



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MÉDICAS,
ODONTOLÓGICAS Y DE LA SALUD

“EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS
MÉDICAS: CARACTERÍSTICAS DE LOS SUSTENTANTES, TIPO DE
ESCUELA Y DESEMPEÑO OBTENIDO”

TESIS
QUE PARA OPTAR POR EL GRADO DE
MAESTRO EN CIENCIAS SOCIOMÉDICAS
Campo disciplinario Educación en Ciencias de la Salud

PRESENTA

EL ALUMNO:

Miguel Angel Gaxiola García

TUTOR:

Dr. Adrián Martínez González

Director de Evaluación Educativa

Coordinación de Universidad Abierta, Innovación Educativa
y Educación a Distancia (CUAIEED)

Ciudad Universitaria. Ciudad de México, Agosto 2020



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

TABLA DE CONTENIDO

1 RESUMEN	6
2 INTRODUCCIÓN	9
3 ANTECEDENTES	9
3.1 CARACTERÍSTICAS DE LOS PROGRAMAS DE ESTUDIOS DE LA LICENCIATURA EN MEDICINA EN MÉXICO	9
3.2 EXÁMENES DE ALTAS CONSECUENCIAS	12
3.3 SELECCIÓN PARA EL SISTEMA NACIONAL DE RESIDENCIAS EN MÉXICO.....	14
3.3.1 SISTEMA DE SELECCIÓN PARA RESIDENCIAS: FORTALEZAS.....	17
3.3.2 SISTEMA DE SELECCIÓN PARA RESIDENCIAS: DEBILIDADES.....	17
3.4. INFERENCIAS A PARTIR DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.....	19
3.5. PERSPECTIVAS SOBRE LA UTILIDAD DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.....	21
3.6. VALIDEZ DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.....	22
3.6.1. Contenido del Examen.	23
3.6.2. Proceso de respuesta.....	23
3.6.3. Estructura interna.....	24
3.6.4. Relación con otras variables.....	25
3.6.5. Consecuencias del examen.....	27
MATERIAL Y MÉTODOS	28
4. PREGUNTA DE INVESTIGACIÓN.....	28
5. PLANTEAMIENTO DEL PROBLEMA	28
6. JUSTIFICACIÓN	29
7. HIPÓTESIS	29

8. OBJETIVOS.....	32
8.1 OBJETIVO GENERAL.....	32
8.2 OBJETIVOS ESPECÍFICOS.....	32
9. METODOLOGÍA.....	33
9.1 CONSIDERACIONES ÉTICAS.....	33
9.2 CONSIDERACIONES FINANCIERAS Y POTENCIALES CONFLICTOS DE INTERÉS	33
9.3 ESTRATEGIA DE BÚSQUEDA BIBLIOGRÁFICA	33
9.4 DISEÑO DEL ESTUDIO.....	34
9.5 DEFINICIÓN DE LAS VARIABLES.....	36
9.5.1 Criterios de inclusión.....	36
9.5.2 Criterios de exclusión	36
9.5.3 Variables sociodemográficas.....	37
9.6 DEPURACIÓN DE LA BASE DE DATOS.....	37
9.7 ESTADÍSTICA.....	38
9.7.1 Tamaño de muestra.....	38
9.7.2 Análisis estadístico.....	39
10. RESULTADOS.....	42
10.1 RESULTADOS GLOBALES.....	42
10.1.1 Rendimiento en el ENARM y su relación con el promedio en la licenciatura.....	43
10.1.2 Estatus de acreditación del programa académico y rendimiento en el ENARM.....	45
10.1.3 Tipo de escuela y rendimiento en el ENARM.....	45
10.1.4 Región geográfica y rendimiento en el ENARM.....	46
10.2 RESULTADOS EN SUSTENTANTES DE PRIMERA VEZ.....	47
10.2.1 Rendimiento en el ENARM y su relación con el promedio en la licenciatura.....	48
10.2.2 Estatus de acreditación del programa académico y rendimiento en el ENARM.....	49

10.2.3 Tipo de escuela y rendimiento en el ENARM.....	49
10.2.4 Región geográfica y rendimiento en el ENARM.....	50
10.3 COMPARACIÓN ENTRE LOS DOS GRUPOS.....	51
10.3.1 Rendimiento en el ENARM	51
11. DISCUSIÓN	54
12. CONCLUSIONES.....	63
13. CRONOGRAMA DE ACTIVIDADES.....	64
14. REFERENCIAS BIBLIOGRÁFICAS	65

1 RESUMEN

INTRODUCCIÓN. La mayoría de los egresados de la carrera de medicina realizan el Examen Nacional de Aspirantes a Residencias Médicas (ENARM) para ganar su acceso a una residencia médica; de éstos sólo una minoría lo consigue (22% en 2016). Para el momento en que se realiza este examen, de acuerdo a las cronologías de los planes de estudio en las etapas últimas del servicio social o una vez terminado este, muchas de las Escuelas de Medicina del país ya habrán realizado exámenes de titulación a sus egresados. A estos parámetros se les puede contrastar el promedio general de egreso obtenido por el estudiante que refleja no solo las evaluaciones sumativas al final de cada ciclo escolar sino también las evaluaciones formativas parciales realizadas desde el primer año de ingreso. Puesto que todos estos parámetros reflejan de manera intuitiva la calidad de la educación médica proporcionada en una escuela, se han estudiado de manera fragmentada y con la más diversa metodología para poder arrojar conclusiones. Sin embargo, hasta el momento no existe una forma de determinar la correlación entre el ENARM y el desempeño durante la carrera de medicina general expresada como un promedio de calificaciones. Además de esta variable (teóricamente atribuible al egresado), existen características de la escuela de medicina que potencialmente pueden afectar el desempeño en el ENARM, estas son: el estatus de acreditación del programa educativo de la escuela, la categorización de la escuela de medicina como pública o privada y la región geográfica donde se encuentra la escuela de medicina de procedencia.

OBJETIVO. Identificar la relación entre el promedio obtenido en la licenciatura y el desempeño en el ENARM (calificaciones y tasa de selección), segmentando de manera particular al grupo de médicos recién graduados, y determinar las características de la escuela de medicina que impactan este desempeño.

MATERIAL Y MÉTODOS. Estudio en dos etapas; primera etapa: estudio observacional no-experimental (descriptivo); segunda etapa: estudio analítico inferencial. Población: La

población estudiada fueron todos los médicos sustentantes del ENARM en el periodo que comprende del año 2014 al año 2018.

RESULTADOS. La correlación de Pearson ($r=0.35$ para toda la muestra y $r=0.50$ para sustentantes de primera vez) indicó una relación entre el desempeño obtenido en el ENARM y el promedio de calificaciones de la licenciatura. Las diferencias en las tasas de selección para cada uno de los rangos de calificaciones fueron estadísticamente significativas al aplicar la prueba de X^2 ($p<0.001$). Tanto de manera global como para los sustentantes de primera vez se determinó un mayor desempeño (t de Student; $p<0.001$) y tasa de selección (X^2 ; $p<0.001$) para escuelas acreditadas así como para escuelas públicas. Las escuelas de medicina segmentadas por regiones geográficas mostraron un desempeño significativamente mayor en Noreste y Noroeste en comparación con Sur-Sureste en la muestra global (ANOVA; $p<0.001$), y un desempeño mayor para Centro-Occidente en sustentantes de primera vez (ANOVA; $p<0.001$). Asimismo, la tasa de selección ajustada por regiones también mostró diferencias, siendo significativamente mayor en Noreste y Noroeste en la muestra global (X^2 ; $p<0.001$) y mayor en Centro-Occidente cuando se analizaron sustentantes de primera vez (X^2 ; $p<0.001$).

CONCLUSIONES. Los hallazgos del estudio muestran que el rendimiento heterogéneo en el ENARM guarda relación con el número de veces que se ha presentado el examen, con el promedio obtenido en la licenciatura, con el tipo de programa académico (acreditado o no acreditado), con el tipo de escuela de procedencia (pública o privada) y con la región geográfica donde se localiza la escuela de procedencia.

2 INTRODUCCIÓN

Los programas de formación médica tienen como propósito principal formar médicos calificados, entendidos como los profesionales que cuentan con los conocimientos, destrezas y actitudes necesarias para diagnosticar y resolver con tratamientos y procedimientos médicos la mayoría de los padecimientos que el ser humano sufre en su vida (Vázquez Martínez & Ortiz-León, 2017) (Petra-Micu, Cortés-Morelos, Talayero-Uriarte, & Fouilloux-Morales, 2012). Sin embargo, el médico general tiene pocas oportunidades para ejercer profesionalmente en nuestro país, ya que existe una marcada orientación hacia la medicina especializada en la práctica de las instituciones de salud. Por ello, la mayoría de los egresados de la carrera de medicina aspiran a formarse como especialistas (Frenk-Mora J, Robledo-Vera C, Nigenda-López G, Ramírez-Cuadra C, Galván-Martínez O, 1990).

Las especialidades médicas se refieren la educación médica que se recibe al término de los estudios de la licenciatura como un requisito para la especialización; el objetivo de estas es ampliar y profundizar en un campo médico específico de atención a la salud, los conocimientos, actitudes y habilidades médicas logradas en la licenciatura; asimismo, fortalecer la capacidad de resolver problemas médicos complejos y de alta competencia profesional, de desarrollar el juicio clínico y el pensamiento crítico aplicado a su ámbito laboral. En México más del 90% de los egresados de licenciatura aspira a esta continuidad (Akaki Blancas & López Bárcena, 2018).

3 ANTECEDENTES

3.1 CARACTERÍSTICAS DE LOS PROGRAMAS DE ESTUDIOS DE LA LICENCIATURA EN MEDICINA EN MÉXICO.

Las escuelas de Medicina alrededor del mundo varían con respecto a sus estándares educativos, currículo, y métodos de evaluación. Muchos países tienen sus propios programas

de certificación. En Estados Unidos, la herramienta que se utiliza para la certificación de los médicos graduados es el examen USMLE (United States Medical Licensing Examination) que consiste en una evaluación en tres pasos necesaria para obtener la licencia profesional en medicina de Estados Unidos. Este examen es llevado a cabo por la FSMB (Federation of State Medical Boards) en conjunto con la NBME (National Board of Medical Examiners). Se requiere que los médicos con un grado MD (Doctor of Medicine; MD del latín *Medicinae Doctor*) aprueben este examen antes de iniciar su práctica clínica; es decir se requiere su aplicación y aprobación aunque no se vaya a realizar una residencia médica de especialidad (Rubright, Jodoin, & Barone, 2018). Los médicos graduados en escuelas fuera de Estados Unidos y que desean integrarse al sistema estadounidense homologan sus estudios médicos de pregrado a través de la ECFMG (Educational Commission for Foreign Medical Graduates); el propósito de esta certificación es evaluar la aptitud y pertinencia de los médicos graduados fuera de Estados Unidos para someterse a entrenamiento clínico en programas de residencia o alta especialidad (fellowship) en Estados Unidos.

En México existen evaluaciones similares al USMLE estadounidense; una de ellas es el Examen General para el Egreso de la Licenciatura en Medicina General (EGEL) administrado por el CENEVAL (Centro Nacional de Evaluación para la Educación Superior) una asociación civil sin fines de lucro cuya actividad principal, de acuerdo a su perfil institucional, es el diseño y la aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias, así como el análisis y la difusión de los resultados que arrojan las pruebas. El EGEL de medicina general consiste en una prueba de tipo “criterial”, es decir, basada en indicadores de dominio profesional (conocimientos y habilidades) previamente determinados por el Consejo Técnico del Área Médica adscrito al CENEVAL. La prueba tiene un formato de opción múltiple y su propósito exclusivo es determinar el grado cuantitativo de conocimientos que los estudiantes, en lo individual, han alcanzado al término de la carrera (“Informe Anual

de Resultados 2014. EXAMEN GENERAL PARA EL EGRESO DE LA LICENCIATURA EN MEDICINA GENERAL (EGEL). Centro Nacional de Evaluación para la Educación Superior, A. C. (CENEVAL),” 2015).

3.1.1 Acreditación de los programas educativos. Tipo de escuela: pública o privada.

En México, la calidad de la atención médica usualmente se asocia con el estatus de acreditación de las escuelas de medicina, considerando que los procesos de acreditación pueden fomentar tanto prácticas de auto-evaluación como de mejora y en consecuencia asegurar que los estudiantes de medicina reciban experiencias educativas de alta calidad basadas en estándares establecidos. (Van Zanten, Norcini, Boulet, & Simon, 2008) Un programa académico acreditado para una escuela de medicina es aquel que cumple con los requerimientos del Consejo Mexicano para la Acreditación de la Educación Médica (COMAEM); es decir, alcanzar un mínimo de 80% en su instrumento de evaluación.

Las cifras actuales del COMAEM revelan que de 158 escuelas de medicina del país (67 públicas y 91 privadas) solo están acreditadas el 47% (75 escuelas; de las cuales 43 son públicas) (Consejo Mexicano para la Acreditación de la Educación Médica A. C. (COMAEM), 2018).

Se ha estudiado la relación entre el desempeño global de una universidad en el ENARM y su estatus de acreditación por parte del COMAEM (Consejo Mexicano para la Acreditación de la Educación Médica, A. C.), observándose (sin significado estadístico) que la probabilidad de ser seleccionado en el ENARM es mayor para los egresados de los programas acreditados que para los egresados de programas no acreditados (Vázquez Martínez & Ortiz-León, 2017).

3.1.2 Regiones geográficas

Otro posible determinante de la calidad en los programas educativos sería la región

geográfica donde se localiza la escuela de medicina o a donde pertenece el examinado; esto se relaciona con la disponibilidad de recursos, tanto económicos como académicos. Un modelo de segmentación geográfica útil es aquel propuesto por la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES); una organización no gubernamental que reúne a las 191 instituciones públicas y privadas más reconocidas de nuestro país. Estas instituciones se distribuyen alrededor las 32 entidades federativas en México y juntas representan casi el 60% de los estudiantes de educación superior y realizan el 90% de la investigación científica. (ANUIES & Asociación Nacional de Universidades e Instituciones de educación Superior ANUIES, 2018) De acuerdo con la organización mencionada, la regiones de México son Noroeste, Noreste, Centro-Occidente, Área metropolitana (Ciudad de México y sus alrededores), Centro-Sur y Sur-Sureste; estas podrían constituir un predictor geográfico del desempeño.

3.2 EXÁMENES DE ALTAS CONSECUENCIAS

Generalmente, los exámenes de altas consecuencias son conocidos por la población general gracias a la difusión realizada por los medios de comunicación. Asimismo, el tipo de exámenes más fácilmente identificados por la sociedad son aquellos que tratan con la evaluación de la educación básica. En este último caso se encuentran por ejemplo las pruebas ENLACE, Excale y Pisa (“Tendencias en el aprendizaje de la educación en México: Una comparación entre ENLACE, Excale y Pisa | Nexos,” 2014.). Este tipo de exámenes por lo regular revisten un carácter ambiguo ya que algunos sectores de la sociedad ven con buenos ojos su ejecución, mientras otros sectores los consideran herramientas de fiscalización y de control político.

De acuerdo a la *Enciclopedia of Science Education*, los exámenes de altas consecuencias se etiquetan de tal manera ya que “acarrear serias consecuencias para los estudiantes y/o

educadores. Normalmente se realizan en forma de pruebas de desempeño estandarizadas y son pruebas comunes dentro de países, estados o ciudades. Los resultados del examen impactan a los estudiantes en términos de si obtienen un diploma o certificado especial o logran el pase o un nivel educativo mayor. Existe un impacto en las escuelas, en los maestros o educadores incluyendo compensaciones económicas, financiamientos, percepciones públicas, estatus social, o inclusive sanciones.” (Cheng, 2014).

Al respecto de la definición, “en español se han usado los términos «examen de alto impacto» o de «altas consecuencias», por el importante efecto y costo que tiene para las personas el acreditar o no un resultado suficiente” (Sánchez Mendiola, 2017).

Una definición más sencilla de “examen de alto impacto”, alineada con estándares de evaluación internacional es la utilizada por el Instituto Nacional para la Evaluación de la Educación en México en la que un examen de esta naturaleza “se indica cuando los resultados del instrumento tienen consecuencias importantes para las personas o las instituciones; por ejemplo, en los procesos de admisión o certificación” (AERA, APA, & NCME, 2014; Sánchez Mendiola & Delgado Maldonado, 2017).

En el ámbito de la educación superior, y específicamente, en Evaluación en Educación Médica, los exámenes de altas consecuencias se han identificado en: exámenes de ingreso a licenciatura, exámenes departamentales, exámenes de certificación en inglés, exámenes profesionales para egreso de licenciatura, exámenes de ingreso a residencia médica (especialidad) y exámenes de certificación de especialidad (Sánchez Mendiola & Delgado Maldonado, 2017).

Aunque no se menciona de manera explícita en las definiciones previas, existe también un impacto general a la sociedad en tanto los exámenes y la educación que estos evalúan está pagada con los impuestos de todos los mexicanos. Esta cuestión sobre la utilización de recursos en educación es añeja y variada; en ocasiones se ha discutido incluso la pertinencia

de ciertas técnicas educativas para la utilización óptima de recursos que, debemos recordar, son limitados (Bowles, 1967). Además de lo anterior, en mi opinión personal, estos exámenes son el único elemento, después de la licenciatura que permite una ponderación homogénea de los graduados de distintas universidades ya que, como ocurre hasta el momento en el país, además de los nombres de las carreras, los requisitos de admisión, los planes y programas de estudio, los requisitos de egreso, los criterios de promoción de semestre o año, los requisitos de graduación y los rangos de calificaciones otorgadas son distintos entre las escuelas de medicina del país (Birnbaum, 1977). De hecho, un aspecto que a simple vista y por sentido común no debería contemplar gran variación entre escuelas, como la duración de la carrera, reviste una variación que puede ser tan amplia como de un año. Esta función de patrón de referencia de los exámenes de altas consecuencias sobre el que se pueden comparar distintas escuelas es una de las premisas sobre las que se basa su reconocimiento.

3.3 SELECCIÓN PARA EL SISTEMA NACIONAL DE RESIDENCIAS EN MÉXICO.

Un estudio reciente del Programa de Seguimiento de Alumnos y Egresados de la Facultad de Medicina de la Universidad Nacional Autónoma de México (UNAM) caracterizó a alumnos que concluían su servicio social; de estos el 91.6% manifestó interés en realizar una especialidad, el 6.8% un maestría y el 1% un doctorado. El 30.4% no había intentado ingresar en ninguna especialidad; de los restantes, el 69.6% ya había realizado el Examen Nacional para Residencias Médicas (ENARM), siendo aceptado el 49.6%. Un 2.7% fue seleccionado para maestría y un 1.9% para estudiar en el extranjero (Ángel Fernández-Ortega, Ortiz-Montalvo, Ponce-Rosas, Fajardo-Ortiz, & Mazón-Ramírez, 2016).

En Estados Unidos, aunque no ha sido diseñado para tal fin, los resultados del USMLE son comúnmente utilizados como un parámetro de tamizaje para seleccionar a potenciales residentes en programas de especialidad (Gelinne et al., 2018). Esta postura de utilizar un

examen diseñado originalmente para fines de licencia médica (entiéndase como titulación) con fines de selección para residencia médica ha sido criticada por su “falta de estructura, coherencia y validez” (McGaghie, Cohen, & Wayne, 2011).

En nuestro medio, a pesar de haber obtenido desempeño satisfactorio en el examen de titulación y/o en el EGEL, para ingresar a los cursos de especialidad médica se requiere, además del título de licenciatura en medicina, colocarse en el orden descendente de los sustentantes de un examen cognitivo de selección del total nacional de aspirantes (Examen Nacional de Aspirantes a Residencias Médicas, ENARM), según la especialidad elegida, que aplica una Comisión Interinstitucional e intersectorial para formar recursos humanos para la salud (Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud, CIFRHS) y que corresponden al número de plazas ofrecidas en cada disciplina por las instituciones receptoras de residentes. Al examen concurren anualmente un poco más de 35,000 aspirantes de todo el país. Habitualmente, sólo se ofrecen alrededor de 7,000 plazas, por lo que queda excluido un gran número de aspirantes, en 2016 solo tuvo acceso a las residencias el 22% (Ramiro-H, Cruz-A, Zerón-Gutiérrez, & Arévalo-Vargas, 2017). Este examen está bajo la responsabilidad del Comité de Enseñanza de Posgrado y Educación Continua (CEPEC) de la Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud (CIFRHS); consiste en una evaluación apegada a norma que emite un rango de puntaje mínimo y puntaje máximo sobre el que se basa la selección de los aspirantes; estos puntajes varían año tras año. El objetivo general de dicho examen es “seleccionar a los médicos generales mexicanos y extranjeros que aspiren a realizar una residencia médica en un curso universitario de especialidad” y su objetivo específico es “realizar la selección de los médicos generales aspirantes a cursar una residencia médica mediante un proceso equitativo, transparente, legal y profesional, a través de la aplicación de un instrumento objetivo y válido de medición de conocimientos en medicina general”. En estos objetivos, médico general alude

a una persona que ha realizado los estudios de nivel licenciatura de medicina independientemente de la denominación en la carrera o el título (CIFRHS-CPEC., 2017).

El ENARM es un examen sumativo de altas consecuencias y sus resultados son determinantes, ya que define, de alguna manera, el proyecto de vida profesional de miles de aspirantes. Para ello los médicos se preparan de diversa manera para responder el examen, esto nos permite observar características sobre el desempeño académico de los sustentantes y valorar el grado de conocimientos con el que cuentan los aspirantes para ingresar a la especialidad e identificar los conocimientos en los que tienen fortaleza y aquellos con debilidad (Rodríguez Carranza, Gatica-Lara, Martínez-González, Rosa Medina-Martínez, & Rodríguez-Carranza, 2013).

Año con año, se entrega un informe con los resultados del desempeño académico obtenido en las distintas áreas de conocimiento del ENARM a cada sustentante de manera confidencial. Por su importancia, su elaboración está a cargo del Comité de Posgrado y Educación Continua (CEPEC) de la Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud. Según esta comisión, el propósito primario del ENARM es el seleccionar a los médicos generales mexicanos y extranjeros que aspiren a realizar una residencia médica en un curso universitario de especialidad; también expresa que se trata de un proceso equitativo, transparente, legal y profesional, a través de un instrumento objetivo y válido de medición de conocimientos en medicina general que explora la capacidad para la correcta utilización de los conocimientos médicos de ciencias clínicas, ciencias básicas y salud pública (Dirección General De Calidad y Educación En Salud, 2010). El ENARM explora conocimientos de medicina general clasificados en las áreas de conocimiento de especialidad troncal: Medicina Interna, Pediatría, Gineco-Obstetricia y Cirugía. Desde el año 2014 el ENARM se ha estado conformando por 450 reactivos con formato de casos clínicos que tienen de una a tres preguntas cada uno; con una sola opción correcta basada en la modalidad de la

mejor respuesta y 3 distractores. 405 de las preguntas con casos clínicos en español y 45 preguntas con casos clínicos en inglés que también evalúan conocimientos médicos. El examen está estructurado en una versión diferente para cada día de aplicación, “con las mismas características de objetividad, calidad, balance y nivel académico”. (CIFRHS-CPEC, 2018) Se aplica en formato electrónico con un tiempo de respuesta de un minuto por pregunta (CIFRHS-CPEC., 2017). El ENARM se diseña a partir de una tabla de especificaciones elaborada por un grupo académico, quien toma como base los temas o aspectos comunes de los planes y programas de estudio de las facultades y escuelas de medicina de nuestro país, la morbilidad y mortalidad en México, las tendencias de la medicina y la salud, los programas nacionales de salud, y las observaciones y experiencias obtenidas en los exámenes previos (Rodríguez Carranza et al., 2013). Con base en dicha tabla de especificaciones se determina la distribución del número y porcentaje de preguntas por especialidad troncal, en todos los casos del total de reactivos. La versión definitiva del ENARM es producto de etapas sucesivas de depuración y de generar nueve reordenamientos del instrumento para su posterior aplicación (Rodríguez Carranza et al., 2013).

3.3.1 SISTEMA DE SELECCIÓN PARA RESIDENCIAS: FORTALEZAS.

Una de las fortalezas de este sistema nacional de residencias es que el examen de selección es el mismo para todos los médicos aspirantes y para todas las especialidades médicas de entrada directa. (Vázquez Martínez & Ortiz-León, 2017) (Mendoza-Aguilar, Salazar-Ruiz, Cardiel-Marmolejo, & Roldan-Valadez, 2020)

3.3.2 SISTEMA DE SELECCIÓN PARA RESIDENCIAS: DEBILIDADES.

Las características actuales del sistema nacional de residencias en México lo hacen objeto de crítica, por ejemplo que subordina el proceso educativo al trabajo, e imposibilita el paso hacia la etapa escolar (perspectiva de evolución educativa) y el despliegue de todo el

potencial existente en el país para la formación de médicos especialistas (Vázquez Martínez, 2017). A pesar de ser el único instrumento con el que se cuenta para seleccionar a los médicos que desean especializarse, el ENARM ha sido también objeto de críticas diversas entre las que se mencionan: que no todos los alumnos que lo contestan tienen las mismas posibilidades de ser seleccionados quizá, entre otros factores, porque la distribución de escuelas o facultades de medicina es inequitativa, por la disparidad de la enseñanza de la Medicina, porque sólo 47% de las instituciones que forman médicos están acreditadas, por la poca profesionalización de los docentes, o por la aplicación de técnicas inadecuadas de enseñanza (Prieto-Miranda, Jiménez-Bernardino, Jiménez-Bernardino, & Esparza-Pérez, 2013) (Consejo Mexicano para la Acreditación de la Educación Médica A. C. (COMAEM), 2018). Otro de los aspectos que puede considerarse una debilidad es el alto costo de su aplicación, el cual ronda en 110 millones de pesos anuales (Barajas-Ochoa & Ramos-Remus, 2017) y su poca concordancia con el gasto real de aplicación, el cual se realiza en una modalidad de transferencia de recursos presupuestarios federales con el carácter de subsidios, para la aplicación del examen, que celebran por una parte, el ejecutivo federal, por conducto de la secretaría de salud y por la otra las entidades federativas representadas por el gobernador del estado donde hay sedes de aplicación (Diario Oficial de la Federación, 2018).

Por último, pero no menos importante, la validez misma del examen ha sido cuestionada, encontrándose, entre otros, los siguientes defectos: a) no existe un reporte técnico de su elaboración; b) no participa un editor profesional; c) la inexistencia de prueba piloto de los reactivos en la población blanco; d) la asignación cuasi-arbitraria del grado de dificultad de los reactivos sin utilizar teorías psicométricas; e) la calificación del examen (cuando no hay empates) considera a los reactivos como iguales a pesar de que tienen una dificultad diferenciada; f) no se calcula el error estándar de la medición y g) se compara a los sustentantes sin utilizar procesos de equiparación de pruebas, a pesar de que se utilizan

distintos reactivos cada día de aplicación (Barajas-Ochoa & Ramos-Remus, 2017). Esto último contrasta con la tendencia en otras esferas, por ejemplo el examen MIR de España, siendo que a partir del análisis del examen y sus resultados en varios años se ha podido establecer que existen variaciones en su grado de dificultad y nivel de discriminación dependiendo de la edición de que se trate. Este tipo de análisis permiten conservar la calidad en este tipo de exámenes de altas consecuencias (Baladrón et al., 2018).

3.4. INFERENCIAS A PARTIR DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.

Entendiendo que la educación médica en México es diversa o heterogénea, ya que existen escuelas y facultades de medicina pública y privada, con características particulares, lo que implica diversidad en los programas académicos así como en los procedimientos de elección y de evaluación de los alumnos (Akaki Blancas & López Bárcena, 2018) se ha intentado utilizar un parámetro evaluador que sea aplicado a todas las escuelas del país para establecer conclusiones; tal parámetro ha sido el ENARM ya que es sabido que la tasa de ingreso de los Médicos Cirujanos a una especialidad se ha convertido en un indicador “no oficial” de calidad de los planes y programas de estudio de esta licenciatura (Jiménez-Sastré Alejandro, Quevedo-Tejero Elsy del Carmen, Hernández-Ortega Hilda María, Fócil-Némiga Eduardo, 2015).

En este contexto, en México son escasos los estudios que identifiquen variables académicas cuya influencia puede facilitar el ingreso a una especialidad médica. Este tipo de investigaciones permitiría la realimentación a los candidatos y proporcionaría información relevante para el diseño e implementación de intervenciones educativas que fortalezcan los planes y programas educativos y que fomenten la calidad. A los parámetros proporcionados por el ENARM se les puede contrastar el promedio general de egreso obtenido por el estudiante que refleja no solo las evaluaciones sumativas al final de cada ciclo escolar sino

también las evaluaciones formativas parciales realizadas desde el primer año de ingreso. Puesto que todos estos parámetros reflejan de manera intuitiva la calidad de la educación médica proporcionada en una escuela se han estudiado de manera fragmentada y con la más diversa metodología para poder arrojar conclusiones (Ahumada-Ayala M, Calva-Rodríguez, Galarza Delgado DA, García-Alcalá H, Lustig N, 2006) (Guerrero-Avila, Romero-González, & Noriega, 2013).

Por lo regular, a los alumnos que aprueban el ENARM se les considera los más destacados en relación con el promedio general de la carrera y como los más destacados en los exámenes CENEVAL (en las escuelas de medicina que lo presentan) (Ahumada-Ayala M, Calva-Rodríguez, Galarza Delgado DA, García-Alcalá H, Lustig N, 2006). Sin embargo, hasta el momento no existe una forma de determinar la correlación entre el ENARM y el desempeño durante la carrera de medicina general expresada como un promedio de calificaciones.

Una perspectiva un poco más amplia sería considerar que la educación médica en el país no puede limitarse a ser un inmenso curso propedéutico para el examen nacional de ingreso a las residencias médicas; cuando de antemano conocemos que menos de una cuarta parte de los aspirantes encontrará una plaza como médico residente (García Acosta & Crocker Sagastume, René; Abreu Hernández, Luis Felipe; Parra Acosta, Haydeé; Vázquez Esquivel, Jesús; Ojeda Blanco, Carlos; Alomía, Harold; Contreras Mayén, Gustavo; Ramos Herrera, Igor; Rodríguez Arroyo, Christian; Vázquez Castellanos, 2015). Además la especialización en el terreno asistencial no es la única “vía terminal” para los egresados de la carrera de medicina, puesto que sus egresados pueden optar por dedicarse a la investigación y la docencia, accediendo a los programas de posgrados orientados a ella, de reconocida calidad por el Consejo Nacional de Ciencia y Tecnología, que actualmente suman más de 100 Maestrías y Doctorados en el área de Medicina y Ciencias de la Salud (Jiménez-Sastré Alejandro, Quevedo-Tejero Elsy del Carmen, Hernández-Ortega Hilda María, Fócil-Némiga Eduardo, 2015). ¿Qué

papel juega el ENARM para estos médicos? ¿Podría jugar algún papel como aval de licencia médica?

Además de todo lo anterior, la sociedad tiene derecho a conocer los aspectos sobre la competencia de los profesionistas médicos, de hecho sobre la competencia de todos los profesionistas (Shumway, Harden, Shumway, & Harden, 2009).

3.5. PERSPECTIVAS SOBRE LA UTILIDAD DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.

Al momento, la mayoría de las escuelas de medicina del país son privadas (Consejo Mexicano para la Acreditación de la Educación Médica A. C. (COMAEM), 2018) y se rigen de manera interna. Entre las escuelas públicas, que podrían ser consideradas como parte del Estado (sin embargo desde 1929 se les han concedido grados variables de autonomía frente al Estado) más allá de su dependencia financiera de este último, ejercen su autonomía en su gobierno interno y en la aprobación y certificación de los programas de estudio (Frenk-Mora J, Robledo-Vera C, Nigenda-López G, Ramírez-Cuadra C, Galván-Martínez O, 1990).

Puesto que no existe una manera de homologar el promedio de calificaciones que otorga cada una de las escuelas o facultades de Medicina del país y que esta calificación (suma de evaluaciones sumativas) sea establecida con los mismos parámetros a cada uno de los graduados de medicina del país; lo anterior como resultado de la heterogeneidad, y que este promedio impacta el acceso del estudiante (una vez aprobado el ENARM) a la institución de su elección para formarse como médico especialista y también este promedio funge como criterio de aceptación en caso de aplicar en convocatorias para becas o estudios de posgrado (Consejo Nacional de Ciencia y Tecnología, n.d.) una solución justa sería utilizar un parámetro válido que se aplique a todos los graduados de las escuelas de medicina sin importar la localización geográfica o el programa de estudios, sin que influya si la escuela es pública o privada o el año de titulación; aunque sabemos que la aplicación del ENARM es para

seleccionar médicos para especialidad médica, la constitución de un examen similar sería justa e igualitaria.

Desde otro punto de vista, los requisitos para obtener licencia médica (entendiéndose como titulación u obtención del grado) varían de escuela a escuela; algunas instituciones educativas de otros países, por ejemplo India, han contemplado la utilización de su examen para cursos de especialización y posgrado (National Eligibility-cum-Entrance Test for postgraduate courses) con un triple propósito: i. Examen requisito para todos los cursos de posgrado en todo el territorio de la India, ii. Examen requisito para obtener licencia médica para graduados nacionales y iii. Examen requisito para médicos que desean ejercer en territorio de la India pero graduados en escuelas extranjeras. Uno de los aspectos “negativos” de una evaluación de tal naturaleza es que socavaría el peso de las evaluaciones universitarias ya que no revestirían mayor peso que constituir un requisito para presentar el examen unificado (Mahajan & Singh, 2017). El sentido común, partiendo de que un médico general debe saber lo mismo independientemente de la escuela donde se graduó (ciudad, país e idioma) y de si va a ejercer como médico general o va a especializarse, dicta que un examen unificado es deseable ya que se optimiza el uso de los recursos y se evitan exámenes redundantes que si no evalúan lo mismo deberían hacerlo. Además, debe considerarse que, tras presentar el ENARM menos de una cuarta parte de los aspirantes encontrará una plaza como médico residente: la mayoría terminarán ejerciendo la medicina general (García Acosta & Crocker Sagastume, René; Abreu Hernández, Luis Felipe; Parra Acosta, Haydeé; Vázquez Esquivel, Jesús; Ojeda Blanco, Carlos; Alomía, Harold; Contreras Mayén, Gustavo; Ramos Herrera, Igor; Rodríguez Arroyo, Christian; Vázquez Castellanos, 2015), otro aspecto que resalta la potencial utilidad de un examen unificado.

3.6. VALIDEZ DEL EXAMEN NACIONAL DE ASPIRANTES A RESIDENCIAS MÉDICAS.

De acuerdo a los Standards for Educational and Psychological Testing (AERA, APA, NCME, 2014) existen cinco fuentes de evidencia sobre la validez de una prueba: i) Contenido del examen, ii) Proceso de respuesta, iii) Estructura interna, iv) Relación con otras variables, v) Consecuencias del examen. Es importante mencionar que en la convocatoria 2018 para el ENARM ya no se menciona el atributo de “válido” refiriéndose al examen (CIFRHS-CPEC, 2018).

3.6.1. Contenido del Examen.

El contenido de un examen se refiere a los “temas, fraseo, y formato de los ítems y preguntas” (AERA et al., 2014). De acuerdo a la CIFRHS, en el ENARM, cada reactivo se diseña por al menos 8 profesores expertos, quienes participan conforme a perfiles específicos en las diferentes etapas de elaboración (CIFRHS-CPEC., 2017); sin embargo no se menciona ni la definición conceptual ni operacional de “experto” para participar en el proceso. Asimismo, sobre el contenido explorado, se refiere que “en el ENARM se explora la capacidad para la correcta utilización de los conocimientos médicos de ciencias clínicas, ciencias básicas y salud pública, en el ámbito de la competencia del ejercicio profesional del médico general.”. (CIFRHS-CPEC, 2018) Se ha valorado de manera transversal el contenido del ENARM en cuanto a proporción de reactivos que exploran ciencias básicas. En el ENARM 2005 y 2006, se encontró que (24%) de los reactivos exploraron conocimientos de las ciencias básicas y que este porcentaje es inferior al que corresponde a los créditos que se otorgan a las ciencias básicas en el Plan de Estudios de medicina (31%) (Rodríguez Carranza et al., 2013). Aquí surge la pregunta de en base a cuál plan o programa académico se distribuyen los reactivos de dicho examen. La evidencia de validez orientada al contenido es un componente central del proceso conocido como *alineación* (“alignment”), el cual evalúa la correspondencia entre los estándares de aprendizaje del estudiante y el contenido del examen.

3.6.2. Proceso de respuesta.

Algunas interpretaciones de constructo implican suposiciones más o menos explícitas sobre los procesos cognitivos que se llevan a cabo en los sujetos que responden el examen. El análisis del grado en que ciertas habilidades, irrelevantes o subordinadas al constructo, pueden influenciar diferencialmente el desempeño de los sujetos; por ejemplo las habilidades necesarias para responder un examen en formato electrónico. El registro de otros aspectos del desempeño, como los tiempos de respuesta, también puede ser relevante para ciertos constructos (AERA et al., 2014). De acuerdo a la CIFRHS, el examen ENARM "se aplica en formato electrónico a través de una red local instalada en cada sede de aplicación, la que cuenta con una computadora personal para cada sustentante. Para la aplicación, los sustentantes deberán consultar el instructivo para el registro, ingreso y aplicación del examen"; esto puede homogeneizar en cierta medida el proceso de respuesta. Sin embargo, de acuerdo al mecanismo de inscripción y aplicación del examen, en 2018 se contó con 13 sedes para sustentar el mismo (CIFRHS-CPEC, 2018), poniendo en desventaja a aquellos sujetos que debieron trasladarse una mayor distancia, lo cual se contrapone a los estándares de justicia para la aplicación de un examen: "los sujetos que contestan un examen deben recibir un tratamiento comparable durante su aplicación" (estándar 3.4) (AERA et al., 2014). Esta fuente de evidencia por lo regular se determina con pruebas estadísticas de tipo *confiabilidad interobservador* y *correlaciones intraclase* (Bajwa, Yudkowsky, Belli, Vu, & Park, 2018).

3.6.3. Estructura interna.

Los estudios para valorar la estructura interna se diseñan para mostrar si algunos ítems en particular funcionan de manera diferente para subgrupos identificables de sujetos (p. Ej. subgrupos raciales, étnicos o de género). El *funcionamiento diferencial de ítems* ocurre cuando grupos diferentes de sujetos, con habilidades globales similares, o que comparten un estatus similar de acuerdo a un criterio apropiado, tienen, en promedio, respuestas

sistemáticamente diferentes para un ítem (AERA et al., 2014). Considerando que el ENARM puede clasificarse como un examen esencialmente unidimensional (sin olvidar que evalúa conocimientos médicos y dominio de inglés), su estructura interna podría defenderse por medio de un análisis estadístico multivariado (p. Ej. Análisis factorial) (AERA et al., 2014). No existe evidencia o difusión de los aspectos antes mencionados con respecto al ENARM. Como parte de la estructura interna de un examen se analizan también las *especificaciones psicométricas*. Estas indican las propiedades estadísticas deseadas de los reactivos (p. Ej. dificultad, discriminación y correlaciones inter-ítem), así como las propiedades estadísticas deseadas del examen en conjunto, incluyendo la naturaleza de la escala de reporte, la dificultad y precisión del examen, y la distribución de los reactivos a través de las categorías cognitivas (AERA et al., 2014). Al respecto del ENARM, se ha señalado la asignación la asignación cuasi-arbitraria del grado de dificultad de los reactivos sin utilizar teorías psicométricas y que la calificación del examen (cuando no hay empates) considera a los reactivos como iguales a pesar de que tienen una dificultad diferenciada (Barajas-Ochoa & Ramos-Remus, 2017). Además de esto, en caso de empate, “el criterio de desempate considera el puntaje más alto de acuerdo al área médica troncal, como sigue: solo Medicina Interna, en caso de empate se considera solo Pediatría, en caso de empate se considera solo Gineco-obstetricia y en caso de empate se considera solo Cirugía” (*Comité de Posgrado y Educación Continua XLI Examen Nacional para Aspirantes a Residencias Médicas 2017 Metodología Proceso de Selección*, n.d.). No existen explicaciones que fundamenten el criterio anterior. El *alfa de Cronbach* es de utilidad para establecer este criterio de validez (Bajwa et al., 2018).

3.6.4. Relación con otras variables.

En muchos casos, la interpretación para la que está destinada la prueba implica que el constructo debiera estar relacionado a algunas otras variables y, como resultado, el análisis de la relación entre los puntajes del examen y otras variables externas al examen, pueden

significar una importante fuente de evidencias de validez. Las variables externas pueden incluir mediciones de algunos criterios que se espera que el examen sea capaz de predecir (relaciones predictivas y concurrentes), así como las relaciones con otros exámenes que hipotéticamente miden los mismos constructos (evidencia convergente), y relaciones con exámenes que miden constructos diferentes o no relacionados (evidencia discriminante). Las relaciones entre métodos diferentes para medir el mismo constructo pueden ser especialmente útiles para elaborar la interpretación y el significado de los resultados (AERA et al., 2014). En México se ha estudiado la relación del ENARM con otras variables que hipotéticamente miden el mismo constructo: Examen CENEVAL, Examen de conocimientos generales y Promedio general de la carrera; sin embargo esta comparación se realizó solo para una universidad y expresando el resultado del ENARM de manera categórica (Seleccionado, No seleccionado) encontrando que la categoría “Seleccionado” tenía mayores puntajes en las otras tres variables. (Ahumada-Ayala M, Calva-Rodríguez, Galarza Delgado DA, García-Alcalá H, Lustig N, 2006). Se ha estudiado también, en el Centro Universitario de Ciencias de la Salud de la Universidad de Guadalajara, la relación del ENARM con algunas variables sociodemográficas: edad, género, estado civil, generación de egreso y tiempo entre la conclusión de los estudios y la presentación del ENARM. En base a este estudio se concluyó que el ser varón, soltero y menor de 30 años son factores positivos para ser seleccionado. Asimismo entre más tiempo transcurra entre el ingreso y entre el egreso, y entre el egreso y la presentación del ENARM, más reducidas serán las probabilidades de resultar seleccionado u obtener un promedio competitivo (Romero-Valle & González-Pérez, 2015). El análisis mencionado solo tomó en cuenta las tasas de selección y el promedio general de conocimientos médicos de manera global; no se exploró la relación entre calificaciones de la licenciatura y este promedio o calificación ENARM. La relación entre variables cuantitativas como criterio de validez se puede establecer mediante el *coeficiente correlación de Pearson*

(Bajwa et al., 2018).

3.6.5. Consecuencias del examen.

Algunas consecuencias el examen, se derivan directamente de las interpretaciones de los puntajes del mismo para usos destinados por el desarrollador. La administración de los exámenes comúnmente se realiza con la expectativa de que algún beneficio se obtendrá del uso y la interpretación de los puntajes obtenidos; por ejemplo, en un examen para decisiones de colocación en una empresa o sistema, la evidencia de validez radicaría en que efectivamente los puestos de colocación fueran diferencialmente benéficos para las personas y la institución (es decir, unos más buenos que otros, de otra manera no habría motivo para el examen) (AERA et al., 2014). No existe evidencia en México de que los puntajes del ENARM estén relacionados con plazas de residente en buenos o malos hospitales (haría falta definir estas categorías); lo anterior puede deberse a que en el momento actual no escoge primero el sustentante con el puntaje más alto y así sucesivamente, sino que, de acuerdo a las características del sistema actual, los procesos de selección se dan muchas veces de manera simultánea y la sede definitiva para cursar la especialidad puede determinarse por suerte, cercanía geográfica u otros factores no caracterizados. Podría estudiarse la agrupación de puntajes ENARM similares en hospitales representativos para estudiar este criterio de validez; esto es solo una hipótesis.

Este criterio de validez (consecuencias) es de particular importancia, ya que ni siquiera un examen como el USMLE de Estados Unidos ha logrado fundamentarlo de manera consistente; por ejemplo, un estudio reciente ha demostrado que los puntajes en el USMLE 1 no guardan relación con la producción científica, la obtención de fondos gubernamentales o la pertenencia al *American Board of Neurological Surgery* en egresados de la especialidad de Neurocirugía (Gelinne et al., 2018). Las consecuencias (o predicciones) de un examen por lo regular hacen uso de la estadística proporcionada por la *regresión logística* (Bajwa et al.,

2018).

MATERIAL Y MÉTODOS

4. PREGUNTA DE INVESTIGACIÓN

La pregunta de investigación general contempla múltiples aspectos relacionados con el sustentante al momento de presentar el ENARM:

¿Hay relación entre el desempeño académico obtenido por los sustentantes en el ENARM y variables demográficas y académicas?

De la cual podemos derivar una pregunta más circunscrita:

¿Cuál es la relación entre el promedio obtenido en la licenciatura y el desempeño en el ENARM?

Y agregar una serie de preguntas de investigación relacionadas con las características de la escuela de procedencia (reflejado esto en variables demográficas y académicas):

- ¿Hay diferencia en el desempeño y tasa de selección en el ENARM entre sustentantes provenientes de escuelas *acreditadas* comparado con escuelas *no acreditadas*?
- ¿Hay diferencia en el desempeño y tasa de selección en el ENARM entre sustentantes provenientes de escuelas *públicas* comparado con escuelas *privadas*?
- ¿Hay diferencia en el desempeño y tasa de selección en el ENARM entre sustentantes provenientes de *diferentes regiones* del país?

5. PLANTEAMIENTO DEL PROBLEMA

Hasta ahora no existe una forma de determinar correlación entre el ENARM y el desempeño durante la carrera de medicina general expresada como un promedio de calificaciones.

6. JUSTIFICACIÓN

A pesar de la importancia que reviste para los individuos, las escuelas de medicina, los organismos de acreditación y la sociedad en general, aún no existen maneras de determinar la relación entre los diversos métodos para evaluar a un egresado de la carrera de medicina. Algunos de ellos a pesar de fijar sus puntos de corte de manera poco clara y arbitraria son de altas consecuencias para los egresados de la carrera de medicina general.

El conocer la correlación entre estas evaluaciones sumativas permitirá posiblemente prescindir de algunas de ellas o unificarlas en una sola con el propósito de optimizar recursos y brindar más seriedad a los organismos acreditadores o evaluadores.

Analizar la relación entre el desempeño obtenido en la carrera y la calificación en el ENARM puede constituir evidencia de validez para este último (*relación con otras variables*). Es importante señalar que se puede llevar a cabo el proceso inverso: proporcionar evidencias de validez a las calificaciones otorgadas por las escuelas de medicina del país; es decir, las escuelas de medicina cuyos promedios de egreso tengan mayor correlación con la calificación ENARM estarían otorgando las calificaciones que más se apegan a la realidad del alumno.

Por otro lado, el conocer con detalle la relación que existe entre el ENARM y otras variables, tales como el desempeño obtenido por los sustentantes durante su formación en pregrado o sus características socioeconómicas, impactaría de manera positiva en la opinión que guarda la sociedad y la comunidad médica sobre el examen al respecto de los atributos de equidad, transparencia, objetividad y validez.

7. HIPÓTESIS

Podemos manejar una hipótesis general al respecto de los datos globales a interpretar: Si las diferentes evaluaciones sumativas que analizaremos tienen una correlación adecuada entonces observaremos correspondencia en los datos pertinentes a cada una de ellas y las

inferencias que se pueden realizar en consecuencia (validez en el rubro de *relación con otras variables*).

Existen sin embargo hipótesis que consideramos pueden abarcar de manera más limitada nuestro objeto de estudio:

1. A mayor promedio obtenido en la licenciatura por los médicos sustentantes mayor desempeño obtenido en el ENARM.

2. A mayor promedio obtenido en la licenciatura por los médicos sustentantes mayor tasa de selección obtenida en el ENARM.

Podemos agregar:

3. Los médicos sustentantes provenientes de escuelas y facultades *acreditadas* tendrán mejor desempeño en el ENARM comparados con sustentantes provenientes de instituciones *no acreditadas*.

4. Los médicos sustentantes provenientes de escuelas y facultades *acreditadas* tendrán una mayor tasa de selección en el ENARM comparados con sustentantes provenientes de instituciones *no acreditadas*.

5. Los médicos sustentantes provenientes de escuelas y facultades *públicas* tendrán mejor desempeño en el ENARM comparados con sustentantes provenientes de instituciones *privadas*.

6. Los médicos sustentantes provenientes de escuelas y facultades *públicas* tendrán una mayor tasa de selección en el ENARM comparados con sustentantes provenientes de instituciones privadas.

7. Los médicos sustentantes provenientes de la *región central* del país tendrán un mejor desempeño en el ENARM que aquellos que provienen de las regiones norte y sur.

8. Los médicos sustentantes provenientes de la *región central* del país tendrán una mayor tasa de selección en el ENARM que aquellos que provienen de las regiones norte y sur.

8. OBJETIVOS

8.1 OBJETIVO GENERAL.

Identificar la relación entre el promedio obtenido en la licenciatura y el desempeño en el ENARM (calificaciones y tasa de selección), segmentando de manera particular al grupo de médicos recién graduados, y determinar las características de la escuela de medicina que impactan este desempeño.

8.2 OBJETIVOS ESPECÍFICOS.

- Determinar de manera general el desempeño obtenido en el ENARM para una muestra que comprende cinco años estableciendo un parámetro para realizar comparaciones.
- Determinar la relación entre el promedio obtenido en la licenciatura y las calificaciones del ENARM en médicos recién graduados.
- Determinar la diferencia en el desempeño en el ENARM entre sustentantes egresados de escuelas acreditadas comparado con sustentantes egresados de escuelas no acreditadas.
- Determinar la diferencia en el desempeño en el ENARM entre sustentantes egresados de escuelas públicas comparado con sustentantes egresados de escuelas privadas.
- Identificar diferencias en el desempeño en el ENARM entre sustentantes egresados de escuelas localizadas en diferentes regiones geográficas del país.
- Evaluar como criterio de validez la relación recíproca de las variables calificación ENARM y promedio de licenciatura.

9. METODOLOGÍA

9.1 CONSIDERACIONES ÉTICAS

Nos ceñiremos a la normativa ética vertida en las guías “International Ethical Guidelines for Epidemiological Studies” editadas por el “Council for International Organizations of Medical Sciences” (CIOMS) en 2009; las guías “International Ethical Guidelines for Health-related Research Involving Humans” editadas por el “Council for International Organizations of Medical Sciences” (CIOMS) en 2016, las guías “Ethical Guidelines for Social Science Research in Health” publicadas en el año 2000, así como a las normas y lineamientos éticos de las instituciones educativas y gubernamentales involucradas. El protocolo de estudio fue aprobado por el Comité de Ética del Programa de Posgrado (Programa de Maestría y Doctorado en Ciencias Médicas, Odontológicas y de la Salud) de la Universidad Nacional Autónoma de México (UNAM). Se preservaron el anonimato y la confidencialidad.

9.2 CONSIDERACIONES FINANCIERAS Y POTENCIALES CONFLICTOS DE INTERÉS

Los recursos necesarios para llevar a cabo el proyecto de investigación consisten en conexión a internet, computadora personal, dispositivos de memoria portátil USB (universal serial bus), impresora, papel para impresión, traslados, entre otros. Estos recursos serán solventados mediante aportaciones personales por parte de los investigadores, recursos de COMAEM y recursos de la Dirección de Evaluación Educativa, UNAM. No se ha identificado ningún conflicto de interés.

9.3 ESTRATEGIA DE BÚSQUEDA BIBLIOGRÁFICA

Se realizó una búsqueda sistemática de información en español en los buscadores o bases de datos Pubmed, Google académico y ERIC utilizando las siguientes palabras clave:

ENARM, Examen Nacional, Validez, Psicometría, Residencias médicas, Médicos generales, MIR, Examen Nacional de Residencias. Asimismo, se realizó una búsqueda sistemática de información en inglés en los buscadores o bases de datos Pubmed, Google scholar y ERIC utilizando las siguientes palabras clave: USMLE, MIR, MCCEE, Validity, Medical resident, Medical residency, Correlation. Medical residency admission test. Se seleccionaron aquellos artículos en los que el resumen reflejara específicamente la validez del proceso de selección de residencias y que la metodología fuera pertinente a nuestra pregunta de investigación. Utilizando los artículos seleccionados y sus referencias se identificaron documentos de interés para nuestro proyecto. La bibliografía inicial consistió en 32 fuentes consultadas (22 en español, 10 en inglés) distribuidas de la siguiente forma: 1 libro, 19 artículos, 1 carta al editor, 2 Informes CIFRHS (2017 y 2018), 1 página web CPEC 2017, 1 página web e informe del COMAEM, 1 página web e informe del CONACYT (Consejo Nacional de Ciencia y Tecnología), 1 página web del DOF (Diario Oficial de la Federación), 1 página web e informe de la DGCES/SSA (Dirección General de Calidad y Educación en Salud), 1 informe de la AMFEM 2015 (Asociación Mexicana de Facultades y Escuelas de Medicina), 1 página web e informe del CENEVAL, 1 informe del CUCS/UdeG (Centro Universitario de Ciencias de la Salud/Universidad de Guadalajara), 1 guía AMEE (Association for Medical Education in Europe) Número 25. Esta bibliografía se robusteció durante la sustentación de las pruebas estadísticas y la discusión para constituir finalmente 70 fuentes consultadas.

9.4 DISEÑO DEL ESTUDIO.

Puesto que en el presente proyecto de investigación no se realizarán intervenciones se puede denominar un estudio observacional. Estas observaciones se llevan a cabo después de realizar las descripciones pertinentes del objeto de estudio, por ejemplo las variables demográficas. Sin embargo, esta clasificación dicotómica entre observacional como opuesto a experimental está derivada principalmente del paradigma de los estudios clínicos sobre los

que se definen los diseños de los proyectos de investigación (Manterola & Otzen, 2014). Debido a que nuestro proyecto no se limita solamente a la descripción del fenómeno en cuanto a los datos con los que contamos sino que después de esta descripción y observación sobreviene un análisis de la situación encontrada, podemos abreviar de las clasificaciones utilizadas en otros paradigmas de investigación para otorgar el carácter analítico a nuestro estudio. En este sentido, por ejemplo, en el paradigma de investigación en calidad como el utilizado para definir los ciclos de mejora continua establecidos por Deming, se utiliza una clasificación de estudios analíticos en oposición a enumerativos, definiendo a los primeros como aquellos que establecen relaciones entre las variables, de asociación o de causalidad y en los que pueden probarse hipótesis específicas previas surgidas de un estudio descriptivo (Provost, 2011). Así pues este proyecto es descriptivo, observacional y analítico conforme se avanza en sus fases. Por lo tanto, podemos resumir:

Diseño del estudio – dos etapas; primera etapa: estudio observacional no-experimental; segunda etapa: estudio analítico inferencial.

Contexto – nuestro contexto es el Examen Nacional de Aspirantes a Residencias Médicas (ENARM), un examen de carácter nacional para la admisión a las residencias médicas en México; los candidatos son admitidos a un curso de especialización basado en sus puntajes y la especialidad seleccionada.

Marco temporal y fuente de los datos – la Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud (CIFRHS) proporcionó una base de datos con registros pertenecientes al ENARM en el periodo 2014-2018. Un conjunto específico de datos indicaba el número de respuestas correctas obtenido en el examen.

Instrumento – el ENARM es de aplicación común para cualquier especialidad de entrada directa (el sustentante no requiere cursos de especialización previos).

9.5 DEFINICIÓN DE LAS VARIABLES.

Las variables utilizadas serán aquellas contenidas en la base de datos correspondiente al examen ENARM comprendido entre los años 2014 a 2018, proporcionada por la Dirección General de Calidad y Educación en Salud de la Secretaría de Salud Federal en un acuerdo establecido con el Consejo Mexicano para la Acreditación de la Educación Médica A.C. y la Universidad Nacional Autónoma de México. Dicha base de datos se encuentra bajo resguardo en las oficinas de la Dirección General de Evaluación perteneciente a la Coordinación de Evaluación Educativa e Innovación Curricular de la UNAM. Las variables contenidas en la base de datos corresponden a las respuestas proporcionadas por el sustentante en la cédula de registro respectiva para cada año del examen. Algunas de las variables debieron agregarse y codificarse en la base de datos puesto que consistían en información que no fue proporcionada por el sustentante (V.gr. estatus de acreditación de la escuela de medicina al momento de realizar el examen).

9.5.1 Criterios de inclusión

Sustentantes de cualquier edad, género, estado civil, promedio de licenciatura, escuela o facultad de procedencia (acreditada o no acreditada) y de cualquier nacionalidad; nuestro subgrupo de mayor interés se compone de sustentantes que hayan presentado por primera y única vez el ENARM en los años analizados.

9.5.2 Criterios de exclusión

Sustentantes provenientes de escuelas extranjeras, registros que no pudieran asignarse a alguna escuela en específico de acuerdo al censo de escuelas de medicina proporcionado por el COMAEM y registros que en el conjunto de datos “escuela de procedencia” mostrasen etiquetas de escuelas inexistentes. Se excluyen también aquellos registros que no presenten datos completos para efecto de los cálculos.

9.5.3 Variables sociodemográficas

Variables sociodemográficas		
Variable	Tipo de variable	Nivel de medición
Edad	Cuantitativa discreta	Intervalar
Género	Cualitativa, categórica, dicotómica	Nominal
Estado civil	Cualitativa, categórica	Nominal
Nacionalidad	Cualitativa, categórica	Nominal
Definición conceptual: indicadores sociales, económicos y demográficos que permiten segmentar la población en grupos homogéneos y así dar significado al proceso de interpretación.		
Definición operacional: datos demográficos ingresados por el sustentante en la cédula de registro y concentrados en la base de datos general.		

9.5.4 Variables académicas

Variables académicas		
Variable	Tipo de variable	Nivel de medición
Escuela de procedencia	Cualitativa, categórica	Nominal
Acreditación del programa (acreditada vs no acreditada)	Cualitativa, categórica, dicotómica	Nominal
Tipo de escuela de medicina (pública vs privada)	Cualitativa, categórica, dicotómica	Nominal
Promedio de licenciatura	Cuantitativa continua	De razón
Calificación ENARM	Cuantitativa continua	De razón
Selección en el ENARM	Cualitativa, categórica, dicotómica	Nominal
Definición conceptual: nombre de la escuela de medicina de procedencia, estatus de acreditación de acuerdo al COMAEM, promedio aritmético de las calificaciones obtenidas por el sustentante durante la carrera de medicina incluyendo internado y servicio social, calificación obtenida en el Examen Nacional de Aspirantes a Residencias Médicas en la edición de que se trate.		
Definición operacional: antecedentes académicos ingresados por el sustentante en la cédula de registro y concentrados en la base de datos general; acreditación de acuerdo al Estatus de Acreditación publicado por el COMAEM a diciembre de 2018, promedio aritmético de la carrera de acuerdo a los datos ingresados por el sustentante en la cédula de registro y concentrados en la base de datos general y que coincide con el certificado de calificaciones de la licenciatura, expresada con dos decimales, calificación obtenida en el Examen Nacional de Aspirantes a Residencias Médicas de acuerdo a la base de datos general y que coincide con la constancia ENARM entregada a cada sustentante expresada con dos decimales.		

9.6 DEPURACIÓN DE LA BASE DE DATOS.

La base de datos proporcionada contaba con 154,582 registros, cada registro

correspondiente a un sustentante que presentó el examen en el lustro que comprende del año 2014 a 2018 (ENARM en sus ediciones XXXVII, XXXVIII, XXXIX, XL y XLI). Se identificaron las etiquetas correspondientes a la variable “escuela de procedencia” y tras la revisión cruzada con las páginas web de las escuelas de medicina del país y el censo de escuelas de medicina proporcionado por el COMAEM, se eliminaron algunos registros por no poder asignarse a alguna escuela en específico. Los registros eliminados corresponden a los sustentantes que en la base de datos en “escuela de procedencia” mostraban las siguientes etiquetas (se muestran las etiquetas como estaban contenidas en la base de datos:

1. Univ. Mich. De san nicolás de hidalgo (Zamora)
2. Inst est tec y sup de matatipac (Tepic)
3. Univ. Cuahutemoc (Puebla)
4. Univ. Valle del bravo, u. Cd. Victoria
5. Univ. Valle del bravo, u. U. Reynosa

Se perdieron 928 registros; la n total disminuyó de 154,582 a 153,654 sustentantes.

9.7 ESTADÍSTICA.

9.7.1 Tamaño de muestra.

El Examen Nacional de Aspirantes a Residencias Médicas constituyó su edición XLI (41 ediciones) en el año 2018. Nuestra muestra corresponde a cinco exámenes que comprenden de la edición 2014 a 2018 agrupando en su totalidad a 153,654 sustentantes. De entre los 153,654 aspirantes que presentaron el Examen Nacional de Aspirantes a Residencias Médicas se identificaron 17,351 aspirantes (11.3%) que realizaron el examen en el mismo año en que se graduaron de la licenciatura, por lo que se consideró que de manera lógica esta era la primera vez que sustentaban el ENARM. No se pudieron identificar aquellos sustentantes que

hubiesen realizado el examen por primera vez en años posteriores a su graduación de licenciatura. De esta manera, se asumió que en el periodo 2014-2018, los sustentantes que presentaron el examen ENARM al menos por segunda ocasión ascendió a 136,303.

9.7.2 Análisis estadístico.

Se realizará un análisis utilizando estadística descriptiva no inferencial para determinar los atributos del grupo de estudio. Entre las determinaciones que se realizarán se encuentran medidas de tendencia central y dispersión como la moda, la mediana, la media aritmética y el rango basándonos principalmente en análisis de frecuencias para cada año y para toda la muestra.

Análisis estadístico		
Variable	Tipo prueba	Observaciones
Descriptivas Toda la muestra	Tendencia central Dispersión	Primera etapa Moda, mediana, moda, rango
Cuantitativas Correlación	Correlación de Pearson	Interpretación depende del grupo de datos; expertos
Cuantitativas Diferencia de medias	T de Student	Considerar tamaño del efecto (muestra grande)
Cuantitativas Más de 2 medias	ANOVA	Análisis post-test (Bonferroni)
Catóricas	Chi cuadrada	Considerar tamaño del efecto (muestra grande)
...		
...		

La correlación entre las variables cuantitativas continuas que son de nuestro interés principal (calificación ENARM y promedio de licenciatura) se realizará por medio de una prueba paramétrica coeficiente correlación de Pearson. La correlación es una medida de asociación monotónica entre dos variables. Una correlación monotónica es aquella en la que ya sea a) mientras el valor de una variable se incrementa, lo mismo ocurre con el valor de la otra variable; o b) mientras el valor de una variable se incrementa, el valor de la otra variable

disminuye. En los datos correlacionados, por lo tanto, el cambio en la magnitud de una variable está asociado con el cambio en la magnitud de otra variable, ya sea en la misma dirección o en dirección opuesta. En otras palabras, valores altos de una variable tienden a estar asociados ya sea con valores altos (correlación positiva) o bajos (correlación negativa) de la otra variable, y viceversa (Schober & Schwarte, 2018). Para una prueba de esta naturaleza es necesario que los datos en los que se analizará la correlación cumplan ciertas premisas: los datos deben derivar de una muestra aleatoria, o al menos representativa y ambas variables son continuas, aleatorias con una distribución normal (Schober & Schwarte, 2018). Los datos de la muestra a la que nos referimos cumplen estas premisas. Se realizarán comparaciones entre las medias de subgrupos (V. gr. media de calificaciones ENARM de toda la muestra de sustentantes comparadas con las calificaciones obtenidas por los sustentantes de primera vez). Para dicha comparación se utilizará una prueba estadística t de Student. Una prueba t es un tipo de estadístico que se usa para comparar medias entre dos grupos. Es una de las pruebas estadísticas de hipótesis más utilizadas. Puesto que existen dos tipos de inferencia estadística, paramétrica y no paramétrica, debemos conocer a cuál pertenecen nuestros datos. Los métodos paramétricos se refieren a la técnica estadística en la que la distribución probabilística está definida y se pueden realizar inferencias sobre los parámetros de la distribución. En los casos en los que la distribución probabilística de los datos no puede ser definida se emplean métodos no paramétricos. Las pruebas t son un tipo de método paramétrico; pueden utilizarse cuando las muestras cumplen las condiciones de normalidad, varianzas iguales, e independencia (Kim, 2015). Por el tamaño de nuestra muestra podemos asumir una distribución normal, las comparaciones se harán entre grupos independientes y al momento de analizar los estadísticos se verificará previamente la homocedasticidad (varianzas iguales). Cuando las muestras no sean del mismo tamaño se contemplará la pertinencia de adecuar la prueba t para muestras desiguales o la implementación de una

prueba U de Mann-Whitney que habitualmente se utiliza para comparar medias originadas de datos no paramétricos (Zimmerman, 1987). Tomando en cuenta los supuestos enumerados (normalidad y homocedasticidad), utilizaremos el análisis de varianza (ANOVA) para examinar las diferencias entre más de dos medias, aplicando un análisis post-test de tipo Bonferroni en cada uno de los casos. (Sawyer, 2009) Las variables categóricas serán examinadas por medio de la prueba de Chi cuadrada (X^2), tomando en cuenta el tamaño de nuestro modelo y las interpretaciones adecuadas para este tipo de datos. (Shi, DiStefano, McDaniel, & Jiang, 2018) En los casos en que se considere pertinente se determinará el tamaño del efecto tomando en cuenta la distinción entre significancia estadística y significado “clínico” o real de los resultados. (Sullivan & Feinn, 2012) (Kraft, 2020)

Las pruebas estadísticas se realizarán utilizando paquetería SPSS (Statistical Analysis in Social Science, en su versión 24).

10. RESULTADOS

El análisis inicial de la base de datos permitió determinar que los reactivos están considerados en las disciplinas de *Medicina Interna, Pediatría, Ginecoobstetricia y Cirugía General* sin especificar su ponderación. No existen datos de la calificación de idioma inglés. El número de reactivos varía para cada disciplina troncal: Medicina Interna cubre 168 reactivos, Pediatría 112, Ginecoobstetricia 100 y Cirugía General 70. En la **Tabla 1** se puede observar la estructura del ENARM por área y subárea. Esta distribución es general ya que no contamos con la asignación de reactivos dentro de cada bloque de disciplinas troncales ni la asignación de casos

10.1 RESULTADOS GLOBALES

El estudio comprendió el registro de 153,654 examinados inscritos en el ENARM durante los últimos cinco años, de los cuales el 51.2% fueron mujeres y 49.7% hombres. Para cada uno de ellos se determinó el rendimiento académico obtenido mediante el cociente de los aciertos entre el número total de reactivos. La **Figura 1**, polígono de frecuencias, permite observar una media de desempeño de 62.3, (desviación estándar 8.9) con un rango de 1.11 a 91.11; la mediana (percentil 50) se localizó en 62.4 y la moda en 60.9 indicando una distribución unimodal asimétrica. La mayoría de los candidatos se desempeñó entre 60 y 70, y el 50.8% de la muestra obtuvo puntajes por encima de la media; el percentil 75 se encontró por encima de 68.9. Para una interpretación adecuada de los resultados, debemos subrayar que los candidatos conocen el contenido del examen con anticipación y usualmente dedican cantidades considerables de tiempo preparándose para este examen que se considera de altas consecuencias. (Van Wieren et al., 2014)

El área en la que los examinados obtuvieron mejor rendimiento fue Ginecoobstetricia con 64.9 en contraste con el 59.2 de Medicina Interna, la diferencia resultó estadísticamente significativa ($p < 0.001$).

Tabla 1. Estructura general del ENARM	
Área y sub área	Reactivos (%)
Medicina Interna	168 (37.3)
Cardiovascular	
Dermatología	
Endocrinología	
Gastroenterología	
Geriatría y Gerontología	
Hematología	
Infectología	
Inmunoalergia	
Nefrología	
Neumología	
Neurología	
Psiquiatría	
Reumatología	
Pediatría	112 (24.9)
Crecimiento y desarrollo	
Infectología pediátrica	
Urgencias pediátricas	
Gineco-obstetricia	100 (22.2)
Ginecología	
Obstetricia	
Cirugía	70 (15.6)
Cirugía abdominal	
Oftalmología	
Otorrinolaringología	
Urgencias	
Urología y proctología	
Traumatología y ortopedia	
Total	450 (100.0)

10.1.1 Rendimiento en el ENARM y su relación con el promedio en la licenciatura.

Para determinar la relación entre el rendimiento académico en la licenciatura y el ENARM, expresados como el promedio de calificaciones y el porcentaje de aciertos, respectivamente, se agrupó a los examinados en rangos según el promedio obtenido en la licenciatura. En la 22 se aprecia que los examinados con un promedio de calificaciones de 9.5 a 10 tuvieron el mayor rendimiento en el examen con una media de 71.8 y un porcentaje de

selección de 63.9; en contraste, los examinados con un promedio de 6.5 a 6.9 alcanzaron una media de 53.7 y un porcentaje de selección de 9.5. La correlación de Pearson ($r=0.35$) indicó una relación entre el porcentaje de aciertos obtenido en el ENARM y el promedio de calificaciones. Las diferencias en las tasas de selección para cada uno de los rangos fueron estadísticamente significativas al aplicar la prueba de X^2 ($p<0.001$). Ver **Tabla 2**.

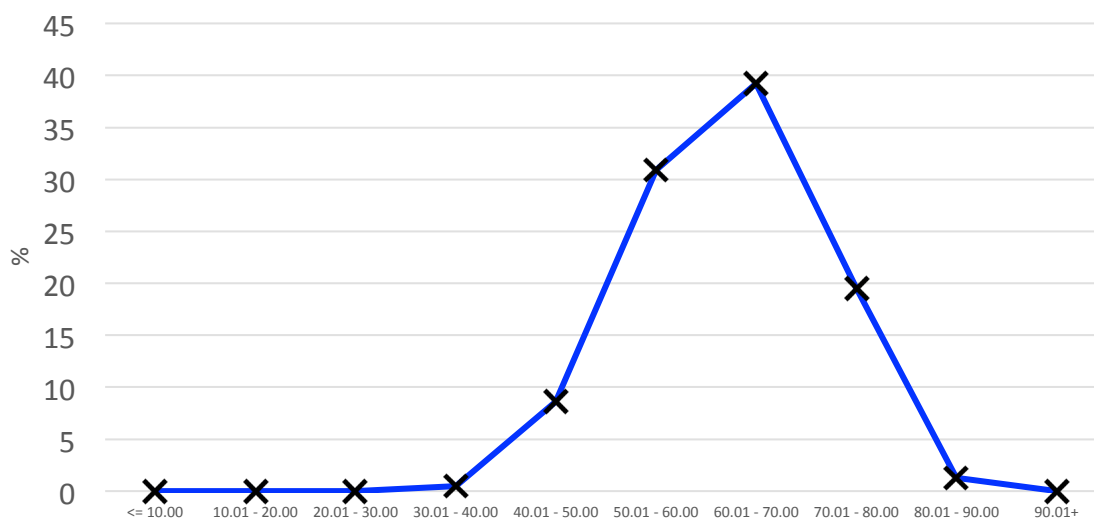


Figura 1. Desempeño en ENARM

Gráfica 1. Resultados obtenidos en cada área por todos los examinados en el ENARM (2014-2018).

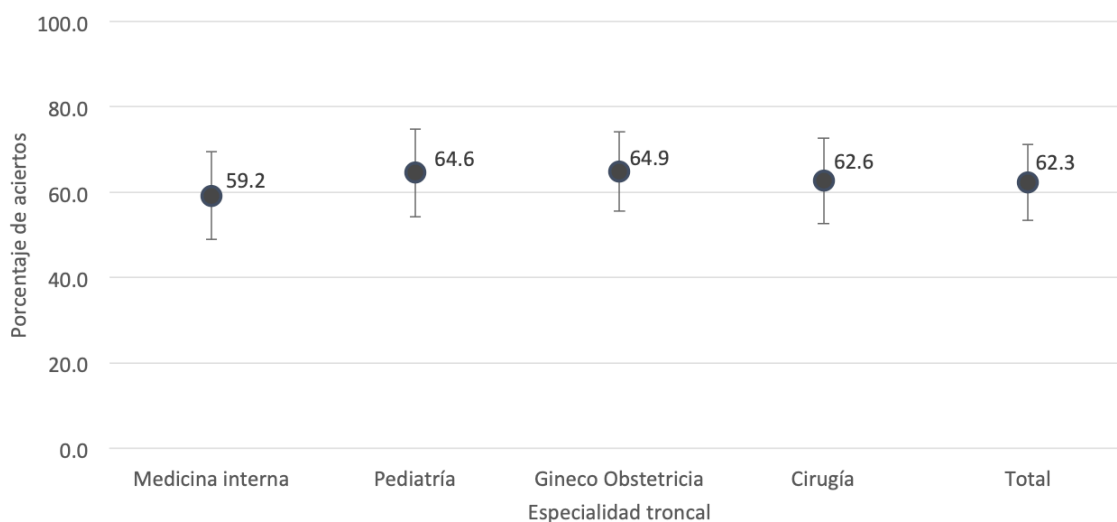


Tabla 2. Desempeño en la licenciatura y desempeño obtenido en el ENARM.
 Todos los sustentantes en el periodo 2014-2018.

	n	ENARM $\bar{x} \pm DE$	Selección %
a) Desempeño			
6.5 – 6.9	799	53.7 ± 8.1	9.5
7.0 – 7.4	8,260	55.5 ± 7.8	11.3
7.5 – 7.9	24,109	58.1 ± 7.9	14.3
8.0 – 8.4	48,231	61.1 ± 8.2	20.6
8.5 – 8.9	46,547	64.0 ± 8.3	29.8
9.0 – 9.4	21,957	67.6 ± 8.2	43.9
9.5 – 10	3,351	71.8 ± 8.1	63.9
Total	153,254	62.3 ± 8.9	26.0

Correlación de Pearson 0.35. Chi cuadrada X^2 $p < 0.001$.
 Datos perdidos: 400.

10.1.2 Estatus de acreditación del programa académico y rendimiento en el ENARM.

Considerando el dictamen de acreditación del COMAEM, agrupamos a los examinados de acuerdo al programa académico de la escuela donde cursaron la licenciatura: acreditados vs. no acreditados. Como lo muestra la **Tabla 3**, la mayoría de los examinados ($n=137272$, 89%) pertenecen a programas acreditados y obtuvieron una media de 62.5 en contraste con la de los examinados pertenecientes a programas no acreditados que fue de 61.4. La prueba t de Student mostró significancia estadística para esta diferencia ($p < 0.001$). La tasa de selección por tipo de programa fue 26.2% para los acreditados vs 22.9% para los no acreditados; la prueba X^2 mostró significancia estadística para esta diferencia ($p < 0.001$). Los tamaños del efecto observados fueron de 0.12 para el desempeño (d de Cohen) y de 1.14 para la tasa de selección (razón de momios).

10.1.3 Tipo de escuela y rendimiento en el ENARM.

La totalidad de los examinados en el quinquenio se agrupó de acuerdo a un criterio dicotómico según la escuela de procedencia. La mayoría de los examinados ($n=115880$, 75%)

cursaron su licenciatura en escuelas públicas en las cuales se observa una mayor media de desempeño en el ENARM (62.3) al compararse con aquellos provenientes de escuelas privadas (62.2). La selección también mostró diferencias, siendo de 26.6% para las escuelas públicas y de 23.6% para las privadas; las diferencias fueron estadísticamente significativas ($p < 0.001$). Los tamaños del efecto observados fueron de 0.01 para el desempeño (d de Cohen) y de 1.13 para la tasa de selección (razón de momios). Ver **Tabla 3**.

Tabla 3. Tipo de escuela de medicina. Periodo 2014-2018.

Variable	n	Desempeño $\bar{x} \pm DE$	Seleccionado*** %
a) Estatus de acreditación			
Programa acreditado ⁺⁺⁺	137,272	62.5 \pm 8.9	26.2
Programa no acreditado	15,236	61.4 \pm 8.9	22.9
Total	152,508	62.3 \pm 8.9	26.0
b) Tipo de escuela			
Pública ⁺⁺⁺	115,880	62.3 \pm 8.8	26.6
Privada	36,628	62.2 \pm 9.1	23.6
Total	152,508	62.3 \pm 8.9	26.0

+++ T de Student $p < 0.001$; *** $\chi^2 p < 0.001$. Sin datos: 1,146.

d de Cohen (acreditadas vs. no acreditadas) = 0.1236

d de Cohen (públicas vs. privadas) = 0.0112

O.R. tasa de selección (acreditadas vs. no acreditadas) = 1.1457 (1.1120 - 1.1805)

O.R. tasa de selección (públicas vs. privadas) = 1.1252 (1.1024 - 1.1485)

10.1.4 Región geográfica y rendimiento en el ENARM.

El rendimiento académico de los examinados se determinó según la región geográfica donde se localiza el plantel educativo en el que cursaron la licenciatura de medicina. Los resultados del estudio se consignan en la **Tabla 4**. Se aprecia que los valores de las medias de calificación mostraron diferencias significativas entre las diversas regiones, cuando se aplicó el análisis de varianza. Sin embargo, los valores entre las regiones Metropolitana, Centro-Sur y Centro-Occidente fueron muy cercanos y no significativos; al igual que cuando se compararon los valores entre las regiones Noreste y Noroeste; fue la región Sur-Sureste la que mostró los

valores menores de rendimiento.

Tabla 4. Desempeño y porcentaje de selección en el ENARM por región*.

Región	n	Desempeño	Selección
		\bar{x}	%
Metropolitana ⁺	37,885	62.6 ± 8.3	24.3
Centro-Sur	22,701	61.2 ± 8.8	22.2
Centro-Occidente ⁺	32,656	62.8 ± 9.3	28.2
Noroeste [§]	14,164	63.3 ± 8.9	31.0
Noreste [§]	22,248	63.8 ± 9.0	31.0
Sur-Sureste	22,854	58.9 ± 8.2	20.7
Total	153,654	62.3 ± 8.9	25.8

Prueba ANOVA < 0.001. +, § diferencia no significativa en el desempeño de las regiones

*Clasificación de acuerdo a ANUIES (2018).

Asimismo, el porcentaje de selección indica que las regiones Noreste y Noroeste no difieren entre sí con 31% (el más alto), les siguen Centro-Occidente con 28.2%, Metropolitana con 24.3%, Centro-Sur con 22.2% y Sur-Sureste con 20.7% (el más bajo). Véase **mapa**.

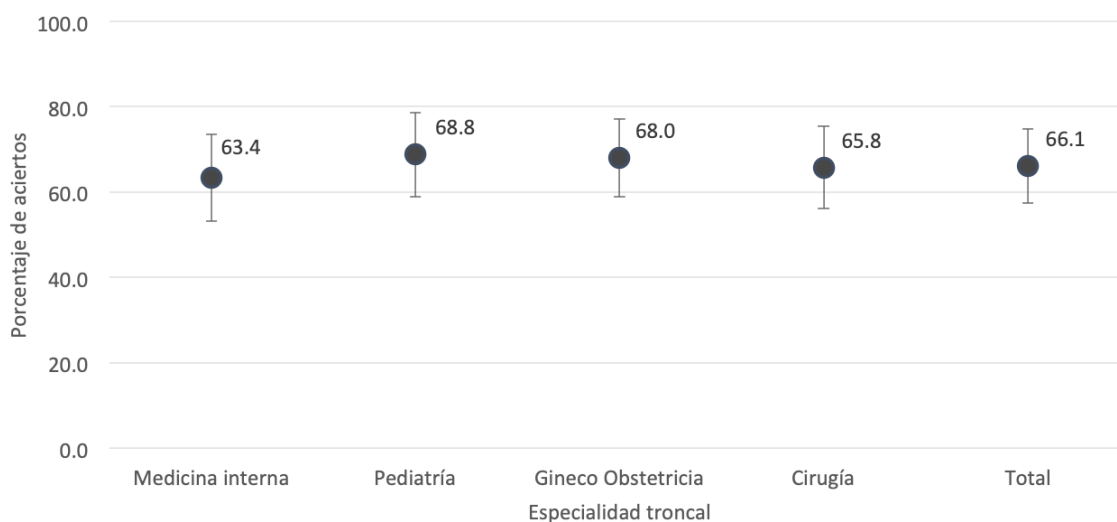


10.2 RESULTADOS EN SUSTENTANTES DE PRIMERA VEZ

El número de sustentantes de primera vez que se logró identificar en el periodo de 2014 a 2018 fue de 17,400 (11.3% del total de aspirantes). De la misma manera

que para el grupo global, en cada uno de ellos se determinó el rendimiento académico obtenido mediante el cociente de los aciertos entre el número total de reactivos Este grupo obtuvo una media de desempeño de 66.1, (desviación estándar 8.73) con un rango de 31.11 a 88.44. El área en la que estos examinados obtuvieron mejor rendimiento fue Pediatría con 68.8 en contraste con 63.4 en Medicina Interna.

Gráfica 2. Resultados obtenidos en cada área por los examinados de primera vez en el ENARM.



10.2.1 Rendimiento en el ENARM y su relación con el promedio en la licenciatura.

Los sustentantes de primera vez también fueron agrupados en rangos según el promedio obtenido en la licenciatura. En la **Tabla 5** se aprecia que los examinados con un promedio de calificaciones de 9.5 a 10 tuvieron el mayor rendimiento en el examen con una media de 74.7 y un porcentaje de selección de 74.5; en contraste, los examinados con un promedio de 6.5 a 6.9 alcanzaron una media de 55.2 y un porcentaje de selección de 22.6. La correlación de Pearson ($r=0.50$) indicó una relación entre el porcentaje de aciertos obtenido en el ENARM y el promedio de calificaciones. Las diferencias en las tasas de selección para cada uno de los rangos fueron estadísticamente significativas al aplicar la prueba de X^2 ($p<0.001$).

Tabla 5. Desempeño en la licenciatura y desempeño obtenido en el ENARM.
Sustentantes de primera vez en el periodo 2014-2018.

	n	ENARM $\bar{x} \pm DE$	Selección %
a) Desempeño			
6.5 – 6.9	31	55.2 ± 7.2	22.6
7.0 – 7.4	261	56.2 ± 7.4	10.0
7.5 – 7.9	1,279	58.6 ± 7.6	11.1
8.0 – 8.4	4,092	62.6 ± 7.8	19.4
8.5 – 8.9	6,158	66.2 ± 7.8	31.3
9.0 – 9.4	4,452	70.1 ± 7.7	49.3
9.5 – 10	1,042	74.7 ± 6.4	74.5
Total	17,315	66.1 ± 8.7	33.9

Correlación de Pearson 0.50. Chi cuadrada X^2 $p < 0.001$.
Datos perdidos: 36.

10.2.2 Estatus de acreditación del programa académico y rendimiento en el ENARM.

Agrupando a los examinados de primera vez conforme al estatus de acreditación del programa académico correspondiente a la escuela donde cursaron la licenciatura comparamos desempeño y tasas de selección. Como lo muestra la **Tabla 6**, la mayoría de los examinados ($n=16062$, 92%) pertenecen a programas acreditados y obtuvieron una media de 66.2 en contraste con la de los examinados pertenecientes a programas no acreditados que fue de 65.8. La prueba t de Student mostró significancia estadística para esta diferencia ($p < 0.001$). La tasa de selección por tipo de programa fue 34.5% para los acreditados vs 33.2% para los no acreditados; la prueba X^2 mostró significancia estadística para esta diferencia ($p < 0.001$).

10.2.3 Tipo de escuela y rendimiento en el ENARM.

Los sustentantes de primera vez en el periodo 2014-2018 se subdividieron en categorías según la escuela de procedencia. La mayoría de los examinados ($n=13507$, 77%) cursaron su licenciatura en escuelas públicas en las cuales se observa una mayor media de desempeño en el ENARM (66.3) al compararse con aquellos provenientes de escuelas

privadas (65.6). La selección también mostró diferencias, siendo de 26.6% para las escuelas públicas y de 23.6% para las privadas; las diferencias fueron estadísticamente significativas ($p < 0.001$). Ver **Tabla 6**.

Tabla 6. Tipo de escuela de medicina. Periodo 2014-2018. Sustentantes de primera vez.

Variable	n	Desempeño $\bar{x} \pm DE$	Seleccionado *** %
a) Estatus de acreditación			
Programa acreditado ⁺⁺⁺	16,062	66.2 \pm 8.7	34.5
Programa no acreditado	858	65.8 \pm 8.3	33.2
Total	16,920	66.0 \pm 8.4	33.9
b) Tipo de escuela			
Pública ⁺⁺⁺	13,507	66.3 \pm 8.6	34.9
Privada	3,804	65.6 \pm 9.2	30.5
Total	17,311	66.1 \pm 8.7	33.9

+++ T de Student $p < 0.001$; *** $\chi^2 p < 0.001$. Sin datos: 431 en a), 40 en b).

10.2.4 Región geográfica y rendimiento en el ENARM.

El rendimiento académico de los examinados se determinó según la región geográfica donde se localiza el plantel educativo en el que cursaron la licenciatura de medicina.

Los resultados del estudio se consignan en la **Tabla 7**. Se aprecia que los valores de las medias de calificación mostraron diferencias significativas entre las diversas regiones, cuando se aplicó el análisis de varianza. Sin embargo, los valores entre las regiones Centro-Occidente, Noroeste y Noreste fueron muy cercanos y sin diferencia significativa; al igual que cuando se compararon los valores entre las regiones Centro-Sur y Sur-Sureste. Asimismo, el porcentaje de selección más alto obtenido corresponde a la región Centro-Occidente con 44.8%, seguida de Noreste con 41.9% y Noroeste con 41.4%. La tasa de selección más baja correspondió a la región Sur-Sureste con 26.1%. La **Tabla 8**, "Regiones Geográficas" muestra los estados mexicanos agrupados de acuerdo a estos criterios.

Tabla 7. Desempeño y porcentaje de selección en el ENARM por región*. Sustentantes de primera vez.

Región	n	Desempeño Selección	
		\bar{x}	%
Metropolitana	7,767	65.2 ± 8.4	29.0
Centro-Sur ⁺	1,790	64.4 ± 9.2	28.9
Centro-Occidente [§]	3,527	68.4 ± 8.6	44.8
Noroeste [§]	789	68.3 ± 7.5	41.4
Noreste [§]	1,878	68.0 ± 8.4	41.9
Sur-Sureste ⁺	1,560	63.9 ± 9.1	26.1
Total	17,311	66.1 ± 8.7	33.9

Prueba ANOVA < 0.001. +, § diferencia no significativa en el desempeño de las regiones

*Clasificación de acuerdo a ANUIES (2018). Datos perdidos: 40.

10.3 COMPARACIÓN ENTRE LOS DOS GRUPOS

Realizando un análisis primario hemos intentado establecer la diferencia entre el desempeño obtenido de manera global por los sustentantes independientemente del año de egreso y el desempeño obtenido por médicos recién graduados (sustentan el examen el mismo año en que egresan de la licenciatura) observando que existen diferencias en las medias obtenidas entre estos dos grupos. Es importante señalar que, como mencionamos previamente, en la base de datos a la que tuvimos acceso **no** se especifica el número de veces que el sustentante presentó el examen; por lo tanto nuestra clasificación se basa sobre la coincidencia entre el año de graduación del sustentante y el año de edición del ENARM.

10.3.1 Rendimiento en el ENARM

En la **Tabla 9** se observan los resultados obtenidos por los examinados en el ENARM tomando al grupo en su totalidad así como la comparación entre estos y aquellos sustentantes de primera vez y sustentantes subsecuentes (presentaron el examen más de una vez o en un año distinto al de su graduación de la licenciatura) de forma anual y para todo el periodo.

Como habíamos señalado en apartados previos, los sustentantes de primera vez presentan mejores desempeños en el ENARM y mayores tasas de selección cuando se comparan con el

Tabla 8. Estados mexicanos agrupados por región geográfica. Número de candidatos por región.

Región	n (%)
Ciudad de México y área metropolitana	37,885 (24.8)
Centro-Sur	22,701 (14.9)
Puebla	10,550 (6.9)
Hidalgo	3,100 (2.0)
Guerrero	2,897 (1.9)
Tlaxcala	1,750 (1.1)
Morelos	1,590 (1.0)
Querétaro	1,455 (1.0)
Estado de México	1,359 (0.9)
Centro-Occidente	32,656 (21.4)
Jalisco	17,192 (11.3)
Michoacán	9,575 (6.3)
Guanajuato	1,928 (1.3)
Nayarit	1,862 (1.2)
Aguascalientes	1,188 (0.8)
Colima	911 (0.6)
Noroeste	14,164 (9.3)
Sinaloa	5,102 (3.3)
Baja California	4,827 (3.2)
Chihuahua	3,018 (2.0)
Sonora	1,217 (0.8)
Noreste	22,248 (14.6)
Nuevo León	8,178 (5.4)
Tamaulipas	6,047 (4.0)
Durango	2,972 (1.9)
Zacatecas	2,510 (1.6)
Coahuila	1,425 (0.9)
San Luis Potosí	1,116 (0.7)
Sur-Sureste	22,854 (15.0)
Veracruz	6,779 (4.4)
Oaxaca	5,076 (3.3)
Chiapas	4,267 (2.8)
Tabasco	3,662 (2.4)
Yucatán	1,879 (1.2)
Campeche	927 (0.6)
Quintana Roo	264 (0.2)
Total	152,508 (100.0)

Valores perdidos: 1146.

Elaborado con datos de Asociación Nacional de Universidades e Instituciones de educación Superior ANUIES (2018).

grupo en su totalidad (global) independientemente del desempeño registrado en la licenciatura, del estatus de acreditación del programa educativo de la escuela, de la categorización de la escuela de medicina como pública o privada y de la región geográfica a donde es asignada la escuela de medicina de procedencia. Ahora bien, el margen de esta diferencia aumenta cuando se comparan a los sustentantes de primera vez con aquellos que hemos catalogado como “subsecuentes”, es decir su categoría complementaria; esta observación se repite de manera consistente para cada uno de los años de aplicación y para el lustro completo: los examinados que presentaron el examen por primera vez tuvieron un mejor rendimiento (66.10) comparado con aquellos que aplicaron en más de una ocasión o en años subsecuentes a su graduación (61.84); la diferencia fue estadísticamente significativa ($p < 0.001$). Del mismo modo, los examinados de primera vez lograron una mayor tasa de selección (33.9%) que quienes aplicaron en más de una ocasión o en años subsecuentes a su graduación (24.8%); esta diferencia fue estadísticamente significativa con la prueba χ^2 ($p < 0.001$).

Aplicación	Primera vez		Subsecuentes		Global	
	n	Desempeño	n	Desempeño	n	Desempeño
2014	2,908	66.25 ± 8.34 ⁺⁺⁺	20,942	61.54 ± 8.23	23,850	62.12 ± 8.39
2015	1,517	62.35 ± 8.63 ⁺⁺⁺	29,126	60.62 ± 8.86	30,643	60.71 ± 8.86
2016	4,202	65.53 ± 8.79 ⁺⁺⁺	27,396	61.06 ± 8.63	31,598	61.66 ± 8.79
2017	4,148	66.24 ± 9.02 ⁺⁺⁺	28,949	62.74 ± 9.10	33,097	63.18 ± 9.17
2018	4,576	67.62 ± 8.26 ⁺⁺⁺	29,890	63.09 ± 8.76	34,466	63.69 ± 8.83
Lustro	17,351	66.10 ± 8.73⁺⁺⁺	136,303	61.84 ± 8.81	153,654	62.33 ± 8.90
Seleccionados (%)	33.9 ⁺⁺⁺		24.8		25.8	

+++ Prueba t < 0.001; *** Prueba χ^2 < 0.001

11. DISCUSIÓN

El estudio evalúa una muestra estratificada que comprende un periodo de cinco años y permite evaluar el desempeño en un examen de altas consecuencias (ENARM) que se aplica a todos los sustentantes que aspiran a realizar una especialización médica (Akaki-Blancas, et. al., 2018 y Barajas-Ochoa, et. al., 2019). Nuestro objetivo principal fue identificar las características demográficas y académicas de los sustentantes relacionadas con su desempeño, así como los atributos de las escuelas de medicina que impactan en este desempeño y las tasas de selección para los cursos de especialización médica.

Panorama general.

Los resultados de la investigación revelan que la población de médicos evaluados es muy heterogénea en cuanto al dominio de los conocimientos médicos, lo que se demuestra al analizar el desempeño en el ENARM; este desempeño está ligado a ciertas variables como son: estatus de acreditación de la escuela de medicina, tipo de escuela de medicina (pública o privada), y región geográfica. Esta heterogeneidad en los programas educativos de las escuelas de medicina también se observa en otras partes del mundo. (Westerman et al., 2019) Esta variabilidad en las condiciones de la educación médica ha sido demostrada y es al parecer obvia; algunos investigadores han buscado identificar a las escuelas de medicina como “mejores” o “peores”, o han intentado asignarles lugares como si se tratara de una competencia. (Barajas-Ochoa, Ramos-Remus, Ramos-Gómez, et al., 2019) Consideramos que el *quid* del asunto es saber por qué suceden estas diferencias.

En nuestro análisis, para la población total examinada, la media de desempeño alcanzó 62.3, lo cual de manera general puede ser considerado como suficiente para cursar estudios de posgrado y además cumple con el requisito de ordenamiento de los examinados en función de asignación de plazas y no la clasificación de su grado de conocimientos

Pudimos observar también un desempeño diferenciado para las que se consideran las

principales áreas clínicas de la profesión médica: Medicina Interna, Ginecoobstetricia, Pediatría y Cirugía General (Abreu, et. al., 2008). En nuestro análisis el desempeño fue mayor en Ginecoobstetricia para la muestra global y en Pediatría para los sustentantes de primera vez; estos resultados son semejantes a los obtenidos en un examen diagnóstico realizado en internos de pregrado (Martínez-González, et. al., 2017). Aunque debemos tener en cuenta que el contenido teórico evaluado es distinto y, por tanto, no del todo equiparable, además de un tamaño de muestra marcadamente menor en el estudio referido, el área de mayor rendimiento también fue Ginecoobstetricia, lo que puede estar relacionado con el diseño del currículo, así como a una mayor exposición a casos en las etapas de formación clínica.

Desempeño en la licenciatura.

De nuestros resultados sobresale la evidencia sobre el impacto entre el desempeño en la licenciatura (relacionado con la formación recibida) y el desempeño obtenido en el ENARM.

Nuestros resultados muestran consistentemente que a mayor promedio en la licenciatura mejor desempeño en el ENARM y mayor probabilidad de ser seleccionado. Este efecto se observa de manera más clara en los extremos de los grupos y la agrupación por rangos (Tabla 4); por ejemplo, el grupo de examinados con el promedio de licenciatura más alto, 9.5 a 10, es el que corresponde al desempeño más alto en el ENARM (71.8) cuando se compara con la media global de toda la población (62.3). Sin embargo, el dato que más llama la atención es que estos examinados tuvieron una tasa de selección casi tres veces mayor que el grupo visto en su conjunto (63.9 vs 26%).

El ENARM ha sido criticado recientemente, de una manera constructiva, en lo que concierne a aspectos de validez y equidad; (Downing, 2003)(Barajas-Ochoa, Ramos-Remus, Castillo-Ortiz, et al., 2019); por ejemplo, la estructura interna del examen no ha sido expuesta o evaluada públicamente y la aplicación de teorías psicométricas establecidas, como la teoría

clásica del test o la teoría de respuesta al ítem para la creación y análisis de reactivos, no se ha llevado a cabo en este examen. (Barajas-Ochoa, Ramos-Remus, Castillo-Ortiz, et al., 2019)

Consideramos que nuestro estudio puede arrojar luz o constituirse en una fuente o evidencia de validez para el ENARM, aquella denominada *relación con otras variables* (AERA, APA, & NCME, 2014; Downing, 2003), característica que debe destacarse ya , como señalamos previamente, otros autores han señalado la ausencia de evidencias de validez de este examen (Barajas-Ochoa, et. al., 2017). Nuestro estudio aporta una evidencia de validez que se robustece por el tamaño de muestra y por el análisis estadístico utilizado, confirmando asimismo nuestra principal hipótesis de trabajo: a mayor promedio de calificaciones obtenido en los estudios de licenciatura de los examinados se obtiene un mayor desempeño en el ENARM, siendo esta relación significativa.

Podemos inferir que el hallazgo de un mayor desempeño en aquellos sustentantes que presentan el ENARM por primera vez probablemente se debe a una mayor capacidad de recordar los contenidos programáticos de la carrera en comparación con los sustentantes que toman el examen de manera subsecuente. Este punto cobra relevancia cuando algunos autores han establecido que además de copioso, el conocimiento médico es fácilmente olvidable. (Taveira-Gomes, Prado-Costa, Severo, & Ferreira, 2015) Las técnicas para disponer del conocimiento almacenado así como la efectividad de las mismas pueden estar relacionadas con el programa educativo que se haya cursado. (Bergman et al., 2015)

En este punto parece pertinente señalar que el desempeño académico no depende exclusivamente del esfuerzo personal, sino también de variables como los antecedentes académicos (personales y familiares), las estrategias de aprendizaje y preparación para el examen, así como factores psicológicos, vocacionales y socioeconómicos. (Ferguson, James, & Madeley, 2002) (Guevara-Guzmán, Galván, & Muñoz-Comonfort, 2007) (Aguilar, León, Morales, Rosas, & Guzmán, 2014)

Acreditación.

Los resultados de este trabajo aportan pruebas de la influencia del proceso de acreditación, dado que los egresados de programas académicos acreditados lograron un mejor desempeño en el examen y una mayor probabilidad de continuar con estudios de posgrado en una especialización médica en relación con los que proceden de programas no acreditados ($p < 0.001$). Esto es similar a lo encontrado por Vázquez Martínez (Vázquez-Martínez, et. al., 2018) para los sustentantes en el ENARM de 2016; estos investigadores identificaron una mayor tasa de selección (sin significancia estadística) para escuelas de medicina acreditadas; sin embargo, no exploraron el desempeño en el examen (media de respuestas correctas).

De esta manera, los resultados obtenidos confirman otra de nuestras hipótesis; consideramos que estos hallazgos pueden tener las siguientes explicaciones:

(a) la autoevaluación para la acreditación favorece una mejor organización y secuenciación de contenidos en los programas educativos; (b) los procesos de acreditación impulsan la formación de mejores médicos y, consecuentemente, la mejora continua de los servicios de salud, de la atención médica que ofrecen y de la seguridad de los pacientes a los que se atiende; (Van Zanten et al., 2008) (Vázquez Martínez & Ortiz-León, 2017) (c) la preparación para la acreditación reanuda actividades académicas y crea un ambiente propicio para los procesos de enseñanza-aprendizaje y; (d) el conocimiento del programa educativo por parte de los estudiantes incrementa su motivación para estudiar y aprender.

Tipo de escuela.

En relación con el tipo de escuela, aun cuando las diferencias no fueron numéricamente considerables, pudimos comprobar, de acuerdo a una de nuestras hipótesis,

que los examinados egresados de escuelas públicas tienen mayor desempeño académico en el ENARM y, asimismo, es proporcionalmente mayor el número de seleccionados para cursar una especialización médica en comparación con los provenientes de escuelas privadas; ambas diferencias fueron estadísticamente significativas ($p < 0.001$).

Los hallazgos en este rubro pueden significar: (a) que las medidas de apoyo o correctivas, llevadas a cabo en el propio plantel o por organismos institucionales/gubernamentales contribuyen a elevar la calidad de la educación médica como parte de un compromiso con la educación pública; (b) el impulso a los estudiantes para trascender sus posibles limitaciones o condiciones de vida y la motivación para lograr la movilidad social como profesionistas mejor preparados y reconocidos; (Daude & Robano, 2015) (c) no obstante que el número de escuelas privadas constituye casi el doble que el de las públicas, puede ser que por requerimientos más laxos para el establecimiento de instituciones privadas y la marcada proliferación de estas (e. g. en el mundo existen alrededor de 3000 escuelas de medicina, el doble de las que existían hace 20 años), (Tackett, Zhang, Nassery, Caufield-Noll, & Van Zanten, 2019) predominen aquellas con más limitaciones estructurales y de instalaciones.

Un fenómeno similar puede ser observado en otros países latinoamericanos. (Biolini, 2002)(Clouet-Huerta, González, & Correa, 2017)(Westerman et al., 2019)

Nuestros resultados contrastan con los de Hernández-Gálvez y colaboradores (Hernández-Gálvez & Roldán-Valadez, 2019) quienes concluyeron un mejor desempeño para escuelas privadas; una posible explicación para esta discrepancia es que a diferencia del estudio referido, basado en informes y reportes históricos, nuestro acceso a datos brutos nos permitió determinar de una forma más precisa las características distintivas de los aplicantes, sus tasas de desempeño y selección; i. e. medias y proporciones.

El tipo de escuela de medicina ha sido estudiado en otros contextos: al menos un

estudio ha mostrado una mayor probabilidad para ser seleccionado en una residencia de medicina familiar si el candidato era egresado de una institución pública. (Wimsatt, Cooke, Biggs, & Heidelbaugh, 2016) Esta análisis por especialidades y su relación con las escuelas de procedencia es una posibilidad para un estudio futuro.

Regiones geográficas.

Como ya hemos mencionado, una publicación reciente (Hernández-Gálvez & Roldán-Valadez, 2019) demostró diferencias en el desempeño y las tasas de selección con relación a las regiones geográficas y factores socioeconómicos; los autores basaron su análisis en datos históricos (el reporte anual del ENARM durante 17 años, desde 2001 hasta 2017) y concluyeron un mejor desempeño en el ENARM para la región Noroeste y para escuelas privadas. En nuestra opinión, el método de segmentación socioeconómica, realizado por los autores, de acuerdo a niveles y el proceso de localización geográfica, aunque sofisticado y realizado a través de paquetería computacional especializada, no es fácil de entender para la comunidad médica general. Por lo anterior, consideramos que la clasificación proporcionada por ANUIES es más adecuada, ya que cuenta con el aval de instituciones de educación superior.

Tomando en cuenta esta clasificación, nuestro análisis, basado en datos brutos, determinó que las regiones geográficas de las escuelas de medicina también impactan el desempeño en el ENARM: el desempeño es mayor en las regiones Noreste y Noroeste, seguidas de las regiones Centro-Occidente, Metropolitana y Centro-Sur situadas a media tabla, en tanto que la región Sur-Sureste se encuentra al final de la lista con escuelas que están rezagadas hasta en cinco puntos en la medición de su desempeño en el ENARM (63.8 vs 58.9) y en diez cuando se trata de ocupación de plazas (31% vs 20.7%); estas diferencias se comprobaron de manera significativa entre las regiones del país ($p < 0.001$). Esta circunstancia es un reflejo del grado de desarrollo desigual en el país así como de la distribución

heterogénea de los recursos monetarios y educativos (e. g. más recursos en el área metropolitana de la Ciudad de México); el número de aplicantes puede darnos una idea, y un análisis de esta situación puede permitir una consideración cuidadosa de las desigualdades en las que se desenvuelven los estudiantes.

Según los datos señalados en la **Tabla 4** y en el mapa, se rebate una de nuestras hipótesis de trabajo, ya que el desempeño académico de los examinados provenientes de la región Metropolitana del país **no** es mayor que el de las demás regiones.

Limitaciones.

Ahora bien, podemos señalar algunas de las limitaciones de nuestro estudio:

En primer lugar, los datos a los que tuvimos acceso son solo un concentrado del desempeño en el examen para cada uno de los sustentantes; el carecer de los reactivos (aunque estos hubiesen estado codificados) y su comportamiento de acuerdo al número de sustentantes que acertaron y su distribución en los grupos de bajo o alto desempeño nos impidió conocer de manera precisa la estructura interna del examen (i. e. dificultad y discriminación). La aplicación de la teoría clásica del test y la teoría de respuesta al ítem al comportamiento de estos reactivos puede, potencialmente, arrojar evidencias de validez al examen. (Himelfarb, 2019)

En segundo lugar, por cuestiones de manejabilidad, en esta etapa del estudio limitamos nuestro análisis a variables agregadas o al menos recurrentes en la literatura sobre el tema (v. gr. región geográfica); sin embargo, existen ciertos grupos que excluimos de nuestro análisis y que pueden ser importantes tanto para mejorar el examen como para su aplicación. Tal es el caso de la comparación en el desempeño entre sustentantes extranjeros y nacionales en el ENARM. (Mendoza-Aguilar et al., 2020)

Otras consideraciones.

Desde un punto de vista económico, no debemos olvidar que en los cursos de

especialización médica se ponen de manifiesto muchas situaciones con impacto en la salud pública: los aplicantes superan numéricamente las posiciones ofrecidas, los programas a los cuales pueden aplicar los candidatos son variables y los costos del proceso de selección son absorbidos por la sociedad en su conjunto. (Ray, Bishop, & Dow, 2018) Además de esto, no podemos olvidar la situación de los médicos que no son seleccionados, ya que, como hemos visto, la mayoría de los candidatos permanecen, por lo tanto, como médicos generales. Algunos reportes estiman que alrededor del 14% de todos los médicos están desempleados. (Nigenda & Muños, 2015)

Debemos también reconocer, que el sistema de selección del que nos hemos ocupado es particular para nuestro país; otros países ejercen variantes en sus políticas de admisión para las especializaciones médicas: en los Estados Unidos y Canadá se utiliza un sistema de pareo (“matching”) cuyo propósito es permitir a los sustentantes y a los programas reclutadores el conocerse mutuamente con vista a realizar sus selecciones. En Australia, Reino Unido, Alemania y Holanda la admisión es a través de una selección abierta. Existen otros aspectos, por ejemplo, en Australia, los candidatos deben poseer experiencia clínica antes de aplicar. (Weggemans, van Dijk, van Dooijeweert, Veenendaal, & ten Cate, 2017)

Debemos señalar que esperábamos valores de p significativos debido al gran tamaño de muestra; sin embargo, aunque los tamaños de efecto pueden considerarse pequeños, aún se puede obtener información muy valiosa relacionada con las variables estudiadas. Por ejemplo, en un escenario real, una diferencia de 4% en las tasas de selección puede significar un impacto en cientos de candidatos para estar dentro o fuera del sistema de residencias.

Asimismo, al usar el ENARM para hacer inferencias sobre el desempeño de las escuelas de medicina, no debemos olvidar que este examen es con referencia a norma; el número de “ganadores” depende directamente del número de lugares disponibles. Además, puesto que al momento ignoramos los índices de dificultad o discriminación de los reactivos y

por lo tanto del examen en general, el clasificar a los sustentantes como insuficientes basándonos solamente en tasas de selección y medias de desempeño es cuando menos riesgoso. Por ejemplo, al analizar los datos, logramos determinar que el área troncal de Medicina Interna es la que presenta la puntuación más baja (59.2 en los resultados globales y 63.4 en los sustentantes de primera vez); sin embargo al desconocer la estructura interna y psicometría del examen (dificultad y discriminación de los reactivos) es difícil recomendar intervenciones educativas con fines de mejora, ya que no podemos asegurar que los reactivos pertenecientes a este bloque exploran lo que es pertinente preguntar a esta población. Lo anterior toma relevancia tomando cuenta que los reactivos de Medicina Interna reciben mayor ponderación para desempatar la posición de dos o más sustentantes.

Por otro lado, la interpretación cuidadosa de la estadística descriptiva puede arrojar datos en otro sentido, como lo muestra la **Figura 1**, las medias de desempeño (e. g. el percentil 75 por arriba de 68.9) podrían indicar que, en su conjunto, los sustentantes están preparados para continuar su entrenamiento como residentes médicos y de alguna forma señalar que el nivel de la educación médica en nuestro país es adecuado.

Otros estudios ya han señalado que con la intención de alinear y aprovechar los recursos de salud escasos, se deben hacer esfuerzos para conocer más sobre “cómo son entrenados los estudiantes de medicina, quiénes son los estudiantes que se gradúan, y los factores económicos, políticos y sociales que facilitan u obstaculizan el establecimiento y operación de las escuelas de medicina”.(Boulet, Bede, McKinley, & Norcini, 2007) (Duvivier, Boulet, Opalek, van Zanten, & Norcini, 2014) De esta manera, el análisis de un examen de altas consecuencias como el ENARM es muy valioso y está fuertemente relacionado con los aspectos mencionados.

12. CONCLUSIONES

Los hallazgos del estudio muestran que el desempeño en el ENARM guarda relación con las variables: número de veces que se ha presentado el examen, promedio obtenido en la licenciatura, estatus de acreditación del programa académico, tipo de escuela de procedencia (pública o privada) y región donde se localiza la escuela de procedencia.

El análisis que aquí se presenta no es exhaustivo; forma parte de un estudio más amplio que permitirá identificar las variables más significativas asociadas a un mayor rendimiento en el ENARM y generar estrategias y políticas educativas que aumenten la calidad en la formación de los médicos generales del país. Algunos de los análisis que nos proponemos hacer en un futuro son:

a) Determinar la diferencia en el desempeño en el ENARM entre quienes aspiran a una residencia de alta demanda comparado con el rendimiento de quienes aspiran a una residencia de baja demanda.

b) Evaluar la variabilidad de promedios de egreso de licenciatura en las carreras de medicina entre las diversas escuelas o facultades del país.

c) Identificar diferencias relacionadas con el género de los sustentantes y su impacto en el desempeño y las tasas de selección en el ENARM.

d) Identificar aquellas escuelas de medicina cuyos promedios de egreso tengan mayor correlación con la calificación ENARM (evidencias de validez de la calificación universitaria).

13. CRONOGRAMA DE ACTIVIDADES

ACTIVIDAD	Agos-Sept 2018	Oct-Nov 2018	Dic-Feb 2019	Marzo-Abril 2019	Mayo-Junio 2019	Jul-Ago 2019	Sept-Oct 2019	Nov-Dic 2019	Enero-Febrero 2020	Marzo-Mayo 2020
BÚSQUEDA BIBLIOGRÁFICA										
Obtención de la base de datos										
Análisis estadístico: demográficos										
Evaluar la calidad de los reactivos										
Búsqueda bibliográfica										
Pruebas estadísticas: correlación										
Elaboración de la tesis										
Difusión de resultados										
Publicación										

14. REFERENCIAS BIBLIOGRÁFICAS

- AERA, APA, & NCME. (2014). *The Standards for Educational and Psychological Testing*.
- Aguilar, M. E. U., León, S. O., Morales, C. F., Rosas, E. R. P., & Guzmán, R. G. (2014). El rendimiento académico en el primer año de la carrera de médico cirujano: Modelo multivariado explicativo. *Gaceta Medica de Mexico*, *150*, 324–330.
- Ahumada-Ayala M, Calva-Rodríguez, Galarza Delgado DA, García-Alcalá H, Lustig N, R.-A. G. (2006). Educación médica de posgrado en México: mitos y realidades. *Medicina Universitaria*, *8*(30), 52–75. Retrieved from www.revistasmedicasmexicanas.com.mx
- Akaki Blancas, J. L., & López Bárcena, J. (2018). Specialist medical training in Mexico. *Educacion Medica*. <https://doi.org/10.1016/j.edumed.2018.03.007>
- Ángel Fernández-Ortega, M., Ortiz-Montalvo, A., Ponce-Rosas, R., Fajardo-Ortiz, G., & Mazón-Ramírez, J. J. (2016). Caracterización de alumnos de la carrera de Medicina. *Investigación En Educación Médica*, *5*, 148–154. <https://doi.org/10.1016/j.riem.2015.11.001>
- ANUIES, & Asociación Nacional de Universidades e Instituciones de educación Superior ANUIES. (2018). Visión y acción 2030 Propuesta de la ANUIES para renovar la educación superior en México: Diseño y concertación de políticas públicas para impulsar el cambio institucional, 184.
- Bajwa, N. M., Yudkowsky, R., Belli, D., Vu, N. V., & Park, Y. S. (2018). Validity Evidence for a Residency Admissions Standardized Assessment Letter for Pediatrics. *Teaching and Learning in Medicine*, *30*(2), 173–183. <https://doi.org/10.1080/10401334.2017.1367297>
- Baladrón, J., Sánchez-Lasheras, F., Romero-Ladrero, J. M., Curbelo, J., Villacampa-Menéndez, P., & Jiménez-Fonseca, P. (2018). Evolución de los parámetros dificultad y discriminación en el ejercicio de examen MIR. Análisis de las convocatorias de 2009 a 2017. *FEM*, *21*(4), 181–193.
- Barajas-Ochoa, A., & Ramos-Remus, C. (2017). Equidad, validez y confiabilidad del Examen

- Nacional para Aspirantes a Residencias Médicas (ENARM): oportunidades para mejorar. *Salud Pública de México*, 59(5, sep-oct), 501. <https://doi.org/10.21149/8769>
- Barajas-Ochoa, A., Ramos-Remus, C., Castillo-Ortiz, J. D., Yáñez, J., Barajas-Ochoa, Z., Sánchez-González, J. M., ... Bustamante-Montes, L. P. (2019). Flaws in the design of the Examen Nacional para Aspirantes a Residencias Médicas produce inequity. *Salud Publica de Mexico*, 61(2), 125–135. <https://doi.org/10.21149/9790>
- Barajas-Ochoa, A., Ramos-Remus, C., Ramos-Gómez, S., Barajas-Ochoa, Z., Sánchez-González, J. M., Hernández-Ávila, M., & Córdova-Villalobos, J. Á. (2019). Desempeño de las escuelas de medicina en México: resultados del Examen Nacional para Aspirantes a Residencias Médicas. *Salud Pública de México*, 61(4, jul-ago), 495. <https://doi.org/10.21149/10042>
- Bergman, E. M., De Bruin, A. B. H., Vorstenbosch, M. A. T. M., Kooloos, J. G. M., Puts, G. C. W. M., Leppink, J., ... Van Der Vleuten, C. P. M. (2015). Effects of learning content in context on knowledge acquisition and recall: A pretest-posttest control group design. *BMC Medical Education*, 15(1), 1–11. <https://doi.org/10.1186/s12909-015-0416-0>
- Birnbaum, R. (1977). Factors Related to University Grade Inflation. *The Journal of Higher Education*, 48(5), 519–539.
- Birolini, D. (2002). Surgery in Brazil. *Archives of Surgery*, 137(3), 352. <https://doi.org/10.1001/archsurg.137.3.352>
- Boulet, J., Bede, C., McKinley, D., & Norcini, J. (2007). An overview of the world's medical schools. *Medical Teacher*, 29(1), 20–26. <https://doi.org/10.1080/01421590601131823>
- Bowles, S. (1967). The Efficient Allocation of Resources in Education. *The Quarterly Journal of Economics*, 81(2), 189–219.
- CIFRHS-CPEC. (2017). *Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud Comité de Posgrado y Educación Continua XLI Examen Nacional para Aspirantes a Residencias Médicas Convocatoria 2017 Contenido*. Retrieved from http://www.cifrhs.salud.gob.mx/site1/enarm/docs/2017/E41_convoy_2017.pdf

- CIFRHS-CPEC. (2018). Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud Comité de Posgrado y Educación Continua XLI Examen Nacional para Aspirantes a Residencias Médicas Convocatoria 2018 Contenido. Retrieved September 20, 2018, from http://www.cifrhs.salud.gob.mx/site1/enarm/docs/2018/E42_convo_2018.pdf
- Clouet-Huerta, D. E., González, B., & Correa, K. (2017). Especialización médica en Chile: Tipos, mecanismos y requisitos de postulación. Una actualización de los procesos para los médicos generales. *Revista Medica de Chile*, *145*(11), 1454–1462. <https://doi.org/10.4067/s0034-98872017001101454>
- Comité de Posgrado y Educación Continua XLI Examen Nacional para Aspirantes a Residencias Médicas 2017 Metodología Proceso de Selección. (n.d.). Retrieved from http://www.cifrhs.salud.gob.mx/site1/enarm/docs/2017/E41_met_proceso_seleccion_2017.pdf
- Consejo Mexicano para la Acreditación de la Educación Médica A. C. (COMAEM). (2018). Estado Global de Acreditación. Retrieved September 20, 2018, from <http://www.comaem.com.mx/wp-content/uploads/2018/09/EDO.-GLOBAL-DE-LA-ACRED..htm>
- Consejo Nacional de Ciencia y Tecnología. (n.d.). Becas nacionales. Retrieved September 23, 2018, from <https://www.conacyt.gob.mx/index.php/becas-y-posgrados/becas-nacionales>
- Daude, C., & Robano, V. (2015). On intergenerational (im)mobility in Latin America. *Latin American Economic Review*, *24*(1). <https://doi.org/10.1007/s40503-015-0030-x>
- Diario Oficial de la Federación. (2018). 2018_09_05_MAT_salud2a.
- Dirección General De Calidad y Educación En Salud. (2010). Examen Nacional para Aspirantes a Residencias Médicas: Características y evolución a su formato electrónico. Retrieved from http://cifrhs.salud.gob.mx/descargas/pdf/enarm_caracteristicas_evolucion.pdf
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830–837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
- Duvivier, R. J., Boulet, J. R., Opalek, A., van Zanten, M., & Norcini, J. (2014). Overview of the

- world's medical schools: An update. *Medical Education*, 48(9), 860–869.
<https://doi.org/10.1111/medu.12499>
- Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: systematic review of the literature. *BMJ*, 324(7343), 952–957.
<https://doi.org/10.1136/bmj.324.7343.952>
- Frenk-Mora J, Robledo-Vera C, Nigenda-López G, Ramírez-Cuadra C, Galván-Martínez O, R.-A. J. (1990). Políticas de formación y empleo de médicos en México 1917-1988. *Salud Pública de México*, 32(1), 440–448.
- García Acosta, V. M., & Crocker Sagastume, René; Abreu Hernández, Luis Felipe; Parra Acosta, Haydeé; Vázquez Esquivel, Jesús; Ojeda Blanco, Carlos; Alomía, Harold; Contreras Mayén, Gustavo; Ramos Herrera, Igor; Rodríguez Arroyo, Christian; Vázquez Castellanos, J. L. (2015). *La educación médica en México: Visión estratégica del cuerpo académico de AMFEM*.
- Gelinne, A., Zuckerman, S., Benzil, D., Grady, S., Callas, P., & Durham, S. (2018). United States Medical Licensing Exam Step I Score as a Predictor of Neurosurgical Career Beyond Residency. *American Association of Neurological Surgeons Annual Scientific Meeting*, 29.
<https://doi.org/10.1093/neuros/nyy313>
- Guerrero-Avila, J. J., Romero-González, J., & Noriega, R. (2013). Análisis de la competencia en medicina visto a través del EGEL-MG. *CULCyT*, 10(49), 26–36. Retrieved from openjournal.uacj.mx/ojs/index.php/culcyt/article/download/157/152
- Guevara-Guzmán, R., Galván, M., & Muñoz-Comonfort, A. (2007). El desempeño académico de los estudiantes de la Facultad de Medicina en el primer año de la carrera. *Gaceta Medica de Mexico*, 143(1), 27–32.
- Hernández-Gálvez, D. C., & Roldán-Valadez, E. (2019). Mexican ENARM: Performance comparison of public vs. private medical schools, geographic and socioeconomic regions. *Salud Publica de Mexico*, 61(5), 637–647. <https://doi.org/10.21149/10078>

- Himelfarb, I. (2019). A primer on standardized testing: History, measurement, classical test theory, item response theory, and equating. *Journal of Chiropractic Education*, 33(2), 151–163. <https://doi.org/10.7899/JCE-18-22>
- Informe Anual de Resultados 2014. EXAMEN GENERAL PARA EL EGRESO DE LA LICENCIATURA EN MEDICINA GENERAL (EGEL-MEDI). Centro Nacional de Evaluación para la Educación Superior, A. C. (CENEVAL). (2015). Retrieved from <http://www.ceneval.edu.mx/acervo-digital>
- Jiménez-Sastré Alejandro, Quevedo-Tejero Elsy del Carmen, Hernández-Ortega Hilda María, Fócil-Némiga Eduardo, Z.-G. M. A. (2015). Desempeño histórico de egresados de una universidad mexicana en el examen de residencias médicas, 2001-2013. *Investigación En Educación Médica*, 4(14), 69–74. [https://doi.org/10.1016/S2007-5057\(15\)30004-1](https://doi.org/10.1016/S2007-5057(15)30004-1)
- Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, 68(6), 540–546. <https://doi.org/10.4097/kjae.2015.68.6.540>
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, (May), 0013189X2091279. <https://doi.org/10.3102/0013189x20912798>
- Mahajan, R., & Singh, T. (2017). *The National Licentiate Examination: Pros and cons*. Retrieved from <http://www.nmji.in>
- Manterola, C., & Otzen, T. (2014). Estudios observacionales. Los Diseños Utilizados con Mayor Frecuencia en Investigación Clínica. *Int.J.Morphol*, 32(2), 634–645. <https://doi.org/10.4067/S0717-95022014000200042>
- McGaghie, W. C., Cohen, E. R., & Wayne, D. B. (2011). Are United States medical licensing exam step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Academic Medicine*, 86(1), 48–52. <https://doi.org/10.1097/ACM.0b013e3181ffacdb>
- Mendoza-Aguilar, A. N., Salazar-Ruiz, S. Y., Cardiel-Marmolejo, L. E., & Roldan-Valadez, E. (2020). Comparison of scores achieved by foreign medical graduates in the National Evaluation for Medical Residency (ENARM) in Mexico. *Postgraduate Medical Journal*,

96(1132), 108–109. <https://doi.org/10.1136/postgradmedj-2019-136893>

- Nigenda, G., & Muños, J. A. (2015). Projections of specialist physicians in Mexico: A key element in planning human resources for health. *Human Resources for Health, 13*(1), 1–13. <https://doi.org/10.1186/s12960-015-0061-z>
- Petra-Micu, I., Cortés-Morelos, J., Talayero-Uriarte, J. A., & Fouilloux-Morales, M. (2012). Representación semántica del concepto " medicina " en estudiantes de la Facultad de Medicina, UNAM. Un estudio longitudinal. *Inv Ed Med Col Copilco Universidad. Deleg. Coyoacán. México D. F. CP, I(04510)*, 49–56. Retrieved from www.elsevier.com.mx
- Prieto-Miranda, S. E., Jiménez-Bernardino, C. A., Jiménez-Bernardino, J. A., & Esparza-Pérez, R. I. (2013). ¿Sirve el examen nacional de residencias médicas de México? *Medicina Interna de Mexico, 29*(5), 518–524.
- Provost, L. P. (2011). Analytical studies: A framework for quality improvement design and analysis. *BMJ Quality and Safety, 20*(SUPPL. 1), i92–i96. <https://doi.org/10.1136/bmjqs.2011.051557>
- Ramiro-H, M., Cruz-A, E., Zerón-Gutiérrez, L., & Arévalo-Vargas, A. (2017). El ENARM y las escuelas y facultades de medicina. Un análisis que no le va a gustar a nadie. The ENARM and the schools and faculties of medicine. An analysis that nobody will like. *Rev Med Inst Mex Seguro Soc, 55*(4), 498–511. Retrieved from <http://www.medigraphic.com/pdfs/imss/im-2017/im174n.pdf>
- Ray, C., Bishop, S. E., & Dow, A. W. (2018). Rethinking the Match: A Proposal for Modern Matchmaking. *Academic Medicine, 93*(1), 45–47. <https://doi.org/10.1097/ACM.0000000000001781>
- Rodríguez Carranza, R., Gatica-Lara, F., Martínez-González, A., Rosa Medina-Martínez, S., & Rodríguez-Carranza, R. (2013). *Presencia de las ciencias básicas en el Examen Nacional de Aspirantes a Residencias Médicas (ENARM)*. *Inv Ed Med* (Vol. 2). Retrieved from www.elsevier.com.mx

- Romero-Valle, S., & González-Pérez, G. J. (2015). *Desempeño de los egresados de medicina en el Examen Nacional de Residencias Médicas: Un estudio en la Universidad de Guadalajara. Universidad de Guadalajara Centro Universitario de Ciencias de la Salud* (1a Edición, Vol. 1). Guadalajara, Jalisco. <https://doi.org/10.1017/CBO9781107415324.004>
- Rubright, J. D., Jodoin, M., & Barone, M. A. (2018). Examining Demographics, Prior Academic Performance, and United States Medical Licensing Examination Scores. *Academic Medicine*, 1. <https://doi.org/10.1097/ACM.0000000000002366>
- Sánchez Mendiola, M. (2017). Mañana tengo examen final , ahora sí tengo que estudiar : ¿ estudiar para aprender o para pasar exámenes? *Investigación En Educación Médica*, 6(21), 1–2. <https://doi.org/10.1016/j.riem.2016.12.002>
- Sánchez Mendiola, M., & Delgado Maldonado, L. (2017). Exámenes de alto impacto: implicaciones educativas. *Investigación En Educación Médica*, 6(21), 52–62. <https://doi.org/10.1016/j.riem.2016.12.001>
- Sawyer, S. F. (2009). Analysis of Variance: The Fundamental Concepts. *Journal of Manual & Manipulative Therapy*, 17(2), 27E-38E. <https://doi.org/10.1179/jmt.2009.17.2.27e>
- Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Shi, D., DiStefano, C., McDaniel, H. L., & Jiang, Z. (2018). Examining Chi-Square Test Statistics Under Conditions of Large Model Size and Ordinal Data. *Structural Equation Modeling*, 25(6), 924–945. <https://doi.org/10.1080/10705511.2018.1449653>
- Shumway, J. M., Harden, R. M., Shumway, J. M., & Harden, R. M. (2009). AMEE Guide No . 25 : The assessment of learning outcomes for the competent and reflective physician AMEE Guide No . 25 : The assessment of learning outcomes for the competent and reflective physician, (25). <https://doi.org/10.1080/0142159032000151907>
- Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the P Value Is Not Enough .

Journal of Graduate Medical Education, 4(3), 279–282. <https://doi.org/10.4300/jgme-d-12-00156.1>

Tackett, S., Zhang, C., Nassery, N., Caufield-Noll, C., & Van Zanten, M. (2019). Describing the Evidence Base for Accreditation in Undergraduate Medical Education Internationally: A Scoping Review. *Academic Medicine*, 94(12), 1995–2008. <https://doi.org/10.1097/ACM.0000000000002857>

Taveira-Gomes, T., Prado-Costa, R., Severo, M., & Ferreira, M. A. (2015). Characterization of medical students recall of factual knowledge using learning objects and repeated testing in a novel e-learning system. *BMC Medical Education*, 15(1), 1–12. <https://doi.org/10.1186/s12909-014-0275-0>

Tendencias en el aprendizaje de la educación en México: Una comparación entre ENLACE, Excale y Pisa | Nexos. (n.d.). Retrieved December 2, 2018, from <https://www.nexos.com.mx/?p=22749>

Van Wieren, A., Palazuelos, L., Elliott, P. F., Arrieta, J., Flores, H., & Palazuelos, D. (2014). Service, training, mentorship: First report of an innovative education-support program to revitalize primary care social service in Chiapas, Mexico. *Global Health Action*, 7(1), 1–9. <https://doi.org/10.3402/gha.v7.25139>

Van Zanten, M., Norcini, J. J., Boulet, J. R., & Simon, F. (2008). Overview of accreditation of undergraduate medical education programmes worldwide. *Medical Education*. <https://doi.org/10.1111/j.1365-2923.2008.03092.x>

Vázquez Martínez, F. D. (2017). Apuntes históricos sobre la formación de médicos especialistas en México desde la evolución educativa. *Investigación En Educación Médica*. <https://doi.org/10.1016/j.riem.2017.03.003>

Vázquez Martínez, F. D., & Ortiz-León, M. C. (2017). Accreditation and results from medical education in Mexico. *Educacion Medica*. <https://doi.org/10.1016/j.edumed.2017.04.016>

Weggemans, M. M., van Dijk, B., van Dooijeweert, B., Veenendaal, A. G., & ten Cate, O. (2017).

The postgraduate medical education pathway: An international comparison. *GMS Journal for Medical Education*, 34(5). <https://doi.org/10.3205/zma001140>

Westerman, M. E., Boe, C., Bole, R., Turner, N. S., Rose, S. H., Gettman, M. T., & Thompson, R. H. (2019). Evaluation of Medical School Grading Variability in the United States: Are All Honors the Same? *Academic Medicine*, 94(12), 1939–1945. <https://doi.org/10.1097/ACM.0000000000002843>

Wimsatt, L. A., Cooke, J. M., Biggs, W. S., & Heidelbaugh, J. J. (2016). Institution-Specific Factors Associated With Family Medicine Residency Match Rates. *Teaching and Learning in Medicine*, 28(3), 269–278. <https://doi.org/10.1080/10401334.2016.1159565>

Zimmerman, D. W. (1987). Comparative power of student T test and Mann-Whitney U test for unequal sample sizes and variances. *Journal of Experimental Education*, 55(3), 171–174. <https://doi.org/10.1080/00220973.1987.10806451>