

Indicadores alimentarios identificados en publicaciones de la red social Twitter y su asociación con la prevalencia de obesidad en Argentina



Haluszka Eugenia^{1,2}, Aballay Laura R¹, Pou Sonia A^{1,2}, Wirtz Baker Julia M^{1,2}, Díaz Oroz Emilia B¹, Pastore Alejandra C¹, Peralta Sparacino Valentina¹, Zonghetti Rocio¹, Niclis Camila^{1,2}

¹Centro de Investigaciones en Nutrición Humana (CenINH), Escuela de Nutrición, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba.

²Instituto de Investigaciones en Ciencias de la Salud (INICSA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) – Facultad de Ciencias Médicas, Universidad Nacional de Córdoba.

Área de conocimiento: Epidemiología y Salud Pública

Resumen

La obesidad representa un problema relevante debido a su creciente prevalencia e impacto sanitario. Las redes sociales podrían ser una fuente válida de información para el estudio de fenómenos relacionados con la salud de las poblaciones. El objetivo de este trabajo fue realizar una aproximación al consumo alimentario poblacional utilizando publicaciones de *Twitter* y evaluar la asociación entre indicadores construidos a partir de ese contenido y el patrón geográfico de carga de obesidad en Argentina.

Se realizó un estudio ecológico multigrupal, utilizando datos de la Encuesta Nacional de Factores de Riesgo 2018 para calcular tasas ponderadas de prevalencia de obesidad (escala provincial), y 6023548 *tweets* georreferenciados recolectados durante 2021–2022. Se construyeron indicadores proxy de la alimentación geolocalizados, empleando procesamiento del lenguaje natural (IA) a partir del contenido compartido en *Twitter*. Se elaboraron mapas y se calculó el Índice de autocorrelación de Moran para detectar patrones espaciales. Además, se estimó la correlación entre los indicadores construidos y la prevalencia de obesidad total, por sexo y grupo etario.

La distribución de la prevalencia de obesidad, la tasa de *tweets* alimentarios y el Índice de Densidad Nutricional-IDN, presentaron una distribución espacial no aleatoria ($p < 0,05$). La frecuencia de mención de algunos alimentos considerados ‘saludables’ (pepino, pomelo, naranja, atún, carne, entre otros) se correlacionó inversamente con la prevalencia de obesidad a nivel provincial, mientras que la mención de alimentos como molleja, queso semiduro, chocolate, morcilla, hamburguesas y golosinas, presentaron una correlación positiva. En algunos casos, estos resultados variaron según sexo y grupo etario. Por último, una mayor tasa de *tweets* alimentarios, se asoció con mejor IDN

promedio a nivel provincial ($p=0,04$).

Concluyendo, este estudio resalta la utilidad de las redes sociales como fuente para explorar tendencias alimentarias relacionadas a obesidad, lo cual favorecería la formulación de políticas efectivas, siendo necesario profundizar esta línea de estudio.

Palabras claves: Obesidad; Twitter; Redes sociales; Indicador alimentario.

Autor de correspondencia: Niclis, Camila. E-mail: cniclis@fcm.unc.edu.ar

INTRODUCCIÓN

La obesidad es una enfermedad crónica no transmisible (ECNT), definida por la Organización Mundial de la Salud (OMS) como una acumulación anormal o excesiva de grasa que puede ser perjudicial para la salud. Esta patología representa un problema sanitario de gran relevancia debido a las múltiples consecuencias físicas, psíquicas y sociales que conlleva, además de generar una alta carga económica para los presupuestos destinados a afrontar los costos de su tratamiento y el de sus enfermedades asociadas (1).

En Argentina, según la 4ta Encuesta Nacional de Factores de Riesgo (ENFR), la prevalencia de obesidad fue del 32,4% en el año 2018 (2). Su incremento en las últimas décadas explica, en gran parte, el elevado crecimiento en la mortalidad y morbilidad por ECNT atribuibles al exceso de peso, principalmente de diabetes, enfermedades cardiovasculares, accidentes cerebrovasculares y diversos tipos de cánceres (2-5). La obesidad, en su etiopatogenia es considerada una enfermedad multifactorial, reconociéndose a los factores genéticos, ambientales, metabólicos y endocrinológicos en su red causal (6).

Entre los factores de riesgo reconocidos para las ECNT, los hábitos alimentarios

poco saludables destacan como uno de los principales contribuyentes a la carga de morbilidad y mortalidad a nivel mundial. Muchos países, especialmente aquellos de ingresos bajos y medios, enfrentan problemas de salud y nutrición debido a la coexistencia de desnutrición, deficiencias de micronutrientes, y sobrepeso u obesidad (7). Diversos estudios permiten describir el consumo alimentario a nivel poblacional, así como otros factores de riesgo asociados al desarrollo de la obesidad. Las encuestas de salud de base poblacional más representativas en Argentina son la ENFR y la Encuesta Nacional de Nutrición y Salud (ENNyS) (2, 8). Ambas encuestas son herramientas fundamentales y valiosas para conocer los factores de riesgo, el estado nutricional y los hábitos de consumo de la población. Sin embargo, se realizan con intervalos prolongados debido a los grandes esfuerzos prácticos y costos que su implementación requiere.

Actualmente, las redes sociales están siendo utilizadas como fuente de información para estudiar diversos fenómenos relacionados a la salud de las poblaciones (9). Las mismas se definen como comunidades formadas por diferentes usuarios y organizaciones que se relacionan entre sí, en plataformas de internet desde diferentes lugares del mundo. Éstas ofrecen múltiples

herramientas para interactuar, convirtiéndose en un medio dinámico que permite compartir información personal, difundir lo que se publica por otros usuarios e interactuar entre ellos (10). La red social *Twitter*, actualmente conocida como X, es una aplicación gratuita que permite a los usuarios mantenerse conectados a través de mensajes de texto breves (*microblogging*) de un máximo de 280 caracteres de longitud. Los datos provenientes de esta red social se han comenzado a utilizar para realizar estudios epidemiológicos y reconocer patrones relacionados a procesos de salud-enfermedad compartidos en los *tweets* (10). De este modo, surge el término infodemiología, definida como una ciencia que se encarga de supervisar y analizar el comportamiento de lo que los usuarios comparten en la web, a fin de predecir, evaluar, e incluso prevenir problemas relacionados con la salud (11). El principal objetivo de este trabajo fue utilizar el contenido compartido en la red social *Twitter* (actualmente denominada X) para realizar una aproximación a los hábitos alimentarios de los usuarios, aportando datos a la vigilancia epidemiológica. Se propone además evaluar la asociación entre indicadores construidos a partir del texto extraído de esa red social y la carga de obesidad de diferentes zonas geográficas de Argentina

(2). De esta manera, se plantea reconocer nuevas fuentes de información, así como herramientas innovadoras y útiles para la identificación de factores alimentarios y otros indicadores de estilos de vida orientados a ampliar la información para la vigilancia epidemiológica de la población argentina.

METODOLOGÍA

El presente trabajo es un estudio ecológico multigrupal, correlacional, el cual se implementó en las etapas o fases que se describen a continuación.

Recolección de datos de Twitter

En la fase inicial se recolectaron *tweets* utilizando la API (Interfaz de Programación de Aplicaciones) correspondiente a la red social *Twitter*, denominada Tweepy (12), la cual permite extraer los *tweets* en tiempo real. Para acceder a la misma fue necesario rellenar un formulario disponible en la página oficial de *Twitter*, indicando el uso de los datos y consintiendo respetar las normas establecidas para el uso de los mismos. El procedimiento de recopilación de *tweets* se llevó a cabo de manera aleatoria utilizando el filtro de geolocalización a través de las coordenadas que engloban en un rectángulo a los límites de Argentina,

durante el periodo 2021–2022. Estos fueron obtenidos a partir de la herramienta online gratuita Bounding Box. Una vez recolectados los *tweets*, se almacenaron en una base de datos no relacional, utilizando MongoDB (13). Todo el proceso de recolección y procesamiento se realizó con el lenguaje de programación Python, y se utilizó un entorno de Jupyter Notebook (14) para su realización.

Una vez recolectados los *tweets* se les aplicó un pre-procesamiento que consiste en la limpieza y curación del cuerpo de texto, eliminando los símbolos y otros elementos que interfieran en el análisis del mismo (emojis, hashtags, entre otros). Cada *tweet* fue geolocalizado a nivel provincial en relación a su locación reportada. Sobre el cuerpo de texto se aplicó la función de nlp provista por Spacy (librería de Python para análisis del lenguaje), la cual permitió identificar sustantivos en cada *tweet*. Entre los sustantivos se identificaron diversos alimentos mencionados en sus flexiones y sinónimos (singulares y plurales, diminutivos, según marcas, entre otros). Para facilitar este proceso, considerando que existen múltiples formas de mencionar cada alimento, además de regionalismos, posibles errores ortográficos o de tipeo y expresiones típicas ampliamente utilizadas en esta red social, se decidió aplicar una metodología

sistemática para la selección de palabras clave para cada alimento. Se aplicó una técnica denominada *Word Embedding* (WE) la cual permitió reconocer palabras que se encuentran cercanas en un espacio multidimensional en función a su relación semántica. El WE utilizado fue entrenado por investigadores de FAMAFA (15) con un corpus de *tweets* de Argentina. De esta manera se identificaron los 10 términos más cercanos por cada alimento y se utilizaron sólo aquellos que cumplían con un criterio de sinonimia, es decir, que podían ser utilizadas en el mismo contexto y no tenían un significado ambiguo. Una vez identificados los alimentos en cada *tweet* se procedió a construir los indicadores *proxy* de la alimentación. Se calculó la tasa de *tweets* alimentarios, que indica la proporción de *tweets* que mencionan alimentos sobre el total de *tweets* recolectados. Además, se calculó la frecuencia de mención de cada alimento y grupo de alimentos por provincia. Por último, se construyó el Índice de Densidad Nutricional (IDN), el cual indica que a mayor valor mejor es la calidad nutricional del alimento. La densidad nutricional de un alimento se construyó a partir de la suma del porcentaje de nutrientes considerados beneficiosos para la salud (proteínas, fibra dietética, vitamina A, vitamina C, vitamina E, calcio, magnesio, hierro y potasio) aportados en 100 Kcal de

ese alimento, teniendo en cuenta la ingesta recomendada para ese nutriente según la Ingesta Dietética Recomendada (IDR). A su vez se le restó el porcentaje aportado de nutrientes considerados perjudiciales para la salud, como las grasas saturadas, sodio y azúcares agregados, en 100 Kcal del alimento, teniendo en cuenta la IDR. El resultado fue una medida de la densidad nutricional de un alimento, que puede tomar valores negativos y positivos (16). Una vez calculado el IDN para cada alimento se estimó la densidad nutricional promedio por *tweet* teniendo en cuenta los alimentos mencionados en cada publicación y la cantidad de veces que aparecieron.

Estimación de prevalencias de obesidad

Para el cálculo y análisis de la prevalencia de obesidad se usaron datos de la 4° ENFR realizada durante el año 2018 (2) y datos poblacionales calculados a partir de proyecciones por interpolación exponencial en años intercensales del censo 2010 (17). Los datos usados tienen un nivel de agregación provincial, siendo las unidades de observación las veintitrés provincias argentinas y la Ciudad Autónoma de Buenos Aires, CABA (veinticuatro jurisdicciones político-administrativas de la República Argentina).

La cuarta ENFR fue llevada a cabo por el Ministerio de Salud de la Nación en conjunto con el Instituto Nacional de Estadísticas y Censos (INDEC) e incluyó localidades con 5000 habitantes o más (17).

Se calcularon las prevalencias (%) de obesidad a nivel provincial, por sexo y grupo etario (menores o mayores a 40 años), utilizando los factores de expansión de la ENFR para ajustar los resultados a la población total.

Construcción de mapas y análisis estadístico

La prevalencia de obesidad, la tasa de *tweets* alimentarios y el IDN, construidos en base a las publicaciones de *Twitter*, fueron utilizados para la elaboración de mapas temáticos en QGIS 3.28, utilizando los datos clasificados según quintiles de su distribución. A fines de detectar patrones espaciales agrupados o dispersos en esos indicadores, se calculó el Índice de autocorrelación de Moran.

Finalmente, todos los indicadores proxy de la alimentación fueron correlacionados con las tasas de prevalencia de obesidad a nivel provincial, tanto para población total como estratificado por sexo y grupo etario, mediante el coeficiente de correlación de Pearson. Un p-valor inferior a 0,05 fue considerado criterio de

significancia estadística. Las pruebas estadísticas fueron realizadas con el software estadístico Stata 17 (18).

RESULTADOS

A partir de las publicaciones recolectadas de *Twitter* Argentina (n=6023548) se observó que los alimentos más

frecuentemente mencionados fueron huevo, mate, cerveza, leche, vino y pasta simple, entre otros (Figura 1A). Al agrupar los alimentos, los grupos más mencionados fueron "azúcares refinados y confituras", "bebidas con alcohol", "carnes y huevos" y "preparaciones con harinas" (Figura 1B).

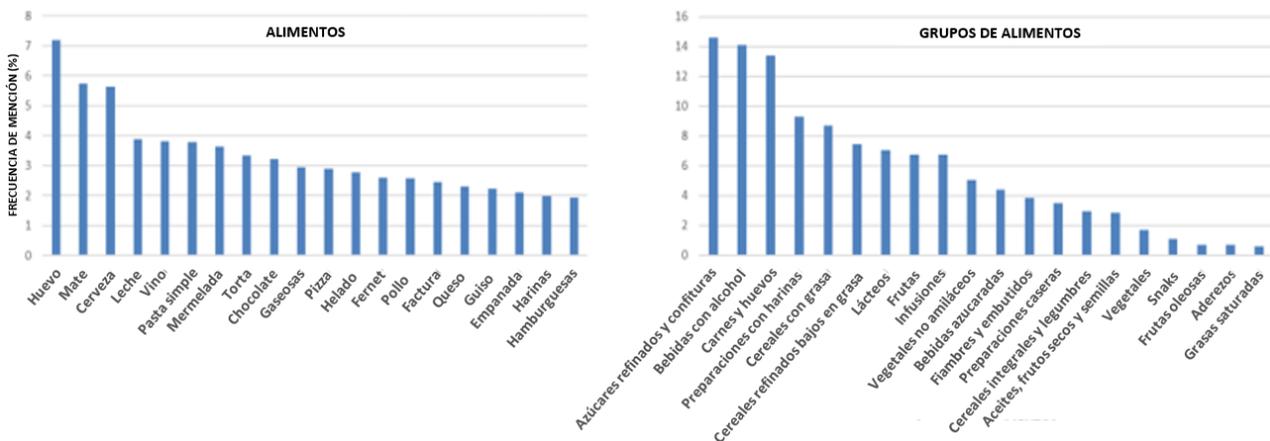


Figura 1. Frecuencia de mención de los 20 alimentos y bebidas (A) y grupos de alimentos y bebidas (B) más nombrados en Twitter Argentina durante el periodo 2021-2022.

Además, se observó que la distribución de la prevalencia de obesidad, la tasa de *tweets* alimentarios y el IDN promedio por *tweet*, presentaron una distribución espacial no aleatoria ($p < 0,05$) (Figura 2A, B y C). Las provincias de Santa Cruz y San Juan (Figura 2A) fueron las que presentaron una mayor prevalencia de obesidad en comparación con el resto de las provincias, mientras que Chaco se destacó por su baja prevalencia. Las provincias que presentaron las tasas de *tweets* alimentarios más elevadas fueron

Tierra del Fuego, Río Negro, Neuquén, La Pampa, San Luis, Entre Ríos y Misiones (Figura 2B). Es decir, en estas provincias los/as usuarios/as generaron mayor cantidad de publicaciones que mencionan alimentos en relación al total de *tweets*. Por otro lado, existieron regiones que presentaron baja tasa de *tweets* alimentarios, como las provincias del noroeste (Jujuy, Salta, Tucumán, Catamarca), las de la región cuyana (La Rioja, San Juan, Mendoza), de la región central (Buenos Aires y Santa Fe) y las que

conforman la región patagónica (Chubut y Santa Cruz). El IDN promedio de los *tweets* por provincias en Argentina mostró valores más elevados en las provincias de

La Rioja y Catamarca seguido por Santa Cruz, Tierra del Fuego, San Juan, Santiago del Estero, Salta, Chaco, Formosa y Corrientes (Figura 2C).

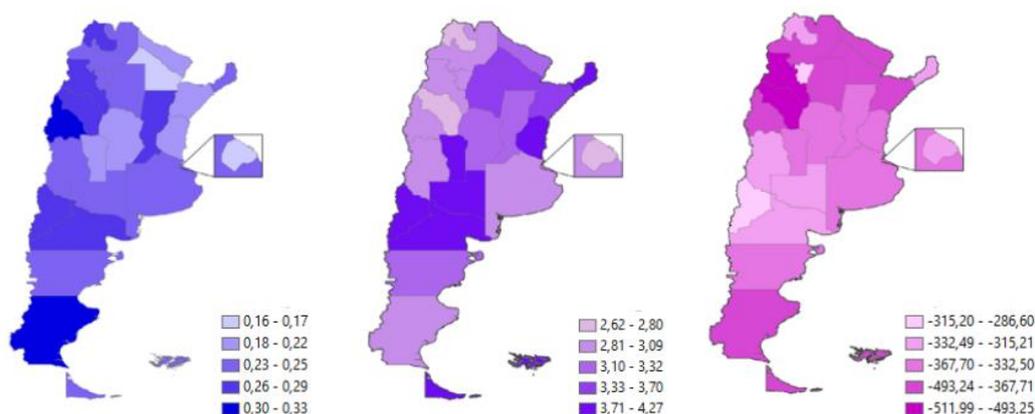


Figura 2. Prevalencia de obesidad en el año 2018 según Encuesta Nacional de Factores de riesgo (A), tasa de *tweets* alimentarios (B) y índice de densidad de nutrientes (C), construidos a partir de publicaciones de la red social Twitter durante el periodo 2021–2022, por provincia en Argentina.

No existió correlación significativa entre la prevalencia de obesidad total y la tasa de *tweets* alimentarios ($p=0,239$). Sin embargo, al relacionar la prevalencia de obesidad con la frecuencia de mención de alimentos seleccionados, de amplio consumo en Argentina, se observó que los términos almendra, melón, avena y molleja se correlacionaron de manera positiva ($p<0,05$). En cambio, las correlaciones fueron inversas con los términos pepino, pomelo, naranja, alfajor, atún, caldos, pan blanco y dulce de leche ($p<0,05$) (Tabla 1).

Además, se observó que a mayor prevalencia de obesidad en varones es mayor la frecuencia de mención de los

alimentos almendra, semillas, queso semiduro, pescado y margarina ($p<0,05$), y es menor la frecuencia de mención de pepino, pomelo, naranja, aguas saborizadas, criollos y papas fritas ($p<0,05$) (Tabla 1). En mujeres la obesidad se correlacionó positivamente con la frecuencia de mención de 'hamburguesas, tartas, molleja, avena y golosinas ($p<0,05$); y de manera inversa con carne de vaca, pepino, alfajor, caldos, queso duro, atún, salchichas, pan blanco, y dulce de leche ($p<0,05$) (Tabla 1).

Tabla 1. Coeficiente de correlación de Pearson entre los alimentos mencionados en la red social Twitter durante el periodo 2021–2022 y la prevalencia de obesidad a nivel provincial en la población argentina en general y según sexo y grupo etario.

Término	Total de la población	Varones	Mujeres	Menores de 40 años	Mayores de 40 años
Almendra	0,4904*	0,6225	0,1617	0,5502*	0,4039*
Pepino	-0,6638**	-0,4537*	-0,6620**	-0,6627**	0,6232**
Pomelo	-0,4494*	-0,4252*	-0,3193	-0,4940*	-0,3143
Naranja	-0,4457*	-0,4690*	-0,2512	-0,4314	-0,3983*
Alfajor	-0,4678*	-0,2901	-0,5086*	-0,2772	-0,5517*
Atún	-0,6168*	-0,3629	-0,6849	-0,6137*	-0,5582*
Caldos	-0,4315*	-0,2648	-0,4690*	-0,4768*	-0,3497
Melón	0,4101*	0,3090	0,3883	0,2994	0,3962*
Champiñón	-0,4811*	-0,3222	-0,4935*	-0,5303*	-0,3711
Avena	0,4000*	0,2082	0,4749*	0,3682	0,4004*
Pan blanco	-0,4608*	-0,2239	-0,5392	-0,4016*	-0,5340*
Molleja	0,4710*	0,2865	0,4942*	0,3890	0,4485*
Dulce de leche	-0,4512*	-0,2278	-0,5189	-0,3867	-0,5361*
Semillas	0,3235	0,5129**	-0,0027	0,5495*	0,1137
Queso semi-duro	0,3507	0,4794*	0,0706	0,4320*	0,2545
Aguas saborizadas	-0,3107	-0,5109*	0,0209	-0,3128	-0,2336
Pescado	0,1762	0,4213*	-0,1577	0,1616	0,1286
Criollos	-0,2902	-0,4870*	0,0269	-0,2909	-0,1749
Papas fritas	-0,2674	-0,4579*	0,0371	-0,4373*	-0,1031
Frutilla	-0,2460	-0,4214*	0,0449	-0,3663	-0,1745
Margarina	0,3325	0,4055*	0,1237	0,3712	0,2906
Hamburguesas	0,1978	-0,0009	0,3461*	0,1361	0,3023
Tartas	0,2494	0,0282	0,4165*	0,1939	0,2564

Golosinas	0,3054	0,0797	0,4541*	0,3163	0,2087
Carne de vaca	-0,1248	0,2405	-0,4935*	0,0862	-0,2895
Alcaucil	-0,3741	-0,0973	-0,5268*	-0,5268*	-0,4358*
Chocolate	0,2348	0,2690	0,1214	0,4256*	0,0310
Legumbres	0,3865	0,2851	0,3606	0,4149*	0,3266
Apio	0,3736	0,3826	0,2291	0,4725*	0,2489
Brócoli	0,1224	0,2827	-0,0902	0,4002*	-0,1475
Morcilla	0,3465	0,3538	0,2143	0,4506*	0,2436
Frutos secos	-0,3698	-0,1056	-0,2423	-0,3728	-0,4055*
Achicoria	-0,3949	0,1722	-0,3655	-0,3297	-0,4869*

* $p < 0,05$ ** $p < 0,001$

Al considerar el grupo etario, en personas menores de 40 años la prevalencia de obesidad se correlacionó positivamente con la frecuencia de mención de almendra, semillas, chocolate, legumbres, apio, queso semiduro, brócoli y morcilla ($p < 0,05$). En tanto, la frecuencia de mención de pepino, pomelo, naranja, caldos, atún, papas fritas y pan blanco se correlacionó de manera inversa ($p < 0,05$). En personas mayores a 40 años la carga de obesidad correlacionó positivamente con la frecuencia de mención de almendra, melón, molleja y avena ($p < 0,05$), y de manera inversa con pepino, naranja, alfajor, frutos secos, achicoria, pan blanco, dulce de leche y atún ($p < 0,05$).

Se evaluó también la frecuencia de mención de alimentos agrupados en relación a la prevalencia de obesidad. Los

resultados mostraron una correlación positiva entre el grupo de aceites y la carga de obesidad ($p = 0,04$), en tanto que el grupo de frutas presentó una correlación inversa ($p = 0,05$). Los demás grupos de alimentos no mostraron correlaciones significativas con la prevalencia de obesidad. El IDN por provincia se correlacionó positivamente con la tasa de *tweets* alimentarios ($p = 0,04$). Sin embargo, en este trabajo no hubo correlación significativa entre el IDN y la prevalencia de obesidad ($p = 0,35$).

DISCUSIÓN

Actualmente, vivimos en una sociedad hiperconectada, donde la forma de comunicarnos ha evolucionado y las redes sociales han adquirido un rol fundamental.

La vasta cantidad y disponibilidad de datos generados por las plataformas virtuales en los últimos años ha impulsado un rápido incremento en la aplicación de técnicas basadas en datos y el análisis de poblaciones a gran escala (16).

La prevalencia de obesidad en la población adulta de nuestro país ha aumentado considerablemente, pasando del 14,6% en 2005 al 25,4% en 2018 de acuerdo con datos provenientes de las ENFR (2). El rápido y generalizado aumento del exceso de peso a nivel mundial ha sido definido como una auténtica epidemia global. En este estudio se destacan diferencias en las tendencias y niveles de prevalencia de obesidad entre las provincias de Argentina, posiblemente vinculadas a procesos transicionales –cambios epidemiológicos, demográficos y nutricionales ocurridos en las últimas décadas– que no se han manifestado de manera uniforme en todo el territorio, como ya fue reportado (19–20). Además, en este trabajo se observó que la prevalencia de obesidad en varones fue mayor en las provincias de San Juan, La Rioja y Santa Cruz. En cuanto a las mujeres esta prevalencia fue mayor en San Juan, seguido de San Luis, Corrientes, Santa Fe, Catamarca, Jujuy y provincias de la región patagónica.

Gore y cols. (16) investigaron la frecuencia con la que las personas mencionan

alimentos en sus *tweets* y observaron que generalmente disminuye a medida que aumenta la tasa de obesidad en un área geográfica. En nuestro estudio, la tasa de *tweets* relacionados con la alimentación no mostró una correlación significativa con la prevalencia de obesidad a nivel provincial. Sin embargo, el coeficiente de correlación de Pearson fue negativo, lo que indica una tendencia de correlación inversa, similar a la observada en el estudio mencionado. Esto podría deberse a que existen múltiples factores que influyen en el comportamiento de las personas en redes sociales al momento de publicar sobre alimentación. Entre ellos, podemos destacar el interés personal: las personas suelen opinar sobre temas que les resultan de interés, y quienes prestan mayor atención a su alimentación probablemente hablen más al respecto. Además, las redes sociales son ampliamente utilizadas por profesionales de la salud para difundir contenido informativo y conocimientos relacionados con hábitos alimentarios saludables. Por último, se debe considerar el contexto social en el que se encuentra inserta cada persona, los gustos y preferencias, como así también comentarios negativos que pueden llevar a que las personas no quieran compartir públicamente sus pensamientos (21). Varios estudios (22, 23) demostraron que la obesidad es una enfermedad

multifactorial y compleja que afecta a personas que están insertas en un contexto socioeconómico, político, cultural, demográfico y epidemiológico particular, y puede generar diferentes maneras de relacionarse con la comida. Muchas veces, estas personas suelen tener sentimientos de culpa y vergüenza, cuadros de depresión, ansiedad, baja autoestima, lo que podría explicar por qué, a veces, las personas que presentan obesidad evitan publicar en sus redes sociales *tweets* referidos a la comida (24). En cuanto a la frecuencia de mención de diversos alimentos, se evidenció que los más nombrados en los *tweets* en Argentina fueron huevo, mate, cerveza, leche, vino, entre otros. En cuanto a los grupos de alimentos, los más frecuentes fueron azúcares refinados y confituras, bebidas con alcohol, carnes y huevos, preparaciones con harinas y cereales con grasa. Estos resultados coinciden con el perfil alimentario tradicional en Argentina, caracterizado por un alto consumo de proteínas y grasas animales, obtenidas principalmente de las carnes rojas, y una baja ingesta de pescado, frutas y verduras (25, 26). Al analizar los alimentos obtenidos de los *tweets* y cómo se correlacionan con la prevalencia de obesidad, observamos que muchos de los alimentos que son considerados saludables (27) como

pepino, pomelo, naranja, champiñón, alcaucil, atún, carne de vaca, frutilla se correlacionaron de manera inversa con la prevalencia de obesidad. Por otro lado, la mención de alimentos de bajo valor nutricional, caracterizados por un alto contenido de azúcares simples, grasas saturadas, sodio, entre otros, como molleja, queso semiduro, chocolate, morcilla, hamburguesas y golosinas, se correlacionó positivamente con la prevalencia de obesidad. Estos resultados son consistentes con la evidencia científica que establece una relación entre el consumo de alimentos considerados poco saludables y la obesidad (28). Sin embargo, algunos resultados no coincidieron con lo reportado por la literatura respecto a los hábitos alimentarios asociados al desarrollo de la obesidad. La mención de alimentos considerados saludables, como almendras, semillas, pescado, avena y melón, mostraron una correlación positiva con la prevalencia de obesidad. En contraste, ciertos alimentos catalogados como poco saludables (27), tales como caldos, alfajores, dulce de leche, aguas saborizadas, criollos y papas fritas, se correlacionaron negativamente. Es importante destacar que, aunque en este estudio se observó la frecuencia de mención de alimentos, no se pudo determinar el contexto en el que fueron

mencionados, es decir, si fueron referidos de manera positiva o negativa en el *tweet*. Además, muchos de los alimentos considerados saludables presentaron una baja frecuencia de mención en los *tweets*, lo que podría limitar su representatividad en relación con la obesidad.

Al analizar el IDN promedio y su relación con la prevalencia de obesidad, no se observó una correlación significativa, al igual que en el estudio de Gore y cols. (16). Sin embargo, en este trabajo, en provincias como Neuquén, donde la prevalencia de obesidad fue elevada, el IDN registrado fue el más bajo, observándose un patrón similar en Río Negro y Jujuy. Es importante señalar que no se consideraron la cantidad ni la preparación de los alimentos consumidos, ya que solo se analizó la mención del alimento. Esto podría explicar la ausencia de una relación significativa entre la densidad de nutrientes de los alimentos mencionados en los *tweets* y la prevalencia de obesidad. Resulta relevante subrayar la importancia de estos factores, dado que tanto la cantidad como la forma de preparación (ingredientes, aditivos, etc.) influyen en el IDN final del alimento (16). Algunas limitaciones se debieron considerar en el presente trabajo. Para la recolección de datos, se utilizó la API de *Twitter*, Tweepy, que impone restricciones en la cantidad de *tweets* que se pueden

recolectar cada 15 minutos. No obstante, esta limitación se abordó aumentando el período de recolección de datos, lo que permitió obtener un mayor volumen de información. Por otro lado, solo un pequeño porcentaje de los *tweets* en *Twitter* están etiquetados geográficamente, lo que restringe el uso de todos los *tweets* en el análisis. Además, algunos términos utilizados en los *tweets* resultaron ser ambiguos, lo que llevó a su descarte. Como consecuencia, parte de la información se perdió durante el proceso de clasificación de los *tweets*. Esto se pudo resolver en parte utilizando artículos o preposiciones en la búsqueda (por ejemplo, la palabra “vino” puede ser usada como bebida y como verbo, entonces se buscó en conjunto con un artículo “de vino” o “con vino”, usos que solo se darían en un contexto de mención del vino como bebida). Es de importancia mencionar que la mayoría de los usuarios de *Twitter* tienen edades comprendidas entre 18 y 49 años, los que representan más del 60% (29). A su vez, si bien el acceso a internet en la población continúa ligado a cierto nivel socioeconómico (30), en varios estudios se ha utilizado como una aproximación del discurso de la población total (16, 31).

Sumado a esto, es importante destacar que ningún alimento o grupo de alimento por sí solo puede asociarse directamente

con la obesidad, el escenario ideal para su análisis es el de patrones de consumo alimentario. Cabe mencionar, además, que este trabajo no aborda el consumo alimentario poblacional, dado que la mera mención de un alimento no implica necesariamente su consumo por parte de quien lo menciona, pero ciertamente puede ser indicativo de prácticas o percepciones en torno a la obesidad que aún requieren ser explicadas.

Finalmente, como fortaleza de este trabajo se destaca que, pese a sus limitaciones, el mismo representa una primera aproximación valiosa del uso de fuentes de información no convencionales para el estudio de la obesidad, empleando una metodología que no presenta antecedentes en la región. El análisis de redes sociales puede potencialmente revelar hallazgos a los que no se puede acceder mediante métodos comúnmente utilizados en la investigación en nutrición y salud, permitiendo obtener información en función de atributos a nivel individual, y/o subgrupos dentro de las redes. Así, los resultados derivados del análisis de redes sociales pueden responder una pregunta de investigación o incorporarse como variables en modelos estadísticos sólidos. En base a estos resultados preliminares, se espera optimizar en adelante el proceso de

construcción de indicadores alimentarios novedosos basados en IA, los que servirán de insumo para profundizar el estudio de los determinantes de la obesidad desde perspectivas poco exploradas.

CONCLUSIONES

La creciente obesidad en Argentina refleja una compleja interrelación entre factores sociales, económicos y culturales que influyen en los hábitos alimentarios. Este estudio resalta la importancia de las redes sociales como fuente de datos para explorar las tendencias alimentarias y su conexión con la obesidad, aunque se evidencian limitaciones en la representatividad y contexto de los *tweets* analizados. La discrepancia entre los resultados sobre la mención de alimentos saludables y no saludables con respecto a la prevalencia de obesidad sugiere cierto nivel de concordancia con resultados obtenidos a partir de estudios de investigación con metodología tradicional. A medida que la tecnología y el acceso a la información continúan evolucionando, se abre una puerta para la aplicación de técnicas innovadoras que podrían enriquecer la investigación en salud pública y la formulación de políticas más efectivas.

BIBLIOGRAFÍA

1. Organización Mundial de la Salud. Obesidad y sobrepeso. 2021. Recuperado de: <https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight>
2. Instituto Nacional de Estadísticas y Censos, Ministerio de Salud de la Nación. 4° Encuesta Nacional de Factores de Riesgo para enfermedades no transmisibles [Internet]. Argentina; 2019 [citado julio, 2024]. Disponible en: https://www.indec.gov.ar/ftp/cuadros/publicaciones/enfr_2018_resultados_definitivos.pdf
3. Aballay LR, Coquet JB, Scruzzi GF, Haluszka E, Franchini G, Carreño P, Raboy E, Román MD, Niclis C, Balangero M, Altamirano N, Barbás MG, López L. Estudio de base poblacional de seroprevalencia y factores asociados a la infección por SARS-CoV-2 en Córdoba, Argentina. *Cadernos de Saude Publica*. 2022;38(4). doi:10.1590/0102-311XES219821.
4. Pou SA, Niclis C, Díaz MP. Socio-territorial configuration of mortality by chronic diseases of greater prevalence. En: *Inequities and Quality of Life in Argentina, Geography and Quality of Life in Argentina*. Cham, Suiza: Springer; 2022. p. 311-353. doi:10.1007/978-3-030-94411-7.
5. Niclis C, Coquet JB, Aballay L. Socio-spatial differences in the prevalence of lifestyle risk factors related to chronic diseases in Argentina from 2005 to 2013. En: *Inequities and Quality of Life in Argentina, Geography and Quality of Life in Argentina*. Cham, Switzerland: Springer; 2022. p. 227-79. doi:10.1007/978-3-030-94411-7.
6. Braguinsky J. Concepto, definición y diagnóstico. En: *Obesidad saberes y conflictos, un tratado de obesidad*. Argentina: Editorial Medica AWWE;2007.
7. Organización Mundial de la Salud, Organización de las Naciones Unidas para la Alimentación y la Agricultura, Fondo de las Naciones Unidas para la Infancia (UNICEF). *Guidance for monitoring healthy diets globally*. Ginebra: Organización Mundial de la Salud; 2024.
8. Secretaría de Gobierno de Salud. 2° Encuesta Nacional de Nutrición y Salud ENNYS 2 [Internet]. 2019 [citado 2023 Oct 18]. Disponible en: <https://cesni-biblioteca.org/2-encuesta-nacional-de-nutricion-y-salud-ennys-2-resumen-ejecutivo/>
9. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. *JMIR Public Health Surveill*. 2020;6(4). doi:10.2196/21978.
10. Matar RP, Esponda EP. Red social Twitter: fortalezas y debilidades. *Infodir (Revista de Información para la Dirección en Salud)*. 2016;12(23):77-86.
11. Wanden-Berghe C, Sanz-Valero J. Infodemiología o infodemia: nuevo reto de la salud global.

Hosp Domic. 2021;5(4):179–83.

12. Tweepy. [Internet]. [citado 2024 Oct]. Disponible en: <https://www.tweepy.org/>
13. MongoDB, Inc. MongoDB: the database for giant ideas. [Internet]. [citado 2024 Oct]. Disponible en: <https://www.mongodb.com/>
14. Jupyter Project. Jupyter Notebook. [Internet]. [citado 2024 Oct]. Disponible en: <https://jupyter.org/>
15. Pérez JM, Luque FM. Atalaya at SemEval 2019 Task 5: robust embeddings for tweet classification. En: Proceedings of the 13th International Workshop on Semantic Evaluation; 2019; Minneapolis, Minnesota, USA: Association for Computational Linguistics; 2019. p. 64–9.
16. Gore RJ, Diallo S, Padilla J. You are what you tweet: connecting the geographic variation in America's obesity rate to Twitter content. PLoS One. 2015;10(9). doi:10.1371/journal.pone.0133505.
17. Instituto Nacional de Estadística y Censos. Censo Nacional de Población y Viviendas 2010 [Internet]. [citado 2024 Jul]. Disponible en: <https://www.indec.gov.ar/indec/web/Nivel4-Tema-2-41-135>
18. tataCorp. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC; 2021.
19. Niclis C, Aballay LR, Coquet JB. Factores de riesgo relacionados a enfermedades crónicas. En: Velázquez GA, Celemin JP, editores. Atlas histórico y geográfico de la Argentina: calidad de vida. 1ra ed. Olavarría: Universidad Nacional del Centro de la Provincia de Buenos Aires; 2020.
20. Pou SA, Aballay LR, Huergo J. Sociodemografía de la obesidad. En: Velázquez GA, Celemin JP, editores. Atlas histórico y geográfico de la Argentina. Calidad de vida I. Tandil, Argentina: Universidad Nacional del Centro de la Provincia de Buenos Aires; 2020. p. 403–59.
21. Cava A. La cultura del gusto: la alimentación entre representaciones y conexiones. Rev Castellano–Manchega Cienc Soc. 2019;26:121–33.
22. Godinho S de Souza. La obesidad desde una mirada social: estudio en Argentina y Brasil (2010 a 2012) [Tesis de Maestría en línea]. Buenos Aires: Pontificia Universidad Católica Argentina; 2013. Disponible en: <https://repositorio.uca.edu.ar/handle/123456789/176>
23. Álvarez Mieres N. Alimentación y salud: la obesidad como factor de riesgo. NPunto. 2019;2(17):1–81.
24. Tamayo D, Restrepo M. Aspectos psicológicos de la obesidad en adultos. Rev Psicología Univ Antioquia. 2014;6(1):91–112.
25. Navarro A, Muñoz S, Lantieri M, Díaz MP, Cristaldo P, Fabro S, et al. Meat cooking habits and risk of colorectal cancer in Córdoba, Argentina. Nutrition. 2004;20(10):873–7.

26. Matos E, Brandani A. Review on meat consumption and cancer in South America. *Mutat Res.* 2002;506-507:243-9. doi:10.1016/s0027-5107(02)00171-9.
27. Ministerio de Salud de Argentina. Manual para la aplicación de las guías alimentarias para la población argentina [Internet]. 2018 [citado 2024 Jun]. Disponible en: https://bancos.salud.gob.ar/sites/default/files/2020-08/guias-alimentarias-para-la-poblacion-argentina_manual-de-aplicacion_0.pdf
28. Organización Panamericana de la Salud. Alimentos y bebidas ultraprocesados en América Latina: tendencias, efecto sobre la obesidad e implicaciones para las políticas públicas [Internet]. 2015 [citado 2024 Oct]. Disponible en: https://bancos.salud.gob.ar/sites/default/files/2020-05/alimentos-bebidas-ultraprocesados-america-latina_obesidad_oms.pdf
29. We Are Social y Meltwater. "Digital 2023 Argentina". [Internet] 2023. [citado 5/jul/2023] Disponible en: <https://datareportal.com/reports/digital-2023-argentina>
30. Ollari M. Redes sociales: entre la expansión hacia un uso activo y las restricciones del acceso diferencial. VIII Jornadas de Sociología de la UNLP; 2014; 3-5 de diciembre; Ensenada, Argentina. En: Memoria Académica [Internet]. Disponible en: http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.4790/ev.4790.pdf
31. Tommasel A, Diaz Pace JA, Godoy DL, Rodriguez JM. Tracking the evolution of crisis processes and mental health on social media during the COVID-19 pandemic. *Behav Inf Technol.* 2021;11:1-20. Taylor & Francis Ltd.