Comparison between Mexican and International Medical Graduates' scores in the ENARM Competing for *Clinical Specialities* in Mexico during 2012-2019: Data Visualization, Trends and Forecasting Analyses

Alexela-Nerey Mendoza-Aguilar[1], Aime Cedillo-Pozos[1] and Ernesto Roldan-Valadez[1,2,*]

[1]*Directorate of Research, Hospital General de Mexico "Dr Eduardo Liceaga", Mexico City, Mexico*

[2]*I.M. Sechenov First Moscow State Medical University (Sechenov University), Department of Radiology, Moscow, Russia*

Abstract: *Objectives*: Because there is heterogeneity in the ENARM scores obtained between Mexicans and International medical graduates (IMG) in the eight clinical specialities with direct-entry (*Anesthesiology, and Emergency Medicine. Geriatrics, Internal Medicine, Medical Genetics, Pediatrics, Pneumology, Psychiatry*), we aimed to evaluate those scores. We hypothesized that Mexican test-takers achieve higher scores than IMG with significant growth trends in their exam scores.

Methods: This study was cross-sectional, used historical data from the annual public report of the ENARM for eight years (2012 to 2019). We compare the minimum (MinSco) and maximum (MaxSco) scores of each speciality using ANOVA. Mexican versus IMG scores were evaluated with an independent student t-test, trends with Spearman's correlation coefficient, and a 5-years forecasting trend.

Results: There was a significant difference among the MinSco for five surgical specialities; F (7, 115) = 26.611, p = < .001; the global mean of MinSco was 69.133; specialities above this mean were *Internal Medicine, Anesthesiology, Pediatrics, and Pneumology*. The global mean for MaxSco was 79.422; five specialities were above: *Internal Medicine, Pneumology, Geriatrics, Psychiatry,* and *Medical Genetics.* We did not find a significant difference in the MinSco between Mexicans and IMG, but a significant difference was found in the MaxSco between both groups.

Conclusions: ENARM represents a market of high-performance test-takers across the clinical specialities. Mexicans and IMG achieved similar entrance scores, but Mexicans showed a higher MaxSco over IMG in all clinical specialities.

Keywords: ENARM, internship and residency, medical education, medicine speciality.

INTRODUCTION

Education of Graduated Doctors

The residence is a critical step in graduated doctors' education since 90% aspire to a postgraduate or medical speciality [1]. In the USA, up to 88% of general practitioners will eventually study a medical speciality; this percentage decrease to 35% in Mexico [2]. The score that a general practitioner (GP) obtains in the National Evaluation for Medical Residency Applicants (ENARM, *Examen Nacional de Aspirantes a Residencias Medicas*) is the entrance door to a specialization course endorsed by a Mexican University [3, 4].

Logistics of the ENARM

The ENARM is a one-step only exam that uses multiple-choice questions and computerized patient cases to assess examinees' knowledge related to foundational science concepts applicable to medical and scientific theories to clinical medicine; details concerning the logistics' of the exam has been published previously [5, 6].

In Mexico, the Interinstitutional Commission issued the reports for Human Resources Training for Health (CIFRHS, *Comisión Interinstitucional para la Formación de Recursos Humanos para la Salud*) is an inter-institutional, consultation, advisory and technical support organization of the Ministry of Public Education and the Ministry of Health [7]; it considers 27 medical specialities with a direct entry [8]. For the Mexican educational institutions, the ENARM scores and the percentages of their graduates' selection are indicators of efficiency and reason of prestige and even of propaganda among the aspirants to study medicine [9].

Conceptual Framework

Conceptual frameworks play an important, essential role in identifying the nature of education problems and in formulating solutions or designing studies [10]. Each year the number of applicants to the Mexican assessment known as ENARM increase; in 2019, there

*Address correspondence to this author at the Directorate of Research, Hospital General de Mexico, Dr. Balmis 148, Colonia Doctores, Delegacion Cuauhtemoc, 06726 Ciudad de Mexico, Mexico; Tel: +52155-2789-2000 ext. 1149; E-mail: ernest.roldan@usa.net

E-ISSN: 1929-6029/21

were more than 57,000 applicants, and only 9,668 Mexican and international medical graduates (IMG) were selected [11]. Several problems about the ENARM have been addressed in recent publications, for example, the *number of Mexican test-takers and accepted GPs* belonging to each Mexican medical school registered in the ENARM [3]; the *logistics and transparency* of the ENARM exam [5]; the *performance of private versus public schools* using a summary measures method, exploring significant differences in the performance based on geographic regions and socioeconomic level of the Mexican states to which each school belongs [3, 12]; and the *assessment of the assumption of equity* in the ENARM [6].

There is an educational problem in Mexico related to the applicant's heterogenous ENARM scores to clinical specialities [1, 13, 14]. We do not know the eight clinical specialities' academic performance with a direct entry: *Anesthesiology, Emergency Medicine, Geriatrics, Internal Medicine, Medical Genetics, Pediatrics, Pneumology,* and *Psychiatry* [14].

We aimed to assess these eight direct-entry clinical specialities' performance and compare the scores of Mexican versus IMG in each speciality; we also included a trend analysis over eight years (2012-2019). We hypothesized that Mexican test-takers achieve higher scores than IMG with significant growth trends in their exam scores.

MATERIALS AND METHODS

Study Design and Data Acquisition

This study was cross-sectional and used historical data that did not require approval by an Institutional Review Board. We based our analyses on the annual public report of the ENARM for eight years from 2012 to 2019 issued by the CIFRHS. The reports contained quantitative information on each medical speciality's academic performance from graduate physicians who took the ENARM; these reports are freely available as PDF files at the CIFRHS website [11]. Original data are included as an online-only supplementary file.

Logistics of ENARM and Assessed Variables

Five test forms are created each year, each comprising 450 multiple-choice single-best answer items; no item is used in more than one test form. All test forms contain the same number of items per area of knowledge (speciality/subspecialty), with an approximate item distribution of 37.5% internal medicine, 25% paediatrics, 22% gynaecology-obstetrics, and 15% surgery. Applicants for each speciality are ranked from highest to lowest according to their total ENARM score. Ranked applicants receive a 'pass' certificate until the quota is met according to that speciality's available positions [6].

For each year (2012-2019), we recorded the minimum and maximum scores (calculated by dividing the absolute number of correct answers by the total number of items) clustered by nationality (Mexican or IMG) and chosen speciality (8 direct-entry specialities) that coincidentally appear in the annual CIFRHS report.

Statistical Analysis and Data Visualization Techniques

Part I, Comparison of the Minimum and Maximum Scores among Surgical Specialities

In the first part of our analysis, we compare the minimum (MinSco) and maximum (MaxSco) scores of the eight direct-entry clinical specialities evaluated by the ENARM (*Anesthesiology, Emergency Medicine, Geriatrics, Internal Medicine, Medical Genetics, Pediatrics, Pneumology, Psychiatry)*; the Kolmogorov-Smirnoff and Shapiro-Wilk tests showed a non-significant p-value for each speciality, which indicated a normal distribution of data in both variables (MinSco and MaxSco). Then, we performed a one-way ANOVA to reveal the differences in the scores achieved by each speciality; variables were tested for homogeneity of variance, and posthoc tests used the LSD (least significant difference) method. To test the assumption that MinSco and MaxSco increase every year, we assessed a significant linear trend for the scores to increase across the specialities. For this assessment, we use the *Polynomial* option (in the ANOVA menu of SPSS); it chose the Degree: Linear (default) option in its Contrast box. Detailed descriptions of the ANOVA test in clinical settings have been previously published by our group [15, 16]. Descriptive statistics were used for each variable and 95% confidence intervals (C.I.) [17]. The effect size assessment (proportion of the variance in the dependent variable that the independent variable can explain) of each result was obtained using the Partial Eta Squared (η^2). Partial eta squared was defined as the ratio of variance associated with an effect, plus that effect and its associated error variance.

$$\eta^2 = SS_{effect} / SS_{effect} + SS_{error} \text{ , where:}$$

SS_{effect} is the sums of squares for the effect the researcher is studying.

The values of η^2 were classified in three groups 0.01 to 0.06 = small effect, 0.06 to 0.14 = moderate impact, and > 0.14 = substantial effect [18].

To visualize the results, we use graph lines showing the evolution of MinSco and MaxSco every year for each speciality. We also drew bar graphs with the global means indicating those specialities whose mean were above or below a global mean for all specialities.

Part II, Comparison of the Minimum and Maximum Scores between Mexican and IMG, Correlations, Trend Lines and Forecasting Analyses

For the second part of our analysis, we looked for significant differences between Mexican and IMG in their scores by independently analyzing each speciality.

The Comparison of means was made using the independent T-test. The Pearson's correlation coefficient helped us to reveal direction trends: positive for increasing scores (↑) with every year (2012 to 2019) or negative for decreasing scores (↓).

Linear Trend Lines

We calculated the trend of the MinSco and MaxSco every year for each speciality,

Linear trend lines are lines of best fit used to estimate a linear relationship in the data. They have the following form:

$$Y = \beta_0 + \beta_1 X ,$$

where Y is the dependent variable, and X is the independent variable that affects it. They represent the simplest trend line model in that they estimate a relationship that is increasing or decreasing at a steady rate β_1 and are therefore best used when the trend of the data resembles a linear pattern. We reported the p-values and the R-squared (a measure of how well the trend line fits the data). The latter considered the best indicator of model performance.

Forecasting Analyses

We forecasted our quantitative time-series data using a triple exponential smoothing method, which is also called *Holt-Winters exponential smoothing* [19, 20]. It was applied using ©Tableau software. This method is used for forecasting the univariate time series when the data might have both linear trend and seasonal pattern. In Holt-Winters exponential smoothing, recent observations are given relatively more weight than older observations; it is suitable for short-term forecasting and uses the maximum likelihood function for estimating parameters [21]. We calculated models that captured the evolving trend or seasonality of the data and extrapolated them into the future five-year period with 95% confidence prediction intervals.

The triple exponential smoothing formulas are given by:

$$s_0 = x_0$$

$$s_t = \alpha \frac{x_t}{c_t - L} + (1+\alpha)(s_{t-1} + b_{t-1})$$

$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$

$$c_t = \gamma \frac{x_t}{s_t} + (1-\gamma)c_t - L$$

where,

st = smoothed statistic, it is the simple weighted average of current observation xt

st-1 = previous smoothed statistic

α = smoothing factor of data; 0 < α < 1

t = time period

bt = best estimate of a trend at time t

β = trend smoothing factor; 0 < β <1

ct = sequence of seasonal correction factor at time t

γ = seasonal change smoothing factor; 0 < γ < 1

The model used to generate the forecast had three components: Level, Trend, and Season. The value for each component might be one of the following:

1. *None*: The component is not present in the model.
2. *Additive*: The component is present and is added to the other components to create the overall forecast value.
3. *Multiplicative*: The component is present and is multiplied by the other components to create the overall forecast value.

The QUALITY OF THE MODEL was evaluated with five statistical values:

RMSE: Root mean squared error.

MAE: Mean absolute error.

MASE: Mean absolute scaled error.

MAPE: Mean absolute percentage error.

AIC: Akaike information criterion.

The smoothing coefficients were optimized to weigh more recent data values over older ones, such that within-sample one-step-ahead forecast errors were minimized.

Alpha is the level smoothing coefficient,

Beta is the trend smoothing coefficient, and

Gamma is the seasonal smoothing coefficient.

The closer a smoothing coefficient was to 1.00, the less smoothing was performed, allowing for rapid component changes and heavy reliance on recent data. The closer a smoothing coefficient was to 0.00, the more smoothing was performed, allowing for gradual component changes and less reliance on recent data [22].

The forecasting method calculated a 5-years trend in the MinSco and MaxSco of each speciality; we detected a crossing point between Mexican and IMG for each medical speciality.

We used our previously calculated global means for the MinSco and MaxSco to group the Mexican and IMG in specialities that lay above or below each speciality's mean.

Score comparisons were performed using the IBM® SPSS® Statistics software (version 26.0.0.1 IBM Corporation; Armonk, NY, USA). Data visualization of the scores, trends, and forecasting analyses were performed using ©Tableau software (version 2019.1.3, Seattle, Washington, USA). Statistical significance considered a p-value < 0.05 (two-tailed).

RESULTS

Scores Included in the Analysis

For each score (MinSco and MaxSco), we evaluated 128 measures, 16 for each speciality (8 scores for Mexicans and 8 for IMG for the years 2012 to 2019), with a total of 256 measures included. However, from the 256 total number of scores, we substracted 24 scores corresponding to some years in which some specialities did not have test-takers; then, a total of 232 scores were included in the analysis.

Grouping of Specialities above or below a Global Mean

We calculated a MinSco global mean of 69.133. Specialities above this mean were Internal Medicine, Anesthesiology, Pediatrics, and Pneumology. Specialities below the mean corresponded to Psychiatry, Geriatrics, Medical Genetics, and Emergency Medicine.

The global mean for the *MaxSco* was 79.422, and five specialities were above this mark: *Internal Medicine, Pneumology, Geriatrics, Psychiatry,* and *Medical Genetics.* The other four specialities below the global mean were *Pediatrics, Anesthesiology, and Emergency Medicine.* Figures **1A** and **B** show the scores above or below the global mean for surgical specialities.

Comparison of Minimum and Maximum Scores Achieved by Surgical Specialities

The one-way ANOVA depicted a significant difference among the minimum scores achieved by the eight clinical specialties; F (7, 115) = 26.611, p = < .001; the η^2 = 0.632 indicated a great effect size. Post-hoc tests showed significant differences between each surgical speciality (bonferronni adjusted p-value = .006). Only two pairs of speciality-comparisons were non-significant:

Anesthesiology vs Medical Genetics (p = 0.010), and *Anesthesiology* vs *Pediatrics* (p = 0.039). There was a significant linear trend for the increasing scores with every year F (7, 115) = 4.167, p = < .044; the η^2 = 0.033 indicated a small effect size.

We also found a significant ANOVA test in the Comparison of the MaxSco between surgical specialities, F (7, 115) = 5.561, p < 0.001, which pointed a difference in the MaxSco among the eight specialities; the η^2 = 0.264 indicated a great effect size. Post-hoc tests showed significant differences between seven pairs of specialty: *Anesthesiology vs Emergency Medicine* (p = 0.001), *Anesthesiology vs Medical Genetics* (p = 0.003), *Emergency Medicine vs Internal Medicine* (p < .001), *Emergency Medicine vs Pediatrics* (p = 0.003), *Geriatrics vs Internal Medicine* (p < .001), *Internal Medicine vs Medical Genetics* (p < .001),

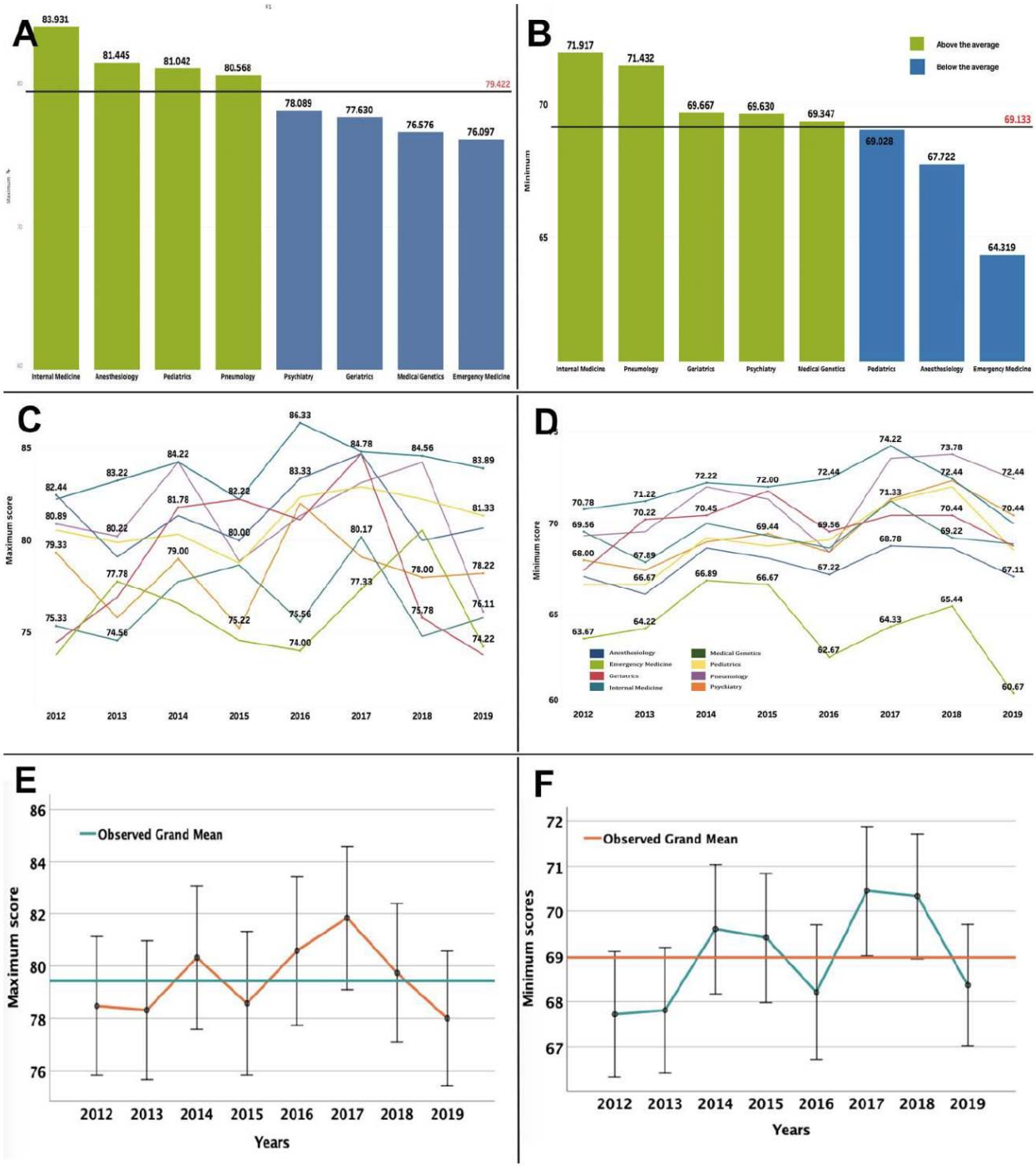

Figure 1: A-B, Scores above or below the global mean for surgical specialities. **C-D**, mean Comparison of surgical specialities showing the trend by year. E-F, global trend of the MinSco and MaxSco for eight years (2012 to 2019).

Internal Medicine vs Psychiatry (p = .001). The test for a linear trend of the MaxSco with every year did not show significance F (7, 115) = .360, p = 0.550; with a small effect size, η^2 = 0.003. Figures **1C** and **D** show a comparison of the means of each clinical speciality by year. Figures **1E** and **F** depict the global performance of the MinSco and MaxSco for eight years (2012 to 2019).

Table **1** depicts the means, standard deviation, standard error, and 95% CI for the MinSco and MaxSco scores in each speciality.

Comparison of Minimum and Maximum Scores between Mexicans and IMG in each Clinical Speciality

For the MinSco, it was very interesting to notice that the IMG got higher scores for all clinical specialities.

Table 1: Means, Standard Deviation, Standard Error, and 95% Confidence Intervals (C.I.) for the Minimum and Maximum Scores in each Speciality

Minimum scores							
Specialities	**Mean**	**Std. Deviation**	**Std. Error**	**95% Confidence Interval for Mean**		**Minimum**	**Maximum**
				Lower Bound	**Upper Bound**		
Anesthesiology	67.722	1.923	0.481	66.698	68.747	63.778	70.666
Emergency medicine	64.319	2.155	0.539	63.171	65.467	60.000	68.000
Geriatrics	69.667	1.482	0.428	68.725	70.608	67.111	71.778
Internal medicine	71.917	1.254	0.313	71.248	72.585	70.000	74.444
Medical Genetics	69.347	1.332	0.333	68.637	70.057	67.556	72.889
Pediatrics	69.028	1.855	0.464	68.039	70.016	66.445	72.000
Pneumology	71.432	2.109	0.703	69.811	73.053	68.444	74.222
Psychiatry	69.630	1.874	0.484	68.592	70.667	67.111	73.778
Maximum scores							
Anesthesiology	81.445	3.126	0.781	79.779	83.110	75.111	86.667
Emergency medicine	76.097	4.916	1.229	73.478	78.717	68.889	83.111
Geriatrics	77.630	6.023	1.739	73.803	81.456	67.778	84.667
Internal medicine	83.931	3.575	0.894	82.026	85.835	77.777	89.333
Medical Genetics	76.576	5.070	1.268	73.875	79.278	69.779	84.444
Pediatrics	81.042	4.748	1.187	78.511	83.572	73.556	89.111
Pneumology	80.568	3.234	1.078	78.082	83.054	74.222	84.223
Psychiatry	78.089	5.049	1.304	75.293	80.885	69.555	86.000

However, *Anesthesiology* was the only speciality with a significant difference between Mexicans and IMG. For Mexicans, the highest score was *Internal Medicine*, but for the IMG was *Pneumology*.

For the MaxSco, we observed exactly the reverse trend, Mexicans got the higher scores in all the specialities, and the differences between scores were all statistically significant. For Mexicans and IMG, the highest score was Internal Medicine, the lowest for Mexicans was Emergency Medicine and IMG *Geriatrics*. Table **2** depicts the means, SD, standard error of the mean between Mexicans and IMG for each clinical speciality; p-values were calculated with the independent t-test.

Positive and Negative Trends in the Minimum and Maximum Scores vs Years (Mexicans and IMG) in each Clinical Speciality

For the MinSco in *Anesthesiology*, only Mexicans showed a positive and significant correlation, R = .849, p = .008; and for IMG R = -.511, p = .196. In *Emergency medicine*, both groups depicted a nonsignificant negative correlation, Mexicans R = -.446, p = .268; IMG R = -.298, p = .474. For *Geriatrics*, both groups showed a positive non-signicant association; Mexicans R = .333, p = .420; IMG R = .059, p = .941. *Internal medicine* showed a similar behaviour than *Geriatrics,* Mexicans R = .225, p = .591; IMG R = .095, p = .823. In *Medical Genetics* Mexicans had a positive non-signicant correlation,

R = .612, p = .107; while in IMG the correlation was negative and nonsignificant, R = -.102, p = .810. *Pediatrics* presented a strong, positive significant correlation in Mexicans R = .743, p = .035; while IMG had a positive non-signicant one, R = .666, p = .071. *Pneumology* the corrrelation for this specialty was calculated only for Mexicans, R = .480, p = .228; the IMG group did not have enough test-takers in different years to calculate those values. Finally, in *Psychiatry*, both groups depicted a significant, positive correlation between the MinSco and the years of exam; Mexicans R = .820, p = .013; IMG R = .866, p = .012.

For the Maximum score, in *Anesthesiology*, both groups had a non-significant correlation, negative in Mexicans R = -.187, p = .657; and positive in IMG R = .271, p = .516. *Emergency medicine*, showed a similar nonsignificant trend; negative in Mexicans R = -.006, p

Table 2: Comparison of Scores between Mexican and International Medical Graduates

Minimum scores							
Specialities	**Mexican**			**IMG**			**p-value**
	Mean	**Std. Deviation**	**Std. Error Mean**	**Mean**	**Std. Deviation**	**Std. Error Mean**	
Internal medicine	71.806	1.293	0.457	72.028	1.292	0.457	0.736
Pneumology	71.083	1.957	0.692	74.222	-	-	0.174
Geriatrics	69.611	1,439	0.509	69.778	1.787	0.894	0.864
Psychiatry	68.917	1.382	0.489	70.444	2.124	0.803	0.118
Pediatrics	68.861	1.928	0.682	69.194	1.896	0.670	0.733
Medical Genetics	68.778	0.742	0.262	69.916	1.585	0.560	0.096
Anesthesiology	66.472	1.626	0.575	68.972	1.306	0.462	0.004*
Emergency medicine	63.639	1.991	0.704	65.000	2.219	0.785	0.218
Máximum scores							
Specialities	**Mexican**			**IMG**			**p-value**
	Mean	**Std. Deviation**	**Std. Error Mean**	**Mean**	**Std. Deviation**	**Std. Error Mean**	
Internal medicine	86.944	1.561	0.552	80.917	2.045	0.723	< 0.000*
Pediatrics	85.222	2.047	0.724	76.861	2.045	0.723	< 0.000*
Anesthesiology	83.278	2.357	0.833	79.611	2.775	0.981	0.013*
Psychiatry	82.222	1.926	0.681	73.365	2.516	0.951	< 0.000*
Geriatrics	81.556	1.671	0.591	69.778	1.787	0.894	< 0.000*
Pneumology	81.361	2.341	0.828	74.222	-	-	0.024*
Medical Genetics	80.944	2.183	0.772	72.209	2.592	0.916	< 0.000*
Emergency medicine	80.000	2.181	0.771	72.195	3.496	1.236	< 0.000*

= .989; and positive in IMG R = .317, p = .444. *Geriatrics* repeated the trend of the two previos specialties; Mexicans R = -.264, p = .528; IMG R = .059, p = .941. *Internal medicine* showed a positive non-signicant correlation for both groups; Mexicans R = .075, p = .860; IMG R = .631, p = .094. *Medical Genetics* Mexicans had a negative, nonsignificant association, R = -.077, p = .856; while a positive, nonsignificant correlation in IMG, R = .251, p = .548. *Pediatrics* presented a strong, positive significant correlation in Mexicans R = .791, p = .019; while IMG had a weak, positive, non-signicant one, R = .016, p = .970. *Pneumology* the corrrelation for this specialty was calculated only for Mexicans, R = -.014, p = .974; the IMG group did not have enough test-takers in different years to calculate those values. Finally, in *Psychiatry*, Mexicans had a positive nonsignificant correlation R = .471, p = .239; while it was negative, and nonsignificant for IMG R = -.237, p = .609. Table **3** shows a table of the correlations between the minimum and maximum scores vs years (up and down arrows) grouped by Mexican or IMG and their statistical significance.

Modelling of Linear Trends

All linear trend models were computed for the median *Maximum* or *Minimum* given years according to the formula:

Type of test-taker * (Year of years + Intercept)

Table **4** shows the R-Squared and p-values of the trend lines for the minimum and maximum scores grouped by the eight selected specialities. Figure **2** depicts the mathematical model for each trend lines grouped by medical speciality.

Figure **3** shows the graphical representation of the observed means and linear trends for both Min and Max scores.

Comparison of 5-Year Forecasting Trends between the Minimum and Maximum Scores of Mexicans and IMG

We identified convergent and divergent forecasting trends between each speciality's minimum and

Table 3: Significant Trends in the Minimum and Maximum Scores between Mexican and International Medical Graduates

Score	Test-taker	Medical specialty			
		Significant	Trend	Non-significant	Trend
Minimum	Mexican	Anesthesiology	↑	Emergency medicine	↓
		Paediatrics	↑	Geriatrics	↑
		Psychiatry	↑	Internal medicine	↑
				Medical Genetics	↑
				Pneumology	↑
	International Medical Graduates	Psychiatry	↑	Anesthesiology	↓
				Emergency medicine	↓
				Geriatrics	↑
				Internal medicine	↑
				Medical Genetics	↓
				Paediatrics	↑
				Pneumology	*
Maximum	Mexican	Emergency medicine	↓	Anesthesiology	↓
		Paediatrics	↑	Geriatrics	↓
				Internal medicine	↑
				Medical Genetics	↓
				Pneumology	↓
				Psychiatry	↑
	International Medical Graduates			Anesthesiology	↑
				Emergency medicine	↑
				Geriatrics	↑
				Internal medicine	↑
				Medical Genetics	↑
				Paediatrics	↑
				Pneumology	*
				Psychiatry	↓

↑ Positive growing trend; ↓ negative growing trend; * insufficient test-takers in different years to calculate the correlation.

Table 4: R-Squared and p-Values of the Trend Lines for the Minimum and Maximum Scores Grouped by the Eight Selected Specialities

Medical Speciality	Linear trend model (Median of scores per given year)	R-Squared	p-value
Anesthesiology	Maximum	.403	.092
	Minimum	.748	< .001
Emergency medicine	Maximum	.696	.002
	Minimum	.229	.354
Geriatrics	Maximum	.930	< .001
	Minimum	.071	.891
Internal Medicine	Maximum	.819	< .001
	Minimum	.038	.922
Medical genetics	Maximum	.800	< .001
	Minimum	.256	.297
Pediatrics	Maximum	.881	< .001
	Minimum	.503	.033
Pneumology	Maximum	.873	< .001
	Minimum	.839	< .001
Psychiatry	Maximum	.843	< .001
	Minimum	.774	< .001

Anesthesiology

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Median Maximum	IMG	0.516546	6	Year of Years	0.0008399	0.0012189	0.689036	0.516546
				intercept	44.1793	51.4328	0.858972	0.42334
Median Maximum	Mexican	0.657594	6	Year of Years	-0.0004925	0.0010566	-0.466087	0.657594
				intercept	104.053	44.583	2.33393	0.0583265
Median Minimum	IMG	0.195733	6	Year of Years	-0.0007457	0.0005123	-1.45566	0.195733
				intercept	100.433	21.6172	4.64599	0.0035184
Median Minimum	Mexican	0.007687	6	Year of Years	0.0015426	0.0003922	3.9329	0.007687
				intercept	1.39315	16.5506	0.0841752	0.935655

Emergency Medicine

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Minimum	IMG	0.473805	6	Year of Years	-0.0007386	0.0009667	-0.764046	0.473805
				intercept	96.1591	40.7899	2.35742	0.0564825
Minimum	Mexican	0.268257	6	Year of Years	-0.000992	0.0008132	-1.21995	0.268257
				intercept	105.49	34.3123	3.0744	0.0218188
Maximum	IMG	0.444152	6	Year of Years	0.001239	0.001513	0.818887	0.444152
				intercept	19.9239	63.844	0.312072	0.765547
Maximum	Mexican	0.98945	6	Year of Years	-1.371e-05	0.000995	-0.0137834	0.98945
				intercept	80.5784	41.9843	1.91925	0.103373

Geriatrics

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Maximum	IMG	0.940742	2	Year of Years	8.256e-05	0.0009834	0.083951	0.940742
				intercept	66.2952	41.5008	1.59744	0.251257
Maximum	Mexican	0.528611	6	Year of Years	-0.0004919	0.0007357	-0.668615	0.528611
				intercept	102.309	31.0451	3.29549	0.0165005
Minimum	IMG	0.940742	2	Year of Years	8.256e-05	0.0009834	0.083951	0.940742
				intercept	66.2952	41.5008	1.59744	0.251257
Minimum	Mexican	0.419588	6	Year of Years	0.0005362	0.000619	0.866351	0.419588
				intercept	46.9886	26.1174	1.79913	0.122099

Internal Medicine

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Maximum	IMG	0.0937266	6	Year of Years	0.0014413	0.0007243	1.98992	0.0937266
				intercept	20.1108	30.5631	0.658008	0.534949
Maximum	Mexican	0.860043	6	Year of Years	0.0001307	0.0007103	0.184042	0.860043
				intercept	81.429	29.9738	2.71667	0.0347962
Minimum	IMG	0.822643	6	Year of Years	0.0001375	0.000587	0.234164	0.822643
				intercept	66.2285	24.7705	2.67369	0.0368449
Minimum	Mexican	0.591173	6	Year of Years	0.0003259	0.0005746	0.567166	0.591173
				intercept	58.0573	24.2449	2.39462	0.0536874

Medical Genetics

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Minimum	IMG	0.81012	6	Year of Years	-0.0001806	0.0007193	-0.251094	0.81012
				intercept	77.5357	30.3505	2.55468	0.0432173
Minimum	Mexican	0.107086	6	Year of Years	0.000507	0.0002677	1.8938	0.107086
				intercept	47.3884	11.2967	4.19488	0.0057182
Maximum	IMG	0.548233	6	Year of Years	0.000728	0.0011446	0.636033	0.548233
				intercept	41.4945	48.2994	0.859111	0.423269
Maximum	Mexican	0.856099	6	Year of Years	-0.000188	0.0009931	-0.189299	0.856099
				intercept	88.8749	41.9035	2.12094	0.0781815

Pediatrics

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Maximum	IMG	0.970326	6	Year of Years	3.618e-05	0.0009329	0.0387779	0.970326
				intercept	75.3351	39.3644	1.91379	0.104159
Maximum	Mexican	0.0192771	6	Year of Years	0.0018107	0.0005709	3.17171	0.0192771
				intercept	8.831	24.0899	0.366585	0.726509
Minimum	IMG	0.0711102	6	Year of Years	0.0014126	0.0006451	2.18964	0.0711102
				intercept	9.60087	27.2215	0.352695	0.736377
Minimum	Mexican	0.034747	6	Year of Years	0.0016006	0.000589	2.71774	0.034747
				intercept	1.33455	24.8515	0.053701	0.958917

Pneumology

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Minimum	IMG	0.134037	6	Year of Years	0.0169313	0.0097772	1.73171	0.134037
				intercept	-705.016	412.559	-1.70888	0.138332
Minimum	Mexican	0.228283	6	Year of Years	0.0010508	0.0007832	1.34156	0.228283
				intercept	26.7541	33.0496	0.809514	0.449123
Maximum	IMG	0.134037	6	Year of Years	0.0169313	0.0097772	1.73171	0.134037
				intercept	-705.016	412.559	-1.70888	0.138332
Maximum	Mexican	0.974085	6	Year of Years	-3.617e-05	0.0010681	-0.0338634	0.974085
				intercept	82.8872	45.0703	1.83906	0.115528

Psychiatry

Panes Row	Column	Line p-value	DF	Coefficients Term	Value	StdErr	t-value	p-value
Maximum	IMG	0.608887	5	Year of Years	-0.000619	0.0011349	-0.545446	0.608887
				intercept	99.4631	47.858	2.0783	0.0922612
Maximum	Mexican	0.238782	6	Year of Years	0.0010142	0.0007754	1.30789	0.238782
				intercept	39.436	32.7203	1.20525	0.273481
Minimum	IMG	0.0116332	5	Year of Years	0.0019107	0.0004923	3.88074	0.0116332
				intercept	-10.113	20.7629	-0.48707	0.64681
Minimum	Mexican	0.0126063	6	Year of Years	0.0012673	0.0003606	3.51424	0.0126063
				intercept	15.4505	15.2171	1.01534	0.349125

Figure 2: Mathematical model of individual trend lines for each medical speciality.

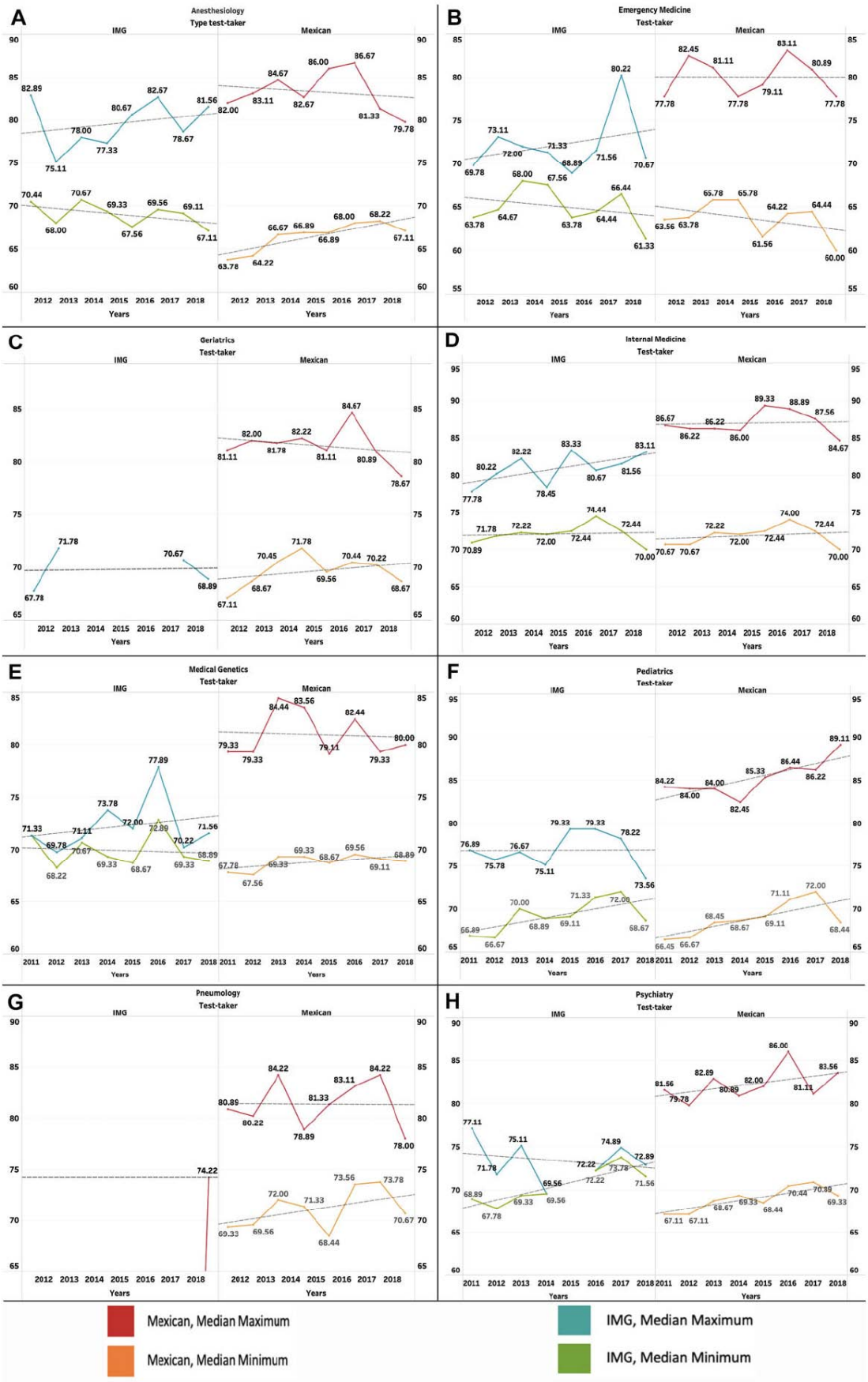

Figure 3: Selected specialities are showing increasing and decreasing trends in the SMinS and SMaxS.

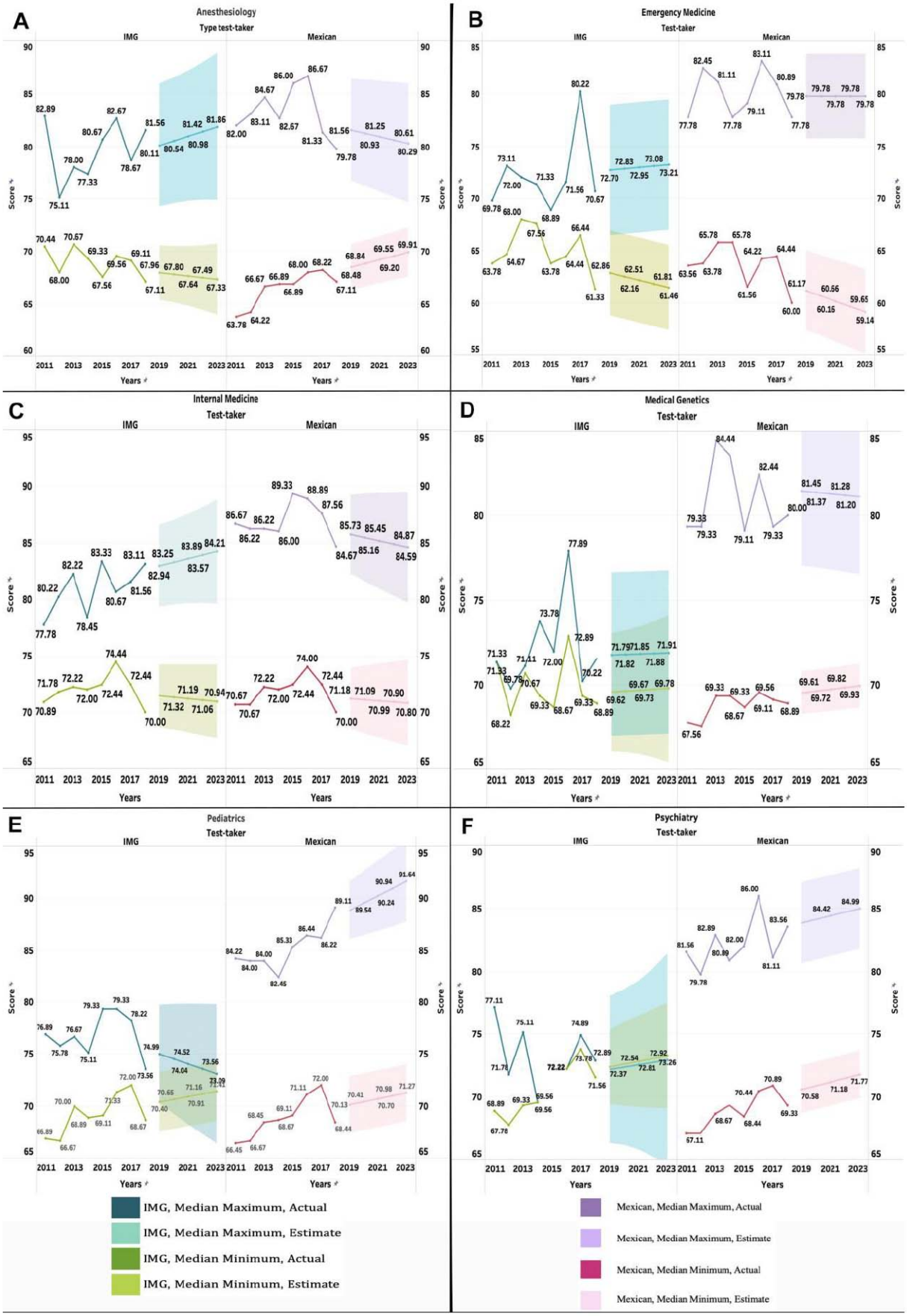

Figure 4: Forecasted trends between the minimum and maximum scores of Mexicans and IMG.

Anesthesiology

Median Maximum

Column Type test-taker	Color Type test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	3	2	0.73	3.1%	27	0.141	0.500	0.000
Mexican	Mexican	Additive	Additive	None	2	2	0.96	2.5%	25	0.298	0.000	0.000

Median Minimum

Column Type test-taker	Color Type test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1	1	0.69	1.8%	14	0.180	0.500	0.000
Mexican	Mexican	Additive	Additive	None	1	1	1.10	1.3%	12	0.100	0.354	0.000

Emergency Medicine

Sum of Maximum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	3	2	0.55	3.0%	29	0.000	0.449	0.000
Mexican	Mexican	Additive	Additive	None	2	2	0.66	2.4%	21	0.000	0.449	0.000

Sum of Minimum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	2.07	1.79	0.77	2.7%	22	0.000	0.147	0.000
Mexican	Mexican	Additive	Additive	None	1.95	1.75	0.89	2.7%	21	0.185	0.000	0.000

Internal Medicine

Sum of Maximum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1.84	1.55	0.60	1.9%	20	0.171	0.500	0.000
Mexican	Mexican	Additive	Additive	None	1.77	1.23	0.99	1.4%	19	0.500	0.000	0.000

Sum of Minimum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1.41	1.09	0.91	1.5%	15	0.322	0.000	0.000
Mexican	Mexican	Additive	Additive	None	1.37	1.12	1.01	1.6%	15	0.500	0.000	0.000

Medical Genetics

Avg. Maximum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	2.47	1.61	0.51	2.2%	24	0.000	0.046	0.000
Mexican	Mexican	Additive	Additive	None	2.26	2.15	0.86	2.7%	23	0.060	0.500	0.000

Avg. Minimum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1.79	1.30	0.58	1.8%	19	0.161	0.500	0.000
Mexican	Mexican	Additive	Additive	None	0.64	0.57	0.95	0.8%	3	0.056	0.500	0.000

Pediatrics

Avg. Maximum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	2.41	1.85	0.96	2.4%	24	0.500	0.000	0.000
Mexican	Mexican	Additive	Additive	None	1.40	1.15	0.90	1.4%	15	0.500	0.000	0.000

Avg. Minimum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1.45	1.29	0.82	1.9%	16	0.000	0.045	0.000
Mexican	Mexican	Additive	Additive	None	1.39	1.09	0.84	1.6%	15	0.000	0.043	0.000

Psychiatry

Avg. Maximum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	2.98	2.68	0.71	3.6%	25	0.228	0.500	0.000
Mexican	Mexican	Additive	Additive	None	1.64	1.28	0.46	1.6%	18	0.000	0.049	0.000

Avg. Minimum

Column Test-taker	Color Test-taker	Model Level	Model Trend	Model Season	Quality Metrics RMSE	MAE	MASE	MAPE	AIC	Smoothing Coefficients Alpha	Beta	Gamma
IMG	IMG	Additive	Additive	None	1.51	1.24	0.93	1.7%	16	0.500	0.000	0.000
Mexican	Mexican	Additive	Additive	None	0.96	0.85	0.84	1.2%	9	0.067	0.100	0.000

Figure 5: Description of the forecasting models grouped by speciality. Geriatrics and Pneumology were not included due to fewer years that did not allow for calculated reliable models.

maximum scores, depending on if the lines will or will not eventually touch each other during or after a 5-year forecasted period (2020-2024 years).

Five specialities showed a convergent pattern for Mexicans between the MinSco and MaxSco: *Anesthesiology, Internal medicine, Medical Genetics, Geriatrics, and Pneumology*, and three a divergent pattern: *Emergency medicine, Pediatrics, Psychiatry.*

In IMG, one speciality depicted a convergent trend: *Pediatrics*; five specialities had a divergent tendency: *Anesthesiology, Internal medicine, Medical Genetics, Emergency medicine, Pediatrics, Psychiatry*. For Geriatrics and Pneumology, because there were not test-takers in all the evaluated years, for that reason, the software could not calculate forecasting graphs. Figure **4** shows the forecasted trends between MinSco and MaxSco for Mexicans and IMG.

Figure **5** presents the description of the forecasted models grouped by speciality (definitions for the different components were described in the methods sections).

Ranking of Specialities between Mexicans and IMG

Additionally, we ranked the specialities based on the MinSco between Mexicans and IMG for each speciality. Adjacent rows with connecting arrows show the displacement in the ranking from the initial rank each speciality reached for Mexicans compared with their position for IMG.

For the MinSco, it was evident that the ranking of medical specialities was different between both groups: three specialities in the Mexican ranking (*Pneumology, psychiatry,* and *Medical genetics*) went up when compared them with the IMG; three moves down (*Internal medicine, Geriatrics,* and *Pediatrics*), and only two (*Anesthesiology* and *Emergency medicine)* depicted the same raking for Mexicans and IMG.

For the MaxSco, the ranking of medical specialities was different in almost all the specialities between both groups: four specialities went up in the Mexican ranking (*Anesthesiology, Pneumology, Emergency medicine,* and *Medical genetics*) after compared them with the IMG; three moves down (*Psychiatry, Geriatrics,* and *Pediatrics*), and only *Internal medicine* depicted the same raking for Mexicans and IMG. Figure **6** showed the ranking displacement in Mexican specialities (MinSco and MaxSco) when we compared them with the scores of IMG.

DISCUSSION

Residency is a critical step in a physician's education; the matching into a residency program is a competitive process of selection by both applicants and program directors [23]. Residency program directors usually do not make a decision based only on the test scores of the applicants. They must have a more comprehensive evaluation and therefore receive large amounts of information about applicants, including academic transcripts, the medical student performance assessment, letters of recommendation and others

Mexican		Minimum score			IMG	
Ranking	Specialty				Specialty	Ranking
1	Internal medicine	71.806		74.222	Pneumology	2
2	Pneumology	71.083		72.028	Internal medicine	1
3	Geriatrics	69.611		70.444	Psychiatry	4
4	Psychiatry	68.917		69.916	Medical Genetics	6
5	Pediatrics	68.861		69.778	Geriatrics	3
6	Medical Genetics	68.778		69.194	Pediatrics	5
7	Anesthesiology	66.472		68.972	Anesthesiology	7
8	Emergency medicine	63.639		65.000	Emergency medicine	8
Mexican		**Maximum score**			**IMG**	
Ranking	Specialty				Specialty	Ranking
1	Internal medicine	86.944		80.917	Internal medicine	1
2	Pediatrics	85.222		79.611	Anesthesiology	3
3	Anesthesiology	83.278		76.861	Pediatrics	2
4	Psychiatry	82.222		74.222	Pneumology	6
5	Geriatrics	81.556		73.365	Psychiatry	4
6	Pneumology	81.361		72.209	Medical Genetics	7
7	Medical Genetics	80.944		72.195	Emergency medicine	8
8	Emergency medicine	80.000		69.778	Geriatrics	5

Figure 6: Ranking displacement in Mexican specialities when compared with the scores of IMG.

[24]; a 2006 survey evinced that 2,528 program directors chose top academic selection criteria based on clinical performance [25].

Thus, the results will benefit four groups of actors interested in the processes of a successful match: ENARM applicants, education department directors, medical school advisors, and medical students who are planning to enter a residency program. The strengths of our study lie in different approaches to analyze the information. We compared the means in eight clinical specialities, the differences between Mexicans and IMG scores, calculated correlations and linear trends, 5-years forecasting, and ranking displacement for Mexicans and IMG in each speciality. Reporting information about a pattern in the assessments across specialities has been considered valuable to residents and program directors [26].

Educational Framework

The preparation for the exam should: motivate the learner through improvement in real-life, final performance; take into account the learner's pre-existing knowledge (learning curve); allow repetition of the skills multiple times; be accompanied by immediate feedback, and be varied (mixed) across content areas. We think the significantly different scores between Mexicans and IMG might primarily represent a lack of practice and direct supervision of skills acquisition (answering previous exam models). Knowing in advance, the clinical field scores are relevant to predicting the performance during the residence. As it was evinced in a recent article of 2019, the performance of USMLE Step 2 CK correlated with higher scores during residence tests with better clinical performance [24].

Publications about the Mexican ENARM have triggered a great interest in the medical community in the last years; some authors have published descriptive reports about the scores of schools and faculties of medicine [3]; other authors have revealed flaws in the design of the ENARM that produce inequity, [6, 27]; a recent study was published about the performance of IMG in the ENARM but without a comparison with Mexicans [8]. To the best of our knowledge, there are no publications about the ENARM that had presented a comparison of scores in clinical specialities between Mexicans and IMG; that situation did not allow us to compare most of our results with others literature.

Grouping of Specialities above or below a Global Mean

The use of an overall mean to compare above or below this mark is helpful to reflect the performance of eight different groups of test-takers that revealed to us which specialities had the students with the best scores. The ENARM global mean for the minimum score (from 2012 to 2019) was 69.133, a score above the previous observation made in a study by de la Garza-Aguilar [4]; this number is also above the mean for the last seven years for the test known as MIR (Medical Intern Resident) in Spain with 57.29 reported by the Ministry of Health [28, 29]. Our findings showed that the clinical specialities whose applicants achieved scores above this mean were Internal medicine, Anesthesiology, Pediatrics, and Pneumology. This observation of high scores at the ENARM contrasts with the matching program results in the USA [30, 31]. The specialities below the mean corresponded to *Emergency Medicine and Anesthesiology*.

Comparison of Minimum and Maximum Scores Achieved by Clinical Specialities

During the eight years assessed, it was evident that the eight clinical specialities' ranking was preserved for the MinSco (Figure **1D**), specialities in the upper values were internal medicine and Pneumology, and in the lower values emergency medicine and anesthesiology. On the contrary, for the MaxSco, although there is an entanglement of scores was evident along the eight years, representing the change of ranking for the clinical specialities at different years, internal medicine and emergency medicine are dominant with the upper and lower scores (Figure **1C**).

Comparison of Minimum and Maximum Scores between Mexicans and IMG in each Clinical Speciality

Our findings revealed that Mexicans and IMG got mostly similar passing grades, which might indicate an equivalent level of education in their medical schools; however, for Pneumology, anesthesiology, and emergency medicine, the IMG got up to 2% points in higher scores (Table **2**). This finding differs from a previous report from the USA observed in 8 years for the orthopaedic surgery residency applicants that national got better scores than IMG [32]. The absence of significant differences in the minimum scores in most specialities comparing Mexican and IMG can also be interpreted as high competitiveness across all specialities (Table **2**). However, MaxSco revealed the

superiority of Mexicans above IMG for all specialities, and all specialities showed a significant difference (Table **2**), which reflected a better level of preparation for this exam. This score revealed a significant gap in knowledge between Mexicans and IMG test-takers [33].

Positive and Negative Trends in the Minimum and Maximum Scores between Mexicans and IMG in each Surgical Specialities

The limited information about trends for applicants matching into USA specialities has been previously addressed. Most foreign articles describe specific specialities' performance without comparing their nationals and IMG [34]. We learned from our findings that there is still missing information, and we do not know which scores at specialities are ruled by the applicants every year and which others by the level of difficulty of the exam; an additional analysis will be necessary to understand how the number of residency positions influences the scores at each medical speciality.

Comparison of 5-Year Forecasting Trends between the Minimum and Maximum Scores of Mexicans and IMG

The predictive graphs help us understand that for Mexicans, the gap between MinSco and MaxSco will decrease for *Anesthesiology, Internal medicine,* and *Medical genetics*. However, for IMG P*ediatrics* and *medical genetics*. It means there are only 3 out of 8 surgical specialities (*Emergency medicine*, *medical genetics,* and *Psychiatry*) between Mexicans and IMG that share the same learning trend.

Ranking of Specialities between Mexicans and IMG

From this analysis, we learned that Mexicans achieved higher scores for MaxSco in the eight clinical specialities; on the contrary, IMG got higher values for their MinSco (Figure **4**). For the MaxSco, the 1st speciality with the highest scores is *Internal medicine*. This fact represents a challenge for future applicants, as they would have to get the best scores to be selected for a residency position. (Figure **4**).

Limitations of the Study

Several limitations need to be acknowledged for this study. With the ENARM, the Mexican Secretariat of Health selects the best candidates each year with reasonable confidence, but a number much higher than the accepted is left without entering a medical speciality; we did not analyze those numbers as this topic was out of the scope of this study. Also, we did not comment on the context regarding the offer and demand of Mexican physicians per number of inhabitants; in 2015, Mexico had 2.2 physicians per 1,000 population, including professionals in the private sector, these numbers represent a significant disparity in the distribution of human health resources in the country. We did not understand which medical schools corresponded the test-takers with the highest scores, as this information was not available in the annual CIFRHS reports. Our assessment did not perform subgroup performance differences considering age, gender, test-takers race, and English as a second language because all these items were not publicly available. The same limitations had been addressed in previous reports for USMLE; residency program directors look in the ENARM results for the best candidates for their programs, considering all aspects of a student's application and an interview; however, we did not take into account intangible factors such as away rotations, personal interactions, membership, and research experience, although all of them might influence the chance of matching [23], these variables were not assessed in the context of this paper. Other topics no included in this study were the need to examine whether there is an ideal applicant-to-position ratio that would allow clinical residency coordinators to remain selective in their choices or whether increasing the number of clinical residency positions would dilute the quality of successful candidates.

In conclusion, our study provides objective and valuable information for residency program directors looking for the best candidates for their programs and also to applicants, revealing that ENARM represents a market of high-performance test-takers across the clinical specialities. Mexicans and IMG achieved similar entrance scores, but Mexicans showed a higher MaxSco than IMG in all clinical specialities. The comparisons using scores will allow program directors to compare academic performance across specialities and understand their competitiveness and evolution in recent years. Future studies are needed to explore if ENARM scores can predict performance on subsequent speciality assessments in training and certification examinations.

REFERENCES

[1] Akaki Blancas JL, López Bárcena J. Formación de médicos especialistas en México. Educación Médica 2018; 19: 36-42. https://doi.org/10.1016/j.edumed.2018.03.007

[2] OCDE. Health at a Glance 2015 OECD INDICATORS 2015. 220 p.

[3] Ramiro HM, Cruz AJ, Zeron-Gutierrez L, Arevalo-Vargas A. [The ENARM and the schools and faculties of medicine. An analysis that nobody will like]. Rev Med Inst Mex Seguro Soc 2017; 55(4): 498-511.

[4] de la Garza-Aguilar J. [Reflections on the quality of medical school curricula in Mexico]. Gac Med Mex 2005; 141(2): 129-41.

[5] Barajas-Ochoa A, Ramos-Remus C. [Equity, validity and reliability of the National Examination for Candidates to Medical Residences (ENARM): opportunities for improvement]. Salud Publica Mex 2017; 59(5): 501-2. https://doi.org/10.21149/8769

[6] Barajas-Ochoa A, Ramos-Remus C, Castillo-Ortiz JD, Yanez J, Barajas-Ochoa Z, Sanchez-Gonzalez JM, *et al*. Flaws in the design of the Examen Nacional para Aspirantes a Residencias Medicas produce inequity. Salud Publica Mex 2019; 61(2): 125-35. https://doi.org/10.21149/9790

[7] Secretaria_de_Salud_Comisión_Interinstitucional_para_la_Formación_de_Recursos_Humanos_para_la_Salud. Reportes academicos. Informacion por año. Mexico City 2018 [Access date: May 16, 2019].

[8] Mendoza-Aguilar AN, Salazar-Ruiz SY, Cardiel-Marmolejo LE, Roldan-Valadez E. Comparison of scores achieved by foreign medical graduates in the National Evaluation for Medical Residency (ENARM) in Mexico. Postgrad Med J 2019. https://doi.org/10.1136/postgradmedj-2019-136893

[9] Blog_ENARM_by_Foro_ENARM.org. Mejores Escuelas y Facultades de Medicina - Ranking 2018 Mexico City: Foro_ENARM.org; 2018 [Date cited: Oct 02, 2019]. Available from: https://comunidadenarm.blogspot.com].

[10] Bordage G. Conceptual frameworks to illuminate and magnify. Med Educ 2009; 43(4): 312-9. https://doi.org/10.1111/j.1365-2923.2009.03295.x

[11] CIFRHS. Examen Nacional para Aspirantes a Residencias Médicas: INFORMACIÓN POR AÑO 2019 [Access: March 05th, 2020]. Available from: http://www.cifrhs.salud.gob.mx/site1/enarm/enarm_ia.html.

[12] Hernandez-Galvez DC, Roldan-Valadez E. Mexican ENARM: performance comparison of public vs. private medical schools, geographic and socioeconomic regions. Salud publica de Mexico 2019; 61. https://doi.org/10.21149/10078

[13] CIFRHS. Examen Nacional para Aspirantes a Residencias Médicas: Puntajes máximos y mínimos por especialidad de médicos mexicanos civiles y lugares de médicos extranjeros seleccionados 2019 [Access: March 05th, 2020]. Available from: http://www.cifrhs.salud.gob.mx/site1/enarm/docs/2019/E43_puntajes_max_min_2019.pdf.

[14] CIFRHS. The Interinstitutional Commission for Human Resources Training for Health. Mexican Ministry of Health. XXXIX Examen Nacional para Aspirantes a Residencias Médicas. Convocatoria 2015. Mexico City: Mexican government; 2015 [Access date October 04 th, 2019]. Available from: http://www.cifrhs.salud.gob.mx/site1/enarm/docs/2015/E39_convo_2015.pdf.

[15] Roldan-Valadez E, Rios-Piedra E, Favila R, Alcauter S, Rios C. Diffusion tensor imaging-derived measures of fractional anisotropy across the pyramidal tract are influenced by the cerebral hemisphere but not by gender in young healthy volunteers: a split-plot factorial analysis of variance. Chin Med J (Engl) 2012; 125(12): 2180-7.

[16] Lopez-Mejia M, Roldan-Valadez E. Comparisons of Apparent Diffusion Coefficient Values in Penumbra, Infarct, and Normal Brain Regions in Acute Ischemic Stroke: Confirmatory Data Using Bootstrap Confidence Intervals, Analysis of Variance, and Analysis of Means. J Stroke Cerebrovasc Dis 2016; 25(3): 515-22. https://doi.org/10.1016/j.jstrokecerebrovasdis.2015.10.033

[17] Pfister R, Janczyk M. Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. Advances in cognitive psychology / University of Finance and Management in Warsaw 2013; 9(2): 74-80. https://doi.org/10.5709/acp-0133-x

[18] Cohen JW. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

[19] Chatfield C. he Holt-Winters Forecasting Procedure. Applied Statistics 1978; 27(3): 264-79. https://doi.org/10.2307/2347162

[20] Adeyinka DA, Muhajarine N. Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models. BMC Med Res Methodol 2020; 20(1): 292. https://doi.org/10.1186/s12874-020-01159-9

[21] Talkhi N, Akhavan Fatemi N, Ataei Z, Jabbari Nooghabi M. Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods. Biomed Signal Process Control 2021; 66: 102494. https://doi.org/10.1016/j.bspc.2021.102494

[22] TABLEAU_SOFTWARE_LLC. Forecast Descriptions. Seattle, Washington. [cited May 12, 2021]. Available from: https://help.tableau.com/current/pro/desktop/en-us/forecast_describe.htm.

[23] Rinard JR, Mahabir RC. Successfully matching into surgical specialties: an analysis of national resident matching program data. Journal of graduate medical education 2010; 2(3): 316-21. https://doi.org/10.4300/JGME-D-09-00020.1

[24] Sharma A, Schauer DP, Kelleher M, Kinnear B, Sall D, Warm E. USMLE Step 2 CK: Best Predictor of Multimodal Performance in an Internal Medicine Residency. Journal of graduate medical education 2019; 11(4): 412-9. https://doi.org/10.4300/JGME-D-19-00099.1

[25] Green M, Jones P, Thomas JX, Jr. Selection criteria for residency: results of a national program directors survey. Academic medicine : journal of the Association of American Medical Colleges 2009; 84(3): 362-7. https://doi.org/10.1097/ACM.0b013e3181970c6b

[26] Sawhill AJ, Dillon GF, Ripkey DR, Hawkins RE, Swanson DB. The impact of postgraduate training and timing on USMLE Step 3 performance. Academic medicine : journal of the Association of American Medical Colleges 2003; 78(10 Suppl): S10-2. https://doi.org/10.1097/00001888-200310001-00004

[27] Barajas-Ochoa A, Ramos-Remus C, Ramos-Gómez S, Barajas-Ochoa Z, Sánchez-González JM, Hernández-Ávila M, *et al*. Desempeño de las escuelas de medicina en México: resultados del Examen Nacional para Aspirantes a Residencias Médicas. Salud Publica de Mexico 2019; 61(4, jul-ago): 495-503. https://doi.org/10.21149/10042

[28] Freire JM, Infante A, de Aguiar AC, Carbajo P. An analysis of the medical specialty training system in Spain. Human resources for health 2015; 13: 42. https://doi.org/10.1186/s12960-015-0038-y

[29] Baladrón J, Sánchez-Lasheras F, Villacampa T, Romeo-Ladrero JM, Jiménez-Fonseca P, Curbelo J, *et al*. El examen MIR 2015 desde el punto de vista de la teoría de respuesta al ítem: Revista de la Fundación Educación Médica 2017; 20: 29-38. https://doi.org/10.33588/fem.201.874

[30] Angus SV, Williams CM, Stewart EA, Sweet M, Kisielewski M, Willett LL. Internal Medicine Residency Program Directors' Screening Practices and Perceptions About Recruitment Challenges. Academic medicine : journal of the Association of American Medical Colleges 2019. https://doi.org/10.1097/ACM.0000000000003086

[31] Gauer JL, Jackson JB. The association of USMLE Step 1 and Step 2 CK scores with residency match specialty and location. Med Educ Online 2017; 22(1): 1358579. https://doi.org/10.1080/10872981.2017.1358579

[32] Schrock JB, Kraeutler MJ, Dayton MR, McCarty EC. A Comparison of Matched and Unmatched Orthopaedic Surgery Residency Applicants from 2006 to 2014: Data from the National Resident Matching Program. J Bone Joint Surg Am 2017; 99(1): e1. https://doi.org/10.2106/JBJS.16.00293

[33] DePasse JM, Palumbo MA, Eberson CP, Daniels AH. Academic Characteristics of Orthopaedic Surgery Residency Applicants from 2007 to 2014. J Bone Joint Surg Am 2016; 98(9): 788-95. https://doi.org/10.2106/JBJS.15.00222

[34] Manthey DE, Hartman ND, Newmyer A, Gunalda JC, Hiestand BC, Askew KL, *et al*. Trends in NRMP Data from 2007-2014 for U.S. Seniors Matching into Emergency Medicine. West J Emerg Med 2017; 18(1): 105-9. https://doi.org/10.5811/westjem.2016.10.31237

Received on 15-04-2021 Accepted on 19-05-2021 Published on 25-05-2021

https://doi.org/10.6000/1929-6029.2021.10.05