

# Atención a la Salud de la Comunidad

Tema 6. Variables del Estudio y Estrategia de Análisis.



**Celia Nespral Gaztelumendi**  
**Miguel Santibáñez Margüello**  
**Montserrat Bustamante Fonfría**

DPTO. DE ENFERMERÍA

Este tema se publica bajo Licencia:

[Creative Commons BY-NC-SA 3.0](https://creativecommons.org/licenses/by-nc-sa/3.0/)

## Fases de la planificación de la investigación

### 1-Conceptualización del problema de investigación:

- Pregunta de investigación
- Revisión bibliográfica
- Formulación objetivos e hipótesis

### 2-Elección del diseño

- Tipo de estudio

### 3-Planificación de la población de estudio

- Criterios de selección
- Reclutamiento sujetos
- Tipo muestreo
- Cálculo del tamaño muestral

## 4-Planificación variables de estudio

**-Selección**

**-Clasificación.**

**-Definición**

**-Medición**

## Esquema de un Protocolo FIS

-Antecedentes y estado actual del tema

-Bibliografía

-Hipótesis, objetivos

-**Metodología:** Diseño, sujetos de estudio, **variables**, recogida y análisis de datos y limitaciones

-Revisión bibliográfica

-Formulación objetivos e hipótesis

-Plan de trabajo

-Experiencia del equipo investigador sobre el tema:

-Plan de difusión

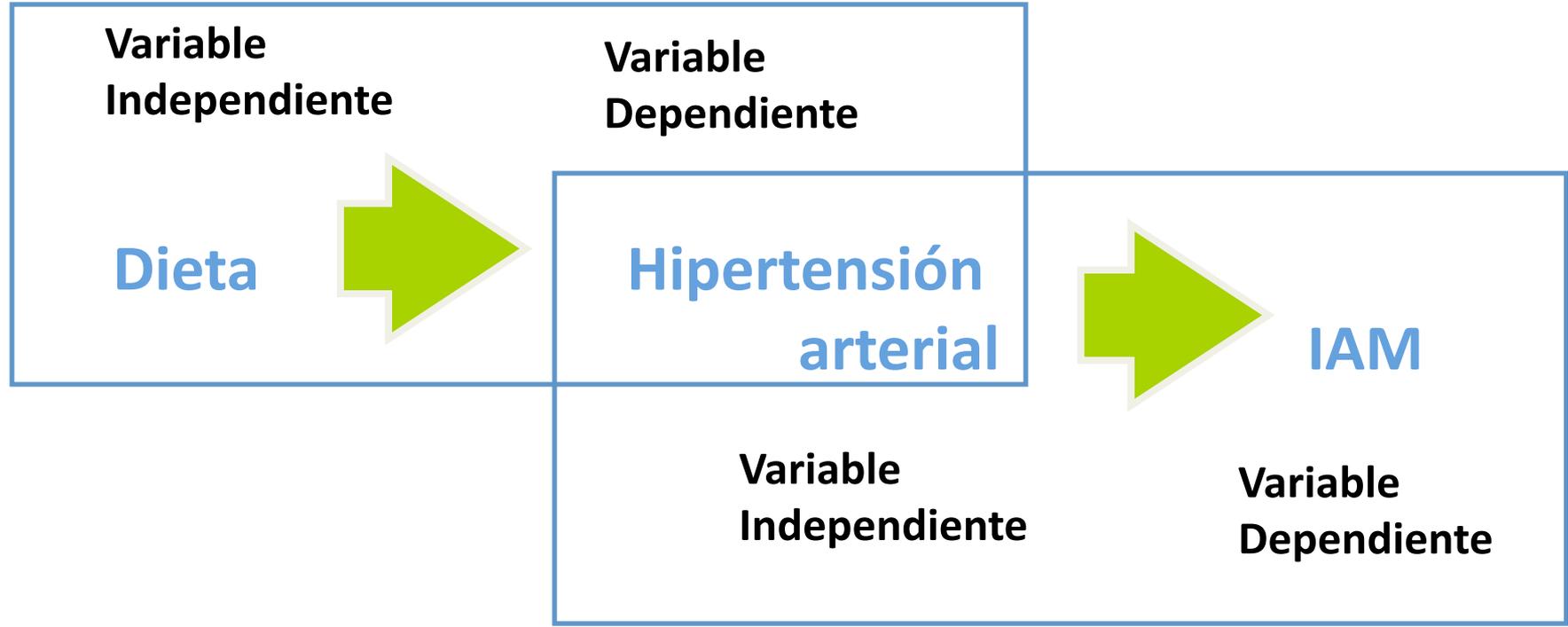
-Medios disponibles

-Justificación detallada

## Clasificación funcional de las variables

- En investigación se clasifican en tres grandes grupos:
  - Variable de respuesta
  - Factor de Estudio
  - Otras variables
- Esta clasificación funcional se hace en base a los objetivos del estudio=en base a la relación causa-efecto que se esté estudiando
- En términos estadísticos la cosa se simplifica aún mas en:
  - Variables dependientes
  - Variables independientes

**Variables dependientes e independientes**



## Clasificación funcional de las variables

- **Variable de respuesta (“response”)**= v. resultado (“outcome”). Endpoint en ensayos clínicos. Estadísticamente = variable dependiente.
- **Factor de estudio**= v. **independiente** principal=exposición
- **Otras variables= covariables (“covariates”)** (estadísticamente serían también v independientes)
  - **Factores de confusión**
  - Variables modificadoras del efecto (interacción)

## Selección de las variables

- Tantas como sean necesarias y tan pocas como sea posible.
- Para ello es fundamental conocer suficientemente el tema objeto de la investigación
- Para ello es necesario recurrir a publicaciones previas que nos orientarán sobre las variables a incluir
- Las variables seleccionadas deben representar adecuadamente los fenómenos que se pretenden medir, ya que de otro modo el estudio carecerá de validez

## Selección de las variables

- Tendencia habitual a tratar de recoger el mayor nº posible de variables (no es posible volver para atrás)
- En la fase previa al inicio siempre se debe repasar el listado de variable incluidas y responderse a las siguientes preguntas:
  - Verdaderamente son de interés para el estudio?
  - Pueden recogerse adecuadamente?
  - Permiten alcanzar los objetivos planteados?
  - Es necesario recoger otras variables no consideradas inicialmente?

Cuando mi colega Petr Skrabanek trabajaba como científico de laboratorio, alguien comentó: *“necesitamos desesperadamente una balanza nueva”*. Su respuesta fue: *“No necesitamos una balanza nueva, necesitamos una chaise-longue, necesitamos pensar”*. La investigación que realmente tiene algún valor se empieza pensando, no haciendo.

James McCormick, 1994

### V RESPUESTA:

- Peso al nacimiento: Variable continua medida en gramos.

### V EXPOSICION:

- Exposición en casa: variable dicotómica nominal (si/no) según la pareja sea o no fumadora.

•Solo se utiliza la variable consumo de tabaco por su pareja como forma de medición → Problema de validez en su definición?

- Parejas fumadora que no fuman en casa en presencia de su mujer embarazada?
- Nº medio de horas de exposición a la semana en casa?
- Debería medirse exposición a otras fuentes (lugares trabajo, ocio...) Definición mas válida
- Niveles de cotinina en saliva ??

## Definición de las variables

- Es conveniente adoptar definiciones estándar, utilizadas y validadas por otros investigadores (publicaciones previas).
- Permite la comparación de los resultados obtenidos con los de otros estudios con una definición similar de las variables de interés
- **Definición operativa:** Explicar claramente lo que realmente se va a medir y con una sola posible interpretación (sin ambigüedades).

### Medición de las variables

- Tras la definición de las variables se debe especificar su escala de medida:
- Solo hay dos tipos de escalas: cualitativa= categorica y cuantitativa
- La escala de medida determinará el análisis estadístico que podrá realizarse

## Tipos de escalas de medidas de las variables

- 1) Escala categórica o cualitativa:
  - 1.1) Nominal (valores=categorías no numéricas bien definidas)
    - 1.1.1) Dicotómica o binaria (solo dos posibles valores excluyentes)
      - Ej: Hipertensión (si / no)
    - 1.1.2) Politómica (>2, mutuamente excluyentes y carecen de orden interno)
      - Ej: Estado Civil (soltero, casado...)
  - 1.2) Ordinal (las categorías pueden ordenarse por rangos)
    - Ej: Clase social (alta, media, baja)
    - Dolor (grave, leve, moderado, no dolor)

## Tipos de escalas de medidas de las variables

- 2) Escala cuantitativa:
  - 2.1) Discreta (solo pueden adoptar ciertos valores, en general numeros enteros)
    - Ej: nº de hijos (1, 2, 3, 4...)
    - Ej: nº de ingresos hospitalarios
  - 2.2) Contínua (cualquier valor numerico, teoricamente infinito aunque generalmente dentro de un rango)
    - Peso, Glucemia basal, consumo de alcohol

### Tipos de escalas de medidas de las variables

- **Cefalea?**
- **Grado de cefalea**
- **Tabaco?**
- **Grado de fumador?**
- **Proteinuria (+, ++, +++, +++++)?**
- **Clase social (I, II, III, IV)?**
- **Edad?**
- **Frecuencia cardiaca?**
- **Bradicardia?**
- **Severidad de Bradicardia?**
- **Presión arterial?**
- **Hipertensión arterial?**

**Matriz de datos**

Id	Sexo	Nivel socio-económico	Consumo de alcohol	Nivel consumo alcohol	Expuesto a plomo
212	0	2	1	2	0
123	1	9	0		1
421	1	1	1	1	
158	0	2	1	0	2
...	...	...	....	...	...
12	0	0	1	0	1
152	1	1	1	1	2

Hombre = 0  
Mujer = 1

Bajo = 0  
Medio = 1  
Alto = 2  
Missing = 9

No = 0  
Sí = 1  
No contesta=8  
Missing=9

Bajo = 0  
Medio = 1  
Alto = 2  
Missing=

No= 0  
Si= 1  
No contesta=8  
NS/NC=88  
Missing=

## Tipos de escala de medida de las variables

- Numerosas variables pueden expresarse como variables nominales dicotómicas, ordinales o cuantitativas continuas.
- Si los datos lo permiten, siempre en escala continúa.
  - Contiene mas información, lo cual permite utilizar tecnicas estadisticas mas potentes (aumento de la potencia del estudio para establecer asociación entre variables)
  - Si el análisis lo requiere puedes agrupar los datos en categorias cualitativas, pero el proceso inverso no es posible.
- Siempre en función de la importancia de la variable en el estudio

## Descripción de las variables en el protocolo FIS

- Esencial la definición operativa de las variables
    - Especificar que es lo que queremos medir (definición)
    - Como se obtiene la información (fuente y forma de recogida)\*
    - Escala de medición (unidades de medida)
    - Establecer las categorías de cada variable (punto de corte).
  - Fundamental para las variables respuesta y exposición que se estimen pertinentes
  - Algunas variables de control socio-demográficas no requieren una definición operativa por su obviedad (ej sexo, edad...).
- \* Se podría poner también en el apartado “recogida de datos”

### V RESPUESTA:

- Peso al nacimiento: Variable continua medida en gramos.

### V EXPOSICION:

- Niveles de cotinína en saliva: variable continúa medida en ng/ml.

### •V RESPUESTA

- Especifica únicamente escala y unidad de medida=incompleta. Es realmente variable respuesta?
- La variable debería llamarse Bajo Peso y definirse como tal=recién nacido con peso <2.500 gr
- La forma de recogida de información y la fuente deberían especificarse=Peso medido en la sala de partos inmediatamente tras el nacimiento y registrado en la hoja de partos de la historia clínica.

### V RESPUESTA:

- Peso al nacimiento: Variable continua medida en gramos.

### V EXPOSICION:

- Niveles de cotinina en saliva: variable continua medida en ng/ml.

### •V EXPOSICION

- En que momento se toma y con que técnica?
- La variable debería referir la técnica utilizada (por ejemplo RIA), el momento de medición (en el tercer trimestre...)
- Si la fuéramos a usar categorizada, deberíamos poner el punto de corte

## Codificación de las variables en la base de datos

- 1º -> diferenciar entre variables categóricas y continuas:
- **Categóricas:**
  - Se debe asignar un código para cada categoría
  - Antes recomendación de no poner 0 (los programas estadísticos no diferenciaban 0 de casilla en blanco=missing). **AHORA SIEMPRE NO EXPTO=0, EXPTO=1.**

## Codificación de las variables en la base de datos

- **Continuas**
  - En la base de datos se plasman con las unidades completas con que se midieron
  
- **Casos especiales**
  - El individuo no contesta (no le da la gana)-> No contesta (8)
  - El individuo no sabe (no contesta porque no sabe la información que se le solicita)-> No sabe (88)
  - Ausencia de información -> por error del entrevistador, ausencia de información en los registros como hª clínica -> missing -> se deja en blanco, o un 9
  - Datos alfanuméricos (nombre, apellidos, nºhistoria [letras+numeros]-> variable cadena
  - Fecha: dia/mes/año -> diferentes ordenes, que hay que comprobar dependiendo en que soporte hagamos el volcado de datos y analicemos. En nuestro país siguiente orden (dd-mm-aaaa). Campos fecha-> años con cuatro cifras

## Comparación entre dos variables

	V Dependiente	
V Inpte	Cuantitativa Normal	Cualitativa Dicotomica
Cuali >2 categorías	ANALISIS DE LA VARIANZA (ANOVA)	Chi <sup>2</sup>
Dicotómica (2 categorías)	T-Student	Chi <sup>2</sup>

## Análisis multivariante

- Hay dos aproximaciones para realizar el ajuste estadístico de la confusión:
  - el análisis estratificado
    - Permite examinar también la **modificación de efecto (interacción)**
  - el análisis multivariante con modelos de regresión

## Análisis multivariante

- Análisis multivariante= análisis de regresión
- Las variables independientes se llaman también covariables (“covariates”) = (“regressor variables”).
- La variable dependiente (“dependent variable”) se llama también (“regressand”)
- Para hacer regresión hay diferentes modelos matemáticos de regresión.
- El uso de uno u otro modelo vuelve a estar determinado por el tipo de variables dependientes e independientes

## Elección del modelo de regresión en función de los tipos de variables

- Variable dependiente = **cuantitativa continua**:
  - modelo de regresión lineal múltiple (“multiple lineal regression”) si  $v$  independientes son cuantitativas (tiempo de reacción y consumo alcohol g/l)
  - Modelo de análisis de la covariancia (ANOVA) si  $v$ . independientes son cualitativas (tiempo de reacción y consumo nulo, bajo y alto de alcohol)

## Elección del modelo de regresión en función de los tipos de variables

- Variable dependiente = binaria
  - (Enfermo o No enfermo) modelo de regresión logística (“logistic regression”) (cáncer de esófago y exposiciones laborales)
  - Cuando se coge tiempo de supervivencia → modelo de riesgos proporcionales de cox (“Cox proportional hazards model”) (supervivencia en cáncer de esófago en función del tamaño del tumor)