



UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E
INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E
INFORMÁTICA



**MODELO DE REGRESIÓN LOGÍSTICA BINARIA PARA
PREDECIR EL RIESGO DE ANEMIA EN NIÑOS MENORES DE 5
AÑOS EN LA RED DE SALUD CHUCUITO – 2023.**

TESIS

PRESENTADA POR:

YUDITH DIANA CHALCO CEREZO

**PARA OPTAR EL TÍTULO PROFESIONAL DE:
INGENIERO ESTADÍSTICO E INFORMÁTICO**

PUNO – PERÚ

2024



Yudith Diana Chalco Cerezo

MODELO DE REGRESIÓN LOGÍSTICA BINARIA PARA PREDECIR EL RIESGO DE ANEMIA EN NIÑOS MENORES DE 5 AÑOS EN LA RED DE SALUD CHUCUITO – 2023.

 Universidad Nacional del Altiplano

Detalles del documento

Identificador de la entrega

trn:oid:::8254:417856369

Fecha de entrega

20 dic 2024, 10:01 a.m. GMT-5

Fecha de descarga

20 dic 2024, 10:02 a.m. GMT-5

Nombre de archivo

TESIS FINAL _ YUDITH DIANA CHALCO CEREZO.pdf

Tamaño de archivo

1.1 MB

75 Páginas

15,671 Palabras

80,102 Caracteres



18% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- ▶ Bibliografía
- ▶ Texto citado
- ▶ Texto mencionado
- ▶ Coincidencias menores (menos de 12 palabras)

Fuentes principales

- 16% Fuentes de Internet
- 3% Publicaciones
- 9% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.



Firmado digitalmente por:
VILLASANTE SARAVIA FREDY
HERIC FIR 01307200 hard
Motivo: En señal de
conformidad
Fecha: 20/12/2024 20:21:17-0500



Firmado digitalmente por JUAREZ
VARGAS Juan Carlos FAU
20145496170 soft
Motivo: Soy el autor del documento
Fecha: 20.12.2024 14:57:13 -05:00



DEDICATORIA

Quiero expresar en primer lugar mi más profundo agradecimiento a Dios, por su gracia y por guiar mis pasos a lo largo de este camino lleno de aprendizajes y desafíos.

Dedicado este logro a mis padres, Jesús Mario Chalco Flores y María Cerezo Chachaque, quienes han sido el motor y la razón principal de cada uno de mis avances. Su amor incondicional y su fortaleza han sido un ejemplo invaluable en mi vida.

Yudith Diana Chalco Cerezo



AGRADECIMIENTOS

En primer lugar, agradezco a Dios quien me ha brindado la fuerza, sabiduría y la serenidad necesaria para superar cada obstáculo en este camino académico.

A mis padres, Jesús y María, les expreso mi más profundo agradecimiento por el amor incondicional, el respaldo constante y las valiosas lecciones que siempre me han dado. Su ejemplo de dedicación y esfuerzo incansable ha sido, sin duda, mi mayor motivación.

Al Dr. Fredy Heric Villasante Saravia, mi asesor de tesis, le agradezco profundamente su orientación, paciencia y consejos clave durante todo el proceso de investigación. Su experiencia y compromiso han sido esenciales para el desarrollo y culminación de este estudio.

Yudith Diana Chalco Cerezo



ÍNDICE GENERAL

	Pág.
DEDICATORIA	
AGRADECIMIENTOS	
ÍNDICE GENERAL	
ÍNDICE DE TABLAS	
ÍNDICE DE FIGURAS	
ÍNDICE DE ANEXOS	
ÍNDICE DE ACRÓNIMOS	
RESUMEN	13
ABSTRACT.....	14
CAPÍTULO I	
INTRODUCCIÓN	
1.1. DESCRIPCIÓN DEL PROBLEMA	17
1.2. FORMULACIÓN DEL PROBLEMA	18
1.2.1. Problema general.....	18
1.3. OBJETIVOS DE LA INVESTIGACIÓN.....	18
1.3.1. Objetivo general	18
1.3.2. Objetivos Específicos.....	18
1.4. HIPÓTESIS DE LA INVESTIGACIÓN	19
1.4.1. Hipótesis General	19
1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN	19
1.6. LIMITACIONES DE LA INVESTIGACIÓN	21



CAPÍTULO II

REVISIÓN DE LITERATURA

2.1.	ANTECEDENTES DE LA INVESTIGACIÓN.....	22
2.1.1.	Antecedentes Internacionales.....	22
2.1.2.	Antecedentes Nacionales.....	26
2.2.	MARCO TEÓRICO.....	30
2.2.1.	Anemia	30
2.2.2.	Factor de riesgo.....	32
2.2.2.1.	Factores determinantes de la anemia.....	32
2.2.3.	Modelo de regresión logística.....	34
2.2.3.1.	Significación de chi-cuadrado de Pearson χ^2	36
2.2.3.2.	R-Cuadrado de Cox y Snell y R-Cuadrado de Nagelkerke.....	37
2.2.3.3.	Porcentaje global correctamente clasificado.....	39
2.2.3.4.	Evaluadores del modelo de regresión logística.....	39
2.3.	DESCRIPCIÓN DE VARIABLES.....	42

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1.	UBICACIÓN GEOGRÁFICA DE LA INVESTIGACIÓN.....	43
3.2.	ÁMBITO DE LA INVESTIGACIÓN.....	43
3.3.	TIPO Y DISEÑO DE INVESTIGACIÓN.....	43
3.4.	POBLACIÓN.....	44
3.5.	INSTRUMENTO DE RECOLECCIÓN Y PROCESAMIENTO DE DATOS.....	44
3.5.1.	Procedimiento de recolección de datos.....	45
3.5.2.	Plan de procesamiento y análisis de datos.....	45



3.5.2.1. Análisis estadístico.....	45
3.6. OPERACIONALIZACIÓN DE VARIABLES	46
CAPÍTULO IV	
RESULTADOS Y DISCUSIÓN	
4.1. IDENTIFICACIÓN DE VARIABLES	48
4.1.1. Prevalencia de anemia.....	48
4.1.2. Selección de variables	49
4.2. ESTIMACIÓN DE MODELO.....	49
4.2.1. Selección de variables	49
4.2.2. Verosimilitud del modelo.....	55
4.2.3. Prueba de bondad de ajuste	56
4.3. PREDICCIÓN DEL RIESGO DE ANEMIA.....	58
4.4. DISCUSIÓN	59
V. CONCLUSIONES.....	62
VI. RECOMENDACIONES	63
VII. REFERENCIAS BIBLIOGRÁFICAS.....	64
ANEXOS.....	69

ÁREA: Modelos predictivos.

TEMA: Modelo de regresión logística binaria para predecir el riesgo de anemia en niños menores de 5 años en la red de salud Chucuito – 2023.

FECHA DE SUSTENTACIÓN: 27 de diciembre del 2024



ÍNDICE DE TABLAS

	Pág.
Tabla 1 Criterios OMS y Rangos Normales para Definir Anemia en Niños Menores de 5 Años	31
Tabla 2 Valores Normales de Concentración de Hemoglobina y Grados de Anemia en Niñas y Niños 6 a 59 Meses (hasta 1000 msnm)	31
Tabla 3 Operacionalización de Variables	46
Tabla 4 Prevalencia de Anemia Según Sexo, en los Niños Menores de 5 años en la Red de Salud Chucuito – 2023	48
Tabla 5 Variables que ingresan al Modelo	49
Tabla 6 Variables que Ingresan al Nuevo Modelo.....	50
Tabla 7 Recodificación de Variables Para el Modelo Logístico.....	51
Tabla 8 Prueba del Coeficiente de Determinación	52
Tabla 9 Tabla de Clasificación	53
Tabla 10 Área Bajo la Curva ROC	54
Tabla 11 Prueba de Coeficiente de Intercepción	55
Tabla 12 Prueba de Bondad de Ajuste (Hosmer y Lemeshow)	56
Tabla 13 Prueba Ómnibus.....	57
Tabla 14 Medidas Simétricas Kappa	57
Tabla 15 Predicción de Riesgo de Anemia en Niños Menores de 5 Años Para Dos Casos	58



ÍNDICE DE FIGURAS

	Pág.
Figura 1 Modelo Causal.....	33
Figura 2 Curvas ROC	41
Figura 3 Gráfica de la Curva COR	54



ÍNDICE DE ANEXOS

	Pág.
ANEXO 1. Registro SIEN, Chucuito - Puno del año 2023.....	69
ANEXO 2. Predicciones de Anemia	70
ANEXO 3. Descripción de Variables del SIEN - DIRESA provincia de Chucuito	71
ANEXO 4. Declaración jurada de autenticidad de tesis	74
ANEXO 5. Autorización para el depósito de tesis en el Repositorio Institucional.....	75



ÍNDICE DE ACRÓNIMOS

SIEN:	Sistema Integral De Estado Nutricional
SIS:	Sistema Integral de Salud
PIN:	Programa Integral de Nutrición
MINSA:	Ministerio de Salud
ENDES:	Encuesta Demográfica y de Salud Familiar
INEI:	Instituto Nacional de Estadística e Informática
OMS:	Organización Mundial de la Salud
C.S:	Centro de Salud
CLAS:	Comité Local de Administración de Salud
JUNTOS:	Programa de Apoyo Directo a los Más Pobres
EESS:	Establecimiento de Salud



RESUMEN

La anemia infantil representa uno de los principales desafíos de salud pública en el país, debido a su alta prevalencia y su significativo impacto epidemiológico y social en la salud. Esta investigación tuvo como objetivo desarrollar un modelo predictivo basado en regresión logística para predecir el riesgo de anemia en niños menores de 5 años en la Red de Salud Chucuito - 2023. La investigación es de tipo aplicada, de nivel correlacional, con un diseño no experimental, transversal-analítico. La población evaluada estuvo compuesta por 1931 niños, abarcando la totalidad de los datos. Para alcanzar los objetivos, se analizaron los registros proporcionados por el Sistema de Información del Estado Nutricional del Niño Menor de Cinco Años y Gestantes (SIEN) Red de Salud Chucuito – Puno, 2023. El modelo predictivo resultante fue: $Ln(y) = -4,548 + 1,090_{x1} + 0,342_{x2} + 0,693_{x3} + 1,266_{x4} - 0,764_{x5} + 0,811_{x6} - 3,253_{x7} - 0,946_{x8}$. Las variables significativas del modelo predictivo fueron: Distrito, Edad, Talla, Hemoglobina, Número de Visitas, Número de Sesiones, Número de Sachets, y el programa JUNTOS, La prueba de bondad de ajuste de Hosmer y Lemeshow, con un nivel de significación del 0.05, tuvo un p-valor de 0.273, y el rendimiento o efectividad del modelo clasificador utilizando la curva ROC, cubre un área de 0.973. El coeficiente de determinación de Nagelkerke fue de 75.5%. El modelo logró una tasa de clasificación de verdaderos negativos (niños no anémicos) del 78.8%, mientras que la tasa de verdaderos positivos (niños anémicos) fue del 95.4%. Demostrando así que el modelo predictivo tiene una alta capacidad predictiva.

Palabras clave: Hemoglobina, Modelo predictivo, Modelos estadísticos, Regresión logística binaria, Riesgo de anemia.



ABSTRACT

Childhood anemia represents one of the main public health challenges in the country due to its high prevalence and significant epidemiological and social impact on health. This research aimed to develop a predictive model based on logistic regression to predict the risk of anemia in children under 5 years old in the Chucuito Health Network - 2023. The study is applied in nature, correlational in level, and uses a non-experimental, cross-sectional analytical design. The evaluated population consisted of 1,931 children, encompassing the entirety of available data. To achieve the objectives, records provided by the Nutritional Status Information System for Children Under Five and Pregnant Women (SIEN) from the Chucuito Health Network – Puno, 2023, were analyzed. The resulting predictive model was: $Ln(y) = -4,548 + 1,090_{x1} + 0,342_{x2} + 0,693_{x3} + 1,266_{x4} - 0,764_{x5} + 0,811_{x6} - 3,253_{x7} - 0,946_{x8}$. The significant variables in the predictive model were: District, Age, Height, Hemoglobin, Number of Visits, Number of Sessions, Number of Sachets, and the JUNTOS program. The Hosmer-Lemeshow goodness-of-fit test, with a significance level of 0.05, yielded a p-value of 0.273, and the performance or effectiveness of the classifier model using the ROC curve covered an area of 0.973. Nagelkerke's determination coefficient was 75.5%. The model achieved a true negative classification rate (non-anemic children) of 78.8%, while the true positive rate (anemic children) was 95.4%. This demonstrates that the predictive model has a high predictive capacity.

Keywords: Hemoglobin, Predictive model, Statistical models, Binary logistic regression, Anemia risk.



CAPÍTULO I

INTRODUCCIÓN

La anemia es un problema global que afecta a 1620 millones de individuos a nivel mundial, lo que equivale al 24,8% de la población mundial; acorde con datos proporcionados por la Organización Mundial de la Salud (OMS). Los niños en etapa preescolar son los más impactados, presentando una prevalencia del 47 %. En el año 2018, el Perú enfrentó una situación preocupante, ya que el 46,6 % de los niños menores de 3 años sufrían de anemia, de acuerdo con los datos proporcionados por la Encuesta Demográfica y de Salud Familiar (ENDES) (Aquino Canchari, 2021).

En el Perú, la anemia representa un problema de salud pública que afecta al 43,6% de los niños menores de tres años. Esta proporción se ha mantenido prácticamente inalterada durante los últimos cinco años, a pesar de las iniciativas implementadas por el gobierno para disminuirla (Zavaleta, 2017).

En 2023, el 43,1 % de los niños y niñas de entre 6 y 35 meses en el país presentó anemia, según los datos del Instituto Nacional de Estadística e Informática (INEI) en los Resultados de los Principales Indicadores de los Programas Presupuestales de la Encuesta Demográfica y de Salud Familiar (ENDES-2023). Este problema de salud tuvo una mayor incidencia en las zonas rurales, alcanzando un 50,3 %, en comparación con el 40,2 % registrado en las áreas urbanas. A nivel departamental, los mayores índices de esta deficiencia de hierro en la sangre se observaron en los niños y niñas de Puno (70,4%), Ucayali (59,4%) y Madre de Dios (58,3%) (INEI, 2024).

En los niños peruanos, la anemia se presenta durante las fases de crecimiento y diferenciación más intensas de las células cerebrales, que abarcan los primeros 24 meses



de vida y el periodo de gestación. En estas fases, los requerimientos nutricionales para el desarrollo del feto y del niño son particularmente altos. Según la OMS, esta circunstancia convierte la anemia en un grave problema de salud pública en Perú. (MINSA, 2017).

La estructura del trabajo de investigación es la siguiente:

Capítulo I, presenta el problema de investigación, justificación, hipótesis general, objetivos a alcanzar y reconocemos las limitaciones.

Capítulo II, contiene el marco teórico que sirve como base para nuestra investigación, donde examinamos el contexto histórico que rodea el tema, establecemos los principios teóricos que sustentan nuestra investigación, Además, proporcionamos referencias a otras tesis que comparten temas similares.

Capítulo III, aborda los materiales y métodos empleados en la investigación y obtención de datos a utilizar, detallando el diseño del estudio, el tipo de investigación, la población y el tratamiento estadístico utilizado. También presenta la validez y confiabilidad de las herramientas.

Capítulo IV, expone los resultados obtenidos y ofrece su interpretación para evaluar el cumplimiento de los objetivos establecidos. Además, proporcionamos una discusión de antecedentes detallada que contextualiza e interpreta estos resultados, ofreciendo una comprensión más profunda de los hallazgos.

Capítulo V, se describen las conclusiones, recomendaciones abarcando todos los objetivos planteados y también presenta la bibliografía consultada y anexos relevantes para la investigación.



1.1. DESCRIPCIÓN DEL PROBLEMA

La anemia es un problema crítico de salud pública que afecta a las poblaciones de países tanto desarrollados como en desarrollo. Este trastorno, reconocido como el más extendido y frecuente a nivel global, se presenta con mayor intensidad en los grupos más vulnerables, incluyendo a los niños pequeños y a las mujeres en edad fértil. Se estima que afecta al 43 % de los niños menores de 5 años, al 38 % de las mujeres gestantes y 29 % de las mujeres que no están embarazadas (Zavaleta y Astete-Robilliard, 2017).

La anemia afecta de manera más significativa a las áreas rurales. En 2022, el 51,5 % de los niños menores de 36 meses padeció esta condición, lo que evidencia la necesidad de una acción prioritaria por parte del Estado. Los departamentos con mayores niveles de anemia en niños de entre 6 meses y 5 años fueron Puno, Ucayali, Huancavelica y Loreto. Este panorama destaca que la anemia está relacionada con una nutrición deficiente, provocada principalmente por la falta de hierro, lo que repercute negativamente en el desarrollo motor, emocional, cognitivo y social de los niños y niñas (Huaylinos Antezana, 2023).

La anemia es una condición prevalente que tiene profundas repercusiones tanto económicas como sociales. En la infancia, no solo perjudica el desarrollo físico y cognitivo de los niños, reduciendo su rendimiento escolar y aumentando su vulnerabilidad a enfermedades infecciosas, sino que también impone una carga económica considerable a las familias y al sistema de salud. Las comunidades más vulnerables son las más afectadas, lo que perpetúa ciclos de pobreza y restringe las oportunidades de desarrollo.

La ciudad de Puno presenta una elevada prevalencia de anemia, reflejando un alto índice de esta condición. A pesar de los esfuerzos realizados por las autoridades de salud



en Chucuito-Juli, los resultados han sido limitados, lo que sugiere que las estrategias actuales no abordan adecuadamente los factores de riesgo específicos.

En este contexto, se necesita un modelo predictivo efectivo para enfrentar el problema de manera más precisa e inmediata. El modelo de regresión logística propuesto permite identificar tempranamente a los niños en riesgo de anemia, optimizando la asignación de recursos y mejorando la eficacia de las intervenciones.

1.2. FORMULACIÓN DEL PROBLEMA

1.2.1. Problema general

¿Cuál es el modelo de regresión logística que mejor predice el riesgo de anemia en niños menores de 5 años en la Red de Salud Chucuito - 2023?

1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. Objetivo general

Determinar un modelo predictivo de regresión logística para predecir el riesgo de anemia en niños menores de 5 años en la Red de Salud Chucuito - 2023.

1.3.2. Objetivos Específicos.

- Identificar y analizar el conjunto de variables que influyen en el riesgo de anemia en niños menores de 5 años en la Red de Salud Chucuito - 2023.
- Desarrollar y entrenar un modelo de regresión logística para predecir el riesgo de anemia en los niños menores de 5 años en la Red de Salud Chucuito - 2023.
- Evaluar el modelo para predecir el riesgo de anemia en los niños menores de 5 años en la Red de Salud Chucuito - 2023.



1.4. HIPÓTESIS DE LA INVESTIGACIÓN

1.4.1. Hipótesis General

El modelo predictivo de regresión logística predice el riesgo de anemia en niños menores de 5 años en la Red de Salud Chucuito - 2023.

1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN

La anemia en la infancia es una preocupación en la salud pública que va en crecimiento en el Perú. Desde un punto de vista fisiológico, el principal impacto de la anemia es la reducción en el suministro de oxígeno a los tejidos. Esto puede generar efectos tanto agudos como crónicos en los infantes, que varían desde un deterioro leve y transitorio hasta un daño severo y permanente en su desarrollo físico (crecimiento) y mental (cognitivo) (Dávila Aliaga et al., 2019).

Según los datos suministrados por el Instituto Nacional de Estadística e Informática (INEI), en 2019 se determinó la prevalencia de anemia en los residentes de Perú mediante la Encuesta Demográfica de Salud Familiar (ENDES). El 69,9 % fue liderado por la región de Puno. Se observa una fluctuación desde 2018 hasta 2019, cuando Puno registró un incremento del 2,2 % en la prevalencia de anemia. En 2017, la zona ya contaba con un 67,7 %, y desde ese momento se ha observado un incremento constante. Esta situación es de preocupación para las autoridades encargadas del sistema de salud de la región, ya que se tiene poca sensibilización de las personas, no asimilan las estrategias, o no pueden acceder correctamente a los programas que plantea el gobierno para disminuirla o erradicarla progresivamente (INEI, 2019).

La regresión logística es el modelo más adecuado para nuestra investigación, basándonos en nuestro objetivo que es, predecir el riesgo de anemia en niños menores de



5 años en la Red de Salud Chucuito – 2023, ya que está diseñado para manejar variables dependientes binarias como la presencia o ausencia de anemia. Este modelo permite interpretar los coeficientes como cambios en la probabilidad del evento, facilitando la comprensión del impacto de variables independientes. Además, no requiere asumir una relación lineal entre variables, es resistente a problemas de heterocedasticidad y evita valores fuera del rango $[0,1]$. A diferencia de la regresión lineal, el análisis discriminante o modelos más complejos como árboles de decisión, la regresión logística ofrece simplicidad, interpretabilidad y capacidad predictiva ajustada a los objetivos del estudio (Hosmer et al., 2013b).

La regresión logística es un método ampliamente utilizado en la investigación y evaluación dentro del ámbito de la salud. Su principal utilidad radica en determinar qué variables explicativas afectan la ocurrencia de un evento y en qué medida lo hacen, lo que resulta especialmente relevante para identificar los factores que influyen en el desarrollo de una enfermedad (Hernández Svendblad, 2019a).

La regresión logística, más que un simple examen estadístico, es un instrumento que facilita la identificación tanto del riesgo multivariado como de la significancia estadística multivariada. Este enfoque permite orientar la investigación hacia la identificación de la causa principal de la enfermedad en un contexto explicativo o, alternativamente, utilizarse para predecir su aparición en estudios de carácter predictivo. La utilidad de la regresión logística ha sido ampliamente demostrada en diversas áreas de investigación dentro de las ciencias médicas, especialmente cuando se busca determinar el efecto de un conjunto de variables, consideradas potencialmente influyentes, sobre la ocurrencia de un proceso específico (Sánchez Martínez et al., 2021).



1.6. LIMITACIONES DE LA INVESTIGACIÓN

Se utilizó únicamente información proveniente del Sistema de Información del Estado Nutricional (SIEN) de la Red de Salud de Chucuito - 2023, lo cual podría limitar la aplicabilidad de los resultados a otras poblaciones o regiones con características distintas. Además, esta investigación, su diseño es transversal no experimental, lo que impide establecer relaciones causales entre las variables analizadas. Al ser una investigación correlacional, solo se pueden observar asociaciones entre los factores de riesgo y la anemia en un momento puntual, sin poder evaluar su evolución a lo largo del tiempo. Asimismo, el modelo no considera factores externos que pueden influir en el riesgo de anemia, como las condiciones ambientales (altitud o calidad del agua), otras enfermedades concurrentes. Estos elementos podrían afectar los resultados, y su omisión puede reducir la precisión del modelo predictivo en contextos diferentes al estudiado.



CAPÍTULO II

REVISIÓN DE LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

2.1.1. Antecedentes Internacionales

Hernández Svendblad (2019b), identificó las variables que incrementan significativamente la probabilidad de prescribir medicamentos off-label en la población pediátrica, utilizando un modelo de regresión logística binaria. La metodología se basó en un estudio transversal, descriptivo, llevado a cabo mediante encuestas tanto en formato físico como online. Se encuestó a pediatras, médicos de familia y médicos generales que atienden a niños. Inicialmente, se planteó la hipótesis de que la principal causa del uso off-label era el desconocimiento sobre la ficha técnica del medicamento. Sin embargo, la regresión logística binaria permitió identificar las principales variables predictoras de esta práctica. Los resultados destacaron dos factores clave que impulsan la prescripción off-label: la falta de alternativas terapéuticas adecuadas para niños y el desconocimiento sobre el uso off-label en la prescripción. El estudio concluye que el modelo de regresión logística binaria; proporcionó un sólido soporte estadístico a los resultados que otros investigadores han reportado en investigaciones previas sobre el uso off-label de medicamentos en la población pediátrica. Abordar ambos factores es factible y necesario para reducir esta práctica, considerando que es una responsabilidad compartida garantizar que los tratamientos médicos en niños sean seguros, eficaces y de calidad.



Sánchez Martínez et al (2021), plantearon como objetivo de la investigación, prever la probabilidad de que la población atendida en un Consultorio Médico en Cuba desarrollara DM2. Se llevó a cabo un estudio observacional, transversal, retrospectivo y analítico con 628 pacientes, utilizando regresión logística binaria en SPSS 25. Los resultados mostraron que la predisposición genética y la hipertensión arterial aumentaron significativamente el riesgo de DM2, mientras que la edad y el sexo no presentaron asociaciones significativas. Como conclusión, se demostró que los individuos con predisposición genética e hipertensión arterial presentaban un riesgo elevado de desarrollar DM2, en contraposición a la edad y el género, cuya correlación no resultó ser significativa. La regresión logística binaria facilitó la identificación del riesgo multivariado y la relevancia estadística, dirigiendo el estudio hacia la auténtica causa de la enfermedad tanto en el ámbito explicativo como predictivo. Se determinó que la regresión logística resulta sumamente beneficiosa en su uso en diferentes campos de la investigación médica cuando se requiere determinar el impacto de un grupo de variables que podrían influir en la manifestación de un proceso particular.

Bermeo E et al (2012), se centraron en analizar la utilidad de los criterios morfológicos de Toki para el diagnóstico de lesiones por VPH en el cuello uterino, en el diagnóstico de lesiones orales. Su metodología fue un estudio retrospectivo de 355 lesiones orales por VPH. La evaluación patológica se realizó tres veces por el mismo examinador con un coeficiente Kappa de 0.780 ($p = 0.01$). Los criterios de Toki asignaron una puntuación de 1 a la presencia de acantosis, 1 a la disqueratosis, 4 a la coilocitosis, 1 a la hiperplasia basal, 2 a las células bi o multinucleadas y 1 a la papilomatosis. La información se analizó utilizando c^2 y



regresión logística binaria. Como resultado obtuvieron que el 58% de los casos (207) correspondieron a verruga vulgar (VV), el 23% (80) a hiperplasia epitelial multifocal (HEM) y el 19% (68) a condiloma acuminado (CA). Según los criterios de Toki, 76/80 lesiones de HEM, 195/207 de VV y 67/68 de CA, tuvieron una puntuación igual o mayor a seis. Una regresión logística mostró que la disqueratosis sugiere la presencia de HEM. La presencia de acantosis y disqueratosis en lesiones orales por VPH no son criterios útiles para el diagnóstico de VV. No se encontraron asociaciones significativas con CA. Concluyeron que los criterios morfológicos estudiados no predicen su posible etiología viral y no permiten un diagnóstico diferencial de las lesiones. Por lo tanto, se requieren estudios de reacción en cadena de la polimerasa para identificar VPHs de alto riesgo o lesiones virales de transmisión sexual.

Rodriguez Rueda et al (2017), se plantearon como objetivo del estudio diseñar un modelo predictivo para el daño hepático en pacientes sépticos. Para realizar el estudio, se llevó a cabo una investigación observacional, de tipo caso-control, retrospectiva y orientada al análisis de crecimiento. Se incluyeron 508 pacientes que fallecieron en la Unidad de Cuidados Intensivos con evidencias clínicas y anatómicas de sepsis y que satisficieron los criterios de intencionalidad, entre enero 2006 y diciembre 2015 del Hospital "Manuel Fajardo Rivero". De este grupo, se seleccionaron 100 casos y controles. Los resultados, tras aplicar un análisis de regresión logística binaria, mostraron que se incorporaron las siguientes variables: Síndrome de disfunción multiorgánica, bilirrubina directa, fosfatasa alcalina, colesterol total, creatinina, y normalización de la razón internacional y plaquetas. El test de Hosmer-Lemeshow dio como resultado un valor de 1,867 con $p=985$. Se observó una sensibilidad del 57,69% y una



especificidad del 100%, con una predicción positiva del 100% y una predicción negativa del 68,57%, alcanzando un índice de validez del 78,00. La zona bajo la curva ROC alcanzó los 0,922, presentando una significancia estadística de 0,000. Para concluir, el modelo evidenció una excelente habilidad para discriminar y resultó ser un predictor eficaz de daño hepático en pacientes con sepsis.

León Sánchez y Linares Guerra (2014), se plantearon como objetivo utilizar la regresión logística binaria; como herramienta matemática para anticipar el conteo de células CD4+; en individuos con VIH/sida durante un año de monitorización. Se realizó un estudio con 118 individuos con VIH/sida (al comienzo y un año posteriormente), los cuales fueron divididos en cuatro subgrupos basándose en la administración de antirretrovirales y el conteo inicial de CD4. Se emplearon variables de nutrición, bioquímicas y antropométricas, junto con los gradientes entre ambas observaciones. Se usó la regresión logística binaria, tomando el recuento de CD4 como variable dependiente, que se emplea como indicador inmunológico en la terapia antirretroviral. Los modelos logísticos adquiridos se modificaron de acuerdo al subgrupo y se contrastaron en cuanto a su precisión y bondad de ajuste. El procedimiento de selección gradual estableció que únicamente el gradiente del índice de masa corporal (IMC) era una variable autónoma. La modificación del modelo facilitó la creación de una ecuación matemática para cada subgrupo. Los colectivos con valores iniciales de CD4 inferiores a 350, incluso con incremento en el IMC, persistieron en el riesgo de deterioro inmunológico tras un año. Todos los modelos mostraron una mayor especificidad en comparación con la sensibilidad, siendo los modelos II y III los más eficaces, tal como se manifiesta en las probabilidades predictivas de las curvas ROC. Se determinó que el gradiente del IMC durante un año podría



emplearse en un modelo logístico; para estimar la necesidad de tratamiento antirretroviral en individuos con VIH/sida.

Orraca Castillo et al (2021), el enfoque principal fue desarrollar un modelo de regresión logística binaria para identificar el riesgo de asma en pacientes pediátricos. La metodología empleada consistió en un estudio observacional, analítico y de casos y controles, llevado a cabo en pacientes asmáticos de la provincia de Pinar del Río, Cuba, entre los años 2015 y 2019. La muestra constaba de 735 casos, mostrando una relación de 1:2 entre los casos y los controles. Se desarrolló un modelo de regresión logística binaria utilizando variables genéticas, infecciosas y ambientales. Los hallazgos revelaron una relación estadística significativa entre los valores observados y las predicciones de la variable dependiente ($p=0,602$). El R^2 de Nagelkerke indicó que los predictores utilizados explicaban el 87% de la variabilidad de la variable. El modelo desarrollado fue capaz de clasificar correctamente el 94% de los casos, con una sensibilidad del 93% y una especificidad del 95%. El modelo se construyó utilizando 11 variables, de las cuales cuatro estaban relacionadas con antecedentes familiares de asma, cuatro con factores infecciosos y tres con factores ambientales. Los investigadores concluyeron que el análisis multivariado permitió crear un modelo para estimar la probabilidad de asma mediante tablas de riesgo en Pinar del Río, lo que facilita la implementación de la medicina personalizada y fomenta la medicina preventiva en la atención primaria de salud.

2.1.2. Antecedentes Nacionales

Navarro Chávez y Oliva Vera (2018), propusieron como meta principal, detectar los elementos asociados a la presencia de anemia en niños menores de



cinco años en Perú durante el año 2018. En esta investigación, se tomaron en cuenta 21 covariables que se examinaron mediante el método multivariante de regresión logística binaria. Se empleó un grupo de 16,499 niños de menos de cinco años, recolectados de la base de datos de la Encuesta Demográfica y de Salud Familiar (ENDES) del 2018. La metodología de la investigación fue cuantitativa y de tipo transversal. Los descubrimientos señalaron que, en Perú, la incidencia de anemia en niños menores de 5 años es del 36,4%, con un 43,2% en áreas rurales y un 33,7% en áreas urbanas. El modelo adquirido categorizó adecuadamente el 69,5% de los casos de anemia en menores de 5 años en todo el país, empleando 14 covariables: Los factores independientes relacionados con el riesgo de anemia incluyen la zona de residencia (OR=1.144), la pertenencia a la primera (OR=1.420) y segunda región natural (OR=1.278), el nivel educativo materno (OR=0.884), la edad de la madre (OR=0.702), la lengua materna (OR=1.181), el nivel de riqueza del hogar (OR=0.834), el sexo del niño (OR=0.824), la edad en meses (OR=0.313), la presencia de episodios recientes de diarrea (OR=1.181), el número de consultas prenatales durante el embarazo (OR=0.852), el consumo de suplementos de hierro durante la gestación (OR=0.859), la presencia de anemia materna durante el embarazo (OR=1.380), la administración de suplementos de hierro en los últimos 12 meses (OR=1.156) y el uso de antiparasitarios en los últimos seis meses (OR=0.824).

Calle Núñez et al (2017), identificaron los factores relacionados con la evolución de infecciones urinarias provocadas por la *Escherichia coli*, generadora de betalactamasas de espectro extendido (BLEE). La investigación se llevó a cabo mediante un estudio de casos, así como también de controles en el Hospital Cayetano Heredia, con una muestra de 150 casos y controles. Los casos en el



estudio correspondieron a pacientes con urocultivo positivo para *E. coli* productora de BLEE, mientras que los controles fueron aquellos con urocultivo positivo para *E. coli* no productora de BLEE. Se realizó un análisis bivariado y una regresión logística binaria para identificar los factores asociados. Los resultados mostraron que, según la regresión logística, los factores relacionados con el riesgo de infecciones urinarias por *E. coli* BLEE incluyen: ser hombre (OR 5,13; IC 95%: 2,37–11,07), tener más de 45 años (OR 2,65; IC 95%: 1,61–4,38) y haber sido hospitalizado previamente (OR 2,57; IC 95%: 1,39–4,75). La investigación concluyó que los pacientes masculinos, mayores de 45 años y con antecedentes recientes de hospitalización tienen un mayor riesgo de desarrollar infecciones urinarias por *E. coli* BLEE, lo cual debe considerarse en el manejo empírico de estos casos.

Bardales Aguirre et al (2021), analizaron los factores biológicos, sociales y clínicos de riesgo de mortalidad en pacientes hospitalizados con COVID-19 en el distrito de Trujillo, Perú. La investigación utilizó un diseño descriptivo, cuantitativo, correlacional, retrospectivo y de corte transversal, basado en datos obtenidos del sistema del Ministerio de Salud. La muestra consistió en 64 pacientes, recolectados entre marzo y mayo de 2020. Los resultados mostraron que el 85,71% de los fallecidos eran hombres, con una edad promedio de 64,67 años, y la ocupación más frecuente fue la de jubilados (28,57%). Los síntomas más prevalentes en los fallecidos fueron dificultad respiratoria (90,48%), malestar general (57,14%), fiebre (80,95%) y tos (52,38%). Los signos más recurrentes fueron disnea y auscultación pulmonar anormal (47,62%). En cuanto a comorbilidades, el 42,86% de los pacientes presentaban enfermedades cardiovasculares y el 14,29% padecían diabetes. El modelo de regresión logística



binaria empleado para predecir la defunción señaló como factores de riesgo la edad, el sexo, la presencia de tos, la dificultad respiratoria y la diabetes. Este modelo demostró una alta precisión al clasificar correctamente el 90,6% de los casos y explicar una proporción significativa de la variabilidad en la mortalidad de los pacientes hospitalizados con COVID-19. Se concluyó que el modelo es útil para identificar los factores de riesgo relacionados con la mortalidad en estos pacientes.

Cueva Román (2023), plantearon como objetivo del estudio el desarrollo de un modelo de regresión logística binaria efectivo, enfocado en los factores sociodemográficos, con el fin de predecir el abandono del tratamiento de tuberculosis en pacientes peruanos. Se trató de un estudio observacional, transversal y analítico, en el que participaron 362 sujetos distribuidos en un grupo de control y un grupo de casos. La información se recolectó mediante una ficha específica, y el análisis se llevó a cabo empleando los programas estadísticos RStudio y IBM SPSS Statistics. Los resultados mostraron que los hombres tienen una mayor incidencia de casos de tuberculosis pulmonar, lo cual eleva el riesgo de experimentar factores asociados. Además, se determinó que el periodo de hospitalización y la falta de datos exactos acerca de los exámenes de laboratorio son elementos que elevan la probabilidad de deserción del tratamiento. Asimismo, el consumo de sustancias alucinógenas también eleva la probabilidad de abandono del tratamiento. Además, se destacó la carencia de evaluaciones con un enfoque predictivo, ya que la mayoría de las investigaciones previas se han centrado en un enfoque meramente explicativo.



2.2. MARCO TEÓRICO

2.2.1. Anemia

La anemia es una condición médica que se evidencia por una baja cantidad y/o el tamaño de los eritrocitos, o en la concentración de hemoglobina en la sangre, lo que afecta la habilidad de la sangre para transportar oxígeno de manera correcta a los tejidos y órganos del cuerpo. Esto puede ocasionar una variedad de síntomas, como fatiga, palidez, debilidad y dificultad para respirar, dependiendo de la gravedad y la causa subyacente de la anemia. La anemia, es un indicador de mal estado en nutrición y de salud (OMS, 2017).

La anemia, es un estado en el que el número de eritrocitos no es suficiente para cubrir las demandas de oxígeno del cuerpo. Estos requerimientos pueden variar según factores como la edad, el sexo, la altitud, el consumo de tabaco y las diferentes etapas del embarazo. Si bien la escasez de hierro es la principal causa de anemia, también pueden afectar otras carencias, como la de folato, vitamina B12 y vitamina A, procesos inflamatorios (agudos o crónicos), infecciones parasitarias y enfermedades hereditarias o adquiridas que interfieren en la producción o supervivencia de los eritrocitos o la síntesis de hemoglobina.

Aunque la concentración de hemoglobina es una medida importante, no es suficiente por sí sola para diagnosticar la carencia de hierro, ya que no todas las anemias son causadas por ferropenia. Es fundamental realizar otros exámenes para determinar la causa exacta de la anemia y su tratamiento adecuado. La prevalencia de la anemia es un indicador clave de salud y, cuando se combina con otras mediciones del estado nutricional relacionado con el hierro, la concentración de

hemoglobina puede ofrecer datos sobre la gravedad de la deficiencia de hierro (OMS, 2011).

La anemia infantil se caracteriza por una reducción de glóbulos rojos y hemoglobina, lo que afecta el transporte de oxígeno en el cuerpo. Comúnmente causada por deficiencia de hierro o problemas de absorción, también puede estar relacionada con ciertas enfermedades. Los síntomas incluyen palidez, cansancio, y falta de energía, y pueden manifestarse en un tono blanquecino en la boca o lengua (Clínica Mayo, 2018).

Tabla 1

Criterios OMS y Rangos Normales para Definir Anemia en Niños Menores de 5 Años

Grupo de edad de niños	Rango normal de hemoglobina (g/dl)	Anémico si la Hb es menor de: (g/dl)
0 - 2 meses	13.5 - 18.5	13.5
2 - 6 meses	9.5 - 13.5	9.5
6 meses - 6 años	11.0 - 14.0	11.0

Fuente: (OMS, 2011b)

Tabla 2

Valores Normales de Concentración de Hemoglobina y Grados de Anemia en Niñas y Niños 6 a 59 Meses (hasta 1000 msnm)

Población	Normal (g/dl)	Anemia por niveles de hemoglobina (g/dl)		
		Leve	Moderada	Severa
Niños de 6 a 59 meses	> 11.0	10.0 - 10.9	7.0 - 9.9	<7.0

Fuente: (OMS, 2011b)



2.2.2. Factor de riesgo

La Organización Mundial de la Salud (OMS, 2016), se entiende por factor de riesgo a toda circunstancia, comportamiento o exposición de un individuo que aumenta la posibilidad de contraer una enfermedad o experimentar una lesión. Entre los factores de riesgo más significativos se incluyen el bajo peso en la infancia, las conductas sexuales no seguras, la hipertensión, el consumo de tabaco y alcohol, el agua contaminada, la falta de saneamiento adecuado y las deficiencias en higiene.

Escalante (2004). Un factor de riesgo se refiere a ciertas variables que están vinculadas a una mayor probabilidad de desarrollar una enfermedad; sin embargo, por sí solas no son suficientes para causarla.

- Factores de riesgo no modificables: Son aquellos que están relacionados con las características propias de cada persona, tales como edad, sexo y predisposición genética.
- Factores de riesgo modificables: Son aquellos que pueden alterarse mediante intervenciones de prevención primaria. A través de acciones preventivas, como dejar de fumar, aumentar la actividad física y controlar la obesidad, estos factores pueden reducirse o eliminarse.

2.2.2.1. Factores determinantes de la anemia

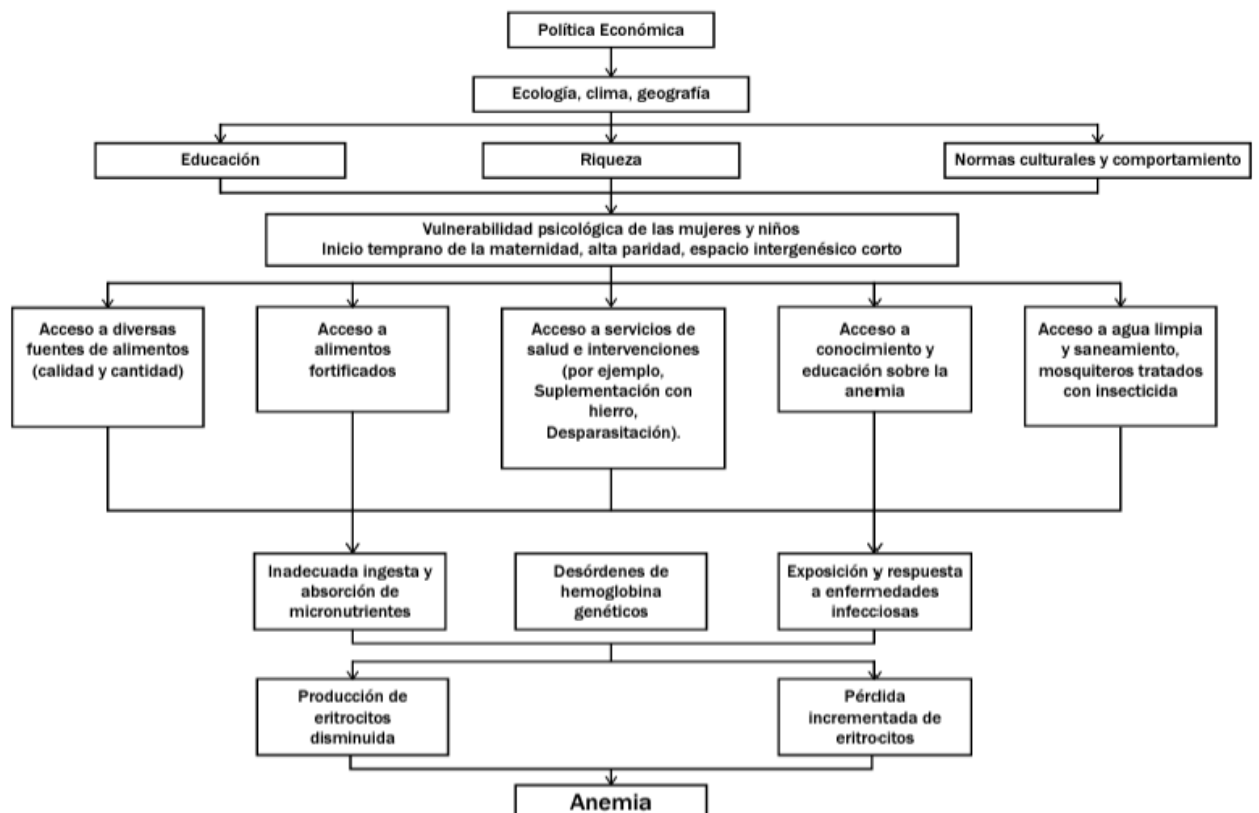
Las causas inmediatas de la anemia abarcan la ingesta inadecuada de hierro y otros micronutrientes esenciales en la dieta, lo que dificulta la correcta producción de glóbulos rojos y hemoglobina. Además, las altas tasas de enfermedades como diarrea, parasitosis y malaria son factores inmediatos, relacionados con prácticas de higiene inadecuadas, la falta de

un correcto lavado de manos y el acceso limitado a agua potable y servicios básicos de saneamiento. Asimismo, se ha demostrado que las vitaminas A, B2, B6, B12 y el ácido fólico son esenciales para la formación de glóbulos rojos en la médula ósea. Las vitaminas A, C y riboflavina favorecerían la absorción del hierro en el intestino, facilitando su movilización desde las reservas, mientras que las vitaminas C y E desempeñan una función antioxidante que protege los glóbulos rojos (MINSA, 2017).

El Plan Nacional toma en cuenta también el modelo causal (Figura 1) del Programa Articulado Nutricional para la reducción de la Desnutrición Crónica Infantil:

Figura 1

Modelo Causal



Fuente: (MINSA, 2017)

La anemia se relaciona con diversos factores sociodemográficos y nutricionales. Un estudio encontró que la frecuencia de anemia en niños de 1 a 5 años en Perú fue del 38,5%. Entre los factores asociados de manera independiente con la anemia se identificaron: vivir en un quintil bajo de riqueza (RP(a): 1,23; IC95%: 1,0-1,4), que la madre carezca de educación o solo cuente con educación primaria (RP(a): 1,25; IC95%: 1,0-1,5), que la madre tenga menos de 19 años (RP(a): 1,34; IC95%: 1,1-1,7), un parto fuera de una institución de salud (RP(a): 1,24; IC95%: 1,1-1,5), la falta de uso de antiparasitarios (RP(a): 1,13; IC95%: 1,0-1,3) y residir a una altitud de 4000 msnm o más (RP(a): 1,45; IC95%: 1,2-1,8). Concluyendo que existen factores sociodemográficos y nutricionales asociados a dicha condición (Al Kassab Córdova et al., 2020).

2.2.3. Modelo de regresión logística

La Regresión Logística se emplea para analizar la relación entre una variable dependiente dicotómica (cualitativa) y una o más variables independientes o explicativas, que pueden ser tanto cualitativas (como las sociodemográficas) como cuantitativas. El objetivo principal es estimar de manera ajustada la probabilidad de que ocurra un evento en función de una o más variables independientes (García Pérez et al., 2010).

Uno de los métodos estadísticos más utilizados en la investigación en salud es la regresión logística. Esta técnica se aplica para modelar la relación entre una variable de resultado binaria y varias variables predictoras, que pueden ser tanto categóricas como continuas. Mediante este método, los investigadores pueden analizar y predecir eventos binarios, como la presencia o ausencia de una

enfermedad, considerando diversos factores clínicos y biológicos. La importancia y utilidad de la Regresión Logística en la investigación clínica se evidencian en la amplia cantidad de literatura científica dedicada a este método (Grigore et al., 2020).

Además, la Regresión logística se ha incorporado de manera exitosa al campo de la inteligencia artificial (IA), lo que ha incrementado significativamente su utilidad e influencia en la investigación y la práctica clínica contemporánea (Issaiy et al, 2023).

Según (Cox, 1958), la regresión logística es un método estadístico utilizado para modelar la relación entre una variable dependiente binaria, que tiene dos posibles resultados (como "sí" o "no"), y una o más variables independientes. Este modelo permite estimar la probabilidad de que ocurra uno de los dos resultados, transformando la relación entre las variables mediante una función logística. Se utiliza ampliamente en situaciones donde se desea predecir la ocurrencia de un evento con base en diferentes factores, garantizando que las probabilidades estimadas estén en el rango de 0 a 1.

El modelo de regresión logística puede formularse como:

$$\text{Log}\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

Donde:

P = Probabilidad de ocurrencia del evento.

Esta probabilidad se obtiene:



$$p = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}$$

Donde:

$$z = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

La variable dependiente, en el modelo lineal de probabilidad, solo puede asumir dos valores:

Si se hace:

$$\text{Log}\left(\frac{p}{1-p}\right) = Y_i$$

- $Y_i = 1$, cuando ocurre determinado caso.
- $Y_i = 0$, si el caso no ocurre.

La regresión logística tiene como objetivo predecir la probabilidad de ocurrencia de un evento en comparación con la probabilidad de que este no ocurra; utilizando las variables independientes. La validez del modelo se evalúa mediante pruebas estadísticas que miden su ajuste.

2.2.3.1. Significación de chi-cuadrado de Pearson χ^2

El coeficiente de Chi cuadrado de Pearson se emplea para contrastar las frecuencias detectadas y las proyectadas; en un contexto binomial, y se establece a través de la fórmula siguiente:

$$\chi^2 = \sum_{i=1}^k \frac{(F_{O_i} - F_{E_i})^2}{F_{E_i}}$$

Donde:

- F_0 : La frecuencia observada en el grupo i (el número de veces que se observó el evento en ese grupo).
- F_e : La frecuencia esperada en el grupo i (la cantidad esperada de eventos si el modelo predice bien)

Esta ecuación calcula la diferencia cuadrática entre la frecuencia observada y la frecuencia esperada, y la divide por la frecuencia esperada. Esto permite comparar las distancias en la misma escala de las cantidades.

Cuando el valor de significación es < 0.05 , indica que el modelo posee capacidad explicativa. Esto significa que las variables independientes son significativas para predecir o explicar el evento o la variable dependiente.

2.2.3.2. R-Cuadrado de Cox y Snell y R-Cuadrado de Nagelkerke

El R-Cuadrado de Cox y Snell; es un coeficiente de determinación utilizado para estimar qué porcentaje de la variabilidad de la variable dependiente puede ser explicado por las variables independientes, y se determina así:

$$R^2 = 1 - \left(\frac{L(b_0)}{L(b_0, b_1, \dots, b_k)} \right)^{2/n}$$

Donde:

- $L(b_0)$ = Verosimilitud con solo la constante.
- $L(b_0, b_1, \dots, b_k)$ = Verosimilitud con más variables.

- n = Tamaño de muestra o número total de observaciones en el conjunto de datos.

Este coeficiente está limitado por el rango $0 \leq R^2 \leq 1$, lo que significa que no puede llegar a ser igual a 1. El R-Cuadrado de Nagelkerke es una versión modificada del R-Cuadrado de Cox y Snell. Este ajusta la escala del estadístico que cubre completamente el rango de 0 a 1. Se define de la siguiente forma:

$$R^2 = \frac{R^2}{R_{max}^2}$$

Donde:

$$R_{max}^2 = 1 - (L(b_0))^{2/n}$$

- R^2 : Coeficiente de determinación, indica qué proporción de la varianza de la variable dependiente es explicada por el modelo.
- R_{max}^2 : Ajuste teórico máximo de R^2 , teniendo en cuenta la verosimilitud del modelo ($L(b_0)$) y el número de observaciones (n).
- $L(b_0)$: Representa la verosimilitud o error del modelo ajustado.
- n : Número de observaciones en los datos, usado para normalizar el ajuste.

De esta manera se puede alcanzar el valor de 1.

Aunque estos coeficientes intentan medir la variabilidad explicada, sus valores suelen ser mucho más bajos que los obtenidos en la regresión lineal. Por lo tanto, es necesario complementarlos con otras métricas de bondad de ajuste.

2.2.3.3. Porcentaje global correctamente clasificado

Presenta la cantidad total de casos que el modelo consigue anticipar de forma acertada. Mediante el uso de la ecuación de regresión y la información recogida, se puede predecir el valor de la variable dependiente basándose en el valor observado; si la predicción es precisa, el caso se categoriza de manera adecuada.

$$\text{Porcentaje Global} = \frac{\sum(C_{\text{correctos}} + C_{\text{incorrectos}})}{\text{Total de casos}}$$

Donde:

- $C_{\text{correctos}}$ = Casos correctamente clasificados
- $C_{\text{incorrectos}}$ = Casos incorrectamente clasificados

Un modelo se considera adecuado si clasifica correctamente más del 50% de los casos. De lo contrario, se debe optar por seleccionar nuevas variables independientes.

2.2.3.4. Evaluadores del modelo de regresión logística

Una vez seleccionado el modelo predictivo definitivo, es fundamental proceder con su validación, lo que consiste en verificar si el modelo; predice de manera correcta la variable dependiente. La validez hace referencia a la medida en que estas predicciones coinciden con los resultados observados, y consta de 2 componentes: la calibración y la discriminación. La calibración se enfoca en comparar la cantidad de eventos predichos con los eventos observados dentro de un grupo de individuos.

La prueba estadística que evalúa la calibración es la de Hosmer-Lemeshow (Hosmer et al, 2013a) . Este es un test de bondad de ajuste que se utiliza para evaluar la diferencia entre los valores observados y los esperados. Su cálculo se realiza mediante la siguiente fórmula:

$$H = \sum_{j=1}^{10} \frac{(o_j - n_j \bar{p}_j)^2}{n_j \bar{p}_j (1 - \bar{p}_j)}$$

Donde:

- n_j = El número total de observaciones en el grupo j (la cantidad total de niños en ese grupo).
- o_j = El número de eventos observados en el grupo j (es decir, cuántas veces ocurrió el evento que estás modelando, en este caso, cuántos niños fueron anémicos en ese grupo).
- \bar{p}_j = La probabilidad promedio predicha de que ocurra el evento en el grupo j (la probabilidad predicha por el modelo de que los niños en ese grupo sean anémicos).
- $1 - \bar{p}_j$ = La probabilidad de que el evento no ocurra en el grupo j (la probabilidad predicha por el modelo de que los niños no sean anémicos).

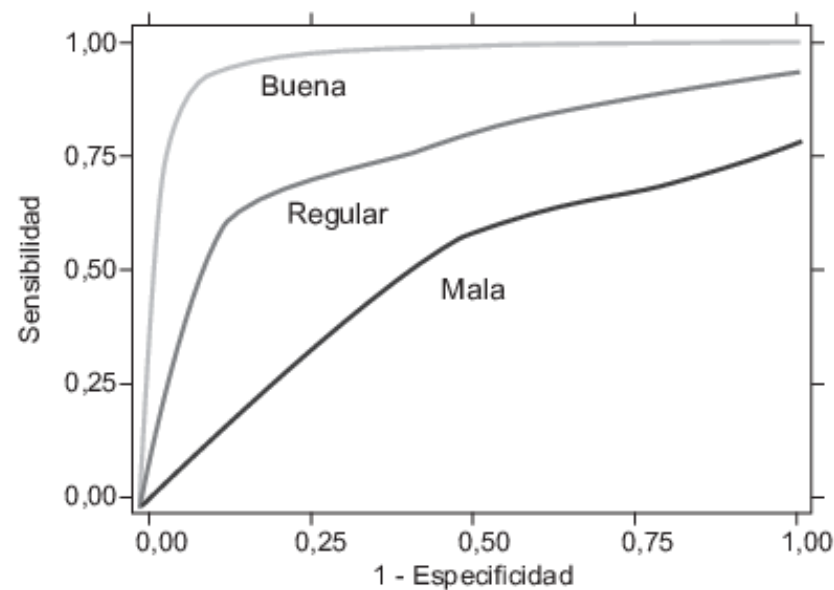
La hipótesis de este test sostiene que un valor superior señala una mayor eficacia del modelo si p es superior a 0.05, el modelo se considera apropiado, lo que indica que no hay discrepancias significativas entre los resultados obtenidos y los estimados. Para valorar la discriminación, se emplea el área bajo la curva ROC (Característica Operacional del

Recibidor), la cual ilustra la correlación entre la sensibilidad (verdaderos positivos) y la especificidad (falsos positivos) en distintos puntos de corte. En una prueba perfecta, la curva tendría un único punto en (0,1), con sensibilidad y especificidad de 1 en cualquier corte. Si el examen no resultara útil, la curva seguiría la diagonal que se extiende de (0,0) a (1,1). Por lo general, las curvas logradas tienden a ser intermedias. Se utiliza el área bajo la curva para evaluar la calidad de la prueba, con valores que oscilan entre 1 (prueba perfecta) y 0,5 (prueba inútil).

La curva ROC mide el equilibrio entre la sensibilidad y la especificidad, mostrando cómo varían estas métricas a medida que se ajusta el umbral de clasificación del modelo. No mide directamente la precisión, pero permite evaluar el rendimiento general del modelo en términos de discriminación.

Figura 2

Curvas ROC



Fuente: Hospital Universitario Ramón y Cajal (2014)



Para poner a prueba la hipótesis, se utiliza el estadístico F, el cual permite medir la capacidad de un conjunto de variables predictoras para explicar la variación en la variable dependiente. Se calcula ese estadístico de la siguiente manera:

$$F = \frac{\frac{SCR}{k}}{\frac{SCE}{n - k - 1}}$$

- $SCR = \sum$ de los cuadrados de la regresión
- $SCE = \sum$ de los cuadrados de los residuos
- $N =$ Número total de observaciones de la muestra
- $k =$ Grados de libertad.

2.3. DESCRIPCIÓN DE VARIABLES

Ya que el estudio busca examinar el efecto de la inclusión de variables en el comportamiento de la variable anemia, se ha empleado la información suministrada por el Sistema de Información del Estado Nutricional del Niño Menor de Cinco Años y Gestantes que acuden a los establecimientos de salud (SIEN) de la Red de Salud Chucuito – Puno, 2023.



CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. UBICACIÓN GEOGRÁFICA DE LA INVESTIGACIÓN

La localidad de Chucuito, situada en el sur de Perú, es una de las trece provincias pertenecientes al departamento de Puno. Al norte se encuentra la provincia de Yunguyo y el lago Titicaca, al este se encuentra la provincia de Yunguyo y Bolivia, al oeste se encuentra la provincia de El Collao y al sur está el Departamento de Tacna. Con una superficie de 3,978.13 km², se calcula que su población es de 111,000 personas.

La provincia de Chucuito se compone de siete municipios: Juli, Desaguadero, Huanculi, Kelluyo, Pisacoma, Pomata y Zepita. La ciudad de Juli es su capital, situada a una altitud de 3,868 metros por encima del nivel del mar.

3.2. ÁMBITO DE LA INVESTIGACIÓN

La Red de Salud Chucuito es una entidad descentralizada dependiente de la Dirección Regional de Puno, responsable de brindar servicios de salud a la población de la provincia de Chucuito, que abarca su área de jurisdicción. En el ámbito de la Nutrición, al igual que en otras Redes de Salud, se emplea el Sistema Informático SIEN, el cual será el enfoque principal de la investigación, junto con el área de Estadística, que se encarga de la gestión de este sistema.

3.3. TIPO Y DISEÑO DE INVESTIGACIÓN

La investigación fue de tipo aplicada porque busca soluciones prácticas a problemas específicos, utilizando el conocimiento existente para generar impacto directo



en la sociedad o el sector productivo, vinculando teoría y práctica (Hernández Sampieri et al., 2014).

La investigación realizada corresponde a un estudio correlacional con un diseño no experimental transversal-analítico (Hernández Sampieri et al., 2010), ya que se utilizaron herramientas y evidencias teóricas vinculadas a los factores de riesgo asociados a la anemia en niños menores de 5 años.

3.4. POBLACIÓN

La población estuvo conformada por 1931 niños menores de 5 años registrados en el SIEN de la Red de Salud Chucuito – Puno, 2023.

El uso de un censo completo en lugar de una muestra es respaldado por la literatura cuando la población es de tamaño manejable y los recursos lo permiten. Según (Hernández Sampieri et al., 2014), el enfoque de censo es preferible en estudios descriptivos y correlacionales cuando se busca una alta precisión en los resultados y cuando la población no es excesivamente grande. Al incluir toda la población, se elimina el error de muestreo, lo que aumenta la precisión y la validez de los resultados obtenidos.

3.5. INSTRUMENTO DE RECOLECCIÓN Y PROCESAMIENTO DE DATOS

El método utilizado fue el análisis de los registros de datos (Anexo 1) proporcionados por el Sistema de Información del Estado Nutricional del Niño Menor de Cinco Años y Gestantes (SIEN), perteneciente a la Red de Salud Chucuito – Puno, correspondiente al año 2023. La información recopilada a nivel local, distrital y provincial es crucial para la toma de decisiones y la planificación de intervenciones, particularmente en el ámbito local.



3.5.1. Procedimiento de recolección de datos

Se logró una coordinación con el jefe del departamento de Nutrición y el responsable de Estadística de la Red de Salud Chucuito, lo que facilitó la obtención de la autorización necesaria para ingresar y examinar la base de datos del Sistema SIEN.

3.5.2. Plan de procesamiento y análisis de datos

Tras la petición de la información, se llevó a cabo el proceso de codificación y traslado a una base de datos, empleando los programas Excel 2016 y SPSS versión 27 para su tabulación. Luego, se llevó a cabo un análisis descriptivo, seguido del análisis de regresión logística, para luego proceder a la interpretación de los resultados. Estos fueron organizados en tablas y figuras estadísticas, en línea con los objetivos de la investigación.

3.5.2.1. Análisis estadístico

La base de datos utilizada en esta investigación, se elaboró con la información obtenida del SIEN de la Red de Salud Chucuito - 2023. Cada registro en la base de datos representa a un niño, y cada columna refleja una variable particular considerada en el análisis (Anexo 1).

Primeramente, se llevó a cabo el análisis descriptivo de los datos y una revisión de los antecedentes de la investigación, esto con el fin de identificar los factores que inciden en el riesgo de anemia.

Posteriormente, se aplicó un Modelo de Regresión Logística para evaluar la relación entre cada factor y la variable dependiente, "presencia de anemia en el niño". Este modelo permitió determinar la fuerza de

asociación (odds ratio, OR) e intervalos de confianza, así como la significancia estadística (p-valor) mediante la prueba Wald y la prueba de Hosmer-Lemeshow. El objetivo del análisis fue desarrollar un modelo de regresión logística que estime la probabilidad de que un niño menor de 5 años presente anemia, centrándose en explicar la variable categórica con dos posibles resultados: anemia presente o anemia ausente.

3.6. OPERACIONALIZACIÓN DE VARIABLES

Los factores seleccionados y su correspondiente operacionalización de variables, provenientes de la base de datos del SIEN, se presentan en la Tabla 3. Se excluyeron algunas variables, ya que se consideran como condiciones de la investigación y no son relevantes para el análisis específico.

Tabla 3

Operacionalización de Variables

Variable	Dimensión	Indicadores	Medición	
FACTORES DE RIESGO	1. Factores Demográficos	Distrito	1. Desaguadero	
			2. Huacullani	
			3. Juli	
			4. Kelluyo	
			5. Pisacoma	
			6. Pomata	
			7. Zepita	
	2. Factores biológicos		Altitud_Loc	m.s.n.m
			Hemoglobina	g/dl
			Sexo	1. Femenino



		2. Masculino	
	Edad	Meses	
	Peso	Kilógramos	
	Talla	Centímetros	
3. Factores de intervención	N_CCred	Nº de Controles	
	N_Sachets	Nº de Sachets	
	N_Consejerias	Nº de Consejerías	
	N_Sesiones	Nº de Sesiones	
	N_Visitas	Nº de Visitas	
4. Factores sociales	JUNTOS	0. No 1. Si	
	SIS	0. No 1. Si	
	PIN	0. No 1. Si	
	VARIABLE	Presencia de anemia en Diagnóstico de	0. No anémico
	DEPENDIENTE	niños(as) menores de 5 años. anemia	1. Anémico

Fuente: Elaboración propia.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. IDENTIFICACIÓN DE VARIABLES

4.1.1. Prevalencia de anemia

Se utilizó el conjunto completo de datos proporcionados por la SIEN-DIRESA, 2023 para calcular la prevalencia de anemia en niños menores de 5 años, considerando un total de 1931 niños registrados en la Red de Salud Chucuito.

Tabla 4

Prevalencia de Anemia Según Sexo, en los Niños Menores de 5 años en la Red de Salud Chucuito – 2023

SEXO	ANEMIA				
	No anémico		Anémico		
	n	%	n	%	n
Femenino	170	48,7%	757	47,9%	927
Masculino	179	51,30%	825	52,1%	1004
TOTAL	349	18,1%	1582	81,9%	1931

Fuente: Elaboración propia.

Los resultados indican que la prevalencia de anemia en niñas; es del 47,9% y en niños del 52,1%. Estos porcentajes son significativos, destacándose que más niños presentan anemia en comparación con las niñas, lo que refleja la problemática de salud en la provincia de Chucuito. En términos generales, considerando ambos sexos, la prevalencia de anemia es del 81,9%.

4.1.2. Selección de variables

En la tabla 5, se observa que de las variables Edad, Sexo, Talla Hemoglobina, Distrito, Altitud_Loc, Juntos, Sis, Pin, Nccred, N_Visitas, N_Consejerías, N_Sesiones y N_Sachets (Anexo 3), solo son significativas para el modelo las siguientes variables: Distrito, Edad, Talla, Hemoglobina, N_Visitas, N_Sesiones, N_Sachets y Juntos.

Tabla 5

Variables que ingresan al Modelo

	B	Error estándar	Wald	gl	Sig.	Exp(B)
EDAD	,346	,137	6,402	1	,011	1,413
SEXO	,181	,213	,720	1	,396	1,198
PESO	-14,989	1861,453	,000	1	,994	,000
TALLA	,748	,308	5,899	1	,015	2,112
HEMOGLOBINA	1,211	,102	141,161	1	,000	3,356
DISTRITO	1,096	,102	115,713	1	,000	2,992
ALTITUD_LOC	-,384	,606	,401	1	,526	,681
JUNTOS	,831	,264	9,933	1	,002	2,296
SIS	,501	,420	1,426	1	,232	1,651
PIN	19,163	7126,990	,000	1	,998	210167481,150
N_CCRED	-,151	,308	,239	1	,625	,860
N_VISITAS	-,762	,300	6,446	1	,011	,467
N_CONSEJERÍA	-,178	,350	,260	1	,610	,837
N_SESIONES	,966	,296	10,666	1	,001	2,628
N_SACHETS	-3,497	1,382	6,404	1	,011	,030
Constante	9,852	1861,453	,000	1	,996	18988,812

Fuente: Elaboración propia.

4.2. ESTIMACIÓN DE MODELO

4.2.1. Selección de variables

Después de analizar y seleccionar los factores según su significancia, se llevó a cabo una estimación preliminar del modelo utilizando todas las variables

seleccionadas, evaluando su significancia en la determinación de la anemia en niños menores de 5 años. Para este análisis, se trabajó con toda la población.

En la tabla 6, se presenta la significancia de las variables distrito, edad, talla, hemoglobina, N_visitas, N_sesiones, N_sachets y Juntos en el nuevo modelo, y se puede observar que todas son significativas, lo que indica que influyen en el modelo. Es importante mencionar que se realizaron pruebas excluyendo estas variables, pero su rendimiento fue inferior, por lo que se decidieron mantener en el modelo.

Tabla 6

Variables que Ingresan al Nuevo Modelo

		B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
								Inferior	Superior
Paso 1 ^a	DISTRITO	1,090	,095	131,649	1	,000	2,973	2,468	3,582
	EDAD	,342	,129	7,013	1	,008	1,407	1,093	1,812
	TALLA	,693	,302	5,256	1	,022	1,999	1,106	3,615
	HEMOGLOBINA	1,266	,100	159,117	1	,000	3,545	2,912	4,315
	N_VISITAS	-,764	,297	6,597	1	,010	,466	,260	,834
	N_SESIONES	,811	,278	8,524	1	,004	2,249	1,305	3,876
	N_SACHETS	- 3,253	1,250	6,772	1	,009	,039	,003	,448
	JUNTOS	,946	,258	13,476	1	,000	2,576	1,554	4,269
	Constante	- 4,548	,594	58,707	1	,000	,011		

Fuente: Elaboración propia.

Tabla 7

Recodificación de Variables Para el Modelo Logístico

Variable	Dimensión	Indicadores	Categorías
FACTORES DE RIESGO DE ANEMIA	1. Factores Demográficos	X1: Distrito	0. Desaguadero
			1. Huacullani
			2. Juli
			3. Kelluyo
			4. Pisacoma
			5. Pomata
	2. Factores biológicos	X2: Edad	6. Zepita
			0. [1 - 12]
			1. [13 - 24]
			2. [25 - 36]
	3. Factores de intervención	X3: Talla	3. [37 - 48]
			4. [49 - 59]
0. > 80.3			
4. Factores sociales	X4: Hemoglobina	1. ≤ 80.3	
		0. ≥ 14.90	
		1. [14.40 - 14.90)	
3. Factores de intervención	X5: N_Visitas	2. [13.50 - 14.40)	
		3. < 11	
3. Factores de intervención	X6: N_Sesiones	0. Uno o más	
		1. Sin visitas	
3. Factores de intervención	X7: N_Sachets	0. Uno o más	
		1. Sin sesiones	
4. Factores sociales	X8: JUNTOS	0. De 0 a 20	
		1. Más de 20	
4. Factores sociales	X8: JUNTOS	0. No	
		1. Si	

Fuente: Elaboración propia.

La tabla 8, muestra el coeficiente de determinación, el cual en nuestro modelo alcanza el 75,5%. Esto implica que el 75,5% de la variabilidad en la variable anemia en niños menores de 5 años (anémicos o no anémicos) puede explicarse por las variables que conforman el modelo, que son las siguientes:

- X1: DISTRITO
- X2: EDAD
- X3: TALLA
- X4: HEMOGLOBINA
- X5: N_VISITAS
- X6: N_SESIONES
- X7: N_SACHETS
- X8: JUNTOS

Tabla 8

Prueba del Coeficiente de Determinación

Paso	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	629,086 ^a	,462	,755

Fuente: Elaboración propia.

En la tabla 9, se presenta la clasificación del modelo, donde se tiene una tasa de clasificación correcta del 92,4%. Esta tabla confirma que el modelo tiene una alta especificidad (78,8%) y una alta sensibilidad (95,4%). Con las variables predictoras: X1: Distrito, X2: Edad, X3: Talla, X4: Hemoglobina, X5: Número de visitas, X6: Número de sesiones, X7: Número de sachets y X8: Juntos, se demuestra que el modelo tiene una precisión del 92,4%, lo que indica que posee

una buena capacidad predictiva con las variables incluidas. Cabe señalar que el punto de corte utilizado es 0,5.

Tabla 9

Tabla de Clasificación

		Pronosticado			
		ANEMIA		Porcentaje correcto	
		No anémico	Anémico		
Paso 1	ANEMIA	No anémico	275	74	78.8
		Anémico	73	1507	95.4
		Porcentaje global			92.4

Fuente: Elaboración propia.

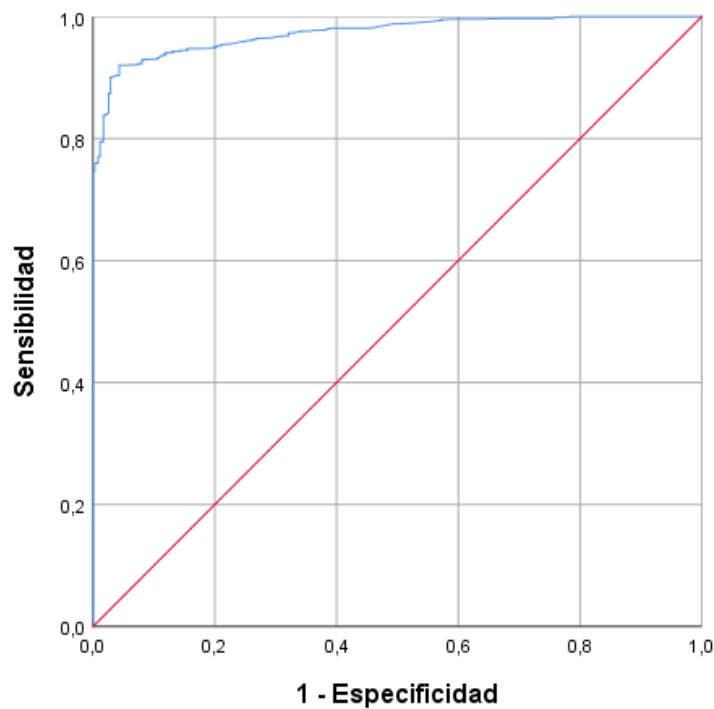
De la tabla de clasificación se puede concluir lo siguiente:

- VP (Verdaderos Positivos): La predicción de que un niño o niña menor de 5 años sea anémico es correcta en 95,4%.
- VN (Verdaderos Negativos): La predicción de que un niño o niña menor de 5 años no sea anémico es correcta en 78,8%.
- FP (Falsos Positivos): La predicción de que un niño o niña menor de 5 años termine como anémico es incorrecta en 4,6%.
- FN (Falsos Negativos): La predicción de que un niño o niña menor de 5 años termine como no anémico es incorrecta en 21,2%.

Para evaluar el rendimiento (o efectividad) del modelo clasificador, se utilizó la curva ROC (Receiver Operating Characteristic). En la figura 3 y la tabla 10 se muestra que el modelo, representado por la curva azul, supera a la curva aleatoria, cubriendo un área de 0.973. Esto indica que el modelo tiene un excelente poder discriminativo para predecir la anemia en niños menores de 5 años.

Figura 3

Gráfica de la Curva COR



Fuente: Elaboración propia

Tabla 10

Área Bajo la Curva ROC

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
0.973	0.003	0.000	0.967	0.979

Fuente: Elaboración propia.

Si el área bajo la curva ROC (COR) es igual a 1, indicaría una clasificación perfecta. A medida que el valor de esta área disminuye, la capacidad del modelo para discriminar entre las clases también se reduce. En este caso, con un área de 0.973, el modelo demuestra una excelente capacidad de clasificación. Al usar un punto de corte de 0.5, se confirma que el modelo sigue siendo muy eficaz, con una

alta capacidad para identificar correctamente la anemia en niños menores de 5 años.

4.2.2. Verosimilitud del modelo

La tabla 11, presenta la significancia estadística alcanzada a través del test de Wald. Considerando que la constante del modelo es relevante, se debe incorporar en el modelo, dado que influye en la variable de anemia (la categorización del niño o niña como anémico o no anémico).

Tabla 11

Prueba de Coeficiente de Intercepción

	B	Error estándar	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	1.510	0.059	651.878	1	0.000	4.527

Fuente: Elaboración propia.

Modelo de regresión logística (log-odds):

$$\text{Ln(odds)} = -4,548 + 1,090x_1 + 0,342x_2 + 0,693x_3 + 1,266x_4 - 0,764x_5 + 0,811x_6 - 3,253x_7 - 0,946x_8$$

En su forma inversa sería de la siguiente manera:

$$P(x) = \frac{1}{1 + e^{-(-4,548 + 1,090x_1 + 0,342x_2 + 0,693x_3 + 1,266x_4 - 0,764x_5 + 0,811x_6 - 3,253x_7 - 0,946x_8)}}$$

El valor de $-4,548$ es el intercepto del modelo. Representa el valor de la log-odds (logaritmo de las odds) cuando todas las variables independientes tienen un valor de cero.

Cada coeficiente asociado a una variable (x_i) indica el cambio esperado en el logaritmo de las odds por cada unidad de incremento en la variable, manteniendo constantes las demás.

- x_1 : Distrito (+1,090): los niños que residen en un distrito particular tienen una mayor probabilidad de anemia en comparación con otros distritos.
- x_4 : Hemoglobina (+1,266): cada incremento en los niveles de hemoglobina aumenta ligeramente la probabilidad de anemia, dado que el coeficiente es positivo.
- x_7 : Número de sachets (-3,253): por cada sachet adicional administrado, la probabilidad de anemia disminuye de manera considerable, ya que un coeficiente negativo reduce significativamente el valor del exponente en la fórmula, reduciendo la probabilidad.

4.2.3. Prueba de bondad de ajuste

La tabla 12, muestra la prueba de bondad de ajuste del modelo. Con un nivel de significancia del 5%, los resultados indican que el modelo es adecuado, lo que sugiere que el modelo de regresión logística se ajusta bien a los datos (p-valor = 2.73).

Tabla 12

Prueba de Bondad de Ajuste (Hosmer y Lemeshow)

Paso	Chi-cuadrado	gl	Sig.
1	9.894	8	0.273

Fuente: Elaboración propia.

En la tabla 13, se tiene la prueba de Ómnibus:

Tabla 13

Prueba Ómnibus

		Chi-cuadrado	gl	Sig.
	Paso	1194.937	8	0.000
Paso 1	Bloque	1194.937	8	0.000
	Modelo	1194.937	8	0.000

Fuente: Elaboración propia.

La hipótesis nula (H_0) se examina en la primera fila, la cual propone que los coeficientes de las variables añadidas son cero. Con un valor de significancia de 0.000, se descarta (H_0), lo que señala que al menos uno de los coeficientes de las variables se encuentra distinto a cero. Esto sugiere que el modelo, con las variables incluidas, presenta un ajuste significativo.

En la tabla 14, se observa que el valor de Kappa es 0,743 es decir en porcentaje 74,3 % el valor de la predicción coincide con el valor real.

Tabla 14

Medidas Simétricas Kappa

		Valor	Error estándar asintótico	T aproximada	Significación aproximada
Medida de acuerdo	Kappa	,743	,020	32,615	,000
N de casos válidos		1931			

Fuente: Elaboración propia.

4.3. PREDICCIÓN DEL RIESGO DE ANEMIA

Tabla 15

Predicción de Riesgo de Anemia en Niños Menores de 5 Años Para Dos Casos

Variables	Casos	
	y_{140}	y_{162}
Y: Anemia	1	0
X1: DISTRITO	4	0
X2: EDAD	4	1
X3: TALLA	0	1
X4: HEMOGLOBINA	2	0
X5: N_VISITAS	1	1
X6: N_SESIONES	1	1
X7: N_SACHETS	0	0
X8: JUNTOS	0	0
Predicción	0,97714	0,03027

Fuente: Elaboración propia

En el caso número 140 (Anexo 2), y_{140} la probabilidad de riesgo de anemia es alta (97.71%), debido a características como pertenecer al distrito de Pisacoma ($x_1 = 4 / \beta_1 = 1,090$), edad entre 49 y 59 meses ($x_2 = 4 / \beta_2 = 0,342$), talla mayor a 80.3 centímetros ($x_3 = 0 / \beta_3 = 0,693$), hemoglobina entre 13.5 a 14.4 g/dl ($x_4 = 2 / \beta_4 = 1,266$), sin visitas médicas ($x_5 = 1 / \beta_5 = -0,764$), sin sesiones educativas ($x_6 = 1 / \beta_6 = 0,811$), de 0 a 20 sachets consumidos ($x_7 = 0 / \beta_7 = -3,253$), y que no pertenece la programa juntos ($x_8 = 0 / \beta_8 = -0,946$). Todos estos respaldados por el intercepto o constante $\beta_0 = -4.568$, llevaron a que el modelo clasificara correctamente este caso como un niño anémico.



En el caso número 162 (Anexo 2), y_{162} la probabilidad de riesgo de anemia es muy baja (3.02%), debido a características como pertenecer al distrito de Desaguadero ($x_1 = 0 / \beta_1 = 1,090$), edad entre 13 y 24 meses ($x_2 = 1 / \beta_2 = 0,342$), talla menor igual a 80.3 centímetros ($x_3 = 1 / \beta_3 = 0,693$), hemoglobina mayor igual a 14.90 g/dl ($x_4 = 0 / \beta_4 = 1,266$), sin visitas médicas ($x_5 = 1 / \beta_5 = -0,764$), sin sesiones educativas ($x_6 = 1 / \beta_6 = 0,811$), de 0 a 20 sachets consumidos ($x_7 = 0 / \beta_7 = -3,253$), y que no pertenece la programa juntos ($x_8 = 0 / \beta_8 = -0,946$). Todos estos respaldados por el intercepto o constante $\beta_0 = -4.568$, llevaron a que el modelo clasificara correctamente este caso como un niño no anémico.

4.4. DISCUSIÓN

La investigación, centrado en la Red de Salud Chucuito – Puno, 2023, logró desarrollar un modelo de regresión logística efectivo para predecir el riesgo de anemia en niños menores de 5 años, alcanzando una precisión del 92.4% en la clasificación de casos. Este resultado se compara favorablemente con investigaciones internacionales y nacionales que emplearon la misma metodología para abordar problemáticas similares.

Los resultados de (Hernández Svendblad, 2019b), al estudiar el uso off-label de medicamentos en población pediátrica, si bien no reportó una precisión global, sí logró identificar las variables predictoras clave de esta práctica: la falta de alternativas terapéuticas para niños y el desconocimiento sobre el uso off-label. Esto resalta el valor de aplicar modelos de regresión logística para identificar factores predictivos en salud pública. En nuestro caso, la regresión logística permitió predecir el riesgo de anemia en niños menores de 5 años, lo que se alinea con los hallazgos sobre la importancia de identificar factores clave en enfermedades prevalentes en la infancia.



Otro estudio internacional, desarrollado por (Rodríguez Rueda et al., 2017) se enfocó en crear un modelo predictivo para daño hepático en pacientes sépticos. Aplicando la regresión logística, obtuvieron un índice de validez del 78%, demostrando la capacidad de la técnica para predecir eventos de salud. Si bien nuestro estudio se centró en una problemática diferente, la alta precisión alcanzada (92.4%) refuerza la eficacia de la regresión logística en la predicción de riesgos en salud.

Asimismo, estudios como el de (Sánchez Martínez et al., 2021), demuestran que la regresión logística es útil para identificar factores de riesgo en enfermedades crónicas. Aunque este estudio se centró en la diabetes, los principios aplicados son similares, puesto que nuestra investigación también incluye múltiples covariables para determinar su influencia en el riesgo de anemia. Esto refuerza la capacidad predictiva de los modelos de regresión logística en la evaluación de condiciones de salud.

A nivel nacional el estudio de (Navarro Chávez y Oliva Vera, 2018), sobre factores de riesgo de anemia en menores de 5 años, utilizando datos de la ENDES 2018, alcanzó una precisión del 69.5% en la clasificación de casos y demostraron que la prevalencia de anemia en Perú es considerablemente alta en áreas rurales, con un 36.4% de anemia en niños menores de cinco años a nivel nacional en el Perú. Este dato es comparable con nuestros hallazgos, donde la anemia afecta a un porcentaje significativo de la población infantil estudiada. El modelo desarrollado en este trabajo confirma la importancia de variables como el nivel de hemoglobina y la residencia en áreas rurales, lo que sugiere que las condiciones socioeconómicas y geográficas juegan un papel clave en la prevalencia de anemia.

Por otro lado, el modelo de (Cueva Román, 2023), para predecir el abandono del tratamiento de tuberculosis también es comparable al nuestro en términos de la



metodología utilizada. Al igual que en este estudio, se utilizaron variables predictivas para identificar el riesgo de un evento de salud, lo que resalta la importancia de los modelos predictivos basados en regresión logística para intervenciones preventivas.

Finalmente, los estudios tanto a nivel nacional como internacional en la investigación coinciden en que el uso de modelos predictivos en salud pública no solo permitiendo identificar los factores de riesgo, sino que también facilita la planificación de intervenciones más eficientes. La aplicación de nuestro modelo de regresión logística para predecir el riesgo de anemia ha demostrado una buena capacidad predictiva, lo que sugiere que este enfoque podría ser utilizado en el futuro para monitorear y reducir la prevalencia de anemia en poblaciones vulnerables.

V. CONCLUSIONES

PRIMERA: Se identificó y analizó los factores que elevan el riesgo de anemia en niños menores de cinco años en la provincia de Chucuito – Puno del año 2023. Las variables que influyeron son las siguientes: Distrito, Hemoglobina, Edad, Talla, Número de visitas, Número de sesiones y Número de sachets y juntos.

SEGUNDA: Se desarrolló el modelo predictivo basado en regresión logística que predice el riesgo de anemia en niños menores de 5 años en la provincia de Chucuito – Puno del año 2023. El modelo resultante fue: $Ln(y) = -4,548 + 1,090_{x1} + 0,342_{x2} + 0,693_{x3} + 1,266_{x4} - 0,764_{x5} + 0,811_{x6} - 3,253_{x7} - 0,946_{x8}$. Este modelo refleja la relación entre los factores de riesgo y la probabilidad de desarrollar anemia.

TERCERA: La prevalencia de anemia en la población estudiada en general fue de 81,9%. El modelo predictivo demostró un excelente ajuste en la prueba de bondad de Hosmer y Lemeshow, con un nivel de significancia de 0.05, su p-valor fue de 0,273, con un coeficiente de determinación de Nagelkerke de 75.5%. La tasa de clasificación de verdaderos negativos (clasificados como no anémicos fue del 78.8%; mientras que la tasa más relevante en este caso, clasificar correctamente a los niños anémicos (verdaderos positivos) fue de 95.4%, demostrando que el modelo tiene un porcentaje global de precisión del 92,4%. Así mismo el rendimiento o efectividad del modelo clasificador utilizando la curva ROC, cubre un área de 0.973, demostrando gran capacidad predictiva del modelo.



VI. RECOMENDACIONES

PRIMERA: Con respecto a la Red de Salud Chucuito, se recomienda focalizar las campañas de sensibilización y/o sesiones demostrativas, así como las visitas a los domicilios de los niños que presentan bajo nivel de hemoglobina y que podrían convertirse en un caso más de anemia, especialmente en los distritos con mayor número de casos. Asimismo, el consumo de sachets multimicronutrientes es muy importante, ya que esta investigación ha demostrado que son factores que ayudan a combatir la anemia.

SEGUNDA: Debido a la gran cantidad de datos generados en las instituciones de salud en nuestras regiones y el país, disponemos de tanta información, que a veces es imposible organizarla con efectividad. Por ello, se recomienda utilizar la técnica de Minería de Datos para descubrir patrones o algoritmos que permitan estudiar la dinámica de las enfermedades que afectan a nuestra niñez en el Perú.

TERCERA: Tomando en cuenta que la anemia en niños menores de 5 años afecta a toda la población del país, la Universidad Nacional del Altiplano y los puestos de salud promuevan la realización de campañas educativas en medios de comunicación; como también la preparación y ejecución de talleres didácticos dirigidos a la comunidad con la finalidad de identificar los factores de riesgo de la anemia en niños y madres gestantes.



VII. REFERENCIAS BIBLIOGRÁFICAS

- Al Kassab Córdova, A., Méndez Guerra, C. I., y Robles Valcarcel, P. (2020). *Factores sociodemográficos y nutricionales asociados a anemia en niños de 1 a 5 años en Perú*. Revista Chilena de Nutricion, 47(6), 925-932.
<https://doi.org/10.4067/S0717-75182020000600925>
- Aquino Canchari, C. R. (2021). Anemia infantil en el Perú: un problema aún no resuelto.
<https://orcid.org/0000-0002-7718-5598>
- Bardales Aguirre, L., Rodriguez Azabache, J., Barros Sevillano, J., Rodríguez Díaz, A., y Yupari, I. L. (2021). *Factores de riesgo de mortalidad por covid-19 en pacientes hospitalizados: un modelo de regresión logística*. Revista de la Facultad de Medicina Humana, 21(1), 19-27.
<https://doi.org/10.25176/rfmh.v21i1.3264>
- Bermeo E, J. R., González L, B. S., Ramón, E. G., y Mendieta, H. Z. (2012). *Modelo de regresión logística binaria para el diagnóstico morfológico de lesiones por virus del papiloma humano (VPH) en la boca*.
- Calle Núñez, A., Colqui Campos, K. A., Rivera Estrella, D. A., y Cieza Zevallos, J. A. (2017). *Factores asociados a la presentación de infecciones urinarias por Escherichia coli productoras de betalactamasas de espectro extendido*. Revista Med Hered (Vol. 28).
<http://dx.doi.org/https://doi.org/10.20453/rmh.v28i3.3180>
- Clínica Mayo. (2018, septiembre 13). Anemia infantil en niños y bebés. Clínica Mayo. <https://clinicamayo.com.ar/2018/09/13/que-es-la-anemia-infantil/>
- Cox, D. R. (1958). *The Regression Analysis of Binary Sequences*. Journal of the Royal Statistical Society Series B: Statistical Methodology, 20(2), 215-232.
<https://doi.org/10.1111/J.2517-6161.1958.TB00292.X>
- Cueva Román, J. A. (2023). Regresión logística binaria en factores asociados al abandono del tratamiento de tuberculosis, Lima 2023.



- Dávila Aliaga, C. R., Paucar Zegarra, R., y Quispe, A. (2019). Anemia infantil. *Revista Peruana de Investigación Materno Perinatal*, 7(2), 46-52.
<https://doi.org/10.33421/inmp.2018118>
- Escalante, P. R. (2004). *Prevención de la enfermedad*.
- García Pérez, R., García Pino, G., González Ballester, D., y García Moreno, R. (2010). *Modelo de regresión logística para estimar la dependencia según la escala de Lawton y Brody*. *Medicina de Familia. SEMERGEN*, 36(7), 365-371.
<https://doi.org/10.1016/J.SEMERG.2010.03.004>
- Grigore, M., Popovici, R. M., Gafitanu, D., Himiniuc, L., Murarasu, M., y Micu, R. (2020). *Logistic models and artificial intelligence in the sonographic assessment of adnexal masses – a systematic review of the literature*. *Medical Ultrasonography*, 22(4), 469-475.
<https://doi.org/10.11152/mu-2583>
- Hernández Sampieri, R., Fernández Collado, C., y del Pilar Baptista Lucio, M. (2010). *Metodología de la investigación, 5ta Ed.*
www.FreeLibros.com
- Hernández Sampieri, R., Fernández Collado, C., María del Pilar Baptista Lucio, D., y Méndez Valencia Christian Paulina Mendoza Torres, S. (2014). *Metodología de la investigación*.
- Hernández Svendblad, P. (2019a). *Uso off-label de medicamentos en la población pediátrica: aplicación del modelo de regresión logística binaria*. *Realidad y Reflexión*, 49(49), 40-65.
<https://doi.org/10.5377/ryr.v49i49.8062>
- Hernández Svendblad, P. (2019b). *Uso off-label de medicamentos en la población pediátrica: aplicación del modelo de regresión logística binaria*. *Realidad y Reflexión*, 49(49), 40-65.
<https://doi.org/10.5377/ryr.v49i49.8062>



- Hosmer, D. W. ., Lemeshow, Stanley., y Sturdivant, R. X. . (2013a). *Applied Logistic Regression Third edition*. Wiley.
- Hosmer, D. W., Lemeshow, S., y Sturdivant, R. X. (2013b). *Applied Logistic Regression: Third Edition*. Applied Logistic Regression: Third Edition, 1-510.
- <https://doi.org/10.1002/9781118548387>
- Hospital Universitario Ramón y Cajal. (2014). *Curvas ROC*.
http://www.hrc.es/bioest/roc_1.html
- Huaylinos Antezana, M. I. (2023). *Desnutrición crónica infantil en Perú: Avances y perspectivas*. *Revista Vive*, 6(18).
<https://doi.org/10.33996/revistavive.v6i18.269>
- INEI. (2019). *Encuesta demográfica y de salud familiar - ENDES*.
- INEI. (2024). *El 43,1% de la población de 6 a 35 meses de edad sufrió de anemia en el año 2023*.
- Issaiy, M., Zarei, D., y Saghazadeh, A. (2023). *Artificial Intelligence and Acute Appendicitis: A Systematic Review of Diagnostic and Prognostic Models*. *World Journal of Emergency Surgery : WJES*, 18(1), 59.
<https://doi.org/10.1186/S13017-023-00527-2>
- León Sánchez, M. A., y Linares Guerra, E. M. (2014). *La regresión logística binaria como instrumento para la predicción de deterioro inmunológico a partir de indicadores nutricionales en personas con vih/sida*. *Revista investigacion operacional* (vol. 35, número 1).
<https://rc.upr.edu.cu/handle/DICT/3328>
- MINSA. (2017). *Documento técnico plan nacional para la reducción y control de la anemia materno infantil y la desnutrición crónica infantil en el Perú*.
<http://www.minsa.gob.pe/>
- Navarro Chávez, Y., y Oliva Vera, J. E. (2018). *Modelo de regresión logística para identificar factores de riesgo y pronóstico de anemia en menores de cinco*



años, Perú – 2018.

<https://repositorio.unprg.edu.pe/handle/20.500.12893/9193>

OMS. (2011a). *Concentraciones de hemoglobina para diagnosticar la anemia y evaluar su gravedad VMNIS | 2.*

OMS. (2011b). *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity.*

OMS. (2016). Factores de riesgo y determinantes de la salud. *Revista de Salud Pública* (Vol. 4).

OMS. (2017). *Metas mundiales de nutrición 2025 - Lo que está en juego.*

Orraca Castillo, O., Amparo León Sánchez, M., Orraca Castillo, M., Lardoeyt Ferrer, R., y Alfredo Miló Valdés, C. (2021). *Modelo de regresión logística binaria para el cálculo de riesgo de asma en Pinar del Río Binary logistic regression model for calculating asthma risk in Pinar del Río.* *Revista Ciencias Médicas*, 25(5), 5301.

Rodríguez Rueda, J. M., Gonzáles Ferrer, V., de Jesús Montero Gonzales, T., y Consuegra Carvajal, Ada Nersys. (2017). Regresión logística binaria para crear un modelo predictivo de daño hepático en el paciente séptico. <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=77321>

Sánchez Martínez, B., Vega Falcón, V., y Nairovys Gómez Martínez, M. (2021). *Prediction of type 2 diabetes mellitus in adult patients using binary logistic regression.* <https://doi.org/10.46377/dilemas.v8i3.2675>

SIEN, Ministerio de Salud, Instituto Nacional de Salud, y Centro Nacional de Alimentación, N. y V. S. (2023). *Informe Gerencial SIEN HIS Estado Nutricional de Niños menores de cinco años que acceden a Establecimientos de Salud.*

<https://www.gob.pe/34823-instituto-nacional-de-salud-centro-nacional-de-alimentacion-y-nutricion-cenan>



Zavaleta, N. (2017). Anemia infantil: retos y oportunidades al 2021. En *Revista Peruana de Medicina Experimental y Salud Publica* (Vol. 34, Número 4, pp. 588-589). Instituto Nacional de Salud.
<https://doi.org/10.17843/rpmesp.2017.344.3281>

Zavaleta, N., y Astete-Robilliard, L. (2017). *Efecto de la anemia en el desarrollo infantil: consecuencias a largo plazo*. *Revista Peruana de Medicina Experimental y Salud Publica*, 34(4), 716-722.
<https://doi.org/10.17843/RPMESP.2017.344.3251>



ANEXOS

ANEXO 1. Registro SIEN, Chucuito - Puno del año 2023

N°	Distrito	Edad	Talla	Hemoglobina	N_Visitas	N_Sesiones	N_Sachets	JUNTOS	Dx_Anemia
130	KELLUYO	18	78	14.2	1	0	0	No	No Anémico
131	KELLUYO	11	83.6	14.5	0	0	0	No	Anémico
132	KELLUYO	18	73.9	14.7	0	0	0	No	Anémico
133	KELLUYO	59	115	14.4	0	0	0	No	No Anémico
134	KELLUYO	19	88.1	16.6	0	0	0	No	Anémico
135	DESAGUADERO	36	84	14.1	0	0	0	No	No Anémico
136	DESAGUADERO	20	71.8	15.2	0	0	0	No	No Anémico
137	DESAGUADERO	39	94.7	14.5	0	0	0	No	No Anémico
138	DESAGUADERO	7	64	14.2	0	0	0	No	No Anémico
139	PISACOMA	52	86.8	14.2	0	0	0	No	Anémico
140	PISACOMA	51	96.2	14.1	0	0	0	No	Anémico
141	DESAGUADERO	27	78.9	14.1	0	0	4	No	No Anémico
142	DESAGUADERO	33	91.9	14.5	0	0	5	No	No Anémico
143	DESAGUADERO	48	102.7	15.7	0	0	2	No	No Anémico
144	DESAGUADERO	44	86.7	14.1	0	0	4	No	No Anémico
145	DESAGUADERO	7	71.6	14.2	0	0	1	No	No Anémico
146	DESAGUADERO	27	88.4	14.8	0	0	4	No	No Anémico
147	DESAGUADERO	47	102	14.4	0	0	0	No	No Anémico
148	PISACOMA	16	72.3	14.3	0	0	0	No	Anémico
149	DESAGUADERO	33	87.8	14.1	0	0	3	No	No Anémico
150	DESAGUADERO	48	102	14.1	0	0	0	No	No Anémico
151	DESAGUADERO	21	83	14.4	0	0	6	No	No Anémico
152	DESAGUADERO	36	82.3	14.1	0	0	2	No	No Anémico
153	DESAGUADERO	6	64.5	15.9	0	0	1	No	No Anémico
154	PISACOMA	13	66	14.7	0	0	1	No	Anémico
155	DESAGUADERO	14	72.3	14.9	0	0	1	No	No Anémico
156	PISACOMA	28	80.8	14.5	0	0	3	No	Anémico
157	DESAGUADERO	7	63.5	14.8	0	0	2	No	No Anémico
158	DESAGUADERO	7	64	14.8	0	0	1	No	No Anémico
159	DESAGUADERO	22	81	14.9	0	0	0	No	No Anémico
160	DESAGUADERO	23	83	14.8	0	0	2	No	No Anémico
161	DESAGUADERO	29	82	14.4	0	0	5	No	No Anémico
162	DESAGUADERO	14	73	15.3	0	0	0	No	No Anémico
163	DESAGUADERO	37	90	15.6	0	0	0	No	No Anémico
164	DESAGUADERO	36	96	14.4	0	0	0	No	No Anémico
165	DESAGUADERO	47	103.2	14.4	0	0	4	No	No Anémico
166	DESAGUADERO	36	93.3	14.5	0	0	4	No	No Anémico
167	DESAGUADERO	8	62.1	14.3	0	0	1	No	No Anémico
168	DESAGUADERO	16	76	14.4	0	0	2	No	No Anémico
169	DESAGUADERO	36	84.2	15.3	0	0	1	No	No Anémico
170	DESAGUADERO	10	57.1	14.6	0	0	2	No	Anémico
171	DESAGUADERO	36	85.8	14.5	0	0	1	No	No Anémico
172	PISACOMA	15	72.4	14.5	0	0	4	No	Anémico
173	DESAGUADERO	36	90.4	14.5	0	0	0	No	No Anémico
174	DESAGUADERO	8	67	17	0	0	1	No	No Anémico
175	DESAGUADERO	6	66.5	14.4	0	0	0	No	No Anémico
176	DESAGUADERO	39	87.4	14.5	0	0	0	No	No Anémico
177	DESAGUADERO	14	70	14.4	0	0	2	No	No Anémico
178	DESAGUADERO	18	81.5	15.1	0	0	1	No	No Anémico
179	DESAGUADERO	12	73.9	14.3	0	0	1	No	No Anémico
180	DESAGUADERO	41	104	14.3	0	0	1	No	No Anémico
181	DESAGUADERO	36	94	14.4	0	0	6	No	No Anémico
182	DESAGUADERO	13	71.6	16.5	0	0	1	No	No Anémico
183	DESAGUADERO	19	71.4	14.3	0	0	0	No	No Anémico
184	DESAGUADERO	6	63.5	14.3	0	0	1	No	No Anémico
185	DESAGUADERO	20	85	14.3	0	0	0	No	No Anémico
186	DESAGUADERO	15	76.4	14.5	0	0	1	No	No Anémico
187	DESAGUADERO	14	67.4	14.5	0	0	0	No	No Anémico
188	DESAGUADERO	24	86	14.5	0	0	0	No	No Anémico
189	DESAGUADERO	12	72.5	14.4	0	0	0	No	No Anémico
190	DESAGUADERO	6	64.8	14.4	0	0	1	No	No Anémico
191	DESAGUADERO	50	93	15	0	0	0	No	No Anémico
192	PISACOMA	7	56.8	15.6	0	0	1	No	Anémico
193	DESAGUADERO	24	86	14.7	0	0	0	No	No Anémico
194	DESAGUADERO	24	78.4	14.5	0	0	3	No	No Anémico
195	DESAGUADERO	36	92.5	15.1	0	0	0	No	No Anémico
196	DESAGUADERO	7	58.1	15	0	0	1	No	No Anémico
197	DESAGUADERO	20	81.8	14.5	0	0	0	No	No Anémico
198	DESAGUADERO	7	68	14.6	0	0	1	No	No Anémico
199	DESAGUADERO	9	63	14.4	0	0	4	No	No Anémico
200	DESAGUADERO	49	86.5	14.6	0	0	0	No	No Anémico



ANEXO 2. Predicciones de Anemia

N°	DISTRITO R	EDAD R	TALLA R	HEMOGLOBINA R	N. VISITAS R	N. SESIONES R	N. SACHETS R	JUNTOS R	ANEMIA R	PRE_1	PGR_2
	X1	X2	X3	X4	X5	X6	X7	X8			
130	3	1	1	2	0	1	0	0	0	,95678	1
131	3	0	0	1	1	1	0	0	1	,50837	1
132	3	1	1	1	1	1	0	0	1	,74419	1
133	3	4	0	1	1	1	0	0	0	,80217	1
134	3	1	0	0	1	1	0	0	1	,29102	0
135	0	2	0	2	1	1	0	0	0	,21638	0
136	0	1	1	0	1	1	0	0	0	,03027	0
137	0	3	0	1	1	1	0	0	0	,09878	0
138	0	0	1	2	1	1	0	0	0	,21801	0
139	4	4	0	2	1	1	0	0	1	,97714	1
140	4	4	0	2	1	1	0	0	1	,97714	1
141	0	2	1	2	1	1	0	0	0	,35569	0
142	0	2	0	1	1	1	0	0	0	,07226	0
143	0	3	0	0	1	1	0	0	0	,02999	0
144	0	3	0	2	1	1	0	0	0	,27983	0
145	0	0	1	2	1	1	0	0	0	,21801	0
146	0	2	0	1	1	1	0	0	0	,07226	0
147	0	3	0	1	1	1	0	0	0	,09878	0
148	4	1	1	2	1	1	0	0	1	,96842	1
149	0	2	0	2	1	1	0	0	0	,21638	0
150	0	3	0	2	1	1	0	0	0	,27983	0
151	0	1	0	1	1	1	0	0	0	,05245	0
152	0	2	0	2	1	1	0	0	0	,21638	0
153	0	0	1	0	1	1	0	0	0	,02170	0
154	4	1	1	1	1	1	0	0	1	,89637	1
155	0	1	1	0	1	1	0	0	0	,03027	0
156	4	2	0	1	1	1	0	0	1	,85892	1
157	0	0	1	1	1	1	0	0	0	,07291	0
158	0	0	1	1	1	1	0	0	0	,07291	0
159	0	1	0	0	1	1	0	0	0	,01537	0
160	0	1	0	1	1	1	0	0	0	,05245	0
161	0	2	0	1	1	1	0	0	0	,07226	0
162	0	1	1	0	1	1	0	0	0	,03027	0
163	0	3	0	0	1	1	0	0	0	,02999	0
164	0	2	0	1	1	1	0	0	0	,07226	0
165	0	3	0	1	1	1	0	0	0	,09878	0
166	0	2	0	1	1	1	0	0	0	,07226	0
167	0	0	1	2	1	1	0	0	0	,21801	0
168	0	1	1	1	1	1	0	0	0	,09964	0
169	0	2	0	0	1	1	0	0	0	,02150	0
170	0	0	1	1	1	1	0	0	1	,07291	0
171	0	2	0	1	1	1	0	0	0	,07226	0
172	4	1	1	1	1	1	0	0	1	,89637	1
173	0	2	0	1	1	1	0	0	0	,07226	0
174	0	0	1	0	1	1	0	0	0	,02170	0
175	0	0	1	1	1	1	0	0	0	,07291	0
176	0	3	0	1	1	1	0	0	0	,09878	0
177	0	1	1	1	1	1	0	0	0	,09964	0
178	0	1	0	0	1	1	0	0	0	,01537	0
179	0	0	1	2	1	1	0	0	0	,21801	0
180	0	3	0	2	1	1	0	0	0	,27983	0
181	0	2	0	1	1	1	0	0	0	,07226	0
182	0	1	1	0	1	1	0	0	0	,03027	0
183	0	1	1	2	1	1	0	0	0	,28177	0
184	0	0	1	2	1	1	0	0	0	,21801	0
185	0	1	0	2	1	1	0	0	0	,16403	0
186	0	1	1	1	1	1	0	0	0	,09964	0
187	0	1	1	1	1	1	0	0	0	,09964	0
188	0	1	0	1	1	1	0	0	0	,05245	0
189	0	0	1	1	1	1	0	0	0	,07291	0
190	0	0	1	1	1	1	0	0	0	,07291	0
191	0	4	0	0	1	1	0	0	0	,04170	0
192	4	0	1	0	1	1	0	0	1	,63424	1
193	0	1	0	1	1	1	0	0	0	,05245	0
194	0	1	1	1	1	1	0	0	0	,09964	0
195	0	2	0	0	1	1	0	0	0	,02150	0
196	0	0	1	0	1	1	0	0	0	,02170	0
197	0	1	0	1	1	1	0	0	0	,05245	0
198	0	0	1	1	1	1	0	0	0	,07291	0
199	0	0	1	1	1	1	0	0	0	,07291	0
200	0	4	0	1	1	1	0	0	0	,13363	0



ANEXO 3. Descripción de Variables del SIEN - DIRESA provincia de Chucuito

Variable	Descripción
EESS	Representa el establecimiento de salud donde se realiza la atención de los niños y gestantes dentro de la Red de Salud Chucuito, conforme a la identificación proporcionada por el Ministerio de Salud.
Fecha	Es la fecha en que el niño menor de cinco años recibe atención, expresada en formato dd/mm/aaaa.
Hist_Clinica	El número otorgado a la historia clínica del niño atendido en el establecimiento de salud correspondiente.
Provincia	Identificación de la provincia en la que reside el niño, en este caso Chucuito, según el INEI.
Distrito	Identificación del distrito dentro de la provincia de Chucuito donde reside el niño, basado en el INEI.
Localidad	Lugar específico de residencia del niño.
Hemoglobina	Resultado de los análisis de hemoglobina obtenidos en la prueba sanguínea, expresado en g/dl.
Altitud_Loc	Altitud de la residencia del niño, expresada en metros sobre el nivel del mar (msnm), de acuerdo con los datos proporcionados por el INEI.
Hbc	Resultado de la hemoglobina ajustado por la altitud de la localidad, expresado en g/dl.
Sexo	Género del niño, que puede ser masculino o femenino.



Fec_Nacimiento	Fecha de nacimiento del infante, expresada en formato dd/mm/aaaa.
Edad	Edad del infante, calculada en meses, con fracción de mes incluida.
Peso	Peso del infante, medido en kilogramos.
Talla	Estatura del niño, medida en centímetros.
Dx_P/T	Índice de peso para la talla, que compara el peso del niño con el esperado según su talla, permitiendo detectar desnutrición aguda o sobrepeso, con base en los valores z-Score.
Dx_T/E	Índice de talla para la edad, que evalúa si el niño presenta retraso en su crecimiento, indicando desnutrición crónica según los valores z-Score.
Dx_P/E	Índice de peso para la edad, que determina si el niño presenta desnutrición global, usando valores z-Score según el patrón OMS.
Dx_Anemia	Clasificación del niño según los niveles de hemoglobina, basada en los estándares de la OMS, los cuales establecen distintos grados de anemia de acuerdo con la concentración de hemoglobina.
N_CCred	Número de controles de Crecimiento y Desarrollo (CRED) que ha recibido el niño, incluyendo el realizado en el momento de la atención.
N_Sachets	Cantidad de sachets de multimicronutrientes consumidos por el niño desde la última consulta hasta la atención actual.



N_Consejerias	Número de consejerías nutricionales recibidas por la madre del niño desde la consulta anterior.
N_Sesiones	Número de sesiones demostrativas que la madre del niño ha recibido, desde la última consulta.
N_Visitas	Número de visitas domiciliarias realizadas a la familia del niño desde la última consulta.
JUNTOS	Indica si la familia del infante está afiliada al programa JUNTOS.
SIS	Indica si el infante está inscrito en el Seguro Integral de Salud (SIS).
PIN	Indica si el infante está inscrito en el Subprograma Infantil del Programa Integral de Nutrición (PIN).

Fuente: (SIEN et al., 2023)



ANEXO 4. Declaración jurada de autenticidad de tesis



Universidad Nacional
del Altiplano Puno



Vicerrectorado
de Investigación



Repositorio
Institucional

DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo YUDITH DIANA CHALCO CEREZO,
identificado con DNI 75277319 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado
INGENIERÍA ESTADÍSTICA E INFORMÁTICA

informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:
" MODELO DE REGRESIÓN LOGÍSTICA BINARIA PARA
PREDECIR EL RIESGO DE ANEMIA EN NIÑOS MENORES
DE 5 AÑOS EN LA RED DE SALUD CHUCUITO - 2023 "

Es un tema original.

Declaro que el presente trabajo de tesis es elaborado por mi persona y **no existe plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

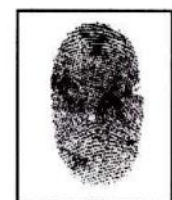
Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como tuyas las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las connotaciones éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Puno 18 de diciembre del 2024


FIRMA (obligatoria)



Huella



ANEXO 5. Autorización para el depósito de tesis en el Repositorio Institucional



Universidad Nacional
del Altiplano Puno



Vicerrectorado
de Investigación



Repositorio
Institucional

AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo YUDITH DIANA CHALCO CEREZO,
identificado con DNI 75277319 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado
INGENIERÍA ESTADÍSTICA E INFORMÁTICA

informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:

“ MODELO DE REGRESIÓN LOGÍSTICA BINARIA PARA PREDECIR
EL RIESGO DE ANEMIA EN NIÑOS MENORES DE 5 AÑOS EN
LA RED DE SALUD CHUCUITO - 2023 ”

para la obtención de Grado, Título Profesional o Segunda Especialidad.

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los “Contenidos”) que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.


En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visita: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

En señal de conformidad, suscribo el presente documento.

Puno 18 de diciembre del 2024


FIRMA (obligatoria)



Huella