

Análisis de sentimientos acerca de la inmunoterapia con alérgenos en *Twitter* mediante un modelo de procesamiento natural de lenguaje

Sentiment analysis on allergen immunotherapy on *Twitter* using a natural language processing model

Alejandro Tarango-García¹, Saul Oswaldo Lugo-Reyes², Aristoteles Alvarez-Cardona³

¹ Universidad Autónoma de Aguascalientes

² Universidad Autónoma de Aguascalientes, Laboratorio de Inmunodeficiencias, Instituto Nacional de Pediatría, Ciudad de México

³ Unidad de Investigación en Inmunología Clínica y Alergia, Universidad Autónoma de Aguascalientes

Fecha de recepción: 27/05/2023

Fecha de aceptación: 19/07/2023

Fecha de publicación: 01/02/2024

*Correspondencia: Saul Oswaldo Lugo Reyes. Email: dr.lugo.reyes@gmail.com; Aristoteles Alvarez Cardona. Email: dr.aristoteles@yahoo.com.mx

Resumen

Objetivo: Analizar los sentimientos acerca de la inmunoterapia alérgeno-específica en *Twitter* mediante el modelo VADER (*Valence Aware Dictionary and sEntiment Reasoner*).

Métodos: Se utilizaron *tweets* relacionados con la inmunoterapia alérgeno-específica obtenidos a través del API (*Application Programming Interface*) de *Twitter*. Se incorporaron las palabras clave "allergy shot" en el período comprendido entre el 1 de enero de 2012 y el 31 de diciembre de 2022. Los datos obtenidos fueron procesados, eliminando las *URL*, nombres de usuarios, *hashtags*, espacios múltiples y *tweets* duplicados. Posteriormente, se realizó un análisis de sentimientos utilizando el modelo VADER.

Resultados: Se recolectaron 34,711 *tweets*, de los que se eliminaron 1928. De los 32,783 *tweets* restantes, se encontró que el 32.41% de los usuarios expresó un sentimiento negativo, el 31.11% un sentimiento neutral y el 36.47% un sentimiento positivo, con una media de polaridad de 0.02751 (neutral) a lo largo de los 11 años.

Conclusiones: La polaridad media de los *tweets* acerca de la inmunoterapia alérgeno-específica es neutral a lo largo de los 11 años analizados. Existe un aumento anual en la polaridad media positiva a lo largo de los años, sobre todo entre 2017, 2018 y 2022. La cantidad de *tweets* disminuyó con el tiempo.

Palabras clave: Sentimiento; Inmunoterapia alérgeno-específica; *Twitter*; Sentimiento positivo; VADER (*Valence Aware Dictionary and sEntiment Reasoner*); API (*Application Programming Interface*).

Abstract

Objective: Analyze sentiments about allergen-specific immunotherapy on *Twitter* using the VADER model VADER (*Valence Aware Dictionary and sEntiment Reasoner*) model.

Methods: *tweets* related to specific allergen immunotherapy were obtained through the *Twitter* Application Programming Interface (API). The keywords "allergy shot" were used between January 1, 2012, and December 31, 2022. The data was processed by removing URLs, usernames, hashtags, multiple spaces, and duplicate *tweets*. Subsequently, a sentiment analysis was performed using the VADER model.

Results: A total of 34,711 *tweets* were retrieved, of which 1928 were eliminated. Of the remaining 32,783 *tweets*, 32.41% expressed a negative sentiment, 31.11% expressed a neutral sentiment, and 36.47% expressed a positive sentiment, with an average polarity of 0.02751 (neutral) over the 11-year period.

Conclusions: The average polarity of *tweets* about allergen-specific immunotherapy is neutral over the 11 years analyzed. There was an annual increase in the average polarity over the years, with 2017, 2018, and 2022 having positive polarity averages. Additionally, the number of *tweets* decreased over time.

Keywords: Sentiments; Allergen-specific immunotherapy; *Twitter*; Positive feeling; VADER (*Valence Aware Dictionary and sEntiment Reasoner*); API (*Application Programming Interface*).

ANTECEDENTES

La rinitis alérgica es la enfermedad alérgica más común en todo el mundo; afecta alrededor del 30% de la población mundial. Quienes la padecen ven afectada su calidad de vida, sufren de morbilidad asociada, y perciben altos costos de atención.^{1,2}

Desde hace más de 100 años se indica la inmunoterapia alérgeno-específica y representa el tratamiento de primera línea modificador de la enfermedad. Sus mecanismos de acción se han descrito en los últimos años y puede administrarse por vía subcutánea o sublingual. Los beneficios de la inmunoterapia incluyen: disminución de los síntomas, menor requerimiento de medicamentos y reducción del riesgo de asma o nuevas sensibilizaciones a futuro. Sin embargo, para un efecto sostenido a largo plazo se requieren, mínimo, tres años de tratamiento por vía subcutánea o sublingual.³

A pesar de su eficacia y seguridad demostradas, el apego al tratamiento con inmunoterapia alérgeno-específica es un problema por resolver y que limita su eficacia. En general, se reconoce que el apego al tratamiento en pacientes con padecimientos crónicos es deficiente (alrededor del 50%).⁴ En diferentes estudios se han encontrado rangos de apego a la inmunoterapia alérgeno-específica entre un 20% y 90%; en una revisión reciente se comentaron varios factores asociados con el apego al tratamiento: edad, médico tratante, eficacia clínica, no conveniencia del tratamiento (duración y efectos adversos), educación del paciente y aspectos económicos (costos, reembolsos y sistema de salud); definiendo diferentes puntos de oportunidad para mejorar el apego, que incluyen: optimizar la educación del paciente acerca del tratamiento y la enfermedad, mejorar la relación médico paciente, contar con mejores datos, estudios de la vida real y estandarizar los resultados de la terapia.⁵

En tiempos recientes, las redes sociales se han vuelto una plataforma de expresión a nivel global para las instituciones y los usuarios, donde se vierten opiniones personales y sentimientos de diversos temas.

En la actualidad, los pacientes a quienes se aplica la vacunación tradicional están sufriendo por miedo, indecisión, incluso por oposición de ciertos sectores de la población, y el conocimiento médico de baja calidad supone un factor importante en este cambio, que está fuertemente influenciado por las redes sociales.⁶

Las redes sociales se han convertido en un elemento importante en la investigación de opiniones, al permitir a los usuarios tener discusiones públicas en tiempo real y desde cualquier parte del mundo.⁷ Como resultado de las redes sociales se generan grandes cantidades de información denominada macrodatos. Una forma de interpretar estos datos, que ha comenzado a ganar tracción, es el análisis de sentimientos, un tipo de minería de textos que corresponde al procesamiento natural de lenguaje, una rama de la inteligencia artificial.

El análisis de sentimientos clasifica cualitativa (categoría) y cuantitativamente (intensidad, numérica) los sentimientos de un texto en: positivo, negativo o neutral dependiendo de su polaridad; esta polaridad es un cálculo que se forma a partir del procesamiento del texto con técnicas de evaluación de

lenguaje natural.⁸ En un abordaje lexicológico, el algoritmo (la máquina) que analiza un texto específico, categoriza y califica cada palabra, mapeándola en el diccionario a un sentimiento (emoción, actitud), para decidir la categoría y calcular la intensidad del enunciado completo. Además de las palabras, el modelo considera elementos como: *emojis*, uso de mayúsculas y signos de exclamación.

El objetivo de este estudio fue: Analizar los sentimientos acerca de la inmunoterapia alérgeno-específica en *Twitter* mediante el modelo VADER.

MÉTODOS

La información se recolectó mediante la API (Interfaz de Programación de Aplicaciones) de *Twitter*, que se programó utilizando el lenguaje de programación *Python*.

A través de la biblioteca Tweepy se extrajeron *tweets* en idioma inglés relacionados con inmunoterapia con alérgenos, empleando la palabra clave de búsqueda "allergy shot" sin restricción geográfica, desde el 1 de enero del 2012 al 31 de diciembre del 2022.

Se eliminaron las URL, nombres de los usuarios, *hashtags* y espacios múltiples. Los datos fueron organizados con la biblioteca Pandas. Los *tweets* duplicados se retiraron, y se excluyeron de la búsqueda los *retweets*.

Para el análisis de sentimientos se utilizó un enfoque de aprendizaje automático no supervisado con el modelo VADER: una herramienta de análisis de sentimientos mediante procesamiento de lenguaje natural (NLP) basado en el léxico y reglas simples, entrenado previamente en textos de redes sociales.

El modelo VADER⁹ analizó el sentimiento detrás de los *tweets*, incluidos los *emojis* y palabras, y generó un puntaje de polaridad de -1 a 1. Para fines del presente estudio se consideró la clasificación expuesta en el **Cuadro 1**.

Para el análisis estadístico se describieron medidas de tendencia central, y se utilizó la prueba de normalidad de Agostino y Pearson, además de la de Kruskal-Wallis para evaluar la diferencia entre las medias de las polaridades estudiadas, mediante el programa GraphPad (Prism, San Diego).

RESULTADOS

Se extrajeron un total de 34,711 *tweets*; de estos se eliminaron 1928 por ser duplicados, para finalmente obtener 32,783 *tweets* publicados entre el 1 de diciembre del 2012 y el 31 de diciembre del 2022.

De los 32,783 *tweets*, 32.41% (10,628) mostraron un sentimiento negativo, 31.11% (10,199) neutral y 36.47% (11,956)

Cuadro 1. Clasificación de acuerdo con el puntaje de polaridad (compound) del modelo VADER.

Positivo	Mayor de 0.05
Neutral	Menor de 0.05 y mayor de -0.05
Negativo	Menor de -0.05

Cuadro 2. Porcentaje de tweets positivos, negativos y neutrales por año.

Año	Negativo	Neutral	Positivo
2012	30.70%	37.94%	31.35%
2013	31.71%	34.66%	33.62%
2014	31.40%	33.79%	34.80%
2015	30.20%	35.29%	34.50%
2016	32.74%	37.68%	29.56%
2017	30.46%	27.68%	41.85%
2018	34.79%	23.40%	41.79%
2019	34.66%	23.64%	41.69%
2020	35.08%	23.33%	41.57%
2021	34.87%	22.96%	42.15%
2022	34.02%	23.91%	42.05%

positivo, para reportar una media de polaridad de 0.02751 (neutral) a lo largo de 11 años. **Cuadro 2**

El mes con mayor cantidad de tweets fue febrero de 2012 con 825 y el de menor cantidad diciembre de 2017 con 66 tweets. La media de polaridad más alta se registró en junio de 2017 con 0.21079 y la más baja en abril de 2016 con -0.10274. **Figura 1**

En el análisis por año se encontró que el 2018 tuvo la media de polaridad más alta con 0.06397 y 2016 la más baja con -0.01172. El año 2012 reportó mayor cantidad de tweets (6136). Los años 2017, 2018 y 2022 tuvieron medias de polaridad positivas, y el resto informaron sentimiento general neutral.

En el análisis estadístico se encontró una diferencia significativa entre las medianas de polaridad durante los 11 años analizados ($p < 0.0001$) mediante la prueba de Kruskal-Wallis. **Figura 2**

DISCUSIÓN

El objetivo de este estudio fue analizar los sentimientos de los usuarios de *Twitter* acerca de la inmunoterapia con alérgenos mediante el modelo VADER. Se observó que a lo largo de 11 años la polaridad se mantuvo neutral; sin embargo, al analizar los datos por año se demostró que la polaridad fue desplazándose hacia la positividad, especialmente en los últimos años, con un predominio de sentimientos positivos durante 2017, 2018 y 2022; además de una disminución en la cantidad de tweets acerca del

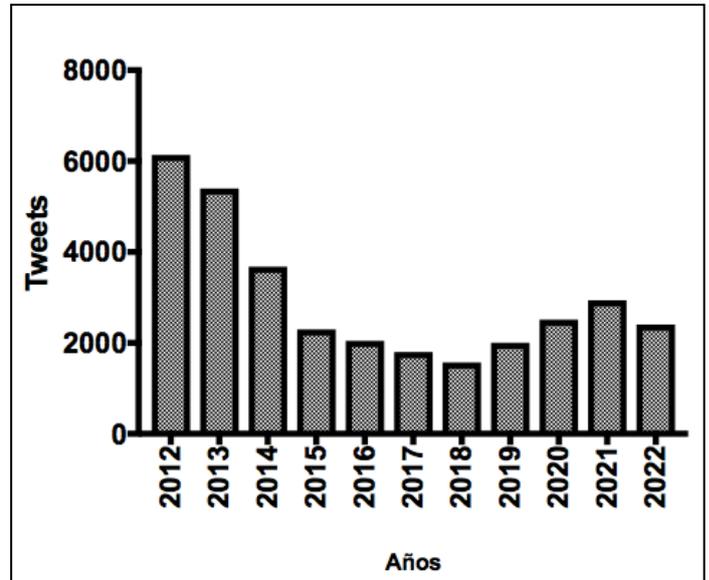


Figura 1. Cantidad de tweets por año (periodo 2012-2022).

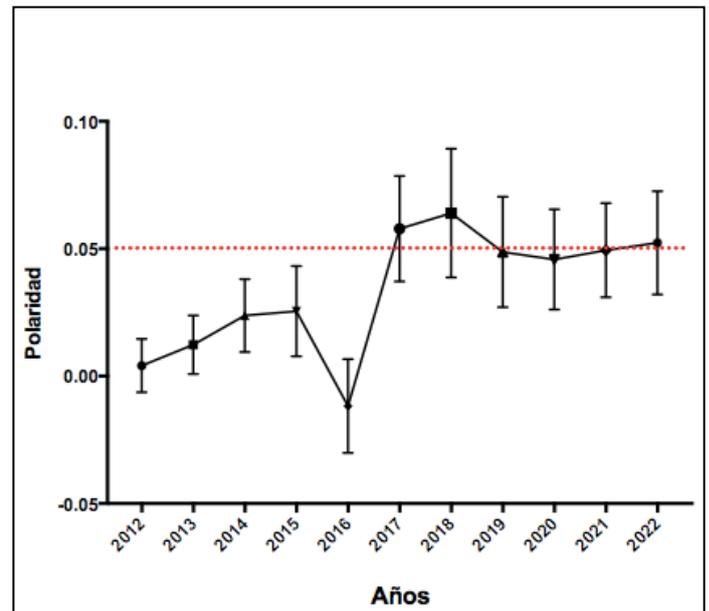


Figura 2. Polaridad graficada con la media por año, con intervalo de confianza del 95%. La polaridad neutral se consideró entre -0.049 a +0.049. La línea punteada representa el límite inferior de la polaridad positiva (favorable para la inmunoterapia alérgeno-específica).

tema. Es decir, las publicaciones de inmunoterapia específica han ido disminuyendo, pero los sentimientos han sido más favorables, quizá debido a las diferentes opciones de posología y vías de administración disponibles en la actualidad.^{3,10}

En los últimos años se han realizado diversos análisis de sentimientos acerca de vacunas contra COVID-19, varios de estos mediante el modelo VADER.^{11,12,13} En el estudio de Liu y Liu reportó que en 92 días, 42.8% de los tweets mostraron un sentimiento positivo, 26.9% fueron neutrales y 30.3% negativos. Dicha investigación, que utilizó los mismos límites de polaridad, incluyó además un análisis temporal y geográfico.¹¹ El estudio de Monselise y colaboradores, en el que analizaron los sentimientos relacionados con la vacuna de COVID-19, asociado con un modelado (*topic modeling*) para

identificar los temas más discutidos; además de utilizar el modelo VADER para clasificar los *tweets*, los autores utilizaron el modelo BERT de minería de textos para clasificar los *tweets* según la emoción que producían, y encontraron que el 46.9% de los *tweets* fueron negativos, casi todos relacionados con miedo.¹²

Las opiniones y actitudes de pacientes y usuarios suelen obtenerse de encuestas y cuestionarios. Minar las redes sociales permite acceder a opiniones y creencias más sinceras y orgánicas, además de incluir a personas renuentes a llenar formatos. A pesar de ser tratamientos seguros y eficaces, la inmunoterapia, las vacunas y otras intervenciones pueden ser controvertidas, afectadas por infodemia, o perder el apego del paciente. El análisis de sentimiento es un abordaje no intrusivo que permite interrogar y evaluar la aceptación o rechazo de dichos tratamientos. Los médicos *influencers* y gobiernos con actividad en redes pueden contribuir a combatir prejuicios e inclinar la balanza en el debate público.

Hasta el momento no existe otro estudio que analice los sentimientos acerca de la inmunoterapia con alérgenos. En el área de inmunología y alergia, este abordaje se ha empleado para investigar actitudes de los pacientes hacia la dermatitis atópica y el asma. Nuestro estudio no incluyó el análisis geográfico, ni se centró en alguna área geográfica específica. La búsqueda de datos se llevó a cabo en idioma inglés, debido a la mayor disposición de información relacionada con el análisis de sentimiento. En futuras investigaciones se pretende efectuar un análisis geográfico en idioma español, con la intención de conocer el sentimiento del uso de inmunoterapia con alérgenos en países hispanohablantes, y su cambio a lo largo del tiempo en diversos lugares.

CONCLUSIÓN

La media de polaridad acerca de la inmunoterapia con alérgenos en 11 años es neutral. Sin embargo, a lo largo de los años se observó un incremento en la polaridad hasta el año 2022, cuando la media de polaridad reflejó un sentimiento positivo. Aunque el sentimiento sea más positivo, la disminución de la cantidad de *tweets*, con el paso de los años, puede disminuir la discusión acerca de la inmunoterapia con alérgenos. Por todo esto, se recomienda a futuro mayor difusión de los efectos benéficos de la inmunoterapia con alérgenos y emprender más estudios enfocados en el análisis de sentimiento. Si bien el apego al tratamiento se asocia con el éxito de la inmunoterapia, es importante insistir en la educación de los pacientes, identificar las tendencias y oportunidades de comunicación, y promover la discusión y aceptación de tratamientos efectivos, para de esta forma tornarse benéfico y oportuno.

Agradecimientos

A cada familia de los autores, a los Dres. Alejandro Tarango Cordova, Alyn María García Chi, Alexa Tarango García y Alfredo García Chi.

Conflicto de interés

Los autores declaran no tener conflicto de interés.

Financiamiento

El presente estudio se ha autofinanciado.

REFERENCIAS

1. Brożek JL, Bousquet J, Baena-Cagnani CE, Bonini S, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines: 2010 revision. *J Allergy Clin Immunol.* 2010; 126 (3): 466-476. DOI: 10.1016/j.jaci.2010.06.047.
2. Brożek JL, Bousquet J, Agache I, Agarwal A, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines-2016 revision. *J Allergy Clin Immunol.* 2017; 140 (4): 950-958. DOI: 10.1016/j.jaci.2017.03.050.
3. Canonica GW, Cox L, Pawankar R, Baena-Cagnani CE, et al. Sublingual immunotherapy: World Allergy Organization position paper 2013 update. *World Allergy Organ J.* 2014; 7 (1): 6. DOI: 10.1186/1939-4551-7-6.
4. Egert-Schmidt A, Kolbe J, Mussler S, Thum-Oltmer S. Patients' compliance with different administration routes for allergen immunotherapy in Germany. *Patient Prefer Adherence.* 2014; 8: 1475-1481. DOI: 10.2147/PPA.S70326.
5. Gehrt F, Xu Q, Baiardini I, Canonica GW, et al. Adherence in allergen immunotherapy: Current situation and future implications. *Allergol Select.* 2022; 6: 276-284. DOI: 10.5414/ALX02318E.
6. Lyu JC, Han EL, Luli GK. COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis. *J Med Internet Res.* 2021; 23 (6): e24435. DOI: 10.2196/24435.
7. Alamoodi AH, Zaidan BB, Al-Masawa M, Taresh SM, et al. Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. *Comput Biol Med.* 2021; 139: 104957. DOI: 10.1016/j.combiomed.2021.104957.
8. Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. *Soc Netw Anal Min.* 2021; 11 (1): 81. DOI: 10.1007/s13278-021-00776-6.
9. Hutto CJ, Gilbert E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media.* 2014; 8 (1): 216-225. DOI: <https://doi.org/10.1609/icwsm.v8i1.14550>
10. Hoyte FC, Nelson HS. Recent advances in allergic rhinitis. *F1000Res.* 2018; 7: 1333. DOI: 0.12688/f1000research.15367.1.
11. Liu S, Liu J. Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis. *Vaccine.* 2021; 39: 5499-5505. DOI: 10.1016/j.vaccine.2021.08.058.
12. Monselise M, Chang CH, Ferreira G, Yang R, et al. Topics and sentiments of public concerns regarding COVID-19 vaccines: social media trend analysis. *J Med Internet Res.* 2021; 23 (10): e30765. DOI: 10.2196/30765.
13. Liew TM, Lee CS. Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672,133 Twitter Posts. *JMIR Public Health and Surveillance.* 2021; 7 (11): e29789. DOI: 10.2196/29789.