**Actividad práctica 1.**

Apellido: Arismendiz Millones

Nombre: Jennyfer Arantxa

Fecha: 11/04/2024

**CASO 1**

Se publicó el artículo:

Medina J, et al (2006). Prevalencia de Sobrepeso y Obesidad en la Población Adulta de Arequipa Metropolitana: Resultados del Estudio PREVENCION. Revista Peruana de Cardiología Vol. XXXII Nº 3.

A continuación se transcriben fragmentos de los acápites de Antecedentes y Métodos. Se adjunta el artículo completo si se desea examinar el mismo.

**Antecedentes**

Los niveles de sobrepeso y obesidad han aumentado progresivamente en las últimas décadas tanto en países desarrollados como en muchos otros en vías de desarrollo. (…). La epidemia de sobrepeso y obesidad es preocupante debido a que existe abundante evidencia que indica que el peso excesivo aumenta el riesgo de padecer diversas patologías que alteran significativamente la expectativa y la calidad de vida. (…). En estudios prospectivos realizados en Norteamérica, Europa y Asia, se ha demostrado que el sobrepeso aumenta la mortalidad debida principalmente a su asociación con eventos cardiovasculares y cáncer. (…). Por lo anterior consideramos que es importante que los estudios epidemiológicos que buscan determinar la prevalencia de sobrepeso y obesidad también analicen la presencia de obesidad abdominal en las poblaciones estudiadas. (…).Es por esto importante evaluar con precisión la prevalencia y distribución poblacional del sobrepeso y la obesidad para poder estimar la magnitud del problema y diseñar estrategias de control adecuadas. En este artículo reportamos los hallazgos en relación a la prevalencia de obesidad, sobrepeso, peso adecuado y bajo peso de acuerdo al índice de masa corporal (IMC) y la prevalencia de obesidad abdominal de acuerdo al perímetro de cintura (PC) en la población adulta de Arequipa Metropolitana, provenientes del estudio de investigación PREVENCION (*Estudio de la Prevalencia de Enfermedades Cardiovasculares y Factores de Riesgo Coronario en Arequipa*) (19).

**Métodos**

Población

Los objetivos generales, diseño, plan y esquema operativo del estudio PREVENCION han sido publicados anteriormente (19). PREVENCION es un estudio poblacional realizado en la ciudad de Arequipa, una de las ciudades más pobladas del Perú. La primera fase de PREVENCION consiste en un estudio poblacional de corte transversal diseñado para determinar la prevalencia de enfermedades cardiovasculares y factores de riesgo cardiovascular en la población adulta de Arequipa Metropolitana. La ciudad de Arequipa fue considerada ideal para un estudio de esta naturaleza, por contar con varias características favorables, incluyendo su geografía relativamente compacta, una población suficientemente grande para proporcionar una muestra adecuada y un número relativamente bajo de hospitales de referencia, todo lo que facilitará el seguimiento prospectivo planeado para la segunda fase del estudio (19). Su población está compuesta principalmente por mestizos, con contribuciones menores de blancos y mínimas de habitantes de raza negra (20,21).

Diseño de la Estrategia de muestreo

El marco muestral para el estudio se obtuvo del Instituto Nacional de Estadística e Informática (INEI) y se basó en el último censo nacional. Se obtuvo una muestra representativa de la población adulta de Arequipa Metropolitana mediante una estrategia de muestreo probabilística, multietápica y estratificada de acuerdo a zonas geográficas y estratos socioeconómicos. Nuestra muestra mínima debía incluir un mínimo de 1600 individuos y como mínimo 200 individuos por género en grupos etáreos predefinidos (20-34, 35-49, 50-64, y 65-80) para poder contar con un poder estadístico adecuado para detectar diferencias entre los subgrupos de edad y género. Esto se logró a través de un exceso deliberado de muestreo de sujetos mayores de 65 años para mejorar la precisión de las estimaciones de este grupo etáreo. Se excluyeron del estudio a las mujeres embarazadas, individuos con diagnóstico previo de alguna neoplasia maligna y personas que vivían menos de 5 años en Arequipa. La tasa de respuesta general fue de 85.3% y la muestra final consistió de 1878 individuos (867 varones y 1011 mujeres) pertenecientes a 626 familias. (…).

Antropometría y Definiciones

El peso se midió usando una balanza clínica (Seca®, Alemania) calibrada y la talla se midió usando un tallímetro con precisión de 0.5 centímetros. En ambos casos se realizaron las medidas con los sujetos descalzos y usando ropa ligera mientras se encontraban de pie con su peso equidistantemente distribuido en ambos pies, con los brazos a los costados y la cabeza mirando al frente. Además, usando una cinta métrica inelástica, se midió el PC [perímetro de la cintura] a nivel umbilical y al final de una expiración normal, con una precisión de 0.1 centímetros. Todas las medidas fueron realizadas por observadores entrenados. Se calculó el IMC dividiendo el peso sobre la talla al cuadrado (kg/m2). Para la clasificación del estado nutricional de cada sujeto se utilizó la clasificación propuesta para adultos por la Organización Mundial de Salud (OMS) basada en diferentes puntos de corte de acuerdo al IMC (1). De esta forma se definieron las siguientes categorías tanto para hombres como para mujeres; bajo peso (IMC < 18.5 kg/m2), peso adecuado (IMC de 18.5 - 24.9 kg/m2), sobrepeso (IMC de 25.0 - 29.9 kg/m2) y obesidad (IMC ≥ 30.0 kg/m2). (…) Además, se clasificó la obesidad en tres clases: clase I (IMC de 30.0 - 34.9 kg/m2), clase II (IMC de 35.0 - 39.9 kg/m2) y clase III u obesidad mórbida (IMC ≥ 40.0 kg/m2). Para evaluar la presencia de obesidad abdominal se utilizó dos clasificaciones basadas en valores del PC: a) la propuesta en el tercer informe del panel de expertos del Programa de Educación Nacional del Colesterol norteamericano (ATP III), que define la obesidad abdominal como la presencia de un PC > 102 centímetros en hombres y > 88 centímetros en mujeres; y b) la propuesta por la Federación Internacional de Diabetes (IDF) para poblaciones de América Central y América del Sur, que utiliza como definición de obesidad abdominal la presencia de un PC ≥ 90 centímetros en hombres ≥ 80 centímetros en mujeres (16,17).”

A partir de este texto:

1. Identificar el objetivo del estudio: (1 p.)

El objetivo del estudio es determinar la prevalencia de sobrepeso, obesidad, peso adecuado y bajo peso según el índice de masa corporal (IMC), así como la prevalencia de obesidad abdominal según el perímetro de cintura (PC) en la población adulta de Arequipa Metropolitana.

1. La población del estudio: (1 p.)

La población del estudio es la población adulta de Arequipa Metropolitana.

1. La muestra: (1 p.)

Se excluyeron del estudio a mujeres embarazadas, individuos con diagnóstico previo de alguna neoplasia maligna y personas que vivían menos de 5 años en Arequipa. La muestra final consistió en 1878 individuos (867 varones y 1011 mujeres) pertenecientes a 626 familias.

1. Identificar las variables que se hayan medido directamente en la unidad de análisis, su tipo y escala de medición. Hay variables **generadas**, como el IMC y las clasificaciones que no se deben considerar dentro de la relación. (2 p.)

|  |  |  |
| --- | --- | --- |
| Variable | Tipo | Escala de medición |
| Peso | Cuantitativa continua | Kg |
| Talla | Cuantitativa continua | Metros |
| Perímetro de cintura | Cuantitativa continua | Centímetros |
| Sexo | Cualitativa dicotómica | Nominal |
| Estado nutricional (Bajo peso, peso adecuado, sobrepeso, obesidad) | Cualitativa  politómica | Ordinal |
| Obesidad | Cualitativa politómica | Ordinal |
| Obesidad abdominal | Cualitativa dicotómica | Ordinal |

**CASO 2**

Se realizó el registro de variables como edad, sexo, peso, talla, hemoglobina y nivel socioeconómico en una muestra de 160 niños entre 8 a 12 años de edad de un centro educativo.

A continuación, se presenta una tabla con la información de las variables registradas en el archivo adjunto de Excel: Base 1.xlsx

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre | Etiqueta | Valor | Etiqueta |
| id | Código de identificación |  |  |
| edad | Edad (años) |  |  |
| sexo | Sexo | 0  1 | Femenino  Masculino |
| peso | Peso (kg) |  |  |
| talla | Talla (cm) |  |  |
| dosajehb | Dosaje hemoglobina (g/dl) |  |  |
| nse | Nivel socioeconómico | 1  2  3 | Bajo  Medio  Alto |

1. Importar el archivo Excel al Stata. (1 p)

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Configurar la etiqueta de variables según la tabla previa. (1 p)

A screenshot of a computer

Description automatically generated

1. Configurar la etiqueta de valores a las variables que corresponda. (1 p)

A screenshot of a computer

Description automatically generated

3.1. Sexo

A screenshot of a computer

Description automatically generated

A screenshot of a table

Description automatically generatedA screenshot of a computer screen

Description automatically generated3.2. Nivel socioeconómico (nse)

1. Con las modificaciones previas guardar la base de datos: Base 1.dta (1 p). (Adjuntar base de datos con todas las modificaciones solicitadas)

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Generar con Stata una descripción de las variables (codebook). (1 p) (Presentar salida en archivo Word y comentar)

5.1. Id

A screenshot of a computer

Description automatically generated

La variable id tiene valores únicos del 1 al 160 sin faltantes. Su media es aproximadamente 80.5, con una desviación estándar de alrededor de 46.3321, indicando cierta dispersión. Los percentiles revelan que el 25% de los valores son menores o iguales a 40.5, la mediana es 80.5, y el 75% son menores o iguales a 120.5. Los percentiles 10 y 90 son 16.5 y 144.5 respectivamente, proporcionando información sobre la distribución de los datos.

5.2. Edad

A screenshot of a computer

Description automatically generated

Los valores de la variable edad están en el rango de 8 a 12, con un total de 5 valores únicos. No hay valores faltantes en esta variable. La distribución de edades muestra que hay 24 personas de 8 años, 45 de 9 años, 48 de 10 años, 36 de 11 años y 7 de 12 años.

5.3. Sexo

A screenshot of a computer

Description automatically generated

La variable llamada sexo representa el género. Los valores codificados son 0 para "Femenino" y 1 para "Masculino". La tabulación muestra que hay 93 registros clasificados como femeninos y 67 como masculinos.

5.4. Peso

A screenshot of a computer

Description automatically generated

Esta variable representa el peso en kilogramos. Los valores de peso están en el rango de 20.3 kg a 57.7 kg, con un total de 135 valores únicos y no hay valores faltantes en esta variable. La media del peso es de aproximadamente 37.147 kg, con una desviación estándar de alrededor de 8.81761 kg, lo que indica cierta variabilidad en los datos. Los percentiles muestran que el 25% de los registros tienen un peso igual o menor a 30.65 kg, la mediana es de 35.5 kg, y el 75% de los registros tienen un peso igual o menor a 43.62 kg. Los percentiles 10 y 90 son 26.15 kg y 49.9 kg respectivamente.

5.5. Talla

A screenshot of a computer

Description automatically generated

La altura varía entre 119.3 cm y 156.2 cm, con una media de aproximadamente 137.251 cm y una desviación estándar de alrededor de 8.43263 cm. No hay valores faltantes en esta variable. Los percentiles indican que el 25% de las observaciones tienen una altura igual o menor a 130.1 cm, la mediana es de 136.9 cm, y el 75% de las observaciones tienen una altura igual o menor a 143.85 cm. Los percentiles 10 y 90 son 126.55 cm y 148.15 cm respectivamente.

5.6. dosajehb

A screenshot of a computer

Description automatically generated

La variable dosajehb indica el nivel de hemoglobina en g/dl. Los valores oscilan entre 9.8 g/dl y 15.1 g/dl, con 44 valores únicos y ninguna ausencia de datos. La media del nivel de hemoglobina es de aproximadamente 12.4362 g/dl, con una desviación estándar de alrededor de 1.12599 g/dl. Los percentiles muestran que el 25% de las mediciones tienen un nivel de hemoglobina de 11.7 g/dl o menos, mientras que el 75% tienen 13.2 g/dl o menos. Los percentiles 10 y 90 son 10.8 g/dl y 13.95 g/dl respectivamente.

5.7. nse

A screenshot of a computer

Description automatically generated

Los valores del nse (nivel socioeconómico) están codificados en un rango de 1 a 3, correspondiendo a las categorías "Bajo", "Medio" y "Alto" respectivamente. No hay valores faltantes en esta variable, con 90 registros clasificados como nivel socioeconómico "Bajo", 51 como "Medio" y 19 como "Alto".

1. Generar la variable agrupada *anemia* (gen) igual a dosajehb y agrupar (recode) la variable dosajehb en dos categorías: (2 p) (Presentar salida en archivo Word)

Anemia (1): dosajehb menor a 11.

Sano (0): dosajehb, mayor o igual a 11.

* + 1. Generar la variable agrupada *anemia* (gen) igual a dosajehb



* + 1. Agrupar (recode) la variable dosajehb en dos categorías:
* Anemia (1): dosajehb menor a 11.
* Sano (0): dosajehb, mayor o igual a 11.



A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated Salida obtenida: Agregamos etiquetas de valor:

1. Generar las tablas de distribución de frecuencias de las variables: edad, sexo, anemia y nse. (1 p) (Presentar salida en archivo Word)

A table with numbers and a number on it

Description automatically generated

A screenshot of a number

Description automatically generated

A number of numbers on a white background

Description automatically generated

A white sheet with black text

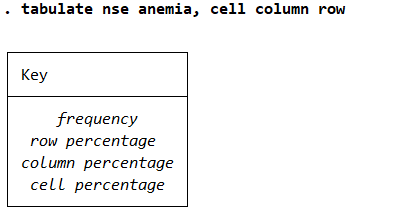
Description automatically generated with medium confidence

1. Generar un gráfico de barra de la variable anemia. Comentar. (2 p) (Presentar salida en archivo Word y comentar)



Podemos observar que más del 85% de la muestra son personas sanas que tienen un nivel de hemoglobina superior a 11 (g/dl), por lo tanto, casi el 15% de la muestra está en un estado anémico.

1. Generar una tabla de doble entrada para distribuir la anemia según nse. (1 p)



A table of numbers and a few words

Description automatically generated with medium confidence

Esta tabla de doble entrada presenta la distribución conjunta entre el nivel socioeconómico y la presencia de anemia en un conjunto de datos. Los valores están organizados en cuatro categorías: "Bajo", "Medio" y "Alto" para el nivel socioeconómico, y "Sano" y "Anemia" para el estado de salud. Cada celda muestra la frecuencia absoluta de observaciones que caen en la intersección de una categoría de nivel socioeconómico y una categoría de estado de salud. Además, se proporcionan porcentajes de fila, columna y celda para ayudar a comprender la distribución de los datos en términos proporcionales. Por ejemplo, podemos observar que el 85.56% de las personas con nivel socioeconómico "Bajo" están sanas, mientras que el 14.44% tienen anemia.

1. Generar un gráfico de barras superpuestas para distribuir la anemia según nse. Comentar la tabla de doble entrada y el gráfico. (2 p) (Presentar salida en archivo Word)





Según la interpretación del gráfico que relaciona la anemia con el nivel socioeconómico, se sugiere que las personas con un estatus socioeconómico bajo tienen una proporción más alta de casos de anemia. En contraste, aquellos con un nivel socioeconómico medio muestran una incidencia de anemia intermedia, mientras que los individuos con un estatus socioeconómico más alto presentan una menor prevalencia de anemia.

1. Generar los histogramas de peso y talla, respectivamente. Comentar. (2 p). (Presentar salida en archivo Word)

El gráfico de barras que representa el peso exhibe los intervalos de peso en kilogramos, acompañados de frecuencias proporcionales expresadas como densidad en el eje vertical. La mayor concentración de densidad se registra en la franja de peso que va de más de 30 kg, pero menos de 37 kg. En contraste, en el histograma del peso en kilogramos, se aprecia una densidad más pronunciada en individuos con una estatura alrededor de los 130 cm.

11.1. Histograma de peso



11.2. Histograma de talla

