

## Investigación original

# Recursos para la evaluación de la política pública en materia de cáncer de pulmón en México

José Luis Manzanares Rivera<sup>1</sup>

## Forma de citar

Manzanares Rivera JL. Recursos para la evaluación de la política pública en materia de cáncer de pulmón en México. Rev Panam Salud Publica. 2020;44:e172. <https://doi.org/10.26633/RPSP.2020.172>

## RESUMEN

**Objetivos.** Estimar las tasas de mortalidad ajustadas por estructura de edad por cáncer de pulmón para los 32 estados del país en el periodo entre 1998 y 2017; contrastar la distribución territorial de la demanda por servicios oncológicos respecto a la disponibilidad de especialistas para atención y determinar la capacidad predictiva de 3 diferentes algoritmos de clasificación supervisada en el contexto de técnicas de aprendizaje automatizado.

**Métodos.** Se realizó un análisis exploratorio y modelación de datos considerando los registros de mortalidad consignados por el sistema nacional de información en salud.

**Resultados.** La tasa de mortalidad por cáncer de pulmón en México observó una reducción de 14,5% entre el periodo previo a la implementación de la Ley General para el Control del Tabaco y el periodo posterior. Una reducción para la población masculina equivalente a 22% en todo el periodo. Se obtiene evidencia de un desequilibrio entre demanda de servicios oncológicos y disponibilidad de especialistas. La fase de modelación permitió demostrar la utilidad de los registros electrónicos de mortalidad disponibles en el país.

**Conclusiones.** A pesar de las reducciones en las últimas dos décadas en los patrones de mortalidad por cáncer de pulmón en México, el análisis, muestra que persisten áreas de oportunidad, principalmente entre la población femenina de los estados de Guerrero, Oaxaca y Puebla. Determinar si estos patrones se asocian con hábitos del tabaquismo o bien a otros determinantes sociales, constituye la principal recomendación para la dirección de los esfuerzos de atención de esta enfermedad oncológica en México derivada de esta investigación.

## Palabras clave

Neoplasias pulmonares; mortalidad; México

Durante el año 2017 en México, 83 687 decesos fueron atribuidos a padecimientos oncológicos, clasificados en 337 causas distintas (1). Dada la complejidad de los factores de riesgo involucrados, este grupo de enfermedades se ha convertido en una amenaza de salud pública, altamente letal que cobra la vida de 10 personas por hora en nuestro país, con una tasa de mortalidad ajustada por estructura de edad de 60 mujeres y 63 hombres por cada 100 000 habitantes (2).

Destaca entre este conjunto de padecimientos, la causa: cáncer de pulmón que registró 6 814 defunciones en este periodo (1). Esta causa de muerte representa, por su magnitud, el principal padecimiento oncológico entre la población masculina en México.

Ante el reconocimiento del impacto adverso de este fenómeno de salud, desde el plano regulatorio en países de América Latina como México, se han implementado iniciativas orientadas a proteger a la población de los efectos nocivos del tabaco y asegurar el derecho de los no fumadores a vivir y convivir en espacios 100% libres de humo de tabaco, como la Ley General para el Control del Tabaco (LGCT), puesta en operación en el país desde 2008.

No obstante, desde la perspectiva del gasto público en salud, las cifras indican que México asigna el monto más bajo dentro de las principales economías de América Latina, con el 2,5% del PIB por debajo de Brasil (3,9%), o Chile que asigna el 4,8%; hasta 2017 México no contaba con registros nacionales de cáncer

<sup>1</sup> El Colegio de la Frontera Norte, Nogales, México. ✉ José Luis Manzanares Rivera, [jmanzanares@colef.mx](mailto:jmanzanares@colef.mx)

de base poblacional (RNCBP) acreditados ante la Organización Mundial de la Salud (OMS) y la Agencia Internacional para la investigación del Cáncer (IARC por sus siglas en inglés).

Por lo anterior, contar con evidencias que contribuyan a la evaluación de instrumentos regulatorios y al diseño de estrategias para hacer un uso eficiente de los recursos disponibles, es una tarea continua relevante, así mismo la aplicación de métodos analíticos que consideran grandes volúmenes de datos es un abordaje que se reconoce como cada vez más útil (3–7).

En este contexto, el presente trabajo tiene como objetivos: 1) estimar las tasas de mortalidad ajustadas por estructura de edad por cáncer de pulmón para los 32 estados del país en un periodo de 20 años que comprende el periodo previo y posterior a la implementación de la LGCT y considera la totalidad de registros disponibles a través del sistema de información en salud (1998-2017); 2) contrastar la distribución territorial de la demanda por servicios oncológicos respecto a la disponibilidad de especialistas para atención; 3) determinar la capacidad predictiva de 3 diferentes algoritmos de clasificación supervisada que utilizan como insumo los datos sociodemográficos actualmente disponibles a través de los registros de mortalidad del sistema nacional en salud.

## MÉTODOS

El análisis se organizó en dos fases: la primera se basó en un análisis exploratorio de datos (AED), mientras la segunda fase se enfocó en modelación empleando técnicas de aprendizaje automatizado (*machine learning*).

En ambas fases, el insumo principal proviene de los microdatos de uso público, en formato csv, no cubos dinámicos, consignados mediante los establecimientos de servicio médico de las instituciones públicas del país. Estos se recopilaron a través del subsistema epidemiológico y estadístico de defunciones, a cargo del Sistema Nacional de Información de Salud (SINAIS)(1). Se consideró el total de bases de datos disponibles a la fecha de esta investigación que comprende el periodo de 20 años entre 1998 y 2017.

En la primera fase, se estimó las tasas de mortalidad por cáncer de pulmón ajustadas por estructura de edad para los 32 estados del país entre 1998 y 2017.

Cáncer de pulmón fue definido a partir de las siguientes causas específicas consignadas por la clasificación internacional de enfermedades (CIE-10): 1) C34: tumor maligno de los bronquios y del pulmón; 2) C340: tumor maligno del bronquio principal; 3) C341: tumor maligno del lóbulo superior, bronquio o pulmón; 4) C342: tumor maligno del lóbulo medio, bronquio o pulmón; 5) C343: tumor maligno del lóbulo inferior, bronquio o pulmón; 6) C348: lesión de sitios contiguos de los bronquios y del pulmón; 7) C349: tumor maligno de los bronquios o del pulmón, parte no especificada<sup>1</sup>.

La estimación fue realizada mediante el método directo considerando como población estándar los registros poblacionales del Consejo Nacional de Población (CONAPO) (8).

En la segunda sección se estima el indicador de concentración relativa denominado: cociente de localización (LQ), un indicador utilizado previamente en estudios empíricos en el área de salud pública (9) y el campo de epidemiología (10). Este

permite contrastar la distribución territorial de la demanda por servicios oncológicos respecto a la disponibilidad de especialistas para atención. La estimación se hace a escala municipal, lo que complementa la perspectiva examinada por las tasas de mortalidad estatales.

El indicador LQ para un momento del tiempo determinado se estima a partir la siguiente especificación:

$$LQ_t = (C/T)_j / (C/T)_N$$

Donde C corresponde al número de registros por defunciones para las siguientes causas C34, C40, C341, C342, C343, C348, C349, T corresponde al número total de registros por defunciones en el año  $t$ ,  $j$  representa una entidad federativa particular de la república mexicana y N indica el escenario nacional para un momento en el tiempo específico.

En ambas escalas geográficas, se compara la proporción entre el número de defunciones por cáncer de pulmón y el total de defunciones. El indicador se estima para el año más reciente disponible en el conjunto de datos.

En la tercera sección, se emplean técnicas de aprendizaje automatizado para determinar la capacidad predictiva de 3 diferentes algoritmos de clasificación supervisada que utilizan como insumo las variables sociodemográficas actualmente disponibles a través de los registros de mortalidad del sistema nacional en salud.

En este enfoque, el interés primordial es evaluar la capacidad predictiva de cada modelo, entendida a partir de su habilidad para clasificar correctamente si el paciente será diagnosticado con cáncer de pulmón, dadas las características del individuo, ejemplo las variables explicativas disponibles: sexo, edad y entidad de residencia de alta incidencia.

La selección de los métodos de aprendizaje automatizado estimados, parte desde los métodos clásicos y relativamente simples hasta un algoritmo más complejo que requiere mayor intensidad en el uso de recursos computacionales.

El primer modelo estimado es la especificación de regresión logística (RL) multivariada, un enfoque paramétrico con el que se pronostican las probabilidades asociadas a la razón de momios para las variables explicativas ya mencionadas. La especificación del modelo es:

$$\text{cancer pulmon } 1 | x_1, \dots, x_n = 1 / (1 - e^{-(\alpha + \beta_i x_i)})$$

Donde,  $x_1$ =Edad,  $x_2$ =Sexo y  $x_3$ =entidad de residencia de alta incidencia<sup>2</sup>.

Enseguida se estiman dos modelos de clasificación supervisada alternativos, uno empleando la técnica de *K-nearest neighbors* (KNN), un modelo ampliamente utilizado en estudios aplicados, basado en un enfoque no paramétrico a partir de una estructura relativamente simple (11), que a diferencia del modelo de regresión logística, no asume una forma funcional particular para el umbral de decisión, lo que permite una mayor flexibilidad.

En este enfoque de clasificación supervisada, para cada observación de la variable explicativa  $x_i$ ,  $i=1, \dots, n$  tenemos asociada una medición de respuesta  $y_i$ , siendo el objetivo estimar un modelo que relaciona la variable de respuesta con las

<sup>1</sup> Se omite la causa D022: Carcinoma in situ del bronquio y del pulmón debido a que representa menos del 0,1% de los casos registrados.

<sup>2</sup> Para efectos de esta investigación se entiende por "entidad de alta incidencia" aquella cuya tasa de mortalidad es superior al percentil 75, en la distribución de las tasas de mortalidad entre los 32 estados del país.

explicativas con la mayor precisión. Específicamente el algoritmo KNN, que intenta estimar la distribución condicional de Y (resultado positivo con diagnóstico de cáncer de pulmón) dado un conjunto de variables explicativas disponibles X.

El tercer método que requiere mayor intensidad computacional es el modelo de bosque aleatorio. Este enfoque requiere la construcción de árboles de decisión que estratifican el espacio de predicción (variable objetivo) a partir de múltiples nodos conectados entre sí por segmentos o ramas. La medida de pureza entre los segmentos es valorada bajo el criterio de entropía. Una discusión técnica de este enfoque fue presentada por Gareth y cols. (11).

La capacidad predictiva de los tres modelos se evalúa utilizando datos fuera de la muestra, para lo cual se particiona la base original en dos subconjuntos, uno denominado de entrenamiento, con el 75% de los datos, sobre el cual se calibrarán los parámetros de los modelos y uno de validación con el 25% restante.

Con el propósito de evaluar la capacidad predictiva de los tres modelos, se estiman los indicadores de sensibilidad y especificidad. El primero expresado como la tasa de verdaderos positivos (TVP =  $vp/vp+fn$ ). Este indicador muestra la proporción de registros con predicción positiva para la causa cáncer de pulmón de entre el total de casos observados como defunciones por esta causa.

Adicionalmente se estima el indicador de especificidad, expresado como la tasa de predicciones verdadero negativo (TVN =  $vn/vn+fp$ ). Este indicador representa la habilidad del modelo para clasificar los pacientes sanos, ya que muestra la proporción de predicciones negativas entre los registros que no presentan cáncer de pulmón como causa de defunción. Donde  $vp$  = verdadero positivo,  $fp$  = falso positivo, (falsa alarma, o en términos estadísticos, un error tipo I, equivalente a rechazar una hipótesis nula verdadera, ej. rechazar,  $H_0$ : el paciente no falleció a causa de cáncer de pulmón;  $vn$  = verdadero negativo,  $fn$  = falso negativo (en términos estadísticos, un error tipo II, ej. fallamos en rechazar una hipótesis nula falsa, ej.  $H_0$ : el paciente no falleció a causa de cáncer de pulmón).

Importante notar que el término positivo o negativo se refiere a la presencia o ausencia de causa de defunción cáncer de pulmón.

En términos teóricos, una predicción perfecta corresponde a una sensibilidad del 100%, en conjunto con una tasa de resultados falsos positivos (TFP)=0.

En este caso, por tratarse de un padecimiento como cáncer, nos interesa contar con un modelo que aporte el mayor nivel de sensibilidad.

## RESULTADOS

### Evolución temporal de las tasas de mortalidad por cáncer de pulmón en México

La estimación busca contar con elementos para evaluar la evolución del cáncer de pulmón en un horizonte temporal antes y después de la entrada en vigor de esquemas de regulación dirigidos a la prevención del tabaquismo (figura 1).

A escala nacional la tasa de mortalidad por cáncer de pulmón observó una reducción de 14,5% entre el periodo previo a la implementación de la Ley General para el Control del Tabaco (LGCT) y la trayectoria observada en el periodo

posterior, entre 2008 y 2017. Sin embargo, la estimación revela un comportamiento diferenciado entre la población masculina y femenina. Considerando el periodo total de estudio, encontramos que la población de sexo masculino observó una reducción en su tasa de mortalidad del 22% en el transcurso de dos décadas.

A escala territorial, observamos que la mayor reducción para el sexo masculino ocurrió en los estados del norte del país: Baja California Sur (54%). Baja California, (50%), Sonora, (47%) y Coahuila con una reducción del 36%. Si bien en la mayoría de los estados, (20 estados), hubo una reducción en las tasas de mortalidad entre esta población, en 9 estados tenemos un incremento, siendo el mayor incremento en Tlaxcala y Querétaro, entidades donde las tasas de mortalidad de hecho se duplicaron en el periodo de referencia.

Mientras las tasas de mortalidad en el periodo de estudio indican una reducción para el caso de la población masculina, para el caso femenino observamos un incremento del 8% entre el escenario para el año de inicio del periodo de estudio y al año más reciente 2017 de la serie. De hecho, en la mayoría de las entidades se observa un incremento de sus tasas de mortalidad entre la población femenina, localizado principalmente en estados pertenecientes a la región sur del país. Los mayores incrementos en las tasas de mortalidad por cáncer de pulmón entre la población femenina se presentaron en los estados de Puebla (con un incremento de 200%), Guerrero (200%), Oaxaca (100%), Yucatán (100%) y Veracruz (50%).

La figura 2 sintetiza las variaciones encontradas con énfasis en las diferencias por sexo para cada entidad federativa.

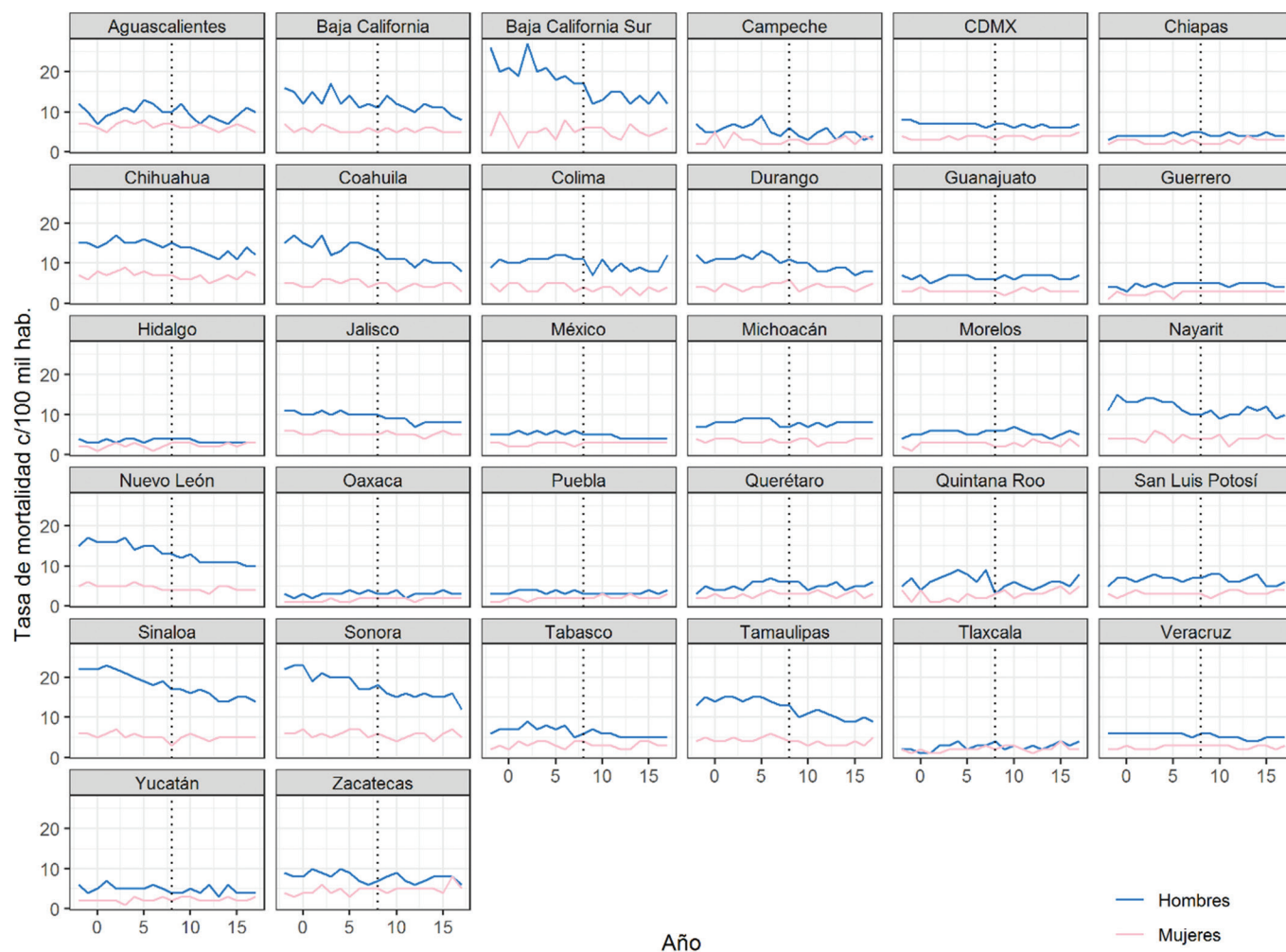
### Distribución territorial de la demanda por servicios oncológicos y disponibilidad de especialistas para atención

Para analizar la distribución territorial de las defunciones por cáncer de pulmón con mayor detalle se estima a continuación el cociente de localización (LQ) con desagregación a escala municipal. Esta medida de concentración relativa permite el contraste de los niveles observados en cada municipio respecto a la concentración observada en el país, representada por el umbral  $LQ=1$ . La figura 3 muestra el resultado.

El patrón de concentración que aporta el cociente de localización LQ para los registros de defunciones por cáncer de pulmón, permite determinar áreas de atención prioritaria, entre ellas diversos municipios al interior de las entidades de alta incidencia como Sinaloa, en los municipios: Choix, Elota o Angostura en donde se observan concentraciones relativas equivalentes a 6, 4,5 y 3,8 veces la que presenta el territorio nacional.

Asimismo, se detecta una importante presencia del fenómeno en municipios del centro occidente y norte del país, incluidos municipios como Teocaltiche, Jalisco, ( $LQ= 6.14$ ), Saucillo y Delicias en el estado de Chihuahua con  $LQ=2.49$  y 3.24 respectivamente y Cananea en Sonora ( $LQ=3.41$ ).

Un segundo rasgo de interés en combinación con la evidencia anterior es la distribución de los médicos con especialidad de oncología registrados. En total la red de médicos oncólogos en centros hospitalarios públicos de México muestra que se tienen 1 039 especialistas activos, con presencia en 303 hospitales (12), no obstante, dentro de estos hospitales encontramos una notable inequidad ya que el 32,4% de los centros hospitalarios concentran el 71,2% de los oncólogos disponibles.

**FIGURA 1. Tasas de mortalidad por cáncer de pulmón ajustadas por estructura de edad y sexo en México. 1998-2017.**

Fuente: elaboración propia con datos de SINAIS, 1998-2017. La línea punteada de referencia marca el año de implementación de la LGCT.

### Proyecciones a partir de herramientas de modelación de aprendizaje automatizado

Se presenta primero el modelo de regresión logística controlando por sexo, edad y entidad de residencia de alta incidencia, la especificación seleccionada incorpora el efecto no lineal observado de la variable edad respecto a la variable objetivo. La variable objetivo indica si el individuo falleció por cáncer de pulmón (1,0), y la variable sexo toma el valor 1: hombre, 0: mujer, la edad de la persona en años y se controla por la residencial del individuo, con valores 1, en caso de residir en entidad de alta incidencia por cáncer de pulmón, cero en caso contrario.

La estimación permite corroborar los resultados encontrados por la fase de análisis exploratorio de datos y el modelo indica mayor probabilidad de defunción por cáncer de pulmón para el sexo masculino,  $OR = 1,67$ ,  $CI (1,59-1,76)$ ,  $p < 0,001$  particularmente en el rango entre los 60 y 80 años. La variable edad indica un  $OR = 1,11$ ,  $CI (1,10-1,13)$ ,  $p < 0,001$ . Mientras la variable que registra si el estado es de alta incidencia con  $OR = 1,54$ ,  $CI(1,46-1,62)$ ,  $p < 0,001$ .

La figura 4 muestra el contraste edad-sexo respecto a estados de alta incidencia.

La evidencia indica que si bien la población de sexo masculino exhibe mayores probabilidades relativas al sexo femenino y estas aumentan a medida que avanza la edad del individuo, estas alcanzan su mayor nivel cercano a 22,5% en aquellos estados considerados como de alta incidencia, valor que contrasta con las probabilidades menores que se hallan para el sexo femenino en estados no considerados de alta incidencia, cuya probabilidad estimada alcanza el 10,2%.

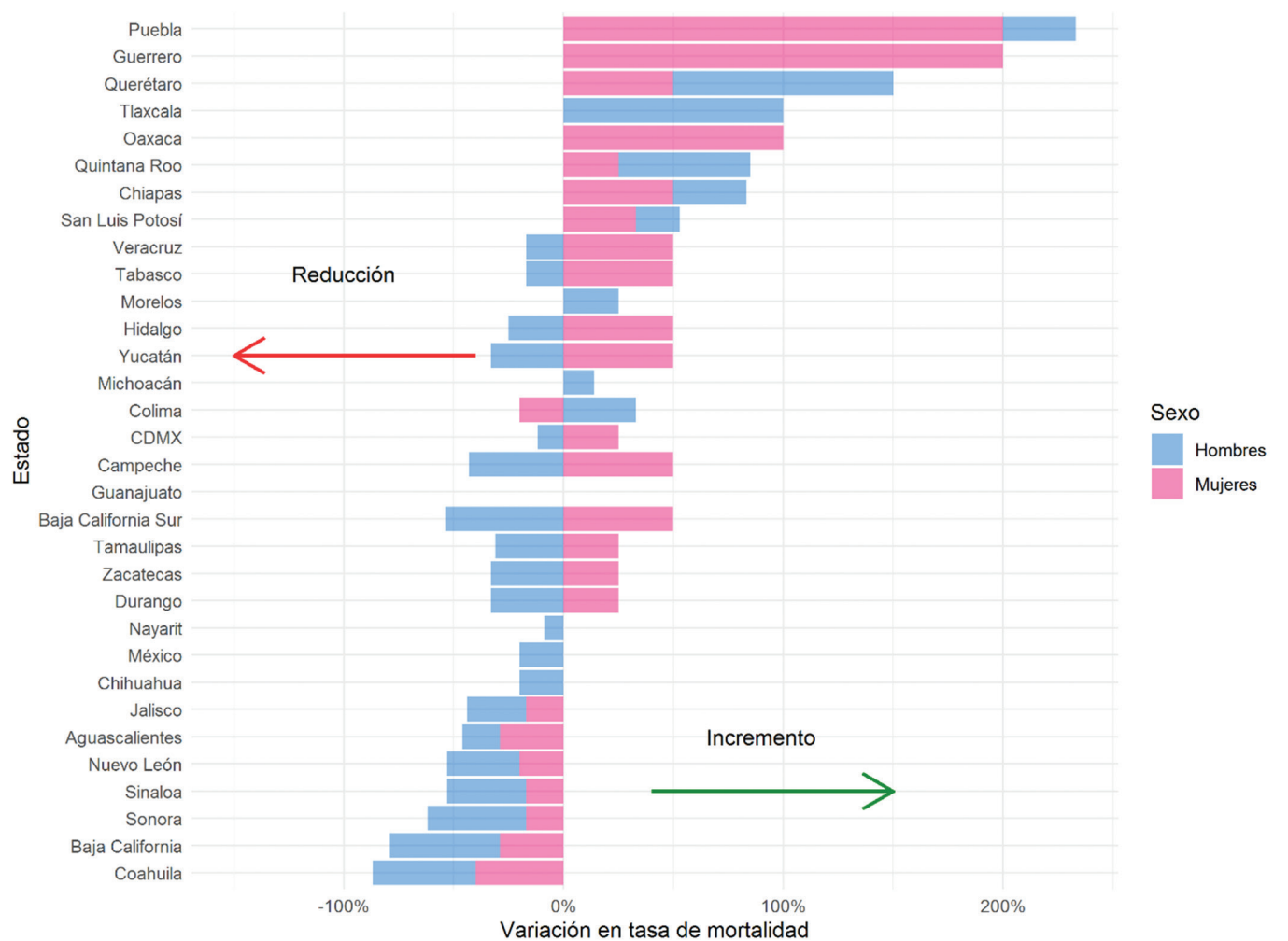
El cuadro 1, sintetiza las predicciones de los tres modelos estimados mediante la representación de los indicadores de desempeño.

### DISCUSIÓN

La evidencia en retrospectiva para 20 años de registros de mortalidad en México, muestra una tendencia contra cíclica dentro del grupo de padecimientos oncológicos, con una disminución causa específica para cáncer de pulmón del orden de 22% en el periodo entre 1998 y 2017.



FIGURA 2. Variación en la tasa de mortalidad por cáncer de pulmón entre 1998 y 2017 en México según entidad federativa.



Fuente: elaboración propia con datos de SINAIS, 1998-2017.

La trayectoria documentada implica una reducción de costos de atención y por ende de la carga para el sistema de salud, lo que sugiere en principio un impacto positivo de los mecanismos regulatorios implementados en respuesta a esta amenaza de salud pública, a pesar de los retos identificados por estudios previos sobre los mecanismos de implementación de la ley general para el control del tabaco (13).

Desde la perspectiva geográfica, llama la atención en esta trayectoria descendente, que se tiene un patrón regional en el cual las entidades ubicadas en el norte del país registraron las mayores reducciones.

Adicionalmente los resultados aportan un contraste relevante en términos de las diferencias por sexo ya que, a excepción del sexo masculino, para las mujeres, se encontró un incremento en estados localizados fundamentalmente en el sur del país, como Guerrero, Oaxaca y Puebla. Esta evidencia indica un contraste fundamental que sugiere roles de género y hábitos distintos.

Se observa también que los estados en donde las mujeres están teniendo una mayor mortalidad por cáncer de pulmón son estados con un importante componente demográfico indígena, un rasgo que ha sido documentado previamente como determinante

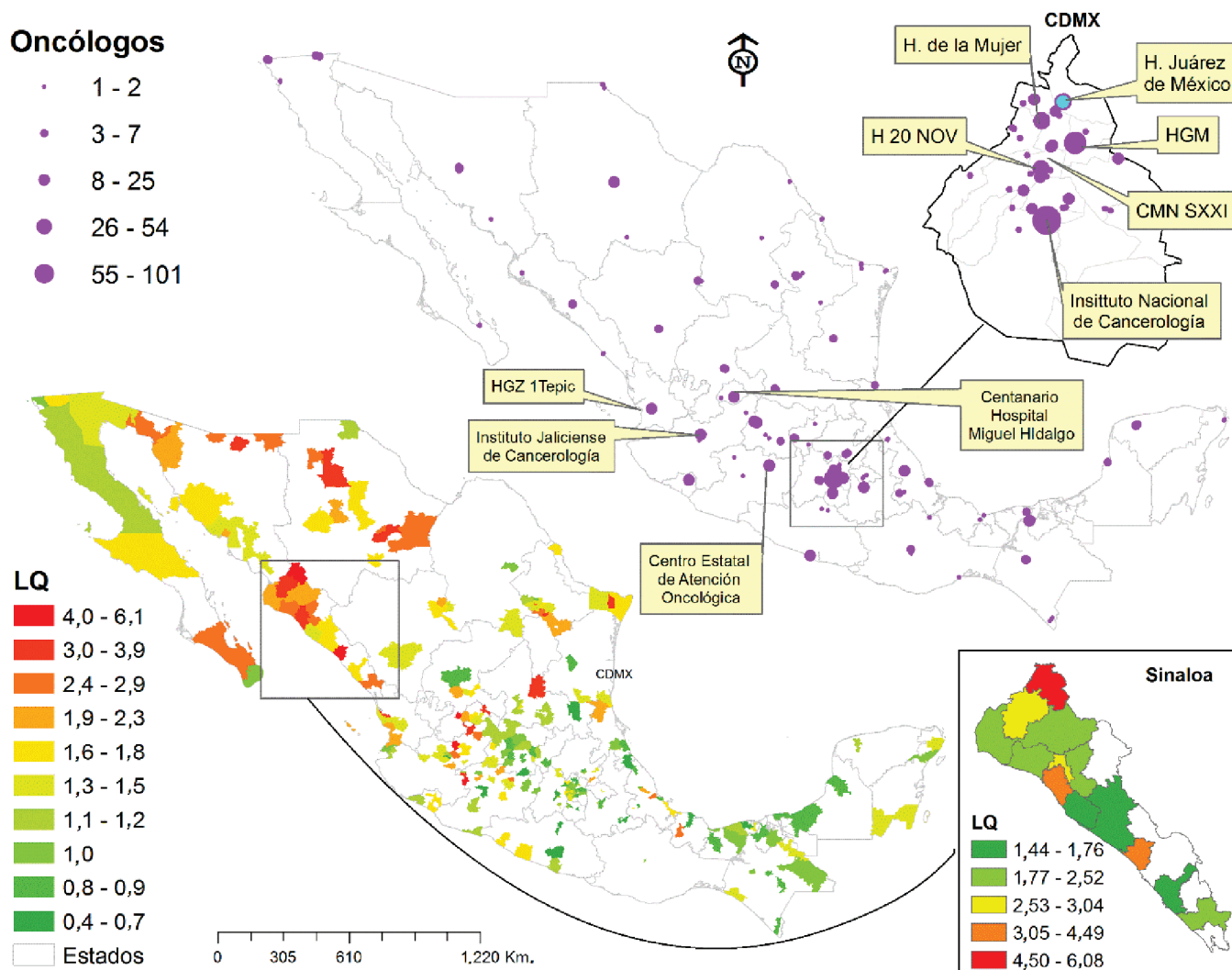
social de afecciones pulmonares, incluidos aquellos de naturaleza oncológica, incluso en ausencia de tabaquismo (14-16).

Desde la perspectiva territorial, un reto identificado por el presente estudio que se traduce en acciones de política pública específicas se refiere al desequilibrio territorial entre la demanda por servicios oncológicos y la distribución de especialistas en unidades hospitalarias en el país.

Evidencia que refuerza los indicadores desarrollados por otros análisis especializados que muestran el rezago de México en esta materia en el contexto de América Latina, donde otros países de la región como Brasil muestran un indicador de oncólogos por cada 100 mil habitantes que duplica el de México con 17 contra solo 8 respectivamente (17).

Por su parte, los resultados en la fase de modelación en el contexto de aprendizaje automatizado demostraron la utilidad de los registros electrónicos de mortalidad que genera el sistema nacional de información en salud, para aportar al entendimiento de la ocurrencia de padecimientos como cáncer de pulmón. En este contexto el modelo de regresión logística permitió una clasificación acertada (sensibilidad) de las defunciones por esta causa del 74% de los casos.

**FIGURA 3. Concentración de casos por municipio según cociente de localización LQ y distribución de oncólogos por establecimiento de servicio de salud (CLUES) en México. 2017.**



Fuente: elaboración propia con datos de (1,18).

Leyenda: H: Hospital; LQ: cociente de localización; H 20 de Nov: Hospital 20 de noviembre; HGM: Hospital General de México; CMN: Centro Médico Siglo XXI; HGZ: Hospital General de Zona 1 Tepic; H. Juárez: Hospital Juárez; H. de la Mujer: Hospital de la Mujer.

El indicador de sensibilidad es particularmente importante en el contexto de afecciones de salud como cáncer, ya que el costo para el paciente de un diagnóstico falso negativo, cuando en realidad la enfermedad está presente, puede tener importantes consecuencias, que incluyen postergar el inicio del tratamiento a etapas avanzadas cuando el cáncer ya se ha propagado a otros órganos.

Adicionalmente, se observó una clara diferencia entre los modelos evaluados, con un rango de 10 puntos porcentuales entre el modelo de mayor y el de menor sensibilidad. No obstante, la incorporación de registros públicos que permitan el seguimiento de pacientes con un mayor grado de detalle persiste como reto para mejorar la capacidad predictiva en este enfoque.

En cuanto a las limitaciones del presente estudio, es importante mencionar que el desfase en la disponibilidad de datos constituye una oportunidad de mejora en los procesos de disseminación de la información en materia de salud en México, ya que a pesar de que el estudio se realiza en 2020, los registros

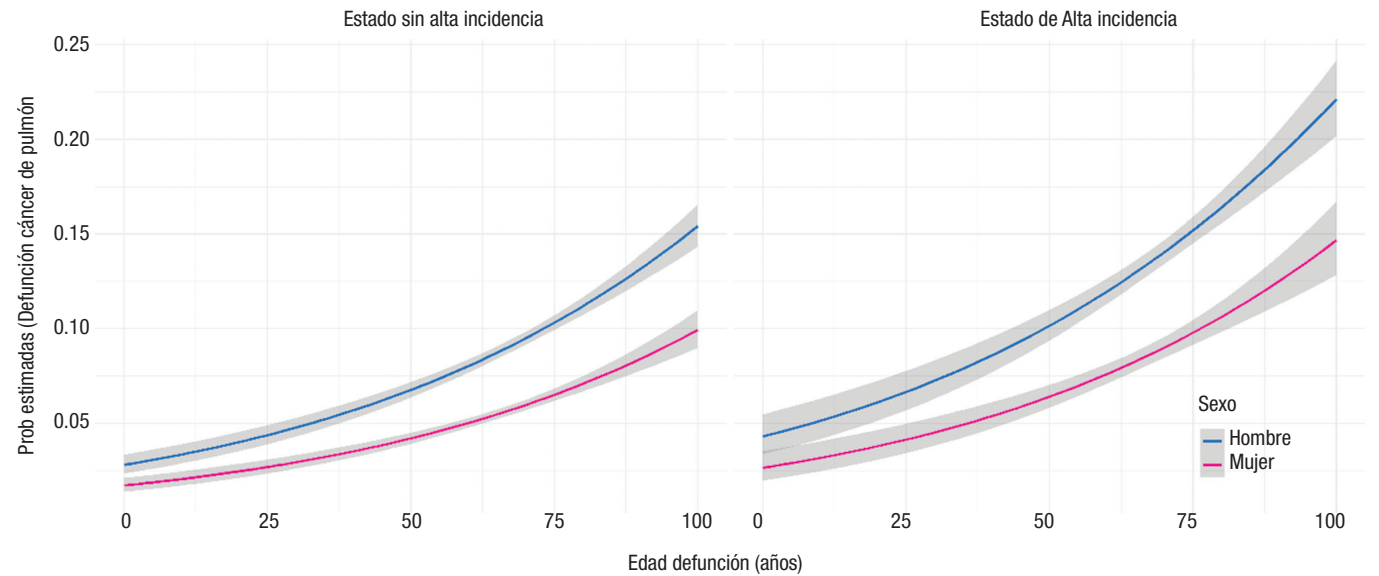
de mortalidad públicos más actualizados disponibles a la fecha corresponde a 2017.

Respecto a la fase de modelación es importante puntualizar que los métodos en el contexto de aprendizaje automatizado empleados se basan en una muestra de entrenamiento de modo que la calidad de los insumos es fundamental para la obtención de predicciones útiles.

En este sentido, las fuentes de información empleadas se basan en protocolos estandarizados que no obstante pueden presentar problemas de subregistro asociados a diferencias regionales. Adicionalmente, si bien, el presente estudio analizó un volumen de datos considerable con 500 mil registros por año en promedio, el número de variables disponible es una seria limitante que representa un área de oportunidad para mejora.

En conclusión, a pesar de las reducciones observadas en las últimas dos décadas en los patrones de mortalidad por cáncer de pulmón en México, el análisis de los registros electrónicos de defunciones muestra que persisten áreas de oportunidad,

**FIGURA 4. Probabilidades estimadas por modelo de regresión logística. Para defunción por cáncer de pulmón en función del sexo y edad según estado de residencia, México 2017.**



Fuente: elaboración propia con datos de (1). R<sup>2</sup> Nagelkerke=0.36

**CUADRO 1. Indicadores de desempeño para los modelos de clasificación de defunciones a causa de cáncer de pulmón. México 2017.**

Modelo	TVP (Sensibilidad) 1-TFN	TVN (Especificidad) 1-TFP	TFP	TFN
RL	0,74	0,40	0,60	0,26
Random Forest	0,67	0,52	0,48	0,33
KNN	0,64	0,56	0,44	0,36

Fuente: elaboración propia con datos de (1).

principalmente entre la población femenina de los estados de Guerrero, Oaxaca y Puebla. Determinar si estos patrones se asocian a los hábitos del tabaquismo o bien a otros determinantes sociales, constituye la principal recomendación para la dirección de los esfuerzos de atención de esta enfermedad oncológica en México derivada de esta investigación

Adicionalmente, el estudio aporta elementos que pueden ser considerados en el proceso de mejora del programa nacional de

atención como la contratación de especialistas en regiones específicas para corregir el desbalance territorial detectado entre la demanda y los recursos disponibles para atención.

**Contribución de los autores.** JLMR concibió el estudio original, recolectó y analizó los datos, interpretó los resultados, escribió y revisó el manuscrito. El autor revisó y aprobó la versión final.

**Agradecimientos.** Al consejo nacional de ciencia y tecnología por el apoyo otorgado mediante el programa de estímulos al sistema nacional de investigadores.

**Conflicto de intereses.** El autor declara no tener conflicto de intereses, incluidos aquellos de tipo financiero, académico, personal, político o religioso, por la filiación institucional o de ninguna otra índole.

**Declaración.** Las opiniones expresadas en este manuscrito son responsabilidad del autor y no reflejan necesariamente los criterios ni la política de la *RPSP/PAJPH* y/o de la OPS

REFERENCIAS

1. SINAIS. Bases de Datos Estándar Defunciones. 2017. Disponible en: [http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std\\_defunciones\\_gobmx.html](http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std_defunciones_gobmx.html). Acceso el día 29 de enero de 2020

2. World Health Organization. Cancer today [Internet]. 2018. Disponible en: [https://gco.iarc.fr/today/online-analysis-table?v=2018&mode=population&mode\\_population=countries&population=900&populations=900&key=asr&sex=0&cancer=39&type=1&statistic=5&prevalence=0&population\\_group=0&ages\\_group%5B%5D=0&ages\\_group%5B%5D=17&group\\_cancer=1&include\\_nmsc=1&include\\_nmsc\\_other=1](https://gco.iarc.fr/today/online-analysis-table?v=2018&mode=population&mode_population=countries&population=900&populations=900&key=asr&sex=0&cancer=39&type=1&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmsc=1&include_nmsc_other=1). Acceso el día 17 de febrero de 2019

3. OECD. Indicator overview: comparative performance of countries and major trends. En: Health at a Glance 2019. OECD; 2019. (Health at a Glance).

4. Kantarjian H, Yu PP. Artificial Intelligence, Big Data, and Cancer. JAMA Oncol. 2015 Aug;1(5):573-4. doi: 10.1001/jamaoncol.2015.1203. PMID: 26181906.

5. Bertucci F, Le Corroller-Soriano AG, Monneur A, Fluzin S, Viens P, Maraninchi D, Goncalves A. Santé numérique et « cancer hors les murs », Big Data et intelligence artificielle [E-health and "Cancer outside the hospital walls", Big Data and artificial intelligence].

- Bull Cancer. 2020 Jan;107(1):102-112. French. doi: 10.1016/j.bulcan.2019.07.006. Epub 2019 Sep 19. PMID: 31543271.
6. Cha HS, Jung JM, Shin SY, Jang YM, Park P, Lee JW, et al. The Korea Cancer Big Data Platform (K-CBP) for Cancer Research. *International journal of environmental research and public health*. 28 de junio de 2019;16(13):2290.
  7. Riba M, Sala C, Toniolo D, Tonon G. Big Data in Medicine, the Present and Hopefully the Future. *Front Med (Lausanne)*. 2019 Nov 15;6:263. doi: 10.3389/fmed.2019.00263. PMID: 31803746; PMCID: PMC6873822.
  8. CONAPO. Proyecciones de la Población de los Municipios de México, 2015-2030 [Internet]. gob.mx. [citado 28 de agosto de 2020]. Disponible en: <http://www.gob.mx/conapo/documentos/proyecciones-de-la-poblacion-de-los-municipios-de-mexico-2015-2030>. Acceso el día 4 de diciembre de 2018
  9. Djira G, Schaarschmidt F, Fayissa B. Inferences for Selected Location Quotients with Applications to Health Outcomes. *Geographical Analysis*. 1 de julio de 2010;42:288-300.
  10. Schuurman N, Hameed SM, Fiedler R, Bell N, Simons RK. The spatial epidemiology of trauma: the potential of geographic information science to organize data and reveal patterns of injury and services. *Can J Surg*. octubre de 2008;51(5):389-95.
  11. Gareth J., T. Hastie, R. Tibshirani, D. Witten. An introduction to statistical learning : with applications in R. 8th ed. G. Casella, S. Fienberg IO, editor. New York: New York : Springer, [2013] ©2013; 2013. 1-426 p.
  12. Secretaría de Salud. Recursos en Salud [Internet]. 2017. Disponible en: [http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std\\_clues\\_gobmx.html](http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std_clues_gobmx.html)
  13. Reynales-Shigematsu LM. El control del tabaco, estrategia esencial para reducir las enfermedades crónicas no transmisibles. En: Políticas para el control del tabaco en México [Internet]. Primera. México: Instituto Nacional de Salud Pública; 2013. (Salud pública y tabaquismo; vol. 1). Disponible en: <https://www.insp.mx/produccion-editorial/publicaciones-anteriores-2010/2729-salud-publica-y-tabaquismo-volumen-i-politicas-para-el-control-del-tabaco-en-mexico.html>. Acceso el día 16 de febrero de 2019
  14. Imaz Olguín V, Green Renner D, Gerson Cwilich R, Alatorre Alexander J, Martos Ramírez G, Green Schneeweiss L, et al. Demographic characteristics of lung cancer and its association with wood smoke exposure in Mexican population. *Anales Médicos de la Asociación Médica del Centro Médico ABC*. 2018;63(3):188-93.
  15. Corrales L, Rosell R, Cardona AF, Martín C, Zatarain-Barrón ZL, Arrieta O. Lung cancer in never smokers: The role of different risk factors other than tobacco smoking. *Critical Reviews in Oncology / Hematology*. 2020;102895.
  16. Ramirez, M., Mireles Lezama, P., Jaimes Ramirez, S., & Gomora Lara B. La experiencia de las estufas ahorradoras de leña en dos comunidades indígenas del Estado de México | Ambiente y Desarrollo. 2012;16(31):91-105.
  17. The IQVIA Institute. Global Oncology Trends 2018 - IQVIA. Parsippany, NJ; 2018.
  18. DGiS. Recursos en Salud [Internet]. Recursos en Salud 2017. Disponible en: [http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da\\_recursos\\_gobmx.html](http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da_recursos_gobmx.html). Acceso el día 9 de mayo de 2018

---

Manuscrito recibido el 5 de mayo de 2020. Aceptado para su publicación, tras revisión, el 18 de septiembre de 2020.

## Resources for the evaluation of public policy on lung cancer in Mexico

### ABSTRACT

**Objectives.** To estimate lung cancer mortality rates adjusted by age distribution for the country's 32 states between 1998 and 2017; to contrast the territorial distribution of demand for oncological services with the availability of specialists to provide care; and to determine the predictive capacity of three different supervised classification algorithms in the context of automated learning techniques.

**Methods.** An exploratory analysis and data modeling were conducted, considering death records from the national health information system.

**Results.** Deaths from lung cancer in Mexico dropped by 14.5% between the period prior to implementation of the General Law on Tobacco Control and the subsequent period. A 22% reduction was observed in the male population by the end of the entire period. There is evidence of an imbalance between the demand for oncological services and the availability of specialists. The modeling phase demonstrated the usefulness of the country's electronic death records.

**Conclusions.** Despite reductions in lung cancer mortality patterns in Mexico in the last two decades, the analysis showed persistent areas of opportunity for improvement, mainly in the female population of Guerrero, Oaxaca, and Puebla states. Based on this research, the main recommendation for focusing efforts to manage this oncological disease in Mexico is to determine whether these patterns are associated with smoking habits or with other social determinants.

### Keywords

Lung neoplasms; mortality; Mexico.



---

## Recursos para avaliação da política pública de câncer de pulmão no México

### RESUMO

**Objetivos.** Estimar o índice de mortalidade por câncer de pulmão, ajustado pela estrutura etária, nos 32 estados do México no período entre 1998 e 2017, comparar a distribuição territorial da demanda por serviços oncológicos com a disponibilidade de especialistas e determinar a capacidade preditiva de 3 algoritmos utilizados na classificação supervisionada no contexto de técnicas de aprendizagem automatizada.

**Métodos.** Realizou-se uma análise exploratória e modelagem de dados dos registros de mortalidade cedidos pelo sistema nacional de informação em saúde.

**Resultados.** Houve uma redução de 14,5% no índice de mortalidade por câncer de pulmão no México entre o período pré-implementação da Lei Geral para o Controle do Tabaco e o período pós-implementação. Verificou-se uma redução de 22% da mortalidade na população masculina em todo o período considerado. Evidenciou-se desequilíbrio entre a demanda por serviços oncológicos e a disponibilidade de especialistas. A fase de modelagem permitiu demonstrar a utilidade dos registros eletrônicos de mortalidade existentes no país.

**Conclusões.** Apesar da redução nos padrões de mortalidade por câncer de pulmão no México nas últimas duas décadas, a análise demonstrou que persistem áreas de oportunidade, sobretudo na população feminina nos estados de Guerrero, Oaxaca e Puebla. Determinar se estes padrões estão associados ao tabagismo ou a outros determinantes sociais é a principal recomendação desta pesquisa para guiar os esforços de atenção ao câncer no México.

### Palabras-clave

Neoplasias pulmonares; mortalidade; México

---