

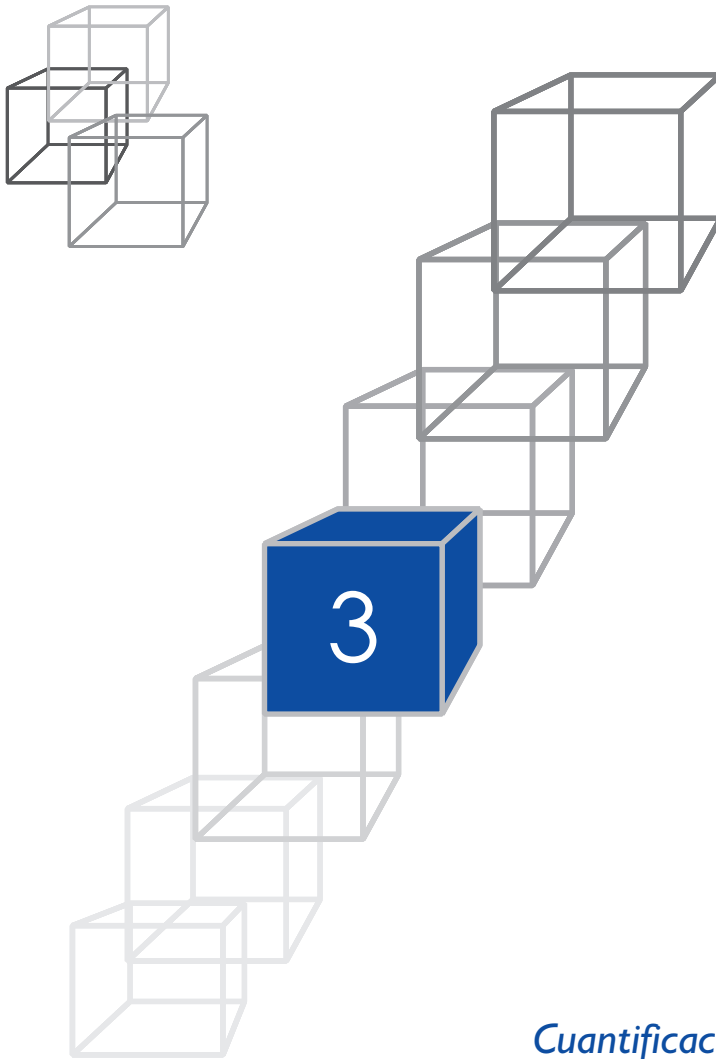
Epidemiología Básica
y
Vigilancia de la Salud

segunda edición

2004



*Módulos de
Epidemiología Básica
y Vigilancia de la Salud*



*Cuantificación de los
Problemas de Salud*



The World Bank

*Todos los derechos reservados.
Este libro no puede reproducirse total o parcialmente por ningún método gráfico,
electrónico o mecánico, incluyendo los sistemas de fotocopia,
registro magnetofónico o de alimentación de datos, sin expreso consentimiento del autor.
Queda hecho el depósito que previene la Ley 11.723
Buenos Aires, República Argentina. 2001
Este material fue realizado con el apoyo financiero del Programa VIGI+A
(Ministerio de Salud-Banco Mundial)*

Autores

Ortiz, Zulma

Médica, Reumatóloga, Universidad de Buenos Aires (UBA)
Gerente de Vigilancia de la Salud
Programa Nacional de Vigilancia de la Salud y Control de Enfermedades (VIGI+A)

Esandi, María Eugenia

Médica, (UBA)
Docente del Curso de Epidemiología Básica e Intermedia,
Centro de Investigaciones Epidemiológicas (CIE),
Academia Nacional de Medicina, de Buenos Aires

Bortman, Marcelo

Médico General (UBA), Epidemiólogo
Coordinador General
Programa Nacional de Vigilancia de la Salud y Control de Enfermedades (VIGI+A)

Revisores

Custer, Silvina
Eiman Grossi, Mirtha

Rico Cordeiro, Osvaldo
Rodríguez Loria, Gabriela

Procesamiento didáctico

Davini, Cristina

Goldenstein, Frida

Lomagno, Claudia

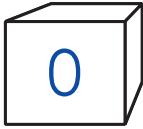
Colaboradores

Abdala, Yamile
Bernardos, Jaime
Bonet, Fernanda
Cabrini, Ana
Carbonelli, Natacha
Córdoba, Patricia
Diana, Anaí
Donnet, M. Isabel

Esandi, Pablo
Galdeano, Emilio
Goizueta, Miguel
Gómez, Julio
Insúa, Iván
Laurynowycz, Alicia
Maidana, Cristina
Ortiz, Carina

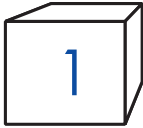
Piccini, Mabel
Ramírez, Rolando
Rulfo, Ana
Tupá, Daniela
Vera del Barco, Pablo
Verdejo, Guadalupe
Videla, Mitha
Yáñez, Loreto

Contenido de los Módulos



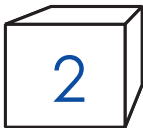
Módulo del Capacitador

Dirigido a quienes desempeñan el rol de Capacitadores. En él se incluye información y orientación para desarrollar actividades como tutor y facilitador del aprendizaje individual y grupal. Al completar este Módulo, estará en condiciones de comprender y programar su trabajo como Capacitador.



Introducción a la Epidemiología

Describe qué es la Epidemiología, los cambios de sus concepciones en la historia y sus tendencias actuales. Presenta los métodos epidemiológicos y contribuciones de la Bioestadística. Al completar el trabajo con este Módulo, dispondrá de elementos para comprender los aportes que la Epidemiología puede brindar para mejorar los procesos de planificación, ejecución y evaluación de los Servicios de Salud.



Tipos de Estudios Epidemiológicos

Introduce el proceso de investigación epidemiológica en sus distintos diseños, analizando las ventajas y las dificultades que se presentan. Al finalizar el trabajo con este Módulo, podrá seleccionar un tipo de diseño apropiado para el estudio que se propone realizar.



Cuantificación de los Problemas de Salud

Recomienda qué datos recolectar y cómo recopilarlos, procesarlos, interpretarlos y presentarlos. Al finalizar el trabajo con este Módulo, podrá elaborar un plan para la cuantificación de un problema de Salud local.



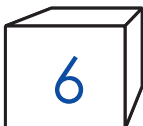
Efecto, Impacto y Fuentes de Error

Presenta cómo cuantificar el riesgo a través de diferentes medidas que permiten establecer la existencia de asociación entre diferentes factores y el evento estudiado. Al completar el desarrollo de este Módulo, se espera que pueda cuantificar el riesgo en el marco de la epidemiología analítica y describir las principales fuentes de error en la cuantificación.



Vigilancia de la Salud

Aborda qué es la Vigilancia de la Salud, sus usos en el área de la Salud Pública, cómo debería funcionar un sistema para alcanzar las metas de Salud deseadas. Al finalizar este Módulo, se espera que pueda reconocer las diferentes estrategias y, fundamentalmente, cómo analizar, interpretar, difundir y utilizar la información.



Investigación de Brote

Presenta cómo realizar una investigación sobre un Brote, saber el tipo de información que se puede necesitar, qué buscar, a qué prestar atención y qué significado tienen los resultados obtenidos. Se espera que al finalizar el Módulo, pueda describir cómo realizar un estudio de Brote y redactar el informe.

Prólogo

Considerando el rol que la Epidemiología debe cumplir en la Salud Pública actual, la capacitación en Epidemiología Básica para los niveles locales es un punto crítico para alcanzar objetivos de Salud para Todos. De forma similar, el desarrollo de la Vigilancia de la Salud, herramienta fundamental para la toma de decisiones en Salud, resultará esencial en la migración del quehacer en Salud hacia una gestión cada vez más basada en la evidencia.

Estos Módulos del Curso de Capacitación en Epidemiología Básica y Vigilancia de la Salud son el producto de una iniciativa de la Representación de OPS/OMS en la Argentina y del Programa Especial de Análisis de Salud, OPS.

Sus contenidos básicos fueron desarrollados como parte del apoyo que esta Representación realizó a la formulación del Programa VIGI+A, que prevé una amplia disseminación de los principios básicos de la Epidemiología y de la Vigilancia de la Salud. Con el trabajo de sus autores y con los aportes de muchos otros colaboradores se han elaborado estos Módulos que esperamos contribuyan significativamente a una mejor capacitación en Epidemiología.

Dr. Juan Manuel Sotelo
Representante de la Organización Panamericana de la Salud
y Organización Mundial de la Salud

Cuantificación de los Problemas de Salud

MÓDULO 3: Cuantificación de los Problemas de Salud

A. Introducción

B. Objetivos del Módulo

C. Contenidos



1. ¿Qué Datos Debemos Recolectar para la Investigación?

1.1 ¿Qué es una Variable?

1.2 Tipos de Variables y Formas de Clasificación

1.2.1 Variables Dependientes y Variables Independientes

1.2.2 Variables de Persona, Tiempo y Lugar

1.2.2.1 Variables Universales

*Edad, Sexo, Nivel Socioeconómico, Grupo
Étnico y Cultural, Ocupación, Estado Civil*

1.3 Definición de Variables

1.4 ¿Cuántas Variables deben Considerarse en un Estudio?

1.5 ¿Cómo se Miden las Variables?

Escalas de medición: nominal, ordinal, numérica

2. ¿Dónde y Cómo Obtener los Datos para la Investigación?

2.1 ¿Dónde Recolectar Datos?

2.1.1 Fuentes Secundarias de Datos

*Fuentes Secundarias Permanentes, Fuentes
Secundarias Eventuales*

2.1.2 Fuentes Primarias de Datos

2.2. ¿Cómo Recolectar los Datos?

Encuestas de Salud

3. ¿Cómo Procesar e los Datos?

3.1 Presentación de Datos en Tablas

3.1.1 Tabla de Distribución de Frecuencias

3.1.2 Tabla de Frecuencias con Intervalos de Clase

3.2 Presentación de Datos en Gráficos

3.2.1 Gráfico de Sectores Circulares

3.2.2 Gráfico de Barras

3.2.3 Histograma

3.2.4 Polígono de Frecuencias

4. ¿Cómo Interpretar los Datos?

4.1 La Sistematización de los Datos

4.2 Las Medidas de Resumen

4.2.1 Medidas de Frecuencia

*Razón, Proporciones o Porcentajes, Tasas, Tasas
Brutas y Tasas Específicas, Comparación de Tasas,
Prevalencia, Incidencia*

4.2.2 Medidas de Tendencia Central

Moda, Media Aritmética, Mediana

4.2.3 Medidas de Orden

Percentiles, Cuartiles

4.2.4 Medidas de Dispersión

Rango, Rango Intercuartílico, Desvío estándar

D. A Modo de Síntesis...

Cuantificación de los Problemas de Salud

A. Introducción

La Organización Panamericana de la Salud (OPS) y el Programa VIGI+A, en un trabajo colaborativo han posibilitado la preparación, difusión y diseminación de este material educativo que contiene conceptos básicos de Epidemiología y Vigilancia de la Salud. En este Módulo, se presentarán los métodos para la cuantificación de los problemas en Salud.

Como ya hemos tratado, la Epidemiología intenta describir la distribución y determinantes de la Salud y enfermedad en poblaciones humanas. Si bien esto puede parecerle algo complicado, en realidad, la metodología de investigación no difiere de la que seguramente Ud. aplica durante su práctica laboral, aun en las cosas cotidianas de la vida. Tal como presentamos en el Módulo anterior, se reconocen una serie de etapas en el proceso de investigación (figura 1), que se inician con la definición del problema y los objetivos. Sin embargo, es necesario seguir avanzando ya que un paso previo e indispensable para interpretar los procesos de Salud y enfermedad es la cuantificación de los datos observados. Probablemente Ud. se pregunte: "Pero, ¿qué datos debo recolectar y cómo?". Qué y cómo recolectar los datos es lo primero que trataremos en este Módulo. Una vez que comprendamos esto, veremos qué hacer con estos datos, es decir, cómo procesarlos e interpretarlos.

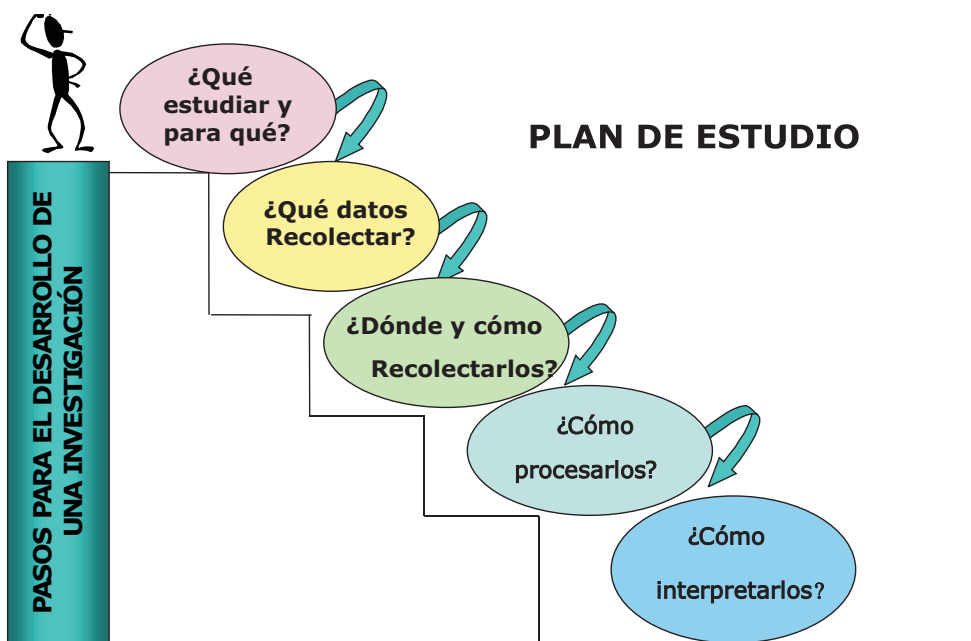
B. Objetivos del Módulo

Se espera que al completar el trabajo en este Módulo, Ud. sea capaz de:

- ▶ Elaborar un plan de recolección de datos epidemiológicos en función de un problema de Salud local: Reconociendo distintos tipos de variables, seleccionando

diferentes fuentes de información, construyendo instrumentos de recolección, procesamiento e interpretación de la información y seleccionando la mejor forma de presentación para sus datos.

Figura 1. Pasos en el proceso de investigación de un evento de Salud



C. Contenidos

1. ¿Qué Datos Debemos Recolectar para la Investigación?

Un dato es el valor que adopta la variable. Los datos que recolectaremos para nuestra investigación dependen de las variables que escojamos. Ahora, ¿qué es una variable? Si bien a menudo utilizamos este término, poco conocemos acerca de él.

1.1 ¿Qué es una Variable?

Como su nombre lo indica, "variable es algo que cambia".

Por ejemplo:

Peso
Estatura
Frecuencia cardíaca
Edad
Temperatura

Definición:

Una variable puede ser definida como cualquier cualidad del organismo, grupo o situación capaz de adoptar valores diferentes en un mismo individuo o entre individuos.

Si pensamos en los datos del ambiente que nos rodea, observaremos que casi todos los aspectos pueden considerarse variables, precisamente porque las personas y las condiciones ambientales varían. En los fenómenos de la Salud y de la enfermedad es necesario investigar no sólo cómo varían (epidemiología descriptiva), sino también por qué varían (epidemiología analítica).

Ya mencionamos que los datos que recolectamos en nuestra investigación dependen de las variables que seleccionamos y, por otra parte, que casi todos los aspectos pueden considerarse variables. Entonces, ¿qué variables debemos seleccionar? Las variables a estudiar se seleccionan sobre la base de su relevancia para los objetivos de la investigación. Para facilitar el proceso de identificación y selección de las variables, nos será útil conocer los tipos de variables y sus diferentes formas de clasificación.

1.2 Tipos de Variables y Formas de Clasificación

Las variables pueden ser clasificadas utilizando diferentes criterios.

FICHA DE ORIENTACIÓN N° 1:

Formas de clasificar las variables



- **Según la función que cumplan en la hipótesis**

- Variables dependientes
- Variables independientes

- **Según la pregunta que responden en el estudio**

- ¿Quiénes?: Variables de Persona
 - ¿Cuándo?: Variables de Tiempo
 - ¿Dónde?: Variables de Lugar
-

1.2.1 Variables Dependientes y Variables Independientes

Las distintas variables que se consideran en un estudio mantienen, en general, relaciones entre sí. Muy frecuentemente un estudio se centra no sólo en las características en las cuales hay realmente interés, sino también en una o más características asociadas a ella.

Por ejemplo, no sólo puede interesarnos investigar la ocurrencia de anemia ferropénica en la población infantil de una región determinada en un determinado período de tiempo, sino también, los factores potencialmente causales de esta enfermedad.

Denominamos variable dependiente a aquella variable que el investigador está interesado en comprender, explicar o predecir, mientras que las variables independientes son aquellas que utilizamos para explicar a la primera. Considerando una asociación entre variables del tipo "causa - efecto", la o las causas serán variables independientes y el efecto será la variable dependiente.

En nuestro ejemplo, la variable dependiente sería la ocurrencia de la anemia ferropénica y las variables independientes, los factores que explicarían dicha ocurrencia de la enfermedad.

FICHA INSTRUMENTO N° 1

Pensar y resolver



A continuación le presentamos un problema de investigación. Formule el/los objetivos e identifique la variable dependiente y la variable independiente.

"En el año 1999, se observó en la Capital Federal un elevado número de suicidios en la población adolescente; pudo saberse a través de fuentes informales que muchos de estos casos ocurrieron en adolescentes adictos a las drogas que en ese último período se habían iniciado en el consumo de una nueva droga importada desde Europa."

1.2.2 Variables de Persona, Tiempo y Lugar

Como se mencionó en el Módulo 1, tres son los determinantes principales de los procesos de salud-enfermedad: persona, lugar y tiempo. Sus atributos o características pueden ser factores o marcadores de riesgo que "determinan" la aparición de un evento, enfermedad

Cuantificación de los Problemas de Salud

o el estado de salud. Por eso, cada vez que desarrollemos un plan de estudio, deberemos preguntarnos en quiénes, cuándo y dónde sucede el evento de Salud o enfermedad.

Las variables que permiten responder concretamente a la pregunta "¿en quiénes sucede el evento?" son conocidas como variables de persona.

Por ejemplo, al investigar una enfermedad podríamos estar interesados en establecer el sexo de las personas afectadas, su edad, profesión, ingresos, hábitos alimentarios, etc. Todas éstas son características de las personas enfermas que permiten caracterizar mejor a quiénes afecta la enfermedad en cuestión.

Las variables que permiten responder concretamente a la pregunta "¿cuándo sucede el evento?" son conocidas como variables de tiempo. Para medir el tiempo en el que ocurre un evento de Salud existen diferentes formas:

Por ejemplo, el tiempo de calendario es el que más a menudo se utiliza, como en los estudios anuales de mortalidad (del 1° de enero al 31 de diciembre del año en curso). Sin embargo, también podemos utilizar el tiempo cronológico si estamos interesados, *por ejemplo*, en el tiempo que transcurre entre la exposición y el daño, tal el caso de las enfermedades infecciosas: ¿cuánto tiempo transcurre entre la exposición de un susceptible a un paciente con varicela y la aparición de enfermedad?

Si queremos describir la presencia de un patrón cíclico en determinado evento, podemos tomar el caso de los accidentes de tránsito, que acontecen con más frecuencia en horario nocturno y durante los fines de semana.

Las variables que permiten responder concretamente a la pregunta: "¿Dónde sucede el evento?" son conocidas como variables de lugar.

Por ejemplo, si quisiéramos investigar la ocurrencia de tuberculosis en una ciudad, será

interesante determinar, entre otras cosas, las características relacionadas con el lugar en el que sucede el evento, como domicilio y zona de residencia de las personas afectadas.

1.2.2.1 Variables Universales

La definición de las características relacionadas con las personas son de singular importancia para cualquier investigación epidemiológica. Existen dentro del grupo de variables de persona ciertas variables llamadas "universales" por la importancia que tienen en las investigaciones sobre grupos poblacionales, tanto que se sugiere considerar siempre la necesidad de su inclusión. Esto no significa que haya que incluirlas automáticamente, sino simplemente que es necesario valorar siempre la posibilidad de hacerlo. Estas variables universales son: edad, sexo, grupo étnico y cultural, ocupación, nivel socioeconómico y estado civil.

Edad

Siempre que desee realizar una investigación debe considerar la descripción de esta variable. Las variaciones en la frecuencia de distintas enfermedades en función de los cambios de la edad de las personas son la regla. En numerosas ocasiones, esas variaciones de acuerdo a la edad le permitirán conjeturar acerca de los factores responsables del desarrollo de una enfermedad.

Por ejemplo, la frecuencia de la enfermedad cardiovascular en las mujeres jóvenes es mucho menor respecto de la de las mujeres mayores de 50 años. Este patrón de la enfermedad corroborado por diferentes estudios, permitió observar que su frecuencia aumenta a partir de los 50 años, probablemente debido a la menopausia y al descenso de los niveles de estrógenos. Este hallazgo hace suponer la presencia de un efecto protector de las hormonas sexuales femeninas en el desarrollo de la enfermedad cardiovascular.

Sin embargo, también se podría pensar que esa diferencia en la frecuencia de la

Cuantificación de los Problemas de Salud

enfermedad podría ser explicada por otros factores: por ejemplo, las mujeres mayores de 50 años están habitualmente menos predispuestas a hacer ejercicios físicos, o no tienen tanto interés en "los kilos de más" y llevan una dieta menos equilibrada, etc. Todos estos factores están también asociados con una mayor frecuencia de enfermedad cardiovascular.

FICHA INSTRUMENTO N° 2

Pensar y resolver



Relea el ejemplo anterior, relativo a las enfermedades cardiovasculares en mujeres mayores de 50 años e indique cuál/es sería/n la/s variable/s dependiente/s y la/s independiente/s, según las distintas posibilidades enunciadas.

Sexo

Al igual que la edad, el sexo es otra variable que se debe considerar. Casi todas las enfermedades ocurren de forma diferente de acuerdo al sexo de las personas. Podríamos citar la mayor frecuencia del cáncer de pulmón en los hombres o del cáncer de mama en las mujeres. A menudo, encontrará en sus investigaciones variaciones en la frecuencia de la enfermedad de acuerdo al sexo.

Sin embargo, siempre debe tener en cuenta que en la mayoría de las poblaciones los dos sexos también difieren en cuanto a la distribución por edades. Siempre debe mirar con cuidado las diferencias encontradas entre las distintas frecuencias de la enfermedad en estudio de acuerdo al sexo y edad.

FICHA INSTRUMENTO N° 3

Pensar y resolver



Interprete la siguiente Tabla 1 identificando variaciones en la distribución de la enfermedad según sexo y edad. Resuma los hallazgos que considere más importantes.

Tabla 1. Distribución por sexo y edad de 1.859 casos de Tuberculosis declarados en la Ciudad (X), durante 1997.

	0-4	5-14	15-24	25-34	35-44	45-54	55-64	65 y más
Hombres 1.326 (71.3 %)	72	57	109	364	263	225	94	142
Mujeres 533 (28.7%)	52	47	99	134	64	37	24	76
Total 1.859 (100%)	124 (6,7 %)	104 (5.6 %)	208 (11.2%)	498 (26.8%)	327 (17.6%)	262 (14.1%)	118 (6.3%)	218 (11.7 %)

Fuente: Departamento de Estadísticas de la ciudad X

Nivel Socioeconómico

El nivel socioeconómico hace alusión, fundamentalmente, a los ingresos percibidos por el individuo o en el grupo familiar. Esta variable se encuentra asociada a muchas y diversas situaciones de Salud, tales como la alimentación, la exposición a factores tóxicos ambientales, el acceso a servicios de Salud, educación, etcétera.

Grupo Étnico y Cultural

El grupo étnico y cultural se refiere al conjunto de personas que guardan similitud respecto de las costumbres (religión, historia, hábitos dietéticos, etc.) y también a la herencia biológica (raza) que los distingue de la población general.

La "raza" (blanca, negra, amarilla) nos permite identificar grupos de personas de color y rasgos físicos comunes. Sin embargo, la definición "grupo étnico" es más amplia y expresa mejor el concepto epidemiológico al que nos referimos, ya que generalmente resulta más útil investigar la frecuencia de una enfermedad para grupos cuyos miembros han vivido juntos y se han casado o entremezclado por un período de tiempo suficiente como para haber adquirido características comunes, tanto por mecanismos biológicos como por mecanismos sociales. No se puede dejar de lado el hecho de que las diferentes etnias se encuentran íntimamente relacionadas con otra variable que es el nivel socioeconómico y muchas veces en realidad lo que se atribuye a determinada raza está en realidad condicionado por el nivel socioeconómico al cual pertenece. Esto se hace aún más evidente con los grupos étnicos que migran, debido a que llevan consigo no sólo el bagaje genético y cultural que le es propio, sino también la carga de los condicionantes que los han llevado a emigrar. En nuestro país, los diversos registros existentes no tienen datos sobre las etnias, por lo que es difícil analizar diferentes perfiles asociados a dicha variable.

Ocupación

Las distintas ocupaciones tradicionalmente se encuentran asociadas a diferentes factores de exposición y por lo tanto a distintas enfermedades. Afortunadamente, se consigna la ocupación en muchos registros llevados en forma sistemática y se usa como variable epidemiológica descriptiva con diferentes objetivos. Esto permite identificar riesgos específicos de Salud asociados a ciertas ocupaciones, como por ejemplo, la mayor incidencia de cáncer de pulmón en los trabajadores de minas de carbón.

FICHA INSTRUMENTO N° 4

Pensar y resolver



Interprete la siguiente tabla identificando las variaciones en la prevalencia de dolor de pecho en las distintas poblaciones, según edad y sexo. A partir de sus observaciones elabore una hipótesis.

Tabla 2. Frecuencia de dolor de pecho en diferentes poblaciones

Lugar	Grupos de edad	Hombres %	Mujeres %
Población general España	45-54	5,3	6,4
	55-64	5,0	7,1
	65-74	5,4	8,8
Total España	45-74	5,2	7,7
Población general Islandia	45-54	3,4	2,0
	55-64	5,1	2,5
	65-74	8,9	3,1
Población general Italia	20-29	0,8	1,7
	30-39	0,4	1,9
	40-49	1,9	3,9
	50-59	5,0	6,7
Estados Unidos Población Negra	25-34	5,1	5,2
	35-54	7,3	6,8
	55-74	6,0	8,5
Estados Unidos Población Blanca	25-34	2,0	5,2
	35-54	3,1	6,1
	55-74	6,9	7,5
Estados Unidos Hispanos	25-34	1,9	4,5
	35-54	2,3	6,0
	55-74	4,5	5,4





Fuente: Revisión Bibliográfica, 1999

FICHA DE REFLEXIÓN N° 1

Para pensar y analizar...



- Le pedimos que observe las siguientes figuras. Cada una de ellas se refiere a distintas ocupaciones.
- ¿Podría Ud. identificar desde sus conocimientos previos cuál/es serían las enfermedades, efectos, cuadros o lesiones que se pueden presentar como consecuencia de las siguientes ocupaciones?

OCUPACIÓN	RIESGO DE
	
	
	
	

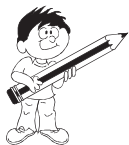
Estado civil

Esta variable ha sido estudiada en relación con el estado de Salud. El estado civil es más fácil de medir pudiéndose efectuar una rápida división de la población en cuatro grupos mayores: soltero, casado, viudo y divorciado.

La mayoría de los estudios ha demostrado que las enfermedades tienen más incidencia en mujeres y hombres solteros que en mujeres y hombres casados. Si bien existen muchas investigaciones que demuestran la relación entre estado civil casado y Salud, su uso está perdiendo vigencia en la actualidad, pues la mayoría de las características de la vida matrimonial, que muy probablemente tienen un efecto positivo sobre la Salud, está extendiéndose a otros tipos como las parejas que viven juntas o en concubinato.

FICHA INSTRUMENTO N° 5

Pensar y resolver



Interprete la Tabla 3 de la página siguiente identificando las variaciones en las tasas de mortalidad de individuos solteros y casados. La tasa de mortalidad es una medida de frecuencia de la mortalidad en los grupos estudiados (varones y mujeres solteros y casados de la Ciudad X).

Veremos en detalle la construcción de las mismas y su significado más adelante en el Módulo.

A partir de sus observaciones elabore una hipótesis.

Tabla 3. Tasa de Mortalidad de individuos solteros y casados en Ciudad X, 1989-1991

Causas seleccionadas de defunción	Hombres		Mujeres	
	Solt.	Casad.	Solt.	Casad.
Tuberculosis respiratoria	96,7	22,7	25,0	11,4
Otras formas de tuberculosis	3,4	1,3	1,8	0,8
Neoplasmas malignos, total	260,0	207,1	270,4	182,5
Diabetes mellitus	21,9	18,3	16,5	29,0
Influenza y neumonía	73,5	27,0	34,1	18,3
Úlcera péptica	21,5	11,8	3,8	2,4
Accidentes automovilísticos	61,9	33,7	11,0	10,8
Otros accidentes	99,2	46,1	36,4	21,1
Suicidio	43,7	20,5	7,7	6,2

Tasa promedio anual por 100.000 habitantes. Fuente: Estadísticas Vitales de la ciudad (X)

1.3 Definición de Variables

Al realizar una investigación epidemiológica, cualquiera sea el tipo de estudio, no sólo es importante identificar las variables que se toman en consideración, sino definir las con la mayor precisión posible. Debe diferenciarse lo que significa definir teóricamente una variable y lo que significa definirla operativamente, proceso que recibe el nombre de operacionalización de una variable. Operacionalizar una variable es simplemente definir la manera en que se observará y medirá cada variable del estudio y tiene un sentido práctico fundamental.

Por ejemplo, podría definirse como desempleo: "Cualquier persona que no trabaja y busca trabajo". Esta es la definición teórica de la variable. Pero si a esta definición le agregamos "en los 6 meses previos al comienzo del estudio", la convertimos en una definición de tipo operacional. Sin embargo, también podríamos haber definido desempleo como "cualquier persona sin empleo fijo en el último año". La definición teórica de la variable no cambia ("desempleo"), pero la definición operacional es diferente de acuerdo a cómo se elija medir dicha variable.

No todos pueden estar de acuerdo con la forma en que el investigador definió sus variables desde el punto de vista teórico u operacional. Sin embargo, comunicar con exactitud lo que dichos términos significan, permite que el estudio pueda ser reproducido por otros investigadores y comparar sus resultados. Si se desconociera la forma en que se definieron las variables, no se podría establecer si se está comparando lo mismo o cosas diferentes.

FICHA DE REFLEXIÓN N° 2

Para pensar y analizar...



A esta altura de nuestro trabajo, es importante que realicemos el análisis de un ejemplo, para reconocer en un caso concreto las variables, sus definiciones operacionales y las relaciones entre variables dependientes e independientes.

Supongamos que Ud. está interesado en conocer la relación que existe entre la pobreza y el incumplimiento del tratamiento para la Tuberculosis:

- ¿Cómo definiría pobreza?
- ¿Es una variable dependiente o independiente?
- La pobreza ¿es un fenómeno uni o multi-dimensional?
- ¿Dónde podría obtener una definición teórica de pobreza que sea ampliamente utilizada?
- ¿ Considera importante, incluir la percepción de la gente pobre respecto de la pobreza para preparar la definición operativa de la misma?

1.4 ¿Cuántas Variables deben Considerarse en un Estudio?

La única respuesta a esta pregunta es: "tantas como sean necesarias y tan pocas como sea posible". No hay un número exacto de variables a considerar en una investigación. Sin embargo, es útil que recuerde que en principio suele seleccionarse una lista enorme de variables, a la que de a poco debemos acotar con sentido común y determinando la relevancia que podría tener la información aportada por cada variable seleccionada.

Por otra parte, deberá tener en cuenta siempre la factibilidad para la recolección y el procesamiento de los datos en el momento de seleccionar las variables de su investigación.

1.5 ¿Cómo se Miden las Variables?

Para poder "medir", cuantificar o clasificar las variables que seleccionamos en nuestra investigación debemos tener en cuenta que existen diferentes escalas de medición. Así por ejemplo, si estamos evaluando el sexo de las personas y su edad, rápidamente, percibiremos que ambas variables son diferentes: mientras que la edad puede expresarse en números, el sexo sólo puede medirse en dos alternativas, masculino y femenino. Se reconocen tres tipos de escalas de medición: nominal, ordinal y numérica.

Escala nominal

Se usa para la forma más simple de medición, cuando los datos se distinguen por un nombre que expresa una cualidad, categoría o atributo. Al utilizar este tipo de escala, simplemente, se cuantifica la cantidad de datos que corresponden a cada una de las categorías de la variable. Un ejemplo es, cuando se utiliza este tipo de escala para clasificar el grupo sanguíneo de las personas, cuyas categorías son A, B, AB y O. Al utilizar la escala nominal, lo que estamos haciendo es simplemente clasificar los individuos de acuerdo a cada uno de los grupos sanguíneos. La presencia o ausencia de

enfermedad también puede ser medida con una escala de tipo nominal.

Los datos valorados en esta escala se llaman observaciones cualitativas, categóricas o atributos. Si el número de categorías posibles es dos, la variable se dice dicotómica o binaria como es el caso del "sexo" (hombre o mujer).

Escala ordinal

Permite establecer una relación de "orden" entre los datos que pertenecen a distintas categorías. La escala ordinal permite clasificar los datos de acuerdo a su posición relativa respecto de otros, en función de una variable determinada. A diferencia de la escala nominal, en la que sólo se clasifican los datos de acuerdo a las categorías de la variable, con la escala ordinal los datos se clasifican en categorías y se ordenan conforme a un determinado criterio. *Por ejemplo*, en la variable "nivel socioeconómico" puede distinguirse un orden entre las categorías: alto, intermedio o bajo. Los individuos con un nivel socioeconómico alto tienen características similares entre sí, pero tienen un nivel socioeconómico mayor que los otros dos grupos. Sin embargo, la medición ordinal no informa acerca de cuánto mayor es un atributo respecto de otro nivel. Volviendo al ejemplo anterior, lo que no puede establecerse es "cuanto más" nivel socioeconómico tiene un individuo de nivel alto con respecto a otro de nivel intermedio u otro de nivel bajo.

Escala numérica

Las observaciones individuales realizadas con una escala numérica corresponden a cantidades que provienen de mediciones o recuentos y reciben el nombre de observaciones cuantitativas. La escala numérica a su vez puede diferenciarse en:

- ▶ **Discreta**, cuando los valores que se obtienen sólo pueden ser números enteros y se originan de recuentos. *Por ejemplo*, número de hijos de una persona; como se imaginará no pueden tenerse 2 hijos y medio, sino 2, 3, 4, etcétera.

► **Continua**, cuando los valores que se obtienen provienen de mediciones y pueden adoptar un intervalo de valores continuo e ininterrumpido, restringido sólo por el grado de exactitud del instrumento de medición. *Por ejemplo*, el peso, la temperatura corporal son ejemplos de variables que pueden medirse con la escala numérica continua.

Como podrá observar, existen variables que sólo pueden medirse con una escala de tipo nominal, *por ejemplo* el sexo, mientras que existen otras variables que pueden medirse con cualquiera de las escalas mencionadas. Es muy importante tener en cuenta que la escala que utilizemos esté estrechamente ligada a la definición operacional dada a cada una de las variables de la investigación.

FICHA DE ORIENTACIÓN N° 2:

Operacionalización de una variable y escalas de medición



Tabla 4. Diferentes formas de medir la presencia del hábito tabáquico.

Distintos tipos de Definiciones Operacionales	Escala	Forma de "medir" el Dato
Se considerará la presencia de tabaquismo cuando el individuo refiera fumar o no	Nominal	Fuma No Fuma
La presencia de tabaquismo se considerará de acuerdo al siguiente criterio <ul style="list-style-type: none"> ► No tabaquismo: cuando el individuo refiere no fumar ► Tabaquismo leve: cuando el individuo refiera fumar menos de 5 cigarrillos/día ► Tabaquismo moderado: cuando el individuo refiera fumar menos de 20 cigarrillos/día ► Tabaquismo severo: cuando el individuo refiera fumar 20 o más cigarrillos/día 	Ordinal	Ausencia de Tabaquismo Tabaquismo Leve Tabaquismo Moderado Tabaquismo Severo
Se valorará la presencia de tabaquismo de acuerdo a la cantidad de cigarrillos que refiere fumar un individuo por día	Numérica	Cantidad de cigarrillos/día

Como podrá observar en la Tabla 4, tenemos más información si sabemos si un individuo es tabaquista leve, moderado o severo, que si sólo sabemos si el individuo fuma o no fuma. Y aún sabemos más respecto del hábito tabáquico de ese individuo si conocemos la cantidad de cigarrillos/día que refiere fumar.

Las escalas de medición tienen una cierta jerarquía, dado que una escala numérica brinda más información que una escala ordinal, y ésta más que una escala nominal. Es muy importante que al momento de definir operativamente sus variables y al procesar los datos obtenidos tenga en cuenta este concepto. Determinar el tipo de escala que se utilizará para la medición de la variable es de fundamental importancia, ya que los análisis estadísticos posteriores dependerán de la escala de medición utilizada para cada variable.

FICHA INSTRUMENTO N° 6

Pensar y resolver



Le pedimos que lea atentamente el caso que se describe a continuación:

" Dos médicos investigaban la prevalencia de enfermedad coronaria y los probables factores de riesgo asociados a esta condición en hombres entre 25 y 65 años de edad de un barrio de la ciudad de Bahía Blanca. Para ello, entrevistaron personas con enfermedad coronaria e indagaron: edad, sexo, antecedentes previos de enfermedad coronaria, tabaquismo (al que clasificaron en leve, moderado y severo de acuerdo a la cantidad de cigarrillos/día) y a la actividad física practicada por ellos (simplemente establecieron dos categorías: actividad física presente; actividad física ausente). Registraron, también, en todos los individuos la presión arterial y les realizaron un análisis de sangre con el objeto de medir la concentración sanguínea de colesterol."

Ahora le sugerimos que:

- 1) Seleccione al menos seis variables.
 - 2) Identifique la variable dependiente.
 - 3) Distinga las variables universales del resto de variables.
 - 4) Reconozca las distintas escalas de medición que requieren los distintos tipos de variables.
-

2. ¿Dónde y Cómo Obtener los Datos para la Investigación?

2.1 ¿Dónde Recolectar Datos?

Como mencionamos al inicio del Módulo, los datos son los valores que adoptan las variables. Estos datos pueden existir o pueden ser provocados por el investigador. De acuerdo a esto, se reconocen dos tipos de fuentes de datos:

- **Fuentes secundarias o documentales**, de las que obtendremos datos ya "existentes" y que son elaboradas por otras personas o instituciones con fines diferentes a los planteados para una determinada investigación. *Como ejemplo*, pueden citarse los certificados de defunción, publicaciones del censo, registros personales, clínicos y de otros tipos.
- **Fuentes primarias o de recolección directa**, constituidas por el conjunto de datos obtenidos por medio de diferentes métodos "planeados y provocados por el mismo investigador de acuerdo a los fines de su estudio". *Como ejemplo* pueden citarse los datos obtenidos por medio de observaciones, encuestas personales y encuestas auto-administradas.

2.1.1 Fuentes Secundarias de Datos

Como mencionamos, los datos secundarios se obtienen de fuentes que han sido creadas con otros fines diferentes a los de la investigación y que han sido registrados por una o

inclusive varias personas. Estas características constituyen las limitaciones más importantes de todos los tipos de fuentes documentales. Sin embargo, resultan muy interesantes y útiles, en determinadas circunstancias, por ser una forma relativamente sencilla de obtener datos.

Lo importante es que podamos reconocer cuáles son las limitaciones propias de cada fuente de datos que utilizamos, para no cometer errores en la interpretación de la información que de ellas se obtiene.

Existen fuentes secundarias de carácter permanente:

- ▶ Estadísticas Hospitalarias.
- ▶ Estadísticas Vitales elaboradas por la Dirección Nacional de Estadísticas de Salud con datos sobre nacimientos, matrimonios y defunciones provistos por el Registro Civil.

También existen fuentes de carácter eventual, como:

- ▶ Los censos.
- ▶ Los empadronamientos.
- ▶ Encuestas Nacionales, de gran utilidad como fuente complementaria de la información obtenida de los registros permanentes. Estos últimos aportan datos de mortalidad y de morbilidad.

Fuentes Secundarias Permanentes

Las estadísticas vitales mantienen un registro sistemático de los nacimientos, defunciones, defunciones fetales y los matrimonios de la población. La calidad de los Sistemas de Registros de Estadísticas Vitales tienen, en general, una relación directa con el grado de desarrollo socio económico de cada país, con mayor cobertura y calidad de información en los países con más alto nivel socio económico.

Cuantificación de los Problemas de Salud

En la Argentina, el Sistema está sustentado en el llenado de los certificados e informes de nacimientos y defunciones vinculados a la obligatoriedad de registrar en los Registros Civiles estos hechos vitales. En los certificados, el médico (o notificador) certifica haber comprobado el nacimiento o defunción de una determinada persona colocando sus datos personales, lugar de ocurrencia y, en el caso de las muertes, hace constar la causa inmediata de la defunción. Esta parte del documento (seis renglones) son completadas por el médico con fines legales y queda retenida en el Registro Civil donde se registra el hecho vital. La segunda parte, más extensa y detallada, es el informe estadístico de nacimiento o de defunción. En el mismo se incluye información más detallada de variables socioeconómicas, factores de riesgo y en el caso de las defunciones la o las causas de muerte, antecedentes e información sobre estados patológicos concomitantes. Este documento se completa precisamente con fines estadísticos y su contenido se encuentra amparado por el secreto estadístico (ley 17.622). Este informe se envía, en la mayoría de las provincias, desde los Registros Civiles locales al Ministerio de Salud Provincial, en otras se envía a Direcciones de Estadísticas Provinciales. Los niveles provinciales codifican y procesan esta información y finalmente las bases de datos son enviadas a la Dirección Nacional de Estadísticas de Salud. Se estima que en la Argentina el registro de las defunciones presenta un subregistro de la mortalidad real de alrededor de un 2% en todo el país. Por otra parte, a pesar de la importancia de estos documentos y de estar amparado por el secreto estadístico, los informes presentan en ocasiones errores o falta de datos:

- ▶ Durante el llenado de los informes de defunción, el médico puede desconocer la verdadera causa de muerte o cómo debe completar el informe; también puede existir desinterés para completarlos, considerando este registro como un hecho administrativo, o bien registrar las causas como "paro cardio-respiratorio", sin entrar en mayores detalles a veces por temores a algún problema de tipo legal.
- ▶ También durante el procesamiento la selección y codificación de las causas básicas de muerte pueden diferir según la metodología utilizada por cada codificador, sin embargo, se han hecho grandes avances en este punto durante la capacitación

para la implementación de la Clasificación Internacional de Enfermedades 10^{ma} revisión, (CIE-10).

En cuanto a los Registros de Morbilidad, los más utilizados son:

Tabla 5. Registros de Morbilidad

Tipos de Registros	¿Qué limitaciones pueden tener este tipo de registros?
<p>Notificación obligatoria: En la mayoría de los países, todo médico tiene la obligación legal de declarar a las autoridades de Salud Pública (denuncia nacional) los casos de determinadas enfermedades. Lo mismo sucede con las denuncias internacionales obligatorias, establecidas en el Código Sanitario Internacional a la OPS/OMS; Paraguay (ARPA), Bolivia (ARBOL), MERCOSUR. Las dos formas de notificación de las enfermedades tienen como objetivo evitar la propagación de enfermedades y la aparición de Epidemias</p>	<p>En general se dan las mismas situaciones que para el certificado de defunción, como el desinterés para completarlos, ya que se ve esta tarea sólo desde el punto de vista administrativo; y tanto los profesionales de la Salud como la población en general desconocen los alcances de este tipo de registro</p>
<p>Registros específicos: En este caso se registran los datos de una enfermedad en particular o de un grupo de ellas de naturaleza similar, <i>como por ejemplo</i> el Registro Nacional de Cáncer en Estados Unidos, y en la Argentina el Registro de Fiebre Hemorrágica Argentina, Tuberculosis, Lepra, entre otros. Otros registros son los registros locales del Ministerio de Salud de la Provincia de Buenos Aires; en el Instituto Nacional Regional de Oncología de Santa Fe: el Registro de Tumores de Concordia y en el Hospital R. De Santamarina de Tandil, Bs. As.</p>	<p>La Argentina carece de registros específicos confiables a nivel nacional, razón por la que es muy difícil realizar estadísticas nacionales</p>
<p>Estadísticas Hospitalarias y demás registros clínicos: En ocasiones, los registros hospitalarios se han utilizado para formular estadísticas de egresos de dichas instituciones agrupados por diagnóstico</p>	<p>Los datos hospitalarios reflejan la morbilidad de sólo una parte de la población, ya que no incluyen a los subsectores Privado ni de Obras Sociales</p>

Fuentes Secundarias Eventuales

El ejemplo más característico de una fuente eventual de datos es el censo. El censo es "el conjunto de operaciones que consiste en reunir, elaborar y publicar datos demográficos, económicos y sociales correspondientes a todos los habitantes de un país o territorio, referidos a un momento determinado" (Naciones Unidas). Estos censos se realizan a intervalos regulares, en nuestro país, cada 10 años.

2.1.2 Fuentes Primarias de Datos

Como mencionamos al inicio, las fuentes primarias están constituidas por el conjunto de datos generados por el propio investigador de acuerdo a los objetivos de su estudio. Este tipo de datos se denominan datos primarios, en contraste con los datos "existentes", denominados secundarios, ya que se originan en este tipo de fuentes.

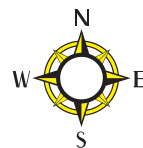
En líneas generales, el investigador puede recabar datos primarios para su investigación observando el evento que desea investigar o preguntando a las personas, para lo cual entonces utilizará la encuesta como forma de recolección de datos.

Como podrá observar, las fuentes primarias son el producto de la aplicación de diferentes formas de recolección de datos por parte del investigador. Sin embargo, es importante diferenciar fuente de forma de recolección de datos.

Lo que diferencia las fuentes primarias de las secundarias no es la forma de recolección de los datos, sino, como ya mencionamos, el hecho de que en el caso de las fuentes primarias, la recolección de los datos se lleva a cabo a los fines de la investigación. En el caso de las fuentes secundarias, si bien los datos pudieron haber sido recolectados con los mismos métodos, dicha recolección fue planeada y ejecutada por otras personas para otros fines diferentes a los del investigador.

FICHA DE ORIENTACIÓN N° 3:

Fuentes y formas de recolección de datos



Para diferenciar estos conceptos le daremos un ejemplo:

Un investigador desea estudiar el tabaquismo en su ciudad, para lo cual decide recabar datos de diferentes tipos de fuentes:

Por un lado, decide consultar los datos "existentes" en la Encuesta Nacional realizada en el último año y por otro, decide planificar y realizar una encuesta en su ciudad con el fin de determinar no sólo la prevalencia del hábito, sino también otras características de la exposición (duración, tipo, cantidad, etcétera).

La primera de las fuentes mencionadas es de tipo secundaria eventual, mientras que la segunda es una fuente primaria, porque el investigador la diseña con el fin de poder recolectar datos que le permitan alcanzar los objetivos de su investigación. Sin embargo, la forma de recolección de datos es la misma en uno y otro caso (encuesta).

2.2 ¿Cómo Recolectar los Datos?

Los datos pueden ser obtenidos mediante tres métodos básicos:

- La Observación: bajo este nombre se incluye el uso de técnicas que van desde la simple observación a otro tipo de mediciones que pueden realizarse tanto in vivo (medición de la presión arterial) como in vitro (en el laboratorio).
- Los Registros: es un método de recolección de datos sistemático, periódico, continuo y en algunas ocasiones permanente. Este método de recolección genera

Cuantificación de los Problemas de Salud

datos generalmente de tipo secundario. El ejemplo característico son las estadísticas vitales.

- ▶ Las Encuestas: es un método de recolección de datos sistemático pero eventual, de corte transversal. Como vimos, cuando se realizan en el total de la población se denominan censos.

Las encuestas producen generalmente datos de tipos primarios, aunque como mencionamos en el apartado anterior también pueden ser utilizadas para generar datos que se constituirán en fuentes secundarias eventuales (Censos y Encuestas Nacionales).

No existe un método o instrumento perfecto para la recolección de datos: todos tienen sus ventajas y sus limitaciones, por lo que no es infrecuente que el investigador combine dos o más de ellos para recoger los datos deseados.

Encuestas de Salud

Las encuestas constituyen una de las formas de recolección de datos más utilizadas por los profesionales de la Salud en sus investigaciones.

En las encuestas, el área de investigación o el problema de interés debe estar claramente definido. Para ello puede ser útil el tratar de responder las siguientes preguntas ¿a quién se va a encuestar y/o entrevistar?, ¿qué objetivo se persigue? ¿cuándo y dónde será conducida la encuesta y/o entrevista? ¿qué se espera encontrar y por qué?

En la encuesta se recolecta información preguntando a los individuos acerca de un área definida de investigación o un problema de interés. La forma de preguntar a los individuos puede ser personalmente (mediante entrevistas personales o en forma telefónica) o por correo.

El instrumento que se utiliza para la recolección de datos es un cuestionario, el que puede

ser leído por el encuestador en entrevistas personales o puede ser autoadministrado (es decir, leído y respondido por el encuestado).

Existen varios métodos para poner en práctica una encuesta, pero los principales son las encuestas personales (entrevistas) y las encuestas auto-administradas (por correo o en forma personal).

Encuestas Personales: Entrevistas

Para la recolección de datos se entrevistan a aquellas personas que serán objetos de estudio. Estas entrevistas pueden hacerse en una forma flexible, no estructurada, de modo tal que el contenido, la formulación y el orden de las preguntas sean diferentes de entrevista a entrevista o, por el contrario, en forma estructurada, con un cuestionario pre-elaborado.

Ventajas y limitaciones de las entrevistas no estructuradas

- Al no imponer límites al entrevistado, permiten mayor profundidad en la información.
- Al incluir temáticas no prefijadas, permiten mayor riqueza.
- Los datos obtenidos son más difíciles de analizar e interpretar.
- Se analizan e interpretan por métodos distintos que no permiten cuantificación.
- Requieren la inversión de un tiempo prolongado para entrevistar a cada persona.
- El número de personas entrevistadas en la mayoría de las situaciones suele ser pequeño.
- Requieren entrevistadores calificados.

Ventajas y limitaciones de las entrevistas estructuradas

- Permiten recoger toda la información requerida en un tiempo determinado.
- Son costosas en tiempo y dinero porque requieren la preparación de un protocolo analítico y un entrevistador por cada entrevistado.

Cuantificación de los Problemas de Salud

- ▶ Deben ser realizadas por entrevistadores entrenados para evitar posibles sesgos y errores.

Encuestas Auto-administradas

Este procedimiento consiste en la entrega a la persona de una serie de preguntas que ella lee y contesta por sí misma sin mediar la acción de un entrevistador.

¿Cuándo debería utilizarse?

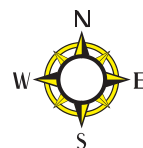
- ▶ Cuando dispone de menos tiempo y recursos económicos.
- ▶ Cuando investigue aspectos delicados o de carácter personal.
- ▶ Cuando no tenga la posibilidad de contar con personal calificado para realizar entrevistas.
- ▶ Cuando necesite obtener respuestas de muchas personas o de personas que viven en lugares diferentes, ya que pueden enviarse por correo.

Limitaciones de encuestas auto-administradas

- ▶ Tienen un número menor de respuestas por parte de las personas.
- ▶ Exige un determinado nivel de comprensión y habilidades por parte de quienes deben cumplimentarlos.

FICHA DE ORIENTACIÓN N° 4

Como diseñar una encuesta



Algunas recomendaciones respecto de las preguntas que se realicen en una encuesta:

- Deben ser claras, estar focalizadas y tener un objetivo preciso.
- Deben ser breves.

-
- No deben exigir recuerdos con demasiada exactitud.
 - No se debe enfatizar demasiado al preguntar algo ya que puede inducir sesgo de respuesta.
 - No se deben usar palabras ambiguas que tengan más de un significado.
 - Es importante reservar las preguntas sobre información personal para un momento avanzado de la encuesta.
 - Los aspectos éticos en relación con la información brindada y la calidad de las preguntas siempre deben ser considerados. Se debe asegurar al encuestado la confidencialidad de la información recolectada
-

3. ¿Cómo Procesar los Datos?

Una vez que se han recolectado los datos necesarios, es importante organizarlos o agruparlos de alguna manera para poder "manejarlos" más fácilmente. Sin la ayuda de la Bioestadística, los datos recolectados para una investigación podrían resultar algo así como una masa caótica de números. La presentación de estos datos en tablas o gráficos permite apreciarlos con mayor claridad, permitiendo explorar la información antes de analizarla. Mientras que las tablas aportan una descripción más detallada de los datos, los gráficos permiten observar los patrones generales. Veremos, entonces, cómo organizar los datos en tablas y, posteriormente, cómo presentarlos en gráficos.

3.1 Presentación de Datos en Tablas

Para organizar los datos y presentarlos en forma de tabla lo primero que tenemos que hacer es agrupar a los individuos o unidades del estudio (personas, viviendas, enfermedades, etc.) según alguna de sus características. La forma de agrupación dependerá de la escala que hayamos utilizado para medir la variable.

Cuantificación de los Problemas de Salud

- ▶ Si la escala que utilizamos es nominal u ordinal deberemos agrupar los datos de acuerdo a las diferentes categorías de la variable.
- ▶ Si la escala que utilizamos es numérica discreta, deberemos observar el rango de valores diferentes que adoptó esa variable. Si este rango es pequeño, entonces los datos se agruparán de acuerdo a cada uno de los valores de la variable. Pero si dicho rango de valores es muy amplio, entonces, deberán construirse intervalos.

Por ejemplo, si la variable es "número de consultas en un año" y el rango va desde individuos con 1 consulta a individuos con hasta 20 consultas en el año, convendrá establecer intervalos como:

1. Individuos con 0-4 consultas al año
2. Individuos con 5-9 consultas al año
3. Individuos con 10-14 consultas al año
4. Individuos con más de 15 consultas al año

Este procedimiento es más apropiado que contar para cada valor de la variable, el número de individuos que reunieron esa característica:

1. Individuos con 1 consulta al año
2. Individuos con 2 consultas al año... hasta llegar al último valor posible de la variable.

- ▶ Si la escala que utilizamos para medir la variable es numérica continua, siempre deben formarse intervalos o clases.

Una vez lograda la agrupación de los datos en diferentes categorías o intervalos, determinar la frecuencia de observaciones en cada categoría o intervalo de la variable y construir la Tabla de Distribución de Frecuencias, en el primero de los casos, y la Tabla de Frecuencias con Intervalos de Clases, en el segundo de los casos.

3.1.1 Tabla de Distribución de Frecuencias

La tabla de distribución de frecuencias permite ordenar el número de individuos que pertenecen a cada categoría de la variable. La frecuencia de observaciones en cada categoría puede expresarse como una frecuencia absoluta (total de observaciones en una determinada categoría) o como una frecuencia relativa (proporción o porcentaje en que un atributo o cualidad se manifiesta dentro de una determinada categoría).

FICHA DE ORIENTACIÓN N° 5

Como confeccionar una Tabla de Frecuencias



Tomemos, por ejemplo, una población formada por 500 personas, que estudiaremos en función de la variable "gravedad de la enfermedad".

Si decimos que, del total de la población (500), 200 personas están enfermas, estamos definiendo la frecuencia absoluta de enfermos ($f = 200$).

La frecuencia relativa puede expresarse como una proporción o como porcentaje, en ambos casos estaremos diciendo lo mismo de maneras diferentes.

- Como proporción: dividiendo el número de casos por el total de los casos examinados. En nuestro ejemplo sería $200/500 = 0,4$ de los sujetos están enfermos.
- Como porcentaje: multiplicando este número por 100, $0,4 \times 100 = 40\%$ de los sujetos están enfermos.

Si de las 200 personas enfermas 50 están graves, podremos elegir dos alternativas diferentes para calcular la frecuencia relativa (fr) de casos graves:

- Frecuencia relativa de casos graves en relación al total de individuos estudiados: dividir la cantidad de casos graves (50) por el total de la población examinada (500), en cuyo caso la frecuencia relativa es igual a 0,1 (proporción) o 10% (porcentaje).

Tabla 6. Cantidad de enfermos graves en el total de la población (n=500)

	Frecuencia absoluta	Frecuencia relativa o Proporción	Porcentaje
Enfermos graves	50	0,1	10%

- Frecuencia relativa de casos graves en relación al total de individuos enfermos: dividir el número de enfermos graves (50) por el total de personas enfermas (200), en cuyo caso la frecuencia relativa es 0,25 (proporción) o 25% (porcentaje).

Tabla 7. Cantidad de enfermos graves en el total de la enfermos (n=200)

	Frecuencia absoluta	Frecuencia relativa o Proporción	Porcentaje
Enfermos graves	50	0,25	25%

La variable "gravedad de la enfermedad" puede agruparse en las categorías "enfermos graves", "enfermos moderados" y "enfermos leves". Los datos correspondientes pueden presentarse en una tabla de distribución de frecuencia (Tabla 8).

Tabla 8. Distribución de enfermos según gravedad de enfermedad (n=200)

“Gravedad de la enfermedad”	Frecuencia absoluta	Proporción	Porcentaje
Enfermos leves	50	0,25	25%
Enfermos moderados	100	0,50	50%
Enfermos graves	50	0,25	25%
TOTAL	200	1,00	100%

Es muy importante tener en cuenta ciertos aspectos al trabajar con porcentajes .

- Cuando exprese la frecuencia de observaciones en forma de porcentaje, siempre informe el dato del numerador y el denominador con el que está trabajando.
- La ventaja de los porcentajes es que permiten comparar grupos de diferentes tamaños con una única medida. La desventaja es que puede perderse perspectiva si sólo se informan los porcentajes.
Por ejemplo, decir que el tratamiento fue eficaz en el 20% de los pacientes tratados, puede ser cierto tanto para 1 paciente entre 5 tratados, como para 1.000 pacientes entre 5.000 tratados.
- El porcentaje suele expresarse junto con el denominador o número total de la muestra encerrados por paréntesis : 20 % (5.000).
- Cada vez que lea porcentajes, verifique los numeradores y denominadores y recalculé cada porcentaje.
- Un problema característico con los porcentajes ocurre cuando los mismos se expresan no para el total de la población, sino para una muestra de la misma. *Por ejemplo,* de cada 1.000 hombres con enfermedad cardiovascular, 800 (80%) tuvieron colesterol alto; de estos 800, 250 (31%) fueron sedentarios. El 31% es 250/800, no 250/1.000.
- No es conveniente utilizar porcentajes cuando el denominador es un número muy pequeño (Denominadores menores que 20).
- Especialmente con muestras pequeñas, los porcentajes pueden ser muy confusos, porque su magnitud puede ser mucho mayor que el número absoluto que

representa: Supongamos un informe efectuado sobre la evolución de 3 pacientes. "En ese estudio, 33% de los pacientes (1 paciente) sobrevivió, 33 % (1 paciente) murió y del tercer paciente se perdieron los datos." El uso de porcentajes en casos como este lleva a confusión y no sirve para simplificar la presentación de los datos.

Otra forma de expresar la frecuencia de observaciones respecto de una variable determinada, es utilizando la "**frecuencia acumulada**". La frecuencia acumulada es el número o porcentaje de observaciones para un valor dado de la variable más todos los valores menores.

Por ejemplo, en un estudio se realizó una encuesta a 120 pacientes diabéticos y se determinó el "número de hospitalizaciones que refería haber presentado en el último año". Los datos fueron volcados en la tabla que se presenta a continuación (Tabla 9).

Tabla 9. Tabla de frecuencias del "Número de hospitalizaciones en el último año"

Número de hospitalizaciones	Nº de individuos (f)	Frecuencia relativa % (fr%)	Frecuencia acumulada % (fa%)
0	22	18,4 %	18,4 %
1	39	32,5 %	50,9 %
2	31	25,8 %	76,7 %
3	16	13,4 %	90,1 %
4	8	6,6 %	96,7 %
5	3	2,5 %	99,2 %
6	1	0,8 %	100,0 %
TOTAL	120	100,0 %	

Si quisiéramos establecer cuántos de los 120 pacientes diabéticos encuestados referían haber presentado 1 o menos de 1 internación en el último año, sabríamos por la tabla, que fue del 50.9%.

FICHA INSTRUMENTO N° 7
Pensar y resolver

Al considerar el grupo sanguíneo de 400 personas elegidas al azar, se observó que 191 tenían grupo O; 156 grupo A; 31 grupo B y 22 grupo AB.

Presente estos datos en una tabla.

1) Una trabajadora social determinó el número de hijos de 120 familias de cierta zona para un estudio que estaba realizando acerca de la población de ese lugar. Para un mejor manejo de los datos que obtuvo, decidió ordenarlos en una tabla de frecuencias, en la que además de representar la frecuencia absoluta y relativa, representó también la frecuencia acumulada. De este modo, organizó la siguiente tabla.

2) Analízela e interprétela. ¿Cuál es el significado de 32,5%; 76,6% y 3? Intente interpretar otros valores de la Tabla 10.

Tabla 10. Tabla de frecuencias: "Número de hijos por familia"

N° de hijos	N° de familias (f)	Frecuencia relativa % (fr%)	Frecuencia acumulada %
0	22	18,3 %	18,3 %
1	39	32,5 %	50,8 %
2	31	25,8 %	76,6 %
3	16	13,4 %	90,0 %
4	8	6,6 %	96,6 %
5	3	2,5 %	99,1 %
6	1	0,8 %	100,0 %
TOTAL	120	100,0 %	

3.1.2 Tabla de Frecuencias con Intervalos de Clase

Cuando la variable cuyos datos pretendemos organizar fue medida con una escala numérica continua o con una escala numérica discreta pero con un rango amplio de

valores deben construirse intervalos y presentar esta información en una tabla de frecuencias con intervalos.

Por ejemplo, en el caso de la variable "nivel de ácido úrico en sangre" utilizaríamos una escala numérica continua para su medición y los datos deberían agruparse en intervalos para facilitar su interpretación (Tabla 11).

Tabla 11. Tabla de frecuencias: "Niveles de ácido úrico en la población X"

Intervalos	Nº de individuos (f)	Frecuencia relativa % (fr%)	Frecuencia acumulada % (fa%)
3.0-3.9	17	6,3 %	6,3 %
4.0-4.9	73	27,2 %	33,5 %
5.0-5.9	101	37,7 %	71,2 %
6.0-6.9	54	20,1 %	91,3 %
7.0-7.9	18	6,7 %	98,0 %
8.0-8.9	4	1,5 %	99,5 %
9.0-9.9	1	0,5 %	100,0 %
TOTAL	268	100,0 %	

Los extremos de estos intervalos son llamados límites, distinguiéndose el límite inferior de cada intervalo (3,0; 4,0... en nuestro ejemplo) y el límite superior (3,9; 4,9...). La diferencia entre el límite superior y el límite inferior se denomina amplitud del intervalo.

Con la agrupación de los datos en intervalos se gana en comodidad, pero se pierde información, ya que una vez agrupados, todos los datos pertenecientes a un intervalo serán identificados para el análisis posterior con un punto medio o marca de clase de ese intervalo. Para calcular dichas marcas de clase se realiza la siguiente fórmula:

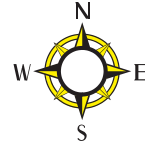
$$\text{Marca de clase del Intervalo} = \frac{\text{Límite superior} + \text{Límite inferior}}{2}$$

Retomando nuestro ejemplo de los niveles de ácido úrico, la marca de clase para el primer intervalo sería:

$$\text{Marca de clase} = \frac{3,99 + 3,0}{2} = 3.5$$

FICHA DE ORIENTACIÓN N° 6

Para construir intervalos



Estas recomendaciones sirven a los fines de orientar la construcción de intervalos, sin embargo, se debe tener en cuenta que no son reglas rígidas, sino que deben adaptarse tanto a los datos que trabajamos como a los propósitos que persigamos en la investigación

El número de intervalos recomendable es aquel que permita presentar los datos de manera resumida sin enmascarar la información esencial. La utilización de demasiados intervalos difiere poco de la tabulación original de los datos, y un número demasiado reducido puede enmascarar información esencial.

- En lo posible construir intervalos de igual amplitud, esto en general es útil si los datos tienen distribución normal, como por ejemplo en el caso de dosaje de albúmina en mujeres mayores de 50 años. Si esto no es así conviene presentar los datos divididos por centilos, cuartiles, etc., es decir, presentando los datos de acuerdo a su dispersión y no a los valores que asume.
- Los intervalos contruidos deben ser mutuamente excluyentes, esto es: si se desea dividir las edades de los individuos participantes en un estudio en intervalos de 10 años, obtendremos lo siguiente:

Grupo a- 0 a 9 años

Grupo b- 10 a 19 años

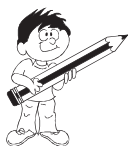
Grupo c- 20 a 29 años

De este modo se evita la mala clasificación de los sujetos, así una persona con 19 años 11 meses y 25 días ingresará en el grupo B

- Para determinar los intervalos es recomendable:
 1. Identificar el dato correspondiente al valor máximo y el correspondiente al valor mínimo
 2. Determinar la diferencia entre estos dos valores (rango)
 3. Dividir esta diferencia por el número de intervalos que se desea obtener para identificar así la amplitud de cada intervalo

FICHA INSTRUMENTO N° 8

Pensar y resolver



Marta, una estudiante de bioquímica, obtuvo los siguientes valores de seroalbúmina en g/l de sangre de 50 mujeres sanas:

42	41	42	44	44	36	38	41	42	44
42	39	49	40	45	32	34	43	37	39
41	39	48	42	43	33	43	35	32	34
39	35	43	44	47	40	39	42	41	46
37	49	41	39	43	42	47	48	51	52

Su compañero de estudios le sugirió que de esa manera no podría interpretar muy bien todos esos datos, sugiriéndole que los agrupe en 7 intervalos de clase o categorías de igual amplitud. También le dijo que considerara al dato con el menor valor como el límite inferior del primer intervalo. Además le sugirió agregar una columna para la "marca de clase" de cada intervalo.

Con estas sugerencias, Marta construyó una tabla de frecuencias con intervalos de clase.

Intente hacerlo usted mismo y, luego, coteje su tabla con las de sus compañeros.
A partir de la tabla elaborada por Ud., haga las siguientes interpretaciones.

- 1) ¿Cuál es el significado de una frecuencia absoluta igual a 9 en la tercera categoría?
 - 2) ¿Cuál es el significado de una frecuencia acumulada de 72% en la cuarta categoría?
-

3.2 Presentación de Datos en Gráficos

En ocasiones, preferirá representar gráficamente sus datos, con el objeto de obtener una rápida impresión visual del conjunto. Para ello podrá utilizar diferentes tipos de gráficos, pero lo que nunca debe olvidar son las siguientes premisas:

- El gráfico debe ser sencillo y explicarse por sí mismo.
- No intente graficar "todos" los datos que tiene en un solo gráfico; por el contrario, es preferible que no contenga demasiada información y su lectura sea fácil.
- Utilice un diseño atractivo, pero sin deformar los hechos que está describiendo.
- Seleccione el gráfico más apropiado de acuerdo al tipo de variable y la escala de medición utilizada para medirla.

El tipo de gráfico está condicionado por el tipo de escala utilizada para medir la variable que desea graficar.



Principales Tipos de Gráficos

- Gráfico de Sectores Circulares o "tortas"
- Gráfico de Barras
- Histograma
- Polígono de Frecuencias

3.2.1 Gráfico de Sectores Circulares

Consiste en un círculo dividido en sectores con áreas proporcionales a la frecuencia con que se presenta cada categoría.

¿Qué tipo de datos suelen presentarse de esta manera?

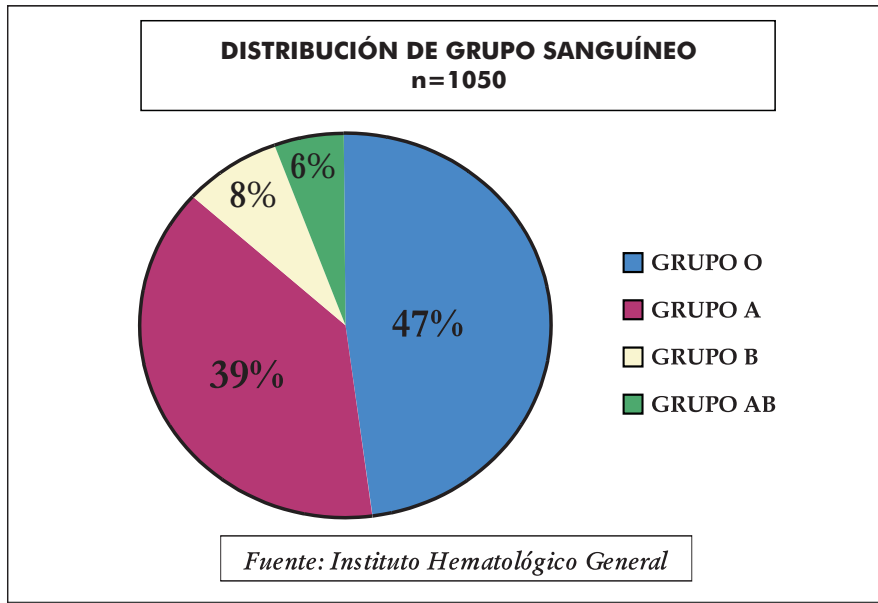
Habitualmente se utiliza este tipo de gráfico para presentar datos que se miden con una escala nominal, pero también puede utilizarse para variables ordinales o numéricas discretas con un rango pequeño de valores.

¿Qué errores puede cometer cuando utilice este tipo de gráfico para presentar sus datos?

- ▶ Representar una variable formada por múltiples categorías con una baja frecuencia de observaciones por categoría; en estos casos resulta más conveniente agruparlas y así facilitar su comprensión.
- ▶ Al igual que en el caso de los porcentajes, omitir el número total de individuos estudiados (n). Este valor es muy importante porque la interpretación del gráfico es muy diferente de acuerdo al tamaño de la población cuyos datos se están representando. En este sentido, *por ejemplo*, no es lo mismo un área correspondiente a un 20% de un gráfico circular que representa los datos de una población formada por 10 personas que la de una formada por 1.000.

Construiremos, ahora, un gráfico circular con los datos del ejemplo 1 sobre grupo sanguíneo.

Gráfico 1. Sectores Circulares



3.2.2 Gráfico de Barras

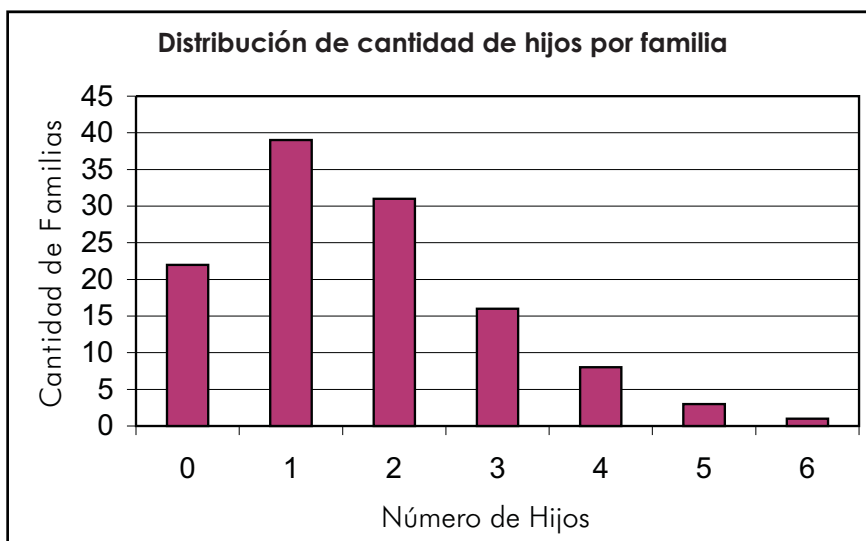
Este tipo de gráficos se construye dibujando barras o rectángulos en los cuales la longitud indica la magnitud o la frecuencia de cada categoría de la variable.

¿Qué tipo de datos suelen presentarse de esta manera?

Se utiliza para representar la distribución de frecuencias de variables en escala nominal, así como para variables en escala ordinal y numérica discreta cuando el rango de valores de la variable es reducido, como en el ejemplo anterior. Como podrá apreciar, en el eje horizontal se indican los valores de la variable y en el vertical sus correspondientes frecuencias.

Tomando los datos de la Ficha Instrumento N° 9, acerca del número de hijos por familia, dibujamos el siguiente gráfico de barras.

Gráfico 2. Gráfico de Barras



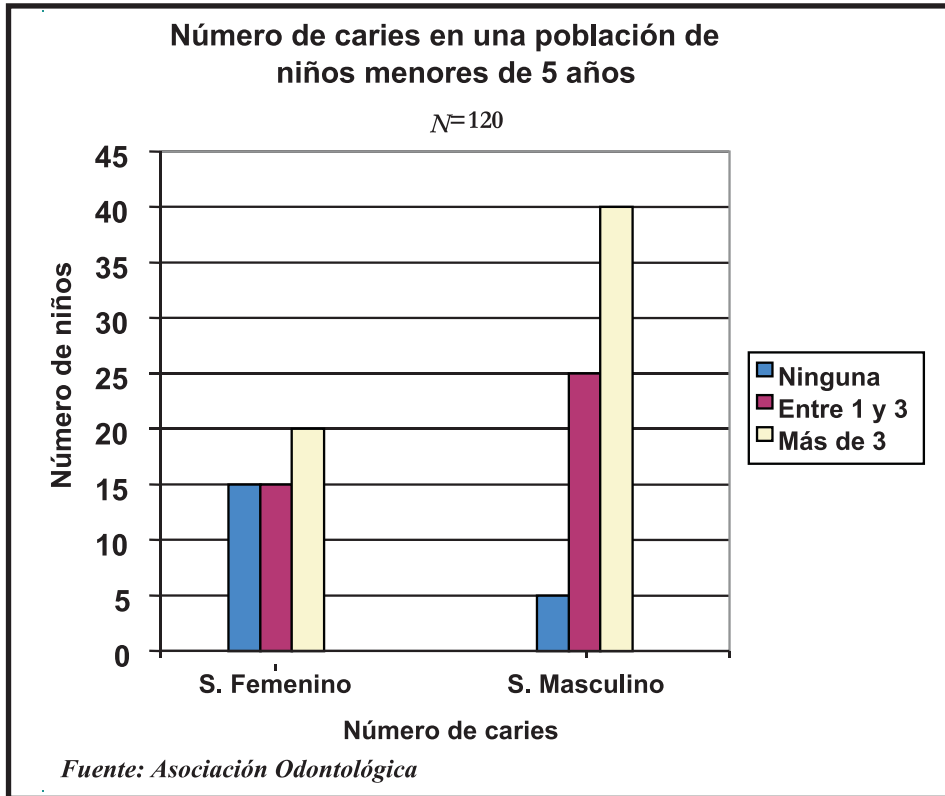
Imagínese ahora, que recolectó datos sobre sexo y número de caries de una población. En este caso, por tratarse de una variable nominal (sexo), en la base de la barra deberá indicar cada una de las categorías de la variable sexo.

El número de caries podrá agruparse en...

1. Ninguna caries
2. Entre 1 y 3 caries
3. Más de 3 caries

... en un gráfico de barras como el que sigue:

Gráfico 3. Gráfico de Barras



Observe que este gráfico es algo más complejo que el anterior, pero en él pudimos representar no sólo la frecuencia de caries en los niños de cada sexo, sino también el número de niños que se encontraron en cada una de las categorías de la variable número de caries.

Como puede observar, el gráfico de barras es muy útil para comparar no sólo la distribución de frecuencias de las diferentes categorías de una variable, sino también la distribución de acuerdo a más de una variable.

3.2.3 Histograma

Los histogramas se utilizan para representar una distribución de frecuencias de una variable medida con una escala numérica continua. Es una representación gráfica análoga al gráfico de barras. Se puede decir que el histograma es un conjunto de rectángulos contiguos levantados sobre los intervalos de la variable y cuya área es proporcional a la frecuencia. A diferencia del gráfico de barras, la información que un histograma comunica es del área de las barras que lo constituyen, y no ya sólo de su altura. Por ello, si la amplitud de los intervalos no varía, las áreas son proporcionales a las alturas y las frecuencias están representadas tanto por las alturas como por las áreas. Si los intervalos fueran de diferente amplitud, la altura de las barras deberán modificarse para conservar el área correcta.

FICHA DE ORIENTACIÓN N° 7

Cómo construir un histograma

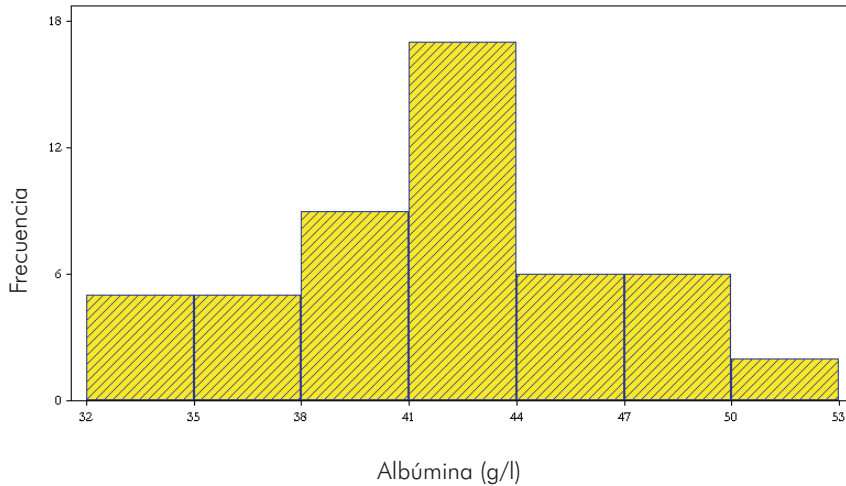


En este punto, le proponemos que repase los datos que Marta, la bioquímica, había recolectado acerca del nivel de seroalbúmina en muestras de sangre de 50 mujeres que presentamos en la Ficha Instrumento N° 9. Marta había mejorado su conocimiento sobre tablas y gráficos y decidió construir un histograma, dado que la concentración de seroalbúmina en la sangre era una variable numérica continua.

- Toda vez que construya un histograma, tenga en cuenta que estos se utilizan para representar la frecuencia de variables medidas con la escala numérica continua y que en este caso la frecuencia de cada intervalo está representado por el área de la barra, no por su altura.

Gráfico 4. Histograma

Concentración de seroalbúmina en sangre en una población de 50 mujeres



3.2.4 Polígono de Frecuencias

Se utilizan para representar una distribución de frecuencias en los mismos casos en que se utiliza el histograma, pero en la mayoría de las veces a fin de comparar dos o más distribuciones en un mismo gráfico. Se construyen uniendo los puntos medios de las bases superiores de los rectángulos del histograma con segmentos rectilíneos. Se debe continuar la línea hasta las marcas de clase de intervalos hipotéticos anterior al primero y posterior al último, de modo que el área encerrada por el polígono sea igual al área encerrada por el histograma. En el gráfico anterior podrá observar el polígono de frecuencias correspondiente a la distribución de seroalbúmina.

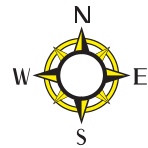
El polígono de frecuencias permite determinar la forma de una distribución de valores numéricos. Estas formas pueden ser de tipos muy variados. Se describen dos aspectos de estas curvas: su simetría y su modalidad.

Se dice que una distribución es **simétrica** si al plegarla sobre sí misma sus dos mitades se superponen aproximadamente. Las distribuciones asimétricas se describen también como **sesgadas**.

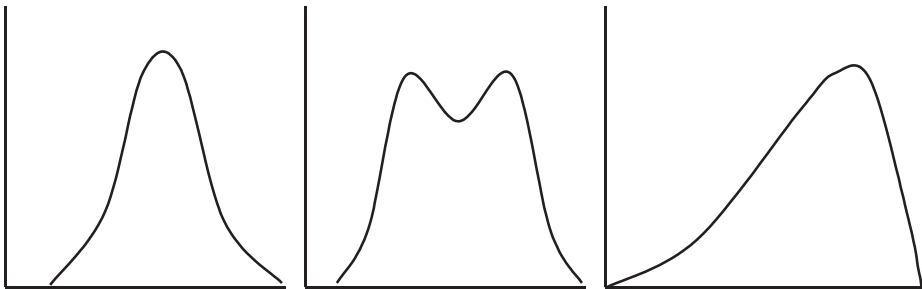
Una distribución **unimodal** es aquella que sólo presenta un pico o punto alto, en tanto que una distribución **polimodal** tiene dos o más picos; si tiene dos recibe el nombre de **bimodal**.

FICHA DE ORIENTACIÓN N° 8:

Diferentes formas de distribución de datos



Observe las siguientes curvas y describa tanto su simetría como su modalidad.



- La primera curva es simétrica, unimodal, con distribución gaussiana o normal.
- La segunda curva es también simétrica pero tiene dos "picos" o modas, o sea, es bimodal.
- La tercera es una curva asimétrica desviada hacia la izquierda (el desvío de la curva se señala hacia el lugar donde está la "cola").

FICHA INSTRUMENTO N° 9:**Pensar y resolver**

Seleccione el gráfico que considere más apropiado para representar la variable "número de hijos por familia" cuya distribución de frecuencias se presenta a continuación. (Tabla 12). Justifique su selección.

Tabla 12. Tabla de frecuencias de la variable " Número de hijos por familia"

N° de hijos por familia	N° de familias (f)	Frecuencia relativa % (fr%)	Frecuencia acumulada % (fa%)
0	22	18,4 %	18,4 %
1	39	32,5 %	50,9 %
2	31	25,8 %	76,6 %
3	16	13,4 %	90,0 %
4	8	6,6 %	96,6 %
5	3	2,5 %	99,1 %
6	1	0,8 %	100,0 %
TOTAL	120	100,0 %	

4. ¿Cómo Interpretar los Datos?

Una vez recolectados los datos del evento en estudio, debemos proceder a su análisis. Dos pasos esenciales para poder llevar a cabo el análisis de los datos son el procesamiento y la interpretación de los mismos. Para ello es fundamental contar con un orden preestablecido. Proponemos una sistemática de trabajo que ayudará a tener mejores resultados, cuyos pasos se encuentran enumerados en la figura 2 de la página 57.

Al describir un evento de Salud, realizamos diferentes mediciones de la frecuencia de aparición de una característica o de los valores de las diferentes variables. Utilizamos para

esto medidas de resumen, que son medidas que precisamente reflejan características de la muestra de datos con las que estoy trabajando en su conjunto. Existen diferentes tipos de medidas de resumen que deben ser seleccionadas de acuerdo a la escala de medición utilizada para medir la variable, tal como se muestra el la figura 2 de la página siguiente.

Nos referiremos a la sistematización del trabajo y resaltaremos algunos puntos clave de las medidas de resumen.

4.1 La Sistematización de los Datos

Le sugerimos que toda vez que deseen procesar los datos obtenidos se sistematice de alguna manera.

FICHA DE ORIENTACIÓN N° 9

Cómo sistematizar los datos

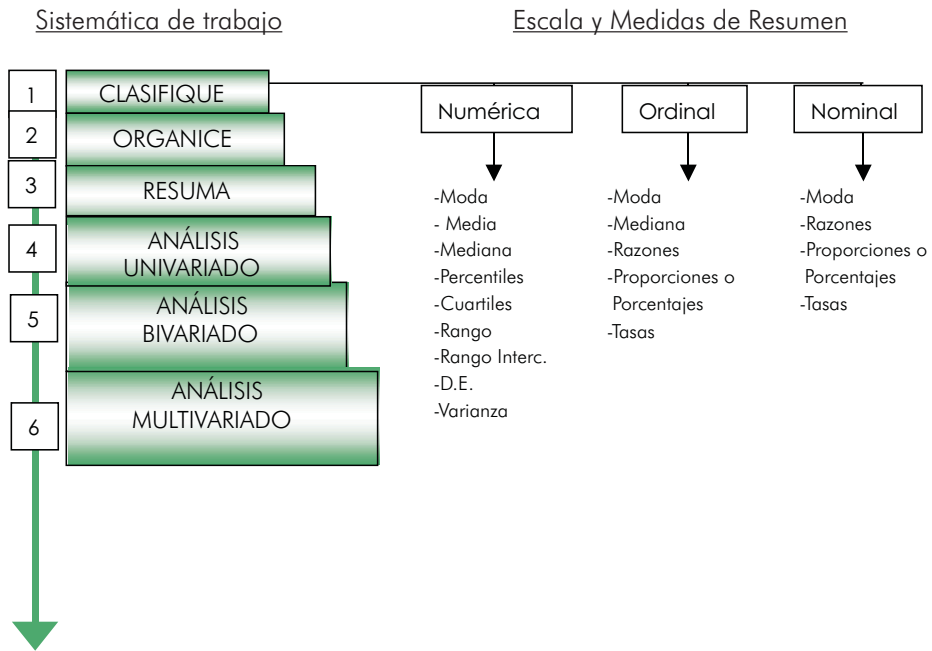


Una forma que nos parece conveniente para sistematizar el trabajo, es la siguiente (figura 2):

1. Clasifique cada una de las variables de su estudio.
2. Organice y presente sus datos en el gráfico que considere más apropiado.
3. Resuma sus datos. Para hacerlo adecuadamente tenga en cuenta el tipo de variable y analice en primer lugar los datos de una sola variable por vez (Análisis Univariado). Una vez que analizó las variables de a una por vez, podrán pasar a analizar la relación entre dos variables (análisis bivariado).

Si superó este desafío, podrá analizar, entonces, la relación entre tres o más variables: análisis multivariado.

Figura 2. Sistemática de trabajo y medidas de resumen a utilizar de acuerdo a la escala empleada



4.2 Las Medidas de Resumen

Las medidas que utilizaremos según las variables que estemos trabajando son las mencionadas en la figura 2 y pueden clasificarse en:

- **Medidas de frecuencia:** razón, proporción o porcentajes y tasas.
- **Medidas de tendencia central:** media, mediana y moda.
- **Medidas de dispersión:** rango, rango intercuartílico, desvío estándar.
- **Medidas de orden:** centiles, cuartiles.

Cuantificación de los Problemas de Salud

Sin embargo, es necesario que también sepa reconocer las diferencias conceptuales entre cada una de las medidas mencionadas.

En caso de que sea numérica, deberán utilizar alguna o varias de las medidas de centralización (media, mediana y moda), de dispersión (rango, rango intercuartílico, desvío estándar y varianza) y/o de orden (percentiles y cuartiles).

En caso de que se trate de una variable nominal, pueden utilizar en ocasiones como medida de tendencia central la moda, pero es conveniente que en estos casos resuma sus datos calculando medidas de frecuencia como razones, proporciones o porcentajes y/o tasas. En caso de que la variable sea ordinal, pueden utilizar en ocasiones la moda y la mediana entre las medidas de centralización para poder describirla, pero más conveniente es el uso de razones, proporciones o porcentajes y/o tasas.

Le proponemos entonces, repasar los conceptos de medidas de frecuencia: razón, proporción y tasa (dentro de estos dos últimos grupos describiremos prevalencia, incidencia acumulada, tasa de incidencia acumulada y densidad de incidencia), medidas de tendencia central, medidas de dispersión y medidas de orden.

Cuadro 1. Medidas de Resumen

Medidas de Frecuencia	<ul style="list-style-type: none">› Razón› Proporción› Tasa (Incidencia, Prevalencia)
Medidas de Tendencia Central	<ul style="list-style-type: none">› Moda› Media› Mediana
Medidas de Orden	<ul style="list-style-type: none">› Percentiles› Cuartiles
Medidas de Dispersión	<ul style="list-style-type: none">› Rango› Rango Intercuartílico› Desvío Estándar

4.2.1 Medidas de Frecuencia

Razón

Esta medida de frecuencia se emplea para variables de tipo nominales.

Razón es una división que no implica ninguna relación específica entre el numerador y el denominador. Esto quiere decir que el numerador y el denominador llevan unidades diferentes y en numerador no se encuentra incluido en el denominador.

Por ejemplo: En un grupo de 300 pacientes con infarto agudo de miocardio (IAM) que estoy estudiando hay 200 varones y 100 mujeres puedo decir que la razón hombre : mujer es en este caso de 2:1. Por cada 2 varones hay una mujer.

En otras palabras, una razón es el número de observaciones del grupo de pacientes con IAM de sexo masculino dividido por el número de pacientes con IAM de sexo femenino; establece la relación de una parte con otra parte.

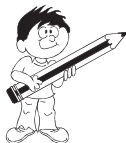
Proporciones o Porcentajes

La proporción es una medida de frecuencia en la que se expresa qué parte del total de observaciones presenta determinada característica. El numerador está incluido en el denominador.

Siguiendo con el ejemplo anterior decimos que del total de pacientes con IAM que en nuestro estudio son 300, 200 son varones vistos por lo tanto la proporción de varones es de $200/300 = 0,66$. Este número puede ser expresado como porcentaje diciendo que el 66 % de los enfermos son varones.

FICHA INSTRUMENTO N° 10

Pensar y resolver



En la siguiente tabla se presentan los datos sobre los casos de SIDA según sexo, desde el inicio de la epidemia hasta diciembre de 1999.

- 1) ¿Cómo se calcula la razón hombre / mujer? ¿Cómo se calcula la proporción de varones expresada como porcentaje?
- 2) ¿Qué significa que la razón hombre: mujer sea de 14,5 en 1.988? ¿Qué significa que el porcentaje de varones sea del 93,6%?
- 3) ¿Cuáles son las posibles causas de las diferencias observadas en la razón hombre: mujer entre el inicio y los años más recientes de la Epidemia?

Tabla 13. Distribución de Frecuencia de casos de Sida según texto 1982-1986

Años	Masculino	% Masculino	Femenino	Sin identificar	Total	Razón hombre: mujer
	N°		N°	N°	N°	
1982 – 86	80	100,0	0	0	80	-
1987	91	97,8	1	1	93	91
1988	189	93,6	13	0	202	14,5
1989	263	88,9	31	2	296	8,5
1990	427	86,1	64	5	496	6,7
1991	646	87,4	90	3	739	7,2
1992	916	80,9	211	5	1.132	4,3
1993	1.146	77,7	315	13	1.474	3,6
1994	1.708	78,1	455	24	2.187	3,8
1995	1.711		451	33	2.195	
1996	1.987		584	18	2.589	
1997*	1.655		561	11	2.227	
1998*	1.252		484	8	1.744	
1999**	588		210	4	802	
Total	12.661		3471	127	16.259	

Tasas

La tasa es una medida que expresa el número de eventos ocurridos y en una población determinada. Las tasas se describen como medidas de frecuencia en las que una medida de tiempo es parte intrínseca del denominador. En el numerador aparecen los eventos (casos nuevos de enfermedad, defunciones, etc) y en el denominador la población de estudio de donde provienen los casos, durante el período en que fueron estudiados. Lo que distingue a una tasa de las proporciones y de las razones es que:

1. En la tasa se relaciona un evento con la población en riesgo de presentar ese evento.
2. Para expresar una tasa debe incluirse en el denominador el tiempo durante el cual las personas estuvieron en riesgo de presentar el evento (tiempo en riesgo).
3. Se multiplican por una constante (10 o múltiplos de 10) que facilita la comparación de tasas de poblaciones diferentes, aún cuando las mismas sean de diferente tamaño. Esta constante nunca debe ser mayor que la población en riesgo. *Ejemplo:* Si la población total es de 3.000 personas, multiplicaremos por 1.000, no deberíamos multiplicar por 10.000 como constante.

En el campo de la salud la tasa es una medida muy utilizada porque, como mencionamos, permite medir de alguna manera la frecuencia de un evento (habitualmente una enfermedad) en una determinada población "expuesta" o "en riesgo" de experimentar dicho evento en un determinado período de tiempo y comparar entre poblaciones.

Tasas Brutas y Tasas Específicas

Las tasas de todas las personas enfermas o muertas se denominan tasas brutas (totales o generales). En este tipo de tasas se consideran todas las causas y características de los sujetos. Se construyen con el número total de eventos (enfermos o muertos) por unidad de tiempo en la población total.

Por ejemplo, la tasa de mortalidad de una población es una tasa bruta.

Sin embargo, normalmente nos interesan más las enfermedades que causan casos o muertes, las características de las personas involucradas, etc. El cálculo de las tasas de sucesos de enfermedades en determinados grupos da por resultado tasas específicas.

Por ejemplo, podemos calcular tasas específicas por enfermedad (tasa de incidencia de cáncer de pulmón); por edad (tasa de mortalidad infantil); por enfermedad y edad (tasa de incidencia de accidentes de hogar en niños menores de 5 años).

Comparación de tasas

Cuando se utilizan tasas para comparar los riesgos de muerte o de enfermedad en dos poblaciones diferentes, por ejemplo, dos hospitales, o dos ciudades, o dos fábricas, es muy importante considerar si estas poblaciones difieren en algún factor que se sabe influye en el riesgo de morir o contraer la enfermedad.

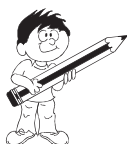
Por ejemplo, si se desea evaluar las tasas de mortalidad por neumonía de dos hospitales diferentes, es muy importante considerar la distribución por edades de las personas con neumonía en uno y otro hospital. Si en uno de los hospitales los pacientes con neumonía son más añosos, las tasas serán seguramente mayores debido a esta diferencia en la edad de la población afectada. Así como la edad también podrían considerarse otros factores que se saben asociados con peor pronóstico pacientes con neumonía (inmunosupresión, bajo nivel socioeconómico, etcétera).

Para evitar esta confusión que genera la presencia de estos factores denominados factores de confusión, al momento de comparar tasas se utiliza un procedimiento denominado estandarización o ajuste de tasas. Por medio de este método se calculan tasas brutas hipotéticas para cada grupo de comparación ("tasas estandarizadas"). Estas tasas, si bien

son hipotéticas, son comparables, y permiten determinar en que medida se diferencian las dos poblaciones respecto al riesgo de enfermar o morir por el evento en consideración.

FICHA INSTRUMENTO N° 11

Pensar y resolver



A continuación le presentamos un estudio y las conclusiones elaboradas por su autor.

En un estudio sobre Alto Peso al nacer (AP), se comunicó que de un total de 51.000 partos ocurridos en una clínica privada de la ciudad de Mendoza, se detectaron 125 casos para los cuales el peso del recién nacido superó los 5.000 gramos. Los 125 casos estudiados se distribuyen según paridad de la madre en la siguiente forma:

Paridad de la madre	N° de Casos	Porcentajes
1	9	7 %
2	24	19 %
3	17	14 %
4	16	13 %
5	9	7 %
6 y 7	19	15 %
8 y 9	9	7 %
10 y +	22	18 %
Total	125	100 %

Al analizar los resultados el autor escribe:

"En resumen, 9 casos, es decir el 7%, son primíparas; el resto, 116, es decir el 93%, eran múltiparas y de estas últimas, 24 eran secundíparas, lo que representa el 19% del total. Este es el grupo más numeroso, lo que nos autorizaría a aceptar que la secundiparidad es el momento más favorable para la producción de fetos gigantes."

- 1) ¿Están de acuerdo o no con estas conclusiones? Justifiquen su respuesta.
- 2) ¿Necesitaría alguna otra información para realizar este análisis? ¿Cuál?

Como mencionamos anteriormente, existen dos medidas de gran relevancia desde el punto de vista epidemiológico, que permiten medir la frecuencia de las enfermedades y que seguramente habrá oído mencionar en más de una oportunidad: **prevalencia e incidencia.**

Prevalencia

Es el número de casos o eventos en una población de una determinada enfermedad en un momento dado en el tiempo. La prevalencia intenta mostrar la frecuencia con la que un evento de Salud o enfermedad existe en el momento, independientemente de cuándo se haya originado.

¿Cómo se calcula la prevalencia?

Se debe determinar el número de sujetos en quienes se presenta el evento estudiado y relacionar dicho número con el total de individuos que forman el grupo observado. Suele expresarse en forma de proporción o porcentaje.

Existen dos formas de calcular la prevalencia:

- Prevalencia puntual:

Se utiliza cuando se desea determinar la prevalencia de una enfermedad en la población en un momento en el tiempo, independientemente de cuándo se haya originado. Se denomina puntual porque la medición se hace en un corto período de tiempo.

$$\text{Prevalencia Puntual} = \frac{\text{Casos de una enfermedad en un momento dado}}{\text{Población total en estudio en ese momento}} \times 100$$

- Prevalencia de período:

Determina la cantidad de casos existentes durante un período.

Es una forma práctica de resolver el problema de la aparición de casos nuevos cuando se está ejecutando el estudio; ya sea por una elevada frecuencia de aparición de la enfermedad o porque el estudio no se puede realizar en un corto período de tiempo. Esta medida informa el número de casos existentes así como los casos nuevos producidos durante ese período.

$$\text{Prevalencia de Período} = \frac{\text{Casos antiguos + casos nuevos en un período dado}}{\text{Población total en estudio en ese período}} \times 100$$

¿Por qué es útil determinar la prevalencia de un evento o enfermedad?

La prevalencia expresa la cantidad de casos de una enfermedad en un momento dado, de ahí su importancia para la elaboración de los programas de Salud preventivos, asistenciales y de control.

Por otra parte, la información suministrada por la prevalencia es, en muchas ocasiones, un punto de partida para la investigación causal.

Incidencia

La incidencia de la enfermedad representa la frecuencia de casos nuevos en una población durante un período determinado de tiempo. La duración de ese período de observación influye en los resultados de la incidencia, pues si éste aumenta ocurrirán más

casos e hipotéticamente pueden llegar a enfermarse todos los individuos y el valor entonces de la incidencia será igual a 1,00 o 100%. Por eso, al usar esta medida, debe siempre expresarse el período de observación.

¿Cómo se calcula la incidencia?

Existen tres formas de calcular la incidencia:

- Incidencia acumulada:

Es una medida que representa la proporción de individuos sanos que contraen la enfermedad en un período de tiempo determinado.

$$\text{Incidencia Acumulada} = \frac{\text{Casos nuevos}}{\text{Número de individuos al comienzo del período}} \times 100$$

- Tasa de Incidencia Acumulada

Mientras que la incidencia acumulada es una proporción, la tasa de incidencia acumulada es una tasa porque el denominador es el tiempo en riesgo de la población estudiada. Sin embargo, no siempre puede calcularse este tiempo en riesgo dado que:

- Frecuentemente, no es posible excluir el tiempo en que los individuos que se enferman ya no están en riesgo.
- En otras ocasiones, es imposible identificar el período de exposición de cada uno de los miembros de la población estudiada.

➤ Por este motivo, el tiempo en riesgo se determina mediante una aproximación, resultado de multiplicar el tamaño promedio de la población observada por la duración del

período, generalmente un año.

$$\text{Tasa de Incidencia Acumulada} = \frac{\text{Casos nuevos}}{\text{Número de individuos a mitad del período}} \times K^*$$

*K es igual a una constante que puede ser cualquier múltiplo de 10, 100, 1.000, 10.000, etc. Esta multiplicación se introduce para facilitar las comparaciones entre poblaciones de diferente tamaño absoluto. Nunca debe ser mayor que la población total en riesgo.

- Densidad de Incidencia:

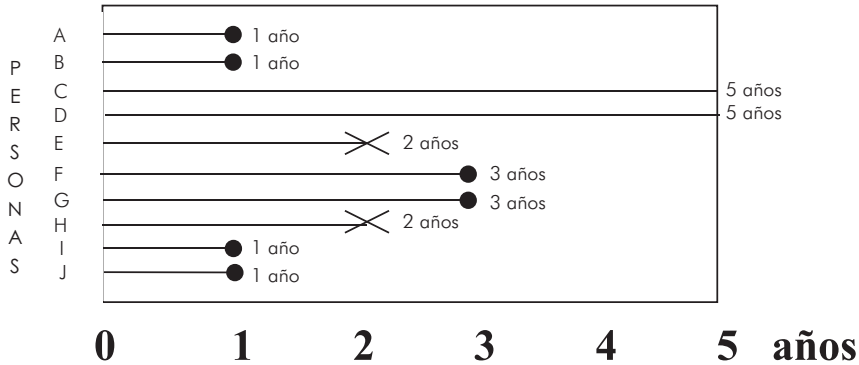
Es el número de nuevos casos registrados, dividido por la suma de los períodos de tiempo en riesgo correspondientes a todos los individuos en estudio. Esta sería la "verdadera tasa de incidencia", porque en este caso se conocen los tiempos en riesgo de cada uno de los individuos y no se recurre a una aproximación como lo es la tasa de incidencia acumulada.

Esta medida se utiliza cuando la población observada es inestable en el tiempo, es decir, cada sujeto ha estado "en riesgo" o expuesto al evento de interés por períodos de tiempo distintos, ya sea por abandono del estudio, por contraer la enfermedad, etc. Es una medida útil para el seguimiento de poblaciones dinámicas en las que ingresan y salen individuos. Su unidad de tiempo es personas-unidad de tiempo utilizada para el estudio (personas-años, personas-días, etc.). Una persona-año representa un individuo en riesgo de desarrollar la enfermedad durante un año.

Para poner un ejemplo, supongamos que seguimos durante 5 años a personas que trabajaron expuestas a radioactividad durante distintos períodos de tiempo. En el gráfico siguiente representamos los distintos períodos de seguimiento. Los puntos representan las

personas que desarrollaron la enfermedad, las cruces las personas que se perdieron para el seguimiento.

Gráfico 5. Duración de exposición a la radioactividad de los trabajadores en seguimiento



La principal utilidad de la incidencia es que permite medir el riesgo que un grupo de individuos desarrolle una enfermedad. Más adelante le explicaremos el enfermedad mientras duró el estudio. Las personas E y H fueron seguidas durante 2 años y no pudieron ser luego ubicadas (se perdieron para el seguimiento). Las personas F y G fueron seguidas durante 3 años y luego desarrollaron la enfermedad.

1) ¿Cuántas personas año expuestas nos plantea este ejemplo? En total 24 personas/ año fueron seguidas en el ejemplo calculamos desde el caso A al J:

$$1 + 1 + 5 + 5 + 2 + 3 + 3 + 2 + 1 + 1 = 24 \text{ personas año.}$$

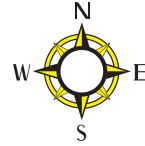
2) ¿Cuál es la densidad de incidencia en este caso? Hubo 6 casos nuevos en 24 personas año expuestas, por lo tanto, la densidad de incidencia fue de:

$$6 \text{ casos nuevos} / 24 \text{ personas año seguidas} = 0,25 \times 10 = 25 \text{ casos por cada 10 personas año expuestas.}$$

3) ¿Cuál sería la incidencia acumulada? Hubo 6 casos nuevos en 10 personas inicialmente expuestas. Por lo tanto, la incidencia acumulada sería:

$$6 \text{ casos nuevos} / 10 \text{ personas inicialmente en riesgo} = 0,6 \times 10 = 6 \text{ casos nuevos cada}$$

10 personas inicialmente expuestas.



FICHA DE ORIENTACIÓN N° 10

Para comprender mejor la densidad de Incidencia

Comparemos la incidencia de accidentes laborales en dos grupos de enfermeros del Servicio de Pediatría de dos hospitales diferentes. El Hospital A informa una tasa de incidencia de accidentes en los enfermeros en el último mes del 20% mientras que el Hospital B informa una tasa del 10%.

Hospital	Enfermeros	Cantidad Accidentados	IA	Cantidad Enfermeros/ Hora expuestos	Turnos laborales	Densidad de Incidencia (accidentes/ personas-horas/mes)
A	100	20	20/100 20%	16.000	8 hs/día	$20/16.000=$ 1,25 por 1.000
B	100	10	10/100 10%	8.000	4 hs/día	$10/8.000$ 1,25 por 1.000

Tabla 14. Comparación de accidentes laborales en los enfermeros de Hospitales A y B. La densidad de incidencia es igual debido a que en el Hospital A los enfermeros llevan a cabo turnos laborales que duplican los del Hospital B y, por lo tanto, duplican también el tiempo de exposición.

A diferencia de las anteriores, esta medida tiene una unidad: personas-unidad de tiempo (personas-año, persona-mes, personas-día). La suma persona- tiempo indica la suma de todos los períodos de tiempo por persona estudiados. El período de tiempo de observación se calcula para cada persona de forma individual, sumándose los de todas las personas del estudio mientras están en riesgo (o sea, mientras está dentro del estudio y no ha contraído la enfermedad).

¿Por qué es útil determinar la incidencia de un evento o enfermedad?

La principal utilidad de la incidencia es que permite medir el riesgo que un grupo de individuos desarrolle una enfermedad. Más adelante le explicaremos el concepto de riesgo, pero por ahora es importante que recuerde que en Epidemiología, incidencia es sinónimo de "riesgo".

FICHA INSTRUMENTO N° 12

Pensar y resolver



Analice las principales diferencias entre los conceptos de incidencia y prevalencia y complete el siguiente cuadro:

CARACTERÍSTICA	INCIDENCIA	PREVALENCIA
Numerador		
Denominador		
Tiempo		
Formas de calcularla		
Utilidad		

Para resumir el conjunto de datos de una variable medida con una escala numérica deberemos utilizar las medidas de tendencia central (media, mediana y modo) y las de dispersión (desvío estándar, rango intercuartílico). En otras palabras, cuando los datos disponibles están dados en una escala numérica, básicamente debemos calcular dos tipos de medidas llamadas medidas de centralización y medidas de dispersión; las medidas de orden son utilizadas en algunas ocasiones.

Los tres tipos de medidas son útiles no sólo para describir de manera más sintética los datos obtenidos, sino para poder comparar de forma más precisa y eficiente las observaciones realizadas.

4.2.2 Medidas de Tendencia Central

Las Medidas de Tendencia Central más utilizadas son:

Moda

Media aritmética

Mediana

Moda (Mo)

Así como "algo" que está de moda, habitualmente es "algo" que se ve con mucha frecuencia. En Bioestadística, la Moda es el valor que se presenta con mayor frecuencia en el conjunto de los datos obtenidos.

Media aritmética (X)

La media aritmética es también conocida como promedio. Se calcula como la suma de las observaciones dividida por el número total de observaciones (n). Supongamos tener n datos que notaremos como $X_1, X_2, X_3, \dots, X_n$. La media aritmética de estos valores se define como la suma de todos ellos, dividida por n y se simboliza con \bar{X} . Ud. puede encontrar ese valor realizando el siguiente cálculo:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

Es conveniente utilizarla sólo si la distribución de los datos tiene una forma simétrica y unimodal (curva normal). Esto es debido a que la media aritmética es muy sensible a la presencia de valores extremos. En estos casos, es conveniente utilizar otra medida de

centralización, llamada mediana.

Mediana (Md)

La mediana es la observación que deja la mitad de los datos (ordenados de menor a mayor) a cada lado. Ud. puede fácilmente encontrar ese valor realizando el siguiente cálculo:

$$\text{Posición de la Md en un conjunto de datos ordenados} = \frac{(n + 1)}{2}$$

(de menor a mayor)

Si el número total de observaciones (n) es impar existirá una única mediana y será el valor que ocupe la posición central. Si el número de observaciones es par, no existe un único valor central y la mediana corresponde a la media de los dos valores centrales.

La mediana es la medida de centralización que se utiliza cuando la variable se mide con una escala numérica u ordinal. Es la medida más apropiada en caso de que la distribución de frecuencia de la variable sea asimétrica, ya que es menos sensible que la media a valores extremos.

4.2.3 Medidas de Orden

Percentiles

Los percentiles son valores que dividen al conjunto de datos, dejando por debajo de ellos determinados porcentajes.

Por ejemplo, el percentil 10 deja por debajo al 10% de los valores observados.

Cuando definimos mediana dijimos que era el valor que separaba el conjunto de los datos en dos mitades iguales, por lo que ahora también sabemos que la mediana es también el percentil 50. Los percentiles se denotan por P1 ,..... P99.

Cuartiles:

Los cuartiles son tres valores que dividen al conjunto de datos observados en cuatro partes de forma tal que el 25% queda por debajo del cuartil 1, el 50 % por debajo del cuartil 2 y el 75% por debajo del cuartil 3. Los cuartiles, por lo tanto, son los percentiles 25, 50 y 75, respectivamente. Los cuartiles se denotan por Q1 , Q2 y Q3.

4.2.4 Medidas de Dispersión

Como mencionamos anteriormente, expresan el grado de variación de los datos. Debe destacarse que dos series de mediciones pueden coincidir en media, mediana y moda y, sin embargo, los datos pueden distribuirse en forma muy diferente alrededor de estas medidas de centralización. Por eso, una medida de centralización siempre debe ir acompañada de una de dispersión. Las más utilizadas son el rango, el rango intercuartílico, el desvío estándar, la varianza .

A continuación describiremos todas estas medidas excepto la varianza.

Rango (R)

Se calcula como la diferencia entre la mayor observación y la menor.

$$R = [\text{valor mayor observado} - \text{valor menor observado}]$$

¿Qué desventajas tiene?

Tiene como desventaja que su cálculo se basa en sólo dos valores que, por ser los extremos, pueden ser muy atípicos. No da información acerca de cómo se dispersan los datos dentro del intervalo limitado por el menor y el mayor valor.

Rango Intercuartílico (RI)

Para solucionar en parte el inconveniente que presenta el rango, se pensó en una medida simple que no se basa en los valores extremos, sino calculando el intervalo entre el primer y el tercer cuartil. A esta medida de dispersión se la denominó Rango Intercuartílico (RI), y se define como la distancia entre los valores entre los cuales se encuentra el 50% central de los datos.

Por lo tanto, la fórmula sería:

$$RI = Q3 - Q1$$

El rango intercuartílico debe utilizarse como medida de dispersión que complementa la información suministrada por la mediana.

Desvío estándar (S ó DE)

Es la medida de dispersión más utilizada aunque su significado y su cálculo pueden parecer algo complejos. El S mide la dispersión de los datos alrededor de la media aritmética. Por lo tanto, su uso es correcto siempre que la media sea la medida de centralización adecuada.

¿Cuándo era apropiado el uso de la media?

En una distribución simétrica.

Le daremos la fórmula del S, que de ninguna manera pretendemos memorice, sino simplemente para que tenga una idea de cómo calcularlo (de todas maneras no se preocupe, hoy con cualquier computadora podrá conocer el S del conjunto de sus datos sin tener que andar haciendo tantos cálculos).

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

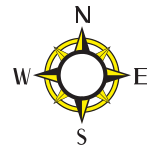
El símbolo \sum significa sumatoria. Ahora, ¿qué sumamos? Las diferencias entre cada valor individual y la media (\bar{x}) elevado al cuadrado. Una vez que hicimos este cálculo, se divide por el número de individuos de la población (n) menos 1 y finalmente calculamos la raíz cuadrada de este resultado.

Existe un concepto muy importante acerca del S, y es que sin importar la forma en que se distribuyen las observaciones, por lo menos el 75% de los valores quedan siempre entre la media menos dos desvíos y la media más dos desvíos. Si la distribución es simétrica y unimodal:

- ▶ 68% de las observaciones están entre la media menos un desvío y la media más un desvío.
- ▶ 95% de las observaciones están entre la media menos dos desvíos y la media más dos desvíos.
- ▶ 99.7 % de las observaciones están entre la media menos tres desvíos y la media más tres desvíos.

FICHA DE ORIENTACIÓN N° 11

Un consejo práctico para determinar la simetría de la curva



- Si bien existen distintos métodos estadísticos para determinar la simetría de la

Cuantificación de los Problemas de Salud

curva, una forma sencilla para conocer si la curva es simétrica o no, es calculando las media, la mediana y la moda.

- Si estas tres medidas son iguales o muy similares, la curva es simétrica o aproximadamente simétrica; si son diferentes, entonces, será asimétrica.

¿Qué medidas de centralización y dispersión utilizar?

- Si la curva es simétrica, es conveniente utilizar la media acompañada del desvío estándar para resumir los datos.
- Si la curva es asimétrica, es conveniente utilizar la mediana acompañada del rango intercuartílico.

D. A Modo de Síntesis...

Si Ud. comprendió los conceptos presentados en éste Módulo, esta listo para cuantificar los problemas de la Salud.

La cuantificación es un proceso absolutamente necesario para la interpretación de los procesos de salud y enfermedad.

Para poder comprender por qué y cómo ocurren los eventos en Salud y enfermedad deberá recorrer además de éste, otros caminos.

Glosario General

Agente: Factor (biológico, físico o químico) cuya presencia es necesaria para que se produzca una enfermedad por exceso, defecto o alteración.

Caso Índice: Es el primer caso diagnosticado de un Brote.

Caso Primario: Es el primer caso de un Brote o Epidemia y en general es reconocido en forma retrospectiva. Se llaman casos co-primarios a los casos que se presentan luego del caso primario y antes de cumplido el período de incubación mínimo y que se supone comparten con él la fuente de infección.

Caso Secundario: Casos generados a partir del caso primario.

Centilos: Ver cuantilos.

CIE-10: Clasificación estadística internacional y problemas relacionados con la Salud. Décima revisión Organización Mundial de la Salud, Washington, 1995.

Confiabilidad: Es el grado en el que mediciones repetidas de un fenómeno relativamente estable caen cerca unas de las otras. Su principal propiedad es la repetitibilidad.

Cuantilos: División de una distribución de datos en subgrupos con igual número de datos y ordenados. Los decilos dividen la muestra en décimos, los quintilos en quintos los teriles en tercios.

Efecto: Resultado de una causa.

"Efecto de Confusión" (confounding): Error introducido en la investigación por la presencia de factores de confusión. Se puede controlar en el proceso de identificación de grupos por medio de la randomización, apareamiento y restricción, o bien, durante el análisis de los datos, por estratificación y análisis multivariado.

"Efecto del Trabajador Sano": Tipo particular de sesgo de selección. Fenómeno que ocurre cuando se comparan ciertas características de un grupo de trabajadores con la población general como consecuencia de que los sujetos empleados son en promedio más sanos que la población general. Por lo tanto, todo exceso de riesgo detectado frente a la exposición del factor en estudio sería minimizado si la comparación se realiza con la población general.

Enfermedad Infecciosa: Es la que deriva de la presencia de un agente infeccioso sea transmisible (ejemplo: sarampión) o no (ejemplo: infección urinaria). Muchas veces se incluyen en este grupo las enfermedades ocasionadas por productos de los agentes infecciosos (toxinas), como por ejemplo, el botulismo o la diarrea por toxina estafilocócica.

Enfermedad Transmisible: Enfermedades producidas por agentes infecciosos o sus toxinas que llegan a un individuo susceptible por transmisión desde otro individuo infectado, animal o reservorio.

Epidemiología Analítica: Parte de la Epidemiología que se ocupa del estudio de los determinantes de las enfermedades. Su principal característica radica en la utilización de un adecuado grupo de comparación o grupo control.

Error: Toda diferencia entre el valor medido, observado o calculado y el verdadero valor. Al realizar cualquier estudio epidemiológico pueden cometerse tres tipos diferentes de errores:

- Error de Medición: Inherente a la precisión de un instrumento de medida.
- Error Sistemático: también llamado sesgo (se describirá en el siguiente apartado), caracterizado por ocurrir siempre en una misma dirección en contraste con el Error por Azar.

Error Aleatorio: El error por azar característicamente sobreviene al realizar mediciones en sólo una porción (muestra) de la población de estudio, también llamada población diana. Recibe también el nombre de error muestral.

Escala: Aparato o sistema empleado en la medición de porciones iguales.

Estandarización: Técnica empleada para remover efectos de diferencias de estructuras de edad u otros factores de confusión cuando se comparan poblaciones. Existen dos métodos el directo y el indirecto cuyo detalle escapan este curso

Estudios Descriptivos: Estudios que describen la distribución de frecuencias de las variables de tiempo lugar y persona asociadas a un evento dado. Permiten generar hipótesis pero no probarlas.

Estudios Analíticos: Estudios que identifican o miden los efectos de diferentes factores de riesgo sobre la Salud. Examinan o miden asociaciones y evalúan hipótesis sobre asociaciones causales. Su característica saliente respecto del diseño es que emplean grupos de comparación.

Estudio de Casos y Controles: Estudio analítico observacional en el que el grupo de

estudio y de control son definidos de acuerdo a la presencia o no de enfermedad.

Estudio de Cohorte: Estudio analítico observacional en el que el grupo de estudio y el grupo control son definidos de acuerdo a la presencia o no de un factor de exposición.

Estudio Longitudinal: Estudio que permite investigar la dinámica de una variable o de un fenómeno de salud a través del tiempo.

Estudio Transversal: Estudio en el que se investiga un fenómeno de Salud en un determinado momento en el tiempo.

Exposición: Proximidad y/o contacto con un agente de enfermedad (o protección) de modo tal que pueda ocurrir la transmisión efectiva del agente. También se usa como la cuantificación de la exposición de un individuo o grupo a un determinado factor.

Factibilidad: Posibilidad de llevar a la práctica un procedimiento, programa, medida de control, estudio, etcétera.

Factor de Confusión: Variable que se asocia tanto con el factor en exposición como con el evento en estudio distorsionando la asociación existente entre estos dos, e introduciendo un error en la investigación llamado "Efecto de Confusión".

Factor de Protección: Un factor de protección es el atributo de un grupo con menor incidencia de una determinada enfermedad en relación con otros grupos, definidos por la ausencia o baja aparición del tal factor.

Factor de Riesgo: Puede ser definido como el atributo de un grupo que presenta mayor incidencia de una determinada patología en comparación con otros grupos poblacionales, caracterizados por la ausencia o baja aparición de tal condición.

Fuente Común: Brote generado a partir de un único foco que usualmente es agua o alimentos.

Fuente Propagada: Es la que ocurre cuando una enfermedad transmisible de persona a persona en una población de susceptibles.

Grupo de Riesgo: Aquel que posee un riesgo mayor de presentar una determinada enfermedad o evento.

Intervención: Modificación intencional del sujeto en alguno de sus aspectos, como por ejemplo, la introducción de algún régimen terapéutico o preventivo.

Incidencia: Número de personas que adquieren una enfermedad en un período dado en una población específica.

Marcadores de Riesgo: Características personales como la edad, el sexo, la raza cuyo efecto se halla fuera de control y son inmodificables.

Media: También llamada promedio. Medida de tendencia central que se obtiene de sumar los valores de un conjunto de datos dados y dividirlos luego por el número total de datos.

Mediana: Medida de tendencia central. Es el valor que asume el dato central de la muestra una vez ordenados los mismos de menor a mayor. Equivale al percentilo 50.

Medidas de Asociación: Ver medidas de efecto.

Medidas de Efecto: Cantidad que refleja la fuerza de la asociación entre variables. Las medidas de este grupo que trabajamos en los módulos son la razón de tasas o riesgos, razón de odds y diferencias de tasas o riesgos. Existen otras que exceden este curso.

Medidas de Impacto: Medidas que permiten cuantificar el impacto que tiene sobre la morbimortalidad de una población la exposición a un determinado factor. Permiten que, sobre la base de su estimación, los administradores de Salud determinen la importancia relativa priorización- de cada factor de riesgo para el cual podría desarrollarse un programa y decidir racionalmente la utilización de los recursos disponibles.

Medida de Resumen: Medidas que agrupan o resumen varias medidas individuales. Pueden corresponder a personas (por ejemplo tasas de mortalidad) o no (promedio de temperatura anual en un determinado lugar). Las medidas de resumen se pueden agrupar en medidas de tendencia central, de dispersión, de orden y de frecuencia.

Modo: Medida de tendencia central. Es el valor de los datos que más se repite.

Modificador de Efecto: Factor que modifica el efecto de un factor causal putativo en estudio. Se trata de un factor de riesgo cuya presencia en diferentes niveles o estratos hace modificar la fuerza de asociación existente entre otro factor de riesgo y el evento en estudio.

Estratificación: División de una población en estudio en subgrupos o estratos de acuerdo a una determinada variable de interés.

Muestreo Aleatorio Simple (MAS): Forma de muestreo probabilístico en el que la selección se realiza al azar, resultando todos los individuos de la población con igual probabilidad de ser seleccionados.

Muestreo Estratificado: Tipo de muestreo probabilístico en el que la población diana se divide en estratos de acuerdo a alguna característica, generalmente demográfica, y posteriormente se realiza un MAS de cada estrato obteniéndose una muestra de la población en la que están representados todos los estratos conformados.

Muestreo por Conglomerados: Tipo de muestreo probabilístico en el que la población diana se divide en conglomerados (familias, barrios, ciudades, etc.) Se elige una muestra aleatoria simple de cada conglomerado, resultando todos ellos representados en la muestra global de la población.

Muestreo Probabilístico: Técnica de muestreo que consiste en extraer una muestra de una población, de tal manera que todos los individuos de la población tengan una probabilidad conocida de ser seleccionados.

Muestreo Sistemático: Tipo de muestreo probabilístico en el que se da al grupo del que se tomará la muestra, una especie de ordenamiento y luego la elección se hace sistemáticamente a lo largo de la serie, por ejemplo, cada segundo, cada centésimo o cada milésimo individuo.

Odds Ratio (OR): Medida de asociación entre un factor de riesgo y la enfermedad que resulta del cociente entre el Odds del grupo expuesto sobre el Odds del grupo no expuesto. Expresa cuantas veces más probable es el riesgo de contraer la enfermedad en presencia de un factor de exposición que en su ausencia.

Patogenicidad: Propiedad de un organismo que determina qué grado se produce enfermedad en la población infectada, dicho de otra manera es la capacidad del organismo para causar enfermedad. Se mide como la razón que se obtiene del número de personas que desarrollan enfermedad clínica sobre los expuestos.

Período de Incubación: Período entre el momento de infección (ingreso del agente al organismo) y la aparición de signos y síntomas (pródromos). La duración de este período puede variar según la definición que se utilice para estos signos y síntomas. En general este período se expresa como un rango (período de incubación máximo y mínimo) y un promedio.

Período Infeccioso: Es el período durante el cual una persona infectada puede transmitir el agente infeccioso. La duración de este período es importante para las medidas de control de la enfermedad. Es importante tener en cuenta que este período puede iniciarse antes de que el paciente tenga síntomas, por lo cual se verificaría transmisión antes de saber que persona está enferma.

Período de Latencia: Período que transcurre entre la infección y el inicio del período infeccioso. La duración de este período será un determinante de los intervalos entre sucesivas infecciones en la cadena de transmisión.

Población de Riesgo: Ver grupo de riesgo.

Población Diana: Población sobre la cual se desea conocer una determinada característica susceptible de ser estudiada.

Prevalencia: Número de casos existentes en una determinada población en un momento determinado en el tiempo.

Prevención Primaria: Medidas y actividades tendientes a la promoción y protección de la salud. Tiende a disminuir la incidencia de enfermedad evitando la aparición de enfermedad y a fomentar el mantenimiento de la salud

Prevención Secundaria: Medidas y actividades tendientes a restaurar la salud toda vez que esta se haya perdido. Tiende a disminuir la prevalencia de la enfermedad acortando la duración de la misma.

Promedio: Ver media.

Proporción: Es una razón en la cual el numerador está incluido en el denominador. Es una parte del todo. Por lo cual la proporción reúne dos características: no puede arrojar un número mayor a 1 (el numerador es menor o a lo sumo igual al denominador) y no tiene unidades ya que el numerador y el denominador tienen la misma unidad y se cancelan entre sí.

Se puede expresar como fracción decimal: 0,2, como fracción $1/5$ o como porcentaje: 20%.

Proporción Atribuible en Expuestos: Se obtiene al dividir el resultado de la diferencia entre las proporciones o tasas de incidencia de individuos expuestos y no expuestos con la incidencia propia del grupo expuesto. Expresa el porcentaje de una enfermedad que puede ser atribuido exclusivamente a dicho factor.

Prospectivo: Todo estudio en el cual al momento de su inicio la enfermedad no

ocurrió y los individuos son seguidos en el tiempo con el fin de detectar la ocurrencia de la misma.

Razón: En términos generales: " Es el valor que se obtiene de dividir una cantidad por otra". Agrupa a las tasas, razones, proporciones, etc. , sin embargo, es un concepto más amplio que estos. La razón es una expresión de la relación existente entre un numerador y un denominador, donde ambos números son cantidades independientes pudiendo presentar unidades de medición diferentes. Además, no es necesario que una incluya a la otra como en el caso de la proporción o porcentaje.

Repetitibilidad: Habilidad de lograr el mismo resultado en distintas medidas realizadas del mismo modo.

Representatividad: Se dice que una muestra es representativa cuando logra exhibir internamente el mismo grado de diversidad que la población diana de la que se tomó.

Reservorio: Especies o poblaciones que tienen la capacidad de mantener al agente por tiempo indefinido.

Retrospectivo: Todo estudio en el cual al momento de su inicio tanto la enfermedad como la exposición han ocurrido.

Riesgo: Se define como la probabilidad de que uno de los miembros de una población definida desarrolle una enfermedad dada en un período de tiempo.

Riesgo Atribuible: Medida de asociación que representa la diferencia de la incidencia de la enfermedad en la población expuesta al factor de riesgo y la incidencia en la población no expuesta a dicho factor. También se la considera una medida de impacto, ya que expresa el número o porcentaje de casos que podrían eliminarse en el grupo expuesto, si se removiera el factor de riesgo en cuestión.

Riesgo Atribuible Poblacional: El RAP representa la diferencia de la incidencia de la enfermedad en toda población y la incidencia en la población no expuesta a dicho factor. Puede expresarse como:

- Riesgo Atribuible Poblacional (en términos absolutos) (RAP): Indica la proporción de casos de una enfermedad en una población que pueden atribuirse exclusivamente a la presencia del factor de riesgo en consideración.
- Riesgo Atribuible Poblacional Porcentual (en términos relativos)(RAPP): representa el porcentaje de la incidencia total de una determinada enfermedad en una población que puede atribuirse exclusivamente a la presencia de dicho factor o bien el porcentaje de la incidencia total de la enfermedad en la población que podría reducirse con la remoción del dicho factor.

Riesgo Relativo: Medida de asociación entre un factor de riesgo y una enfermedad que resulta del cociente de las tasas de incidencia de la población expuesta y la no expuesta. Expresa cuántas veces más probable es el riesgo de contraer la enfermedad en presencia de un factor de exposición que en su ausencia

Sesgo: Error sistemático que afecta la validez de una investigación.

Sesgo de Información: Falla en la medición de los datos de la exposición o evento que resultan en diferencias sistemáticas de la calidad de información entre los grupos de comparación en estudio.

Sesgo de Seguimiento: Falla en la medición de los datos de la exposición o evento resultantes de la pérdida diferencial de sujetos en seguimiento en algunos de los grupos de comparación en estudio.

Sesgo de Selección: Error sistemático que se produce durante la etapa de identificación de los grupos de estudio y de grupo control, que origina una diferencia en los grupos comparados y que introduce un error en los resultados observados.

Susceptible: No infectado que puede infectarse. Los inmunes no son susceptibles por poseer protección celular o humoral (anticuerpos).

Tasa: Medida de la frecuencia con la que ocurre un fenómeno. Todas las tasas son razones, algunas son proporciones. Es la expresión de la frecuencia con que ocurre un evento en una población en un tiempo determinado, sea un período de tiempo o un momento puntual. Los componentes de una tasa son: el numerador, el denominador, el período de tiempo específico en el que puede ocurrir el evento de nuestro interés y habitualmente un coeficiente múltiplo de 10, que convierte la tasa en un número entero permitiendo una interpretación más fácil de la misma. El uso de tasas es esencial para la comparación de poblaciones en distintos momentos, lugares o diferentes grupos dentro de la misma población.

Técnicas de Muestreo: Procedimientos que permiten la selección de una parte de la población diana (muestra).

Transmisión Vertical: Transmisión trans placentaria de madre a hijo.

Validez: Es el grado en que un estudio o de un instrumento miden exactamente lo que desea medir.

Vector: Invertebrado que transporta el agente de un vertebrado a otro.

Vehículo: Medio por el cual el agente llega al huésped.

VIH: Virus de la inmunodeficiencia humana agente causal del síndrome de inmunodeficiencia humana (SIDA).

Virulencia: Grado de patogenicidad de una enfermedad. Capacidad de un microorganismo de generar enfermedad en un huésped susceptible.