente de su utilidad para evaluar el nivel de competencia clínica al final del grado, hay que señalar la influencia favorable del step 2 de habilidades clínicas sobre el funcionamiento curricular en las escuelas de los EE. UU.²⁵. La implantación de este examen, junto con la administración longitudinal de ECOE a lo largo del grado (del mismo modo que se hace en el Reino Unido), ha asegurado la adquisición de un nivel aceptable de competencia clínica de los graduados y ha relanzado la remodelación curricular en todas las escuelas que mejoran y unifican sus mapas de competencias, y ahora prestan más atención al entrenamiento clínico en las rotaciones¹⁶. Al tiempo, se ha evidenciado que disminuye la dispersión de las calificaciones, y ayuda a corregir defectos curriculares de base, como son la escasa observación y retroalimentación del alumno por parte de los profesores y el apoyo

excesivo de estos en los exámenes teóricos, corroborando una vez más que la evaluación dirige el aprendizaje y mejora la enseñanza («assessment drives learning» and «assessment enhances teaching»).

La valoración en la entrevista por el staff del Servicio. Los sesgos de la entrevista han llevado a proponer recientemente un sofisticado algoritmo para mejorar su fiabilidad²⁶. Entre las limitaciones de la entrevista al uso está que en general no todos los miembros del comité de selección entrevistan a todos los candidatos, y un entrevistador «halcón» puede bajar hasta 30 puestos a un candidato en una lista de 150, ocurriendo lo mismo pero en sentido contrario con el entrevistador «paloma». Pero aparte de la variabilidad entre evaluadores, hay otros sesgos, bien de tipo intermitente que se da cuando un mismo evaluador tiene «un día malo vs. un día bueno», o de tipo selectivo; p.ej. el evaluador valora sistemáticamente peor a las mujeres, o los candidatos de color. Los comités de selección dedican mucho tiempo y esfuerzo a conseguir fiabilidad en la selección, pero no siempre logran acertar, por lo que parece imprescindible refinar el método, que podría llegar a utilizarse en otros países donde se recurre ya a la entrevista, o podría utilizarse en el futuro.

Conclusiones sobre los criterios de selección de los residentes en los EE. UU. y sus limitaciones. La «Rank Order List» es un predictor débil-moderado del desempeño clínico real durante la Residencia²⁷; así, la posición alta de un candidato en ella debe tomarse tan solo como la probabilidad de actuar de manera competente, pero con un intervalo de confianza muy amplio. Las limitaciones de la fiabilidad de este instrumento multicriterio pueden relacionarse con: 1) las habilidades y competencias que definen un estudiante excelente no necesariamente se transfieren a su desempeño durante la residencia y tampoco se transmiten de un dominio a otro; p.ej. un estudiante excelente en anatomía no necesariamente llega a ser un buen cirujano; 2) las calificaciones otorgadas durante el Grado sobre conocimientos y competencia clínica siguen siendo variables entre escuelas y no permiten comparar con suficiente fiabilidad los rendimientos de los estudiantes en unas y otras; sin estandarizar el proceso está siendo perfeccionado; 3) los test estandarizados de conocimiento del USMLE (steps 1-2 de conocimiento clínico), diseñados para certificar la graduación, y usados en muchos servicios para seleccionar al residente no necesariamente predicen el desempeño de este durante la residencia; 4) las entrevistas adolecen de diferentes sesgos que limitan su validez para predecir el desempeño profesional; los staffs del comité de selección manejan criterios diferentes y se ha comprobado que si se ciega a los entrevistadores de la información previa sobre el candidato, el valor de la entrevista aporta poco. Además, los entrevistados «típicamente» saben qué y cómo contestar para causar una impresión positiva, y así, lo que dicen y cómo se comportan o «actúan» no se corresponde siempre con lo que después desarrollan como residentes. Por otra parte, ya se ha difundido el «coaching» para «rendir» al máximo en la entrevista y los estudiantes empiezan incluso a deformar positivamente su imagen en las redes sociales para mejorar su paso por la entrevista. En cualquier caso, el empeño de las instituciones académicas de los EE. UU. para superar estas limitaciones es ejemplar y debe ser seguido con atención.

Comentario final

Al comparar los criterios para la selección de los residentes hemos visto que en muchos países se rechaza el uso de los exámenes de ámbito nacional del tipo «disparo único» que, como el MIR, no miden la formación o competencia clínica de los candidatos, y se desaconseja también la utilización del expediente académico a causa de la variabilidad en las calificaciones otorgadas durante el grado en las diferentes escuelas; en estos países se revisa el formulario de solicitud y se recurre a la entrevista personal no estructurada para elegir al residente. En el bando contrario están los países que, considerando la entrevista un criterio inevitablemente arbitrario, se apoyan solo, o fundamentalmente, en test de conocimientos, por lo que reconocen que se trata de una medición parcial del mérito del candidato. Finalmente, hay otros países como el Reino Unido y los EE. UU. que aplican metodologías multicriterio que al parecer nos las más consistentes han sido utilizadas para diseñar otra que mejore la empleada actualmente en España, que además de resultar inadecuada, ejerce una influencia negativa sobre el funcionamiento curricular en el pregrado.

El análisis crítico del método de selección utilizado en nuestro país y la descripción del que podría sustituirlo se hacen en otro artículo que aparecerá en el próximo número de esta revista. En breve, el nuevo método propuesto incluye los siguientes criterios: 1) la revalorización del expediente académico que exigiría homogeneizar los criterios de evaluación utilizados en las diferentes facultades (20% del peso); 2) dos exámenes de conocimiento similares a los steps 1-2 de conocimiento clínico del USMLE, que podrían ser administrados en el 3.º curso para medir conocimientos de ciencias básicas, y justo antes de finalizar el 6.º curso, para medir conocimientos clínicos, eliminando así el año de preparación de la «oposición MIR» (60% del peso); y 3) un test de juicio situacional similar al administrado en el Reino Unido que mide atributos profesionales decisivos para la buena práctica clínica no evaluables con test convencionales (20% del peso).

De los EE. UU. y otros países del norte y centro de Europa podría tomarse la opción de la entrevista, y considerarse también la posibilidad de utilizar la nota obtenida en un ECOE de ámbito nacional que, de aceptarse, no debería pesar más de un 5%. La introducción de los ECOE en las facultades españolas se produce lenta, pero progresivamente, y está en el horizonte la implantación de un ECOE de grado similar al step 2 de habilidades clínicas del USMLE, que podría llegar a aplicarse a nivel nacional tal como contempló la Ley de Ordenación de las Profesiones Sanitarias (LOPS, Art 21) en el año 2011.

Al diseñar el método para ordenar el acceso a la residencia se ha de perseguir la máxima eficiencia en la selección y ha de tenerse en cuenta la influencia del mismo sobre el funcionamiento curricular que será favorable o adversa dependiendo del perfil del propio método. La utilización de uno o más exámenes de ámbito nacional que, como el MIR, prima la adquisición de conocimientos sobre la práctica clínica puede llegar a arruinar la estrategia docente¹¹. Por el contrario, el uso de criterios que impulsen el aprendizaje y la evaluación en el contexto clínico propiciarán la mejor formación práctica del estudiante y una mayor dedicación del profesorado a la

enseñaza en el contexto clínico²⁸⁻³⁰. Por eso urge modificar el actual MIR mejorando no solo las características del test o los test de conocimiento a emplear, sino añadiendo criterios adicionales como las ya mencionados. Pero antes de afrontar la sustitución del MIR por otro examen de ámbito nacional conviene considerar la pertinencia del uso de este tipo de exámenes que son rechazados en la mayoría de los países con desarrollos curriculares más modernos, con la excepción de los EE. UU. A continuación se exponen los argumentos de expertos en docencia defensores y detractores de este tipo de exámenes.

La pertinencia de los exámenes de grado de ámbito nacional para otorgar el grado y/o la selección de residentes. Los defensores argumentan que la utilización de estos exámenes favorece la estandarización de la enseñanza y elimina las diferencias mayores en las evaluaciones entre las escuelas del mismo país, facilitando el establecimiento del currículum nuclear («core curriculum» con «core competences»), y la homogeneidad en la formación de los egresados en todo el país^{12,13,31}. Por otra parte, el examen nacional ahorra recursos al evitar la aplicación de diferentes tests en cada escuela, y de implantarse en los diferentes países europeos facilitaría la comparación entre graduados, la armonización de títulos y la libre circulación de titulados.

Los argumentos en contra se aportan sobre todo por expertos en docencia del Reino Unido y otros países del norte y centro Europa^{11,32}, donde a pesar de que las escuelas del mismo país pueden tener métodos de instrucción diversos, todas aseguran un buen nivel de competencia clínica al final del grado que no necesita ser demostrado con este tipo de examen; por poner un ejemplo, solo 3 de 5833 graduados en el Reino Unido en el 2005 estuvieron por debajo de dicho nivel. Pero quizás, el argumento de mayor peso en contra del examen nacional es que este compromete la diversidad de los currículos y elimina la motivación para innovarlos, una tarea que se considera irrenunciable en las 29 escuelas del Reino Unido, donde encuestas a miles de graduados demuestran que las innovaciones contribuyeron a mejorar su formación.

Se asegura que de existir un examen de este tipo las escuelas perseguirían no ya la innovación y mejora curricular, sino cómo preparar mejor a sus estudiantes para superarlo, y se enfatiza que el objetivo de la enseñanza del pregrado es formar buenos médicos, y no preparar a los estudiantes para superar exámenes de conocimiento como los que se pasan en algunos países europeos. Se considera que, a menos que el sistema de ordenación para el acceso a la residencia sea válido y fiable, no debe usarse, y no hay actualmente ninguno que cumpla estos criterios. Finalmente, se aduce que la selección para una especialidad debe basarse en las habilidades y aptitudes para ejercerla, y no sobre la base de unos conocimientos difusos de todas las especialidades demostrados en un momento dado. La British Medical Association advierte contra la influencia negativa de los exámenes del tipo «disparo único» sobre el aprendizaje y propone recurrir a la evaluación multimodal para ordenar a los estudiantes que acceden al Foundation Programme. En todo caso la selección para la residencia no debe hacerse solo sobre la base de un test de conocimiento³³.

La pertinencia de los exámenes finales de grado paneuropeo como refuerzo del proceso de convergencia y la comparabilidad de títulos. En un encuentro mantenido en Praga, grupos de expertos en

evaluación del aprendizaje clínico debatieron la pertinencia de aplicar un examen final de grado en Europa, que podría ser similar al USMLE y que, al menos en principio, podría contribuir a la armonización de los currículos, la homologación de títulos y la movilidad de estudiantes y profesionales en los países de la Unión³.

Los defensores de la implantación de exámenes de ámbito transnacional argumentaron que estos ayudarían a: 1) homogenizar el nivel de competencia de los egresados desde las diferentes escuelas de medicina europeas; 2) mejorar la convergencia de los currículos, la comparabilidad de los títulos y la movilidad de estudiantes y profesores; 3) mejorar la asistencia al paciente en el contexto de la libre circulación de profesionales; 4) garantizar el control de calidad de los estándares de las cada vez más numerosas escuelas privadas de medicina; y 5) controlar el nivel de competencia y las habilidades psicosociales de graduados no europeos que acceden a la práctica en nuestros sistemas de salud.

Además, un examen paneuropeo favorecería la aplicación de auditorias externas que reforzarían el cumplimiento de los estándares docentes y la transparencia y fiabilidad del propio examen, e impulsaría la innovación curricular al evidenciar el rendimiento de los estudiantes y las deficiencias en su formación, propiciando la mejora progresiva del propio proceso de evaluación. Finalmente, la agregación de los esfuerzos de todas las escuelas y facultades en la configuración y aplicación de esa prueba única, disminuiría los costes de los exámenes múltiples de ámbito nacional, y permitiría aprovechar mejor las prestaciones de los grupos de expertos en evaluación que ahora trabajan dispersos en Europa. La implantación de un examen de ámbito internacional («International Licensing Examination») facilitaría el intercambio y la circulación de graduados no solo dentro de Europa, sino entre Europa, Norteamérica y el resto de los países. Sin embargo, otros expertos igualmente acreditados, o incluso con mayor prestigio, advierten que aparte de las dificultades de tipo logístico y lingüístico que tendría la realización de un examen paneuropeo, existen otros argumentos en su contra, que vienen a ser los mismos aducidos en contra de los exámenes de ámbito nacional.

Así, van der Vleuten³², y Harden¹¹, se preguntan si este tipo de examen realmente en realidad se impediría para emprender cualquier cambio curricular en vez de favorecerlo y advierten que se está produciendo un desplazamiento de la cultura de pasar test (evaluación sumativa para determinar si el candidato pasa/no pasa), a otra de evaluar para formar (evaluación formativa); la primera se aplica en momentos fijos a lo largo del grado, en tanto que la segunda se engrana con el propio aprendizaje en el contexto clínico real, al cual propulsa continuamente. Así, el «assessment of learning» está dejando paso al «assessment for learning» (evaluación con retroalimentación inmediata), y está claro que los exámenes de grado de tipo nacional alejan la evaluación del propio proceso del aprendizaje. De acuerdo con estos autores, los exámenes finales de grado solo permiten evaluar habilidades y competencias fáciles de medir («low profile tasks»), pero no las complejas, con lo cual su validez es tan solo limitada, independientemente de cuáles sean su practicabilidad y costes. Esta clase de exámenes acarrearían una armonización reduccionista de currículos de eficacia ya probada, como los del Reino Unido y

otros países del centro y norte de Europa, y comprometerían la diversidad de programas docentes preconizada por Bolonia.

Como conclusión sobre esta controversia parece claro que el dilema está entre utilizar exámenes centralizados vs. descentralizados, por un lado, y test «disparo único» vs. evaluaciones continuas durante el grado, por el otro. Ya hemos visto que mientras en los

EE. UU. los exámenes de grado están centralizados, en el Reino Unido la evaluación del estudiante está descentralizada en las escuelas que usan examinadores externos. En relación con el tipo de prueba, muchos exámenes de grado usan uno o más test de conocimientos tipo «disparo único» para cualificar al candidato que no será examinado nunca más sobre los mismos contenidos. Como apunta Schuwirth³⁴, tomar la TA del paciente, conocer la anatomía de hígado, o calcular la t de Student, deben conocerse en un momento dado del grado y no necesitan ser mejorados para completar la formación del estudiante, por lo que pueden ser evaluados con un examen de este tipo. Sin embargo, en términos de competencia clínica la estrategia de «disparo único» implica un desalineamiento entre evaluación y enseñanza que no permite medir los parámetros que definen al graduado competente, los cuales han de ser evaluados en diferentes momentos del proceso formativo; y es que la progresión continua en el aprendizaje es inherente al concepto mismo de competencia que ha arraigado ya fuertemente en los currículos modernos.

Toda esta controversia parece ignorarse en nuestro país, donde la existencia del MIR impulsa a los estudiantes a supeditar la adquisición de la competencia clínica al estudio intensivo encaminado a adquirir más y más conocimientos, pero el análisis crítico de las limitaciones de este examen, cómo podría ser modificado, y cómo habría que mejorar las enseñanzas impartidas en nuestras facultades, son objetivos del artículo concatenado con este que se publicará en el próximo número de esta revista.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

BIBLIOGRAFÍA

- Aspa Marco FJ, Rodríguez de Castro F. Evaluación final: ¿sirve el examen MIR? *Educ Med.* 2010;13 Suppl 1:73S-9S.
- Gómez JM, Pujol R, Martínez-Carretero JM, Blay C, Jaurrieta E, Friedman M, et al. El proyecto COMBELL. Un análisis de la competencia clínica médica. *Med Clin.* 1995;105:649-51.
- Lobato RD, Lagares A, Alén JF, Alday R. El desarrollo del proceso de Bolonia y el Grado de Medicina. Situación actual y expectativas para su implantación definitiva. *Neurocirugía.* 2010;21:146-56.
- Saz Perez JV. Bolonia: una oportunidad perdida. La heterogeneidad de los estudios de medicina en España. *Rev Clin Esp.* 2013;213:440-1.
- Vázquez G, Murillo-Cabezas F, Gómez J, Martín C, Chaves J, Peinado JL. El examen MIR, su cambio como una opción estratégica. *Educ Med.* 2008;11:203-6.
- Lorusso N, López-Valcarcel BG. Variabilidad de los procesos selectivos para el acceso a las especialidades médicas en Europa: entre convergencia y divergencia. *Educ Med.* 2013;16:215-8.
- Patterson F, Ashworth V, Zibarras L, Coan P, Kerrin MO, Neill P. Evaluations of situational judgement tests to assess non-academic attributes in selection. *Form Educ.* 2012;46:850-68.
- Patterson F, Carr V, Zibarras L, Burr B, Berkin I, Plint S. New machine-marked tests for selection into core medical training: evidence from two validation studies. *Clin Med.* 2009;9:417-20.
- Rahman M. Succeeding in a Situational Judgement Test. *Student BMJ.* 2012;20:e6953.
- McCrorie P, Boursicot KAM. Variations in medical school graduating examinations in the United Kingdom: Are clinical competence standards comparable? *Medical Teacher.* 2009;31:223-9.
- Harden R. Five myths and the case against a European or national licensing examination. *Medical Teacher.* 2009;31:217-20.
- Hum R, Cooney L. One exam to end them all? *Student BMA News,* June 2014. bma.org.uk/news-views-analysis.
- Ricketts C, Archer J, Noble I. Are national qualifying examinations a fair way to rank medical students? *BMJ.* 2008;337, a1279.
- Weinberg I. After the situational judgment test. *Student BMJ.* 2014;21:f1901.
- Melnick DE. Licensing examination in North America; is external audit valuable? *Medical Teacher.* 2009;31:212-4.
- Anderson MB, Kanter SL. Medical education in the United States and Canada, 2010. *Acad Med.* 2010;85 Supl 9:2S-18S.
- United States Medical Licensing Examination. (USMLE). 2007. USMLE website: Canada Medical Council [consultado Dic 2013]. Disponible en: <http://www.mcc.ca>
- Alexander EK, Osman NY, Walling JL, Mitchell VG. Variation and imprecision of clerkship grading in U.S. medical schools. *Acad Med.* 2012;87:1070-6.
- Champlain A, Swygert K, Swanson DB, Boulet JR. Assessing the underlying structure of the United States Medical Licensing Examination Step2 test of clinical skills using confirmatory factor analysis. *Acad Med.* 2006;81 10 Suppl:17S-20S.
- Zahn CM, Saguil A, Artino AR, Dong T, Ming G, Servey JT, et al. Correlations of national Board examiners scores with the USMLE steps 1 and 2 scores. *Acad Med.* 2012;87:1348-54.
- Tamblyn R, Abrahamowicz M, Dauphinee WD, Wenghofer E, Jackes A. Association between licensure examination scores and practice in primary care. *JAMA.* 2002;288:3019-26.
- McGaghie WC, Cohen E, Wayne D. Are United States Medical Licensing Exam Step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Acad Med.* 2011;86:48-52.
- Papadakis MA, Teherani A, Banach MA, Knettlar TR, Rattner SL, Stern DT. Disciplinary action by medical boards and prior behaviour in medical school. *N Engl J Med.* 2005;353:2673-82.
- Boulet JR, McKinley DW, Whelan GP. Clinical skills deficiencies among first-year residents: Utility of ECFMG clinical skills assessment. *Acad Med.* 2002;77:33S-5S.
- Hauer KE, Teherani A, Kerr KM, O'Sullivan PS, Irby DM. Impact of the United States Medical Licensing Examination Step 2 clinical skills exam on medical school clinical skills assessment. *Acad Med.* 2006;81 10 Suppl:13S-6S.
- Ross DA, Moore EZ. A quantitative experimental paradigm to optimize construction of rank order lists in the National Resident Matching Program. The ROSS-MOORE approach. *Acad Med.* 2013;88:1281-6.
- Baker K. The tip of the iceberg: improving the quality of rank Order Lists for the Match. *Acad Med.* 2013;88:1281-6.

