



Suplemento en Español.

Métodos de evaluación de competencias clínicas en Internados Médicos: revisión de alcance de prácticas actuales y tendencias emergentes -

Nicolás Lavados-Toro^{1*}, Joaquín Salgado-González², Silvana Jiménez-Vera³, Álvaro Herrera-Alcaíno², Oscar Jerez-Yáñez⁵

¹ Facultad de Medicina, Universidad de Chile, Chile, nicolaslavados@ug.uchile.cl, <https://orcid.org/0009-0005-4223-3288>.

² Facultad de Medicina, Universidad de Chile, Chile, joaquin.salgado.g@ug.uchile.cl, <https://orcid.org/0009-0005-3307-1440>.

³ Facultad de Medicina, Universidad de Chile, Chile, silvanajimenez@ug.uchile.cl, <https://orcid.org/0000-0003-3730-7447>.

⁴ Facultad de Medicina, Universidad San Sebastián, Chile, levarito@uchile.cl, <https://orcid.org/0009-0007-4861-2144>.

⁵ Departamento de Educación en Ciencias de la Salud, Facultad de Medicina, Universidad de Chile, Chile, ojerez@uchile.cl, <https://orcid.org/0000-0003-0869-5938>.

* Correspondencia: nicolaslavados@ug.uchile.cl

Tabla S1. Estrategia de búsqueda para cada base de datos.

Bases de datos	Términos utilizados en la búsqueda (limitada a artículos publicados hasta el 27 de Octubre de 2024)	Combinaciones utilizadas
Web of Science	#1 ALL=((“medical student” OR “medical students”) AND (“clinical clerkship” OR “medical clerkship” OR “clinical rotation”))	#4: #1 AND #2 AND #3
	#2 (“competence assessment” OR “competency evaluation” OR “clinical competence assessment” OR “formative assessment” OR “summative evaluation” OR “feedback”)	
	#3 (“workplace-based assessment” OR “portfolio assessment” OR “OSCE” OR “Mini-CEX” OR “Direct Observation of Procedural Skills” OR “DOPS” OR “simulation-based assessment” OR “digital assessment tools” OR “virtual reality” OR “high-fidelity simulators”))	
PubMed	#1 ((“medical student”[All Fields] OR “medical students”[All Fields]) AND (“clinical clerkship”[All Fields] OR “medical clerkship”[All Fields] OR “clinical rotation”[All Fields]))	#4: #1 AND #2 AND #3
	#2 (“competence assessment”[All Fields] OR “competency evaluation”[All Fields] OR “clinical competence assessment”[All Fields] OR “formative assessment”[All Fields] OR “summative evaluation”[All Fields] OR “feedback”[All Fields]))	
	#3 (“workplace-based assessment”[All Fields] OR “portfolio assessment”[All Fields] OR “OSCE”[All Fields] OR “Mini-CEX”[All Fields] OR “Direct Observation of Procedural Skills”[All Fields] OR “DOPS”[All Fields] OR “simulation-based assessment”[All Fields] OR “digital assessment tools”[All Fields] OR “virtual reality”[All Fields] OR “high-fidelity simulators”[All Fields]))	
Scopus	#1 TITLE-ABS-KEY ((“medical student” OR “medical students”) AND (“clinical clerkship” OR “medical clerkship” OR “clinical rotation”))	#4: #1 AND #2 AND #3
	#2 (“competence assessment” OR “competency evaluation” OR “clinical competence assessment” OR “formative assessment” OR “summative evaluation” OR “feedback”)	
	#3 (“workplace-based assessment” OR “portfolio assessment” OR “OSCE” OR “Mini-CEX” OR “Direct Observation of Procedural Skills” OR “DOPS”	



	OR "simulation-based assessment" OR "digital assessment tools" OR "virtual reality" OR "high-fidelity simulators")	
--	--	--

Tabla S2: Ítems extraídos

Categoría	Detalles
Desenlace primario	
Métodos evaluativos	Métodos específicos utilizados para evaluar la competencia clínica, como ECOE, Mini-CEX, DOPS, WBA y evaluaciones basadas en simulación.
Áreas de competencia	Áreas de competencia evaluadas, incluyendo razonamiento clínico, habilidades comunicacionales, profesionalismo, trabajo en equipo y habilidades técnicas.
Mecanismos de retroalimentación	Tipos de retroalimentación entregada, ya sea formativa o sumativa, y los mecanismos mediante los cuales se proporcionó.
Herramientas tecnológicas	Uso de herramientas tecnológicas en las evaluaciones, como plataformas digitales, realidad virtual, simuladores de alta fidelidad y telemedicina.
Variables adicionales	
Autores y año	Autores del estudio y el año de publicación.
Entorno	El escenario clínico o rotación en el que se realizó el estudio.
Tamaño muestral y nivel de los estudiantes	Número de participantes y su nivel de formación médica.
Componentes innovadores	Aspectos únicos o innovadores de los métodos de evaluación utilizados.
Otras variables	
Características de los Participantes	Detalles sobre los participantes, como su año en la carrera de medicina y la rotación clínica en la que participaron.
Percepciones de los Estudiantes	Datos cualitativos y cuantitativos sobre las percepciones de los estudiantes.
Implicaciones en la Práctica	Si se requieren recomendaciones para la formación médica o para futuras investigaciones.

Tabla S3. Características de los estudios

Autor(es)	Año	País	Diseño del estudio	Tamaño muestral	Nivel de los participantes	Escenario
Kasai et al.	2020	Japón	Estudio de cohorte prospectivo	75	Estudiantes en internado clínico	Internado clínico
Shikino et al.	2023	Japón	Estudio observacional prospectivo	79	Estudiantes de 5º año de medicina	Internado clínico
Haruta et al.	2024	Japón	Encuesta longitudinal	831	Estudiantes de medicina	Sesiones de entrevista médica en práctica clínica
Bord et al.	2015	EE. UU.	Estudio transversal	80	Estudiantes de 2º a 4º año de medicina	Internado en Medicina de Urgencias
Malone et al.	2024	EE. UU.	No aleatorizado, 2 grupos	216	Estudiantes de 4º año de medicina	Internado en Medicina de Urgencias
Olupeliyawa et al.	2014	Australia	Estudio transversal	25	Estudiantes de último año de medicina	Rotación clínica final
Qureshi & Zehra	2020	Pakistán	Ensayo clínico aleatorizado (ECA)	80	Estudiantes de último año de medicina	Internado de último año
Gran et al.	2016	Noruega	Estudio cualitativo	30	Estudiantes de 5º año de medicina	Internado en Medicina General
Bozzo et al.	2020	Chile	Estudio longitudinal	462	Estudiantes de 6º año de medicina	Internado en Medicina Interna
Rouse et al.	2024	EE. UU.	Estudio pre y post intervención	200	Estudiantes de 3º año de medicina	Internado en Medicina Interna
Torre et al.	2021	EE. UU.	Estudio de validación	477	Estudiantes de 3º año de medicina	Internado en Medicina Interna
Perrig et al.	2016	Suiza	Estudio pre y post intervención	48	Estudiantes de 4º año de medicina	Internado en Medicina Interna
Patel et al.	2024	EE. UU.	Estudio piloto de intervención	13	Estudiantes de 2º año de medicina	Internado médico
Reid et al.	2021	EE. UU.	Estudio piloto	32	Estudiantes de 3º año de medicina	Internado en Ginecología y Obstetricia
Okubo et al.	2014	Japón	Estudio de cohorte prospectivo	68	Estudiantes de 5º y 6º año de medicina	Unidad ambulatoria, Univ. Médica Femenina de Tokio
Costich et al.	2024	EE. UU.	Estudio piloto	131	Estudiantes en internado clínico	Atención primaria pediátrica
Klapheke et al.	2022	EE. UU.	Estudio piloto	109	Estudiantes de 3º año de medicina	Internado en Psiquiatría
Parikh et al.	2015	EE. UU.	Estudio transversal	389	Estudiantes de 3º año de medicina	Internado en Cirugía
Sullivan et al.	2016	EE. UU.	Estudio de métodos mixtos	98	Estudiantes de 3º año de medicina	Internado en Cirugía
Luo et al.	2023	China	Ensayo clínico aleatorizado (ECA)	38	Estudiantes de 4º año de medicina	Internado en Cirugía
Kim et al.	2016	EE. UU.	Estudio pre y post intervención	110–123	Estudiantes de 3º año de medicina	Varios internados
Martinsen et al.	2021	Noruega	Ensayo clínico aleatorizado (ECA)	38	Estudiantes de 5º año de medicina	Varios internados
Phinney et al.	2022	EE. UU.	Estudio cualitativo	35	Estudiantes de 3º y 4º año de medicina	Varios internados
Rogausch et al.	2015	Suiza	Ánalysis multinivel	165	Estudiantes de 4º año de medicina	Varios internados



Ryan et al.	2021	EE. UU.	Estudio transversal	220	Estudiantes de 3º año de medicina	Varios internados
Ryan et al.	2024	EE. UU.	Estudio transversal	1810	Estudiantes en nivel de internado clínico	Varias instituciones

Tabla S4. Evaluación de calidad

Estudio	Diseño del estudio	Muestreo	Tipo de datos	Validez del instrumento de evaluación	Ánálisis de datos	Desenlace	Puntaje Total
Bord et al. (2015)	3	3	3	3	3	3	18
Bozzo et al. (2020)	1	1	3	3	1	1,5	10,5
Costich et al. (2024)	2	3	3	3	3	3	17
Gran et al. (2016)	1	1	3	3	1	1,5	10,5
Haruta et al. (2024)	1,5	1	3	3	2	1,5	12
Kasai et al. (2020)	1,5	1	3	3	2	1,5	12
Kim et al. (2016)	1,5	2	3	3	2	1,5	13
Klapheke et al. (2022)	1	1	3	3	1	1,5	10,5
Luo et al. (2023)	1,5	1	3	3	2	1,5	12
Malone et al. (2024)	1,5	1	3	3	2	1,5	12
Martinsen et al. (2021)	2	3	3	3	3	3	17
Okubo et al. (2014)	3	3	3	3	3	3	18
Olupeliyawa et al. (2014)	1,5	1	3	3	2	1,5	12
Parikh et al. (2015)	1	1	3	3	1	1,5	10,5
Patel et al. (2024)	1,5	1	3	3	2	1,5	12
Perrig et al. (2016)	2	1	3	3	2	1,5	12,5
Phinney et al. (2022)	1	1	3	3	1	1,5	10,5
Qureshi & Zehra (2020)	2	1	3	3	2	1,5	12,5
Reid et al. (2021)	1,5	1	3	3	2	1,5	12
Rogausch et al. (2015)	1,5	1	3	3	2	1,5	12
Rouse et al. (2024)	3	3	3	3	3	3	18
Ryan et al. (2024)	2	3	3	3	3	3	17
Ryan et al. (2021)	1	1	3	3	1	1,5	10,5
Shikino et al. (2023)	1,5	1	3	3	2	1,5	12
Sullivan et al. (2016)	1,5	1	3	3	2	1,5	12
Torre et al. (2021)	1,5	1	3	3	2	1,5	12

Tabla S5. Resultados y medición de efecto

Estudio	Año	Resultado	Medida(s) del efecto
Bord et al.	2015	Desarrollo de un ECOE para la evaluación de habilidades clínicas en el internado de urgencias	Puntaje promedio: 70,5% (DE = 7,2%), Dificultad de ítem: Baja (>80% correcto), Media (50-80% correcto), Alta (<50% correcto), Discriminación del ítem: Buena ($r_{\text{pb}} > 0,3$), Regular ($r_{\text{pb}} = 0,1-0,3$), Pobre ($r_{\text{pb}} < 0,1$), Correlación biserial puntual (r_{pb}): 0,24 (promedio)
Bozzo et al.	2020	Mejoría en habilidades clínicas durante el internado de medicina interna	Medidas de tendencia central, pruebas de normalidad, pruebas de Wilcoxon pareadas ($p < 0,05$), modelos de regresión lineal múltiple, correlación de Spearman ($r = 0,739$, $p < 0,0001$), alfa de Cronbach (0,8), valores p ($p \leq 0,05$)
Costich et al.	2024	Implementación y percepción docente de las EBA basadas en EPA en atención primaria ambulatoria durante el internado de pediatría	Retroalimentación docente: Aumento en la entrega de retroalimentación específica y orientada a tareas ($p = 0,006$), mayor satisfacción con las oportunidades de retroalimentación (no significativo), más retroalimentación dentro de 24 h (no significativo). Temas de entrevistas: Beneficios de las EBA, barreras a la retroalimentación, sugerencias de mejora
Gran et al.	2016	Experiencias de retroalimentación durante el internado	Ánálisis temático para identificar temas clave y experiencias de retroalimentación, consistencia de los temas entre entrevistas
Haruta et al.	2024	Validación del Simulated Patient Assessment Tool (SPAT) para evaluar desempeño clínico de estudiantes con pacientes simulados	Análisis factorial exploratorio: reveló dos factores (comunicación y desempeño médico), varianza explicada 60,47%. Consistencia interna: alfa de Cronbach global 0,929. ANOVA de un factor: diferencias significativas entre pacientes simulados ($F(34,760)=16,79$, $p<0,001$) y escenarios ($F(20,774)=11,39$, $p<0,001$). Validez convergente: correlación moderada ($r=0,212$, $p<0,05$) entre SPAT y ECOE post-CC
Kasai et al.	2020	Mejoría en desempeño clínico y profesionalismo en estudiantes de internado	Mini-CEX: Mejoras significativas en anamnesis (pre: $5,52 \pm 1,05$, post: $6,57 \pm 0,75$, $p<0,001$), examen físico (pre: $5,38 \pm 0,82$, post: $6,78 \pm 0,87$, $p<0,001$), profesionalismo (pre: $5,87 \pm 1,00$, post: $7,13 \pm 0,81$, $p<0,001$), juicio clínico, consejería, organización/eficiencia y competencia global. P-MEX: Mejoras en relación médico-paciente (pre: $3,00 \pm 0,34$, post: $3,38 \pm 0,24$, $p<0,001$) y habilidades reflexivas (pre: $3,15 \pm 0,32$, post: $3,50 \pm 0,26$, $p<0,001$). Test de rangos con signo de Wilcoxon ($p<0,05$)
Kim et al.	2016	Implementación del requisito de mini-CEX en todos los internados de tercer año y su impacto en observación directa y habilidades clínicas	Cumplimiento: 92% de formularios completados, 78% con retroalimentación específica. Reporte estudiantil: aumento significativo en observación directa de examen físico (ej: cirugía 49%→87%, $p<0,0001$). AAMC GQ: más observación docente en anamnesis y examen físico. ECOE: disminución de reprobación de 12% a 2% ($p=0,0046$). Valores p : <0,0001 (observación en cirugía), 0,0046 (ECOE)
Klapheke et al.	2022	Estudio piloto de EBA usando EPAs y modelo RIME en internado de psiquiatría	Promedios en EPAs y RIME: perfil >4.0 en EPA1, EPA6 y rol de reportero en RIME. Retroalimentación estudiantil: 44% la consideró útil, 32,2% evaluaciones justas. Retroalimentación docente: facilidad de completar EPA = 3,89 (DE=1,27), claridad en explicación = 4,33 (DE=0,5)
Luo et al.	2023	Mejoría en autoconfianza y competencia clínica en internos de cirugía	Evaluación de autoconfianza (SCA): mejoras en todas las estaciones (ej. anamnesis pre: $2,68 \pm 0,82$, post: $3,47 \pm 0,84$, $p<0,01$; examen físico pre: $2,47 \pm 0,84$, post: $3,42 \pm 0,77$, $p<0,01$; RCP, razonamiento clínico, asepsia quirúrgica). Mini-CEX: mejora significativa (grupo ECOE: $6,59 \pm 0,62$ vs. control: $5,31 \pm 1,09$, $p<0,01$). DOPS: sin diferencia significativa (grupo ECOE: $5,15 \pm 0,58$ vs. control: $4,96 \pm 0,63$, $p>0,05$)
Malone et al.	2024	Habilidades en atención urgente y emergente	Prueba de rangos con signo de Wilcoxon ($T = 337,5$, $p = 0,001$), prueba de McNemar ($p \leq 0,001$), análisis temático para datos cualitativos, significancia determinada por $p \leq 0,05$
Martinsen et al.	2021	Habilidades clínicas medidas por evaluaciones mini-CEX	ANOVA ($F = 3,603$, $p = 0,066$), ANCOVA ($F = 1,884$, $p = 0,179$), puntajes promedio (3,5–3,6 sobre 4), desviaciones estándar (0,55–0,63)

Okubo et al.	2014	Evaluación de un programa de internado clínico en atención ambulatoria para mejorar el razonamiento clínico de estudiantes	Mini-CEX: más alto en todas las áreas para estudiantes del programa. SCT: significativamente más alto post-curso (pre: $77,4 \pm 10,5$; post: $86,8 \pm 10,2$). ECOE: mayores puntajes para grupo de intervención vs control
Olupeleyawa et al.	2014	Evaluación del impacto educativo del T-MEX en la colaboración de estudiantes en equipos de salud	Coeficiente de generalización: 0,62 con tres formularios, 0,80 predicho con ocho formularios. Análisis de contenido de retroalimentación y reflexiones. Análisis temático de grupos focales y entrevistas
Parikh et al.	2015	Comunicación, empatía y confianza en cuidados al final de la vida	Estadísticas descriptivas de puntajes promedio (media = 89,0%, DE = 6,7%), desviaciones estándar, correlación con puntajes de confianza ($r = 0,325$, $p < 0,01$) y habilidades comunicacionales ($r = 0,383$, $p < 0,01$)
Perrig et al.	2016	Mejoría en habilidades de examen musculoesquelético	Prueba de Friedman ($p < 0,001$ para CS, $p < 0,001$ para MSES, $p < 0,01$ para IPS), prueba de rangos de Wilcoxon (CS: $p < 0,001$, MSES: $p < 0,001$, IPS: $p < 0,001$), prueba U de Mann-Whitney (CS: $p < 0,01$, MSES: $p < 0,01$, IPS: $p < 0,01$), alfa de Cronbach (0,47–0,83), valores p ($p \leq 0,05$)
Phinney et al.	2022	Retroalimentación y autorreflexión en entornos clínicos	Análisis temático para identificar temas clave y tensiones, consistencia de temas entre entrevistas
Qureshi & Zahra	2020	Habilidades comunicacionales usando retroalimentación de pacientes simulados	Coeficiente de correlación de Pearson ($r = 0,83$) para consistencia interna de LCSAS (alfa de Cronbach = 0,83), mejora de puntajes pre/post-test (diferencia media = 1,5, $p \leq 0,05$)
Rogausch et al.	2015	Influencia de habilidades clínicas previas y características del contexto sobre puntajes mini-CEX en internados	Análisis de regresión: posición clínica del docente fue el predictor más influyente (coeficiente de regresión = 0,55, $p < 0,001$ para residentes vs jefes de departamento). Complejidad de tarea y tamaño de clínica: predictores significativos. ECOE: no predictor significativo. Correlación: débil entre ECOE y mini-CEX ($r = 0,26$ global, $r = 0,27$ por dominio). Valores p: $p < 0,001$ posición docente, $p < 0,05$ complejidad y tamaño
Rouse et al.	2024	ECOE mejorado para neutralizar inflación de notas y evaluación más completa de habilidades clínicas	Comparación ECOE pre y post-intervención (pre: M = 94,25%, DE = 5,65%, post: M = 81,00%, DE = 6,88%), encuestas estudiantiles (promedio: 4,4 Reporter, 4,4 Interpreter, 4,2 Manager, 3,5 tiempo asignado, 3,5 dificultad)
Ryan et al.	2021	EPAs centrales medidas por escala O-SCORE	Teoría de generalización para confiabilidad (coeficiente Phi = 0,19–0,44), puntajes promedio (3,48–3,62), componentes de varianza (estudiante = 3,5%–8%, evaluador = 29,6%–50,3%)
Ryan et al.	2024	Evaluación de confiabilidad de EBA para decisiones de confianza sumativa en educación médica	Coeficiente Phi: >0,7 umbral de confiabilidad aceptable. Varianza atribuida al aprendiz: <10% en la mayoría de análisis. Número de observaciones necesarias: rango 3 a >560, mediana 60
Shikino et al.	2023	Evaluación de efectividad de retroalimentación SRS en razonamiento clínico durante encuentros simulados	Precisión diagnóstica: grupo SRS mayor (pre: 51,3%, post: 89,7%) vs IC recorder (pre: 57,5%, post: 67,5%) ($p=0,037$). Mini-CEX: mejoras significativas en anamnesis, examen físico, profesionalismo, organización/eficiencia y competencia global ($p<0,001$). Checklist: puntaje total mayor SRS (pre: 12,2, post: 16,1) vs IC recorder (pre: 13,1, post: 13,8) ($p<0,001$). Tiempo de retroalimentación: SRS $22,6 \pm 2,1$ min vs IC $27,7 \pm 2,1$ min ($p=0,04$)
Sullivan et al.	2016	Mejoría en habilidades de toma de decisiones clínicas	ANOVA de medidas repetidas (diverticulitis aguda: Pillai's Trace = 0,807, F(2,36) = 75,279, $p<0,000$; hemorragia GI: Pillai's Trace = 0,822, F(2,19)=43,941, $p<0,000$), análisis temático, valores p ($p \leq 0,05$)
Torre et al.	2021	Validación del Multistep Exam (MSX) para evaluar razonamiento clínico analítico en internado de medicina interna	Correlación: MSX correlación positiva significativa con Step 2 CS ICE ($r=0,26$, $p<0,01$). Regresión lineal múltiple: MSX predictor significativo de Step 2 CS ICE ($\beta=0,19$, $p<0,001$), explicando 4% adicional de varianza más allá de NBME Medicina y ECOE Medicina. Confiabilidad: alfa de Cronbach 0,70–0,80
Reid et al.	2021	Evaluación de ECOE telemedicina para manejo de menopausia en internado de Ginecología/Obstetricia	Puntajes de nota post-encuentro (mediana: 20/45), encuestas estudiantiles (78% incomodidad con telemedicina, 66% valor educativo excelente o superior), precisión diagnóstico diferencial (100% identificaron menopausia/perimenopausia, 84% hipertiroidismo)
Patel et al.	2024	Evaluación de doble coaching en habilidades de anamnesis de estudiantes de internado	Puntajes minicard (sin mejora significativa), encuestas estudiantiles (promedio 1,43; 1=Excelente, 5=Pobre), encuestas pacientes (promedio 1,23; 1=Excelente, 5=Pobre), encuestas docentes (promedio 1,69; 1=Excelente, 5=Pobre)

Tabla S6. Área de Competencia, Método de Evaluación, Tecnología y Herramientas, Componentes Innovadores, Hallazgos Clave, Conclusiones, Implicaciones prácticas

Autor(es)	Año	Área de Competencia	Método de Evaluación	Tecnología & Herramientas	Componentes Innovadores	Hallazgos clave	Conclusiones	Implicaciones prácticas
Kasai et al.	2020	Desempeño clínico y profesionalismo	Mini-Clinical Evaluation Exercise (Mini-CEX), Professionalism Mini-Evaluation Exercise (P-MEX)	Juego de roles, revisión entre pares	Combinación de juego de roles y revisión entre pares durante rondas clínicas, múltiples perspectivas en la atención al paciente	El OSCE discrimina eficazmente entre estudiantes de alto y bajo rendimiento, con una amplia distribución de calificaciones.	El OSCE ofrece una herramienta útil para evaluar los conocimientos y habilidades en medicina de urgencias, y proporciona información sobre el desempeño de los estudiantes.	Recomendado para su uso en rotaciones de medicina de urgencias para evaluar y mejorar las habilidades clínicas; puede adaptarse a otras instituciones.
Shikino et al.	2023	Razonamiento clínico y toma de decisiones	Encuentros simulados con pacientes (Mini-CEX, listas de verificación)	Sistema de Reconocimiento de Voz (SRS), grabadora IC	SRS para generar transcripciones de entrevistas y proporcionar retroalimentación específica y precisa	Se demostró una mejora de la competencia a lo largo de las sesiones; alta correlación entre observadores.	Los OSCE mejoran eficazmente la competencia en rotaciones clínicas de alta exigencia.	La integración regular del OSCE con retroalimentación mejora la competencia clínica; respalda el uso de pacientes estandarizados para retroalimentación formativa en tiempo real.
Haruta et al.	2024	Habilidades de entrevista médica	Simulated Patient Assessment Tool (SPAT)	Ninguna	Desarrollo y validación de SPAT, uso de escala Likert de 6 puntos con 13 ítems	Mejor especificidad y oportunidad de la retroalimentación, mayor satisfacción con las oportunidades de recibir retroalimentación.	Las WBAs son factibles en el entorno ambulatorio y mejoran la entrega de retroalimentación.	Se necesita mayor desarrollo y capacitación docente; se recomienda explorar las perspectivas de los estudiantes sobre el impacto de las WBAs.
Bord et al.	2015	Habilidades clínicas	Objective Structured Clinical Examination (OSCE)	Simulaciones con maniquí, encuentros con pacientes estandarizados	Uso de EM Milestones, escenarios con fases estables e inestables, retroalimentación inmediata	Los médicos generales y los estudiantes enfatizaron la importancia de la confianza mutua y la retroalimentación oportuna para el desarrollo profesional.	Herramientas estructuradas como StudentPEP pueden mejorar los procesos de retroalimentación.	Recomendaciones para mecanismos de retroalimentación más estructurados en las rotaciones clínicas.
Malone et al.	2024	Habilidades en atención urgente y emergente	Evaluaciones del programa, encuestas, retroalimentación de docentes	Simulaciones de alta fidelidad (HF), simulaciones de Realidad Virtual (VR)	Comparación de simulaciones HF y VR para evaluar competencia en atención urgente y	Se confirmó la validez y fiabilidad del SPAT, con diferencias significativas en las puntuaciones entre	El SPAT es efectivo para evaluar el desempeño clínico, aunque se requiere estandarización en	Estandarizar el proceso de evaluación de pacientes estandarizados, considerar la selección de escenarios para exámenes de alta exigencia.

					emergente	pacientes estandarizados y escenarios.	las evaluaciones de pacientes estandarizados.	
Olupeliyawa et al.	2014	Trabajo en equipo y colaboración	Teamwork Mini-Clinical Evaluation Exercise (T-MEX)	Ninguna	Enfoque en encuentros clínicos, reflexión estructurada, escala de respuesta según etapa	Mejoría significativa en las habilidades de comunicación, entrevistas médicas, exámenes físicos y profesionalismo.	El juego de roles y la revisión entre pares mejoran el desempeño clínico y el profesionalismo.	Estrategia educativa efectiva con recursos limitados, promueve una atención integral del paciente.
Qureshi & Zehra	2020	Habilidades comunicacionales en interacciones clínicas	OSCE con pacientes simulados (SPs)	Liverpool Communication Skills Assessment Scale (LCSAS)	Retroalimentación estructurada de los SPs tras la interacción	Mayor observación directa de exámenes físicos y anamnesis, disminución en las tasas de fracaso en el OSCE.	El mini-CEX es factible en todas las rotaciones clínicas, mejora la observación directa y las habilidades clínicas.	Recomendado para su uso en todas las rotaciones clínicas para mejorar la observación, la retroalimentación y las habilidades clínicas.
Gran et al.	2016	Retroalimentación y supervisión	Entrevistas estructuradas	StudentPEP	Uso de StudentPEP para retroalimentación estructurada y reflexión	El perfil promedio de habilidades de los estudiantes sugirió que ya no necesitaban supervisión directa en EPA1 y EPA6.	El marco EPA/RIME tuvo éxito con un compromiso adicional mínimo de tiempo por parte de los docentes.	Se recomienda un mayor desarrollo de habilidades en las EPAs 2-5, con más evaluaciones e instrucción docente.
Bozzo et al.	2020	Razonamiento clínico, anamnesis, comunicación	Objective Structured Clinical Examinations (OSCEs)	Pacientes simulados estandarizados	Diseño iterativo de OSCE con retroalimentación formativa y evaluación estructurada en múltiples escenarios	Mayor autoconfianza y competencia clínica en anamnesis, examen físico y razonamiento clínico, pero no en habilidades procedimentales.	El OSCE formativo con retroalimentación inmediata potencia la autoconfianza y la competencia clínica.	Recomendado para la formación previa a la rotación clínica, con el fin de mejorar el desempeño y la confianza.
Rouse et al.	2024	Habilidades clínicas y razonamiento	OSCE multiestación	Ninguna	Integración del marco RIME, formato de tres estaciones, entrega en video para la estación Manager	Las simulaciones de alta fidelidad (HF) fueron mejor valoradas (4.6/5) que la realidad virtual (VR) (4.4/5); la competencia fue alcanzada por el 91.7% en HF frente al 65.5% en VR ($p \leq .001$).	La realidad virtual puede evaluar eficazmente las habilidades en cuidados de urgencia, aunque requiere aclimatación; la simulación de alta fidelidad sigue siendo superior para la evaluación sumativa.	Se aconseja precaución al usar realidad virtual para evaluación sumativa; se requiere práctica adicional en entornos de VR.
Torre et al.	2021	Razonamiento clínico	Multistep Exam	Ninguna	Formato por pasos	No hubo mejoría	El mini-CEX es	Se sugiere el uso continuo del

			(MSX)		para evaluar razonamiento clínico analítico, formulario estructurado para cada paso	significativa en los puntajes de OSCE y pruebas, cierta mejora en habilidades autoinformadas; la retroalimentación fue valorada pero escasa.	factible y valorado, pero no se observó un impacto educativo significativo en los puntajes.	mini-CEX con mayor capacitación de los evaluadores en la entrega de retroalimentación accionable para lograr impacto.
Perrig et al.	2016	Examen musculoesquelético, comunicación y comportamiento profesional	Objective Structured Clinical Examination (OSCE)	No especificado	Enseñanza interactiva en pequeños grupos con retroalimentación sistemática de pacientes, pares e instructores	Mejoría significativa en las habilidades de razonamiento clínico, con mayores puntajes en mini-CEX, SCT y OSCE.	La práctica reflexiva en un entorno ambulatorio mejora las habilidades de razonamiento clínico.	Recomendado para su uso en rotaciones clínicas para mejorar las habilidades de razonamiento clínico; adaptable a otros entornos educativos.
Patel et al.	2024	Habilidades de anamnesis	Entrevistas a pacientes (herramienta Minicard)	Zoom, grabaciones de audio y video	Doble coaching por médicos internos y docentes, retroalimentación inmediata	Mejoría en las competencias colaborativas, retroalimentación efectiva y autoevaluación informada.	El T-MEX facilita el aprendizaje mediante retroalimentación estructurada, reflexión y aprendizaje situado.	Recomendado para evaluar y desarrollar habilidades de trabajo en equipo en entornos clínicos.
Reid et al.	2021	Habilidades en telemedicina, manejo de menopausia	Telemedicine OSCE	Zoom, checklist SP	Encuentro virtual de telemedicina, retroalimentación de SP sobre comunicación digital, nota post-encuentro	Correlación positiva entre el desempeño en OSCE y las puntuaciones de confianza/comunicación.	Los OSCE son efectivos para entrenar la comunicación y la empatía.	Los OSCE de cuidados al final de la vida en las rotaciones clínicas pueden mejorar habilidades profesionales tempranas en escenarios desafiantes.
Okubo et al.	2014	Habilidades de razonamiento clínico	Objective Structured Clinical Examination (OSCE), Script Concordance Test (SCT)	Método SNAPPS, mini-CEX, preceptor de 1 minuto	Uso de práctica reflexiva, retroalimentación diaria, múltiples oportunidades de reflexión, integración de SNAPPS y mini-CEX	No hubo mejoría significativa en las habilidades de anamnesis, pero se recibió retroalimentación positiva de estudiantes y pacientes.	La tutoría dual es beneficiosa para la retroalimentación, aunque se pueden requerir más sesiones para mejorar las habilidades.	Incorporar sesiones de retroalimentación más frecuentes y longitudinales; considerar la revisión en video para discusiones más profundas.
Costich et al.	2024	Actividades Profesionales Confiables (EPAs)	Workplace-Based Assessments (WBAs)	Qualtrics™, escala de confianza modificada de Chen	Entrenamiento basado en desempeño (PDT) y entrenamiento de marco de referencia (FORT) para docentes	Mejoría significativa en habilidades clínicas (CS) y musculoesqueléticas (MSES) inmediatamente después de la intervención y mejora sostenida en el seguimiento (T2); las	Intervenciones adicionales dirigidas en pequeños grupos tienen un impacto duradero en las habilidades clínicas.	Incluir sesiones estructuradas orientadas a la retroalimentación en las rotaciones clínicas mejora la retención de habilidades, particularmente en exámenes prácticos como las habilidades musculoesqueléticas.

						habilidades IPS aumentaron inicialmente pero disminuyeron con el tiempo.		
Klapheke et al.	2022	Habilidades clínicas	Evaluaciones en el lugar de trabajo (WBAs)	Escala supervisora EPA/RIME	Integración de EPAs y modelo RIME para educación basada en competencias	Las evaluaciones basadas en dispositivos móviles (WBA) aumentaron la aceptación, facilitaron la retroalimentación en tiempo real; se observaron tensiones respecto al impacto de la retroalimentación sumativa percibida.	Efectivo para la evaluación formativa, aunque persisten tensiones respecto al peso percibido de la retroalimentación.	Necesidad de un cambio cultural para considerar las WBAs como evaluaciones de bajo riesgo; apoyo continuo del supervisor y mayor autonomía de los estudiantes en su realización.
Parikh et al.	2015	Comunicación, empatía y confianza en cuidados al final de la vida	OSCE con pacientes estandarizados (SPs)	Escenarios de cuidados al final de la vida basados en simulación	Integración de cuidados paliativos en cirugía usando retroalimentación de SP	Mejoría significativa en las habilidades de comunicación ($p \leq 0.05$).	La retroalimentación de los pacientes estandarizados es efectiva para mejorar las habilidades de comunicación.	La retroalimentación de pacientes estandarizados debe integrarse en los OSCE de las rotaciones clínicas para el entrenamiento en comunicación.
Sullivan et al.	2016	Razonamiento clínico y toma de decisiones	Casos quirúrgicos virtuales de pacientes (VSPCs)	Casos virtuales de pacientes (escenarios quirúrgicos simulados)	Modelo de aprendizaje cognitivo-aprendizaje usando VSPCs con retroalimentación en cada intento	Los estudiantes diagnosticaron correctamente la menopausia pero tuvieron dificultades con las estrategias de manejo, encontrando desafiante el formato de telemedicina.	El OSCE en telemedicina es efectivo para practicar encuentros virtuales con pacientes y el manejo de la menopausia.	Considerar la integración temprana de la formación en telemedicina y proporcionar instrucción sobre la menopausia antes del encuentro.
Luo et al.	2023	Autoconfianza y competencia clínica	OSCE formativa con retroalimentación inmediata	Ninguna	Retroalimentación inmediata después de cada estación OSCE, método de "feedback sandwich"	Las puntuaciones de mini-CEX estuvieron más influenciadas por las características del contexto que por las habilidades clínicas previas.	Las características del contexto impactan significativamente en las puntuaciones del mini-CEX.	Considerar el enfoque en la retroalimentación narrativa o mejorar el diseño de las WBAs para aumentar la validez de las puntuaciones.
Kim et al.	2016	Habilidades clínicas	Mini-Clinical Evaluation Exercise (mini-CEX)	Ninguna	Implementación en todos los internados de 3er año, sistemas de órganos prescritos para observación	El OSCE mejorado proporcionó una evaluación más completa, redujo la inflación de calificaciones y	El OSCE multiescena con el marco RIME es efectivo para evaluar las habilidades clínicas.	Considerar ajustes adicionales en la asignación de tiempo y el contenido de las estaciones; se requiere evaluación continua.

						mejoró la distribución de notas.		
Martinsen et al.	2021	Habilidades clínicas, anamnesis, examen del paciente	Mini-Clinical Evaluation Exercise (mini-CEX)	Sesiones de retroalimentación observacional	Modelo de retroalimentación estandarizado, mínimo de 8 evaluaciones mini-CEX, OSCE, pruebas escritas y encuestas post-práctica	Fiabilidad modesta en el uso del O-SCORE para la evaluación de estudiantes de medicina; alta variabilidad atribuida a los evaluadores.	El O-SCORE mostró una fiabilidad limitada en el contexto de evaluación de estudiantes de medicina.	Se sugiere una capacitación más sólida de los evaluadores y posiblemente reducir el número de evaluadores o implementar nuevas escalas específicas para las EPAs.
Phinney et al.	2022	Retroalimentación y autorreflexión en entornos clínicos	Herramienta de evaluación en el lugar de trabajo (WBA)	Códigos QR, acceso móvil	Dos iteraciones de WBA: herramientas de escritorio y móviles; autocumplimiento por estudiantes, uso de teoría de actividad histórico-cultural (CHAT) para examinar tensiones	Baja fiabilidad de las WBAs para decisiones sumativas, altas puntuaciones en todas las escalas y mayor observación directa.	Las WBAs se utilizan mejor para retroalimentación formativa, no son fiables para decisiones sumativas de confianza.	Se necesita más investigación para desarrollar instrumentos fiables en decisiones sumativas; usar las WBAs para mejorar la retroalimentación y la observación.
Rogausch et al.	2015	Habilidades clínicas	Mini-Clinical Evaluation Exercises (mini-CEX)	Ninguna	Ánalysis de características del contexto y habilidades clínicas previas sobre puntajes mini-CEX	Mejor precisión diagnóstica, mayores puntuaciones en Mini-CEX y listas de verificación con SRS.	La retroalimentación mediante SRS conduce a una mayor precisión diagnóstica y mejor desempeño clínico.	La retroalimentación basada en SRS es efectiva y eficiente, recomendada para mejorar la formación clínica.
Ryan et al.	2021	Actividades Profesionales Confiables (EPAs) incluyendo anamnesis, examen físico, diagnóstico diferencial, etc.	Ottawa Surgical Competency Operating Room Evaluation (O-SCORE)	Sistema WBA compatible con móviles	Solicitudes de WBA basadas en móviles impulsadas por estudiantes, integrando EPAs centrales para retroalimentación formativa; evaluación frecuente para alcanzar confiabilidad	Mejoría en el razonamiento clínico con los intentos, pero persistencia de errores en algunas áreas.	Los VSPCs son beneficiosos para el razonamiento clínico y la toma de decisiones.	Se recomiendan más casos y puntos de decisión variados para un aprendizaje mejorado.
Ryan et al.	2024	Actividades Profesionales Confiables (EPAs)	Workplace-Based Assessments (WBAs)	Ninguna	Estudio de generalización multiinstitucional, diversas estrategias de implementación, diferentes escalas y tipos de evaluadores	Correlación positiva significativa entre MSX y los puntajes de Step 2 CS ICE, siendo MSX Step 1 el más predictivo del puntaje ICE.	El MSX es útil para evaluar y proporcionar retroalimentación sobre el razonamiento clínico.	Se necesitan más estudios en diferentes tipos de estudiantes y escuelas de medicina.

Tabla S7. Tipo de retroalimentación, Mecanismo de Retroalimentación, Percepción de los Estudiantes

Autor(es)	Formativo/Sumativo	Mecanismo de retroalimentación	Percepción de los estudiantes
Bord et al.	Formativo	Retroalimentación inmediata de los docentes observadores y de los pacientes estandarizados	Los estudiantes recibieron retroalimentación breve después de cada caso y consideraron el OSCE valioso para la evaluación.
Bozzo et al.	Formativo y sumativo	Retroalimentación verbal inmediata de los pacientes estandarizados y retroalimentación escrita para evaluaciones sumativas	Los estudiantes encontraron valor en los escenarios de pacientes estructurados y realistas, así como en la retroalimentación.
Costich et al.	Formativo	Retroalimentación específica, orientada a tareas, con respuesta abierta	No mencionado directamente en el artículo.
Gran et al.	Formativo	Retroalimentación oportuna y constructiva	Los estudiantes valoraron la retroalimentación estructurada, pero la consideraron demandante en tiempo.
Haruta et al.	Formativo	Retroalimentación oral por PEs, SPAT scores	Los estudiantes encontraron valiosa y constructiva la retroalimentación de los pacientes estandarizados.
Kasai et al.	Formativo	Retroalimentación entre pares durante juego de roles y sesiones de revisión entre pares	Los estudiantes reconocieron la importancia de múltiples perspectivas en la atención del paciente.
Kim et al.	Formativo	Retroalimentación inmediata documentada en formularios (mini-CEX)	Los estudiantes informaron mayores tasas de observación directa y retroalimentación.
Klapheke et al.	Formativo	Retroalimentación formativa individualizada vía mail	Cerca de la mitad consideró útil la retroalimentación, pero la mayoría no quiso que su desempeño fuera compartido con los directores de residencia.
Luo et al.	Formativo	Retroalimentación inmediata después de cada OSCE(ECOE)	Los estudiantes reportaron un aumento en su autoconfianza tanto a corto como a largo plazo.
Malone et al.	Formativo	Encuestas de evaluación, preguntas abiertas, retroalimentación de los docentes	Los estudiantes encontraron la realidad virtual inmersiva y una buena preparación, aunque señalaron la necesidad de más práctica en VR.
Martinsen et al.	Formativo	Retroalimentación estructurada después de las sesiones de mini-CEX, OSCE y pruebas escritas	Los estudiantes valoraron el mini-CEX por la retroalimentación directa, pero desearon retroalimentación más detallada sobre cómo mejorar.
Okubo et al.	Formativo	Retroalimentación inmediata de los docentes utilizando el método de preceptor de 1 minuto y mini-CEX	Retroalimentación positiva sobre la utilidad del entrenamiento práctico y de razonamiento clínico.
Olupeliyawa et al.	Formativo	Retroalimentación estructurada de los evaluadores, reflexiones de los estudiantes y planes de acción	Los estudiantes valoraron la retroalimentación estructurada y las oportunidades de reflexión.
Parikh et al.	Formativo	Retroalimentación inmediata de pacientes estandarizados y docentes	Los estudiantes valoraron la experiencia, destacando mayor comodidad para manejar conversaciones delicadas.

Patel et al.	Formativo	Retroalimentación inmediata de pacientes hospitalizados y docentes a través de Zoom	Los estudiantes apreciaron la retroalimentación específica y oportuna; los pacientes disfrutaron participar.
Perrig et al.	Formativo	Retroalimentación inmediata del instructor, compañeros y pacientes	Los estudiantes valoraron la práctica en pequeños grupos con pacientes reales y la retroalimentación detallada.
Phinney et al.	Formativo	Observación del supervisor, formularios móviles de WBA, autoevaluación del estudiante	Los estudiantes valoraron la retroalimentación en tiempo real; consideraron la autoevaluación más fácil, aunque a veces menos auténtica.
Qureshi & Zehra	Formativo	Retroalimentación inmediata de pacientes estandarizados después del OSCE	Los estudiantes encontraron valiosa la retroalimentación de los pacientes estandarizados para el perfeccionamiento de habilidades.
Reid et al.	Formativo	Retroalimentación verbal inmediata de pacientes estandarizados, puntuación de la nota posterior al encuentro	Los estudiantes apreciaron la práctica en telemedicina, la consideraron incómoda pero útil.
Rogausch et al.	Formativo	Comentarios narrativos, retroalimentación estructurada	No mencionado directamente en el artículo.
Rouse et al.	Sumativo	Retroalimentación estructurada de los docentes, revisión en video	Los estudiantes consideraron el OSCE justo y representativo de las experiencias clínicas.
Ryan et al.	Formativo	Observación directa y sistema de retroalimentación O-SCORE	A los estudiantes se les asignó iniciar las WBAs, a veces calificando las sesiones de forma estratégica.
Ryan et al.	Formativo	Retroalimentación formativa inmediata, tutoría formal	Los estudiantes percibieron un aumento en la frecuencia de la observación directa.
Shikino et al.	Formativo	Retroalimentación inmediata utilizando SRS o grabadora IC	Los estudiantes encontraron la retroalimentación mediante SRS más efectiva y eficiente.
Sullivan et al.	Formativo	Retroalimentación inmediata, discusiones dirigidas por el instructor	Los estudiantes apreciaron un entorno seguro para ensayo y error.
Torre et al.	Formativo	Retroalimentación inmediata, rúbrica de evaluación estructurada	Los estudiantes se beneficiaron de pasos explícitos en el proceso de razonamiento.