



Indicadores demográficos en la incidencia de la covid-19 en Santiago de Cuba

Larisa Zamora Matamoras*
Nelsa María Sagaró del Campo**
Luis Eugenio Valdés García***
Iliana Benítez Jiménez****

Los indicadores demográficos han sido empleados por algunos investigadores para estimar el número de personas infectadas por la covid-19. El presente trabajo tiene como primer objetivo determinar en qué medida la incidencia de casos con covid-19 en los municipios de la provincia de Santiago de Cuba puede ser explicada a partir de determinados indicadores demográficos. El segundo objetivo es construir una jerarquía de grupos de municipios de acuerdo al comportamiento diferenciado de los indicadores demográficos seleccionados. Se desarrolló un estudio ecológico, exploratorio, de grupos múltiples, comparando los nueve municipios de la provincia Santiago de Cuba según variables del nivel global, supuestamente relacionadas con la cantidad de casos con covid-19 confirmados desde el 15 de octubre de 2020 hasta el 16 de enero de 2021. Se aplicó el análisis de regresión lineal múltiple para seleccionar el modelo que describiera mejor el comportamiento de los datos y el análisis de clúster para visualizar la agrupación de los municipios. Se evidenció una correlación significativa entre la cantidad de casos con covid-19, la densidad de población y el grado de urbanización. En cambio, en el modelo de regresión solo resultó significativa la densidad poblacional cuando se consideraron los nueve municipios y el índice de masculinidad, cuando se excluyó el municipio atípico, Santiago de Cuba. El índice de masculinidad resultó ser una variable espuria condicionada por la densidad poblacional como variable confusora. El análisis de clúster reveló la formación de tres grupos de municipios, quedando Santiago de Cuba aislado del resto de los municipios.

Palabras clave: Covid-19. Regresión lineal múltiple. Clúster. Indicadores demográficos. Santiago de Cuba.

* Universidad de Oriente, Dpto. Matemática, Santiago de Cuba, Cuba (larisa@uo.edu.cu; <https://orcid.org/0000-0003-2210-0806>).

** Universidad de Ciencias Médicas, Dpto. Posgrado, Santiago de Cuba, Cuba (nelsa@infomed.sld.cu; <https://orcid.org/0000-0003-2210-0806>).

*** Universidad de Ciencias Médicas, Santiago de Cuba, Cuba (valdez@infomed.sld.cu; <https://orcid.org/0000-0003-1613-4305>).

**** Universidad de Oriente, Dpto. Sociología, Santiago de Cuba, Cuba (ilianabj@uo.edu.cu; <https://orcid.org/0000-0002-4753-3416>).

Introducción

La pandemia de covid-19 ha progresado de manera diferenciada en los países y regiones en el mundo y dentro de países según sus divisiones geográficas. Hay investigadores que han empleado indicadores demográficos para estimar el número de personas infectadas en realidad por el actual coronavirus (BOHK-EWALD *et al.*, 2021). Algunos como Patrick Garland y colaboradores (2020) proponen el estudio de la densidad poblacional como uno de los principales diferenciales de la incidencia de la enfermedad. Así, se encontraron diversos estudios que relacionan la expansión de la covid-19 con la densidad de la población de forma aislada (KADI; KHELFAOUI, 2020; DESAI, 2020), mientras otros la relacionan con variables meteorológicas (RASHED *et al.*, 2020; KODERA *et al.*, 2020; COŞKUN *et al.*, 2021) y con la movilidad espacial representada en la entrada de viajeros a un territorio, lo cual fue evidenciado por Larisa Zamora y colaboradores (2020). Esta es, de hecho, una de las variables demográficas más relacionada con el comportamiento de la epidemia por covid-19 en el mundo.

Otros investigadores relacionan la incidencia de la enfermedad con el grado de urbanización de los territorios, vinculándola también con la densidad poblacional, dado que los territorios más urbanizados suelen tener el mayor número de habitantes en espacios geográficos menores; sin embargo, afirman que en estos casos se cuenta con mejores infraestructuras de salud para enfrentar las epidemias (FANG; WAHBA, 2020).

Dos indicadores demográficos de los estudiados más tempranamente con relación al comportamiento de la covid-19 han sido el sexo y la edad de las personas infectadas por el virus, y en Cuba existe cierta inmediatez en estos datos que puede apreciarse en el sitio oficial del Ministerio de Salud Pública (MINSAP, 2021) para la divulgación del comportamiento de la epidemia. Igualmente, se han hecho estudios más detallados sobre estos diferenciales en cada una de las provincias del país (CEDEM, 2020), en los que se identifica la distribución más o menos equitativa de los casos por sexos en todas las provincias y en los que el lugar de ocurrencia de los eventos de transmisión es uno de los elementos que podría definir la mayor presencia de hombres o mujeres en cada caso. En tal sentido, se escoge el índice de masculinidad, indicador demográfico que establece la razón entre ambos sexos por territorios, para conocer su posible influencia en el número de casos confirmados con la enfermedad.

La edad, por su parte, también ha sido estudiada en relación con la incidencia de la enfermedad, pero, en mayor medida, en relación con su evolución, ya que los mayores de sesenta años son los de mayor riesgo de agravamiento y de muerte (CEDEM, 2020). En el mundo, estudios resaltan la afectación desproporcionada que ha tenido la covid-19 en la población de adultos mayores en todo el orbe debido a sus comorbilidades y al deterioro funcional de esta, pero también por estar más expuestos a la fragilidad de los sistemas de salud (OMS/OPS; 2020).

Resultaría útil conocer la posible influencia del grado de envejecimiento de la población, como indicador de nivel macro, en el número de casos confirmados de personas con covid-19, teniendo en cuenta que estudios afirman que los adultos mayores, además de presentar síntomas graves de la enfermedad, son especialmente susceptibles de contraerla (ESPIN *et al.*, 2020). Se puede valorar en este caso también la posibilidad de su no detección en personas de menor edad debido a que no presentan síntomas. De ahí el estudio del envejecimiento de la estructura por edades de la población como otro de los indicadores demográficos a analizar.

Este panorama y la posible influencia de cada uno de estos indicadores demográficos comentados en la propagación de la covid-19 motivó un estudio para verificar esa posible influencia en la distribución no homogénea de la cantidad de casos en los diferentes municipios de la provincia Santiago de Cuba. El colectivo de autores de este trabajo es parte del Grupo Asesor del Consejo de Defensa del Gobierno Provincial para atender la epidemia de la covid-19, por lo que se escoge esta provincia para hacer el estudio y proponer recomendaciones para el enfrentamiento a la epidemia.

En esta dirección, el presente trabajo tiene como primer objetivo determinar en qué medida la incidencia de casos con covid-19 en los municipios de la provincia de Santiago de Cuba puede ser explicada a partir de determinados indicadores demográficos. El segundo objetivo es construir una jerarquía de grupos de municipios de acuerdo con el comportamiento diferenciado de los indicadores demográficos seleccionados y el número de casos con covid-19.

Material y método

Se llevó a cabo un estudio ecológico exploratorio, de grupos múltiples, comparando varias regiones en un mismo momento en el que las unidades de análisis fueron nueve poblaciones geográficamente bien delimitadas que se corresponden con los nueve municipios en que se divide la provincia Santiago de Cuba. Según la actual división políticoadministrativa, estos municipios son Santiago de Cuba (SC), Palma Soriano (PS), San Luis (SL), Songo La Maya (SM), Segundo Frente (SF), Tercer Frente (TF), Guamá (G), Mella (M) y Contramaestre (C). El período considerado para el estudio se extendió desde el 15 de octubre de 2020 hasta el 16 de enero de 2021, cuando se hizo un primer corte exploratorio de la segunda ola de covid-19 en la provincia.

Los indicadores demográficos considerados fueron la densidad poblacional (DP), el índice de masculinidad (IM), el grado de urbanización (GU) y el grado de envejecimiento poblacional (GE), los cuales se han estudiado en el orden estadístico como covariables. Estas representan el comportamiento a niveles territoriales de indicadores de tamaño y estructura de la población. La definición de cada una de ellas se presenta a continuación:

- densidad poblacional: número de habitantes residentes en un territorio por km²;
- índice de masculinidad: número de hombres por cada cien mujeres;
- grado de urbanización: porcentaje de población del territorio que vive en zonas urbanas;
- grado de envejecimiento poblacional: porcentaje de personas con sesenta años y más de edad con respecto al total de la población.

Como variable dependiente o de respuesta se consideró el número de casos confirmados con covid-19 (*casos*) por municipios en el período de estudio.

Los datos demográficos fueron obtenidos directamente o calculados a partir de los reportes anuales de la Oficina Nacional de Estadística e Información (ONEI, 2021) y del Ministerio de Salud Pública (MINSAP, 2021), ambos de Cuba, referidos a la población del año 2019, edición 2020. Los datos asociados a la enfermedad se extrajeron de los partes informativos de Salud Pública al Consejo de Defensa Provincial sobre la vigilancia de la covid-19 y se encuentran disponibles en el sitio oficial de Gobierno del MINSAP (2021).

Análisis estadístico

Al ser el tamaño de la muestra muy pequeña (nueve) se verificó la normalidad de los datos para decidir qué coeficiente emplear para analizar la correlación entre las variables. Como las variables no siguen una distribución normal se trabajó con el coeficiente de correlación de Spearman.

Para determinar los indicadores demográficos que más influyeron en la incidencia de la covid-19 por municipios en Santiago de Cuba se aplicó la regresión lineal múltiple (RLM) (WOLFGANG; LÉOPOLD, 2015; STOCKEMER, 2019). En esta técnica se emplearon todos los métodos de selección del mejor modelo implementados en el Minitab, versión 19.0 (*ninguno, paso a paso, selección hacia adelante, eliminación hacia atrás, criterio de información hacia adelante*, empleando el criterio de información de Akaike – AIC y el criterio de información bayesiano – BIC), *software* empleado para el procesamiento estadístico de los datos.

Se analizó la existencia de valores atípicos. Se eliminó como valor atípico el municipio Santiago de Cuba y se obtuvo un nuevo modelo de regresión.

Mediante la aplicación del análisis de clúster (KARL *et al.*, 2015) se determinaron los municipios más similares según los indicadores estudiados.

Se mantuvo una conducta bioética acorde a lo dispuesto en la Declaración de Helsinki de la Asociación Médica Mundial. Debido a que no se emplearon datos personales no fue aplicable el consentimiento informado.

Resultados

En la Tabla 1 se muestran los valores de los cuatro indicadores estudiados y la cantidad de casos confirmados con covid-19 en la provincia de Santiago de Cuba y en cada uno de sus municipios.

TABLA 1
Valores de los indicadores demográficos y cantidad de casos confirmados con covid-19
Santiago de Cuba – 15 octubre 2020-16 enero 2021

Municipios	Casos	IM	DP	GU	GE
Total provincia	1.307	98,44	188,10	70,89	19,50
Contramaestre	40	102,84	154,40	59,25	19,70
Mella	46	106,23	104,20	71,00	17,40
San Luis	35	102,08	115,20	63,87	18,40
Segundo Frente	10	104,40	75,40	32,25	18,40
Songo la Maya	78	101,07	127,60	44,16	18,00
Santiago de Cuba	970	94,31	494,20	89,17	20,30
Palma Soriano	107	98,69	131,40	61,89	20,60
Tercer Frente	4	110,30	82,30	38,97	16,90
Guamá	17	108,08	36,00	27,81	16,40

Fuente: Oficina Nacional de Estadística e Información (ONEI, 2021) y Ministerio de Salud Pública (MINSAP, 2021).

Nota: DP: densidad poblacional, IM: índice de masculinidad, GU: grado de urbanización, GE: grado de envejecimiento.

Durante el período estudiado, que corresponde a la segunda ola de la epidemia en la provincia de Santiago de Cuba, se detectaron un total de 1307 casos positivos, para un promedio por municipios de 145 casos y una dispersión de 311.

Se obtuvo la matriz de correlación de Spearman y se analizó la significación de las asociaciones entre las variables con un nivel de confiabilidad del 95%.

La cantidad de casos con covid-19 se correlacionó de forma significativa y directa con la densidad poblacional ($r = 0,800$; $p = 0,010$) y el grado de urbanización ($r = 0,717$; $p = 0,030$). Lo anterior significa que los municipios con valores altos en estas covariables tendrán una cantidad mayor de casos confirmados con covid-19 que aquellos municipios con valores bajos de estas. Lo contrario ocurre con el índice de masculinidad, covariable correlacionada de forma negativa con la cantidad de casos con covid-19 ($r = -0,850$; $p = 0,004$).

La cantidad de casos con covid-19 se correlacionó con el grado de envejecimiento con un 90% de confiabilidad ($r = 0,636$; $p = 0,066$).

Los pares de covariables correlacionadas fueron: densidad poblacional e índice de masculinidad ($r = -0,817$; $p = 0,007$); grado de envejecimiento e índice de masculinidad ($r = -0,837$; $p = 0,005$); grado de urbanización y densidad poblacional ($r = 0,717$; $p = 0,030$); grado de envejecimiento y densidad poblacional ($r = 0,795$; $p = 0,010$).

Al aplicar los métodos de selección del mejor modelo de regresión, se obtuvieron dos modelos diferentes:

- modelo 1: Casos = 1362,623 – 7,450 IM + 2,514 DP – 1,776 GU – 39,110 GE;
- modelo 2: Casos = –185,423 + 2,253 DP.

El modelo 1 se obtuvo con el método *ninguno* según el Minitab, el cual incorpora todas las covariables seleccionadas en el modelo. En este modelo solo la densidad poblacional resultó significativa con un 95% de confiabilidad ($p = 0,002$).

Al aplicar el resto de los métodos programados siempre se obtuvo el mismo modelo, el modelo 2, en el cual solo resultó significativa una covariable, la densidad poblacional.

Se asumió el modelo 2 como el mejor modelo de regresión lineal hasta el momento.

Calidad del modelo de regresión obtenido (modelo 2)

Para analizar la calidad del modelo de regresión obtenido se empleó un conjunto de indicadores: la desviación estándar de la distancia entre los valores observados y los ajustados (S), los porcentajes de variación en la variable respuesta que es explicada por el modelo (R^2) y ajustado para el número de covariables relativo al número de observaciones (R^2_{ajus}). Aunque el objetivo de la investigación no fue obtener un modelo para hacer predicciones, se calculó el R^2 pronosticado (R^2_{pron}), el cual permite determinar qué tan bien predice el modelo la respuesta para nuevas observaciones.

El R^2 , R^2_{ajus} y R^2_{pron} toman valores entre 0 y 100%, mientras más alto sea su valor, mejor se ajustará el modelo a los datos. En el caso del indicador S , mientras más bajos sean sus valores, mejor se ajustará el modelo a los datos.

El modelo obtenido explica el 94,94% de la variabilidad contenida en la variable dependiente, pero con una baja capacidad de predicción (31,76%) que, aunque se analiza, no fue objetivo del estudio como se explicó antes.

Ecuación de regresión para el modelo 2

La Tabla 2 muestra las estimaciones de los coeficientes del modelo (coef), el error estándar de estas estimaciones (EE coef), el intervalo de confianza al 95% de confiabilidad (IC de 95%) y el valor de estadígrafo t de student (T) que permite docimar las hipótesis acerca de la significación o no de los coeficientes del modelo, junto con su p -valor.

TABLA 2
Coeficientes del modelo 2 de regresión e indicadores de bondad de ajuste

Término	Coef	EE del coef.	Valor T	Valor p	S	R^2	R^2_{ajus}	R^2_{pron}
Constante	-185,423	35,605	-5,208	0,001*	69,970	95,57%	94,94%	31,76%
DP	2,253	0,183	12,291	0,000*				

Fuente: Elaboración propia.

*significativas al 1% ($p < 0,01$).

Para los datos procesados se concluyó que existe suficiente evidencia para afirmar, con un nivel de confiabilidad superior al 99% ($p < 0,01$), que el coeficiente de regresión es diferente de cero y, por lo tanto, la covariable asociada a este coeficiente ejerce una influencia significativa sobre el número de casos confirmados con covid-19.

Análisis de los residuos del modelo 2

El cumplimiento de los supuestos de independencia y normalidad de los residuos se verificó por medio de las pruebas de Durbin-Watson y Ryan Joiner, respectivamente. La homocedasticidad fue verificada gráficamente.

La prueba de Durbin-Watson arrojó un valor del estadígrafo $DW = 1,24$. Este valor cae en una zona de incertidumbre ($d_1 = 0,82$ y $d_2 = 1,32$ para el nivel de confiabilidad del 95%), por lo que no se puede garantizar la independencia o no de los errores.

La prueba de Ryan Joiner mostró un valor de 0,95 y un p-valor mayor que 0,10; por lo que se concluye con un nivel de confiabilidad del 95% que los residuales o errores siguen una distribución normal.

Ajustes y diagnósticos para todas las observaciones en el modelo 2

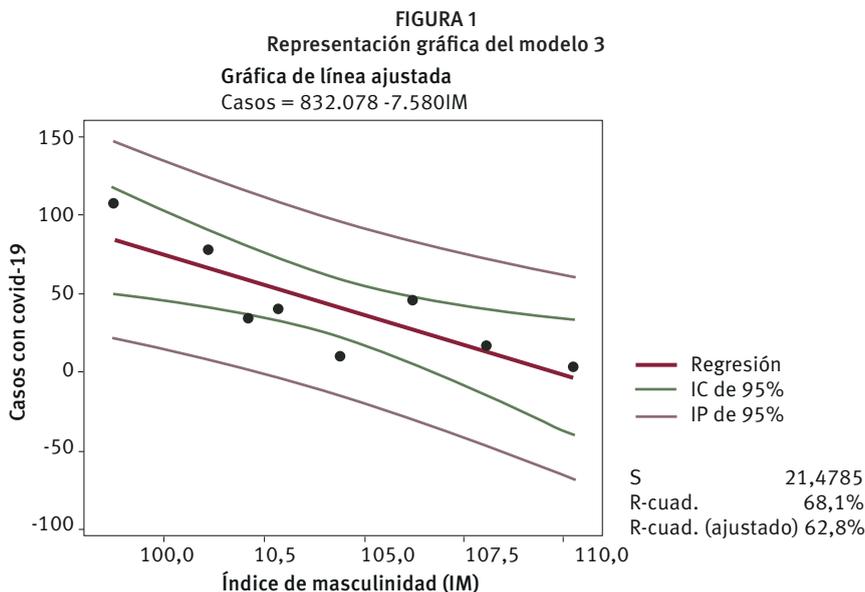
En la Tabla 1 del Anexo se muestran los estadígrafos de ajuste y diagnóstico para todas las observaciones. La observación 6, correspondiente al municipio Santiago de Cuba, muestra un valor elevado de la distancia de Cook, lo cual indicó que este municipio tiene un peso importante en la estimación. Igual comportamiento mostró el estadístico de influencia DFITS y el valor de apalancamiento (AA) indicando un cambio importante en el pronóstico de este caso cuando es eliminado de la ecuación de regresión.

La no garantía del cumplimiento del supuesto de errores independientes y la existencia de un municipio atípico (Santiago de Cuba) conllevó a hacer un nuevo análisis, en el cual fue eliminado dicho municipio.

De esta forma, la correlación entre los casos con covid-19 y el índice de masculinidad ($r = -0,825$; $p = 0,012$) y entre los casos con covid-19 y la densidad poblacional ($r = 0,714$; $p = 0,047$) resultaron significativas, al igual que en el primer análisis. Sin embargo, la densidad poblacional y el nivel de urbanización, que en el primer análisis mostraron estar correlacionadas con la cantidad de casos con covid-19, en este nuevo análisis no lo están.

El comportamiento de las correlaciones entre las covariables se mantuvo como en el primer análisis. Lo mismo ocurrió al aplicar la regresión lineal e introducir todas las covariables al modelo, ninguna de ellas resultó significativa.

Con los otros métodos de selección del mejor modelo de regresión se obtuvo el modelo 3, que tiene como ecuación: $\text{casos} = 832,078 - 7,580 \text{ IM}$, y cuya representación gráfica se muestra en la Figura 1.



Fuente: Elaboración propia.

Calidad del modelo 3

En la Figura 1 se muestran algunos de los indicadores de bondad de ajuste considerados en este estudio. Este modelo explica el 62,81% de la variabilidad contenida en la variable dependiente, el cual es inferior al del modelo 2. Sin embargo, el valor del estadígrafo S es menor, lo cual es un indicador de un mejor ajuste de este modelo a los datos.

Ecuación de regresión del modelo 3

La Tabla 3 muestra las estimaciones de los coeficientes del modelo (*coef*), el error estándar de estas estimaciones (EE del coef), el intervalo confidencial al 95% (IC de 95%) y el valor de estadígrafo *t* de student (valor *T*) con su *p*-valor (valor *p*).

TABLA 3
Coefficientes del modelo 3 de regresión indicadores de bondad de ajuste

Término	Coef	EE del coef.	Valor <i>T</i>	Valor <i>p</i>	S	<i>R</i> ²	<i>R</i> ² _{ajus}	<i>R</i> ² _{pron}
Constante	832,078	220,744	3,769	0,009*	21,478	68,12%	62,81%	44,06%
IM	-7,580	2,117	-3,581	0,012**				

Fuente: Elaboración propia.

* significativas al 1% (*p* < 0,01); ** significativas al 5% (*p* < 0,05).

Análisis de los residuos del modelo 3

La prueba de Durbin-Watson arrojó un valor del estadígrafo $DW = 1,83$. Este valor cae en la zona de ausencia de correlación ($d_1 = 1,320$ y $4 - d_2 = 2,68$ para el nivel de confiabilidad del 95%), por lo que se puede garantizar, con este nivel de confiabilidad, la independencia de los errores.

La prueba de Ryan Joiner mostró un valor de 0,975 y un p-valor mayor que 0,10, por lo que se concluye con un nivel de confiabilidad del 95% que los residuales o errores siguen una distribución normal.

Ajustes y diagnósticos para todas las observaciones en el modelo 3

La Tabla 2 del Anexo muestra los ajustes y diagnósticos para todas las observaciones para el modelo 3. De esta tabla se concluyó que no existen valores atípicos.

Debido a que estas autoras no encuentran una explicación lógica a la relación obtenida en el modelo 3, teniendo en cuenta que la presencia de las mujeres y los hombres en la cantidad de casos confirmados con covid-19 en los municipios considerados en este modelo ha sido homogénea (veáse la Tabla 3 en el Anexo), se pasó a analizar la posible existencia de una relación espuria entre las variables índice de masculinidad y cantidad de casos confirmados.

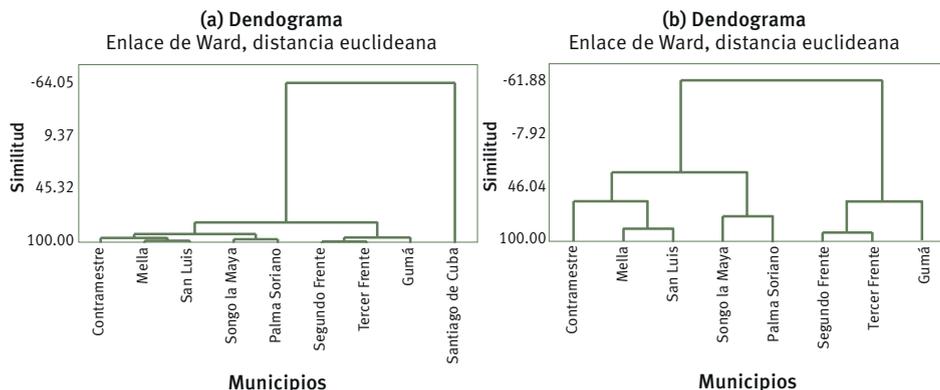
Se calcularon los coeficientes de correlación de orden cero entre el índice de masculinidad y la cantidad de casos ($r = -0,83$; p-valor = 0,012) y el coeficiente de correlación parcial entre ellas, considerando como variable de control a la densidad poblacional ($r = -0,69$; p-valor = 0,087). Como se observa, el valor del coeficiente de correlación ha cambiado y su significación también, lo cual es indicativo de que existe una relación espuria entre el índice de masculinidad y la cantidad de casos, determinada por la densidad poblacional, la cual actúa como variable confusora.

Análisis de clúster

El análisis de clúster permitió visualizar la formación de tres grupos de municipios (Figura 2a). El primer grupo concentra los municipios Contramaestre, Mella, San Luis, Songo La Maya y Palma Soriano. El segundo a los municipios Segundo Frente, Tercer Frente y Guamá. El municipio cabecera, Santiago de Cuba, queda aislado del resto y forma el tercer grupo.

Cuando se eliminó del estudio al municipio Santiago de Cuba se evidenció la formación de los dos primeros grupos antes descritos (Figura 2b) y la subdivisión de los municipios coincidió con niveles de urbanización mayores y menores del 40%, respectivamente.

FIGURA 2
Agrupaciones de las provincias según indicadores sociodemográficos
Con el municipio Santiago de Cuba (a) y sin el municipio Santiago de Cuba (b)



Fuente: Elaboración propia.

Discusión

Para el período analizado, la densidad poblacional fue el indicador demográfico que más se relacionó con la transmisión de la covid-19 en la provincia Santiago de Cuba, lo cual coincide con la mayoría de los autores consultados.

La hipótesis de que los territorios con alta densidad poblacional son más vulnerables a las epidemias debido a la mayor posibilidad de contactos interpersonales ha incentivado múltiples investigaciones. Sin embargo, algunos autores están a favor del efecto beneficioso que ejerce la densidad poblacional en la lucha contra las epidemias al reconocer que, dado un cierto umbral de densidad de población, las ciudades brindan instalaciones y servicios de mayor calidad a sus residentes (VERBRUGGE; TAYLOR, 1980).

Davhal Desai (2020) afirma que la densidad de población de las megaciudades como Nueva York, Londres y Mumbai crea un ambiente ideal para el desarrollo y la rápida expansión de infecciones como la covid-19, pero también facilita el uso eficaz de recursos escasos y servicios que pueden distribuirse a un número grande de las personas en períodos más cortos de tiempo. Lois Verbrugge y Ralph Taylor (1980) consideraron, por el contrario, la relación de la densidad poblacional con actitudes sociales negativas y comportamientos indeseables. Esto podría relacionarse en el contexto actual de la epidemia por covid-19 con respuestas no deseadas de incumplimiento de medidas de restricciones para evitar la propagación de la enfermedad.

El estudio de Nadjat Kadi y Mounia Khelfaoui (2020), muy similar al nuestro en su diseño, mostró la relación entre la densidad de población y la propagación de la covid-19 en ciudades argelinas. Los modelos de regresión obtenidos sugieren que existe una fuerte correlación entre la densidad de población y el número de infecciones por covid-19, y el

análisis de conglomerados les permitió agrupar las ciudades con mayor número de casos infectados por covid-19 y mayor densidad de población.

Felipe Carozzi y colaboradores (2020) estimaron el vínculo entre la densidad de población y la propagación y gravedad de la covid-19 en condados contiguos. Encontraron que la densidad ha influido en el momento de aparición del brote y que los condados más densamente poblados son los que tienen más probabilidades de tener un brote temprano.

Lo anterior confirma que la densidad sí influye en la transmisión de la epidemia, pero que puede interactuar igualmente con otros indicadores. Respecto a esto se encuentran estudios que afirman que la urbanización acerca a las personas tanto espacial como económicamente y que la conjunción de mayor densidad, menor distancia y mejor integración del mercado permite un desarrollo económico que garantiza instituciones mejor diseñadas, infraestructuras de alta calidad e intervenciones eficaces para llevar a cabo el distanciamiento social que hace a las ciudades más fuertes contra las enfermedades infecciosas (FANG; WAHBA, 2020).

Por otra parte, Andrés Olivera (2020) plantea que las ciudades han sido el principal escenario de vulnerabilidades y riesgos ante la pandemia de la covid-19, indistintamente del grado de desarrollo económico y social de los países afectados y del entorno social y cultural en que se desenvuelven. Para este autor, en el entorno urbano cubano, la elevada densidad habitacional, la creación de asentamientos humanos periféricos y la centralización de servicios genera mayor movilidad de los ciudadanos para acceder a productos e insumos de primera necesidad, lo cual favorece la transmisibilidad de la covid-19.

En el presente estudio el nivel de urbanización no resultó ser una variable significativa en el modelo de regresión. Esto puede estar relacionado con las particularidades que tiene Cuba en la clasificación de las zonas en rurales o urbanas, que la diferencian del resto de los países, ya que la construcción de viales, pavimentación de calles, electrificación, entre otras, en las zonas poco pobladas ha llevado a cambiar su clasificación de rural a urbano (LEÓN; HARDOY, 1972).

Sin embargo, al hacer el análisis de clúster se evidencia la presencia del nivel de urbanización en la formación de los conglomerados. Se obtienen dos grupos de municipios y un municipio aislado, Santiago de Cuba, cuya principal diferencia es precisamente el grado de urbanización. El segundo de esos grupos contiene a los municipios de menor grado de urbanización combinados con bajas densidades de su población.

Al excluir del estudio el municipio Santiago de Cuba, por resultar una observación atípica, el índice de masculinidad fue el único indicador que se relacionó significativamente con la transmisión de la epidemia, de forma negativa, aunque se demostró que dicha relación es espuria. La variable densidad poblacional nuevamente muestra su alta incidencia al actuar como una variable confusora en dicha relación.

Estudios sobre la epidemia en Cuba han relacionado la variable sexo con la incidencia de la covid-19 y muestran una ligera mayoría del sexo masculino entre los casos positivos con la enfermedad, aunque se afirma que la mayor presencia de uno u otro sexo dependerá

de las características de los lugares donde se produzcan los eventos de transmisión de la covid-19 (CEDEM, 2020). Otras investigaciones muestran el análisis por sexo a través del índice de masculinidad. Tal es el resultado de un estudio realizado en la provincia La Habana, Cuba, en el cual se obtiene que aunque a nivel territorial el índice de masculinidad de los infectados es de 101 hombres por cada 100 mujeres, existe mayor participación femenina en la transmisibilidad por coresidencia y redes familiares en una proporción de 88 hombres por cada 100 mujeres (MOLINA *et al.*, 2020). En ninguno de los casos se ha demostrado la existencia de vínculo entre el índice de masculinidad de los territorios y la incidencia de la covid-19, lo cual explica su comportamiento en el presente estudio como una variable espuria.

Aunque no constituye un objetivo del presente trabajo, se analizan aquellas correlaciones que aparecen entre los indicadores demográficos con un 95% de nivel confiabilidad. Se debe tener en cuenta que dichas correlaciones se corresponden con las particularidades de la dinámica de la población de la provincia Santiago de Cuba que podrían no tener igual comportamiento en otros territorios. Esta dinámica se caracteriza por una sostenida baja fecundidad, baja mortalidad y saldos migratorios negativos, fundamentalmente de su población joven, que han provocado un alto grado de envejecimiento (BENÍTEZ *et al.*, 2020), por encima del 16% (nivel III) en todos sus municipios, como lo muestra la Tabla 1.

La correlación más significativa se da, de forma negativa, entre el IM y el GE ($r = -0,84$; $p\text{-valor} = 0,005$), lo cual se explica por la mayor presencia femenina en los grupos de edades de los mayores de sesenta años debido a la sobremortalidad masculina, que hace alcanzar a la mujer mayor esperanza de vida que los hombres (75,89 en hombres y 80,33 en mujeres) (ONEI, 2020) y el alto grado de envejecimiento existente en estas poblaciones.

Existe también correlación negativa entre el IM y la DP ($r = -0,82$; $p\text{-valor} = 0,007$), lo cual podría explicarse por la correlación positiva existente entre el GE y la DP ($r = 0,79$; $p\text{-valor} = 0,010$) y negativa entre el IM y el GE, como se vio anteriormente. Es decir, que en la provincia Santiago de Cuba, los municipios de mayor nivel de envejecimiento son los de mayor densidad poblacional y a su vez, los de menor IM.

Por último, también existió correlación positiva entre el GU y la DP ($r = 0,72$; $p\text{-valor} = 0,030$), lo cual se explica a partir del comportamiento típico en el resto de las poblaciones, siendo las zonas urbanas las más densamente pobladas por ser las ciudades los espacios de mayor atracción de población.

Conclusiones

Se puede concluir que la densidad poblacional es el indicador demográfico que influye en el número de casos confirmados con covid-19 en la provincia Santiago de Cuba, y que puede interactuar con otros indicadores demográficos como el grado de urbanización, debido a la movilidad e interacción que se genera dentro de los territorios, un elemento que resulta consustancial al modo de vida urbano.

Las medidas a adoptar para el control de la epidemia deberán atender de forma rápida la aparición de casos en los territorios con mayor densidad poblacional y con mayor grado de urbanización, como las variables que de forma combinada podrían incidir más en la propagación de la epidemia por covid-19 en la provincia. Lograr una menor movilidad dentro de las ciudades más densamente pobladas, como lo es el municipio cabecera, Santiago de Cuba, además de mantener los protocolos de detección y control de casos positivos de covid-19 sigue siendo una de las formas para prevenir la propagación de esta epidemia en la provincia.

Se recomienda hacer estudios similares en otras provincias del país y contrastar los resultados obtenidos, para aportar así a la adopción de medidas en el enfrentamiento a la epidemia de covid-19 en el país.

Referencias

BENÍTEZ, I.; NARANJO, W.; GARZÓN G. Dinámica demográfica de la provincia Santiago de Cuba: un reto para el desarrollo territorial. **Novedades en Población**, v. 16, n. 31, p. 1-17, 2020. Disponible en: <http://www.novpob.uh.cu>. Acceso en: 27 mayo 2021.

BOHK-EWALD, C.; DUDEL C.; MYRSKYLÄ, M. A demographic scaling model for estimating the total number of covid-19 infections. **medRxiv**, Preprint, 2020. Disponible en: <https://doi.org/10.1101/2020.04.23>. Acceso en: 19 feb. 2021.

CAROZZI, F.; SANDRO P.; SEFI, R. **Urban density and Covid-19**. Bonn, Germany: IZA – Institute of Labor Economics, July 2020. (Discussion Paper, n. 13440). Disponible en: <https://ssrn.com/abstract=3643204>. Acceso en: 19 feb. 2021.

CEDEM – Centro de Estudios Demográficos. **Demografía y COVID-19: diferenciales sociales de una pandemia**. La Habana: CEDEM, 2020. Disponible en: <https://cuba.unfpa.org/es/publications/demograf%C3%ADa-y-covid-19-diferenciales-sociales-y-epidemiol%C3%B3gicos-de-una-pandemia>. Acceso en: 10 feb. 2021.

COŞKUN, H.; YILDIRIM, N.; GÜNDÜZ, S. The spread of COVID-19 virus through population density and wind in Turkey cities. **Science of the Total Environment**, v. 751, 141663, 2021. Disponible en: <https://doi.org/10.1016/j.scitotenv.2020.141663>. Acceso en: 19 feb. 2021.

DESAI, D. **Urban densities and the Covid-19 pandemic: upending the sustainability myth of global megacities**. Observer Research Foundation, 2020 (ORF Occasional Paper, n. 244). Disponible en: https://www.orfonline.org/wp-content/uploads/2020/05/ORF_OccasionalPaper_244_PandemicUrbanDensities.pdf. Acceso en: 10 feb. 2021.

ESPIN, J.; CARDONA, A.; MIRET, L.; GONZÁLEZ, M. La COVID-19 y su impacto en la salud del adulto mayor. **Archivos del Hospital Universitario "General Calixto García"**, v. 8, n. 3, 2020. Disponible en: <http://www.revcalixto.sld.cu/index.php/ahcg/article/view/568/546>. Acceso en: 13 may 2021.

FANG, W.; WAHBA, S. Urban density is not an enemy in the coronavirus fight: evidence from China. **World Bank Blog**, 2020. Disponible en: <https://blogs.worldbank.org/sustainablecities/urban-density-not-enemy-coronavirus-fight-evidence-china>. Acceso en: 15 feb. 2021.

GARLAND, P.; BABBITT, D.; BONDARENKO, M.; SORICHETTA, A.; TATEM, A. J.; JOHNSON, O. The COVID-19 pandemic as experienced by the individual. arXiv: 01167 [physics.soc-ph], 2020. Disponible en: <https://arxiv.org/abs/2005.01167>. Acceso en: 15 feb. 2021.

KADI, N.; KHELFAOUI, M. Population density, a factor in the spread of COVID-19 in Algeria: statistic study. **Bull Natl Res Cent**, v. 44, n. 1, p. 1-7, 2020. Disponible en: <https://doi.org/10.1186/s42269-020-00393-x>. Acceso en: 19 feb. 2021.

KARL, H.; RARDLE, W.; SIMAR, L. **Applied multivariate statistical analysis**. Springer-Verlag, 2015.

KODERA, S.; RASHED, E.A.; HIRATA, A. Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity. **Int J Environ Res Public Health**, v. 17, n. 15, 2020. Disponible en: <https://doi.org/10.3390/ijerph1715>. Acceso en: 19 feb. 2021.

LEÓN, A. M.; HARDOY, J. E. La urbanización en Cuba. **Demografía y Economía**, v. 6, n. 1, p. 41-67, 1972. Disponible en: <www.jstor.org/stable/40601951>. Acceso en: 15 feb. 2021.

MOLINA, M.; VÁZQUEZ, M.; DUJARRIC, G. La co-residencia familiar y la COVID-19 en La Habana. **Novedades en Población**, v. 16, n. 32, p. 33-54, 2020. Disponible en <http://www.novpob.uh.cu>. Acceso en: 12 mayo 2021.

MINSAP – Ministerio de Salud Pública. Dirección de Registros Médicos y Estadísticas de Salud. **Anuario de Salud 2019**. La Habana, 2020. Disponible en <https://salud.msp.gob.cu/?tag=anuario-estadistico-de-salud>. Acceso en: 19 feb. 2021.

MINSAP – Ministerio de Salud Pública. **Coronavirus en Cuba**. Información Oficial del MINSAP, 16 de febrero de 2021/11:59 pm. Disponible en: <https://salud.msp.gob.cu/>. Acceso en: 17 feb. 2021.

OLIVERA R. A. Globalización, urbanización y salud: impactos de la COVID-19. **Arquitectura y Urbanismo**, v. 41, n. 3, p. 06-16, 2020. Disponible en: <https://www.redalyc.org/jatsRepo/3768/376865021002/html/index.html>. Acceso en: 15 feb. 2021.

ONEI – Oficina Nacional de Estadística e Información. **Anuario Demográfico de Cuba 2019**. La Habana, 2020. Disponible en: <http://www.onei.gob.cu/publicaciones-tipo/Anuario>. Acceso en: 19 feb. 2021.

OMS – Organización Mundial de la Salud/OPS Organización Panamericana de la Salud. **La COVID-19 y adultos mayores, 2020**. Disponible en: <https://www.paho.org/es/temas/envejecimiento-saludable/covid-19-adultos-mayores>. Acceso en: 13 mayo 2021.

RASHED, E. A.; KODERA, S.; GOMEZ-TAMES, J.; HIRATA, A. Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: multi-prefecture study in Japan. **Int J Environ Res Public Health**, v. 17, n. 15, 2020. Disponible en: <https://doi.org/10.3390/ijerph17155354>. Acceso en: 19 feb. 2021.

STOCKEMER, D. Multivariate regression analysis. *In*: STOCKEMER, D. **Quantitative methods for the social sciences**. Cham: Springer, 2019. p. 163-174.

VERBRUGGE, L. M.; TAYLOR, R. B. Consequences of population density and size. **Urban Affairs Quarterly**, v. 16, n. 2, p. 135-160, 1980. Disponible en: <http://hdl.handle.net/2027.42/68954>. Acceso en: 15 feb. 2021.

WOLFGANG, K.; LÉOPOLD, S. **Applied multivariate statistical analysis**. 4. ed. Springer-Verlag Berlin Heidelberg, 2015.

ZAMORA, M. L.; SAGARÓ DEL CAMPO, N. M.; VALDÉS, G. L. E.; BENÍTEZ, J. I. Entrada de viajeros y densidad poblacional en la propagación de la COVID-19 en Cuba. **Rev. Cubana Med**, v. 59, n. 3, 2020. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0034-75232020000300003&lng=en. Acceso en: 19 feb. 2021.

Sobre los autores

Larisa Zamora Matamoros es doctora en Ciencias Matemática. Profesora titular en la Universidad de Oriente.

Nelsa María Sagaró del Campo es doctora en Ciencias Médicas. Profesora auxiliar en la Universidad de Ciencias Médicas, Santiago de Cuba.

Luis Eugenio Valdés García es doctor en Ciencias Médicas. Investigador auxiliar adjunto del Instituto de Vacunas Carlos J. Finlay. Cuba.

Iliana Benítez Jiménez es doctora en Ciencias Sociológicas. Profesora titular en la Universidad de Oriente.

Dirección para correspondencia

Larisa Zamora Matamoros
Universidad de Oriente
Avenida Patricio Lumumba s/n
90500 – Santiago de Cuba, Cuba

Nelsa María Sagaró del Campo
Universidad de Ciencias Médicas
Calle E s/n y Avenida de Las Américas
90100 – Santiago de Cuba, Cuba

Luis Eugenio Valdés García
Universidad de Ciencias Médicas
Calle E s/n y Avenida de Las Américas
90100 – Santiago de Cuba, Cuba

Iliana Benítez Jiménez
Universidad de Oriente
Avenida Patricio Lumumba s/n
90500 – Santiago de Cuba, Cuba

Resumo

Indicadores demográficos na incidência da Covid-19 em Santiago de Cuba

Indicadores demográficos têm sido usados por alguns pesquisadores para estimar o número de pessoas infectadas pela Covid-19. O primeiro objetivo deste estudo é determinar até que ponto a incidência de casos de Covid-19 nos municípios da província de Santiago de Cuba pode ser explicada por certos indicadores demográficos. O segundo objetivo é construir uma hierarquia de grupos de municípios de acordo com o comportamento diferenciado dos indicadores demográficos selecionados. Foi desenvolvido um estudo ecológico, exploratório e multigrupo, comparando os nove municípios da província de Santiago de Cuba de acordo com variáveis de nível global, supostamente relacionadas ao número de casos de Covid-19 confirmados entre 15 de outubro de 2020 e 16 de janeiro de 2021. A análise de regressão linear múltipla foi aplicada para selecionar o modelo que melhor descrevia o comportamento dos dados e a análise de agrupamento para visualizar o agrupamento dos municípios. Foi encontrada uma correlação significativa entre o número de casos de Covid-19, a densidade populacional e o nível de urbanização. Por outro lado, no modelo de regressão, apenas a densidade populacional

era significativa quando os nove municípios foram considerados e o índice de masculinidade, quando o município atípico, Santiago de Cuba, foi excluído. O índice de masculinidade revelou-se uma variável espúria condicionada pela densidade populacional como uma variável confusa. A análise de agrupamento revelou a formação de três grupos de municípios, com Santiago de Cuba sendo isolado do resto dos municípios.

Palavras-chave: Covid-19. Regressão linear múltipla. *Cluster*. Indicadores demográficos. Santiago de Cuba.

Abstract

Demographic indicators in the incidence of Covid-19 in Santiago de Cuba

Some researchers have used demographic indicators to estimate the number of people infected by COVID-19. The first goal of this study is to determine to what extent the incidence of cases of COVID-19 in the municipalities of the province of Santiago de Cuba can be explained by certain demographic indicators. The second goal is to construct a hierarchy of groups of municipalities according to the differentiated behavior of the selected demographic indicators. An ecological, exploratory, multi-group study was developed, comparing the nine municipalities of Santiago de Cuba province according to global level variables, supposedly related to the number of cases with COVID-19 confirmed from October 15, 2020 to January 16, 2021. Multiple linear regression analysis was applied to select the model that best described the behavior of the data and cluster analysis to visualize the grouping of the municipalities. A significant correlation was found between the number of cases with COVID-19, population density and urbanization level. On the other hand, in the regression model, only population density was significant when the nine municipalities were considered and the masculinity index, when the atypical municipality, Santiago de Cuba, was excluded. The masculinity index turned out to be a spurious variable conditioned by population density as a confounding variable. The cluster analysis revealed the formation of three groups of municipalities, with Santiago de Cuba being isolated from the rest of the municipalities.

Keywords: COVID-19. Multiple linear regression. Cluster. Demographic indicators. Santiago de Cuba.

Recebido para publicação em 07/04/2021

Aceito para publicação em 31/05/2021

Anexo

TABLA 1
Ajustes y diagnósticos para todas las observaciones bajo el modelo 2

Mpio.	Casos	Ajuste	EE de ajuste	IC de 95%	Resid.	Resid est.	Resid elim.	AA	D de Cook	DFITS
C	40	162,472	23,367	(107,218; 217,726)	-122,472	-1,857	-2,413	0,112	0,216	-0,855
M	46	49,361	24,594	(-8,795; 107,517)	-3,361	-0,051	-0,048	0,124	0,006	-0,018
SL	35	74,146	24,031	(17,322; 130,970)	-39,146	-0,596	-0,566	0,118	0,024	-0,207
SF	10	-15,531	26,741	(-78,764; 47,702)	25,531	0,395	0,370	0,146	0,013	0,153
SM	78	102,086	23,587	(46,311; 157,861)	-24,086	-0,366	-0,342	0,114	0,009	-0,122
SC	970	928,110	67,832	(767,713; 1 088,506)	41,890	2,438	5,812	0,940	46,324	22,947
PS	107	110,648	23,494	(55,094; 166,202)	-3,648	-0,055	-0,051	0,113	0,000	-0,018
TF	4	0,016	26,146	(-61,809; 61,841)	3,984	0,061	0,057	0,140	0,000	0,023
G	17	-104,307	30,922	(-177,427; -31,187)	121,307	1,933	2,620	0,195	0,453	1,291

Fuente: Elaboración propia.

Mpio.: municipio, EE: error estándar, IC: intervalo de confianza, Resid.: residual, Resd est.: residual estandarizado, AA: valor de apalancamiento, D: distancia, DFITS: diferencia en los valores ajustados.

TABLA 2
Ajustes y diagnósticos para todas las observaciones en el modelo 3

Mpio.	Casos	Ajuste	EE de ajuste	IC de 95%	Resid.	Resid est.	Resid elim.	AA	D de Cook	DFITS
C	40	52,505	8,128	(32,616; 72,395)	-12,505	-0,629	-0,594	0,143	0,033	-0,243
M	46	26,802	8,717	(5,473; 48,131)	19,198	0,978	0,974	0,165	0,094	0,432
SL	35	58,278	8,833	(36,665; 79,891)	-23,278	-1,189	-1,241	0,169	0,144	-0,560
SF	10	40,727	7,604	(22,121; 59,333)	-30,727	-1,530	-1,788	0,125	0,168	-0,677
SM	78	65,899	10,087	(41,217; 90,582)	12,101	0,638	0,603	0,221	0,058	0,321
PS	107	84,011	13,946	(49,885; 118,136)	22,989	1,407	1,570	0,422	0,722	1,340
TF	4	-4,030	14,960	(-40,637; 32,577)	8,030	0,521	0,487	0,485	0,128	0,473
G	17	12,807	11,167	(-14,518; 40,132)	4,193	0,229	0,210	0,270	0,010	0,128

Fuente: Elaboración propia.

Mpio.: municipio, EE: error estándar, IC: intervalo de confianza, Resid.: residual, Resd est.: residual estandarizado, AA: valor de apalancamiento, D: distancia, DFITS: diferencia en los valores ajustados.

TABLA 3
Cantidad de casos confirmados con COVID-19, por sexo y por municipio, excluyendo Santiago de Cuba

Municipio	Hombres	Mujeres
Contramaestre	23	19
Mella	24	22
San Luis	18	22
Segundo Frente	6	6
Songo la Maya	43	35
Palma Soriano	57	49
Tercer Frente	0	3
Guamá	10	7

Fuente: Elaboración propia.