Estimación de la probabilidad de mortalidad de la enfermedad por Ébola mediante un meta-análisis con efectos aleatorios

Ruth Narai Ble-Lezama, Addy Bolívar-Cimé

ruthnblez@gmail.com, addy.bolivar@ujat.mx
Universidad Juárez Autónoma de Tabasco

Resumen

La enfermedad por el virus del Ébola fue detectada por primera vez en 1976 y puede ser grave y mortal. El 17 de mayo de 2026, a raíz de la aparición de nuevos casos del virus en la Rep. Dem. del Congo y Uganda, la OMS declaró una "emergencia de salud pública de interés internacional". Ya que en cada brote la probabilidad de mortalidad puede variar debido a diversos factores, una forma de estimar la probabilidad de mortalidad global p de esta enfermedad es mediante un meta-análisis con efectos aleatorios. Utilizando los datos históricos hasta el 2026, en este trabajo usamos un enfoque de máxima verosimilitud en el meta-análisis con efectos aleatorios.

Introducción

La enfermedad causada por el virus del Ébola (por sus siglas en inglés, EVD) es una enfermedad aguda grave asociada a tasas de letalidad altas. Es ocasionada por los virus del género *Orthoebolavirus*, y existen cuatro especies que afectan a los seres humanos: Zaire, Sudán, Tai Forest y Bundibugyo.

En la actualidad, el brote activo de EVD de la especie Bundibugyo es el de mayor magnitud reportado en la República Democrática del Congo, presentando un ritmo de transmisión acelerado. La probabilidad de mortalidad por EVD ha sido calculada por [3], [4]. Sin embargo, dadas las condiciones de salud pública actuales, es necesario actualizar dicha probabilidad, incluyendo los datos más recientes no considerados en los anteriores meta-análisis.



Figura 1. Trabajador sanitario en RDC. Fuente: Council on Foreign Relations (2026)

Meta-análisis de efectos aleatorios

El meta-análisis es un método sistemático cuantitativo que permite analizar los resultados de estudios similares independientes, con el objetivo de producir una estimación del efecto global, [2].

Existen dos modelos estadísticos que permiten realizar inferencias acerca del efecto global de los estudios: el modelo de efectos fijos y el de efectos aleatorios. En este trabajo se utiliza el modelo de efectos aleatorios, el cual asume la existencia de heterogeneidad entre los efectos de los estudios. De esta forma, el efecto del i-ésimo estudio θ_i sigue una distribución de efectos con media μ y con un error denotado por δ_i. Además, el efecto observado $\hat{\theta}_i$ del estudio i es un estimador del efecto θ_i, con un error de muestreo ϵ_i. Así, para $i = 1, 2, ..., n$,

$$\theta_i = \mu + \delta_i \qquad \hat{\theta}_i = \theta_i + \epsilon_i. \tag{1}$$

Sustituyendo, podemos expresar el modelo de efectos aleatorios de la siguiente manera:

$$\hat{\theta}_i = \mu + \delta_i + \epsilon_i. \tag{2}$$

Asumiendo normalidad, $\epsilon_i \sim N(0, v_i)$, $\delta_i \sim N(0, \tau^2)$, donde τ^2 representa la heterogeneidad intraestudios y v_i la variación que existe dentro de cada estudio. Por consiguiente, tenemos que

$$\theta_i \sim N(\mu, \tau^2) \qquad \hat{\theta}_i \sim N(\mu, v_i + \tau^2). \tag{3}$$

Estimación puntual de los parámetros

Nuestros parámetros de interés son la media global μ y la varianza entre los estudios τ^2 . Es posible estimar la media global mediante una media ponderada, la cual está dada por la siguiente expresión

$$\hat{\mu} = \frac{\sum w_i \hat{\theta}_i}{\sum w_i} \tag{4}$$

donde el peso de cada observación (w_i) se define como el inverso de la varianza de los efectos observados, es decir

$$w_i = \frac{1}{v_i + \tau^2} \tag{5}$$

con v_i conocidos. En la práctica, estos deben ser estimados con los datos recabados.

Utilizando el método de máxima verosimilitud, podemos encontrar expresiones para los estimadores de interés. La función de logverosimilitud está dada por

$$l(\mu, \tau^2) = -\frac{n}{2}\ln 2\pi - \sum(\ln(v_i + \tau^2)) - \frac{1}{2}\frac{(\hat{\theta}_i - \mu)^2}{v_i + \tau^2}. \tag{6}$$

Llevando a cabo la maximización de la función con respecto a ambos parámetros, tenemos que sus estimadores satisfacen lo siguiente

$$\hat{\mu} = \frac{\sum w_i \hat{\theta}_i}{\sum w_i} \tag{7}$$

$$\hat{\tau}^2 = \text{máx}\left\{0, \frac{\sum w_i^2(\hat{\theta}_i - \hat{\mu})^2 - v_i}{\sum w_i^2}\right\}. \tag{8}$$

Obtenemos los valores de los estimadores eligiendo un valor inicial adecuado para $\hat{\tau}^2$, y posteriormente, iterando sobre $\hat{\tau}^2$ y $\hat{\mu}$ hasta que converjan, [6].

Intervalos de confianza de los parámetros

Intervalos de confianza por verosimilitud perfil

La logverosimilitud perfil se fundamenta en la función de logverosimilitud de uno de los parámetros de interés, tomando en consideración que el otro parámetro es desconocido y debe ser estimado. El estadístico de razón de verosimilitud de Wilk es

$$-2\left[l(\mu, \tau^2) - l(\hat{\mu}, \hat{\tau}^2)\right] \sim \chi_1^2. \tag{9}$$

Sea $\hat{\tau}^2(\mu)$ el estimador de máxima verosimilitud de τ^2 al variar el valor de μ y, análogamente, sea $\hat{\mu}(\tau^2)$ el estimador de máxima verosimilitud de μ al variar τ^2. Así, los intervalos de confianza del $95\,\%$ para los dos parámetros están dados por todos los valores que satisfacen:

$$l(\mu, \hat{\tau}^2(\mu)) > l(\hat{\mu}, \hat{\tau}^2(\hat{\mu})) - \frac{\chi^2_{0.05,1}}{2} \tag{10}$$

$$l(\hat{\mu}(\tau^2), \tau^2) > l(\hat{\mu}(\hat{\tau}^2), \hat{\tau}^2) - \frac{\chi^2_{0.05,1}}{2} \tag{11}$$

para μ y τ^2 respectivamente, y donde $\chi^2_{0.05,1}$ es el punto porcentual superior 0.05 de la distribución χ_1^2, [1].

Intervalos de confianza tipo Wald

Por la normalidad asintótica de los estimadores,

$$\hat{\mu} \sim N(\mu, SE(\hat{\mu})^2) \qquad \hat{\tau}^2 \sim N(\tau^2, SE(\hat{\tau}^2)^2) \tag{12}$$

donde

$$SE(\hat{\mu}) = \frac{1}{\sqrt{\sum w_i}} \tag{13}$$

$$SE(\hat{\tau}^2) = \sqrt{2\left(\sum w_i^2\right)^{-1}}. \tag{14}$$

Por lo que los intervalos de confianza del $95\,\%$ para μ y τ^2 están dados por las siguientes expresiones

$$\hat{\mu} \pm z_{0.025}(SE(\hat{\mu})) \tag{15}$$

$$\hat{\tau}^2 \pm z_{0.025}(SE(\hat{\tau}^2)) \tag{16}$$

donde $z_{0.025}$ es el punto porcentual superior 0.025 de la distribución $N(0, 1)$, [5].

Resultados

Se realizó un estudio de meta-análisis para calcular la tasa de letalidad global de la enfermedad por el virus del Ébola. Los datos fueron recuperados del artículo [3] y de las páginas web de la OMS y la CDC el 01 de agosto de 2026. Se excluyeron los datos referentes a brotes de la especie Reston, dado que no es patogénica para los seres humanos. Asimismo, se utilizó el logit de la tasa de letalidad p_i como el efecto θ_i para el estudio i-ésimo, es decir,

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1 - p_i}\right).$$

Por lo que una estimación de la tasa de letalidad global de los brotes está dada por

$$\hat{p} = \frac{1}{1 + \exp(-\hat{\mu})}.$$

La varianza dentro de cada estudio se estima por el cuadrado del error estándar del logit de la proporción, es decir

$$v_i = \frac{1}{n_i p_i (1 - p_i)}$$

donde n_i el tamaño muestral del estudio i-ésimo.

Asimismo, se aplica una corrección en los casos en los que $p_i = 0$ o $p_i = 1$. Los parámetros estimados fueron μ y τ^2 junto con sus intervalos de confianza, utilizando el software R.

Estimación de la probabilidad de mortalidad

A continuación se presentan tasas de letalidad de la enfermedad causada por el virus del Ébola.

	Estimación puntual	Wald	Perfil
Zaire	64.3	(56.7, 71.3)	(56.2, 71.3)
Sudán	47.5	(39.5, 55.7)	(37.9, 56.8)
Bundibugyo	38.5	(27.6, 50.8)	(19.1, 53.5)
Global	57.5	(50.7, 64.0)	(50.4, 64.0)

La tasa de letalidad global de la enfermedad causada por EVD resultó ser > 0.5, mostrando una alta probabilidad de mortalidad.

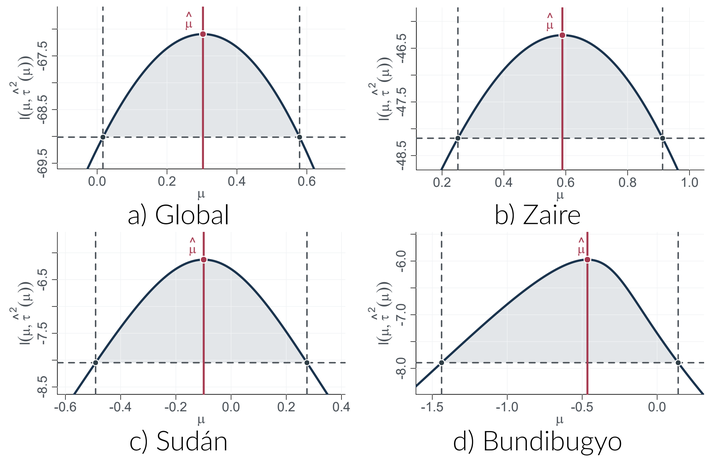


Figura 2. Función de logverosimilitud perfil de μ

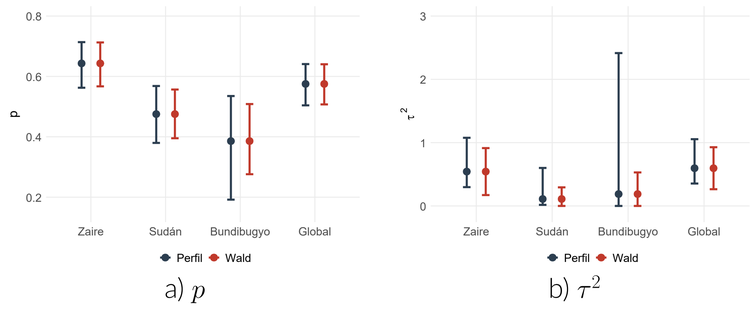


Figura 3. Intervalos de confianza para p y τ^2

Se puede observar que Zaire es la especie más letal, con intervalos de confianza por encima de 0.5. Es seguida por la especie Sudán, y posteriormente por la Bundibugyo, con la tasa de letalidad más baja.

En el caso de la especie Bundibugyo, a pesar de que sus intervalos de confianza indican la posibilidad de homogeneidad, este resultado no es representativo ya que la muestra de brotes es pequeña (5). Dado que los intervalos de confianza de τ^2 en el estudio global no contienen al cero, podemos afirmar que existe heterogeneidad entre los estudios, justificando el uso del meta-análisis de efectos aleatorios.

Conclusiones

- El meta-análisis mostró que hubo una disminución de la probabilidad de mortalidad del Ébola con respecto al estudio [3] del año 2022, de $60.6\,\%(51.6, 69.4)$ a $57.5\,\%(50.4, 64.0)$.
- La tasa de letalidad de la enfermedad varía por especie, siendo la de mayor probabilidad la especie Zaire, y la de menor porcentaje la especie Bundibugyo.
- Existe heterogeneidad entre los brotes, lo que sugiere que la letalidad varía según el tiempo y el lugar en que sucedió el brote.

Referencias

[1] R. J. Hardy and S. G. Thompson. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, 15(6):619–629, 1996.

[2] M. Harrer, P. Cuijpers, T.A. Furukawa, and D.D. Ebert. *Doing Meta-Analysis with R: A Hands-On Guide*. Chapman & Hall/CRC Press, London, 2021.

[3] J. Izudi and F. Bajunirwe. Case fatality rate for Ebola diseases, 1976-2022: A meta-analysis of global data. *Journal of Infection and Public Health*, 17(1):25–34, Jan 2024.

[4] A. Lefebvre, C. Fiet, C. Belpois-Duchamp, M. Tiv, K. Astruc, and L. S. Aho Glélé. Case fatality rates of Ebola virus diseases: A meta-analysis of World Health Organization Data. *Médecine et Maladies Infectieuses*, 44(9):412–416, 2014.

[5] Y. Pawitan. *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, 2001.

[6] A. A. Veroniki, D. Jackson, W. Viechtbauer, R. Bender, J. Bowden, G. Knapp, O. Kuss, J. P. Higgins, D. Langan, and G. Salanti. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, 7(1):55–79, 2016.