
Una propuesta metodológica para la selección de umbrales de clasificación en procesos de seroconversión y transmisión de anticuerpos a recién nacidos en Bogotá-Colombia

A methodological proposal to select a classification threshold in seroconversion and antibodies transmission processes to newborns in Bogotá-Colombia

Edwin R. Hernández^a
edwin.hernandez@usantotomas.edu.co

Diego F. Lemus^b
dflemusp@gmail.com

Wilmer D. Pineda^c
wilmerpineda@usantotomas.edu.co

Resumen

En el presente artículo se propone un método para seleccionar umbrales de clasificación que permita establecer si vacunar una madre gestante contra la tos ferina permite la transmisión de los anticuerpos necesarios para evitar que un neonato se contagie de esta enfermedad antes de recibir su primer vacuna. Adicionalmente, se propone la regresión gamma como medio para determinar los factores asociados a la producción de anticuerpos en la madre y la transmisión de éstos al recién nacido. Finalmente, como resultado de relevancia distrital, se obtuvieron los perfiles de las madres que poseen mayor probabilidad de transmisión de anticuerpos a sus hijos durante el periodo de gestación, información que puede ser empleada por la Secretaría Distrital de Salud (SDS) en el planteamiento de campañas preventivas y políticas públicas que tengan impacto en la población más vulnerable a esta enfermedad en la ciudad de Bogotá.

Palabras clave: Umbrales de clasificación, Tos ferina, Regresión Gamma, Regresión Logística, Curva ROC.

Abstract

In this article a method to select one classification threshold to establish whether vaccinating a pregnant mother against whooping cough allows the transmission of

^aProfesional en estadística

^bMagíster en Ciencias Estadística

^cDocente, Facultad de Estadística, Universidad Santo Tomás

the necessary antibodies to prevent a neonate from getting this disease before receiving their first vaccine is proposed. Additionally, gamma regression is proposed to identify mother's antibodies production associated factors and the transmission of these to the newborn. Finally, as a result of this study, the profiles of mothers who have a higher probability of transmitting antibodies to their children during the gestation period were obtained, information that can be used by the District Department of Health (SDS) in the approach of preventive campaigns and public policies that have an impact on the population most vulnerable to this disease in the city of Bogotá.

Keywords: Classification thresholds, Whooping cough, Gamma regression, Logistic regression, ROC curve.

1. Introducción

La mortalidad infantil es un evento de gran relevancia para todas las instituciones públicas de salud a nivel nacional (Ministerio de Salud 2014). A nivel distrital, la Secretaría Distrital de Salud (SDS) es la institución pública del Distrito Capital que busca disminuir la ocurrencia de este evento sin importar su causa (Secretaría Distrital de Salud 2001). En el caso de la tos ferina, la SDS se encuentra interesada en establecer los factores asociados a la producción de anticuerpos en la madre después de recibir la vacuna contra la enfermedad y la transmisión de éstos al recién nacido durante su periodo de gestación. De la misma manera, busca establecer umbrales en el proceso de generación de anticuerpos de las maternas que permitan identificar los perfiles de las gestantes (en Bogotá) que aumenten la probabilidad de transmisión de estos a los neonatos.

Para tal fin, se propone un método iterativo que permite determinar un umbral de seroconversión en las madres, un umbral de transmisibilidad de anticuerpos de la gestante al neonato y un umbral para identificar la tasa de pérdida de títulos en recién nacidos. Adicionalmente, se sugiere el uso de la regresión gamma para la identificación de los factores asociados a la producción de anticuerpos de tos ferina en la madre y la transmisión de éstos al recién nacido en la ciudad de Bogotá.

Los umbrales de clasificación obtenidos y las variables identificadas como factores asociados a la producción de anticuerpos en las gestantes pueden ser empleadas como herramientas de vigilancia epidemiológica a nivel distrital y permitirá la creación, promoción y seguimiento de campañas preventivas por parte de la SDS contra la tos ferina y así, de manera implícita, controlar y/o disminuir la tasa de mortalidad por esta enfermedad en recién nacidos sobre todo en la población más propensa o vulnerable de adquirirla. Sumado a esto, los resultados de la presente investigación pueden ser extrapolados a la población de gestantes en Colombia con el fin de establecer un método de clasificación nacional inicial y así evitar el uso de los estándares de la Organización Mundial de la Salud que fueron construidos para la población europea o estadounidense la cual difiere física y morfológicamente de la nuestra (Instituto Nacional de Salud 2017).

El resto del artículo está organizado de la siguiente forma: en la Sección 2 se presenta una breve contextualización de la enfermedad bajo estudio. En la Sección 3 se presenta la teoría estadística empleada para dar respuesta a los objetivos planteados en el estudio. En la Sección 4 se encuentran los conceptos inmersos en la construcción de los umbrales de clasificación y la metodología propuesta. En la Sección 5 se enuncian los factores y escenarios considerados en el estudio de simulación y los resultados del mismo. Finalmente, se presentan las conclusiones de la presente investigación.

2. Tos ferina

En la presente sección se presentan los aspectos generales, epidemiológicos, clínicos, de diagnóstico y de tratamiento de la tos ferina con el objetivo de contextualizar al lector del problema bajo análisis.

2.1. Aspectos generales y clínicos

Autores como (Ministerio de Salud 2014), (Paisley et al. 2012) y (Singh & Lingappan 2006) afirman que la tos ferina es una enfermedad aguda y altamente contagiosa causada por la bacteria *Bordetella parapertusis*. Su único huésped es el ser humano y su transmisión se da a través de las gotas que se producen al toser o estornudar; también se puede dar con el contacto con las secreciones respiratorias de una persona infectada (Frumkin 2013). La tos ferina posee una duración aproximada de ocho semanas y se desarrollan en tres importantes fases, la catarral, la paraxística y finalmente, la convaleciente (Kent & Heath 2014).

2.2. Aspectos Epidemiológicos

(Van Esso 2014) afirma que esta infección puede atacar a personas de cualquier cualquier grupo de edad pero los casos más graves se han detectado en neonatos y lactantes en los primeros meses de vida. Según (Stinson et al. 2015) la vacuna contra esta enfermedad no confiere inmunidad al paciente pues fue diseñada para proporcionar una protección (anticuerpos) que va desapareciendo paulatinamente en el transcurrir del tiempo.

2.3. Prevención

Según (Altunaiji et al. 2012) el principal medio para contrarrestar la adquisición de la tos ferina es la vacunación. En Colombia, la (Secretaría Distrital de Salud 2018) afirma que, para los niños entre los dos meses y seis años de edad, el cuadro de vacunación está dado por cinco dosis: la prevalente a los 2, 4, y 6 meses, y la DPT a los 18 meses y 5 años de edad. Después de la administración de un

esquema completo de vacunación, con cinco dosis en niños y tres en adultos, se ha encontrado que el 95% de las personas vacunadas tienen un nivel óptimo de inmunidad frente a la enfermedad (Qin 2015).

3. Marco Teórico

En esta sección se presenta un corto resumen de las metodología y modelos estadísticos que fueron empleados en el desarrollo del presente trabajo.

3.1. Modelo de regresión Gamma

(Agresti 2015) y (McCullagh & Nelder 1989) afirman que los modelos lineales generalizados (MLG) tienen tres componentes: una función de enlace $g(\cdot)$, un predictor lineal y un componente aleatorio tal como se presenta a continuación.

$$g(\mu_i) = X_i\beta + \epsilon_i, \quad (1)$$

donde $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ es el i -ésimo vector de p covariables para la i -ésimo valor observado de la variable de interés Y_i , y $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ es el vector de parámetros desconocidos que serán estimados. En el caso de la regresión gamma, se asume que $Y_i|\mu_i, \phi \sim \text{Gamma}(\mu_i, \phi)$ y la función de densidad de probabilidad está dada por:

$$f(Y_i|\mu_i, \phi) = \frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\mu_i}\right)^\phi \exp^{-\phi y_i/\mu_i} Y_i^{\phi-1}, \quad Y_i > 0. \quad (2)$$

La esperanza y la varianza de las variables aleatorias Y_i están dadas por $E(Y_i) = \mu_i$ y $\text{Var}(Y_i) = \frac{\mu_i^2}{\phi}$. Reescribiendo la ecuación (2) en términos de la familia exponencial se obtiene:

$$f(Y_i|\mu_i, \phi) = \phi \left\{ Y_i \left(-\frac{1}{\mu_i} \right) - \ln(\mu_i) \right\} - \ln(\Gamma(\phi)) + \phi \ln(\phi Y_i) - \ln(Y_i) \quad (3)$$

Bajo la ecuación (2), se puede identificar el parámetro natural $-1/\mu_i$ y el parámetro de dispersión ϕ . Adicional a esto, al asumir medias diferentes se obtiene que el coeficiente de variación es el mismo para todas las variables aleatorias y es igual a $CV(Y_i) = \phi^{-1/2}$ (Faraway 2016). Finalmente, resulta importante señalar que para la regresión gamma, las funciones de enlace consideradas en el presente estudio fueron la canónica $g(\mu_i) = -\frac{1}{\mu_i}$ y la logarítmica $g(\mu_i) = \log(\mu_i)$.

3.2. Modelo de regresión logística

Sean $n_i Y_i$, una secuencia de n variables aleatorias con distribución binomial de parámetros n_i y π_i . Si $n_i Y_i|\pi_i, n_i \sim \text{bin}(\pi_i, n_i)$ - la función de densidad de proba-

bilidad está dada por:

$$f(Y_i|\pi_i, n_i) = \binom{n_i}{n_i Y_i} \pi_i^{n_i Y_i} (1 - \pi_i)^{n_i - n_i Y_i} \tag{4}$$

Si se define $\theta_i = \log\left(\frac{\pi_i}{1-\pi_i}\right)$ se puede reescribir la ecuación (4) en términos de la familia exponencial de la siguiente manera:

$$f(Y_i|\pi_i, n_i) = \exp\left[\frac{Y_i \theta_i - \log[1 + \exp(\theta_i)]}{1/n_i} + \log\left(\binom{n_i}{n_i Y_i}\right)\right] \tag{5}$$

Según (Dobson & Barnett 2018) y (McCulloch & Searle 2004), en la ecuación (5) se puede identificar que el parámetro natural está dado por θ_i y el parámetro de dispersión por $\phi = 1/n_i$. La esperanza y la varianza de las variables aleatorias Y_i están dadas por $E(Y_i) = \pi_i$ y $Var(Y_i) = \frac{\pi_i(1-\pi_i)}{n_i}$.

De forma similar a la regresión Gamma, en este modelo se tienen tres componentes: una función de enlace $g(\cdot)$, un predictor lineal y un componente aleatorio tal como se presentó en la ecuación 3 donde $X_i^T \beta$ representa el predictor lineal y ϵ_i el i-ésimo residual (componente aleatorio); $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ es el i-ésimo vector de p covariables para la i-ésimo valor observado de la variable de interés Y_i , y $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ es el vector de parámetros desconocidos que serán estimados (Hardin & Hilbe 2018). Las funciones de enlace consideradas para la regresión logística en este estudio, el enlace canónico $g(\mu_i) = -\frac{1}{\mu_i}$ y el enlace logarítmico $g(\mu_i) = \log(\mu_i)$.

3.3. Umbrales de clasificación

Según (Balsa 2017) los modelos de regresión logística son ampliamente utilizados como sistemas de clasificación pues permiten generar un arreglo numérico conocido como matriz de confusión, la cual tiene la siguiente estructura (Figura 1):

		Pronóstico	
		Predicción positiva	Predicción negativa
Condición Verdadera	Condición positiva	Verdadero positivo	Falso negativo
	Condición negativa	Falso positivo	Verdadero negativo

Figura 1: Tabla de confusión

En la metodología propuesta para el proceso de identificación de los umbrales de seroconversión, la matriz de confusión y el área bajo la curva ROC (denominada en la literatura relacionada como AUC) resultan de vital importancia ya que los resultados que reporta permiten determinar el grado de fiabilidad del clasificador de discernir adecuadamente entre los dos posibles estados (seroconversión y

no seroconversión), es decir, indica la probabilidad de aciertos previstos para el modelo predictivo analizado. Los valores referenciados por (Balsa 2017) para la interpretación del AUC se muestran a continuación:

- Valores entre 0.5 y 0.6 implican que el modelo ajustado no es el adecuado.
- Valores entre 0.6 y 0.75 implican que el modelo ajustado tiene una tasa de clasificación regular.
- Valores entre 0.75 y 0.9 implican que el modelo ajustado tiene una buena tasa de clasificación.
- Valores entre 0.9 y 0.97 implican que el modelo ajustado tiene una tasa de clasificación muy buena.
- Valores entre 0.97 y 1 implican que el modelo ajustado tiene una excelente tasa de clasificación.

Cabe destacar que este tipo de herramientas gráficas son muy utilizadas para seleccionar modelos de diagnóstico más óptimos, dado que, según (Qin 2015) la curva ROC es independiente de la distribución de las clases de la población y se relaciona directamente con el costo/beneficio en la toma de decisiones diagnósticas.

3.4. Sensibilidad

En (Bjornstad 2018) se especifica que la sensibilidad es el parámetro que permite calcular la probabilidad de obtener un verdadero positivo en el diagnóstico de la unidad de observación bajo estudio. Para el cálculo de la sensibilidad se emplea la siguiente ecuación (Peng & Dominici 2008):

$$\text{sensibilidad} = \frac{VP}{CP} \quad (6)$$

donde VP denota la suma de verdaderos positivos y CP se define como condición positiva, que corresponde a la suma de verdaderos positivos y los falsos negativos. En el presente estudio, se consideran verdaderos positivos a las gestantes que seroconvierten, los neonatos que recibieron un número adecuado de anticuerpos y los neonatos que los conservan durante su primer trimestre de vida.

3.5. Especificidad

La especificidad es el parámetro que permite determinar la probabilidad de obtener un verdadero negativo en el diagnóstico de la unidad de observación bajo estudio. Según (Chan 2015) y (Merrill 2009) la expresión teórica de la especificidad está dada por:

$$\text{especificidad} = \frac{VN}{CN} \quad (7)$$

donde VN representan la suma de verdaderos negativos y CN la condición negativa, que corresponde a la suma de los falsos positivos y los verdaderos negativos. En el presente estudio, se consideran verdaderos negativos a las gestantes que no

seroconvierten, los neonatos que no recibieron un número adecuado de anticuerpos y los neonatos que no conservan estos durante su primer trimestre de vida.

4. Resultados y discusión

En la presente sección se hace una descripción detallada de las consideraciones metodológicas empleadas para el desarrollo del problema y los resultados de la aplicación propuesta.

4.1. Análisis descriptivo

Los datos empleados en la presente investigación fueron tomados del sistema de vigilancia epidemiológica de la SDS para la tos ferina y la base de nacidos vivos de la misma entidad. A continuación, se listan las variables más relevantes para el desarrollo de la presente investigación:

- **Variable interés:** Número de anticuerpos desarrollados por cada una de las unidades de observación (madre gestante y neonato) en los diferentes momentos del estudio.
- **Covariables:** Número de embarazos, tipo de afiliación al sistema de salud, número de consultas prenatales, peso según edad gestacional, tiempo intergenésico, tiempo de gestación, tipo de parto, multiplicidad en el parto, edad quinquenal de la madre y número de hijos vivos, entre otras.

El estudio fue desarrollado en cuatro periodos con el fin de dar cumplimiento a los siguientes objetivos:

- **Periodo 1:** Medir el número de títulos que presentó la madre gestante en el tercer mes de embarazo antes de recibir la vacuna DTPA.
- **Periodo 2:** Medir el número de títulos que presentó la madre gestante en el sexto mes de embarazo después de recibir la vacuna DTPA.
- **Periodo 3:** Medir el número de títulos que presentó el neonato en el momento del nacimiento; muestra tomada del cordón umbilical.
- **Periodo 4:** Medir el número de títulos que presentó el neonato a los tres meses de nacido.

A partir de la información recolectada se generan las variables de interés en el estudio de la siguiente manera:

- **Diferencia entre periodo 2 con respecto al periodo 1:** variable asociada al número de títulos desarrollado por la madre gestante después de la vacunación.
- **Periodo 3:** variable asociada al número de títulos heredados por el neonato.
- **Diferencia entre periodo 3 y periodo 4:** variable asociada al número de títulos que el neonato desarrolla durante sus primeros meses de vida.

En la Figura 2 y la Figura 3 se observan las densidades de las variables de interés en el presente estudio:

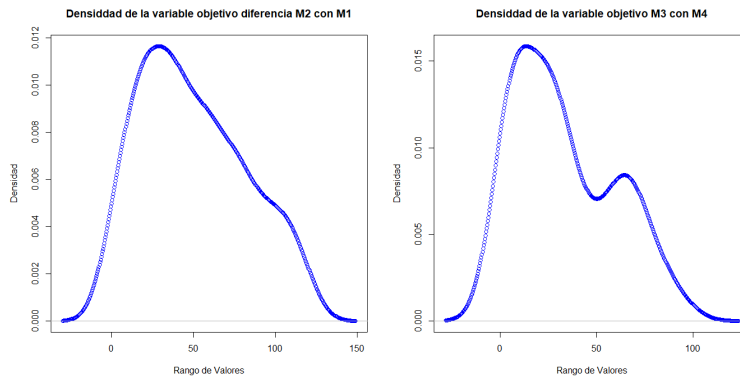


Figura 2: Densidad de la diferencia de títulos en la madre gestante y en el neonato

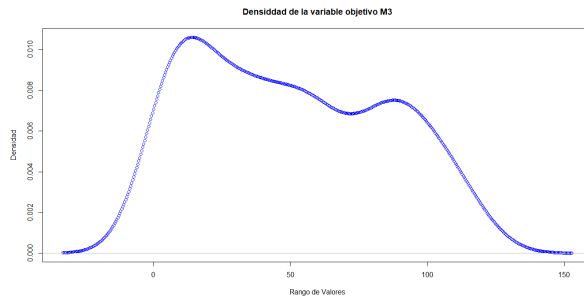


Figura 3: Densidad del número de títulos heredados por el neonato

Se realizaron pruebas de bondad de ajuste (Kolmogorov) y se determinó que las variables de interés se ajustan adecuadamente a una distribución gamma.

4.2. Modelamiento estadístico

En esta subsección se presenta el ajuste de los modelos de regresión Gamma, con el fin de identificar las variables que inciden en el comportamiento de las variables de interés del estudio.

4.2.1. Modelo para el número de títulos desarrollado por la madre gestante después de la vacunación

En la Tabla 1 se presenta el modelo de regresión Gamma ajustado para el número de títulos desarrollado por la madre gestante después de la vacunación después de emplear diferentes métodos de selección de variables. Las variables que quedaron

en el modelo fueron el número de embarazos (n_{emb}), el tipo de afiliación (t_{afil}), el número de consultas prenatales (numconsul) y la edad de la madre (Edad).

Tabla 1: Modelo ajustado para el número de títulos desarrollados por las madres

Variable	Estimador	Error estándar	Estadístico t	Pr(> t)
Intercepto	4.01	0.18	21.99	< 2e-16
n_{emb}	-0.08	0.03	-2.57	0.01
t_{afil} (Especial)	0.33	0.19	1.72	0.09
t_{afil} (Subsidiado)	0.29	0.11	2.63	0.01
numconsul (Cero)	-0.32	0.32	-1.00	0.32
numconsul (6-10)	0.11	0.10	1.15	0.25
numconsul (más de 10)	-0.75	0.24	-3.16	0.00
Edad	-0.19	0.09	-2.05	0.01

Como parte de la validación del modelo en la Figura 4 se observa que, a pesar de la presencia de colas pesadas, los residuales tienen un buen comportamiento.

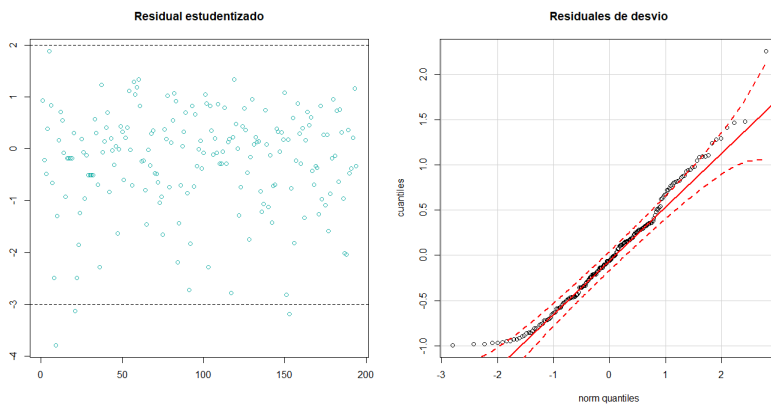


Figura 4: Residuales del modelo ajustado para el número de títulos desarrollado por la madre

En la Figura 5 se observan los perfiles obtenidos para el modelo ajustado, donde se evidencia que a medida que aumenta la cantidad de embarazos de la mujer se espera que en promedio el número de anticuerpos empiece a disminuir. Adicional a la condición anterior, una mujer que tiene de 6 a 10 controles prenatales en su primer embarazo tiende a tener en promedio mayor de número de títulos que una mujer que tiene de 1 a 5 controles y más de 10 controles prenatales.

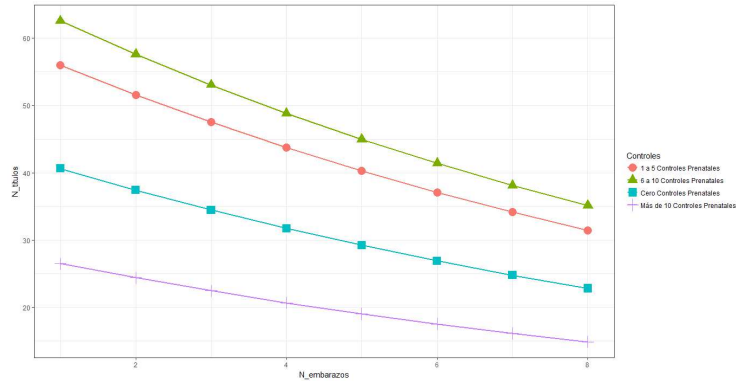


Figura 5: Perfiles del modelo ajustado para el número de títulos desarrollado por la madre

4.2.2. Modelo para la transmisión de anticuerpos al neonato

En la Tabla 2 se presentan los resultados del modelo de regresión Gamma ajustado para analizar el comportamiento de la transmisión de anticuerpos al neonato. Después del proceso de selección de variables se determinó que el tipo de parto (t_{parto}), el tiempo intergenésico (t_{inter}), el número de embarazos de la madre (n_{emb}), el número de anticuerpos desarrollados por la madre (difuml) y el tiempo entre la vacunación contra la DTP y el día del nacimiento (t_{vacu}) son las variables relevantes en la transmisión de anticuerpos al neonato.

Tabla 2: Modelo ajustado para la transmisión de anticuerpos al neonato

Variable	Estimador	Error estándar	Estadístico t	$\Pr(> t)$
Intercepto	3.10	0.24	13.01	$< 2e - 16$
t_{parto} (Espontáneo)	0.08	0.11	0.72	0.47
t_{parto} (Instrumentado)	-1.47	0.43	-3.40	0.00
t_{inter} (1 a 5 años)	-0.43	0.22	-0.86	0.03
t_{inter} (5 a 7 años)	-0.57	0.25	-1.89	0.01
t_{inter} (7 a 8 años)	-0.99	0.33	-2.99	0.00
t_{inter} (8 a 9 años)	-1.57	0.41	-3.83	0.00
t_{inter} (9 o más años)	-0.86	0.45	-1.91	0.03
n_{emb}	0.11	0.04	2.45	0.02
difuml	0.004	0.001	4.74	$4.64e - 06$
t_{vacu}	0.005	0.002	2.39	0.02

Como parte de la validación del modelo ajustado para la transmisión de anticuerpos al neonato en la Figura 6 se observa que, a pesar de la presencia de colas pesadas, los residuales tienen un buen comportamiento.

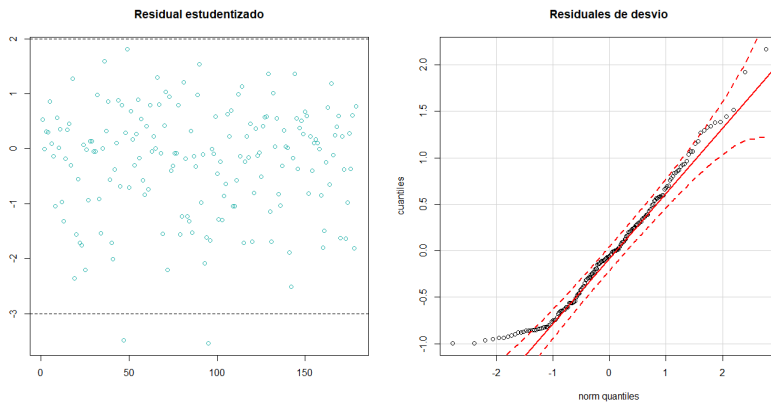


Figura 6: Residuales del modelo ajustado para la transmisión de anticuerpos al neonato

En la Figura 7 se visualizan los perfiles obtenidos para el modelo ajustado; se evidencia que, en promedio, el número de títulos que trasmite la mamá al neonato son pocos cuando el parto es instrumentado, sin importar la cantidad de embarazos que ha tenido la madre gestante ni tampoco el tiempo intergenésico. Ocurre lo contrario si el embarazo es espontaneo o por cesárea.

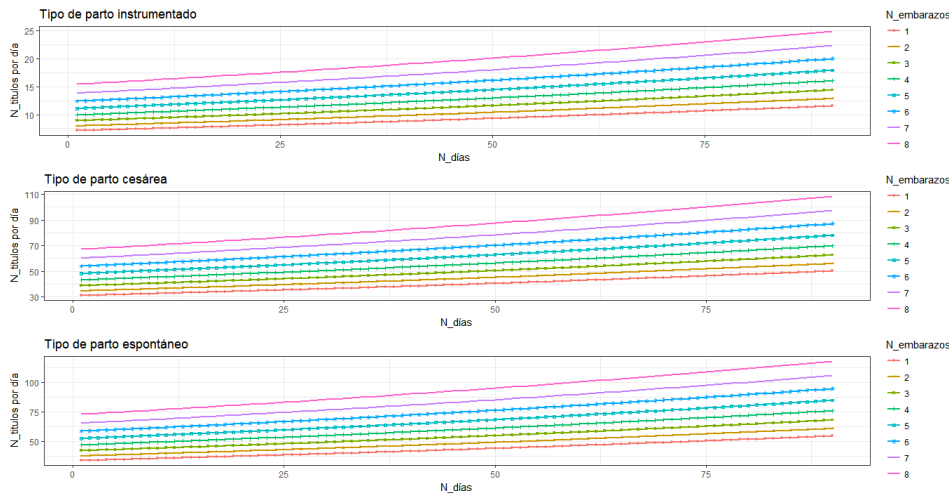


Figura 7: Perfiles modelo de P3

4.2.3. Modelo para el número de títulos desarrollados por el neonato

En la Tabla 3 se presentan los resultados del modelo de regresión Gamma ajustado para el número de títulos desarrollados por el neonato. Después del proceso de selección de variables se determinó que el tiempo de gestación (t_{ges}), la edad quinquenal de la madre (Edad) y el número de consultas prenatales (numconsul)

son las variables relevantes en la evolución de los anticuerpos durante los tres primeros meses de vida del neonato.

Tabla 3: Modelo ajustado para el número de títulos desarrollados por el neonato

Variable	Estimador	Error estándar	Estadístico t	$\Pr(> t)$
Intercepto	3.23	0.23	14.06	$< 2e - 16$
t_{ges} (Pretérmino)	-0.88	0.25	-3.55	$6.35e - 05$
Edad (20 a 29 años)	0.56	0.26	2.13	0.04
Edad (30 a 34 años)	0.46	0.22	2.09	0.02
Edad (35 años o más)	-0.73	0.36	-2.03	0.02
numconsul (Cero)	-3.20	0.73	-4.39	$3.38e - 05$
numconsul (más de 5)	-0.61	0.33	-1.84	0.03

Como parte de la validación del modelo ajustado para la transmisión de anticuerpos al neonato en la Figura 8 se observa que, a pesar de la presencia de colas pesadas, los residuales tienen un buen comportamiento.

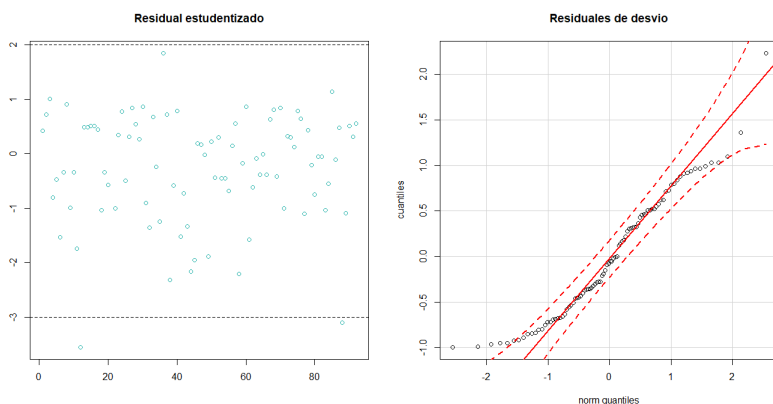


Figura 8: Residuales del modelo ajustado para el número de títulos desarrollados por el neonato

4.3. Umbrales de clasificación

La metodología que se propone en esta sección consiste ajustar un modelo de regresión logística y estimar el AUC respectivo a partir de la variable de clasificación obtenida (seroconversión o no seroconversión) a partir del umbral simulado.

4.4. Seroconversión en madres gestantes

En la Figura 9 se observa el comportamiento de las curvas ROC con los diferentes umbrales de clasificación para un intervalo entre 10 y 80 en el número de títulos desarrollados por las madres gestantes. Mediante este gráfico y los cálculos del AUC para cada una de las curvas ROC se determinó que el área bajo la curva es máxima para los umbrales entre 38 y 53 títulos.

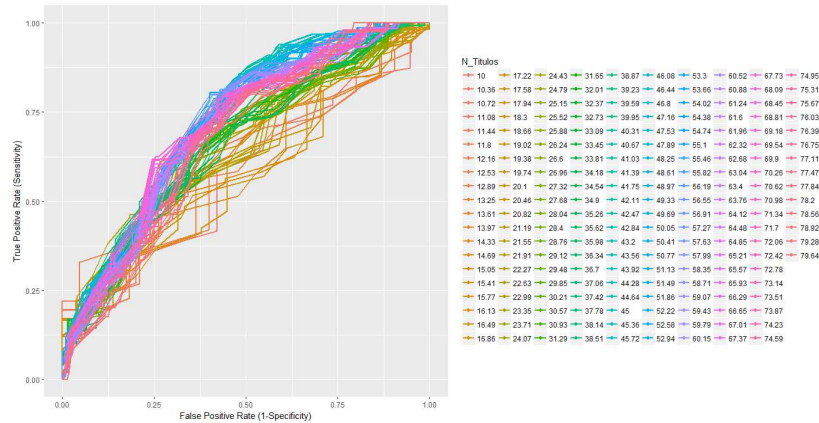


Figura 9: Curva ROC con totalidad de umbrales

De forma más específica en la Figura 10 se presentan las curvas ROC para los grupos de umbrales entre 30-40 títulos y entre 40-50 que son los intervalos donde dicha curva presenta mejor convexidad. Nótese que en el primer grupo de curvas hay una gran variabilidad en la capacidad explicativa del modelo, mientras que en el segundo grupo esta capacidad es mucho más estable; por este motivo se buscó el umbral de interés en este grupo. Finalmente, el umbral seleccionado corresponde a la curva ROC con mayor AUC que, en este caso, es el de 41.03 títulos.

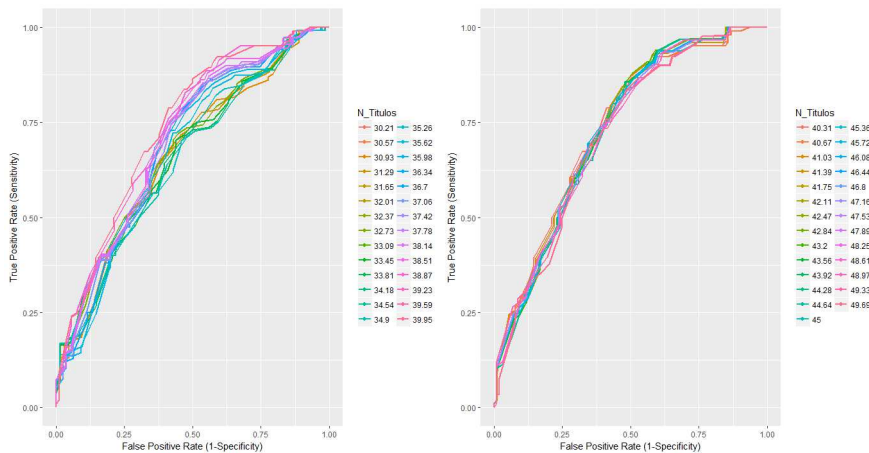


Figura 10: Curva ROC con umbrales de 30-40 títulos y 40-50 títulos

4.5. Trasmisión a neonatos

En la Figura 11 se observa el comportamiento de las curvas ROC con los diferentes umbrales de clasificación para el número de títulos que transmiten las madres gestantes a los neonatos. Mediante este gráfico y los cálculos del AUC para cada una de las curvas ROC se determinó que el área bajo la curva es máxima para los umbrales entre 20 y 45 títulos.

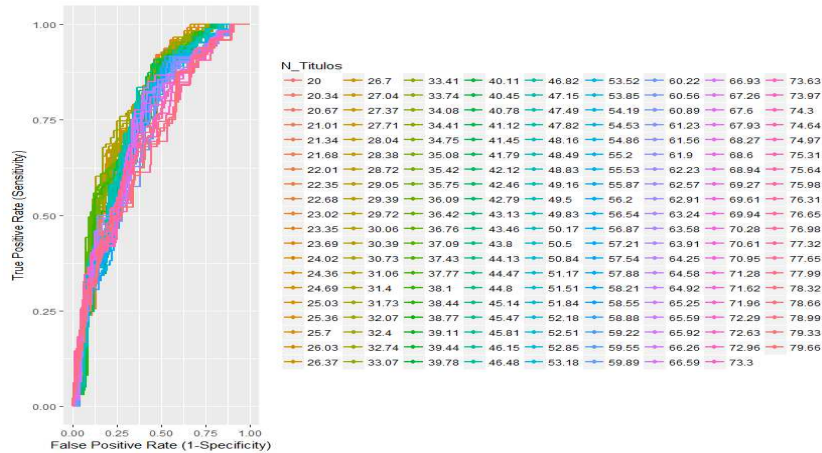


Figura 11: Curva ROC con totalidad de umbrales

De forma más específica en la Figura 12 se presentan las curvas ROC para los grupos de umbrales entre 20-30 y 30-40 títulos, que son los intervalos donde dicha curva presenta mejor convexidad. En este caso particular, la variabilidad de la capacidad explicativa del modelo es similar para ambos grupos de curvas ROC y en ese sentido la selección del umbral dependió exclusivamente del valor del AUC. Finalmente, el umbral seleccionado fue el de 30.73 títulos.

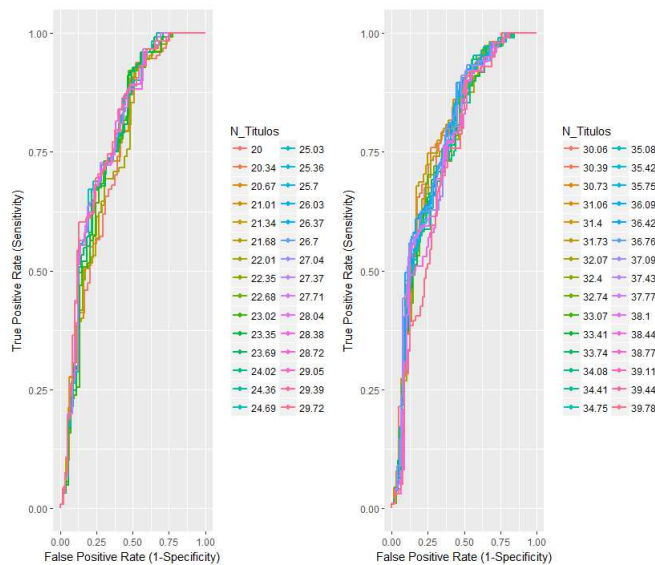


Figura 12: Curva ROC con umbrales de 20-30 títulos y 30-40 títulos

Este resultado indica que si la medición sobre el neonato es mayor a 30.73 títulos la vacuna aplicada sobre la madre fue transmitida adecuadamente al recién nacido. Esta medida es de gran relevancia porque permite determinar aquellos recién nacidos que son más susceptibles de contraer la enfermedad antes de su primera vacuna.

contra dicha enfermedad.

4.6. Evolución de los anticuerpos en los neonatos

En la Figura 13 se observa el comportamiento de las curvas ROC con los diferentes umbrales de clasificación para la evolución de los anticuerpos en los neonatos. Mediante este gráfico y los cálculos del AUC para cada una de las curvas ROC se determinó que el área bajo la curva es máxima para los umbrales entre 68 y 80 títulos.

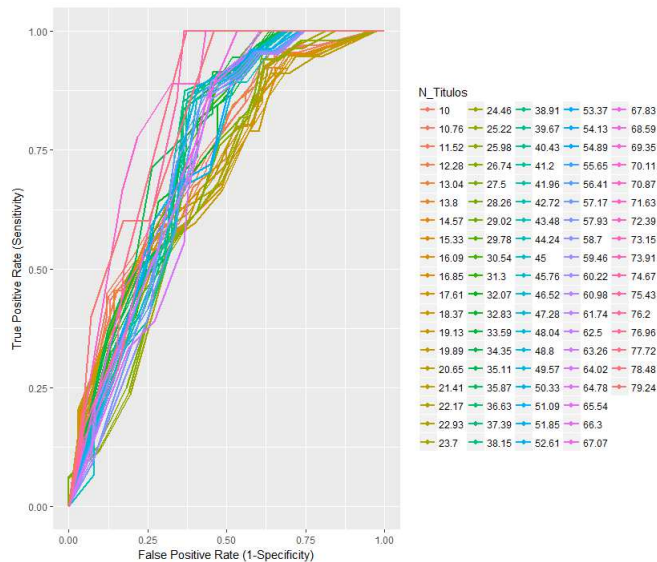


Figura 13: Curva ROC con totalidad de Umbrales

En la Figura 14 se presentan las curvas ROC para los grupos de umbrales entre 60-70 y 70-80 títulos, que son los intervalos donde dicha curva presenta mejor convexidad. En este caso particular, la variabilidad de la capacidad explicativa del modelo es similar para ambos grupos de curvas ROC y en ese sentido la selección del umbral dependió exclusivamente del valor del AUC. Finalmente, el umbral seleccionado fue el de 69.35 títulos.

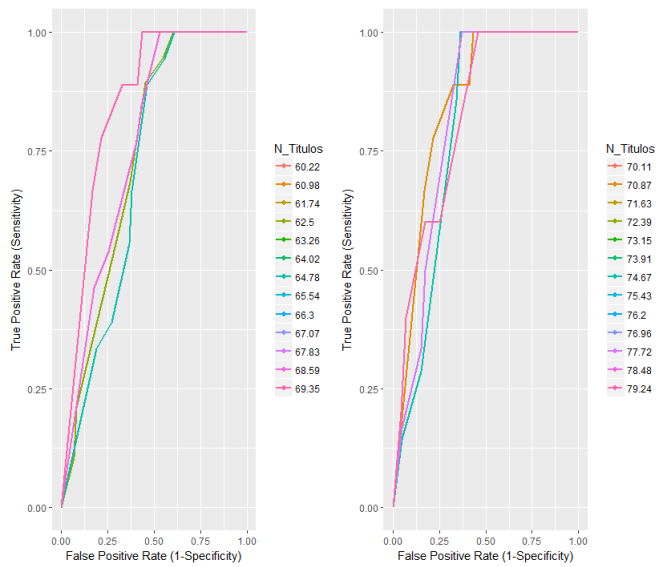


Figura 14: Curva ROC con umbrales de 60-70 títulos y 70-80 títulos

Finalmente, en la Figura 15 se encuentran las curvas ROC que poseen mayor AUC, para cada uno de los escenarios, es decir para el control de la madre gestante, relación entre la madre y el neonato, y por último el neonato.

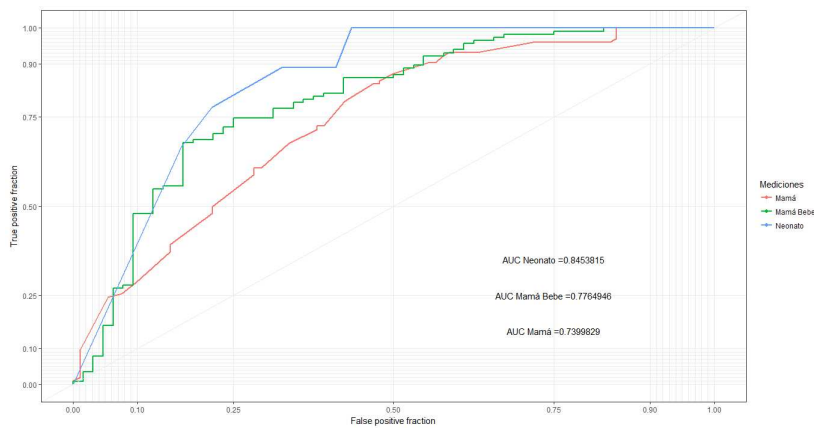


Figura 15: Curva ROC con los mejores umbrales

5. Conclusiones

- Los resultados obtenidos de las estimaciones del modelo gamma, permiten evidenciar que los factores de riesgo influyentes en la producción de anticuerpos de la madre gestante están relacionados con características físicas de la madre y su relación con el periodo gestacional. Bajo esta consideración, las personas a cargo de la toma de decisiones públicas en el sector salud podrían generar un programa de control y seguimiento para las madres gestantes y sus neonatos, con el fin de intervenir de manera temprana dichos pacientes y lograr la prevención de la adquisición de la tos ferina.
- Determinar los valores de los umbrales en escenarios propios del proceso gestacional para la población colombiana específicamente en la ciudad de Bogotá, permitirá controlar e identificar a la población más vulnerable en la adquisición de la tos ferina. Resulta importante señalar que los umbrales establecidos para la madre y el neonato presentan una diferencia en la cantidad de títulos observados, esto debido a que los primeros meses de vida el neonato es más vulnerable a cualquier tipo de infección y requiere de una mayor capacidad de defensa antes de su primer vacuna contra la enfermedad.

Recibido: Noviembre de 2019

Aceptado: Enero de 2020

Referencias

- Agresti, A. (2015), *Foundations of Linear and Generalized Linear Models*, 1 edn, Wiley, Hoboken, New Jersey.
- Altunaiji, S., Kukuruzovic, R., Curtis, N. & Massie, J. (2012), 'Antibiotics for whooping cough (pertussis)', *Evidence-Based Child Health* 7(3), 893–956.
- Balsa, C. (2017), Un paquete r para análisis masivo de modelos predictivos de regresión logística multivariante y sus medidas de discriminación y de clasificación asociadas, Master's thesis, Universitat Oberta de Catalunya, Catalunya, España.
- Bjornstad, O. N. (2018), *Epidemics: Models and Data using R*, 1 edn, Springer, New York, New York.
- Chan, B. (2015), *Biostatistics for Epidemiology and Public Health Using R*, 1 edn, Springer, New York, New York.
- Dobson, A. & Barnett, A. (2018), *An Introduction to Generalized Linear Models*, 4 edn, Wiley-Interscience, Londres, Inglaterra.
- Faraway, J. (2016), *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*, 2 edn, Chapman y Hall, Londres, Inglaterra.

- Frumkin, K. (2013), 'Pertussis and persistent cough: Practical, clinical and epidemiologic issues', *The Journal of Emergency Medicine* **44**(4), 889–895.
- Hardin, J. & Hilbe, J. (2018), *Generalized Linear Models and Extensions*, 4 edn, Stata Press, Hoboken, New Jersey.
- Instituto Nacional de Salud, I. (2017), 'Protocolo de vigilancia en salud pública: Tos ferina'.
*<https://www.ins.gov.co>
- Kent, A. & Heath, P. (2014), 'Pertussis', *Medicine* **42**(1), 8–10.
- McCullagh, P. & Nelder, J. (1989), *Generalized Linear Models*, 2 edn, Chapman y Hall, Londres, Inglaterra.
- McCulloch, C. & Searle, S. (2004), *Generalized, Linear, and Mixed Models*, 1 edn, Chapman y Hall, Hoboken, New Jersey.
- Merrill, R. (2009), *Introduction to Epidemiology*, 5 edn, Jones y Bartlett Publishers, Burlington, Massachusetts.
- Ministerio de Salud, M. (2014), 'Guía de práctica clínica para la identificación y el manejo clínico de la tos ferina en menores de 18 años de edad'.
*<http://gpc.minsalud.gov.co>
- Paisley, R., Blaylock, J. & Hartzell, J. (2012), 'Whooping cough in adults: An update on a reemerging infection', *The American Journal of Medicine* **125**(2), 141–143.
- Peng, R. & Dominici, F. (2008), *Statistical Methods for Environmental Epidemiology with R: A Case Study in Air Pollution and Health*, 1 edn, Springer, New York, New York.
- Qin, X. (2015), 'Resurgence of pertussis and its laboratory diagnosis', *Clinical Microbiology Newsletter* **37**(9), 69–76.
- Secretaría Distrital de Salud, S. (2001), 'Tos ferina'.
*<http://www.saludcapital.gov.co>
- Secretaría Distrital de Salud, S. (2018), 'Distrito intensifica vacunación'.
*<https://bogota.gov.co>
- Singh, M. & Lingappan, K. (2006), 'Whooping cough: The current scene', *Chest* **130**(5), 1547–1553.
- Stinson, C., Hooper, G. & Oliver, J. (2015), 'An evidence-based protocol for protecting newborns from pertussis', *Nursing for Women's Health* **19**(5), 402–409.
- Van Esso, D. (2014), 'Actualización en tos ferina', *Pediatría Integral* **18**(2), 101–107.