

Revisión del artículo: Análisis de patrones espacio-temporales sobre la evaluación de la expansión de infecciones humanas con el virus de influenza aviar A(H7N9) en China 2013-2014

Wen Dong, Kun Yang, Quanli Xu, Lin Liu & Juan Chen

3 de octubre, 2019

Revisión hecha por: Pedro Montenegro & Pablo Vivas

Dong, W., Yang, K., Xu, Q., Liu, L., Chen, J. (2017). Spatio-temporal pattern analysis for evaluation of the spread of human infections with avian influenza A(H7N9) virus in China, 2013–2014. *BMC Infectious Diseases*, 17(1). <https://doi.org/10.1186/s12879-017-2781-2>.

Sobre la revista: *BMC Infectious Diseases* es una revista de acceso abierto que publica artículos de investigación originales revisados por pares en todos los aspectos de la prevención, el diagnóstico y el tratamiento de enfermedades infecciosas y de transmisión sexual en humanos, así como la genética molecular, la fisiopatología y la epidemiología relacionadas. Esta revista está indexada, rastreado y cubierta por PubMed, MEDLINE, CAS, EMBASE, Scopus, Thomson Reuters (ISI) y Google Scholar.

1. Introducción

Los autores del artículo realizan un análisis espacio-temporal sobre la propagación del virus aviar A(H7N9) en China a través de dos técnicas de estadística espacial: elipse de desviación estándar y agrupaciones espacio-temporales. Se plantea hacer la caracterización de la propagación de esta enfermedad, pues en la provincia de Zhejiang se registra la primera persona infectada con el virus en el país. Esta situación reviste de importancia debido a que se da una propagación en lugares aledaños; todo esto ocurre en un lapso menor a 2 años y alcanza el estatus de epidemia.

Según diversos estudios, el virus A(H7N9) se considera como la sucesión o mutación de virus antecedentes como A(H7N3) y A(H9N2), estos tuvieron sus inicios con la propagación de la enfermedad a través de las aves afectando severamente la salud humana.

El estudio plantea indagar el posible patrón de transmisión de esta enfermedad; además de información como la dirección y distribución de los eventos en el espacio-tiempo de China. Los investigadores buscaban explicar la forma de expansión de este tipo de gripe, para ello, hacen uso de la información de los pacientes infectados con este mal y que han sido reportados a la Organización Mundial de la Salud (OMS), se indaga si dicha relación presenta una distribución aleatoria o hay indicios de agrupaciones. El principal propósito de un estudio de esta naturaleza es aprender sobre este tipo de enfermedades, su patrón de propagación y otras características útiles para que se pueda prevenir y controlar los contagios.

En resumen, los objetivos de este estudio fueron:

(1) investigar las características epidémicas y el patrón direccional de los

brotes del virus A(H7N9), y

(2) identificar la presencia de agrupaciones espacio-temporales en humanos infectados con el virus A(H7N9).

El presente documento tiene como objetivo realizar una revisión del artículo, haciendo énfasis en las técnicas de procesos puntuales utilizadas, al igual que en la presentación de información espacial.

2. Resumen

Con información de 460 personas contagiadas con el virus A(H7N9), entre marzo del 2013 a diciembre 2014, los autores emprenden el análisis para alcanzar los objetivos ya mencionados. Para el desarrollo de las técnicas estadísticas emplean 3 herramientas: R, Satscan y ArcGIS, la primeras dos de acceso libre.

La técnica de elipse de desviación estándar se basa en las distancias de los puntos para crear elipses que concatenan los eventos y poder caracterizarlos con medidas de posición (el centro de la elipse), variabilidad (el área encerrada por la elipse) y dirección.

Para la detección de agrupaciones espacio-temporales, se utilizan simulaciones de Monte Carlo. En el artículo no se centran en la explicación de esta técnica, sin embargo, dan la referencia correspondiente. En la guía de usuario Satscan (2018) [1] se detalla el siguiente procedimiento: se definen para cada grilla o cuadrante los casos en riesgo o incidencias de contagio, posteriormente, se establece un radio fijo que puede ser el mismo cuadrante o ser ajustado, y se asume que el total de sujetos dentro del radio tienen la misma probabilidad de estar en riesgo, aleatoriamente se realizan las simulaciones reemplazando los casos reales con una nueva

distribución, hasta obtener la distribución esperada en los cuadrantes que fueron previamente definidos.

En el análisis temporal se construyen permutaciones de ventanas temporales, a cada ventana se estima la probabilidad log-verosimilitud para conformar agrupaciones, siendo la máxima log-verosimilitud la agrupación más probable, ya que es la que obtiene la mayor probabilidad de ocurrencia de contagio, las otras agrupaciones que cumplan con el criterio del *valor p* al 5 % de significancia conformarán agrupaciones secundarias. El test compara la distribución observada con la esperada, a través de las múltiples ventanas evaluando el número de casos esperados entre los observados y mediante una variable indicadora $I()$, que asume el valor de 1 cuando el número de casos es mayor al esperado en la ventana, en otros casos 0, estableciendo como hipótesis nula (H_0) la conformación de agrupaciones, como hipótesis alternativa que no hay suficiente evidencia estadística para formar agrupaciones.

La sección de resultados inicia con la representación de la distribución espacio-temporal de los casos reportados con este virus, esta descripción se presenta en forma de mapa en la Figura 1. Un aspecto fundamental que caracteriza los análisis espacio-temporales de este artículo es que se dividen por fases (3 fases en total). Esto debido a que los resultados de la curva epidémica¹ muestran 2 fases muy bien marcadas y una posible tercera, cuyo comportamiento no se podrá saber porque se sale de la ventana temporal de la información utilizada para el estudio. Estas fases las definen de la siguiente forma:

¹Representación gráfica del número de casos epidémicos de acuerdo a la fecha de la aparición de la enfermedad

Fase I: 13 de Marzo - 31 de Mayo, 2013 ($n = 130$)

Fase II: 1 de Junio - 31 de Mayo, 2013 - 2014 ($n = 294$)

Fase III: 1 de Junio - 31 de Diciembre, 2014 ($n = 36$)

Por esta razón aplican las dos técnicas estadísticas a cada una de las fases y grafican su resultado.

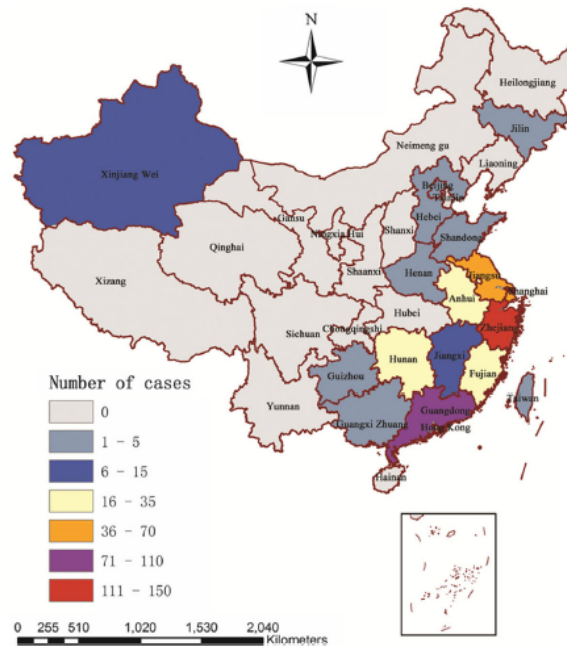


Figura 1: Distribución Espacial de casos reportados de personas con influenza A(H7N9) China entre Marzo 2013 y Diciembre 2014 ($n = 460$)

A partir de este punto el artículo se vuelve una descripción de los resultados obtenidos para cada una de las fases. Una abstracción de los resultados se puede observar en la Figura 2, donde el primer mapa muestra el resultado obtenido al aplicar la técnica de elipse de desviación estándar y los dos restantes la agrupación espacio-temporal.

Los principales resultados para cada una de las fases son:

Fase I: el centro medio de la elipse está situado en la provincia de Xuan-cheng, Anhui, y la elipse se concentró en el este de China, principalmente

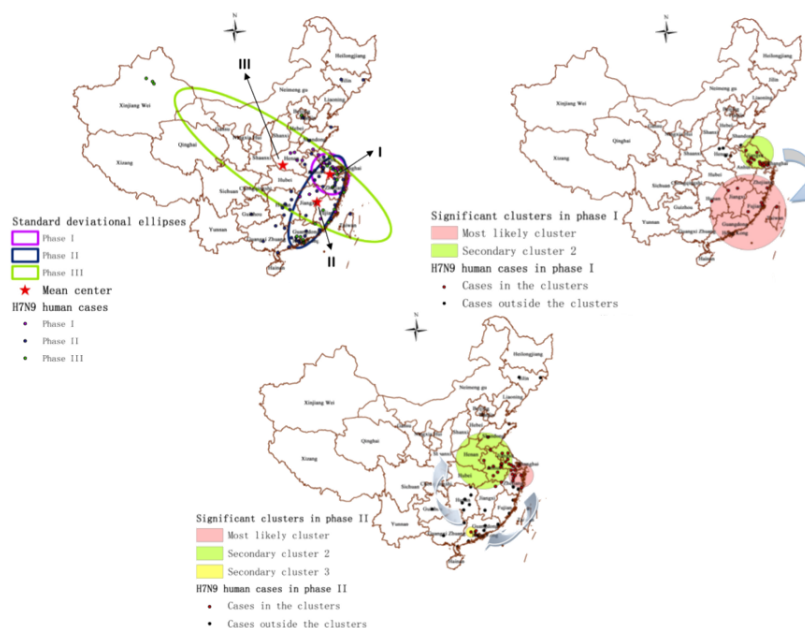


Figura 2: Resultados de los análisis estadísticos, para cada una de las fases

incluyendo Jiangsu, Anhui, Zhejiang y Shanghai. Por otra parte, la elipse mostró un alargamiento noroeste-sureste. En cuanto a las agrupaciones, se identifican dos agrupaciones significativas ($P < 0,05$) del brote: una siendo la más probable y otra siendo una agrupación secundaria.

Fase II: la distribución espacial de A(H7N9) se extendió a áreas costeras del sudeste y el centro medio de la elipse se movió hacia el sudoeste y finalmente situado en Fuzhou, provincia de Jiangxi, que estaba a punto de 476,01 km al suroeste del centro de la media de la fase I. Consecuentemente, la cobertura de la elipse en la fase II también se extendió desde el este a las zonas costeras sureste de China, principalmente, incluyendo Jiangsu, Anhui, Zhejiang, Shanghai, Jiangxi, Fujian, Guangdong, Hunan, y así sucesivamente. Además, la elipse presenta un alargamiento suroeste-noreste. Cuando se realizaron las agrupaciones, se identificaron 3 de ellas, una siendo la más probable y dos secundarias.

Fase III: el centro medio de la elipse se movió hacia el noroeste y se situó en Nanyang, provincia de Henan, que es alrededor de 779.27 km al noroeste del centro medio de la fase II. La extensión de la elipse en la fase III alcanza el máximo de las tres fases, incluyendo principalmente Taiwan, Fujian, Jiangxi y otras zonas importantes de China. Además, la elipse mostró un alargamiento noroeste-sureste. En esta fase no existieron agrupaciones estadísticamente significativas ($\alpha = 5\%$) de casos de A(H7N9)

3. Crítica

La pregunta de investigación se muestra con claridad, los objetivos están bien definidos, sin embargo, las limitaciones del estudio y el alcance de los resultados son poco claros. En la aplicación de técnicas estadísticas no describen algunas decisiones tomadas, no se mencionan distribuciones teóricas, supuestos ni alcances. Por ejemplo, para realizar la técnica de elipse de desviación estándar, utilizan una desviación estándar para el parámetro del tamaño de la elipse, pero no clarifican esta decisión. También, cuando realizan la técnica de agrupaciones espacio-temporales, no hacen el paso previo que es verificar la hipótesis de interacción espacio-temporal con pruebas como: Knox, Mantel o Jacquez, entre otras[3]. Por otra parte, aunque utilizan dos softwares de libre acceso, no se menciona la opción de reproducibilidad, herramienta trascendental en la ciencia moderna.

Asimismo, el artículo presenta un ordenamiento cronológico poco adecuado en relación a los cuadros y gráficos, reflejado principalmente en la asignación de espacio y la línea del artículo, en reiteradas ocasiones se

debe avanzar hasta dos páginas para observar la figura que se menciona en el texto, esto hace que los lectores se pierdan tratando de encontrar la figura correcta de la cual se derivan características importantes para la línea narrativa. Esto aunado a figuras muy extensas, hace que el artículo sea difícil de leer.

El artículo omite mencionar que la estimación de la función de máxima verosimilitud por defecto asume una distribución de Poisson, esta se define de la siguiente forma:

$$\left(\frac{c}{E(c)}\right)^p \left(\frac{C-c}{C-E(c)}\right)^{C-c} I() \quad (1)$$

donde,

C: es el número total de casos.

c: es el número de casos observados dentro de la ventana.

E(c): valor esperado de casos dentro de la ventana bajo la H_0

I(): es la variable indicadora.

La división en 3 fases por medio de las curvas epidemiológicas, carece de más fundamento. Además, como esta decisión afecta todo el desarrollo del artículo, se podría invertir más tiempo clarificando las decisiones tomadas. Por ejemplo, en la fase III el centro de la elipse está influenciado por valores atípicos y gran dispersión espacial, y se llegan a conclusiones algo precipitadas, únicamente por la disponibilidad de los datos y no por la naturaleza de los brotes de este virus.

Es común que en el análisis de patrones espaciales esté presente la exploración, la distribución y las condiciones de la población transmisora del virus, según Martha I. Nelson et al (2011) la zona de convergencia y los recorridos de contagio de los transmisores de enfermedades virales

aumentan las probabilidades de contagio en seres humanos, además, no se detalla que otros factores o covariables afectan el patrón direccional, los resultados del análisis muestran de manera simple que los cambios de direcciones únicamente se atribuyen a la elipse de desviación estándar, cuya explicación se limita a la distribución de dos variables la longitud y la latitud.

El artículo omite advertir las inconmensurables discusiones en el ámbito científico sobre el uso del *valor p*, el planteamiento de la significancia estadística al 5 % para determinar la conformación de agrupaciones no cuenta una justificación sustancial, según Sullivan et al (2012) [2] con una muestra suficientemente grande la prueba estadística casi siempre demostrará una diferencia significativa, a menos que no haya ningún efecto, sin embargo, las diferencias muy pequeñas, incluso si son significativas, a menudo no tienen sentido. Entonces, no hay suficiente propiedad para concluir que la agrupación secundaria de la fase II es significativa cuando el valor *p* es 0.044.

4. Conclusiones

Los autores confeccionan tres elipses de desviación estándar y cinco agrupaciones significativas ($P > 5\%$). La introducción explica la poca claridad presentada en la distribución del virus, cuya ausencia y altos picos del brote inducen ruido al análisis y no hay una propuesta para resolver dicha preocupación, la revisión bibliográfica recomienda incluir covariables en la distribución del brote, esto puede ser útil comprender dicha intermitencia.

Sobre el uso de las técnicas, se exhorta que el artículo debe presentar

mayor desarrollo teórico de las técnicas como no omitir supuestos o la definición de los parámetros, aunque se utilicen los parámetros por defecto, asimismo, hay omisión de pruebas para determinar el correcto uso de agrupaciones, también se encuentra el uso deliberado de la significancia estadística que finalmente concluye con la conformación de agrupaciones.

Referencias

- [1] Martin Kulldorff. *Satscantm user guide for version 9.6*, 2018.
- [2] Gail M. Sullivan and Richard Feinn. “using effect size – or why the p value is not enough.”. 2012.
- [3] Toshiro Tango. *Statistical Methods for Disease Clustering*. 2009.