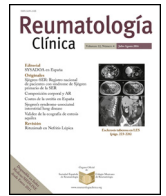




Sociedad Española
de Reumatología -
Colegio Mexicano
de Reumatología

Reumatología Clínica

www.reumatologiaclinica.org



Original

El método para establecer el punto de corte en el examen clínico objetivo estructurado define el desempeño de los candidatos a la certificación como reumatólogo

Virginia Pascual-Ramos*, Ana Guilaisne Bernard-Medina, Diana Elsa Flores-Alvarado, Margarita Portela-Hernández, María del Rocío Maldonado-Velázquez, Luis Javier Jara-Quezada, Luis Manuel Amezcua-Guerra, Nadina E. Rubio-Judith López-Zepeda, Everardo Álvarez-Hernandez, Miguel Ángel Saavedra, César Alejandro Arce-Salinas y Consejo Mexicano de Reumatología (CMR)

Consejo Mexicano de Reumatología, Ciudad de México, México

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 11 de julio de 2016
Aceptado el 25 de noviembre de 2016
On-line el xxx

Palabras clave:

Competencias basadas en educación
Certificación

R E S U M E N

Antecedentes: El Consejo Mexicano de Reumatología certifica candidatos mediante una evaluación teórica (ET) y un examen clínico objetivo estructurado (ECOFE). En 2013 y 2014, el punto de corte para acreditar el ECOFE se estableció por criterio (PC ≥ 6); a partir del 2015, también se estableció por el método del desempeño limítrofe (PDL). Se comparó el desempeño de los candidatos con ambos puntos de corte y examinó su impacto en la correlación entre la ET y el ECOFE.

Material y métodos: En 2015 y 2016, respectivamente, 43 y 37 candidatos aplicaron; ambos ECOFE se integraron con 15 estaciones; un evaluador por estación calificó la lista de cotejo y el desempeño global del candidato mediante una escala de Likert (inferior, limítrofe y superior). A cada candidato se le asignó una calificación global del ECOFE.

Resultados: El promedio (\pm DE) de la ET fue de 6,6 ($\pm 0,6$) en el 2015 y de 6,4 ($\pm 0,6$) en el 2016; 5 candidatos/año no acreditaron. El promedio (\pm DE) del ECOFE fue de 7,4 ($\pm 0,6$) y 7,3 ($\pm 0,6$), respectivamente; todos acreditaron de acuerdo con el PC, mientras que 21 (49%) y 19 (51%) no lo hicieron de acuerdo con el PDL (7,3 en 2015 y 7,4 en 2016). Los PDL para cada estación variaron.

La ET correlacionó con el ECOFE. Los candidatos con ECOFE acreditado (por PDL) calificaron mejor en la ET que su contraparte.

Conclusiones: El método para establecer el punto de corte del ECOFE afecta al desempeño de los candidatos a una certificación, pero no impacta a la correlación entre la ET y el ECOFE.

© 2016 Elsevier España, S.L.U. y Sociedad Española de Reumatología y Colegio Mexicano de Reumatología. Todos los derechos reservados.

The method used to set the pass mark in an objective structured clinical examination defines the performance of candidates for certification as rheumatologists

A B S T R A C T

Background: The Mexican Accreditation Council for Rheumatology certifies trainees (TR) on an annual basis using both a multiple-choice question (MCQ) test and an objective structured clinical examination (OSCE). For 2013 and 2014, the OSCE pass mark (PM) was set by criterion referencing as ≥ 6 (CPM), whereas overall rating of borderline performance method (BPM) was added for 2015 and 2016 accreditations. We compared OSCE TR performance according to CPM and BPM, and examined whether correlations between MCQ and OSCE were affected by PM.

Keywords:

Education-based competence
Accreditation

* Autor para correspondencia.
Correo electrónico: virtichu@gmail.com (V. Pascual-Ramos).

Methods: Forty-three (2015) and 37 (2016) candidates underwent both tests. Altogether, OSCE were integrated by 15 validated stations; one evaluator per station scored TR performance according to a station-tailored check-list and a Likert scale (fail, borderline, above range) of overall performance. A composite OSCE score was derived for each candidate. Appropriate statistics were used.

Results: Mean (\pm standard deviation [SD]) MCQ test scores were 6.6 ± 0.6 (2015) and 6.4 ± 0.6 (2016) with 5 candidates receiving a failing score each year. Mean (\pm SD) OSCE scores were 7.4 ± 0.6 (2015) and 7.3 ± 0.6 (2016); no candidate received a failing CPM score in either 2015 or 2016 OSCE, although 21 (49%) and 19 (51%) TR, respectively, received a failing BPM score (calculated as 7.3 and 7.4, respectively). Stations for BPM ranged from 4.5 to 9.5; overall, candidates showed better performance in CPM.

In all, MCQ correlated with composite OSCE, $r=0.67$ (2015) and $r=0.53$ (2016); $P \leq .001$. Trainees with a passing BPM score in OSCE had higher MCQ scores than those with a failing score.

Conclusions: Overall, OSCE-PM selection impacted candidates' performance but had a limited affect on correlation between clinical and practical examinations.

© 2016 Elsevier España, S.L.U. and Sociedad Española de Reumatología y Colegio Mexicano de Reumatología. All rights reserved.

Introducción

Los consejos que certifican especialistas de la salud se comprometen con la sociedad, para garantizarle que los médicos certificados cuentan con las habilidades clínicas y los conocimientos necesarios para ejercer su profesión. Como parte de ese compromiso desarrollan y aplican instrumentos de evaluación, establecen puntos de corte y finalmente, elaboran decisiones cruciales, no solamente para los implicados sino también para la comunidad médica y la sociedad. Es por ello por lo que la selección del punto de corte debe ser sólida desde el punto de vista metodológico, pues este varía de manera importante en función del método aplicado para su cálculo¹.

El Consejo Mexicano de Reumatología (CMR) desde hace años y con periodicidad anual, certifica a todos los residentes quienes han cumplido su formación como reumatólogos; la certificación se lleva a cabo al final de su periodo de entrenamiento y consiste en un examen escrito de opción múltiple que evalúa conocimientos teóricos (ET) y un examen de habilidades clínicas. Éste, tradicionalmente, se basaba en el caso clínico único; debido a las limitaciones inherentes a dicha modalidad de evaluación^{2,3} desde el año 2013 se implementó y posteriormente aplicó un examen clínico objetivo estructurado (ECO) como parte del proceso de certificación⁴.

Durante la certificación de los años 2013 y 2014, el punto de corte del ECOE para su acreditación se estableció «por criterio» en cuando menos 6 (en una escala del 0 al 10, donde 0 implicaba el peor desempeño); para ello se discutió y acordó en una sesión *ex profeso*, la mínima cantidad de elementos de una lista de cotejo que debían ser correctos para así considerar las habilidades del reumatólogo como adecuadas y seguras. A partir del año 2015, los evaluadores asignados a cada estación del ECOE fueron instruidos y entrenados para llenar la lista de cotejo y además evaluar el desempeño general del candidato mediante una escala de Likert, con la finalidad de determinar el punto de corte del ECOE por el método del desempeño global límite⁵ (PDL).

Los objetivos del ejercicio descrito fueron:

1. - Comparar el desempeño en el ECOE de los candidatos a la certificación como reumatólogo, usando puntos de corte establecidos por 2 métodos diferentes, por criterio (PC) y por el método del PDL.
2. - Examinar si la correlación entre la ET y el ECOE se afectaba por el punto de corte seleccionado.

Material y métodos

En nuestro país contamos con 16 centros aprobados para la formación de especialistas en reumatología. En los años 2015 y 2016, 43 y 37 candidatos, respectivamente, aplicaron a la certificación en Reumatología en 2 días consecutivos; todos los candidatos habían completado un programa de entrenamiento en sus respectivos centros de formación de cuando menos 4 años y habían sido recomendados por sus profesores.

Ambas versiones de la ET contaron con 222 y 200 preguntas respectivamente, integradas básicamente en un formato de casos clínicos y habían sido elaboradas por reumatólogos certificados, con experiencia y revisadas por un sub-comité del Consejo formado por 4 reumatólogos certificados. Ambos circuitos del ECOE constaron de 15 estaciones diseñadas por miembros del Consejo; cada estación fue validada por un subcomité de cuando menos 6 reumatólogos certificados, ajenos al diseño de las estaciones, toda vez que cuando menos un 80% de los validadores aprobaran la inclusión de cada rubro en la lista de cotejo; cada estación tuvo una duración de 8 minutos y los circuitos incluyeron 4 estaciones de descanso; se asignó un evaluador externo a cada estación, debidamente entrenado para calificar la lista de cotejo y la escala de Likert de desempeño global del candidato (desempeño inferior, límite y superior). Las listas de cotejo de cada estación incluyeron un número de rubros que varió de 5 a 21.

En ambos años se establecieron los siguientes puntos de corte; para el proceso de certificación del año 2015 como $\geq 5,7$ para la ET y ≥ 6 para el ECOE; para el proceso de certificación del año 2016 como ≥ 6 para ambas evaluaciones, ET y ECOE. En todos los casos la máxima puntuación posible fue de 10.

El PDL se calculó de la siguiente manera: para cada estación se seleccionaron las listas de cotejo de los candidatos calificados como «límite» en la escala de Likert del desempeño global; después, se determinó el promedio de la puntuación obtenida en dichas listas de cotejo, lo cual constituyó el PDL para la estación evaluada. Finalmente se obtuvo el PDL para todo el ECOE promediando los PDL de las 15 estaciones.

A cada candidato se le asignó una calificación para la ET y una para el ECOE (global), promediando la calificación de las 15 estaciones.

Análisis estadístico: se realizó estadística descriptiva y se usaron las pruebas adecuadas a la distribución de las variables. Para la validez de constructo se correlacionó la puntuación de la ET y la del ECOE mediante el coeficiente de correlación de Pearson; así mismo, se comparó el puntaje de la ET entre quienes tuvieron una calificación en el ECOE por encima y por debajo del punto de

corte mediante la prueba de U de Mann Whitney. Los análisis se realizaron usando el paquete estadístico SPSS/PC v20.

Resultados

La **tabla 1** resume los datos relevantes de ambas certificaciones; el promedio de la ET fue similar en el 2015 y 2016 y cercano al 6,5, al igual que el porcentaje de candidatos con calificación por debajo del punto de corte, 12 y 14%, respectivamente.

La puntuación de los candidatos en el ECOE fue superior a la de la ET en ambos años y cercana al 7,4; ningún candidato tuvo una calificación reprobatoria en el ECOE de acuerdo con el PC. En ambas certificaciones, el PDL del ECOE se calculó en 7,3 (en el año 2015) y 7,4 (en el año 2016), respectivamente y para ello se usaron las listas de cotejo de 30 y 26 candidatos, respectivamente, calificados con desempeño limítrofe. De acuerdo al PDL, el porcentaje de candidatos con calificación reprobatoria incrementó a casi el 50% en ambos años (**tabla 1**).

La **tabla 2** resume el número de rubros de la lista de cotejo para cada estación, la puntuación de los candidatos en cada estación, el PDL para cada estación y el desempeño de los candidatos de acuerdo con el PC y el PDL para la certificación del 2015. Se encontraron datos similares durante la certificación del 2016 (no se muestran). En general, se apreció una variación importante en el PDL de cada estación, de 4,5 para la estación número 2 a un valor de 9,5 para la número 10. El desempeño de los candidatos en cada estación fue mejor cuando se usó el PC que cuando se usó el PDL, salvo en el caso de la estación número 2.

Las **figuras 1 y 2** muestran como la ET y el ECOE correlacionaron de manera significativa y moderada en ambas certificaciones ($r=0.67$ y $r=0.53$, para los años 2015, y 2016, respectivamente, $p=0.001$ en ambos casos). De manera similar, los candidatos con

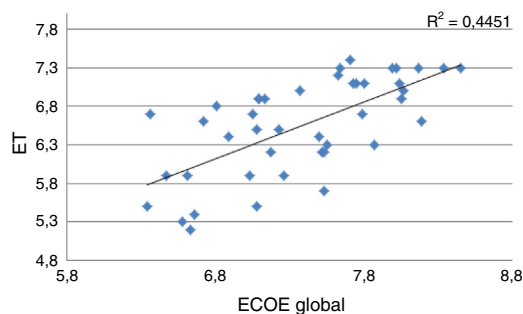


Figura 1. Correlación entre la ET y el ECOE (certificación del 2015).

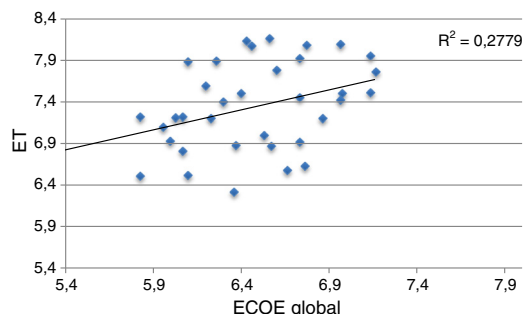


Figura 2. Correlación entre la ET y el ECOE (certificación del 2016).

puntuación en el ECOE \geq PDL tuvieron mejor puntuación en la ET que aquellos con puntuación en el ECOE $<$ PDL, como se muestra en la **tabla 1**.

Discusión

La sociedad actual tiene un creciente interés por el desempeño de los médicos y exige que los trabajadores de la salud estén adecuadamente preparados para ejercer su profesión. Este proceso cuando se define de manera regulada se conoce como certificación de las competencias médicas; implica evaluar, que la evaluación sea sumativa y también requiere establecer puntos por encima de los cuales el candidato acreditará (o no) la certificación. Existen numerosos métodos para definir dichos puntos de corte, sin embargo, la disponibilidad de recursos humanos y económicos es un elemento clave en la selección de alguno en particular. En la actualidad no hay un estándar de oro para evaluar las competencias médicas, si bien el ECOE ha ganado adeptos en años recientes.

El presente trabajo muestra que los candidatos a una certificación en Reumatología tuvieron un desempeño en el ECOE drásticamente variable, en función de cómo se estableció *a priori* el punto de corte para definir la acreditación en la evaluación de las competencias clínicas. Lo anterior fue evidente cuando se consideró el ECOE global para cada candidato, como han propuesto algunos autores, en vez de analizar el desempeño de cada estación⁵. Además del PC se calculó el PDL, para lo cual cada evaluador asignado a una estación, de manera directa y repetida, determinó el desempeño general de los candidatos mediante una escala de Likert; esta estrategia se ha propuesto como un elemento imprescindible para no afectar a la confiabilidad del método de evaluación⁶. La ventaja adicional de usar escalas que evalúan el desempeño general del candidato (a diferencia de las listas de cotejo) es que pueden incluir aspectos cualitativos como la eficiencia y soltura en las actividades realizadas por los candidatos, aspectos importantes en sí mismos y que difícilmente se examinan mediante las listas de cotejo⁷. De hecho, existe evidencia de que las evaluaciones globales son más confiables que las listas de cotejo^{6,7}, sin embargo, se ha señalado que es conveniente integrar varios métodos de evaluación cuando

Tabla 1
Resultados de las certificaciones del 2015 y del 2016

	2015 N=43	2016 N=37
Evaluación teórica (promedio \pm DE)	6,6 \pm 0,6	6,4 \pm 0,6
Candidatos con calificación reprobatoria en la ET/n (%)	5 (12)	5 (14)
ECOE global (promedio \pm DE)	7,4 \pm 0,6	7,3 \pm 0,6
Candidatos con puntuación del ECOE $<$ PC/n (%)	0	0
PDL global	7,3	7,4
Candidatos con puntuación del ECOE $<$ PDL/n (%)	21 (49)	19 (51)
Puntuación de la ET en los candidatos con ECOE \geq PDL/mediana (Q25-Q75)	7,1 (6,4-7,3)	7,8 (7,5-8)
Puntuación de la ET en los candidatos con ECOE $<$ PDL/mediana (Q25-Q75)	6,3 (5,9-6,7)	6,9 (6,5-7,2)

Tabla 2
Desempeño de los candidatos por estación de acuerdo con el PC y al PDL durante el proceso de certificación del año 2015

Estaciones	Puntuación: media \pm DE	PDL	N (%) candidatos con puntaje \geq PC	N (%) candidatos con puntaje \geq PDL
1 (5)	7,6 \pm 1,1	7,7	40 (93)	21 (49)
2 (5)	4,8 \pm 1,6	4,5	11 (26)	18 (42)
3 (6)	6,8 \pm 1,8	6,7	35 (81)	27 (63)
4 (14)	6,1 \pm 1	6,0	21 (49)	13 (30)
5 (14)	7,4 \pm 1,3	6,9	37 (86)	28 (65)
6 (21)	7,7 \pm 1,3	7,6	35 (81)	26 (61)
7 (9)	7,7 \pm 2,3	8,6	26 (61)	23 (53)
8 (15)	6 \pm 1,7	6,8	19 (44)	13 (30)
9 (14)	7,7 \pm 1,8	7,5	31 (72)	26 (61)
10 (19)	9,7 \pm 0,6	9,5	43 (100)	31 (72)
11 (10)	7,7 \pm 1,5	7,7	36 (84)	23 (53)
12 (8)	9,0 \pm 1,3	8,6	43 (100)	22 (51)
13 (8)	7,3 \pm 1,8	7,2	36 (84)	22 (51)
14 (6)	8,7 \pm 1,4	8,2	42 (98)	35 (81)
15 (17)	6,6 \pm 0,9	6,5	30 (70)	19 (44)

se trata de competencias clínicas con altas consecuencias⁸⁻¹⁰. Las diferencias resaltadas de cada instrumento de evaluación explican el variable desempeño de nuestros candidatos con ambos puntos de corte.

El punto de corte seleccionado no afectó a la correlación entre la ET y el ECOE. Nuestro trabajo confirma estudios previos que han mostrado una correlación de moderada a buena entre ambas evaluaciones^{11,12}. Algunos autores han propuesto que la adecuada correlación entre la ET y el ECOE apoya la validez de constructo del ECOE; aunado a lo anterior, observamos que los candidatos que no acreditaron el ECOE de acuerdo con el PDL tuvieron un peor desempeño en la ET. Finalmente, se puede inferir la validez de constructo del PDL por el hecho de que el desempeño calificado como limítrofe de los candidatos varió para cada estación como se mostró en la [tabla 2](#); existía la preocupación de que los evaluadores usaran la puntuación de la lista de cotejo para evaluar el desempeño global del candidato; de haber sido así y consistente entre los diferentes evaluadores, el PDL de las diferentes estaciones hubiera estado cercano al 5. Merece la pena mencionar que la correlación entre la ET y el ECOE no implica necesariamente que ambas herramientas evalúen el mismo conocimiento y por ende sean redundantes; si bien una parte importante de lo que un ECOE evalúa corresponde a la capa del «saber» de la pirámide de Miller¹³, consideramos que el uso de ambas herramientas incrementa la información relativa a las competencias de los candidatos. Teniendo en cuenta que dicho incremento es limitado y que el ECOE consume importantes recursos humanos y económicos, es necesario optimizarlo para mayormente evaluar la capa del conocimiento del «muestra cómo se hace».

El trabajo tiene limitaciones que resaltar; la implementación y elaboración de un ECOE es compleja y requiere de una experiencia amplia y consolidada, de la cual el Consejo está en el proceso de obtención. Un único evaluador por estación (en vez de 2) puntuó la lista de cotejo y el desempeño global del candidato. Si bien encontramos una adecuada correlación entre la ET y el ECOE, no se siguieron las recomendaciones de Matsell et al.¹⁴, quienes sugieren clasificar las estaciones del ECOE en dominios (conocimiento, solución de problemas, habilidades clínicas y manejo del paciente) y correlacionarlas con sus equivalentes en la ET. No pudimos correlacionar la puntuación del ECOE con las evaluaciones obtenidas durante la formación de residentes, por no tener acceso a las mismas. El CMR ha desarrollado un ECOE conformado por un número limitado de estaciones, las cuales además son de corta duración (8 minutos), todo lo cual afecta a la evaluación de aquellos problemas que requieren de la integración de diferentes habilidades, objetivo primordial dentro de un proceso de certificación. Finalmente, en cada certificación se aplicaron 2 ECOE consecutivos debido a la limitación en el número de consultorios, lo cual pudo haber afectado a la reproducibilidad de los resultados¹⁵; no obstante, el espacio físico, los pacientes y los evaluadores fueron los mismos en ambos circuitos.

En nuestra especialidad existen pocos estudios que analizan el desempeño de un ECOE para evaluar las competencias clínicas¹⁶⁻²⁰. Aún son más limitados aquellos que lo hacen en el contexto académico y profesional de una certificación^{5,17}. La más reciente tendencia en educación (médica) se basa en competencias, por lo que el ECOE tiene el potencial de ser la herramienta idónea para evaluar las habilidades clínicas²¹. Sin embargo, en aquellos procesos cuyos resultados tienen un impacto importante en los candidatos y la sociedad, se recomienda el uso de herramientas adicionales y complementarias. Debe hacerse también un cuidadoso análisis de los resultados, los cuales dependen del método que se use para definir el punto de corte, como se aprecia en el presente trabajo.

Finalmente, la reflexión sobre herramientas que evalúan las habilidades clínicas (como el ECOE) debe extenderse a la de la enseñanza de dichas habilidades en nuestra especialidad; en la

actualidad prima el modelo tradicional de educación médica, poco uniforme a lo largo de los diferentes centros de formación de nuestro país, los cuales se caracterizan por la diversidad de las poblaciones que atienden, de los servicios de salud a los cuales pertenecen y de los recursos físicos y humanos con los que cuentan. Este modelo de enseñanza muestra un rezago en la incorporación de tecnologías de educación innovadoras, como la educación médica basada en simulación, la cual ha mostrado favorecer la adquisición de una amplia gama de habilidades clínicas mejor que el modelo tradicional²².

En conclusión, el CMR ha invertido tiempo, compromiso y recursos para otorgar al proceso de certificación en reumatología transparencia, profesionalismo y confiabilidad; para ello ha desarrollado instrumentos de evaluación sólidos y ha cuestionado su desempeño. Consideramos que este es el camino a seguir, a pesar de las dudas y frustración que asoman cuando informamos resultados no aprobatorios y que la misma reflexión y esfuerzos deben extenderse al modelo de formación de nuestros especialistas.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que en este artículo no aparecen datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Conflicto de intereses

No existe conflicto de intereses que declarar.

Agradecimientos

Agradecemos a todos los pacientes quienes año con año, con entusiasmo, esfuerzo y compromiso nos ayudan en el examen nacional de certificación en reumatología.

Bibliografía

1. Downing SM, Haladyna TM. Validity threats: Overcoming interface with proposed interpretations of assessment data. *Med Educ.* 2004;38:327-33.
2. Wass V, van der Vleuten C. The long case. *Med Educ.* 2004;38:1176-80.
3. Nweble DI. The observed long-case in clinical assessment. *Med Educ.* 1994;25:369-73.
4. Pascual Ramos V, Medrano Ramírez G, Solís Vallejo E, Bernard Medina AG, Flores Alvarado DE, Portela Hernández M, et al. Performance of an objective structured clinical examination in a national certification process of trainees in rheumatology. *Reumatol Clin.* 2015;11:215-20.
5. Wilkinson TJ, Newble DI, Frampton CM. Standard setting in an objective structured clinical examination: Use of global ratings of borderline performance to determine the passing score. *Med Educ.* 2001;35:1043-9.
6. Regehr G, Mac Rae H, Reznik RK, Szalay D. Comparing the psychometric properties of checklists and global ratings scales for assessing performance on an OSCE-format examination. *Acad Med.* 1998;73:993-7.
7. Ilgen JS, Ma IWY, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ.* 2015;49:161-73.
8. Schuwirth LW, van der Vleuten CP. The use of clinical simulations in assessment. *Med Educ.* 2003;37:65-71.
9. Wass V, van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet.* 2001;357:945-9.
10. Newble D. Techniques for measuring clinical competence: Objective structured clinical examination. *Med Educ.* 2004;38:199-203.
11. Van der Vleuten CP, van Luyk SJ, Beckers HJ. A written test as an alternative to performance testing. *Med Educ.* 1989;23:97-107.
12. Kissin EY, Grayson PC, Cannella AC, Demarco PJ, Evangelisto A, Goyal J, et al. Musculoskeletal ultrasound objective structured clinical examination: An assessment of the test. *Arthritis Care Res.* 2014;66:2-6.

13. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65 Suppl 9:S63-7.
14. Matsell DG, Wolfish NM, Hsu E. Reliability and validity of the objective structured clinical examination in paediatrics. *Med Educ.* 1991;25:293-9.
15. Van der Vleuten CP, Swanson DB. Assessment of clinical skills with standardized patients: State of the art. *Teach Learn Med.* 1990;2:58-72.
16. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med.* 1975;1:447-51.
17. Raj N, Badcock LJ, Brown GA, Deighton CM, O'Reilly SC. Design and validation of 2 objective structured clinical examination stations to assess core undergraduate examination skills of the hand and knee. *J Rheumatol.* 2007;34:421-4.
18. Hassell AB, West Midlands Rheumatology Services and Training Committee. Assessment of specialist registrars in rheumatology: Experience of an objective structured clinical examination (OSCE). *Rheumatology.* 2002;41:1323-8.
19. Ryan S, Stevenson K, Hassell AB. Assessment of clinical nurse specialists in rheumatology using an OSCE. *Musculoskeletal Care.* 2007;5:119-29.
20. Criscione-Schreiber LG, Solana RJ, Hawley J, Jonas BL, O'Rourke K, Bolster MB. Expert panel consensus on assessment checklists for a rheumatology objective structured clinical examination. *Arthritis Care Res.* 2015;67:898-904.
21. Carraccio C, Englander R. The objective structured clinical examination: A step in the direction of competency-based evaluation. *Arch Pediatr Adolesc Med.* 2000;154:736-41.
22. Mc Gaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A Meta-Analytic comparative review of the evidence. *Acad Med.* 2011;86:706-11.