

Manual de estadística e indicadores epidemiológicos en odontología

UNIVERSIDAD DE LA REPÚBLICA
COMISIÓN SECTORIAL DE ENSEÑANZA

Susana Lorenzo-Erro
Ramon Alvarez-Vaz
Fernando Massa
Mariana Musto-Viscailuz
Patricia Olmos
Anunzziatta Fabruccini
Matias Muñoz



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



comisión sectorial
de enseñanza



Manual de estadística e indicadores epidemiológicos en odontología

Susana Lorenzo-Erro, Ramón Álvarez-Vaz, Fernando Massa,
Mariana Musto-Viscailuz, Patricia Olmos,
Anunzziatta Fabruccini, Matías Muñoz

COLABORADORES

Laura Cosetti, Ernesto Andrade, Luis Bueno,
Matías Muñoz, Martín Sanguinetti, Andrés Rodríguez
Raúl Riva

COMPAGINACIÓN

Fiorela Apelo

REVISIÓN

Valentina Colistro

Manual de estadística e indicadores epidemiológicos en odontología / Susana Lorenzo Erro, Ramón Álvarez Vaz, Fernando Massa, Mariana Musto Viscailuz, Patricia Olmos, Anunzziatta Fabruccini, Matías Muñoz
- Montevideo: Universidad de la República. Comisión Sectorial de Enseñanza, 2023.
154 p. -- (Manuales Didácticos / Comisión Sectorial de Enseñanza).

ISBN: 978-9974-0-2059-7

1. ODONTOLOGIA
2. EPIDEMIOLOGIA
3. DATOS ESTADISTICOS
4. ENSEÑANZA SUPERIOR
5. UNIVERSIDADES

I. Lorenzo Erro, Susana II. Álvarez Vaz, Ramón III. Massa, Fernando IV. Musto Viscailuz,
Mariana V. Olmos, Patricia VI. Fabruccini, Anunzziatta VII. Muñoz, Matías

CDD: 617.6

La publicación de este libro fue realizada con el apoyo de la Comisión Sectorial de Enseñanza (CSE) de la Universidad de la República.

Este manual es el producto de la presentación de un proyecto en el llamado de la Comisión Sectorial de Enseñanza (CSE) de la Universidad de la República (Udelar) «Elaboración de manuales didácticos para la enseñanza de grado», 2017, que fuera aprobado en el mismo año y financiado durante 2018.



Atribución-NoComercial-CompartirIgual 4.0, Susana Lorenzo-Erro; Ramón Álvarez-Vaz; Fernando Massa; Mariana Musto-Viscailuz; Patricia Olmos; Anunzziatta Fabruccini; Matías Muñoz, 2024. Facultad de Odontología.

<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Comunicación y Publicaciones, CSE, Udelar.

Alberto Lasplaces 1620, 11600 Montevideo

Tel.: (+598) 2400 8393

www.cse.udelar.edu.uy comunicacion@cse.udelar.edu.uy

ISBN: 978-9974-0-2059-7

Coordinación editorial: Vanesa Sanguinetti

Corrección de estilo y diseño editorial: Felipe Correa

TABLA DE CONTENIDOS

Introducción

¿Por qué un manual sobre estadística aplicada e indicadores de salud bucal?	9
La epidemiología, la estadística y la salud colectiva: un desafío de integración y posicionamiento	11

Parte I

Estadística descriptiva y aspectos básicos de estadística inferencial

Descripción univariada – capítulo 1

Presentación de datos y variables	15
Introducción	15
Escala	16
Clasificación	18
Ejercicios	21

Descripción univariada – capítulo 2

Distribución de frecuencias	23
Introducción	23
Distribución de frecuencias (DF)	23
Tipos de frecuencia	24
DF para variables continuas	25
Representación gráfica	27
Ejercicios	31

Descripción univariada – capítulo 3

Medidas de resumen	33
Introducción	33
Medidas de posición	34
Medidas de dispersión	43
Representación gráfica	45
Ejercicios	49

Descripción bivariada – capítulo 4	
Análisis de dos variables cuantitativas	51
Introducción	51
Medidas de asociación lineal	51
Ejercicios	55
Descripción bivariada – capítulo 5	
Análisis de dos variables cualitativas	57
Introducción	57
Tabla de contingencia	57
Medidas de asociación	64
Consideraciones finales	68
Ejercicios	69
Breve introducción a la estadística inferencial – capítulo 6	
Introducción a la inferencia estadística	71
Introducción	71
Definiciones	71
Procedimientos inferenciales	76
Ejercicios	79

Parte II

Indicadores epidemiológicos en salud bucal

Indicadores y medidas de frecuencia – capítulo 7	
Concepto y características de los indicadores epidemiológicos	83
Introducción	83
Indicadores al realizar un diagnóstico sociocomunitario	83
Indicadores y medidas de frecuencia – capítulo 8	
Los indicadores en el proceso de operacionalización de variable en una investigación	87
Indicadores y medidas de frecuencia – capítulo 9	
Medidas de frecuencia, relaciones entre variables y comparación de poblaciones	89
Proporción	90
Prevalencia	91
Incidencia	91

Proporción de incidencia acumulada	93
<i>Odds ratio</i> / chance	94
Tasa	96
Tasas estandarizadas	98
Razones	98
Medidas de asociación	100
Indicadores de enfermedades bucales – capítulo 10	
Enfermedad periodontal	105
Breve descripción clínico-epidemiológica	105
Indicadores de enfermedad periodontal usados en epidemiología	106
Ejercicio 10.1	114
Abordaje y uso de los índices de EP en la clínica	116
Indicadores de enfermedades bucales – capítulo 11	
Caries dental	119
Breve descripción clínico-epidemiológica	119
Criterios para el diagnóstico de caries	120
Unidades de medida de la caries dental	123
Principales indicadores epidemiológicos de caries dental	124
Índice significativo de caries (SIC)	128
Índice de cuidados odontológicos	128
Otros índices (OSD y DS)	129
International Caries Detection and Assesment Sistem (ICDAS)	130
Medidas de frecuencia y características relevantes para medición de caries dental	131
Ejercicios	133
Indicadores de enfermedades bucales – capítulo 12	
Cáncer bucal	135
Breve descripción clínico-epidemiológica	135
Indicadores de cáncer bucal	137
Indicadores de enfermedades bucales – capítulo 13	
Trastornos témporo mandibulares (TTM)	139
Breve descripción clínico-epidemiológica de los TTM y el bruxismo	139
Principales indicadores epidemiológicos	140
Referencias	143

Este manual está dedicado a todos los estudiantes de grado de la salud, y muy especialmente a los estudiantes de Odontología, como futuros integrantes de los equipos de salud. Nuestro anhelo es que les sea útil no solo durante la carrera, sino también en el ejercicio de su práctica sanitaria. Práctica sanitaria de la cual deseamos fervientemente que pueda beneficiarse la población uruguaya, ejerciendo así su derecho a la salud.

¿Por qué un manual sobre estadística aplicada e indicadores de salud bucal?

La iniciativa para la realización de este manual parte de un grupo de docentes del Servicio de Epidemiología y Estadística de la cátedra de Odontología Social de la Facultad de Odontología (FO) de la Universidad de la República (Udelar), que viene trabajando en estos temas desde hace varios años en el ámbito de la enseñanza y la investigación. En el plan de estudios de la FO figura como una de las competencias generales del perfil del odontólogo: «analizar e interpretar los resultados relevantes de las investigaciones experimentales, epidemiológicas y clínicas», y entre las competencias específicas: «interpretar la situación social y epidemiológica en que se inserta su actuación profesional, identificando los problemas de salud bucal que afectan a las personas y a la comunidad, con el objetivo de mejorar la calidad de vida de la población» (Facultad de Odontología, 2014). Se trata de una institución en la que la investigación en salud bucal está en un proceso de desarrollo creciente, en un contexto sanitario en que se está desarrollando el Sistema Nacional Integrado de Salud (SNIS), donde los futuros profesionales de la salud serán actores fundamentales. En este sistema hay un constante flujo de datos e información, es necesario tener la capacidad de analizarlos para implementar un proceso de vigilancia epidemiológica, evaluar su desarrollo y realizar mejoras continuas.

El plan de estudios vigente —en particular el área de Salud Colectiva, a la que pertenecen los docentes de la cátedra de Odontología Social— propone como objetivo general:

conocer los aportes de la epidemiología general y bucodental para el estudio de las causas y distribución del proceso salud-enfermedad así como para la planificación, ejecución, vigilancia y evaluación de las intervenciones sanitarias. Manejar la aplicación del método científico a una situación socio-epidemiológica concreta [... y] conocer la estadística como herramienta descriptiva y analítica a ser utilizada en el desarrollo, análisis e interpretación de trabajos científicos.

Teniendo en cuenta los objetivos académicos del área de Salud Colectiva y las competencias generales que se busca desarrollar en el egresado de la carrera de Doctor en Odontología, se advierte que es necesario el conocimiento teórico y práctico de la epidemiología y la estadística como disciplinas científicas que contribuyen a su formación integral.

Los cursos del plan de estudios de la carrera de Doctor en Odontología en que se abordan contenidos de estadística y epidemiología son Iniciación a la Investigación —ubicado en el primer año de la carrera, con una carga de 60 horas— y Salud Colectiva I —ubicado en el tercer año, con 90 horas—.

El equipo de docentes del servicio decidió elaborar este manual luego de comprobar la inexistencia de textos en español que traten sobre indicadores epidemiológicos y estadística aplicada a la salud bucal con la extensión y profundidad que requieren los cursos Iniciación a la Investigación y Salud Colectiva I. Este libro está dirigido especialmente a estudiantes de ambos cursos; asimismo puede servir como material de apoyo para quienes desarrollan su trabajo final de carrera (TFC).

El objetivo general de este manual es integrar los aspectos esenciales de la descripción estadística con aquellos de la investigación clínico-epidemiológica y favorecer el proceso de enseñanza-aprendizaje de ambas disciplinas, respondiendo a las necesidades del estudiante de Odontología, por lo tanto se adapta a los programas de los cursos de grado mencionados.

Como objetivos específicos se proponen dos:

1. Generar un material didáctico que reúna los principales contenidos sobre estadística descriptiva, aspectos básicos de estadística inferencial e indicadores epidemiológicos de salud bucal previstos en los programas de los cursos de Iniciación a la Investigación, Salud Colectiva y Trabajo Final de Carrera.
2. Ofrecer una herramienta que pueda ser utilizada a lo largo de toda la carrera, cada vez que la formación del estudiante o profesional así lo requiera.

La epidemiología, la estadística y la salud colectiva: un desafío de integración y posicionamiento

Es posible conceptualizar la epidemiología como la disciplina que estudia el proceso salud-enfermedad en las colectividades humanas, analizando la distribución y los factores determinantes de las enfermedades, daños a la salud y eventos asociados, proponiendo por un lado medidas específicas de prevención, control y erradicación de enfermedades, y por otro aportando indicadores que den soporte al planeamiento, la administración y la evaluación de las acciones en salud (Rouquayrol *et al.*, 2014).

Para estudiar las colectividades humanas y analizar sus perfiles de salud y enfermedad se trabaja permanentemente articulando datos que provienen de las personas, construyendo sistemas de información y generando teorías que buscan explicar la realidad, darle sentido para poder transformarla. En este marco las personas son consideradas en tanto sujetos sociales. Los datos son materia prima que cuando se procesa adquiere sentido y se convierte en información. Una teoría puede entenderse como un sistema intelectual que integra conceptos que refieren a uno o más paradigmas.

Los paradigmas son los lentes con los que percibimos la realidad, son los conjuntos de ideas, formas de sentir y valores que nos condicionan a la hora de aproximarnos a ella. Los elementos que componen un paradigma científico son, según el profesor Jaime Breilh (2010), objeto, sujeto y praxis. De la relación entre el sujeto y el objeto surge el conocimiento científico. Las experiencias perceptivas no están dadas únicamente por estímulos de la naturaleza, sino que dependen de la cultura y las estructuras simbólicas que usamos para conocer, una de estas es el método científico. Este no es un movimiento exclusivamente inductivo (de objeto a sujeto, de observación a teoría), ni tampoco un movimiento exclusivamente deductivo (de sujeto a objeto, de teoría a conclusión). El método científico es un movimiento dialéctico que se da en un campo práctico e histórico determinado. Teniendo en cuenta lo anterior es posible entender que al acercarnos a la realidad en forma fragmentada, observando variables y juntándolas —inscriptos en un enfoque de la epidemiología tradicional y realizando un abordaje empírico-analítico—, tendremos un conocimiento distorsionado de esta, no lograríamos comprenderla en su totalidad.

Pero yendo desde la epidemiología tradicional hacia un paradigma más integrador, que será necesariamente complejo y crítico, se conceptualiza la salud como un proceso vital, biológico, determinado socialmente, un objeto polisémico donde

puede distinguirse un objeto real, uno pensado y la praxis (Breilh, 2013). Diversas disciplinas científicas sociales y biológicas contribuyen para ayudar a entender la salud, entre ellas la epidemiología y la estadística, y desarrollan permanentemente la triangulación de métodos, investigadores e información. El objetivo final es abordar el sujeto en estudio y la población, entenderlos, interpretarlos y elaborar soluciones a sus problemas de salud. Desde esta perspectiva integradora, la estadística se presenta como una disciplina científica auxiliar de la epidemiología, a la que ayudará aportando datos e información cuantitativa; esta deberá ser integrada a la información cualitativa disponible para así mejorar el armado el rompecabezas de situaciones y procesos de salud en las poblaciones.

De esta forma, al analizar datos y convertirlos en información con sentido, siendo conscientes o no, nos estaremos inscribiendo en uno o varios paradigmas y esto condicionará la forma en que explicaremos la porción de realidad sobre la que trabajamos. Por lo tanto es importante tomar conciencia y —en lo posible— explicitar el paradigma o los paradigmas utilizados, de modo de que quede claro cuál es nuestro punto de vista y poder tomar distancia de este cuando sea necesario.

En suma, siendo conscientes del papel que desempeñamos como académicos, epidemiólogos, estadísticos, investigadores, y de nuestra relación con el conocimiento como fuente de poder, es necesario que nos responsabilicemos de nuestras opciones teórico-metodológicas. Sean conscientes o no, los paradigmas en los que el investigador se inscribe condicionan lo que ve y lo que no ve, lo que muestra y lo que no muestra, y esto tiene consecuencias directas sobre las acciones sanitarias con y sobre los sujetos con los que se trabaja. En definitiva, cada vez que se investiga está en juego nuestro papel ético-político como académicos.

PARTE I

Estadística descriptiva y aspectos básicos
de estadística inferencial

Presentación de datos y variables

Introducción

Para poder estudiar el mundo que nos rodea es necesario hacer mediciones. No todo puede medirse, ni todo lo que medimos sirve.

En este capítulo se estudiará la definición estadística de *variable*, los diferentes *tipos de variables* y las distintas *escalas de medición*.

Definimos *variable* como toda aquella característica de un objeto que puede ser medida o categorizada (Kim & Dailey, 2008a).

Como toda definición, esta incluye y excluye elementos, por ejemplo, ¿se puede identificar cuáles de los siguientes ítems cumplen con la definición de variable y cuáles no?

- Dientes.
- Concentración de flúor.
- La película que vi ayer.
- Nombre del paciente.
- Dolor.

El primer elemento de la lista claramente no refiere a una variable, sino a objetos de los cuales puede interesar medir distintas características, como el color, el tamaño, la cantidad presente en la boca, si presentan caries o no, etc.

Siguiendo la lista, se encuentra la concentración de flúor, en este caso sí se está ante una variable, tal vez se debería definir el objeto al que se le va a medir esta variable (por ejemplo, distintas marcas de pastas dentales), las unidades en que se va a realizar la medición y el instrumento que se utilizará, pero la concentración de flúor es una característica de un objeto que puede ser medida.

Luego aparece «la película que vi ayer». Aquí hay que preguntarse, ¿puede tomar distintos valores? Claramente no. (Se podría convertir esta expresión en una variable si se la incorpora a un cuestionario en forma de pregunta: ¿qué película vio

ayer? En ese caso el nombre de la película sería una característica que variaría entre los diferentes encuestados, tomando tantos valores como películas se hayan visto).

«Nombre del paciente» es una característica de los sujetos que puede tomar distintos valores, por lo que es la segunda variable que encontramos en este listado.

Por último, aparece «dolor». Si se puede medir, es una variable. Esto da lugar a la pregunta: ¿es posible medir las sensaciones humanas? Lo que a su vez conduce a la discusión sobre el significado del término *medida* y a las diferentes escalas de medición que se verán a continuación.

Escalas

Las mediciones existen en una variedad de formas. Las escalas de medición caen dentro de ciertos tipos definidos. Estos tipos están determinados por las operaciones empíricas aplicadas en el proceso de medición y por las propiedades formales de las escalas. Además (y esto es de gran interés a varias ciencias), las operaciones estadísticas que pueden aplicarse a un conjunto de datos empíricos dependen del tipo de escala en la cual los datos están ordenados (Stevens, 1946).

Puede decirse que *medición*, en sentido amplio, se define como la asignación de valores a objetos o eventos de acuerdo a reglas. El que los valores puedan ser asignados siguiendo distintas reglas conduce a diferentes tipos de escalas y, así, a diferentes tipos de medida (Stevens, 1946). Entonces resulta indispensable hacer explícitas:

- Las reglas para la asignación de valores.
- Las propiedades matemáticas de las escalas resultantes.
- Las operaciones estadísticas aplicables a las mediciones hechas con cada tipo de escala.

De acuerdo a los puntos antes mencionados se encuentran los tipos de escalas de medición descritas en la tabla 1.1.

Tabla 1.1 Escalas de medición de variables, operaciones empíricas básicas y estadísticas

Escala	Operaciones empíricas básicas	Operaciones estadísticas permitidas
Nominal	Determinación de igualdad	Número de casos Moda Tablas de contingencia
Ordinal	Determinación de mayor o menor	Mediana Percentiles
Intervalos	Determinación de igualdad de intervalos o diferencias	Media Desviación estándar Correlación
Razón	Determinación de igualdad de proporciones	Coefficiente de variación

Escala nominal

La escala nominal es la que impone menos restricciones para la asignación de números. Los números se usan únicamente como etiquetas (de hecho, en lugar de números pueden ser letras o palabras). Algunos ejemplos de esta escala son las variables:

- Número de camiseta de un jugador de fútbol.
- Número de cédula de identidad.
- Color de ojos.
- Línea de ómnibus.

En estos casos las únicas operaciones estadísticas permitidas son contar cuál es la clase que más se repite (moda) y la cantidad de casos en cada clase. *Clase* (o *categoría*) refiere a los grupos en que se divide la variable estudiada («ojos verdes» es una clase dentro de la variable «color de ojos»).

En esta escala la única regla es no asignar diferentes etiquetas a la misma clase o una misma etiqueta a clases diferentes (identificar «ojos verdes» con el número 3 y asignar el mismo número a «ojos claros» impediría una correcta interpretación de un «3» en el registro).

Escala ordinal

La escala ordinal surge al asignar un orden a los valores de la variable. Un ejemplo de esta escala puede ser el puesto de llegada en una maratón: sabemos que 1, 2 y 3 corresponden al primer, segundo y tercer puesto; sin embargo, la distancia entre

quien llegó primero y quien llegó segundo difícilmente sea igual a la distancia entre el segundo y el tercero (y este dato no está presente).

Escala de intervalo

Encontramos la escala de intervalo en las variables de tipo cuantitativo. En esta escala podemos establecer un orden entre los valores, aplicar comparaciones de igualdad y desigualdad y medir distancias entre los valores. Sin embargo, no se pueden realizar multiplicaciones ni divisiones, porque estas escalas carecen de un cero absoluto. En esta escala el «valor cero» se define por conveniencia o por convención.

Un clásico ejemplo de este tipo de escala son dos escalas de temperatura: medida en grados Celsius y medida en grados Fahrenheit. Ambas escalas tienen un valor cero arbitrario que no significa ausencia de la variable medida (en este caso, temperatura). Otro ejemplo de este tipo de escala es la fecha del calendario: diferentes culturas y religiones sitúan el cero del año de acuerdo a distintos criterios. Otro ejemplo posible es la elevación o altitud, en esos casos se coloca —arbitrariamente— un nivel del mar histórico como correspondiente al cero.

Escala de razón

Las variables de esta escala presentan todas las características de las variables de intervalo, pero además cuentan con un cero absoluto que representa la ausencia total de la medida.

La escala de razón nos permite hacer todas las operaciones estadísticas disponibles para las escalas anteriores y además se pueden realizar cálculos de proporciones, razones y coeficientes de variación. Un ejemplo de variable expresada en escala de razón puede ser el salario medido en pesos, en este caso el cero implica la ausencia del salario.

Clasificación

Es posible agrupar las escalas de medición mencionadas de manera que por un lado estén las correspondientes a variables que registran atributos o cualidades y por otro las correspondientes a variables que miden cantidades. Esta agrupación genera una clasificación de variables en *cualitativas* y *cuantitativas*.

Se habla de variables *cualitativas* cuando estas registran la presencia (o ausencia) de cierto atributo. Por otro lado, las variables *cuantitativas* registran qué tanto se presenta un atributo. Dicho en palabras más sencillas: las variables cualitativas expresan la ocurrencia o posesión de una cualidad específica y las variables cuantitativas expresan una cantidad (medida a través de números con los cuales tiene sentido operar).

En resumen: las variables cuya escala de medición es nominal u ordinal se clasifican como cualitativas; las variables de intervalo o razón se clasifican como cuantitativas. La figura 1.1 expresa esto de manera gráfica.



A modo de ejemplo, considere la variable nominal «color de ojos». Algunas posibles clases de esta variable serían «negros», «marrones» o «verdes». Claramente se hace referencia a cualidades y no se está cuantificando nada en particular, debido a esto se trata de una variable cualitativa.

Un ejemplo de una variable de razón es «estatura de una persona». Si la unidad de medida son centímetros, algunos posibles valores serían 171, 156, 182. En este caso se está registrando qué tanto se expresa la variable que se está midiendo, por este motivo se dice que la estatura es una variable cuantitativa.

Adicionalmente, debe señalarse que el grupo de las variables cuantitativas se puede subdividir en dos: las cuantitativas *discretas* y las cuantitativas *continuas*. Las *discretas* son aquellas variables tales que es posible encontrar dos valores entre los que no existen más valores. Por el contrario, en el caso de las variables cuantitativas *continuas* esto no se cumple, entre dos valores cualesquiera siempre habrá infinitos valores.

Para ilustrar esta distinción considere la variable «número de hermanos», claramente esta variable adoptaría los valores 0, 1, 2, 3, etc. Para determinar si se trata de una variable discreta o continua basta con tomar dos valores sucesivos (por ejemplo 2 y 3) y notar que entre ellos no existe otro valor posible de la variable en cuestión. Por este motivo, «número de hermanos» se trata de una variable cuantitativa discreta.

Un ejemplo de una variable cuantitativa continua es «temperatura ambiente». Suponga que se está registrando la temperatura en grados Celsius. Posibles valores de esta variable son 21,3; 15,8; 4,8. Si bien los valores presentados tienen un solo dígito decimal, la temperatura no salta desde 15,7 a 15,8, sino que aumenta de manera gradual. Si contásemos con el instrumental adecuado se podría observar que entre 15,7 y 15,8 existen posibles valores de temperatura, por ejemplo 15,72. Asimismo, entre 15,7 y 15,72 también existen valores de temperatura, por ejemplo 15,7182. Este proceso podría repetirse infinitamente, y así decimos que la temperatura es una variable cuantitativa continua.

Por último, y para no generar confusiones, es importante ahondar un poco en la relación entre estos tipos de variable y las escalas de medición. Las variables cuantitativas deben tener escalas de medición de razón o intervalo (el carácter discreto o continuo de las variables se asocia con ambos tipos de escala). Una variable cuantitativa discreta puede ser de razón o de intervalo en su escala de medición; del mismo modo, una variable cuantitativa continua puede ser de razón o de intervalo en su escala de medición.

Ejercicios

Ejercicio 1.1

Un estudiante se propone estudiar los destinos de vacaciones de sus compañeros de clase. Para esto entrevista a algunos compañeros sobre su destino vacacional favorito. Identifique:

- Sujeto.
- Variable.
- Escala de medición y clasificación.

Ejercicio 1.2

En la figura 1.2 se presenta una encuesta realizada a doce estudiantes de tercer año de Facultad de Odontología.

Sexo	F/M	<input type="checkbox"/>
Edad (años)		
Nivel de instrucción alcanzado por los padres		
i	<input type="checkbox"/>	Ninguno
ii	<input type="checkbox"/>	Primaria
iii	<input type="checkbox"/>	Secundaria
iv	<input type="checkbox"/>	Terciario no universitario
v	<input type="checkbox"/>	Terciario universitario
Estado civil		
i	<input type="checkbox"/>	Soltero/a
ii	<input type="checkbox"/>	Casado/a
iii	<input type="checkbox"/>	Unión libre
iv	<input type="checkbox"/>	Viudo/a
v	<input type="checkbox"/>	Separado/a
Cantidad de cursos aprobados		<input type="checkbox"/>

Figura 1.2 Encuesta realizada a estudiantes de Facultad de Odontología.

Identifique cada una de las variables, clasifíquelas e indique su escala de medición.

Ejercicio 1.3

- 1) Explique la diferencia entre variables cuantitativas y cualitativas.
- 2) Explique la diferencia entre una variable cuantitativa discreta y una variable cuantitativa continua.

Distribución de frecuencias

Introducción

Debido a la variabilidad y complejidad inherente a los problemas biomédicos, la distribución de frecuencias es una forma de estudio muy común en las ciencias de la salud.

Este método utiliza a una tabla donde se presentan los valores (para variables cualitativas: *clases*) de una variable junto con la cantidad de veces que estos se repiten. Este último dato, que indica qué tan frecuente es un valor, es llamado *frecuencia*. A través de esta tabla es posible resumir y describir las observaciones de una variable.

Distribución de frecuencias (DF)

En el caso de una variable nominal u ordinal, la distribución de frecuencias (DF) se construye en una tabla donde cada fila presenta una de las clases de la variable y en la celda siguiente el número de veces que se observó en la muestra.

De manera general se podría decir que, considerando una variable con k clases (C_1, C_2, \dots, C_k) mutuamente excluyentes y exhaustivas, a cada una le corresponde una *frecuencia absoluta*. A partir de estas frecuencias absolutas es posible generar las *frecuencias relativas*, que se calculan a través del cociente entre la frecuencia para cada clase y el total de observaciones. La DF de la variable considerada es entonces la colección de parejas $\{(C_1, f_1), (C_2, f_2), \dots, (C_k, f_k)\}$.

Tipos de frecuencia

Frecuencia absoluta

La frecuencia absoluta es el número de veces que aparece un determinado valor o clase en un estudio estadístico. Se representa por f_a . La suma de las frecuencias absolutas es igual al número total de datos, identificado como n .

Frecuencia relativa

La frecuencia relativa es el cociente entre la frecuencia absoluta de un determinado valor y el número total de datos. Este cociente es una proporción, siempre es un número entre 0 y 1 (y la suma de las frecuencias relativas es igual a 1). También es posible multiplicar estas proporciones por 100 para expresarlas como porcentajes (en tal caso, la suma de las frecuencias relativas es igual a 100 %).

Ejemplo 1

Considere un grupo de treinta estudiantes a los cuales se les realiza una encuesta donde una de las variables registrada es su género. La siguiente tabla presenta la DF de la variable «género».

Tabla 2.1 Distribución de frecuencias absolutas y relativas según género en una encuesta.

Género	frec. abs.	frec. rel.
Masculino	13	43,3%
Femenino	17	56,7%
Total	30	100%

A partir de esta DF es posible decir que el 57% de los alumnos son mujeres y que el 43% son hombres.

Estructura de la tabla

Al momento de construir la tabla, deben tenerse en cuenta tres elementos: título, estructura y contenido.

Título. Debe expresar en forma sintética y clara lo que se presenta en la tabla. De existir más de una, es una buena práctica numerarlas consecutivamente.

Estructura de la tabla. Estará formada por filas (renglones horizontales) y columnas (disposición vertical).

Contenido. La primera columna indica cada clase de la variable bajo estudio. La segunda columna presenta las frecuencias absolutas (número de casos) de cada clase. A partir de estas, se calculan los valores de la tercera columna, las frecuencias relativas. Estas últimas pueden presentarse como un porcentaje o como una proporción. En cuanto a las filas, la primera fila corresponde al encabezamiento de las columnas. Las filas siguientes presentan cada clase, su frecuencia absoluta y su frecuencia relativa. En la última fila se indica el número total de observaciones.

DF para variables continuas

Hasta aquí se presentó la construcción de las tablas de frecuencia correspondientes a las variables nominales u ordinales. Esta metodología también es adecuada para variables cuantitativas discretas, siempre y cuando la muestra presente pocos valores diferentes.

Para variables cuantitativas continuas también es posible construir DF, pero se deben crear *intervalos de clase* (IDC), que ocupan el lugar de las clases de una variable ordinal. La creación de estos intervalos solo es posible luego de determinar el número de clases y su amplitud. Estas decisiones son importantes, si se usan pocas clases se puede perder mucha información de los datos, si se usan demasiadas clases las frecuencias pueden resultar muy pequeñas (lo que las vuelve inestables).

Existen varios métodos que pueden ser empleados para confeccionar los IDC, y se puede emplear el más sencillo. Se presenta a continuación uno de estos procedimientos.

1) Ordenar los n datos de menor a mayor. Se identifica como X_{\min} al menor de los datos y como X_{\max} al mayor.

2) Calcular el *rango* de los datos. El rango es la diferencia entre la observación más grande y la más pequeña

$$r = X_{\max} - X_{\min}$$

3) Aproximar el número de clases (denotado por k) empleando la fórmula

$$k = \sqrt{n}$$

Cuando la raíz cuadrada del número de datos no resulte en un número entero, se deberá redondear el resultado al entero más cercano.

4) Calcular la amplitud (o longitud) de los IDC (L) empleando la fórmula

$$L = \frac{r}{k}$$

Ya se determinaron el número de IDC a emplear y su amplitud.

Por último, se determinarán los límites de cada uno de los IDC.

5) Definir X_{\min} como límite inferior del primer IDC y $X_{\min} + L$ como límite superior. De esta manera, el primer IDC se conforma por los valores $[X_{\min}; X_{\min} + L)$.

El uso de los paréntesis recto y curvo responde a que el intervalo incluye al valor más pequeño, pero no al valor más grande.

6) Calcular los límites de los demás IDC repitiendo el procedimiento. El segundo IDC tendrá como límite inferior al límite superior del intervalo anterior, o sea, del primero. De esta manera el límite inferior del segundo IDC será $X_{\min} + L$. Para determinar el límite superior simplemente se le suma L al inferior, en este caso $X_{\min} + 2 \times L$. A diferencia de los demás IDC, el último intervalo contiene a ambos extremos: $[X_{\min} + (k - 1) \times L; X_{\min} + k \times L]$.

7) Disponiendo de los IDC, se confecciona la DF. Para obtener la frecuencia absoluta de un IDC se cuenta la cantidad de observaciones de la variable continua que caen dentro de dicho IDC.

Ejemplo 2

Retomando al grupo de treinta estudiantes que contestaron una encuesta (ejemplo 1), suponga que se registró su peso y estatura. A continuación, se presenta un cuadro con las treinta estaturas observadas (en metros).

Cuadro 2.1 Datos de estatura (m) de estudiantes, de una encuesta de la Facultad de Odontología

1,71	1,76	1,71	1,68	1,54
1,75	1,57	1,52	1,52	1,71
1,63	1,80	1,75	1,65	1,56
1,57	1,71	1,62	1,65	1,73
1,62	1,55	1,64	1,67	1,66
1,75	1,65	1,59	1,68	1,50

Al aplicar el procedimiento descrito anteriormente se observa que los datos de los extremos son 1,50 y 1,80, lo cual resulta en un rango r de 0,30. Por otro lado, teniendo en cuenta que $n = 30$, el número de clases k puede aproximarse a 5 (5,48).

Luego, disponiendo de los valores de r y k (que son 0,30 y 5) es posible calcular la amplitud de las clases mediante su cociente $L = 0,06$. La construcción del primer IDC sería $[1,50; 1,50 + 0,06)$ resultando así en $[1,50; 1,56)$. El segundo IDC va de 1,56 a $1,56 + 0,06$. De esta manera, sus límites son $[1,56; 1,62)$. La construcción de los demás IDC continúa de la misma manera hasta que el último es $[1,74; 1,80]$. Nótese que el paréntesis recto indica la inclusión del valor inmediato, mientras que el curvo indica que el valor inmediato está excluido.

Por último, se confecciona la DF, asignando a las frecuencias absolutas el número de observaciones que pertenecen a cada IDC.

Tabla 2.2 Distribución de frecuencias de la estatura de 30 estudiantes encuestados

IDC	frec. abs.	frec. rel.
[150 ; 156)	5	16,7%
[156 ; 162)	4	13,3%
[162 ; 168)	9	30,0%
[168 ; 174)	7	23,3%
[174 ; 180]	5	16,7%
Total	30	100%

A partir de esta tabla es posible observar que la mayoría relativa de los alumnos (más concretamente el 30,0%) tiene una estatura de entre 1,62 y 1,68 m.

Representación gráfica

La DF es una forma efectiva de organizar y presentar los datos. Los gráficos son capaces de transmitir la misma información, pero de una manera visual. De este modo pueden ayudar a hacer evidentes tendencias y patrones en los datos. En esta sección se presentan gráficos que posibilitan la representación de una DF.

La selección del tipo de representación depende en primera instancia del tipo de variable o variables bajo estudio y de su nivel de medición. Es imprescindible identificar adecuadamente la clasificación de la variable para elegir el gráfico apropiado.

Por lo general, al trabajar con variables o datos cualitativos (nominales u ordinales) se emplean gráficos de barras o gráficos sectoriales («gráficos de torta»). Cuando se trabaja con variables cuantitativas, el gráfico a usar depende de si la variable cuantitativa es discreta o continua. Si se trata de una variable cuantitativa discreta, se utiliza el gráfico de barras, mientras que cuando la variable cuantitativa es continua se emplea un histograma.

Gráfico de barras

Se trata de un gráfico donde a través de barras se representan las frecuencias (ya sean absolutas o relativas) correspondientes a las clases de una variable cualitativa. En el eje horizontal se disponen las clases utilizadas, mientras que en el eje vertical se presentan valores de referencia sobre los que representar las frecuencias. Una vez dispuestos los ejes, las barras tienen su base en el eje horizontal y su altura se determina por la frecuencia correspondiente a cada clase. Al tratarse de una variable cualitativa, las barras deben estar separadas; tanto el ancho de la separación como el de las barras es indiferente, siempre y cuando no varíe.

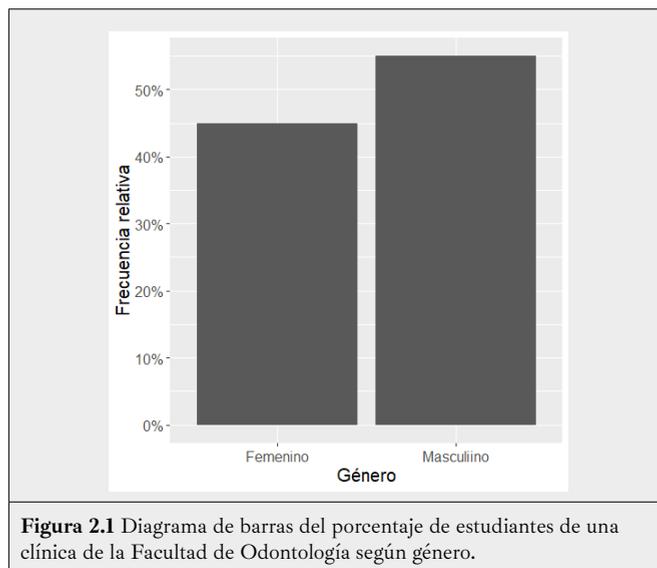
Si la variable a representar fuese ordinal, es una buena práctica que la ubicación de las clases en el eje horizontal respete el ordenamiento de las clases.

Ejemplo 3

En una clínica de la Facultad de Odontología se lleva a cabo un relevamiento para conocer el perfil de los estudiantes que la integran. Una de las preguntas refiere al género de los estudiantes. A continuación, se presenta la DF (tabla 2.3) así como el gráfico de barras (figura 2.1) que representa las frecuencias relativas.

Tabla 2.3 Distribución de frecuencias de estudiantes de una clínica de la Facultad de Odontología según género

Género	f. abs.	f. rel.
Masculino	11	55%
Femenino	9	45%
Total	20	100%



Queda como ejercicio la confección del gráfico de barras para representar las frecuencias absolutas. Podrá observarse cómo ambos gráficos se diferenciarán en los valores del eje vertical, pero la relación de altura entre las barras será la misma.

Histograma

Un histograma es similar en apariencia y construcción a un gráfico de barras, excepto que se utiliza para representar variables cuantitativas continuas, sea su nivel de medición de intervalo o de razón. Se trata de un caso especial de un gráfico de barras, su construcción se basa en los datos de una DF de variables cuantitativas. Algunos detalles a tener en cuenta a la hora de construir un histograma:

- Debido a que la variable representada es cuantitativa continua, las barras se presentan juntas.
- En caso de que las clases (intervalos de clase) tengan la misma amplitud, las barras tendrán el mismo ancho. Si los intervalos de clase tuviesen amplitudes diferentes, el ancho de las barras deberá ser proporcional a la amplitud de cada intervalo.

En general, los histogramas son sensibles a las elecciones que se realizan al construir la DF a representar; el número de IDC y su amplitud pueden impactar en la presentación del histograma.

Ejemplo 4

Continuando con el ejemplo anterior, el relevamiento de los estudiantes de la clínica de Facultad de Odontología también registró la distancia (medida en kilómetros) desde la casa de cada uno hasta la facultad. A continuación se presenta un cuadro con los datos.

Cuadro 2.2 Datos de la distancia (km) del hogar de cada estudiante hasta la Facultad de Odontología

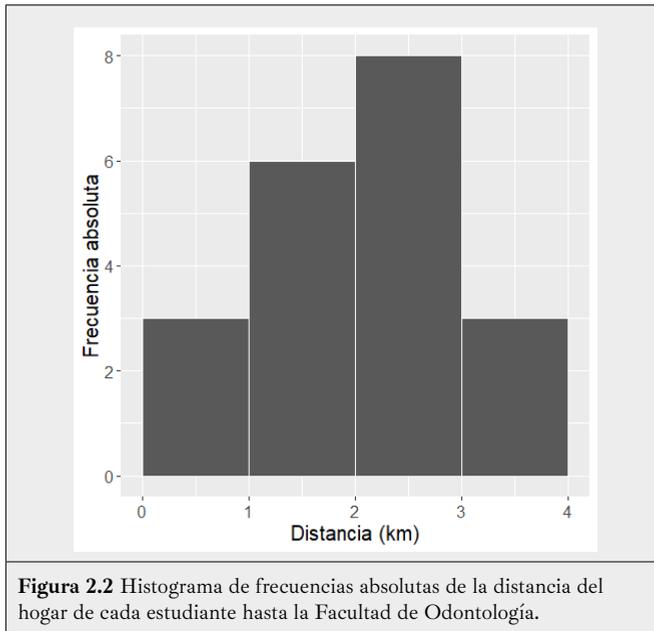
id	distancia (km)						
1	1,6	6	2,1	11	1,8	16	3,7
2	0,4	7	2,0	12	1,3	17	2,5
3	3,5	8	1,7	13	2,5	18	2,5
4	2,2	9	0,4	14	1,5	19	1,9
5	3,5	10	2,1	15	0,3	20	2,4

A partir de estos datos, considerando cuatro IDC es posible construir la siguiente distribución de frecuencias (tabla 2.4).

Tabla 2.4 Distribución de frecuencias de la distancia (km) del hogar de cada estudiante hasta la facultad

Distancia (km)	frec. abs.	frec. rel.
[0 ; 1)	3	15%
[1 ; 2)	6	30%
[2 ; 3)	8	40%
[3 ; 4)	3	15%
Total	20	100%

A partir de esta tabla es posible construir el histograma (figura 2.2). Nótese cómo los cuatro IDC poseen la misma amplitud, por lo tanto, las barras tendrán el mismo ancho.



Ejercicios

Ejercicio 2.1

En el transcurso de un mes concurrieron 450 personas a una clínica dental. Doscientas de ellas asistieron una vez, ciento cincuenta asistieron dos veces y el resto asistió tres veces o más.

Construya la distribución de frecuencias de la variable «número de veces que el paciente concurre a la clínica».

Ejercicio 2.2

Once estudiantes de un curso de estadística obtuvieron las siguientes calificaciones finales: 71, 79, 81, 73, 97, 68, 61, 77, 88, 65, 87. Utilizando el procedimiento expuesto anteriormente, construya la distribución de frecuencias de las notas.

Ejercicio 2.3

Vuelva sobre los datos de los dos ejercicios anteriores y represente gráficamente las distribuciones de frecuencias calculadas en esos ejercicios.

Medidas de resumen

Introducción

En el capítulo anterior se presentó la distribución de frecuencias y diferentes gráficos para resumir los datos. En este capítulo, al trabajar con variables cuantitativas, se verá cómo representar el conjunto de datos bajo estudio a través de un solo número. Este número, que llamaremos *medida de resumen*, implica distintos procedimientos de cálculo y de interpretación.

La pluralidad de medidas de resumen se debe a que no es posible sintetizar en un solo número los diferentes valores observados. Las diferentes medidas de resumen son importantes y se complementan a la hora de resumir la información. Es importante conocer la mecánica de su cálculo para poder entenderlas, y más importante aún es poder explicar qué nos están mostrando con respecto a los datos bajo estudio.

Las medidas estadísticas que describen características tales como el centro de los datos, se llaman *medidas de tendencia central o de posición*. El término *tendencia central* refiere al valor alrededor del cual la distribución tiende a centrarse.

Las medidas de tendencia central más comúnmente utilizadas son: media aritmética (o promedio), mínimo, máximo, mediana, moda y cuartiles.

Sin embargo, las medidas de tendencia central no bastan para describir los datos de una manera adecuada, esto se debe a que en todo resumen se pierde información. También es necesario conocer cómo se distribuyen los datos, qué tan concentrados o dispersos están.

Las medidas que determinan el nivel de dispersión se llaman *medidas de dispersión o de variabilidad*. Las medidas típicas de dispersión son: rango, rango intercuartílico, varianza y desviación estándar.

Medidas de posición

Media

La medida de tendencia central más utilizada es la *media aritmética* (también conocida como *promedio*). Consiste en la suma de todos los valores (X_i mediciones), dividida por el número de observaciones (n).

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

Donde \bar{X} es el símbolo para representar la media del conjunto de datos. La letra X representa la variable, los n valores de los datos de la muestra son representados por X_1, X_2, \dots, X_n .

La suma de las n mediciones se representa con la letra sigma mayúscula: Σ , así se representa de manera compacta la suma de n valores, desde el primero (X_1) hasta el último (X_n).

Ejemplo 5

Supongamos que un cirujano examinó a cinco pacientes con inflamación gingival. Midió el diámetro de la inflamación de cada uno y obtuvo las siguientes medidas: 5,5; 7,0; 4,5; 8,5 y 10 mm. La media de los datos se puede calcular utilizando la fórmula anterior.

$$\bar{X} = \frac{\sum_{i=1}^{i=5} X_i}{5} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5} = \frac{5,5 + 7,0 + 4,5 + 8,5 + 10}{5} = 7,1$$

Pese a que el resultado de la fórmula es un valor que no fue observado, este se interpreta como el diámetro de inflamación que tendrían los pacientes, si todos presentaran el mismo valor.

La desventaja de esta medida de resumen es que es sensible a valores extremos. Es decir, la presencia de unos pocos valores muy grandes (o muy pequeños) pueden afectar el resultado de esta medida, aumentándola o disminuyéndola considerablemente, haciendo que el resultado no represente ni a la mayoría de las observaciones ni a los datos extremos.

Como medida de resumen, la media debe ser complementada con el mínimo y el máximo.

Mínimo y máximo

Volviendo al ejemplo de los pacientes a los que se les mide el diámetro de la inflamación gingival, la media de 7,1 se da en un contexto donde hay valores por debajo y por encima de dicho valor. Interesa entonces poner atención al valor más pequeño (que llamaremos *mínimo*) y al valor más grande (que llamaremos *máximo*).

¿Por qué son importantes estos valores? Si en lugar de contar con los valores 4,5 y 10, hubiésemos observado los valores 2,5 y 12,5, la media seguiría siendo 7,1, pero la situación sería muy diferente.

Mediana

La mediana (Md), se define como el valor de la variable que divide la distribución de datos en dos partes iguales. Al usarla en conjunto con el valor mínimo permite decir entre qué valores está el primer 50% de nuestros datos, en conjunto con el máximo permite decir entre qué valores se encuentra el otro 50% de nuestra distribución (Kim & Dailey, 2008a).

Para calcular la mediana, lo primero que se debe hacer es ordenar los datos en orden creciente, es decir, de menor a mayor. Luego, dependiendo del número de observaciones, se emplea una de las dos fórmulas que se presentan a continuación (Verga & Cáceres, 2004).

$$Md = X_{\frac{n+1}{2}} \quad \text{si } n \text{ es impar.}$$

$$Md = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} \quad \text{si } n \text{ es par.}$$

La primera de las dos fórmulas indica que cuando se dispone de una cantidad impar de observaciones, la mediana es el valor central. No obstante, no siempre es sencillo determinar cuál es el valor central.

En el caso de que se cuente con un número par de observaciones, el valor de la mediana se calcula como el promedio de los dos valores centrales de la serie de observaciones ordenadas.

Vale la pena agregar que, a diferencia de la media, la mediana es una medida de resumen robusta a datos extremos. Es decir, la mediana posee un comportamiento más estable que la media ante la aparición de unos pocos valores extremos.

Ejemplo 6

Para ilustrar el concepto de *valor central* y el cálculo de la mediana, se retoma el ejemplo de las cinco mediciones del diámetro de la inflamación gingival. Los valores observados eran 5,5; 7,0; 4,5; 8,5 y 10.

Según lo expuesto anteriormente, el primer paso consiste en ordenar la serie de datos de mayor a menor. Las observaciones ordenadas son 4,5; 5,5; 7,0; 8,5 y 10. El siguiente paso consiste en determinar si el tamaño de la muestra (n) es par o impar, en este caso se dispone de cinco observaciones, por lo que la fórmula a utilizar es la que corresponde al caso impar. La fórmula dice que la mediana es el dato central. En una serie de cinco datos, es sencillo observar que el dato central es el tercero, ya que es el dato que deja igual número de observaciones hacia su izquierda y hacia su derecha. De este modo, la mediana es 7,0.

Pero si fuesen 1837 observaciones, ¿cuál sería el dato central? En este caso, donde $n = 1837$, la determinación del dato central se lleva a cabo sumando una unidad al tamaño muestral ($n + 1$) y dividiendo el resultado de dicha suma entre dos. De esta manera, el dato central sería el dato en la posición 919.

Por último, supongamos que ingresa un nuevo paciente con inflamación gingival al consultorio donde se registraron las cinco mediciones. Su diámetro de inflamación resulta ser 8,2. Al calcular la mediana de los seis datos se deberían seguir los mismos pasos.

En primer lugar, se deben ordenar, obteniendo la siguiente serie: 4,5; 5,5; 7,0; 8,2; 8,5 y 10. Luego, al disponer de un n par, la fórmula a utilizar es la segunda, la que dice que la mediana es el promedio de los dos datos centrales (como se trata de pocas observaciones, es fácil ver que los datos centrales son el tercero y el cuarto, ya que hay igual número de observaciones a la izquierda del tercer dato que a la derecha del cuarto). Finalmente, se deben promediar el tercer dato (que es 7,0) y el cuarto (que es 8,2), lo que resulta en el valor 7,6.

En caso de que el número de observaciones fuese par y más elevado, digamos 3472, la determinación de los datos centrales se llevaría a cabo de la siguiente manera: el menor de los datos centrales sería aquel que resulta de dividir entre dos al tamaño muestral ($n/2$), esto nos lleva a que el menor de los datos centrales sería el que se encuentra en el lugar $3472/2 = 1736$, a su vez, con esto hallamos el mayor de los datos centrales, que sería el siguiente, el 1737.

Moda

Otra medida utilizada para describir el centro de un conjunto de datos es la moda, que se define como el valor que ocurre con mayor frecuencia (Kim & Dailey, 2008a).

La moda es la única medida de tendencia central que puede ser utilizada tanto para variables cualitativas como cuantitativas. Se obtiene construyendo la distribución de frecuencias (ver capítulo 2), el valor que aparece con mayor frecuencia (absoluta o relativa) es la moda.

Algunos autores destacan como ventaja de la moda su simplicidad, ya que no es necesario realizar cálculos para obtenerla.

Una desventaja es que un conjunto de datos puede presentar varias modas, siendo bimodal, trimodal o multimodal, dependiendo del número de modas que presente. Otra desventaja es que puede no tener moda, si ningún valor ocurre más de una vez (Kim & Dailey, 2008a).

Cuando las distribuciones son multimodales se habla de que la población puede estar conformada por varios grupos homogéneos internamente y heterogéneos entre uno y otro (Verga & Cáceres, 2004).

Los siguientes ejemplos han sido extraídos del libro *Biostatistics for Oral Healthcare* (Kim & Dailey, 2008a).

Ejemplo 7

Un grupo de pacientes fue examinado en un ensayo clínico para determinar si un aumento en la ingesta de golosinas en un corto período causó un aumento en el nivel de *Streptococcus mutans* en la dentición general. El índice gingival (GI) de los pacientes fue una de las variables que observaron los odontólogos.

El cuadro 3.1 presenta el número identificador y el valor del GI de cada uno de los dieciocho participantes del ensayo.

Cuadro 3.1 Datos del índice gingival (GI) de pacientes de un ensayo clínico

Subject No.	GI	Subject No.	GI	Subject No.	GI
1	2	7	0	13	0
2	1	8	0	14	2
3	0	9	2	15	1
4	1	10	1	16	1
5	3	11	1	17	3
6	1	12	2	18	1

A partir de los datos del GI, se construyó la tabla 3.1.

Tabla 3.1 Distribución de frecuencia del índice gingival (GI) de pacientes de un ensayo clínico.

GI	frec. abs.	frec. rel.
0	4	22,2%
1	8	44,4%
2	4	22,2%
3	2	11,1%
Total	18	100%

Es fácil ver que la moda es 1, ya que es el valor de GI registrado en el mayor número de participantes.

Ejemplo 8

El temor a recibir atención odontológica ha demostrado ser uno de los principales factores que aleja a las personas de la consulta. Muchas personas con problemas dentales prefieren el dolor y la incomodidad oral severa a buscar atención.

Para evaluar el efecto del temor a la atención odontológica, se registró la frecuencia del pulso de los pacientes mientras estaban sentados en el sillón odontológico. Los valores observados fueron 92, 88, 90, 94, 92, 86, 95, 94, 93, 92, 82, 90, 96, 94, 92, 89, 94. Hay cuatro pacientes con una frecuencia de pulso de 92 y otros cuatro con 94, por lo que 92 y 94 son los valores que ocurren con mayor frecuencia. Por lo tanto, hay dos modas, una es 92 y la otra 94.

Cuartiles

Anteriormente se presentó la mediana como medida de tendencia central que divide las observaciones en dos conjuntos con igual número de datos. De modo similar, existen medidas de resumen capaces de particionar la muestra en más de dos conjuntos con igual número de datos. En el caso de que se consideren cuatro conjuntos iguales, las medidas que los generan son los *cuartiles*.

De la misma manera que es necesaria una mediana para particionar la muestra en dos mitades, se necesitan tres cuartiles para particionar la muestra en cuatro cuartos. En ese sentido, se hace necesario calcular el primer, segundo y tercer cuartil.

A la hora de definir y calcular estas medidas, al igual que con la mediana, es imprescindible ordenar los datos de la variable de menor a mayor.

- El primer cuartil es el valor que supera al 25% de las observaciones y que es superado por el 75% restante. Se denomina Q_1 .
- El segundo cuartil es el valor que supera al 50% de las observaciones y que es superado por el 50% restante. Es decir que es igual a la mediana. Se denomina Q_2 .
- El tercer cuartil es el valor que supera al 75% de las observaciones y que es superado por el 25% restante. Se denomina Q_3 .

En este manual se presenta una metodología de cálculo que pretende ser clara y proporcionar un procedimiento sencillo. Existen varias otras fórmulas, el lector es libre de indagar sobre ellas.

Antes de presentar la metodología de cálculo se recuerdan los pasos correspondientes al cálculo de la mediana.

1. *Ordenar* las observaciones de menor a mayor.
2. *Calcular* la posición del dato central.
3. Determinar el valor que *ocupa* la posición del dato central.

El cálculo de los cuartiles sigue estos mismos pasos, con algunas modificaciones en el segundo y en el tercero, que se presentan a continuación.

Paso 2

Posición del cuartil

Al momento de determinar la posición que ocupa el cuartil entre los datos ordenados es importante contar con el tamaño de la muestra (n) y con la proporción de observaciones que serían menores al cuartil que se desea calcular. Si se está calculando la posición de Q_1 , esta supera al 25% de las observaciones, Q_2 supera al 50% y Q_3 al 75%. La ecuación que sigue utiliza estos datos para calcular la posición de los cuartiles.

$$posición = (n + 1) \times p$$

Donde p representa los porcentajes descritos anteriormente.

A modo de ejemplo, considere una situación donde se dispone de 20 observaciones correspondientes al tiempo hasta la falla de un implante. Si se quisiera calcular los tres cuartiles, se debería utilizar la ecuación anterior sustituyendo n por 20 y considerar los tres valores de p , uno para cada cuartil.

$$\text{posición } Q_1 = (20 + 1) \times 0,25 = 21 \times 0,25 = 5,25$$

$$\text{posición } Q_2 = (20 + 1) \times 0,50 = 21 \times 0,50 = 11,5$$

$$\text{posición } Q_3 = (20 + 1) \times 0,75 = 21 \times 0,75 = 15,75$$

Antes de continuar con el tercer paso, es importante puntualizar que *no se debe redondear el resultado de estos cálculos*, a la hora de determinar los cuartiles las cifras a la derecha de la coma son tan importantes como las de la izquierda de la coma. En este sentido, y para facilitar la ecuación del tercer paso, las cifras a la izquierda de la coma se denominan *parte entera de la posición* (E), mientras que las cifras de la derecha se denominan *parte decimal de la posición* ($0,d$). A modo de ejemplo, considere la posición del primer cuartil: la parte entera es 5 y la parte decimal 0,25.

Paso 3

Valor que ocupa el cuartil

En última instancia, disponiendo de la posición que ocupa cada cuartil en la serie de observaciones, se procede a calcular el valor de los cuartiles, para ello se introduce la siguiente ecuación.

$$\text{cuartil} = X_E + 0,d \times (X_{E+1} - X_E)$$

Donde X_E representa el valor de la observación que ocupa la posición indicada por la parte entera de la posición (E). El valor $0,d$ indica la parte decimal de la posición y, por último, X_{E+1} representa el valor de la observación que ocupa la posición $E + 1$, es decir, la que sigue inmediatamente a X_E .

Ejemplo 9

Un equipo de investigadores de la Facultad de Odontología quiso poner a prueba la durabilidad de un material empleado en implantes dentales. A tales efectos, empleando registros clínicos, estudiaron el tiempo hasta la falla de una muestra 20 implantes.

Cuadro 3.2 Datos del tiempo de durabilidad de implantes dentales, de registros clínicos

15,9	26,9	29,7	29,9
24,7	24	20,7	29,3
20,1	18,8	16,1	26,3
22,1	24,7	16,7	22,7
29,1	19,7	29,1	20,6

Debido a que se trata de 20 observaciones, las posiciones de los cuartiles son las calculadas en la descripción del segundo paso: 5,25; 10,5 y 15,75.

Antes de proceder al tercer paso, los datos se deben ordenar de menor a mayor. A continuación se presentan los datos ordenados.

Cuadro 3.3 Datos ordenados del tiempo de durabilidad de implantes dentales, de registros clínicos

15,9	20,1	24	29,1
16,1	20,6	24,7	29,1
16,7	20,7	24,7	29,3
18,8	22,1	26,3	29,7
19,7	22,7	26,9	29,9

A la hora de calcular Q_1 , los valores de E y $0,d$ son 5 y 0,25, por lo que la ecuación que determina el valor final de Q_1 es:

$$Q_1 = X_5 + 0,25 \times (X_6 - X_5)$$

Los valores correspondientes a X_5 y X_6 son 19,7 y 20,1 respectivamente. Por lo que la ecuación adopta los siguientes valores:

$$Q_1 = 19,7 + 0,25 \times (20,1 - 19,7) = 19,8$$

Recuerde que al realizar los cálculos de la ecuación, primero se realiza la resta (20,1 - 19,7), su resultado es multiplicado por 0,25 y por último se adiciona el valor 19,7.

Empleando la ecuación, los valores correspondientes a Q_2 y Q_3 son:

$$Q_2 = X_{10} + 0,50 \times (X_{11} - X_{10}) = 22,7 + 0,50 \times (24,0 - 22,7) = 23,4$$

$$Q_3 = X_{15} + 0,75 \times (X_{16} - X_{15}) = 26,9 + 0,75 \times (29,1 - 26,9) = 28,6$$

Finalmente, es posible decir que los valores 19,8; 23,4 y 28,6 particionan los datos en cuatro conjuntos con igual número de observaciones. A la hora de interpretar los resultados, y teniendo en cuenta el contexto del problema, se puede decir que el 75% de los implantes duró más de 19,8 años, que la mitad superó los

23,4 años y que solo el 25 % superó los 28,6 años. También es posible afirmar que el 50 % central de los datos se encuentra entre 19,8 y 28,6 años.

A modo de resumen, a continuación, se presenta un ejemplo donde se plantea el uso de todas las medidas de posición vistas hasta aquí.

Ejemplo 10

Se conoce que los antidepresivos tricíclicos y antihipertensivos pueden influir en el flujo salival. La disminución del flujo salival disminuye la limpieza natural de la cavidad oral, aumentando así la incidencia de enfermedad periodontal y caries. Para explorar qué efectos podría tener el nivel de ácido ascórbico (NAA) en saliva en la enfermedad periodontal y caries, un periodoncista tomó 30 muestras salivales de sus pacientes y midió su NAA.

Cuadro 3.4 Datos del nivel en saliva de ácido ascórbico (NAA) de un conjunto de pacientes

id	NAA (mg%)	id	NAA (mg%)	id	NAA (mg%)
1	0,08	11	0,17	21	0,26
2	0,1	12	0,19	22	0,28
3	0,1	13	0,21	23	0,28
4	0,1	14	0,21	24	0,29
5	0,11	15	0,22	25	0,31
6	0,12	16	0,22	26	0,31
7	0,13	17	0,23	27	0,33
8	0,13	18	0,24	28	0,33
9	0,14	19	0,25	29	0,36
10	0,15	20	0,26	30	0,37

El cálculo muestra que la media aritmética del NAA es 0,216 (resultado de sumar todos los valores y dividir el resultado entre 30).

Para el cálculo de la mediana se aprovecha que los valores ya están ordenados de menor a mayor y, teniendo en cuenta que se dispone de un número par de datos, se calcula la mediana como el promedio de los dos datos centrales. En este caso, el menor de los datos centrales es el número 15 (resultado de dividir el tamaño muestral entre dos) y el mayor es el 16, los cuales corresponden a los valores 0,22 y 0,22. El promedio de estos dos, y por lo tanto la mediana, es 0,22.

El NAA de 0,10 se observa con mayor frecuencia (tres veces), es por tanto, la moda del conjunto de datos.

La media y la mediana se encuentran en el centro de la distribución de los datos, pero la moda se encuentra lejos, a la izquierda del centro. En el este caso,

la media y la mediana son mejores medidas de ubicación que la moda. De hecho, en este ejemplo, la moda resulta engañosa como medida de tendencia central.

Por último, se calculan el primer y tercer cuartil del NAA. Queda como ejercicio para el lector determinar por qué no se presenta el cálculo del segundo cuartil. Debido a que se cuenta con 30 observaciones, las posiciones de Q_1 y Q_3 son:

$$\text{posición } Q_1 = (30 + 1) \times 0,25 = 31 \times 0,25 = 7,75$$

$$\text{posición } Q_3 = (30 + 1) \times 0,75 = 31 \times 0,75 = 23,25$$

Disponiendo de estos valores (y descomponiéndolos en E y $0,d$) se procede a utilizar la segunda ecuación.

$$Q_1 = X_7 + 0,75 \times (X_8 - X_7) = 0,13 + 0,75 \times (0,13 - 0,13) = 0,13$$

$$Q_3 = X_{23} + 0,25 \times (X_{24} - X_{23}) = 0,28 + 0,25 \times (0,29 - 0,28) = 0,29$$

Observe cómo con estos datos se puede ver que el 50 % central de los valores de NAA se encuentra entre 0,13 y 0,29.

Medidas de dispersión

Rango

La medida de variabilidad más simple es el rango, se define como la distancia entre la observación más grande y la más pequeña (Kim & Dailey, 2008a).

Usamos la letra X para simbolizar la variable bajo estudio, denominando cada una de las n observaciones como X_1, X_2, \dots, X_n . Si se le llama X_{\max} al mayor valor entre esas observaciones y X_{\min} al menor valor, entonces el rango (r) se define como:

$$r = X_{\max} - X_{\min}$$

El rango es fácil de calcular y es adecuado para variables cuantitativas, sean de intervalo o de razón. Representa una distancia (calculada a través de una resta), por lo tanto no debe ser aplicado a variables cualitativas.

Tiene las desventajas de ser inestable e insuficiente, ya que al depender solamente de los datos de los extremos (X_{\min} y X_{\max}) ignora los $n - 2$ valores comprendidos entre ellos. Por tanto, puede no reflejar la dispersión de los valores entre las dos observaciones extremas.

Ejemplo 11

A continuación se introducen tres casos que tienen observaciones extremas comunes, por lo que presentan el mismo rango (ya que este es independiente de los valores intermedios). El rango de los tres casos es 500, sin embargo, los valores se distribuyen de manera diferente:

- Caso 1: 30, 150, 250, 280, 310, 410, 530.
- Caso 2: 30, 30, 30, 30, 30, 30, 530.
- Caso 3: 30, 530, 530, 530, 530, 530, 530.

Rango intercuartílico

El *rango intercuartílico* (RI) es una medida de variabilidad que intenta compensar los problemas del rango. El RI se define como la distancia entre el primer y el tercer cuartil.

Al igual que el rango, es adecuado para variables cuantitativas, ya sean de intervalo o de razón. Representa una distancia calculada a través de una resta, por lo que no debe ser aplicado a variables cualitativas.

Aunque solo depende de dos observaciones, es un poco más estable que el rango. Es fácil de calcular a partir de los cuartiles (Verga & Cáceres, 2004):

$$RI = Q_3 - Q_1$$

Varianza muestral (s^2) y desviación estándar (s)

Para comprender el nivel de dispersión entre las observaciones de la muestra es necesario medir su variabilidad. Las medidas de variabilidad más utilizadas son la varianza muestral (s^2) y la desviación estándar (s) debido a que presentan características estadísticas que las hacen más solventes que el rango o el rango intercuartílico. A diferencia del r y el RI, estas dos medidas de dispersión usan todas las observaciones de un conjunto de datos (Kim & Dailey, 2008a). Debido a que su cálculo involucra restas, sumas, potencias y divisiones, la varianza y la desviación estándar solo pueden ser utilizadas para variables cuantitativas. La varianza (s^2) se define como el promedio de las distancias al cuadrado entre las observaciones y su media (\bar{X}).

$$s^2 = \frac{\sum_{i=1}^{i=n} (X_i - \bar{X})^2}{n - 1}$$

Al tratarse de un promedio, posee las mismas características de la media aritmética. Tiene la ventaja de que utiliza toda la información, pero, al igual que la media, es muy sensible a datos extremos.

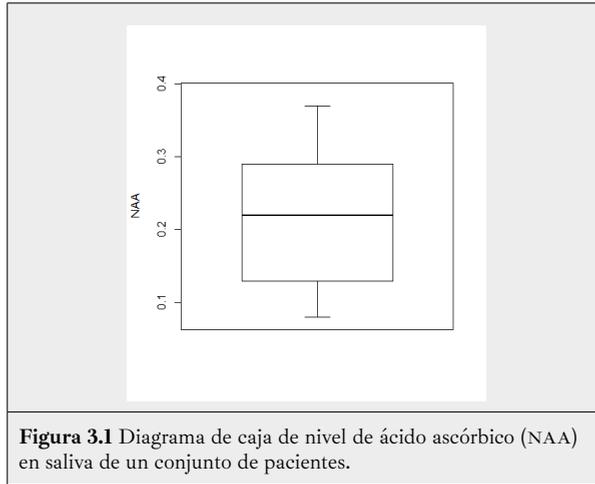
Debido a que el cálculo de la varianza involucra elevar diferencias al cuadrado, la unidad de medida del resultado final es el cuadrado de la unidad de medida de los datos originales. Considere, a modo de ejemplo, que si se contara con observaciones que registren el número de caries en un grupo de n personas, la varianza del número de caries estaría expresada en «caries al cuadrado», lo cual no es muy fácil de interpretar.

La desviación estándar (s) soluciona este defecto de la varianza, ya que es la raíz cuadrada de la varianza. Por lo tanto, tiene la misma unidad de medida que las observaciones originales.

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Representación gráfica

Si bien existen diversos tipos de representaciones para variables cuantitativas, este manual presenta el gráfico de caja (*box-plot*) debido a que permite obtener información visual sobre tendencia central, posición y variabilidad de los datos recabados. Otros gráficos que representan la distribución de valores de una variable cuantitativa son el histograma (ver capítulo 2), el polígono de frecuencias y el diagrama de tallo y hojas, entre otros. El *box-plot*, ideado por John Tukey, presenta la distribución de los datos de una variable empleando cinco valores, tres de los cuales son los cuartiles. Los otros dos valores son el límite superior (LS) y el límite inferior (LI), e intentan describir la variación de los datos. La siguiente figura presenta el gráfico de caja correspondiente al nivel ácido ascórbico (NAA) presentado en el ejemplo 10.



La caja representa la parte central de la distribución. Esta se construye a partir de los cuartiles (calculados previamente). La base de la caja corresponde al primer cuartil (en el ejemplo se calculó un valor de $Q_1 = 0,13$). La tapa superior de la caja corresponde al tercer cuartil (en este caso, $Q_3 = 0,29$).

La línea que divide la caja horizontalmente es la mediana (en este caso, $Q_2 = 0,22$). Conste que la mediana no tiene por qué ubicarse en la mitad de la caja. Debido a la definición de los cuartiles, es posible afirmar que la caja representa el 50% central de los datos, de los cuales hay un 25% entre el primer cuartil y la mediana y otro 25% entre la mediana y el tercer cuartil.

A continuación se presenta la metodología utilizada para el cálculo de los límites del gráfico.

1. Se calcula el rango intercuartílico (RI). La longitud de las líneas verticales del gráfico se calcula como $1,5 \times \text{RI}$.
2. El *valor candidato a límite superior* (LS_c) será el resultado de añadir una línea vertical (con la longitud $1,5 \times \text{RI}$) sobre el Q_3 . La ecuación resultante es:

$$LS_c = Q_3 + 1,5 \times \text{RI}$$

Una vez calculado este valor, se compara con el máximo valor observado en la muestra (X_{\max}). El valor final del LS será el mínimo de estos dos valores.

$$LS = \min(LS_c; X_{\max})$$

3. El *valor candidato a límite inferior* (LI_c) será el resultado de agregar una línea vertical (con longitud $1,5 \times RI$) hacia abajo desde Q_1 . La ecuación resultante es:

$$LI_c = Q_1 - 1,5 \times RI$$

Una vez calculado este valor, se compara con el mínimo de la muestra (X_{\min}). El valor final del LI será el máximo de estos dos valores.

$$LI = \max(LI_c; X_{\min})$$

4. En caso de que los valores finales de LS y LI sean iguales a los valores candidatos, aquellos valores que excedan los límites se marcarán en el gráfico con un punto y serán denominados *atípicos*.

Esta corrección que se realiza sobre los valores límites tiene por motivo evitar el cálculo de límites excesivamente grandes o biológicamente imposibles.

A partir de estas fórmulas fue posible calcular los valores de los límites para el ejemplo 10. En primer lugar, el $RI = 0,29 - 0,13 = 0,16$, por lo que la longitud de las líneas fue $1,5 \times 0,16 = 0,24$.

Disponiendo de este valor, los valores candidatos de cada límite fueron:

$$LS_c = 0,29 + 0,24 = 0,53$$

$$LI_c = 0,13 - 0,24 = -0,11$$

En último lugar, teniendo en cuenta que los valores más extremos de los NAA son 0,08 y 0,37, los valores finales de los límites resultaron:

$$LS = \min(LS_c; X_{\max}) = \min(0,53; 0,37) = 0,37$$

$$LI = \max(LI_c; X_{\min}) = \max(-0,11; 0,08) = 0,08$$

Ejemplo 12

Una empresa encargada de manufacturar enjuague bucal realizó un experimento en un grupo de 20 participantes para determinar el tiempo que los usuarios de su producto lo mantienen en la boca. Los valores observados (y ordenados) fueron 28, 30, 30, 33, 34, 35, 35, 35, 38, 38, 40, 40, 42, 43, 45, 45, 46, 50, 53 y 65.

El lector puede comprobar que los tres cuartiles correspondientes a estos datos son 34,25; 39; 45. Por lo tanto el rango intercuartil es 10,75 y los valores

candidatos a límites son 61,1 para el LS y 18,1 para el LI . Debido a que el LI_c es menor al valor mínimo, el valor final del LI es 28. Por otro lado, el LS_c no supera al máximo valor de la muestra, por lo tanto $LS = 61,1$ y el valor 65, que excede al LS , es clasificado como valor atípico y se marca en el gráfico con un punto.

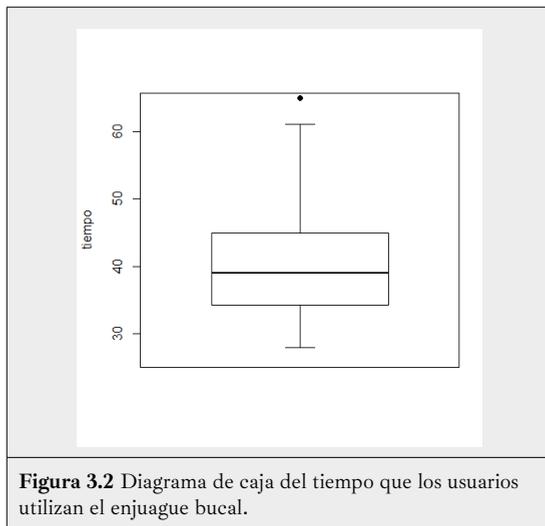


Figura 3.2 Diagrama de caja del tiempo que los usuarios utilizan el enjuague bucal.

Ejercicios

Ejercicio 3.1

Como parte de su proyecto de tesis para la obtención del título de Magíster en Ciencias Odontológicas, un estudiante se propuso investigar si la distancia intercanina (DI) difiere entre hombres y mujeres. Para ello realizó mediciones en 20 participantes; 10 hombres y 10 mujeres. Las mediciones de los hombres fueron 5,1; 4,9; 5,2; 5,7; 5,0; 5,5; 5,8; 5,4; 5,7 y 4,5. Las mediciones de las mujeres fueron 4,3; 4,4; 4,2; 4,8; 4,1; 4,7; 4,7; 4,4; 3,9 y 4,5.

Seleccione una medida de resumen y calcúlela tanto para hombres como para mujeres, para determinar si ambos géneros presentan o no similar DI.

Ejercicio 3.2

A continuación, se presentan valores de temperatura en la Isla Gorriti y en la Cuchilla del Perdido. Los valores de la Isla Gorriti son:

21, 16, 30, 22, 20, 36, 30, 20, 19, 15, 24, 17, 25, 22, 20, 36, 29, 28, 19, 31, 23, 27, 24, 23, 21, 27, 19, 21, 34 y 18.

Mientras que los valores de la Cuchilla del Perdido son:

25, 26, 23, 25, 27, 25, 25, 25, 14, 24, 25, 21, 25, 23, 24, 24, 25, 35, 24, 23, 24, 25, 25, 24, 25, 29, 25, 24 y 25.

Evalúe la dispersión de la temperatura en ambos lugares.

Calcule el rango y el rango intercuartil para cada uno.

Interprete los valores obtenidos e indique en qué lugar se dan temperaturas más homogéneas.

Ejercicio 3.3

Con el fin de evaluar la implementación de una nueva campaña publicitaria para promover la asistencia odontológica preventiva en niños en edad escolar, se registró el número de visitas realizadas por un grupo de 20 niños de entre 6 y 12 años de edad durante los últimos cinco años. Se obtuvieron los siguientes datos: 0, 3, 5, 6, 7, 10, 11, 2, 1, 5, 4, 8, 7, 13, 3, 0, 2, 1, 9, 4.

Realice un gráfico de caja que le permita interpretar la distribución del número de visitas.

Análisis de dos variables cuantitativas

Introducción

En el análisis de los estudios clínicos y epidemiológicos surge la necesidad de determinar la relación entre dos variables cuantitativas (número de caries, profundidad de sondaje, edad) en una muestra dada. Los objetivos de dicho análisis suelen ser:

- Determinar si las variables están relacionadas, es decir, si los valores de una variable tienden a ser más altos o más bajos con valores más altos o más bajos de la otra variable.
- Poder predecir el valor de una variable, dado un valor determinado de la otra variable.
- Cuantificar el grado de concordancia entre los valores de las dos variables.

En este capítulo se presentan dos medidas de asociación lineal, las más utilizadas en la práctica para cuantificar el grado de asociación entre dos variables cuantitativas. Estas medidas de asociación son la covarianza y el coeficiente de correlación de Pearson.

Medidas de asociación lineal

Covarianza

En el capítulo 3 se introdujo la varianza, que cuantifica la variabilidad de una sola variable; la covarianza generaliza el concepto de *varianza*, mide el grado de asociación lineal entre dos variables cuantitativas. A continuación, se presenta la fórmula del cálculo:

$$\sigma_{XY} = \sum_{i=1}^n \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

Obsérvese la similitud entre esta fórmula y la correspondiente al cálculo de la varianza (cuya interpretación no era muy clara debido a las unidades de medida del indicador).

En cuanto a la interpretación del valor obtenido mediante el cálculo de la covarianza, solo puede interpretarse en cuanto al *signo* del valor obtenido.

- Si la covarianza es *positiva*, esto quiere decir que valores elevados en una de las variables suelen estar acompañados de valores elevados en la otra y viceversa (los valores bajos en una suelen estar acompañados por valores bajos en la otra).
- Si la covarianza es *negativa*, esto quiere decir que valores elevados en una de las variables suelen estar acompañados por valores bajos en la otra y viceversa.

Ejemplo 13

En un estudio epidemiológico realizado por la Facultad de Odontología se quiso indagar sobre la posible asociación entre la frecuencia de cepillado y el número de caries. Para ello se encuestaron 10 pacientes, estos concurrieron a la facultad, allí se le preguntó a cada uno cuántas veces por día se cepillaba y se le realizó un examen bucal para determinar el número de caries. En el cuadro 4.1 se presentan los datos obtenidos.

Cuadro 4.1 Frecuencia de cepillado y caries dental, estudio epidemiológico de la Facultad de Odontología

id	1	2	3	4	5	6	7	8	9	10
frecuencia	4	4	3	3	2	2	1	1	0	0
caries	0	1	2	4	3	4	4	5	6	7

Para cuantificar la relación entre estas variables, se propone utilizar la covarianza. En el cuadro 4.2 X indica los valores de frecuencia de cepillado, mientras que Y indica los valores de caries.

Cuadro 4.2 Cálculos intermedios para hallar la covarianza de la frecuencia de cepillado y caries dental, estudio epidemiológico de la Facultad de Odontología

X	$(X - \bar{X})$	Y	$(Y - \bar{Y})$	$(X - \bar{X})(Y - \bar{Y})$
4	2	0	-3	-6
4	2	1	-2	-4
3	1	2	-1	-1
3	1	4	1	1
2	0	3	0	0
2	0	4	1	0
1	-1	4	1	-1
1	-1	5	2	-2
0	-2	6	3	-6
0	-2	7	4	-8
				-2,7

El valor final de la covarianza es $-2,7$ y, siendo negativo, indica que cuanto mayor es la frecuencia de cepillado, menor es el número de caries.

Correlación

El mayor inconveniente de la covarianza es que puede adoptar cualquier valor, sea positivo, negativo o incluso cero. Debido a que su rango de variación es potencialmente infinito, no se puede hablar de un valor «grande» o «pequeño» de asociación entre variables. Para subsanar este problema se propone el *coeficiente de correlación*.

Covarianza y coeficiente de correlación son similares, pero difieren en lo que se puede interpretar de ellos. Los valores de covarianza no están estandarizados, por lo que resulta difícil determinar la fuerza de la relación entre las variables en cuestión. El coeficiente de correlación acota la variación al rango comprendido entre los valores -1 y 1 , con lo que expresa tanto la dirección como la fuerza de la relación lineal entre dos variables. Una fuerte relación lineal positiva resulta en un valor cercano a 1 , mientras que una fuerte relación lineal negativa resulta en un valor cercano a -1 . En última instancia, un valor del coeficiente de correlación (y de la covarianza) cercano a 0 indica ausencia de asociación lineal entre las variables.

La virtud del coeficiente de correlación radica en que es independiente de las unidades de medida de las variables.

El coeficiente de correlación (denominado r) se calcula a partir de la covarianza. Su valor resulta de dividir la covarianza entre el producto de la desviación estándar de cada variable.

$$r = \frac{\sigma_{XY}}{S_x \times S_y}$$

Por lo tanto, una covarianza positiva siempre producirá una correlación positiva y una covarianza negativa siempre producirá una correlación negativa.

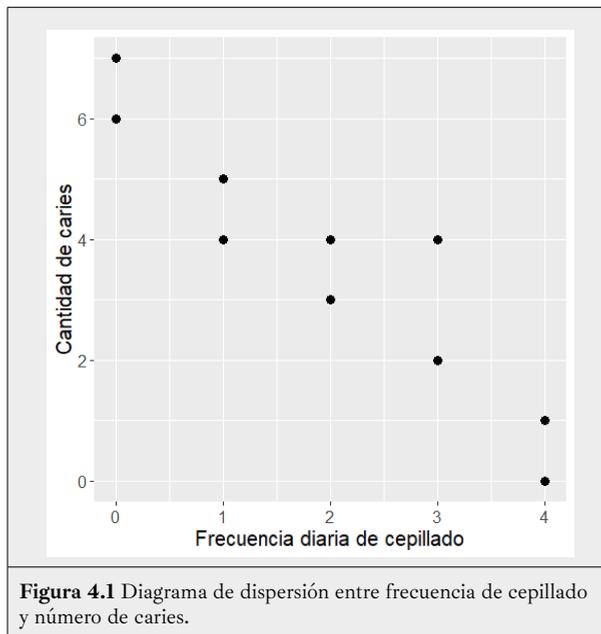
Un detalle que es importante tener en cuenta a la hora de interpretar el coeficiente de correlación es que si bien expresa la relación que puede existir entre dos variables, no permite afirmar que una sea causa de la otra. Por ejemplo, aun en el caso de que la asociación sea fuerte, esta se puede deber a que ambas tengan una misma causa (lo que explicaría su coocurrencia habitual).

Representación gráfica

El coeficiente de correlación mide la relación lineal entre dos variables. En un caso ideal (teórico) se puede representar a través de una línea recta. Por lo tanto, al momento de analizar las relaciones entre dos variables es recomendable representarlas gráficamente y luego calcular el coeficiente de correlación.

La representación gráfica es llevada a cabo mediante el *diagrama de dispersión*. Para su realización se debe considerar un eje cartesiano por cada variable, lo cual se traduce en que los valores de una variable se representan en un eje horizontal y los de la otra se representan sobre un eje vertical.

Retomando el ejemplo 13, de frecuencia de cepillado y número de caries, es posible representar estos 10 datos mediante el gráfico de dispersión en la figura 4.1.



En este caso se puede observar cómo al aumentar la frecuencia de cepillado disminuye el número de caries.

Ejercicios

Ejercicio 4.1

Además de la frecuencia de cepillado, también se consultó a los 10 participantes del estudio sobre sus hábitos alimenticios. Se registró el número de días por semana que consumen productos azucarados. Los datos ampliados se presentan en el cuadro 4.3.

Cuadro 4.3 Datos ampliados del ejercicio 4.1

id	1	2	3	4	5	6	7	8	9	10
frec. cepillado	4	4	3	3	2	2	1	1	0	0
frec. prod. azuc.	1	3	2	4	6	5	6	5	7	6
caries	0	1	2	4	3	4	4	5	6	7

Calcule la covarianza entre el número de caries y la frecuencia de consumo de productos azucarados.

Ejercicio 4.2

A partir de los cálculos realizados en el ejercicio 4.1, calcule el coeficiente de correlación entre la frecuencia de consumo de productos azucarados y el número de caries.

¿Le parece que la asociación entre ambas variables es fuerte, moderada o baja?

Ejercicio 4.3

A partir de los datos analizados en los ejercicios 4.1 y 4.2, realice el gráfico de dispersión entre la frecuencia de consumo de productos azucarados y número de caries. ¿Es consistente con el valor del coeficiente de correlación calculado en el ejercicio 4.2?

Análisis de dos variables cualitativas

Introducción

En este capítulo se extenderán las nociones presentadas sobre frecuencias y proporciones a cuando estas son registradas en dos o más grupos de interés para el investigador. Tal como anteriormente el cálculo de frecuencias (absolutas o relativas) apuntaba a la descripción de un fenómeno medido a través de una variable cualitativa, ahora, al segmentar los datos en dos o más grupos, el interés se centrará, además, en la comparación de dichas medidas de resumen.

El tipo de análisis realizado en este capítulo se lleva a cabo sobre dos variables cualitativas, una de ellas hace referencia al fenómeno a describir, mientras que la otra indica cómo se agrupan los datos a analizar. Este tipo de situaciones son muy comunes en la práctica, cuando, por ejemplo, se quiere investigar sobre el efecto de un medicamento o de un cierto factor de riesgo sobre una enfermedad.

Tabla de contingencia

Así como al trabajar con una única variable cualitativa se introdujo el concepto de *distribución de frecuencias* para organizar la información, a continuación se presenta la *tabla de contingencia* (o *tabla de doble entrada*), el homólogo bivariado de la distribución de frecuencias.

Supongamos que en un estudio se desea comparar el grado de acuerdo de un grupo de participantes ante la instauración de una nueva política según tengan o no estudios terciarios. Supongamos que la variable «a favor» se dividía en las clases «de acuerdo» y «en contra». Una vez recolectados los datos, son volcados en la tabla 5.1.

Tabla 5.1 Estructura de una tabla de contingencia del grado de acuerdo con la instauración de una nueva política según los encuestados posean estudios terciarios

		Grados de acuerdo		
		A favor	En contra	Total
Estudios terciarios	Con			
	Sin			
Total				

Esta tabla organiza la información correspondiente a la ocurrencia simultánea (o conjunta) y marginal de las clases de las variables relevadas. Los dos pares de celdas con rayas azules verticales corresponden a las frecuencias absolutas de cada una de las variables por separado, la suma de cada par debe dar lo mismo —ya que se trata del mismo número de participantes— y ser igual al total, que se registra en la celda punteada. Estas frecuencias se denominan *marginales*, son las mismas que se habrían registrado en dos distribuciones de frecuencias separadas.

La verdadera importancia de esta tabla reside en las celdas con rayas verdes diagonales, donde se sitúan las frecuencias simultáneas o «conjuntas», las cuales indican el número de casos que corresponden a determinada clase de cada variable al mismo tiempo.

Lo expuesto hasta aquí (y lo que sigue) también es válido si las variables involucradas cuentan con más de dos clases.

Ejemplo 14

Siguiendo con la premisa de que en un estudio se desea comparar el grado de acuerdo de un grupo de participantes ante la instauración de una nueva política según tengan o no estudios terciarios, en la tabla 5.2 se presentan las frecuencias que podrían haberse volcado en la tabla 5.1.

Tabla 5.2 Grado de acuerdo con la instauración de una nueva política según los encuestados posean estudios terciarios

		Grado de acuerdo		
		A favor	En contra	Total
Estudios terciarios	Con	13	4	17
	Sin	5	7	12
Total		18	11	29

A partir de las distribuciones marginales de cada variable es posible verificar que:

- La mayoría de los participantes del estudio estuvieron a favor de la nueva política (18 a favor y 11 en contra).
- La mayoría de los participantes contaba con estudios terciarios (17).
- El número total de encuestados fue de 29, que corresponde al total de frecuencias absolutas en ambas variables.

Al interior de la tabla es posible apreciar que trece participantes estuvieron a favor de la nueva política y contaban con estudios terciarios; solo cuatro participantes con estudios terciarios estuvieron en contra. De los doce participantes que no contaban con estudios terciarios, cinco estuvieron a favor y siete en contra.

Continuando con el análisis de las tablas de contingencia, de la misma manera que con las distribuciones de frecuencia se calcularon frecuencias relativas (para cada frecuencia absoluta con respecto del total de casos), estas tablas también admiten el cálculo de porcentajes. No obstante, dado que existen tres totales en la tabla (el general, el de las columnas y el de las filas) se pueden construir tres tipos de frecuencias relativas, a los que se les suele llamar *perfiles*.

Perfiles

Los porcentajes de la tabla de contingencia más sencillos de calcular son los que resultan de relativizar cada frecuencia (tanto las marginales como las conjuntas) respecto del total general. Para su construcción basta con realizar una regla de tres para cada elemento de la tabla. La tabla 5.3 presenta el resultado de aplicar este procedimiento a los valores de la tabla 5.2.

Tabla 5.3 Frecuencias relativas conjuntas del grado de acuerdo y estudios terciarios

		Grado de acuerdo		
		A favor	En contra	Total
Estudios terciarios	Con	44,8%	13,8%	58,6%
	Sin	17,2%	24,1%	41,4%
Total		62,1%	37,9%	100%

En última fila (distribución marginal de las columnas) se puede observar que el 62,1% de los encuestados se manifestaron a favor de la nueva política. En la distribución marginal de las filas (última columna) se observa que el 58,6% de los participantes poseían estudios terciarios.

Vale la pena realizar dos observaciones. La primera es que, al provenir de las distribuciones marginales, ambas afirmaciones son llevadas a cabo de manera independiente entre sí. La segunda es que así como las frecuencias absolutas marginales de la tabla 5.2 sumaban el total general, las frecuencias relativas marginales de la tabla 5.3 suman 100 %.

Los valores al interior de la tabla reflejan el peso relativo de cada combinación de clases de ambas variables respecto del total. Por ejemplo, el 24,1% de los participantes se manifestaron en contra de la nueva política y no contaban con estudios terciarios. En cuanto a la suma de los valores, su cálculo asegura que su suma sea 100 %.

A modo de ejemplo se presenta el cálculo de algunos valores de la tabla 5.3.

%(A favor-Estudios terciarios)	%(En contra)
29.....100%	29..... 100%
13.....x	11..... x
$x = \frac{13 \times 100}{29} = 44,8$	$x = \frac{11 \times 100}{29} = 37,9$

No obstante, este tipo de análisis no provee mucha más información que la contenida en la tabla original en tanto que no permite la comparación de la frecuencia de las clases de una variable respecto de las clases de la otra. Con este propósito es que se plantea la construcción del *perfil fila* (horizontal) y el *perfil columna* (vertical). El motivo principal por el que se plantea la construcción de estos perfiles es que *no es posible la comparación de las frecuencias conjuntas de la tabla si los valores marginales no son iguales*.

El perfil fila es la tabla de contingencia que resulta del cálculo de porcentajes de cada valor de la tabla respecto del total de su propia fila. En el caso de la tabla 5.2, al leer la primera fila se observó que de los 17 participantes con estudios terciarios, 13 se manifestaron a favor. Plantear la regla de tres correspondiente permite concluir que el 76,5% de los participantes con estudios terciarios se manifestaron a favor. El porcentaje de participantes que se manifestaron a favor en la segunda fila de la tabla (participantes sin estudios terciarios) se construye del mismo modo, relativizando mediante una regla de tres las frecuencias conjuntas de dicha fila con respecto al total de la fila. De esta manera es posible calcular que de los 12 participantes sin estudios terciarios (que representan el 100% de la fila) el 41,7% se manifestó a favor (5 participantes). La tabla 5.4 presenta el perfil fila.

Tabla 5.4 Perfil de grado de acuerdo según estudios terciarios

		Grado de acuerdo		
		A favor	En contra	Total
Estudios terciarios	Con	76,5%	23,5%	100%
	Sin	41,7%	58,3%	100%
Total		62,1%	37,9%	100%

Cada fila presenta un valor de 100 % que facilita la lectura de la tabla y la última fila presenta las frecuencias relativas marginales correspondientes a las clases de la variable «grado de acuerdo». A partir de esta tabla resulta posible comparar las proporciones de respuestas a favor de la nueva política según se tengan o no estudios terciarios. En este caso se puede observar que los participantes con estudios terciarios son mucho más propensos a manifestarse a favor (76,5% contra 41,7%). Esta última afirmación es fundamental, ya que nos permite observar que *en un perfil fila, la comparación de valores debe realizarse dentro de cada columna*.

Análogamente al caso anterior, el perfil columna es la tabla de contingencia que resulta del cálculo de porcentajes de cada valor de la tabla respecto del total de su propia columna. Retomando la tabla 5.2, la primera columna presenta las frecuencias de «estudios terciarios» (13 con estudios terciarios y 5 sin) en el grupo de participantes que se manifestaron a favor (18 participantes). La metodología del cálculo del perfil columna solo requiere el cálculo de reglas de tres en cada columna. En la primera columna se calcula que 72,2% de los encuestados que se manifestaron a favor tenían estudios terciarios, mientras que el 27,8% no los tenía. Al completar los demás cálculos, se obtiene la tabla 5.5.

Tabla 5.5 Perfil de estudios terciarios según grado de acuerdo

		Grado de acuerdo		
		A favor	En contra	Total
Estudios terciarios	Con	72,2%	36,4%	58,6%
	Sin	27,8%	63,6%	41,4%
Total		100%	100%	100%

Debido a que esta tabla de contingencia tiene el valor 100 % a lo largo de la última fila, *en el perfil columna, la comparación de porcentajes debe realizarse entre los valores de cada fila*. En este sentido es posible observar que las personas que se manifestaron a favor fueron mucho más propensas a contar con estudios terciarios (72,2%) que los que se manifestaron en contra (36,4%). En este caso se están comparando las proporciones de estudios terciarios en los grupos generados por la variable «grado de acuerdo».

La importancia de los perfiles fila y columna reside en que a partir de ellos se hace posible la comparación de proporciones de un fenómeno entre grupos generados por otra población. El aspecto importante a tener presente es que cada situación particular requiere el cálculo de solo uno de ellos.

En el ejemplo presentado la intención era comparar el grado de acuerdo según se tuvieran (o no) estudios terciarios, debido a esto y a la forma en la que se planteó el primer cuadro (con «estudios terciarios» en las filas y «grado de acuerdo» en las columnas), el perfil fila es el que plantea la respuesta a la pregunta.

Si la variable cuyo comportamiento interno se quiere estudiar está en las filas, ese dato se encontrará en el perfil fila. Si la variable cuyo comportamiento interno se quiere estudiar está en las columnas, ese dato se encontrará en el perfil columna.

Representación gráfica

La representación gráfica de las tablas de contingencia no presenta mayores dificultades, siempre y cuando se vea a la tabla como una representación bivariada de una distribución de frecuencias. Al tratarse de variables cualitativas, solo es posible graficar las frecuencias de cada clase.

Como se mencionó anteriormente, existen distintos tipos de frecuencias que resultan de los distintos perfiles de la tabla. Los dos gráficos más utilizados para graficar el perfil fila y el perfil columna de una tabla de contingencia son:

- El *gráfico de barras agrupadas*, que ofrece una lectura de uno de los perfiles de la tabla de contingencia agrupando (horizontalmente) los valores de una clase que sumen 100% y separándolos de los valores de otra clase que también sumen 100%.
- El *gráfico de barras estratificado*, donde el procedimiento es similar, pero la agrupación se da verticalmente, de modo que las barras tienen la misma altura (con divisiones dentro de cada barra).

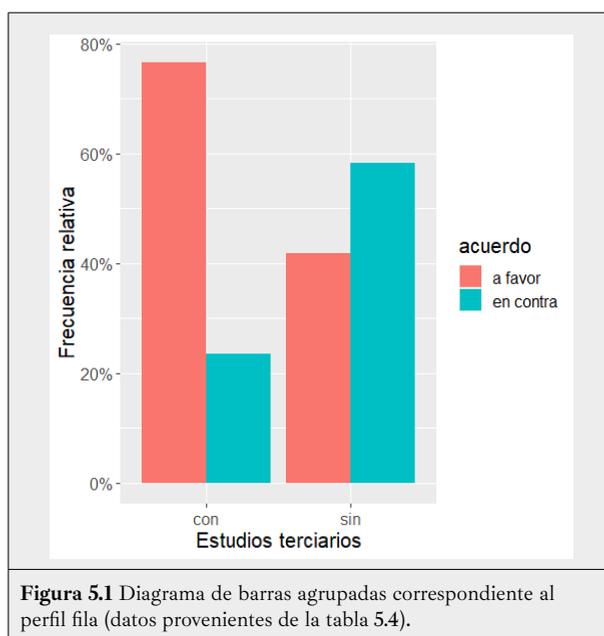
A continuación se presentan los dos gráficos correspondientes al perfil fila del ejemplo, correspondiente al estudio entre «grado de acuerdo» y «estudios terciarios». El motivo del estudio es comparar el grado de acuerdo según se tuvieran estudios terciarios y las clases «con estudios terciarios» y «sin estudios terciarios» se ubicaron en las filas, por esto la construcción del perfil fila permite satisfacer el estudio (la comparación de las frecuencias absolutas no es posible debido a que el número de personas con estudios terciarios es diferente al número de personas sin estudios terciarios).

Para la construcción de cualquiera de los dos gráficos es necesario identificar cuáles son los valores que deben agruparse. En la tabla 5.6 se presenta el perfil fila correspondiente a la tabla 5.4 y se señalan los pares de valores que se agruparían en la confección de las barras.

Tabla 5.6 Valores necesarios para la construcción de los gráficos

		Grado de acuerdo		Total
		A favor	En contra	
Estudios terciarios	Con	76,5%	23,5%	100%
	Sin	41,7%	58,3%	100%
Total		62,1%	37,9%	100%

En el caso del gráfico de barras agrupadas, deberían crearse dos grupos de barras; uno correspondiente a la clase «con estudios terciarios» y otro para la clase «sin estudios terciarios». Cada uno de estos grupos consistiría de dos barras situadas una junto a la otra; una para la frecuencia relativa «a favor» y otra para la frecuencia relativa de la clase «en contra». La figura 5.1 presenta el gráfico de barras agrupadas.



En el caso del gráfico de barras estratificadas, el procedimiento es el mismo, con la salvedad de que en lugar de situarlas una junto a la otra, las barras correspondientes a las clases «a favor» y «en contra» se sitúan una sobre la otra. La figura 5.2 presenta el gráfico de barras estratificadas del ejemplo 14.

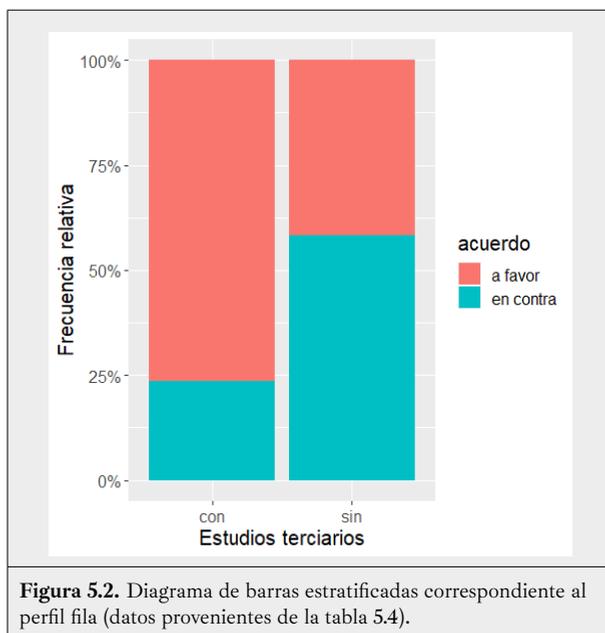


Figura 5.2. Diagrama de barras estratificadas correspondiente al perfil fila (datos provenientes de la tabla 5.4).

Medidas de asociación

El análisis de dos variables cualitativas expuesto hasta aquí se ha basado en tablas de doble entrada. Estas fueron presentadas como una distribución de frecuencia bivariada donde la información perteneciente a una muestra de n sujetos es resumida en una tabla con cuatro frecuencias conjuntas (en el caso de dos variables con dos clases cada una). Más allá de esto, de la misma forma en que el coeficiente de correlación resume toda la información de dos valores en un solo valor, el investigador puede plantearse resumir la tabla en el valor de un único *estadístico* (ver «Definiciones» en el capítulo 6).

Estos estadísticos, además de reducir la tabla de contingencia a un único valor, pretenden cuantificar lo que en estadística se denomina *la fuerza de la asociación entre dos variables*. De aquí en más, se hablará de *asociación* como la intensidad con la que una clase (o más) de una variable tiende a aparecer en la muestra conjuntamente con otra clase (o más) de la otra variable considerada. Se entenderá

por *asociación perfecta* que una clase de la variable X aparece siempre con una clase de la variable Y , mientras que la otra clase de X aparece siempre con la otra clase de Y . Se entenderá por *asociación nula* que las clases de las dos variables coexisten independientemente la una de la otra.

Los estadísticos que se plantean a continuación pretenden comparar dos valores de la tabla de contingencia, pero de diferente manera, retomando el ejemplo del grado de acuerdo según estudios terciarios. Mediante el perfil fila (tabla 5.4) se pudo ver que los participantes con estudios terciarios estaban a favor de la nueva política en el 76,5% de los casos, mientras que aquellos que no contaban con estudios terciarios estaban a favor solo en el 41,7% de los casos. Una primera forma de comparar estos valores sería a través de la diferencia, se podría decir: «El grado de acuerdo con la nueva política difiere 34,8% entre las personas con y sin estudios terciarios». Sin embargo, no es posible saber qué tan significativo resulta esto (no sabemos cuánto apoyo hay para cada opción), por lo que puede conducir a interpretaciones erróneas. Debido a que la diferencia entre proporciones no parece ser la mejor opción, a continuación se plantean dos métodos que trabajan con cocientes.

Método 1

El primer método, llamado *riesgo relativo* o *razón de prevalencia*, cuantifica qué tan grande es una de las proporciones de la tabla de contingencia en comparación con la proporción de la misma clase en la otra variable. La forma de cálculo es a través de un cociente, por lo que no tiene unidades de medida. Este estadístico indica qué tanto mayor es el valor del numerador con respecto al del denominador. Adicionalmente, y como se ilustrará a continuación, este valor admite una interpretación porcentual.

Tanto la tabla 5.4 como la figura 5.1 presentan el perfil fila del ejemplo mencionado, en ambos se puede observar que la proporción de personas a favor de la nueva política es de 76,5% y 41,7% en personas con y sin estudios terciarios respectivamente. Al realizar el cociente entre estas cifras, se obtiene la fuerza de la asociación:

$$\frac{76,5 \%}{41,7 \%} = 1,83$$

La interpretación de este valor indica que matemáticamente el valor 76,5 es 1,83 veces mayor que el valor 41,7 o, lo que es lo mismo, 83 % más grande. Esto quiere decir que, sea cual sea la proporción (o la frecuencia absoluta) de participantes a favor entre las personas sin estudios terciarios, la proporción de personas a favor entre los participantes con estudios terciarios será 83 % superior.

Si el valor calculado hubiese sido exactamente igual a 1 (para lo cual las proporciones en que se dividen deben ser iguales), la fuerza de la asociación sería nula, ya que la proporción de personas a favor —con o sin estudios terciarios— sería la misma.

Si el cociente hubiese sido realizado de manera inversa,

$$\frac{41,7 \%}{76,5 \%} = 0,54$$

como las proporciones comparadas siguen siendo las mismas, el resultado refleja el mismo fenómeno, solo que corresponde a la otra cara de la misma moneda. En este caso el valor indica que la proporción de personas a favor entre quienes no tienen estudios terciarios es 46 % menor a la proporción entre los que sí tienen estudios terciarios.

Es importante aclarar que, al igual que el 83 % anterior, el 46 % surge de restar 1 al valor del estadístico y multiplicar el resultado por 100.

Resulta equivalente afirmar «las personas con estudios terciarios están a favor en una proporción 83 % mayor» o «las personas sin estudios terciarios están a favor en una proporción 46 % menor».

Método 2

Este método, llamado *odds ratio*, también desarrolla la comparación a través de un cociente, pero no de proporciones, sino de *chances (odds)*.

La *chance de ocurrencia de un evento* es qué tan posible es la ocurrencia de ese evento respecto de su no ocurrencia. Para ilustrar este concepto se utiliza el ejemplo de arrojar un dado y registrar el número de la cara superior.

Si interesara la *probabilidad* de que salga el número 3, esa probabilidad es 1/6, debido a que solo uno de los posibles 6 resultados genera el evento «sale el 3».

El caso de la chance es distinto, si interesa la chance de que salga el número 3, la división se realiza respecto al número de eventos en contra del evento «sale el 3» y no respecto del total de eventos, por lo tanto la operación es 1/5.

Finalmente, para el cálculo del estadístico en cuestión se comparan las chances del evento de interés, una para cada clase de la variable que define los grupos a comparar. A continuación se presentará, paso a paso, el cálculo de este estadístico en el ejemplo de «grado de acuerdo» según «estudios terciarios», utilizando el perfil fila.

Como se mencionó anteriormente, el primer paso consiste en el cálculo de las chances; es posible calcularlas a partir de las probabilidades. La chance de un evento puede ser calculada dividiendo la probabilidad de la ocurrencia de dicho evento entre la probabilidad de la no ocurrencia de ese mismo evento:

$$chance(\text{evento}) = \frac{p(\text{evento})}{[1 - p(\text{evento})]}$$

De esta manera, a partir del perfil fila es posible calcular la chance de estar a favor en los dos grupos a comparar.

Grupo con estudios terciarios:

$$chance(\text{a favor}) = \frac{76,5 \%}{(100 \% - 76,5 \%)} = 3,25$$

Grupo sin estudios terciarios:

$$chance(\text{a favor}) = \frac{41,7 \%}{(100 \% - 41,7 \%)} = 0,72$$

Obsérvese que las chances no poseen unidades de medida.

Finalmente, el valor de la fuerza de la asociación se calcula comparando estos valores a través de un cociente de chances (*odds ratio*).

$$\frac{3,25}{0,72} = 4,51$$

Por lo tanto, es posible afirmar que la chance de estar a favor de la nueva política entre los participantes con estudios terciarios es 351% mayor a la chance de estar a favor entre los participantes sin estudios terciarios. La obtención del 351% fue llevada a cabo del mismo modo que se interpretó el valor de la fuerza de la asociación con el método 1 (restar 1 al valor del estadístico y multiplicar el resultado por 100).

En caso de que el cociente de chances fuese realizado de manera inversa, el resultado sería inverso, reflejaría el mismo fenómeno visto desde el punto de vista del otro grupo.

$$\frac{0,72}{3,25} = 0,22$$

Como en este caso el numerador lo ocupa la chance del grupo «sin estudios terciarios», este resultado indica que la chance de estar a favor entre los participantes «sin estudios terciarios» es 78% menor que entre participantes «con estudios terciarios».

Resulta equivalente afirmar «las personas con estudios terciarios tienen 351% más chance de estar a favor» o «las personas sin estudios terciarios tienen 78% menos chance de estar a favor».

Consideraciones finales

En el ámbito de la epidemiología, las comparaciones llevadas a cabo mediante los valores planteados a lo largo de este capítulo son llamadas riesgo relativo (RR) y *odds ratio* (OR). Tanto el RR como el OR no poseen unidades de medida y su rango de valores posibles va de cero hasta infinito, pasando por el valor 1, que equivale a igualdad de chances, también equivale a igualdad de riesgos, que a su vez equivale a la situación en que la fuerza de la asociación entre las dos variables consideradas es nula.

En el ámbito de la epidemiología la elección del estadístico a utilizar suele estar determinada por el tipo de estudio a partir del cual se generaron los datos. El cálculo del RR no se recomienda en estudios de tipo retrospectivo, debido a que la asignación inicial de los participantes es a partir de la presencia/ausencia de la enfermedad estudiada y el carácter aleatorio está dado por la presencia/ausencia del factor de riesgo considerado (por lo que las proporciones de enfermedad que surgen del perfil adecuado no representan las proporciones de enfermedad en la población general).

El RR es recomendado en estudios de carácter prospectivo, donde los participantes (con y sin el factor de riesgo bajo estudio) están sujetos a desarrollar o no la enfermedad.

En el capítulo 9 de este manual se profundiza en estos dos estadísticos debido a su extenso uso en estudios epidemiológicos.

Ejercicios

Ejercicio 5.1

Se llevó a cabo un estudio para evaluar el efecto de dos nuevos tratamientos sobre la gripe. El total de 100 participantes fue dividido en dos grupos de 50. En el tratamiento A se recuperaron 35 participantes, mientras que en el B solo se recuperaron 18.

Sitúe en una tabla de contingencia los valores correspondientes.

Ejercicio 5.2

Una clínica odontológica realizó una encuesta a 202 pacientes con el objetivo de evaluar el impacto del hábito de fumar sobre la enfermedad periodontal. De los 84 fumadores, 62 presentaron enfermedad periodontal. En total se encuestaron 124 pacientes que no presentaban enfermedad periodontal, de los cuales 102 eran no fumadores.

- Plantee una tabla de contingencia que resuma la información proporcionada. Tenga en cuenta que para completarla deberá calcular algunos valores por diferencia atendiendo a que los totales marginales resultan de la suma de las frecuencias conjuntas.
- Seleccione y plantee el perfil adecuado para saber si los fumadores son más propensos a desarrollar enfermedad periodontal que los no fumadores.

Ejercicio 5.3

La tabla 5.7 es una tabla de contingencia donde se volcó información proveniente de un estudio para evaluar el efecto del consumo de dulces sobre la presencia de caries.

Tabla 5.7 Tabla de contingencia, estudio sobre el efecto del consumo de dulces

	tiene caries	no tiene caries	Total
come dulces	145	32	?
no come dulces	?	69	86
Total	162	?	?

Calcule las frecuencias faltantes en la tabla y, si corresponde, calcule el perfil que le permita analizar si la proporción de personas con caries es distinta entre las personas que consumen dulces y las que no. Represente sus hallazgos gráficamente.

Introducción a la inferencia estadística

Introducción

La inferencia estadística es una etapa crucial en el proceso de informarnos acerca del mundo que nos rodea, ya que nos ayuda a entender nuestro entorno y a tomar decisiones acerca de cómo actuar frente a determinadas circunstancias.

La *inferencia* es el conjunto de procedimientos estadísticos utilizados para deducir propiedades de determinadas cantidades (*parámetros poblacionales*) en la población, a partir de cierta información obtenida mediante algún esquema de muestreo.

Formalmente, se trata de la selección de un modelo probabilístico que represente el fenómeno investigado, la evaluación de dicho modelo y los resultados que provee.

Definiciones

Para obtener un manejo adecuado de los procesos inferenciales es fundamental conocer (y dominar) los modelos probabilísticos en los que se basará el proceso inferencial. También es necesario conocer ciertos conceptos como *población*, *muestra* y *parámetro*.

El concepto de *población* se refiere a un conjunto de elementos, sujetos o eventos con ciertas características de interés, sobre los cuales se quiere responder alguna pregunta o llevar a cabo algún experimento. Estas poblaciones pueden consistir de elementos reales (por ejemplo, el conjunto de personas de una ciudad, las piezas dentales de una persona) o de un conjunto, que incluso puede ser infinito, de elementos hipotéticos (por ejemplo, todos los posibles valores de la temperatura ambiente en un momento dado, todos los posibles partidos de ajedrez).

Un objetivo habitual de la inferencia estadística es obtener algún tipo de información de la población de interés.

Cuando contar con la totalidad de la población es imposible (o poco práctico), la investigación es llevada a cabo sobre una *muestra*. La muestra es un subconjunto (de tamaño manejable) de la población. Dentro del ámbito de la inferencia estadística es deseable contar con muestras obtenidas mediante algún tipo de esquema aleatorio (al azar) que, *a priori*, otorgue a cada miembro de la población una probabilidad positiva de ser seleccionado para integrar la muestra. De esta manera se evitan eventuales errores (*sesgos de selección*), que podrían invalidar los resultados inferenciales.

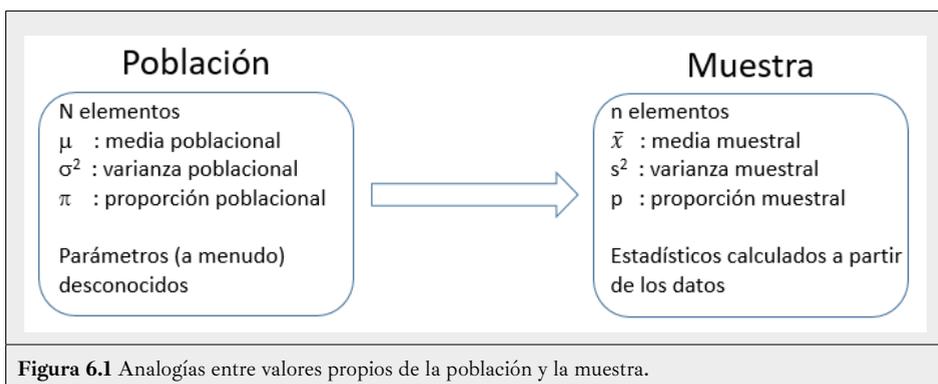
Habiendo definido estos dos conceptos, es posible comparar los papeles que desempeñan la estadística inferencial y la estadística descriptiva:

- La estadística inferencial tiene como propósito *inferir* (extraer conclusiones) acerca de las propiedades de la población bajo estudio.
- La estadística descriptiva intenta *describir* los datos recabados en la muestra, sin siquiera asumir que esta proviene de una población mayor o si fue hecha mediante un procedimiento probabilístico.

Una definición técnica de lo que es un *parámetro* requiere de conocimientos de probabilidad y modelos probabilísticos que exceden el objetivo de este manual. Por lo tanto, se definirá a un *parámetro* como «aquella cantidad o valor que caracteriza a cierto conjunto de datos». Al referirse a un valor que resume una cierta característica al nivel de la población, se habla de *parámetros poblacionales* (o sencillamente *parámetros*).

La contraparte muestral de los parámetros poblacionales se define como *estadístico*, que puede verse como una función de los datos observados.

De esta manera, la inferencia estadística intenta realizar afirmaciones sobre los parámetros a partir de los estadísticos (es decir, de los datos observados y procesados).



En la figura 6.1 se presentan algunas características de interés en una población y en una muestra. En primer lugar, se presenta la cantidad de elementos de cada una, vale aclarar que el tamaño muestral n siempre es menor o igual al tamaño de la población N (en algunas ocasiones incluso asume que la población cuenta con una cantidad infinita de elementos).

Luego se presentan algunos valores comúnmente utilizados, los parámetros son valores que caracterizan la totalidad de la población (se suelen escribir con letras griegas) mientras que los estadísticos son sus equivalentes muestrales (se suelen escribir con letras latinas).

En resumen, la mayor diferencia entre los parámetros y los estadísticos es que los valores de los estadísticos pueden ser calculados una vez que la muestra es relevada, mientras que los parámetros no son conocidos (ya que se trabaja con una muestra por la imposibilidad de trabajar con la población; si se trabaja con toda la población no es una *muestra*, sino un *censo*).

La siguiente figura presenta un ejemplo sencillo que pretende aclarar los conceptos anteriores. Se presenta un experimento donde una bolsa contiene bolillas numeradas del uno al cinco y se extraen tres de ellas, sin reponerlas.

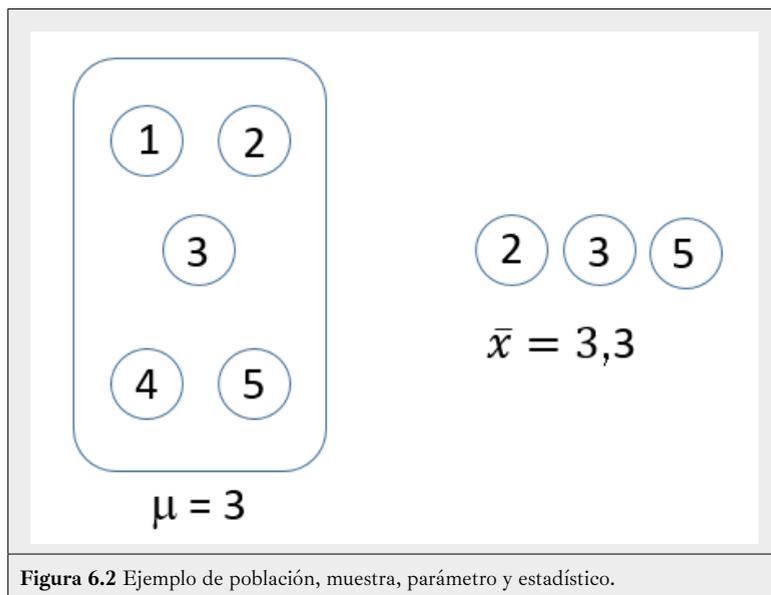


Figura 6.2 Ejemplo de población, muestra, parámetro y estadístico.

La selección de la muestra es realizada al azar, de esa manera todos los elementos de la urna pueden ser seleccionados. En este caso, los elementos seleccionados resultaron ser el dos, el tres y el cinco.

Pese a que el valor poblacional de la media era 3, el valor del estadístico equivalente fue 3,3, la diferencia entre ambos se debe a las variaciones en el mecanismo aleatorio de selección de la muestra. Es debido a este carácter aleatorio de la muestra que el valor de los estadísticos calculados a partir de ella resulta aleatorio.

En el ejemplo anterior se seleccionó una de las posibles diez muestras. La tabla 6.1 expone las diez posibles muestras junto con el valor de la media muestral en cada una de ellas:

Tabla 6.1 Distribución en el muestreo de la media muestral

muestra
(1 , 2 , 3) = 2,0
(1 , 2 , 4) = 2,3
(1 , 2 , 5) = 2,7
(1 , 3 , 4) = 2,7
(1 , 3 , 5) = 3,0
(2 , 3 , 4) = 3,0
(1 , 4 , 5) = 3,3
(2 , 3 , 5) = 3,3
(2 , 4 , 5) = 3,7
(3 , 4 , 5) = 4,0

El valor del estadístico es desconocido hasta seleccionar la muestra, pero gracias al cuadro anterior es posible observar que los valores 2,7; 3,0 y 3,3 son más probables a los valores 2,0 y 4,0. Esta colección de valores es llamada *distribución en el muestreo*, e indica cuáles son los valores más (o menos) probables de un estadístico.

La distribución en el muestreo de un estadístico dado juega un papel fundamental en la inferencia estadística, ya que presenta la variación de un estadístico a través de todas las posibles muestras. En el ejemplo que se expuso anteriormente, por motivos didácticos, se obtuvo una tabla de valores relativamente pequeña, pero en experimentos más complicados la distribución en el muestreo no se expone en una tabla, sino que es una fórmula a partir de la cual se pueden obtener propiedades probabilísticas de un estadístico.

En la práctica se cuenta con una sola muestra, conformada por n elementos, sin embargo, existen desarrollos teóricos que permiten llevar a cabo procedimientos inferenciales a partir de esta única muestra y el cumplimiento de ciertos supuestos.

A continuación, empleando el ejemplo 4.7.7 del libro *Biostatistics for Oral Healthcare* (Kim & Dailey, 2008b), se describirán muy someramente los procedimientos inferenciales más comunes (Borrell & Crawford, 2012). El ejemplo plantea

una situación donde se pretende estudiar el impacto del hábito de fumar sobre el éxito de la colocación de implantes dentales. Al tratarse de un estudio que busca conocer la asociación entre dos variables cualitativas, los datos se presentan en forma de tabla de contingencia (tabla 6.2) (Kim & Dailey, 2008b).

Tabla 6.2 Número de pacientes según hábito de fumar y resultado del implante

Fumador	Resultado del implante		
	Fracaso	Éxito	Total
Si	9	38	47
No	10	166	176
Total	19	204	223

En el capítulo 5, correspondiente al análisis de tablas de contingencia, se propuso el cálculo de perfiles fila y perfiles columna como una forma de comparar las proporciones de las clases de una variable según las clases de la otra. Ante esta tabla podría plantearse cuál es la proporción de éxito de los implantes cuando los pacientes son fumadores o no. En este sentido el cálculo del perfil fila resulta adecuado.

Tabla 6.3 Distribución del resultado del implante según hábito de fumar

Fumador	Resultado del implante		
	Fracaso	Éxito	Total
Si	19,1%	80,9%	100%
No	5,7%	94,3%	100%
Total	8,5%	91,5%	100%

A partir de la tabla 6.3 es posible observar que los pacientes que fuman poseen una proporción de fracasos de implante más elevada que los que no fuman. Una de las maneras en las que se puede cuantificar la relación entre ambas proporciones es a través del *odds ratio* (OR), que en esta tabla es de 3,93. Según este resultado, los pacientes que fuman tienen una chance de que su implante falle que equivale a casi cuatro veces la chance de fallo que tienen los pacientes que no fuman (ver capítulo 5).

Procedimientos inferenciales

Tanto las proporciones presentadas en la tabla 6.3 como el valor del OR han sido calculados a partir de la muestra de 223 pacientes, por lo cual son estadísticos que *estiman* sus equivalentes poblacionales (la población estaría conformada por el grupo de pacientes con implantes, hayan sido exitosos o no).

Al calcular el valor de un estadístico, se está llevando a cabo el procedimiento inferencial más sencillo: la *estimación puntual*.

Los siguientes procedimientos inferenciales surgen al considerar que las estimaciones puntuales son realizadas a partir de una muestra aleatoria de la población bajo estudio, por lo que dichas estimaciones no tienen por qué coincidir (por lo general no coinciden) con los parámetros poblacionales que intentan representar. De esta manera surge la idea de cuantificar el grado de incertidumbre que poseen dichas estimaciones.

Para cuantificar el grado de incertidumbre es necesario tener en cuenta la definición considerada en el muestreo que se expuso anteriormente (tabla 6.3).

El procedimiento que se expone a continuación tiene como objetivo, en lugar de un único valor (estimación puntual), proponer un rango de valores en el cual se encuentre el parámetro poblacional desconocido. Este método se denomina *estimación por intervalos* o *cálculo de intervalos de confianza*.

Retomando al ejemplo anterior, el valor del OR se estimó en 3,93. A partir de este valor se procede a calcular un intervalo de confianza y se obtiene el rango de valores comprendido entre 1,49 y 10,44. Es decir que este procedimiento de estimación por intervalos estaría indicando que el valor de OR, que es desconocido, puede llegar a tener valores tan pequeños como 1,49 o tan grandes como 10,44.

En este caso puede observarse que el intervalo calculado a partir de esta muestra tiene una amplitud considerable, pero no contiene al valor 1, descartándolo como posible valor del OR poblacional. El valor 1 tiene especial importancia al estimar un OR, ya que representa una situación donde las chances estimadas en ambos grupos serían indistinguibles.

La amplitud del intervalo depende de varios factores; uno de ellos es el tamaño muestral, al aumentar el tamaño muestral, la muestra se aproxima más a la población y las estimaciones se vuelven más precisas, lo cual se traduce en intervalos más angostos.

Otro factor relevante es la confianza del intervalo. El intervalo calculado anteriormente posee una confianza del 95%, esto quiere decir que, si se repitiera infinitamente el experimento consistente en

1. recolectar los datos de personas que reciben implantes;
2. registrar si fuman o no;
3. registrar si el implante tuvo éxito o falló;
4. calcular un intervalo de confianza;

entonces el 95% de dichos intervalos incluiría el verdadero valor del parámetro poblacional desconocido (en este caso el OR). La teoría probabilística que sustenta estos resultados y posibilita los métodos de cálculo de los intervalos está más allá del cometido de este manual.

Por último, se plantea el método de *pruebas de hipótesis*, donde, a partir de la formulación de una hipótesis que tenga el investigador, se trata de contrastar qué tanto conforman los datos con dicha hipótesis. Para ello es necesario formular la hipótesis del investigador en términos estadísticos, es decir, empleando un parámetro.

Ejemplo 15

En la tabla 6.2 se presentó una tabla de contingencia donde se resumen los datos de un estudio donde se investigó la relación entre el hábito de fumar y la falla de un implante. Una hipótesis natural que diera origen a este estudio podría ser: «El que una persona sea fumadora tiene relación significativa con la propensión a que sus implantes fallen».

Para abordar este problema mediante una prueba de hipótesis, se debería identificar un parámetro que indique en qué casos se estaría cumpliendo (o no) la hipótesis del estudio. Como se mencionó anteriormente, dos variables cualitativas (con dos clases cada una) no están asociadas si su OR es igual a uno. De esta manera, las dos hipótesis que componen la prueba serían «el OR es igual a uno» y «el OR es distinto de uno». En términos estadísticos se escribiría de la siguiente manera:

$$H_0: \text{OR} = 1 \quad H_0) \vee 1$$

$$H_0: \text{OR} \neq 1 \quad H_1) \neq 1$$

De estas dos, la primera es llamada *hipótesis nula*, ya que hace referencia a una situación donde la relación entre las variables consideradas es nula. En contraposición, la otra hipótesis es llamada simplemente *hipótesis alternativa*.

El procedimiento en sí consta de diversos pasos, pero en líneas generales: se parte de una situación en la que se adopta como cierta a la hipótesis nula, luego se calcula un estadístico y a partir de este, empleando un criterio de decisión, se toma la decisión de rechazar o no rechazar la hipótesis nula. El criterio utilizado para tomar esta decisión es la comparación de dos cantidades: el *p-valor* y el *nivel de significación* de la prueba.

El nivel de significación (comúnmente denominado *error de tipo I* y denotado por α) indica la probabilidad de rechazar la hipótesis nula. Este valor es fijado por el investigador (en general es del 5%). Esto implica que: partiendo de la base de que la hipótesis nula es cierta, si el experimento que dio lugar a los datos se repitiese infinitamente y la prueba de hipótesis se realizara en cada una de dichas repeticiones, el 5% de las veces que se rechazara la hipótesis nula se estaría cometiendo un error (si el investigador eligiera un nivel de significación menor al 5%, estaría siendo más exigente en cuanto al error que está dispuesto a tolerar).

Por otro lado, el *p-valor* (que se calcula a partir de los datos observados) indica cuál es la probabilidad de observar un valor del estadístico aún más extremo al observado. Alternativamente, el mismo criterio de decisión puede formularse a partir de la comparación del valor del estadístico y un cierto *valor crítico*. Este último es obtenido empleando el nivel de significación previamente establecido. En la mayoría de las pruebas los criterios para rechazar la hipótesis nula serían:

- Que el *p-valor* sea menor al nivel de significación.
- Que el valor del estadístico (típicamente en valor absoluto) sea mayor al valor crítico.

Ambos criterios son equivalentes, ya que es posible pensar en ellos como dos caras de una misma moneda.

Ejemplo 16

Continuando con el ejemplo anterior, tras definir las hipótesis a partir del OR, el siguiente paso es el cálculo de un estadístico. Sin entrar en los pormenores estadísticos del asunto (y yendo directamente a la etapa del criterio de decisión) el resultado de la prueba consistiría en comparar el valor del estadístico (que es 2,77) con el valor crítico (que con un nivel de significación del 5% es 1,96). De esta manera, se estaría en condiciones de rechazar la hipótesis nula de que el OR sea 1 (situación de no asociación) en favor de la hipótesis alternativa de que el OR sea distinto de uno. De esta forma se llegaría a la conclusión de que la asociación entre las variables hábito de fumar y falla del implante es significativa.

Ejercicios

Ejercicio 6.1

Un nutricionista quiere estimar cuánto sodio consumen los niños de menos de diez años en Uruguay. Para ello, a partir de una muestra de 75 niños, obtiene un consumo promedio de 2993 miligramos de sodio al día.

¿Sería capaz de identificar el parámetro y el estadístico en esta situación?

Ejercicio 6.2

Si bien este manual no ahonda en los fundamentos probabilísticos sobre la estadística inferencial, indique cuáles de los siguientes aspectos podrían reducir la amplitud de un intervalo de confianza.

- a.* Disminuir la confianza del intervalo.
- b.* Que la variable analizada presente menor variabilidad.
- c.* Trabajar con una muestra de mayor tamaño.
- d.* Todas las anteriores.

Ejercicio 6.3

En el capítulo 4 se presentó al coeficiente de correlación de Pearson como una medida de la asociación lineal entre dos variables cuantitativas. ¿Cómo formularía una prueba de hipótesis para comprobar si existe o no evidencia que permita respaldar la hipótesis de que personas con más años de edad presentan mayores niveles de caries? (Recuerde que la dirección de la correlación no interfiere en el cálculo).

PARTE II

Indicadores epidemiológicos en salud bucal

Concepto y características de los indicadores epidemiológicos

Introducción

Los indicadores son herramientas que utiliza la epidemiología para resumir, generalmente desde el punto de vista cuantitativo, información referida a condiciones de vida y nivel de salud en una comunidad o grupo de personas. Desde el punto de vista estadístico se conocen como *medidas de frecuencia*, pueden ser tasas, proporciones, *odds*, medidas de riesgo, razones o diferencias.

Utilizando el método epidemiológico podremos utilizar estas herramientas en diversos momentos y procesos de trabajo. En algunos casos serán nuestro punto de partida para elaborar diagnósticos de salud comunitarios y en otros estarán al final del proceso de diseño de una investigación epidemiológica, al realizar lo que llamaremos *operacionalización* de las variables en estudio.

En este capítulo se considera la utilización de indicadores a la hora de realizar un diagnóstico epidemiológico y se presenta la forma en que se calculan en general. En los capítulos 10 a 13 se abordarán los indicadores específicos de cada una de las enfermedades más prevalentes o graves en odontología: enfermedad periodontal, caries, trastornos témporo mandibulares, cáncer oral.

Indicadores al realizar un diagnóstico sociocomunitario

Los indicadores serán útiles a la hora de realizar un diagnóstico sociocomunitario (diagnóstico poblacional), ya que brindan la información necesaria para facilitar la selección de qué problemas de salud abordar, identificar alternativas, recursos e intervenciones y, sobre todo, para monitorear los cambios. Para todo esto será estratégico que estén integrados en un sistema de información, un conjunto de datos que al procesarlos y analizarlos pasan a tener sentido (es como si los datos nos hablaran, nos dijeran cosas sobre ese grupo de personas).

Partiendo del concepto de *epidemiología* de Rouquayrol (Rouquayrol *et al.*, 2014), expresado ya al comienzo de este manual, queda claro que el propósito al trabajar en epidemiología con las poblaciones es conocer la distribución y los determinantes del proceso de salud-enfermedad para que sirva como base para las intervenciones en salud.

Si además nos declaramos inscriptos en el paradigma de la epidemiología crítica, y acompañamos la conceptualización que hace el profesor Breilh de dicho paradigma y de la determinación social de la salud (entendida como derecho humano y como proceso multidimensional, socialmente determinado, donde intervienen aspectos socioeconómicos, culturales, psicológicos y biológicos, entre otros), nuestro foco estará en una batería de indicadores y procesos que buscaremos integrar para luego intervenir en favor de la salud.

Entendemos que los indicadores socioeconómicos y culturales están en la base de la explicación de los procesos de salud y enfermedad, por lo que tendrán crucial importancia a la hora de realizar el análisis.

Según los aspectos de la realidad sobre los cuales resumen información los indicadores, encontraremos varios tipos:

- Sociodemográficos: cuantifican procesos vitales, como los nacimientos y las muertes (por ejemplo, la tasa de natalidad, la de mortalidad).
- Estructurales: geográficos, políticos, de estructura social y clases sociales.
- Epidemiológicos (de morbilidad y mortalidad): permiten conocer las enfermedades existentes (por ejemplo, la proporción de enfermos de diabetes en la población, el CPOD, la tasa de mortalidad infantil).
- Socioculturales: permiten conocer las características socioculturales de los grupos humanos (por ejemplo, trabajo, educación, alimentación).
- Ambientales: características de una zona (por ejemplo, nivel de polución, nivel de agroquímicos tóxicos, número de basurales). Los indicadores ambientales pueden incluir factores económicos cuando refieren al conjunto de las características naturales y sociales de existencia.
- De los servicios de salud: permiten caracterizar y monitorear los servicios de salud existentes (por ejemplo, número de policlínicos en una zona determinada, número de consultorios odontológicos, personal de salud en la zona o institución considerada).

Cuando el diagnóstico que se propone realizar es a nivel de una comunidad concreta, se hará referencia a un *diagnóstico local* (mesodiagnóstico), y cuando se trabaja a nivel de las poblaciones se estará realizando un *análisis de la situación de salud* (diagnóstico macro). El primero terminará con la instalación de un programa de salud y el segundo con la instalación de una política en salud.

De acuerdo al concepto expuesto, el personal de salud que utilice indicadores deberá tener claro su propósito, ya sea para seleccionar alguno de los indicadores antes mencionados, ya sea para construir aquellos que mejor describan la situación de salud que le interese mostrar, evaluar, o sobre la cual se trabajará.

Los indicadores en el proceso de operacionalización de variable en una investigación

Al desarrollar un proceso de investigación epidemiológica para conocer las características de una comunidad o población podemos desarrollar investigaciones cuantitativas (descriptivas, analíticas o experimentales) o cualitativas (observación participante, grupos focales, entre varios otros).

En estos casos (especialmente en una investigación cuantitativa) el proceso de investigación epidemiológica incluye *operacionalizar* las variables en estudio. La operacionalización incluye al menos dos etapas, una fundamentalmente teórico-conceptual y otra metodológica. En la primera se define conceptualmente la variable y en la segunda se lleva el concepto a una dimensión observable que podrá ser resumida con uno o más indicadores.

Cuando se estudian situaciones de vulnerabilidad de grupos humanos expuestos a procesos malsanos que incluyen *factores de riesgo* corresponden estudios prospectivos (de cohorte) o retrospectivos (de caso-control) (procedimientos desarrollados por la epidemiología tradicional). Cuando se busca evaluar la eficacia de una terapia preventiva o medicamento se desarrollarán estudios experimentales. En todos los casos, al desarrollar el diseño metodológico del estudio (pensar cómo se realizará, con qué personas, procedimientos, instrumentos, personal, etc.) se deberá realizar la operacionalización de las variables.

Para cualquiera de los tipos de estudio mencionados se deberá tener claro el objetivo y la conceptualización de los procesos a estudiar. Luego se intentará descomponer esos procesos, en general complejos, en características observables y medibles. Estas características cuantificadas son los indicadores.

Desde el punto de vista metodológico, los diagnósticos comunitarios son muy similares a un estudio epidemiológico descriptivo, solo que se le agregan, muchas veces, técnicas de investigación cualitativa (entrevistas, grupos focales) que aportan otra dimensión de los procesos; además, se pretende que la comunidad se involucre para buscar en conjunto la priorización y resolución de los problemas.

Ejemplo 17

En una investigación en pacientes adultos que concurren a la Facultad de Odontología demandando atención, se quiere conocer el nivel de enfermedad bucal. El estudio se centra en una primera etapa en la enfermedad caries dental.

En la etapa del diseño metodológico se seleccionará, por ejemplo, el criterio de esmalte cavitado, propuesto por la OMS para estudios poblacionales (WHO, 2013), dejando fuera del conteo las etapas previas de la caries dental.

A su vez, si se quiere conocer el número de personas que tienen una o más caries con cavidad en la población, se estará trabajando con un indicador que reporta el número relativo de personas con caries para esa población (*prevalencia*, ver capítulo 9); si lo que interesa es saber el número de caries en cada individuo, se trabajará con un indicador de conteo (1, 2, 3 caries). En el primer caso se utilizarán fundamentalmente la proporción o la tasa (medidas de frecuencia, capítulo 9), y en el segundo la media o la mediana (medidas de resumen, capítulo 2).

Medidas de frecuencia, relaciones entre variables y comparación de poblaciones

Uno de los objetivos de la epidemiología es cuantificar en frecuencia y magnitud los fenómenos en las poblaciones humanas. Cuando se estudia un fenómeno en una población se puede estar interesado en ver con qué frecuencia ocurre y, si se trata de una enfermedad, saber si se relaciona con alguna característica de esa población. Uno puede seguir profundizando el estudio e indagar qué tan fuerte es esa relación entre la enfermedad y la característica seleccionada.

Para estudiar la periodontitis en la población de estudiantes de tercer año de la Facultad de Odontología, en primer lugar se estudiaría la cantidad de enfermos, en segundo lugar se indagaría sobre la relación entre la periodontitis y el hábito de fumar, y en tercer lugar se buscaría qué tan fuerte es esa relación. En este capítulo se estudiarán indicadores que permitan abordar los tres niveles.

La medida más básica para estudiar la frecuencia de un fenómeno de salud es la frecuencia absoluta (número) de casos, pero tiene poca utilidad a la hora de dimensionar el fenómeno en estudio. La frecuencia absoluta permite cuantificar el fenómeno pero no permite comparar (concepto ya visto en el capítulo 5, donde se presentaron los perfiles fila y columna). No es lo mismo considerar el hallazgo de 5 casos de caries incipiente en una población de 50 niños (10 %) que en una población de 1000 (0,5%). Es así que se consideran otras medidas de frecuencia que son cocientes, donde importa el numerador pero también el denominador.

Las medidas de frecuencia se establecen a partir de dos conceptos epidemiológicos fundamentales: *prevalencia* e *incidencia*. La prevalencia se relaciona con el número de casos de un fenómeno en la población y la incidencia refiere a la velocidad de avance de ese fenómeno; estos conceptos se abordarán en profundidad más adelante.

Las relaciones entre las variables y comparación de poblaciones se realizan utilizando proporciones, tasas y razones.

Proporción

Número de casos de una cierta clase comparado con una muestra (o población) determinada. Cuando se compara con una población se expresa π , cuando se compara con una muestra se expresa p .

Es una relación entre dos eventos posibles, donde el numerador forma parte del denominador.

$$\text{proporción} = \frac{A}{A + B} \times 100$$

Normalmente se expresan en porcentaje. Por ejemplo, si de 30 alumnos (A+B), 17 son mujeres (A) y 13 son hombres (B), la proporción de mujeres es:

$$p(\text{mujeres}) = \frac{17}{17 + 13} \times 100 = \frac{17}{30} \times 100 = 56,7\%$$

En el capítulo 5 se usan proporciones para construir los perfiles de las tablas de contingencia y calcular frecuencias relativas.

Ejemplo 18

Porcentaje de aprobación según el grado de estudio (perfil fila):

	aprueba	reprueba	Total
poco	55.9%	44.1%	100%
mucho	75.0%	25.0%	100%
Total	58.8%	41.3%	100%

Analizando el cuadro, se puede afirmar que la proporción de alumnos que estudia poco y aprueba el examen es casi del 56%, mientras que la proporción de alumnos que estudia mucho y aprueba es del 75%. Parece acertado entonces afirmar que, en la población de alumnos estudiada, el estudiar mucho aumenta la probabilidad de aprobar un examen.

Prevalencia

La *prevalencia* de una enfermedad se expresa como proporción.

$$\text{prevalencia} = \frac{\text{Nro. de casos}}{\text{Población}}$$

Prevalencia puntual: Número de personas enfermas (o con la condición en estudio) en un momento del tiempo en relación con la población considerada.

Prevalencia por período: número de personas enfermas (o con la condición en estudio) en un período determinado (normalmente a la mitad del período), en relación con la población considerada. Puede expresarse en valores absolutos o relativos.

Incidencia

Implica considerar la velocidad a la que ocurre el fenómeno en estudio, por lo que involucra la dimensión tiempo. Se trata de contar los nuevos casos en un período de tiempo determinado y se vincula con el concepto de *riesgo*, que nos permite cuantificar la probabilidad del evento en un tiempo determinado (diferente de *riesgo relativo*).

Esta medida no ofrece información individual. Es una medida adecuada cuando se busca probar una hipótesis etiológica y se trata de una enfermedad crónica.

Si las personas consideradas participan en el estudio durante el mismo período, la incidencia se calcula como el número de nuevos casos en la población en el período de tiempo elegido (por ejemplo, en un año).

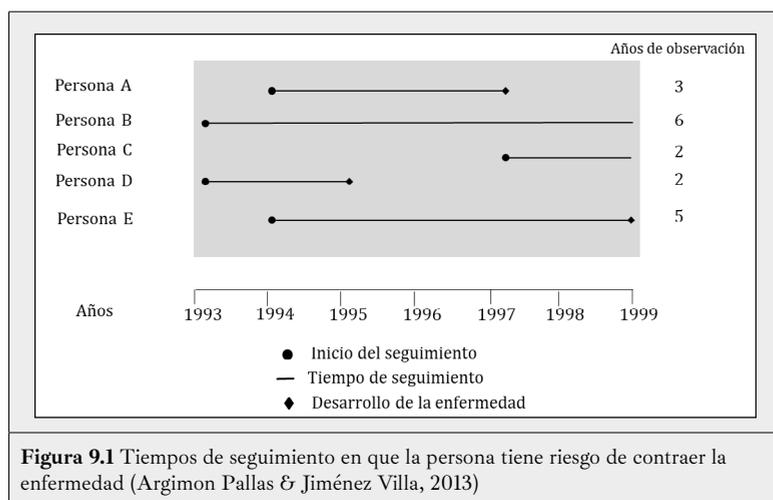
Pero esto no sucede en la práctica, por eso se calcula la *densidad de incidencia*: el número de nuevos casos/persona-tiempo. Esto implica considerar cuánto tiempo estuvo cada individuo en el estudio (¿un mes?, ¿dos?, ¿un año?) y ese valor se coloca en el denominador. Dado esto, el tiempo de seguimiento solo representa una parte del tiempo en que la persona corre el riesgo de contraer la enfermedad.

Ejemplo 19

En un estudio de seis años de duración se observaron 5 personas (figura 9.1) con diferentes tiempos de seguimiento y 3 desarrollaron la enfermedad. De haberse observado las cinco personas durante todo el período, la incidencia acumulada sería de tres nuevos casos en cinco personas en estudio por seis años ($3/5 \times 6$); el riesgo de contraer la enfermedad sería de un 10% cada persona-año.

Sin embargo, solo la persona B fue observada durante el período completo, en el resto los períodos de observación fueron menores: la persona A estuvo tres años en el estudio, la persona B estuvo seis años, las personas C y D estuvieron dos años y la persona E estuvo cinco años. Por otra parte, una vez que una persona contrajo la condición, según de que condición se trate, puede ser contraproducente continuar el seguimiento.

Si se suma el tiempo que estuvieron en el estudio las 5 personas, se obtiene el valor de 18 años. Se dice entonces que el denominador de la tasa es 18 persona-año. La densidad de incidencia será $3/18 = 0,16$ persona-año. Si se multiplica la densidad de incidencia por 100, se puede expresar como porcentaje. Se dice que la incidencia fue de 16,6% cada persona-año.



Esta afirmación es válida bajo dos asunciones: el riesgo de contraer la enfermedad no varía a lo largo del tiempo y la tasa de incidencia en las personas que se pierden durante el seguimiento es la misma que en aquellos que siguen bajo observación. Además, el denominador deberá ajustarse al problema en estudio (hay ocasiones en que será necesario considerar unidades de tiempo menores para medir los períodos de seguimiento, como un mes, o medio año).

Proporción de incidencia acumulada

La proporción de incidencia acumulada (TIA) se calcula en los estudios de cohorte, donde se parte de individuos sanos, y da idea del riesgo promedio de un individuo de contraer la enfermedad en el período de tiempo considerado.

La TIA mide «nuevos casos» cada «cantidad de personas sanas al comienzo», no tiene unidades y varía entre 0 y 1. Al principio del estudio vale 0, y si toda la muestra contrae la enfermedad pasa a valer 1.

Si el objetivo de un estudio es predecir el cambio del estado de salud de un individuo en función de alguna característica, será útil estimar el riesgo de que se contraiga la enfermedad.

Ejemplo 20

Se considera un estudio longitudinal realizado en el año 1987 sobre candidiasis y personas portadoras del virus HIV-1, involucró 123 personas libres de candidiasis al comienzo del estudio. Cada una de las personas fue acompañada durante tres años, 21 desarrollaron candidiasis al menos una vez. Entonces se podrá calcular la TIA:

$$TIA = \frac{21}{123} = 0,17$$

o 17% en tres años.

Para calcular la TIA de esta forma no debería haber pérdidas (abandono de individuos en el estudio) durante todo el período. De haber pérdidas, el cálculo se deberá realizar de otra forma, considerando el peso relativo de cada individuo según el tiempo que permaneció en el estudio y si se enfermó o no. Los individuos que abandonan el estudio se denominan *censurados*.

**

Considerando la relación entre prevalencia e incidencia, se puede afirmar que la prevalencia depende de la incidencia (velocidad de ocurrencia) y de la duración de la enfermedad (días, meses, años). No es lo mismo una enfermedad crónica que en una aguda en cuanto al tiempo de duración y la velocidad de avance. Por ejemplo, la pérdida de inserción periodontal es de avance lento y se acumula en el tiempo, en cambio la inflamación gingival es de avance más rápido y puede remitir a los pocos días con un buen cepillado.

Contar exclusivamente con datos de prevalencia suele ser insuficiente. Para conocer las características de la enfermedad es necesario contar con ambos indicadores.

Odds ratio / chance

El *odds ratio* (chances) ya fue visto en este manual como medida de asociación de variables cualitativas (capítulo 5). Ahora se hará foco en su utilidad para comparar poblaciones en estudio. Se trata de una medida utilizada sobre todo en la cultura anglosajona.

Ejemplo 21

Se dará un ejemplo de la vida deportiva para facilitar su comprensión. Peñarol y Nacional jugaron 10 partidos. En 6 de ellos ganó Peñarol y Nacional ganó en 4 (tabla 9.1). Basados en los datos que tenemos, ¿quién tiene mayor probabilidad de ganar en un próximo partido?

Se puede responder que Peñarol tiene 6 en 10 (o sea 60%, usando proporciones) y también se puede decir que Nacional tiene 4 en 10 (40%). Otra forma de expresar esta relación compara las veces a favor que tiene un equipo con las veces en contra. Para ello se usan las chances: se dice que Peñarol tiene 6 resultados en que gana, contra 4 que tiene Nacional.

Expresando en forma matemática la relación entre la proporción o probabilidad de que un fenómeno ocurra y la chance, sería así:

$$\text{chance}(\text{evento}) = \frac{p(\text{evento})}{[1-p(\text{evento})]}$$

Esto significa que el numerador y el denominador son proporciones complementarias, es decir que ambas refieren a un mismo total (en este ejemplo, los 10 partidos).

En este caso, la chance de que gane Peñarol, aplicando la fórmula anterior, será:

$$\text{chance}(\text{evento}) = \frac{p(\text{evento})}{[1-p(\text{evento})]} = \frac{6/10}{1-0,6} = \frac{0,6}{1-0,6} = \frac{0,6}{0,4} = 1,5$$

De acuerdo a lo ya visto en el capítulo 5, esto se interpreta como que Peñarol tiene 1,5 chances de ganar, o lo que es lo mismo, Peñarol tiene 50% más chance de ganar que Nacional.

También, cuando se conoce la chance se puede calcular la probabilidad:

$$\text{probabilidad} = \frac{\text{odds}}{1 + \text{odds}}$$

Entonces la probabilidad de que gane Peñarol será en este caso:

$$p(\text{P.gana}) = \frac{1,5}{1 + 1,5} = \frac{1,5}{2,5} = 0,6$$

Como ya se vio, en el 60% de los casos gana Peñarol.

Continuando con ejemplos futbolísticos (tabla 9.1), puede observarse que cuando se trabaja con tablas de contingencia que relacionan dos variables cualitativas, también se pueden aplicar las chances:

Tabla 9.1 Partidos ganados según la condición de locatario del club de fútbol Peñarol

	Ganador		Total
	Peñarol	Nacional	
Visitante Peñarol	2	3	5
Locatario Peñarol	4	1	5
Total	6	4	10

Si no se toma en cuenta quién es locatario, la chance de ganar de Peñarol será:

$$\text{chance}(\text{P.gana}) = \frac{6/10}{[1 - 6/10]} = \frac{0,6}{1 - 0,6} = \frac{0,6}{0,4} = 1,5$$

Como se puede observar, la chance no es una probabilidad. Se puede avanzar un poco más y ponerle condiciones a esa chance, en este ejemplo se puede comparar la chance de ganar de Peñarol de acuerdo a si es o no es locatario.

Para esto se aplicarán los conceptos del capítulo 4 y se trabajará con los perfiles fila de la tabla:

- La chance de que gane Peñarol, dado que es locatario es

$$\frac{4/5}{1 - 4/5} = \frac{4}{1} = 4$$

- La chance de que gane Peñarol, dado que es visitante es de

$$\frac{2/5}{1 - 2/5} = \frac{2}{3} = 0,66$$

Más adelante se volverá sobre estos conceptos al hacer referencia a medidas de asociación.

Tasa

La tasa es un caso especial de proporción (puede ser de incidencia o de prevalencia), mide el número de veces que se produce un fenómeno en una población en un tiempo y lugar dados. Sirve para comparar distintas poblaciones o incluso para comparar la misma población antes y después de determinadas acciones médicas (por ejemplo, vacunación). Refiere a un fenómeno que ocurre con respecto al total de una población.

Se calcula mediante un cociente que tiene en el numerador la cantidad de casos de la enfermedad y en el denominador la cantidad de población, considerando un tiempo y espacio determinados. Este cociente se multiplica por una constante para que resulte fácilmente entendible.

La tasa general (TG) se calcula de la siguiente manera:

$$TG = \frac{\text{Hecho que ocurre en la población}}{\text{Población a la mitad del período en esa área}} \times K$$

Donde K es un valor que depende de la magnitud del cociente. Cuando $K = 100$, la tasa se expresa como un porcentaje (o casos cada 100, que es lo mismo). Cuando $K = 1000$, la tasa se expresa en casos cada 1000 personas; cuando $K = 10000$, la tasa se expresa en casos cada 10.000 personas.

Las TG consideran a toda la población en el denominador, mientras que las tasas específicas (TE) solo a la porción que se quiere estudiar por algún motivo en particular.

La razón por la cual se toma la población a la mitad del período es arbitraria, es una determinación demográfica que permite hacer comparaciones internacionales.

Ejemplo 22

Si se considera la tasa de mortalidad infantil (TMI), que es una tasa de prevalencia entre el número de niños fallecidos antes del primer año de vida sobre el total de los niños nacidos vivos en ese año, esta tasa específica se expresa de la siguiente manera:

$$TMI = \frac{\text{N. de niños fallecidos antes del primer año de vida}}{\text{Total de niños nacidos vivos}} \times 1000$$

Es importante señalar que la TMI se calcula de manera diferente a otras tasas puesto que es una tasa específica pautada internacionalmente.

Ejemplo (año 2018):

$$TMI = \frac{269}{40139} = 0,007 = 7\text{‰}$$

Se interpreta: cada 1000 niños que nacen en un año, en ese año mueren 7 niños antes de cumplir un año.

Ejemplo 23

En el cuadro 9.1 se considera el número de defunciones y la población del Uruguay en el año 2011. A partir de esta serán calculadas la tasa bruta de mortalidad y las tasas específicas de mortalidad en Montevideo y en el interior.

Cuadro 9.1 Número de defunciones y población del Uruguay en el año 2011

	Defunciones	Población
Montevideo	14288	1319108
Interior	18778	1967206
Total	33066	3286314

La tasa bruta de mortalidad será:

$$TG = \frac{33066}{3286314} \times 1000 = 10,1\text{‰}$$

Se interpreta que fallecen 10,1 personas cada mil habitantes.

$$TE(\text{Montevideo}) = \frac{14288}{1319108} \times 1000 = 10,8\text{‰}$$

$$TE(\text{Interior}) = \frac{18788}{1967206} \times 1000 = 9,5\text{‰}$$

Tasas estandarizadas

Las tasas estandarizadas se usan cuando se quieren comparar dos o más poblaciones que difieran en su composición por factores que necesariamente han de tenerse en cuenta para que dicha comparación sea correcta (sexo, edad, etnia). No es lo mismo saber que murieron 100 personas en el año 1985 en la población de Perú que en la población de Alemania. Para poder compararlas es necesario realizar un procedimiento matemático complejo directo e indirecto que permita tener en cuenta las características mencionadas y tomar una población de referencia que habilite las comparaciones. Esto escapa al propósito de este manual, por lo que no se describe aquí el procedimiento de tal cálculo.

Razones

En algunos casos hay cantidades que no pueden ser sumadas porque refieren a distintas entidades. En estos casos se utilizan razones, que son un cociente:

$$\text{razón} = \frac{A}{B}$$

Por ejemplo, si en una ciudad hay 400.000 hogares y 1.200.000 habitantes, la razón de habitantes por hogar es:

$$\text{razón}(\text{hab./hogar}) = \frac{1200000}{400000} = 3$$

Se interpreta como «3 habitantes cada hogar».

Tabla 9.2 Resumen de las medidas de frecuencias más usadas en epidemiología

Proporciones, razones y tasas			Ejemplos		
	Operación matemática	Unidades	Indicador	Descripción	
Cociente	$\frac{A}{X}$	Un número dividido entre otro.	No corresp.	No corresponde.	No corresponde.
Proporción	$\frac{A}{A+B}$	El numerador está incluido en el denominador (A+B).	No tiene unidades.	Individuos libres de caries.	10 individuos con caries en una población de 100: 10/100 = 0,1 10 % de la población total son individuos libres de caries.
Razón	$\frac{A}{B}$	El numerador (A) no está incluido en el numerador (B).	Puede tener unidades.	Relación entre mujeres y hombres en una población.	Hay 10 hombres y 5 mujeres: 10/5. La razón es 2 a 1; hay 2 hombres por cada mujer.
Tasa	$\frac{A}{N\Delta_t}$	En el denominador se especifica la duración del período Δ_t y el tamaño de la población en la mitad del período \bar{N} .	Tiene unidades.	Tasa de mortalidad infantil (TMI).*	TMI, año 2018: 269/40139 = 0,0067 0,0067 × 1000 = 67 Se expresa como 67 ‰: 67 por mil.

* La TMI se calcula colocando en el numerador el total de defunciones de menores de un año ocurridas en un año dado y en el denominador el número de nacidos vivos en ese mismo año.

Medidas de asociación

En los estudios analíticos o experimentales, además de cuantificar la frecuencia con la que se presenta determinado evento en el grupo con el factor de riesgo (grupo expuesto) y en el grupo que no lo presenta (grupo control), interesa conocer qué tan fuerte es la relación entre el evento estudiado y la presencia o ausencia de dicho factor de riesgo.

Una manera de hallar la fuerza de asociación es comparando ambos grupos mediante cocientes; los más utilizados son el riesgo relativo (RR) y el *odds ratio* (OR). Como se vio en el capítulo 5, ambas medidas pueden calcularse a partir de tablas de contingencia.

Tabla 9.3 Estructura genérica de una tabla de contingencia

Exposición	Enfermos	Sanos	Total
Si	A	B	A+B
No	C	D	C+D
Total	A+C	B+D	A+B+C+D

Riesgo relativo (razón de incidencias)

El riesgo relativo indica qué tanto más probable es que se desarrolle una enfermedad en el grupo expuesto en relación con el grupo control. Se calcula mediante el cociente entre la incidencia en el grupo expuesto respecto de la incidencia en el grupo no expuesto, por lo que también se denomina *razón de incidencias* o *razón de riesgos*.

Considerando la disposición de valores de la tabla 9.3, se podrían definir las incidencias entre los expuestos (I_e) y los no expuestos (I_o) de la siguiente manera:

$$I_e = \frac{A}{A+B} \quad I_o = \frac{C}{C+D}$$

Al comparar estos valores mediante un cociente (y no una resta, ver capítulo 5) se obtiene el RR:

$$RR = \frac{I_e}{I_o}$$

Ejemplo 24

Se realizó un estudio basal en 180 escolares que iban a ser beneficiados con un programa educativo para prevenir la aparición de caries. Del total de niños, se registró que 60 presentaban un consumo elevado de bebidas azucaradas, mientras que los restantes 120 mantenían un consumo bajo.

Si se quisiera saber si el consumo elevado de bebidas azucaradas es un factor de riesgo de caries en esta población, se podría comenzar aplicando el perfil fila (ver capítulo 5). En este caso se pretende, además, con un solo indicador dar una idea de qué tan fuerte es la asociación entre el alto consumo de bebidas azucaradas (variable independiente) y la caries dental (variable dependiente).

Estos datos se sintetizan en la tabla 9.4.

Tabla 9.4 Presencia de caries según consumo de bebidas azucaradas

Consumo	Caries	sin Caries	Total
Elevado	20	40	60
Bajo	20	100	120
Total	40	140	180

Si se calcula la incidencia en los niños que presentaban un consumo elevado de bebidas azucaradas (grupo expuesto) y en el grupo de niños que presentaban un consumo bajo (grupo control), se observa que esta resulta mayor en el grupo de los expuestos.

$$I_e = \frac{20}{60} = \frac{1}{3} \quad I_o = \frac{20}{120} = \frac{1}{6}$$

Para expresar qué tanto mayor (fuerza de la asociación) es la incidencia, utilizando el RR:

$$RR = \frac{0,33}{0,17} = 2$$

Esto significa que en el grupo de los niños expuestos, la incidencia de caries era el doble de la del grupo de niños no expuestos.

Cuando este cociente es 1, el riesgo es el mismo en ambos grupos (expuesto y no expuesto). Como el grupo de los expuestos se coloca en el numerador, cuando el cociente es mayor que uno se interpreta que pertenecer al grupo de los expuestos es un factor de riesgo, y cuando el cociente es menor que uno se interpreta como que el pertenecer al grupo de los expuestos es un factor protector (el riesgo de desarrollar enfermedad es menor en ese grupo con respecto al grupo no expuesto).

Odds ratio (OR)

El *odds ratio* o *cociente de chances* es la medida de asociación, puede utilizarse en estudios analíticos (de cohorte), experimentales o transversales. Este indicador relaciona mediante un cociente la chance de enfermar entre los expuestos con la chance de enfermar entre los no expuestos.

En los estudios de caso-control no se recomienda utilizar el RR debido a que el diseño de estudio (el investigador selecciona por un lado una muestra de casos y selecciona a los controles de acuerdo a sus posibilidades logísticas) implica que los grupos no provienen de una misma cohorte dividida en un grupo de expuestos y otro de no expuestos, lo que hace que no sea correcto calcular incidencias con esos datos.

Para ejemplificar los cálculos de este indicador, se presenta el proceso de cálculo del OR según los valores de la tabla 9.3.

1. Cálculo de la chance de enfermedad entre los expuestos.

$$chance(\text{expuestos}) = \frac{A}{B}$$

2. Cálculo de la chance de enfermedad entre los no expuestos.

$$chance(\text{no expuestos}) = \frac{C}{D}$$

3. Cálculo del OR a partir de las chances de los dos pasos anteriores.

$$\frac{chance(\text{expuestos})}{chance(\text{no expuestos})} = \frac{A/B}{C/D} = \frac{AD}{BC}$$

Luego de este razonamiento, es posible observar que el cálculo involucra una multiplicación cruzada de los elementos de la tabla de contingencia. No obstante, la mecanización de esta operación no se recomienda, ya que se debe ser cuidadoso en cuanto a la interpretación final del indicador (si la tabla se confecciona de manera diferente los resultados serán incorrectos).

Ejemplo 25

Un grupo de investigadores trabaja sobre la influencia del ambiente en la salud humana y se propuso investigar si la presencia de plátanos (*Platanus occidentalis*) en los barrios se asocia con el aumento de casos de enfermedades respiratorias. Para ello se conformó una muestra de 25 personas, de las cuales 15 vivían en una zona de Montevideo donde había plátanos y 10 en una zona donde no había.

Tabla 9.5 Casos de enfermedad respiratoria según la presencia de plátanos

Zona	Enf. Resp.	s/ Enf. Resp.	Total
c/plátanos	10	5	15
s/plátanos	6	4	10
Total	16	9	25

Si sobre los valores de la tabla 9.5 se lleva a cabo el razonamiento indicado anteriormente, se obtienen las siguientes cifras:

1. Cálculo de la chance de contraer una enfermedad respiratoria entre los que viven en una zona con plátanos.

$$\text{chance}(\text{c. plátanos}) = \frac{10}{5} = 2$$

2. Cálculo de la chance de contraer una enfermedad respiratoria entre los que no viven en una zona con plátanos.

$$\text{chance}(\text{s. plátanos}) = \frac{6}{4} = 1,5$$

3. Cálculo del OR a partir de las chances de los dos pasos anteriores.

$$\frac{\text{chance}(\text{c. plátanos})}{\text{chance}(\text{s. plátanos})} = \frac{2}{1,5} = 1,33$$

El OR de enfermar entre los expuestos es de 1,33.

De esta manera, el cociente de chances de presentar un caso de enfermedad respiratoria entre quienes viven en una zona con plátanos respecto de quienes no es de 1,33.

También se puede expresar como que quienes viven en barrios con plátanos tienen 33% más chance de contraer una enfermedad respiratoria.

Si se analiza la situación considerando como referencia el grupo de personas que viven en un barrio sin presencia de plátanos, los pasos 1 y 2 serán iguales y el paso 3:

3. Cálculo del OR de contraer una enfermedad respiratoria para las personas que viven en zonas sin plátanos, con respecto a los que viven en zonas con plátanos:

$$\frac{\text{chance}(\text{s. plátanos})}{\text{chance}(\text{c. plátanos})} = \frac{1,5}{2} = 0,75$$

El OR de contraer enfermedad respiratoria quienes viven en zonas sin plátanos es 0,75. Como el resultado es menor de 1, se considera que vivir en un barrio sin plátanos tiene un efecto protector con respecto a contraer una enfermedad respiratoria.

Se puede afirmar que quienes viven en barrios sin plátanos tienen 25% menos chance de desarrollar una enfermedad respiratoria con respecto a los sí viven en zonas con plátanos.

Enfermedad periodontal

Breve descripción clínico-epidemiológica

Se puede conceptualizar a la enfermedad periodontal como una patología que afecta a alguno de los tejidos de soporte del diente (encía, ligamento periodontal, hueso).

Desde el punto de vista epidemiológico, puede afirmarse que la EP es una enfermedad crónica cuya distribución está modelada socialmente mediante la política y la economía (Sheiham & Nicolau, 2005). Desde el punto de vista fisiopatológico se constata que los tejidos periodontales pueden verse afectados por la presencia de un biofilm que produce una respuesta inmunoinflamatoria que puede dañarlos (Page *et al.*, 1997). Dos características clínicas relevantes de la EP son la pérdida de inserción ósea (PIP) y la presencia de bolsa periodontal (BP). La bolsa periodontal —un surco gingival profundizado por un proceso patológico— es una de las características clínicas más importantes de la EP (Genco & Borgnakke, 2013).

Mientras que la investigación clínica se concentra en las llamadas *causas proximales*, como microorganismos actuantes y mecanismos patológicos e inmunológicos del huésped, el que la EP tenga un patrón social de producción y desarrollo ha sido abordado y demostrado en varias revisiones sistemáticas (Sheiham & Nicolau, 2005; Thomson *et al.*, 2012; Borrell & Crawford, 2012; Lencova *et al.*, 2006).

Día a día surge evidencia acerca de la asociación entre las enfermedades periodontales y las enfermedades generales, de ahí la importancia de un diagnóstico periodontal preciso (Khader *et al.*, 2006; Chávarry *et al.*, 2009; Lockhart *et al.*, 2012; Graziani *et al.*, 2018; Suvan *et al.* 2011).

A pesar del creciente desarrollo de técnicas microbiológicas e inmunológicas, estas aún no van más allá de la identificación de signos clínicos en el diagnóstico periodontal (Research, Science and Therapy Committee, 2003).

Cuando se analizan los factores asociados a la EP, algunos son sistémicos y otros locales (Genco & Borgnakke, 2013). Aquellos sobre los que hoy existe suficiente evidencia son: hábito de fumar, presencia de enfermedades como diabetes tipo 1 y 2 con ausencia de control médico, y osteoporosis; posiblemente también se encuentren asociados la obesidad y el estrés (Genco & Borgnakke, 2013).

Al estudiar la EP desde el punto de vista individual se piensa en el diagnóstico del paciente para su posterior tratamiento. Sin embargo, a nivel epidemiológico interesa la situación general de las poblaciones, por lo tanto se utilizan estadísticos que resuman dicha situación para poder luego realizar intervenciones o comparaciones entre poblaciones o grupos.

Con tal de confeccionar dichos estadísticos, es importante tener un criterio claro para clasificar a los individuos en sanos o enfermos, es decir, para la *definición de caso*. Pese a esto, no se ha estandarizado aún la definición de caso de EP desde el punto de vista epidemiológico, existen algunos indicadores que toman en cuenta algunas piezas índice y otros que toman la boca completa, lo cual dificulta la comparación. Se suele usar la profundidad al sondaje o bolsa periodontal (BP) y la pérdida de inserción (PIP) mayor o igual a 4 mm.

El elemento más crítico no parece ser el número de cuadrantes o sextantes elegidos, sino los sitios de los dientes examinados (Beltran-Aguilar *et al.*, 2012).

Ya sea usando registros parciales o totales, entre un 20 % y un 40 % de los adultos sufre EP. Si analizamos los casos más graves, por ejemplo, bolsas mayores a 6 mm (índice CPI), se constata que los individuos son entre el 10 % y el 15 % de los adultos (Petersen & Beahni, 2012; Oppermann *et al.* 2015).

Indicadores de enfermedad periodontal usados en epidemiología

Índice periodontal comunitario (CPI)

El índice periodontal comunitario (CPI, por su sigla en inglés) resume la situación de las siguientes condiciones periodontales: sangrado, sarro y bolsa.

Este índice fue propuesto por la OMS en 1997, en sustitución del índice periodontal comunitario de necesidades de tratamiento (CPITN, por su sigla en inglés), que estaba en uso en la década de los ochenta. El CPITN no solo informaba sobre la presencia de enfermedad periodontal, sino también sobre las necesidades de tratamiento. Las necesidades de tratamiento se retiraron porque el tratamiento de la enfermedad periodontal es individual y en diferentes pacientes con un mismo tratamiento la evolución puede ser diferente.

Es un índice de boca parcial, ya que la divide en sextantes y selecciona dientes índices que representarán a los diferentes sextantes (figura 10.1).

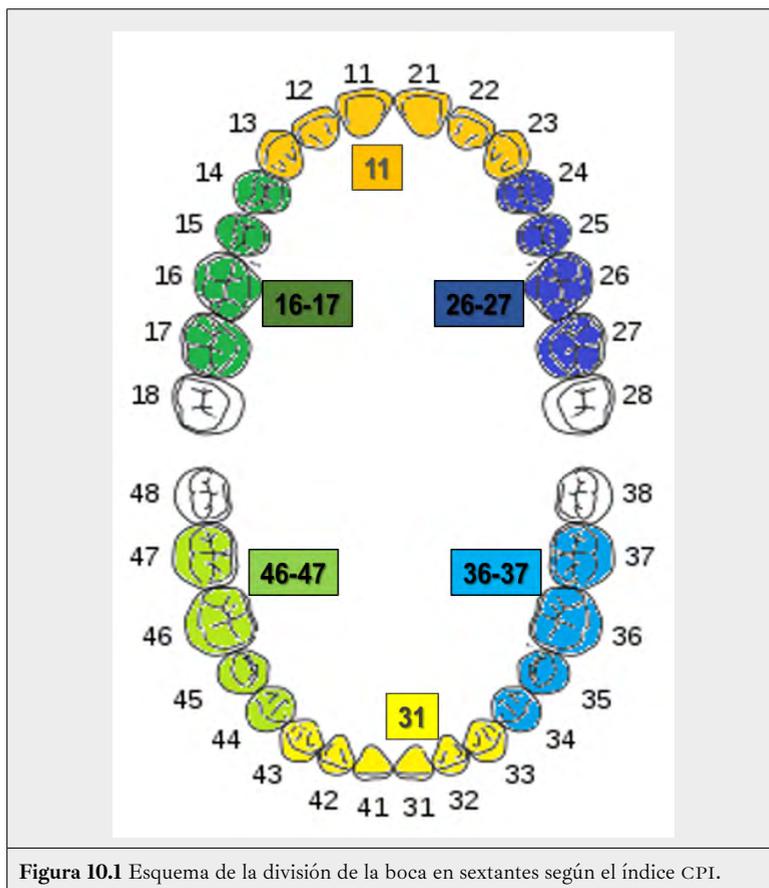


Figura 10.1 Esquema de la división de la boca en sextantes según el índice CPI.

Examen clínico y registro

Deben examinarse los dientes índice, o todos los dientes que queden en el sextante cuando no existe el diente índice. Si en el sextante hay menos de dos piezas, el sextante se considera excluido (figura 10.1).

Para el examen clínico se utiliza una herramienta especialmente diseñada: la sonda CPI de la OMS. Es una sonda que tiene una esfera de 0,5 mm en su extremo y está graduada con tres bandas negras. La primera banda negra va desde los 3,5 a los 5,5 mm, la segunda va desde los 5,5 a los 8,5 mm y la tercera va desde los 8,5 a los 11,5 mm (figura 10.2).

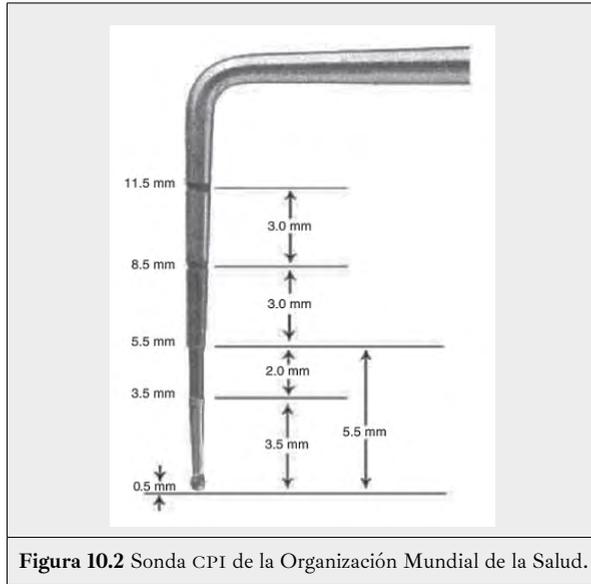


Figura 10.2 Sonda CPI de la Organización Mundial de la Salud.

Medición y resumen del sangrado

Cada diente índice se examina recorriendo todo el perímetro y observando si existe sangrado al sondaje (sonda OMS, con punta redondeada.).

Medición de la bolsa periodontal

En cada diente índice se miden seis puntos (tres vestibulares y tres palatinos o linguales) y se anota la medida del sitio que tiene la peor condición. La presencia de bolsa periodontal se mide como la distancia entre el margen gingival y el fondo del surco.

Cuando la primera banda negra de la sonda es visible, se considera que no hay bolsa periodontal (0,5 a 3,5 mm). Cuando el margen gingival queda dentro de la primera banda negra (3,5 a 5,5 mm) estamos en presencia de *bolsa periodontal incipiente* (4 a 5 mm). Cuando el margen gingival queda por fuera de la primera banda negra —cuando esta ya no se ve— se considera que la bolsa es *profunda* (6 o más mm).

Registro de la enfermedad periodontal con el criterio CPI

Este indicador codifica la peor condición observada en cada diente índice en el siguiente orden: bolsa (peor condición), sarro y sangrado.

Para el indicador CPI, un diente índice se codifica con la siguiente escala: 0 (sano), 1 (sangrado), 2 (sarro), 3 (bolsa incipiente), 4 (bolsa profunda); si no existen piezas índice para examinar y quedan menos de dos piezas en el sextante, se codifica como 5 (sextante excluido). Los códigos para cada condición son entonces:

0. Sano

1. Hemorragia observada, directamente o utilizando el espejo bucal, después de la exploración.
2. Sarro observado durante la exploración con la sonda, la primera banda negra de la sonda es completamente visible.
3. Bolsa de 4-5 mm (margen gingival dentro de la primera banda negra de la sonda).
4. Bolsa de 6 mm o más (no se llega a ver la primera banda negra de la sonda).
5. Sextante excluido (existen menos de dos dientes).

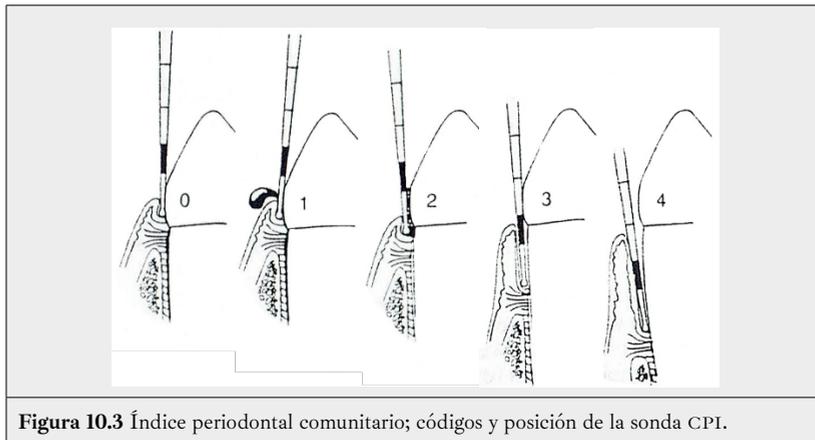


Figura 10.3 Índice periodontal comunitario; códigos y posición de la sonda CPI.

Pérdida de inserción periodontal (PIP)

Hay otra condición que no está incluida en el CPI y que es necesario valorar para saber si hay EP: la pérdida de inserción periodontal (PIP). Se valora en cada diente índice como la distancia entre el límite amelocementario y el fondo del surco. Se usan las mismas bandas que en el CPI (figura 10.2): 0 (hasta 3 mm); 1 (4-5 mm); 2 (6 mm o más).

Modificaciones al CPI para relevamientos nacionales

Ya que el CPI registra solo la peor situación del sextante, en el año 2003 un grupo de investigadores brasileiros (Ministério da Saúde, 2009) diseñó un sistema de registro donde se anota la situación de cada diente índice con respecto a las tres condiciones relevadas, sarro, hemorragia y bolsa (figura 10.4). Esto permite obtener la prevalencia de cada condición en la población a partir de la situación de los dientes índice.

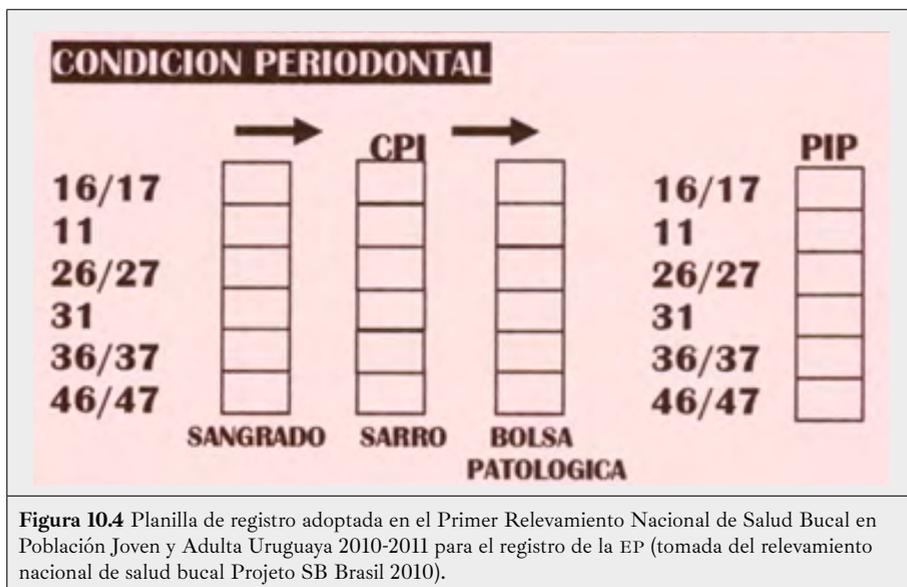


Figura 10.4 Planilla de registro adoptada en el Primer Relevamiento Nacional de Salud Bucal en Población Joven y Adulta Uruguay 2010-2011 para el registro de la EP (tomada del relevamiento nacional de salud bucal Projeto SB Brasil 2010).

Ejemplo 26

A los efectos de entender las diferencias entre ambos criterios, se presenta el siguiente caso:

Un individuo presenta sangrado y sarro en todos los sextantes. Además, en el 36 presenta bolsa incipiente (4-5 mm) y profunda (6 mm o más); en el 37 presenta bolsa incipiente. En el 46 tiene también bolsa profunda. El sextante 26-27 tiene solo una pieza y el sextante que contiene al 11 no tiene piezas dentales.

Tabla 10.1 Comparación entre el sistema CPI OMS, 1997, y la Modificación Brasil, 2003, aplicada al ejemplo 26

Situación de los dientes índice	Criterio OMS 1997	Modificación Brasil 2003			
		Sangrado	Sarro	Bolsa patol.	Excluido
Sextante 16-17	2	X	X	-	-
Sextante 11	5	-	-	-	X
Sextante 26-27	5	-	-	-	X
Sextante 31	2	X	X	-	-
Sextante 36-37	4	X	X	X	-
Sextante 46-47	4	X	X	X	-

Al finalizar el examen periodontal, la OMS propone al menos reportar el siguiente resumen:

- % de sujetos con tejidos periodontales sanos.
- % de sujetos con hemorragia.
- % de sujetos con sarro (cálculo dental).
- % de sujetos con bolsas de 4-5 mm como peor condición.
- % de sujetos con bolsas profundas (> 6) como peor condición.
- promedios de sextantes excluidos del examen.
- % de sujetos con pérdida de fijación en el grado más alto.

Recordar que con el sistema de registro propuesto por Brasil podemos registrar la prevalencia de cada condición.

Ventajas y desventajas del CPI

Ventajas: rápido, reproducible, internacionalmente reconocido.

Desventajas:

- Mucha de la información recogida no se utiliza (solo se registra la peor situación de cada uno de los seis sextantes).
- Se subestima, en términos generales, la prevalencia de la EP.
- El CPI presenta los valores de profundidad de bolsa en forma categórica (0, 1, 2), por lo que no es posible obtener una dimensión real de la enfermedad.
- Presenta valores de profundidad de bolsa y no de pérdida de inserción. Al no medir las secuelas de la enfermedad, esto lleva a subestimarla.

Algunos autores plantean que el CPI tiende a subestimar la prevalencia de la patología, mientras que otros entienden que la sobreestima. Según Ainamo (1975), este indicador tiende a sobreestimar la severidad y a subestimar la prevalencia en edades avanzadas. Más recientemente, Kingman y Albandar (2002) establecieron que los registros parciales tienden en general a subestimar la prevalencia, dependiendo de si se desprenden de un muestreo de la boca completa (como el CPI) o de algún sector de la boca en particular y del número de sitios examinados. Los autores señalan que debe tenerse especial atención en la interpretación de los hallazgos provenientes de estudios epidemiológicos de la EP que utilicen el CPI, ya que los errores sistemáticos pueden producir graves errores de clasificación de los individuos, con consecuentes falsos negativos que pueden distorsionar las asociaciones de la EP con otros problemas de salud.

Un estudio en adultos del año 2007 en el que se compararon varios protocolos de registros parciales, incluido el CPI, confirmó la sobreestimación en general (bolsas medias y severas) de la EP en un 25% con respecto al examen de boca completa (Vettore *et al.*, 2007). Este estudio coincide con otros en los que en general los dientes índice llevaron a una sobreestimación de la media de la profundidad de bolsas periodontales mayores de 4 mm.

Sin embargo, al examinar individuos jóvenes o una población en que la prevalencia era baja, los registros de bolsas moderadas y profundas fueron similares a los de boca completa (Agerholm & Ashley, 1996). Otro estudio en adolescentes y adultos jóvenes, del año 2012, concluyó que el CPI subestimaba la prevalencia de EP (Peres *et al.*, 2012).

En conclusión, parece ser que en adultos este indicador sobreestima la prevalencia de enfermedad periodontal (excepto las bolsas incipientes). Sin embargo, si la prevalencia es baja o se trata de individuos jóvenes, la prevalencia de bolsas incipientes y profundas es similar al de boca completa.

CPI (OMS, 2013)

La última publicación de la OMS con respecto al CPI data del año 2013 e incluye algunas modificaciones. Mide hemorragia gingival, bolsas periodontales y pérdida de inserción periodontal (ya no considera el sarro). Se sigue utilizando la misma sonda.

No se divide la boca en sextantes, salvo para la medición de la pérdida de inserción periodontal (PIP). La medición de sangrado y bolsas periodontales se realiza en todos los dientes presentes en boca, menos los terceros molares. Se miden seis puntos en cada diente (tres vestibulares y tres palatinos o linguales) y se anota la medida del sitio que tiene la peor condición.

Ejercicio 10.1

Del conjunto de los pacientes que concurren a sección Registro de la Facultad de Odontología se seleccionó un grupo de adultos de 35 a 44 años en los cuales se estudió la periodontitis y se obtuvieron los datos que se ven en los cuadros 10.1 y 10.2.

Calcule la prevalencia de:

- Periodontitis (definida como bolsa + PIP al menos en un sextante) y de bolsas incipientes y profundas.
- Calcule el promedio de bolsas incipientes y profundas por sextante en esta población.
- Interprete los resultados obtenidos y realice un informe epidemiológico.

Cuadro 10.1 Medidas de bolsa periodontal en un grupo de adultos de 35 a 44 años

id	Sextante 1	Sextante 2	Sextante 3	Sextante 4	Sextante 5	Sextante 6
1	más de 6 mm	4-6 mm	hasta 4 mm	más de 6 mm	hasta 4 mm	4-6 mm
2	hasta 4 mm	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	4-6 mm
3	más de 6 mm	hasta 4 mm	más de 6 mm	4-6 mm	hasta 4 mm	4-6 mm
4	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm
5	más de 6 mm	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	más de 6 mm
6	hasta 4 mm	más de 6 mm	más de 6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm
7	hasta 4 mm	hasta 4 mm	más de 6 mm	hasta 4 mm	4-6 mm	hasta 4 mm
8	hasta 4 mm	más de 6 mm				
9	más de 6 mm	4-6 mm	más de 6 mm	más de 6 mm	4-6 mm	hasta 4 mm
10	hasta 4 mm	4-6 mm	4-6 mm	4-6 mm	más de 6 mm	4-6 mm
11	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	más de 6 mm
12	hasta 4 mm					
13	hasta 4 mm	hasta 4 mm	4-6 mm	4-6 mm	hasta 4 mm	4-6 mm
14	más de 6 mm	más de 6 mm	hasta 4 mm	hasta 4 mm	más de 6 mm	más de 6 mm
15	hasta 4 mm	hasta 4 mm	hasta 4 mm	4-6 mm	hasta 4 mm	hasta 4 mm
16	hasta 4 mm					
17	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm
18	más de 6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm
19	hasta 4 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	más de 6 mm	hasta 4 mm
20	4-6 mm	hasta 4 mm	hasta 4 mm	hasta 4 mm	más de 6 mm	4-6 mm

Cuadro 10.2 Medidas de pérdida de inserción periodontal en un grupo de adultos de 35 a 44 años

id	Sextante 1	Sextante 2	Sextante 3	Sextante 4	Sextante 5	Sextante 6
1	profunda	profunda	sano	profunda	incipiente	sano
2	sano	sano	sano	profunda	incipiente	profunda
3	sano	sano	sano	incipiente	sano	sano
4	incipiente	sano	incipiente	sano	sano	incipiente
5	incipiente	sano	incipiente	incipiente	sano	sano
6	sano	sano	incipiente	sano	sano	sano
7	incipiente	profunda	profunda	profunda	sano	sano
8	profunda	sano	sano	sano	sano	profunda
9	profunda	profunda	sano	sano	sano	sano
10	sano	incipiente	sano	profunda	incipiente	sano
11	sano	incipiente	sano	sano	sano	sano
12	profunda	profunda	sano	sano	sano	profunda
13	profunda	sano	sano	incipiente	profunda	sano
14	incipiente	sano	incipiente	sano	profunda	incipiente
15	incipiente	sano	profunda	incipiente	incipiente	sano
16	sano	profunda	sano	sano	sano	sano
17	sano	sano	sano	incipiente	sano	profunda
18	profunda	sano	incipiente	incipiente	sano	sano
19	sano	profunda	sano	sano	sano	sano
20	sano	profunda	sano	sano	sano	incipiente

Abordaje y uso de los índices de EP en la clínica

El examen periodontal completo incluye: valoración de inflamación gingival, pérdida de inserción, profundidad de sondeo, aumento de corona clínica – recesión gingival, biofilm y sarro supragingival; y se complementa con la valoración de la reabsorción ósea, la movilidad y la presencia de furcas (Page & Eke, 2007).

Se presentarán a continuación los indicadores más usados en la clínica.

Indicadores de valoración del periodonto superficial

Índice de placa (O'Leary, 1972)

Cada diente erupcionado es subdividido en cuatro sectores: vestibular, lingual, mesial y distal. Se aplica entonces una solución reveladora y se determinan las superficies teñidas en la unión dentogingival.

Es un indicador dicotómico e identifica la presencia de biofilm con 1 y la ausencia con 0. En la ficha clínica periodontal se registrarán solo las superficies coloreadas. Al final se obtiene un porcentaje de superficies que presentan biofilm:

$$\text{total de superficies con biofilm} = \frac{\text{superficies con score 1}}{\text{total de superficies examinadas}} \times 100$$

Índice de sangrado gingival (Ainamo & Bay, 1975)

Como simplificación del índice de inflamación gingival de Silness y Løe (1964), Ainamo y Bay (1975) desarrollaron un indicador dicotómico (ausencia 0 / presencia 1), a partir de un sondaje suave del surco gingival. Cada diente erupcionado se divide en cuatro superficies (mesial, vestibular, lingual y distal). El porcentaje total se obtiene a partir de la siguiente fórmula:

$$\text{total de superficies con sangrado al sondaje} = \frac{\text{superficies con score 1}}{\text{total de superficies examinadas}} \times 100$$

Sarro supragingival (O'Leary, 1967)

En este caso la ficha clínica contempla la valoración por vestibular o lingual/palatino, 0 y 1 según corresponda. El cálculo del total de superficies se obtiene a partir de la siguiente fórmula:

$$\frac{\text{total de superficies con sarro supragingival}}{\text{total de superficies examinadas}} = \frac{\text{superficies con score 1}}{\text{total de superficies examinadas}} \times 100$$

Indicadores para la valoración del periodonto profundo

En este caso consideramos la pérdida de inserción (PIP) y la profundidad de sondeo (PS) a partir de una sonda periodontal milimetrada. La diferencia entre ambas medidas se conoce como *aumento de corona clínica* (en adelante ACC).

De los indicadores mencionados, la PIP representa la mejor estimación de enfermedad y se calcula de forma indirecta, a partir de la suma de la PS + ACC de cada sitio examinado (Research, Science and Therapy Committee, 2003).

En este caso todos los dientes erupcionados deben ser examinados en seis sitios: mesiovestibular, vestibular, distovestibular, mesiolingual/palatino, lingual/palatino, distolingual/palatino (Susin *et al.*, 2005).

En 1999 se presentó la clasificación de las EP, donde se establece que el cuadro destructivo más común es la periodontitis crónica (Armitage, 1999).

Definir cuándo una persona tiene periodontitis es complejo (Savage *et al.*, 2009). La bolsa periodontal determina el diagnóstico de periodontitis. Sin embargo, con un solo sitio en el que se mida PS mayor o igual a 4 mm no se llega con exactitud al diagnóstico periodontal. Debe haber por lo menos dos superficies con PS mayor o igual a 4 mm y que no sean sobre el mismo diente, además de conjugar hallazgos clínicos y radiográficos (Eke *et al.*, 2012).

La clasificación de *localizada* se determina a partir de la existencia de hasta 30% de sitios con PI; por encima de este valor se considera *generalizada*.

La severidad se establece de la siguiente forma, a partir de la PIP, leve = 1-2 mm, moderada = 3-4 mm, severa \geq 5 mm (Armitage, 1999). En este caso, independientemente de la cantidad de sitios, primará el *score* de mayor grado por sobre el resto.

Si bien tanto para la extensión como para la gravedad partimos de medidas registradas en sitios, nuestro diagnóstico final será sobre el individuo.

Caries dental

Breve descripción clínico-epidemiológica

La caries dental es el mayor problema de salud pública en la mayoría de los países industrializados, afecta a entre un 60 y 90 % de los escolares y a la mayoría de los adultos (Petersen, 2003).

Desde el punto de vista fisiopatológico, es una enfermedad compleja y dinámica localizada en la superficie dentaria, caracterizada por un desequilibrio en los procesos de desremineralización, producido por la fermentación bacteriana de los carbohidratos de la dieta. Si a lo largo de un período de tiempo predominan los momentos de pérdida mineral de los tejidos duros dentarios hacia el biofilm y la saliva, esta pérdida resultará en el establecimiento de una lesión de caries (Fejerskov, 2004; Kidd & Fejerskov, 2004).

Diversos factores intervienen para determinar el proceso de la caries dental, algunos actúan directamente en la superficie del diente (saliva, biofilm, dieta y flúor), mientras que otros influyen en la forma en que las personas actúan (conocimiento, actitud, comportamiento, grado de escolaridad, nivel socioeconómico, renta) (Fejerskov & Manji, 1990).

Para entender la causalidad de caries dental en una población se debe ampliar el concepto de *causalidad* para que incluya las variables no biológicas, estas no deben ser consideradas como factores de confusión de relaciones causales, sino como partes de una red de causas posibles. En forma breve, podemos decir que la caries dental se desarrolla y progresa como resultado de estructuras sociales que desencadenan comportamientos y reacciones psicológicas, que a su vez aceleran el proceso biológico (Baelum *et al.*, 2008).

Por lo tanto, desde el punto de vista de la epidemiología y la salud colectiva, puede considerarse una enfermedad crónica, explicada a través del análisis de la dieta, la exposición a fluoruros y patrones de conducta, que se manifiesta con cambios microbiológicos en la composición del biofilm y es afectada por el flujo salival (Selwitz *et al.*, 2007).

Los factores socioculturales mencionados pueden considerarse integrantes de los determinantes de la salud (Dahlgren & Whitehead, 2007), estos contextualizan y permiten explicar las diferentes patologías bucales, incluida la caries dental. Y es por esto que el proceso de la enfermedad debe ser sujeto a un control permanente a lo largo de la vida, con el fin de evitar consecuencias irreversibles (Maltz *et al.*, 2010).

Para entender la medición de la caries dental en las poblaciones (Alves *et al.*, 2016) es necesario explicar algunos conceptos particulares de esta enfermedad bucal y repasar conceptos estadísticos como la *unidad de medida* y las *medidas de frecuencia* de caries dental (capítulo 9).

Criterios para el diagnóstico de caries

Dado que la caries dental ocurre en forma continua, desde la desmineralización más temprana hasta la cavidad, es importante manejar criterios claros para las condiciones bajo las cuales se considera que la enfermedad está presente. Esto permitirá tomar las mejores decisiones de tratamiento y lograr así el mejor resultado de salud para el paciente. Estos criterios deberán describir las etapas clínicamente relevantes del proceso de la enfermedad.

Para caries dental existen múltiples conjuntos de criterios, muchos difieren en su umbral diagnóstico. A la hora de considerarlos deberá tenerse en cuenta el objetivo del diagnóstico asociado; por ejemplo, en ciertas ocasiones el objetivo no es el tratamiento inmediato, sino un diagnóstico epidemiológico (poblacional) que podrá tener diferentes propósitos (como una primera aproximación al problema y la planificación de acciones de salud colectiva).

El *umbral diagnóstico* de la caries dental es la fase de su historia natural a partir de la cual se considera a una pieza o superficie dentaria, efectivamente, con caries. Existe una gran controversia sobre aquello que debe considerarse *lesión de caries*. La polémica principal se da con el registro de las lesiones iniciales que aún no se han convertido en cavidad. Estas lesiones iniciales aparecen como fisuras descoloridas y sin pérdida de sustancias, o como una mancha blanca en superficies lisas visibles, o como sombras interproximales detectables radiográficamente. El problema es que no todas estas lesiones sin cavidad progresan a lesiones con cavidad en dentina (que requieren tratamiento quirúrgico restaurador), de hecho, con un tratamiento de control de mínima intervención (o incluso sin él), una buena proporción de ellas se mantendrán estáticas o incluso se remineralizarán (Pitts, 1993).

Tradicionalmente los investigadores europeos han registrado caries en una escala que se extiende a través un rango completo de la enfermedad, desde la lesión inicial más temprana detectable hasta la lesión con la afectación pulpar (Pitts & Fyffe, 1988). Esta es conocida como la escala C_1 - C_4 (tabla 11.1) y se basa en criterios publicados por la OMS en 1979 (WHO, 1979).

Empero, los investigadores del Reino Unido y Estados Unidos, entre otros, han registrado la caries como presente o ausente. Este criterio dicotómico se basa en una inspección visual (en las mejores condiciones posibles de limpieza e iluminación del diente) realizada por un examinador entrenado y calibrado. Se diagnostica caries cuando la lesión afecta la dentina (cavidad franca), es decir C_3 (Cortés, 1999).

La OMS actualmente utiliza un criterio similar al dicotómico: registra un diente como cariado «cuando en una lesión en una fosa o fisura, o superficie dental lisa, se observa la presencia de una cavidad evidente, un socavado en el esmalte, o un reblandecimiento en el tejido dentario de las paredes o piso de la cavidad» (WHO, 2013), dando más importancia a la presencia de cavidad que a la penetración en la dentina. La OMS propone el uso de la sonda CPI para confirmar la evidencia de caries en la superficie dentaria. Señala, también, que ante la duda de la existencia de caries debe optarse por considerar sana la pieza.

Tabla 11.1 Comparación de los criterios para el diagnóstico de caries (escala C₁-C₄ y escala dicotómica)

Diagnóstico a través del rango completo de caries Escala C ₁ -C ₄	Diagnóstico desde la lesión de dentina (C ₃) Escala dicotómica
<p>0: <i>Sano</i>, no hay evidencia de caries clínica tratada o no tratada (ligera tinción permitida en una fisura que de otro modo sería sana).</p>	
<p>C₁: <i>Caries inicial</i>, no hay pérdida clínicamente detectable de sustancia. Para fosas y fisuras: puede haber manchas significativas, decoloración o puntos ásperos en el esmalte que no atrapan al explorador, pero la pérdida de sustancia no puede ser diagnosticada con certeza. Para superficies lisas: pueden ser áreas blancas y opacas con pérdida de brillo.</p>	<p>0: <i>Sano</i>, no hay evidencia de caries clínica tratada o no tratada (ligera tinción permitida en una fisura que de otro modo sería sana).</p>
<p>C₂: <i>Caries de esmalte</i>, pérdida evidente de sustancia dental en fosas o fisuras, o en superficies lisas, sin piso o pared blanda o esmalte socavado. La textura del material dentro de la cavidad puede ser calcárea o desmenuzable, pero no hay evidencia de que la cavitación haya penetrado en la dentina.</p>	
<p>C₃: <i>Caries de dentina</i>, piso blando detectable, esmalte debilitado o una pared ablandada, o el diente tiene un relleno temporal. En superficies proximales la punta del explorador debe ingresar a la lesión fácilmente.</p>	<p>1: En fosas y fisuras de superficies oclusales o superficies lisas, cuando al pasaje suave del explorador se capta una discontinuidad y se presenta uno o más de los siguientes:</p>
<p>C₄: <i>Involucración pulpar</i>, cavidad profunda con probable afectación de la pulpa. Esta lesión no debe sondearse (generalmente en el análisis de datos de esta se incluye con C₃).</p>	<p>a. Ablandamiento en la base del área. b. Opacidad adyacente al área que proporciona evidencia de socavamiento o desmineralización. c. Ablandamiento del esmalte en el área adyacente que el explorador puede raspar.</p>

Recientemente, en 2002, en un intento por conseguir un criterio diagnóstico de caries universal, un grupo de reconocidos investigadores propuso un sistema de evaluación y detección de caries internacional (ICDAS, por su sigla en inglés). A partir de varios estudios clínicos y epidemiológicos este sistema registra seis puntos de escala ordinal sobre el grado de profundidad de las lesiones (yendo de lesiones superficiales hasta las más profundas) (Pitts, 2004), combinados con una escala dicotómica que evalúa la progresión de caries dental (activa/inactiva) (Nyvad *et al.*, 1999). Este método necesita un meticuloso entrenamiento y calibrado de los examinadores, así como muy buena iluminación, un campo operatorio seco y un examen clínico más largo. Como resultado, las puntuaciones del índice CPO donde la caries se mide sobre una gama completa de estadios son más altas que aquellos que solo consideran el componente C desde la presencia de cavidad en dentina (desarrollado más adelante en este capítulo). Los requerimientos logísticos mencionados pueden dificultar su utilización en relevamientos poblacionales.

Sin embargo, aún no existe consenso sobre si el registro de lesiones sin cavidad debe ser utilizado en las encuestas de vigilancia epidemiológica. Algunos autores

argumentan que, al llevarse a cabo la vigilancia en intervalos de varios años, sería posible saber si la caries dental está aumentando o disminuyendo a través del registro de las lesiones más avanzadas. A su vez recomiendan solo la medición de lesiones con cavidad, porque el método diagnóstico es más simple y fiable para los examinadores, además no requiere de un campo seco e iluminado, por lo que resulta de menor costo (Burt *et al.*, 2008). Por otro lado, cuando el objetivo es fundamentalmente monitorear medidas preventivas poblacionales de salud bucal (por ejemplo, el impacto de un programa de fluoración del agua), se argumenta que las lesiones iniciales dan cuenta de todo el proceso de caries, y en períodos más cortos estas sí podrían captar la variación de las estimaciones de caries de una población con una prevalencia más baja de la enfermedad, que sería imposible de captar si solo se toman en cuenta las lesiones avanzadas. Por lo tanto, sería más adecuado registrar ambos tipos lesiones (con y sin cavidad).

Si bien la discusión persiste en la literatura, trabajando en salud colectiva se entiende que la decisión final del investigador (incluir o excluir las lesiones iniciales o moderadas de la caries dental) debe basarse en el objetivo del estudio y los recursos disponibles, por lo que cobra especial importancia la identificación del criterio diagnóstico para definir el componente cariado C del CPO en el estudio.

Unidades de medida de la caries dental

La medición de caries dental puede ser realizada en diferentes unidades de medida (Chaves, 1962), pueden verse simplificadas en la figura 11.1:

1. *Unidad individuo*: Podemos tomar a los miembros de una población y clasificarlos en dos grupos: los individuos que presentan o presentaron caries en el pasado y los que no presentan ni presentaron caries.
2. *Unidad diente*: Podemos contar los dientes, y clasificar la población en varios grupos: los que no tienen y nunca tuvieron dientes cariados, y aquellos que tienen o tuvieron 1, 2, 3 o más dientes con caries.
3. *Unidad superficie dentaria*: Podemos contar las superficies dentarias y clasificarlas en varios grupos: las que no tienen y nunca tuvieron caries y aquellas que tienen o tuvieron 1, 2, 3 o más superficies con caries.
4. *Unidad lesiones*: Podemos contar las lesiones que presenta cada individuo, esto es muy familiar para el paciente, que generalmente pregunta «¿cuántas [lesiones de] caries tengo?». Pero lo más interesante de clasificar la población

no es la cantidad de lesiones presentes, sino la severidad de estas, puesto que la cantidad de lesiones suele coincidir con las superficies afectadas en la mayoría de los casos.

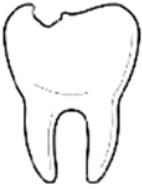
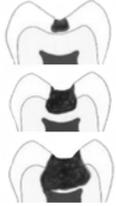
INDIVIDUO	DIENTE	SUPERFICIE DENTARIA	LESIONES/ SEVERIDAD
			
1 diente con caries	1 diente cariado	1 superficie cariada	1 lesión moderada y 2 lesiones severas

Figura 11.1 Unidades de medida para caries dental (modificado de Chaves, 1962).

Principales indicadores epidemiológicos de caries dental

Índice CPO

El índice CPO es el más utilizado para medir la enfermedad caries dental. Fue descrito por Klein y Palmer en 1937 y adoptado por la OMS (Cortés, 1999). Este índice mide el presente y el pasado de la caries dental en un individuo o una población, y puede aplicarse en la dentición permanente (CPO) o, luego de las modificaciones introducidas por Gruebbell en 1944, en la dentición decidua (CEO).

C refiere a dientes con lesiones de caries dental no tratadas, P indica la pérdida dentaria por causa de caries dental y O cuenta las piezas con obturación o restauración como consecuencia de caries dental. Para la dentición temporaria, en el componente P antiguamente se registraba la extracción indicada por caries y solo debe usarse cuando el niño se encuentra en una edad en la que la exfoliación normal no sería una explicación suficiente para la ausencia de la pieza dental y se registra con la letra *e* (WHO, 2018).

El índice CPO es el resultado de la suma de estos valores. Cuando la unidad de observación es el diente (tabla 11.2), el índice se expresa como CPO-D o ceo-d,

entretanto si la unidad observada es la superficie dentaria, el índice será CPO-S o ceo-s (Arana *et al.*, 2005).

Tabla 11.2 Índices de caries (modificado de Antunez & Peres, 2013)

Índice	Unidad	Objeto	Componentes		
ceod	dientes	dentición temporaria	«cariados»	«extraído por caries»	«obturados»
ceos	superficies dentarias	dentición temporaria	«cariados»	«extraído por caries»	«obturados»
CPO-D	dientes	dentición permanente	«cariados»	«perdido por caries»	«obturados»
CPO-S	superficies dentarias	dentición permanente	«cariados»	«perdido por caries»	«obturados»

En el caso de un individuo adulto, el índice puede adoptar valores de 0-28 (excluyendo terceros molares) cuando la unidad es el diente (CPO-D), o valores de 0-128 cuando la unidad es la superficie dentaria (CPO-S) (12 dientes anteriores con 4 superficies y 16 dientes posteriores con 5 superficies).

En una población el CPO o ceo se calculan como el promedio del grupo, por lo que pueden asumir valores decimales (figura 11.2), o sea la suma de los valores de cada individuo dividido entre el número total de individuos examinados (Burt *et al.*, 2008).

Además es un índice irreversible, una vez producida la lesión de caries dental, esta no revertirá, manteniéndose en ese estadio (podrá ser obturada o extraída). En consecuencia, el índice puede incrementar o permanecer estable, y con la variación de la constitución de los componentes ya afectados (cariado, perdido u obturado) no variará el valor total del CPO (Arana *et al.*, 2005).

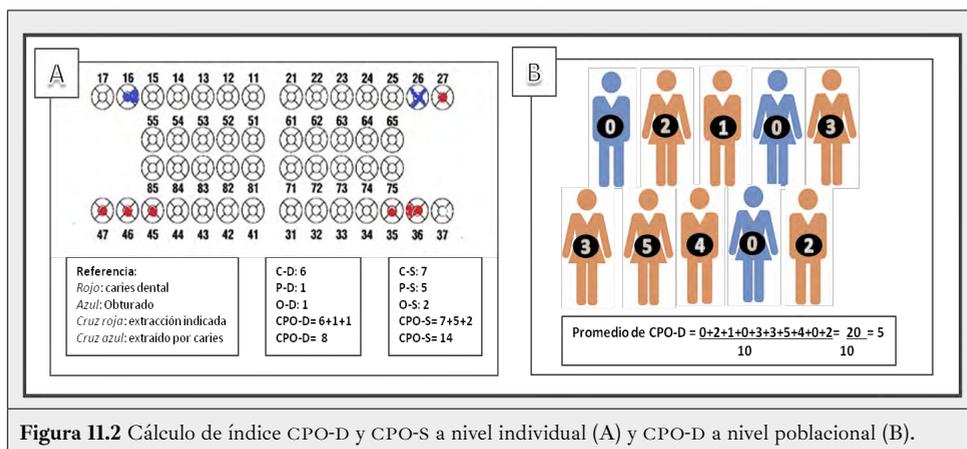


Figura 11.2 Cálculo de índice CPO-D y CPO-S a nivel individual (A) y CPO-D a nivel poblacional (B).

Limitaciones del índice CPO

El CPO es simple, versátil, estadísticamente manejable y fiable cuando los examinadores han sido entrenados y calibrados. Sin embargo, presenta limitaciones:

1. Sus valores no están relacionados con el número de dientes en riesgo, un valor de CPO es un recuento de aquellos dientes que el examinador juzgó afectados por caries, no tiene denominador. El valor de CPO, por tanto, no muestra directamente la intensidad del daño. Por ejemplo, en un niño de 7 años de edad con un valor de $CPO-D = 3,0$ y con 9 dientes permanentes en su boca, resulta que el 33% de sus dientes han sido atacados por la caries dental en un corto tiempo. Mientras que en un adulto con 32 dientes y con el mismo $CPO-D = 3,0$, el valor presenta apenas un 11% de los dientes afectados, y en un período de tiempo mucho más largo.
2. En individuos adultos su valor se satura rápidamente, sobre todo si se reporta sobre la unidad diente (si el individuo presenta un $COP-D = 32$ y en un período posterior presenta una nueva lesión, esta no podrá ser registrada).
3. El CPO no mide la gravedad de la enfermedad caries, da el mismo valor a un diente con una lesión leve como a uno con lesión severa.
4. El CPO da el mismo valor a las piezas dentarias con lesiones de caries no tratadas, a las piezas perdidas y a las obturadas; no discrimina entre la historia y el presente de la enfermedad.
5. Los datos del CPO no expresan la funcionalidad del sistema estomatognático. Si el individuo presenta 3 piezas perdidas y estas están en el mismo sextante, el individuo podrá tener una masticación deficiente. En este sentido el CPO también es de poca utilidad para la estimación de las necesidades de tratamiento (el índice no fue diseñado para tal fin).
6. Cuando contamos los dientes perdidos en el componente P solo es válido contar aquellas piezas perdidas a causa de caries dental. Los dientes pueden perderse por razones periodontales en adultos mayores, y por razones de ortodoncia en adolescentes. Se requieren entonces reglas de decisión para determinar qué hacer frente a estos casos.

7. El índice CPO puede sobrestimar la experiencia de caries en dientes a través de restauraciones preventivas y sellantes de fosas y fisuras. En un estudio epidemiológico existe el riesgo de contar estas restauraciones dentro del componente O del CPO, a pesar de que estos dientes son sanos. En estos casos, el valor del CPO será inflado (Burt *et al.*, 2008). Las restauraciones de resina compuesta que se han colocado por razones estéticas y los sellantes no deberían ser incluidos en el CPO, deben ser contabilizados por separado.

8. Los valores de CPO no pueden compararse de un grupo a otro sin tener en cuenta el criterio diagnóstico de caries, ya que no existe un criterio universal de lo que es un diente con caries dental. La comparación entre grupos donde caries dental fue registrada como lesión cavitada con otro donde se registró desde lesiones sin cavidad es claramente inválida (a menos que el estudio que utilizó el criterio no cavitado hubiera registrado separadamente el número de dientes con lesiones cavitadas, en cuyo caso se podrían comparar los dientes con lesiones cavitadas en ambos estudios).

9. Desde el punto de vista estadístico, cuando se trabaja a nivel poblacional (promedio de dientes C, P y O) debe tenerse en cuenta la distribución de los valores de la variable en la población estudiada, ya que muchas veces los valores no presentan una distribución homogénea y presentar solo el valor promedio podría dar una idea equivocada del problema en la población. En estos casos resulta útil acompañar el valor promedio del CPO-D con el del desvío estándar y también calcular el valor de la mediana del CPO-D.

10. El CPO debería presentarse siempre con sus componentes discriminados. Es cuestionable usar un índice para la enfermedad que es tan dependiente de los juicios de tratamiento de muchos profesionales. También es cuestionable que combine un tratamiento previo (componentes P y O) con necesidad de tratamiento actual (componente C), esto no se usa en otras áreas de la vigilancia en salud pública (Burt *et al.*, 2008).

A pesar de estas limitaciones, con algunas especificaciones el CPO sigue siendo el principal índice utilizado para expresar el estado de caries dental en una población.

Índice significativo de caries (SIC)

El índice significativo de caries (SIC, por su acrónimo en inglés) fue introducido por Brathall (2000) con el fin de llamar la atención sobre las personas con los valores más altos de caries dental en cada población. El SIC es una nueva forma de presentación de los datos del CPO-D; da una imagen más precisa de la distribución asimétrica de caries dental en la población, especialmente en los países desarrollados —que habitualmente llegan a la conclusión errónea de que la situación de caries está controlada en toda la población, mientras que un grupo carga con la enfermedad— (Burt *et al.*, 2008). El SIC se calcula clasificando a los individuos de acuerdo a sus valores CPO-D, se ordena del valor menor al mayor, se divide la población en tercios, se selecciona el tercio de la población con los puntajes más altos de caries y se calcula el promedio CPO-D de ese subgrupo (Brathall, 2000).

El SIC advierte a las autoridades de las necesidades especiales de intervención preventiva y de tratamiento de caries dental en el subgrupo con mayores niveles de enfermedad.

Al ser una extensión del índice CPO, presenta las mismas desventajas que aquel. Es importante considerar que este índice es más significativo cuando la prevalencia de la caries dental es baja y tiene distribución asimétrica (por ejemplo, distribución de frecuencias con cola a la derecha, que es la que predomina hoy en el mundo), es en esos casos que resulta mayor la diferencia entre el promedio de CPO y SIC (Burt *et al.*, 2008; Mehata, 2012).

Índice de cuidados odontológicos

El índice de cuidados odontológicos fue propuesto por Walsh (1970) y para su cálculo son utilizados los componentes del CPO-D. Este índice resulta del cociente de la proporción del número de dientes obturados (componente O) y el número total de dientes con experiencia de caries (componentes C, P y O). Se aplica para evaluar y comparar la efectividad de programas de atención odontológica.

Es un índice que no está definido para la evaluación individual, es decir, se usa solamente para datos agregados. Su valor procura reflejar la capacidad del sistema de salud para dar respuesta a las necesidades de tratamiento odontológico de grupos específicos.

Otros índices (OSD y DS)

En un intento de medir la salud oral y no la enfermedad, Sheiham y colaboradores en 1987 propusieron dos índices nuevos mediante la ponderación de los dientes sanos y obturados, uno sobre la pérdida dentaria y el otro sobre los dientes cariados. El primero toma en cuenta el *número de dientes funcionales* (suma de dientes obturados y dientes sanos) y asume un valor de 0-28 (excluyendo terceros molares). A este índice de dientes obturados y sanos se lo identifica con la sigla FS-T (*filled and sound teeth*; en español sería OSD, obturados y sanos dientes).

El segundo busca medir la cantidad de tejido dental sano y asume un valor de 0-128. Toma en cuenta el *número de dientes equivalente a sanos* y va por la sigla DS (T-Health es la denominación anglosajona). Para calcular el DS se le asigna un peso numérico a cada diente dependiendo de su estado: 4 (diente sano), 2 (diente obturado), 1 (diente con caries). Al dar peso al diente saludable, este índice intenta medir la influencia de la prevención primaria.

Cabe señalar que el CPO-D aumenta con la experiencia de caries, mientras que los indicadores OSD y DS disminuyen según aumenta la experiencia de caries (Benigeri *et al.*, 1998).

Schuller y Holst (2001) realizaron un estudio para examinar la relación entre CPO-D y OSD en diferentes poblaciones. Demostraron que el CPO-D es más adecuado para describir la variación en poblaciones con menores niveles de enfermedad, mientras que el OSD es más adecuado para describir la variación en poblaciones con altos niveles de enfermedad.

En conclusión, estos dos índices (OSD y DS) con el enfoque de medir la salud bucal y la funcionalidad (en lugar de la enfermedad) merecen más atención de la que han recibido. En este sentido, se recomienda utilizar tanto el CPO-D como el OSD para describir el estado bucal de una población (Burt *et al.*, 2008).

International Caries Detection and Assessment System (ICDAS)

Como mencionamos anteriormente, el ICDAS (sistema internacional de evaluación de detección de caries) fue creado en un intento de proponer un sistema de detección de caries internacionalmente aceptado. Su versión inicial fue creada en 2002 por un grupo de cariólogos y epidemiólogos (Pitts, 2004), y luego de pequeños ajustes su versión final fue publicada en el 2007 (Ismail *et al.*).

El ICDAS utiliza un sistema de identificación de tres dígitos [X-Y-Z] para el registro de las superficies dentarias. El primer lugar, X, se dedica al estado de las superficies (no restaurada, sellada, obturada). El segundo lugar, Y, se dedica a la presencia y profundidad de las lesiones de caries, abarcando desde los primeros cambios visuales en el esmalte hasta la cavitación extensa (tabla 11.1).

Tabla 11.1 Y – Descripción de los criterios de profundidad del ICDAS

Puntuación	Criterio
0	Sano. Luego de un breve secado (5 segundos) no hay evidencia de actividad de caries (cambio cuestionable en la translucidez, mancha blanca o acastañada del esmalte). Se registrarán como sanos los siguientes defectos: hipoplasias del esmalte, fluorosis, desgaste dental (desgaste, abrasión y erosión) y manchas extrínsecas o intrínsecas.
1	Primer cambio visual del esmalte. Luego de un breve secado (5 seg.) se observa una opacidad blanca o acastañada que no es consistente con la apariencia clínica del esmalte sano (diferente, además, de las características mencionadas en el código 0). En superficies oclusales este cambio de color se limita al fondo de las fosas y fisuras.
2	Cambio visual evidente del esmalte. Incluso con el diente húmedo se observa una opacidad blanca o acastañada. En superficies oclusales este cambio de coloración no se difunde a las paredes laterales de las fosas y fisuras.
3	Discontinuidad del esmalte sin exposición de dentina o sombra subyacente. Incluso con el diente húmedo se observa opacidad blanca o acastañada. Luego de un breve secado (5 seg.) se observa una pérdida de la estructura dentaria. Esto se verá como una microcavidad (pero la dentina no es visible en las paredes o en la base de la cavidad). En las superficies oclusales la entrada de las fosas y fisuras se ve más ancha de lo normal.
4	Sombra oscura subyacente de la dentina con o sin ruptura del esmalte. Esta lesión aparece como una sombra de dentina oscura visible a través de una superficie de esmalte (aparentemente intacta) que puede mostrar o no signos de rotura localizada. La apariencia de sombra a menudo se ve más fácilmente cuando el diente está mojado. El área oscurecida es una sombra intrínseca que puede aparecer en color gris, azul o marrón.
5	Cavidad detectable con dentina visible, involucra menos del 50% de la superficie. Cavitación en esmalte opaco o descolorido que expone la dentina subyacente. Luego de un breve secado (5 seg.) se observa evidencia de pérdida de la estructura dental: cavitación franca (coloración amarilla, marrón o marrón oscura). La sonda CPI puede usarse para confirmar la presencia de una cavidad aparente en la dentina.
6	Cavidad detectable extensa con dentina visible, involucra el 50% o más de la superficie. Se observa pérdida obvia de la estructura del diente, la cavidad es profunda y amplia y la dentina es claramente visible en las paredes y en la base. Una cavidad extensa implica al menos la mitad de la superficie de un diente y puede llegar a la pulpa.

El tercer lugar, Z, se dedica a la evaluación de la actividad de caries, utilizando el sistema Nyvad (tabla 11.2). Antes de comenzar con el examen clínico, los dientes deben limpiarse cuidadosamente y se debe contar con muy buena iluminación; además, las superficies dentarias deben estar secas (puede emplearse suavemente aire comprimido o gasas), y cuando se requiera verificación táctil solo podrá emplearse una sonda CPI para explorar la lesión (Braga *et al.*, 2009).

Tabla 11.2 Z – Características de la actividad de caries según Nyvad para códigos de profundidad de ICDAS

Código ICDAS	Signos de lesiones activas	Signos de lesiones detenidas
1, 2 o 3	La superficie del esmalte es blanco-amarillenta; opaca con pérdida de brillo, se siente áspera cuando la punta redonda de la sonda es pasada suavemente por la superficie. La lesión está en una zona de retención de placa, es decir, en la entrada de fosas y fisuras, cerca del margen gingival o, para las superficies proximales, por debajo o por encima del punto de contacto. La lesión puede estar cubierta por placa gruesa antes de la limpieza	La superficie del esmalte es de color blanco, café o negro. El esmalte puede ser brillante y se siente duro y liso cuando la punta redonda de la sonda es pasada suavemente por la superficie. Para superficies lisas, la lesión de caries normalmente se encuentra a cierta distancia del margen gingival. La lesión puede no estar cubierta por placa gruesa antes de la limpieza.
4, 5 o 6	La dentina se siente suave o con consistencia de cuero al sondaje suave.	La dentina es brillante y dura al sondaje suave.

Medidas de frecuencia y características relevantes para medición de caries dental

Si bien las medidas de frecuencia fueron mencionadas en el capítulo 8, la existencia de particularidades en la medición de caries amerita hacer algunas especificaciones.

Prevalencia

La *prevalencia* es la proporción de individuos afectados por una enfermedad en un momento dado, en nuestro caso se trata de individuos con por lo menos una lesión de caries, medida que se expresa con un valor de 0 a 100 %.

La prevalencia de caries en la odontología (figura 11.2B) cuantifica en una población a todos los individuos que tienen uno o más dientes afectados (por ejemplo, si de un grupo de 10 personas, 7 tienen caries, la prevalencia es del 70 %). Esta medida es obtenida a través de una única recolección de datos, en una unidad de medida, en un momento del tiempo (medida característica de los estudios

transversales). La prevalencia diferencia individuos afectados por la enfermedad e individuos libres de caries, sin embargo no es posible distinguir individuos afectados de forma leve, moderada o severa. Historicamente en odontología se empleó el índice de Knutson, que cuantifica el inverso de la prevalencia de caries, o sea, la proporción de individuos con CPO igual a 0 (libres de caries). En el ejemplo anterior el libre de caries es 3 de 10 personas, es decir el 30%.

Incidencia

La *incidencia* refiere a la velocidad de avance de la enfermedad. Representa la proporción de individuos que desarrollan la enfermedad en cierto período de tiempo, contabilizándose solo los nuevos casos desarrollados en ese período, medida que se expresa con un valor de 0 a 100 %.

La incidencia de la caries representa la proporción de individuos libres de caries que desarrollan por lo menos una lesión de caries a lo largo de un período de acompañamiento. Se obtiene a partir de por lo menos dos recolecciones de datos, en una unidad de medida, en diferentes momentos del tiempo (medida característica de los estudios longitudinales). Para considerar a un individuo como un nuevo caso (*caso incidente*), este debe presentarse como libre de caries en el momento de la primera recolección de datos, o sea, en riesgo de desarrollarla.

Severidad

La *severidad* se refiere al tipo de tejido dentario involucrado, así como a la presencia o no de cavidad. La severidad aumenta según el grado de involucramiento de los diferentes componentes del órgano dentino-pulpar.

De manera general, las lesiones de caries coronaria pueden clasificarse en lesiones no cavitadas, lesiones cavitadas en esmalte y lesiones cavitadas en dentina. Además, pueden ser clasificadas de acuerdo a su actividad (activas/inactivas).

Extensión (o experiencia)

La *extensión* es el número de dientes (índice CPO-D) o superficies (índice CPO-S) que presentan lesiones o tratamiento por caries. En las poblaciones esta medida suele calcularse como el promedio del índice CPO-D y CPO-S (ver índice CPO al inicio de la sección «Principales indicadores...»).

Ejercicios

Ejercicio 11.1

Los siguientes datos provienen de una guardería de una ONG para niños de hasta cinco años de edad, para la cual se está planificando un programa de atención a la salud bucal. Además de las lesiones tardías (lesiones de caries con cavidad, nivel c_3) se registraron lesiones iniciales (lesiones de caries sin cavidad, nivel c_{1-2}).

Pac	Edad	Sexo	c_{1-2}/d	c_3/d	e/d	o/d	$c_{1-2}eo/d$	c_3eo/d	presencia $/c_{1-2}eo$	presencia $/c_3eo$
1	2	M	3	2	0	0	3	2	1	1
2	3	M	1	0	0	0	1	0	1	0
3	4	F	4	2	2	1	7	5	1	1
4	2	M	1	0	1	1	2	3	1	1
5	4	F	0	0	0	0	0	0	0	0
6	3	F	1	0	0	0	1	0	1	0
7	3	F	2	1	0	1	3	2	1	1
8	2	M	0	0	0	0	0	0	0	0
9	3	F	1	1	0	0	1	1	1	1
10	4	F	3	1	1	2	6	4	1	1

- Calcule la prevalencia de caries de esta población (historia y presente de la enfermedad caries dental) a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías. ¿Cómo interpreta los resultados?
- Calcule la extensión de caries (historia y presente) a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías (criterio OMS). ¿Cómo interpreta los resultados?
- ¿Cuáles son los pacientes que presentan mayor c_3eo (criterio OMS)/d y $c_{1-2}eo/d$?

Ejercicio 11.2

Tomando en cuenta los datos del ejercicio 11.1.

- a) Calcule la extensión de caries presente (componente c) a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías. ¿Cómo interpreta los resultados?
- b) Calcule la extensión de tratamientos realizados (componente o) en la población.
- c) ¿Cómo interpreta los resultados?
- d) Prepare un reporte epidemiológico con los resultados.

Ejercicio 11.3

Tomando en cuenta los datos del ejercicio 11.1.

- a) Calcule el SIC (historia y presente) a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías (criterio OMS). ¿Cómo interpreta los resultados?
- b) Calcule la extensión de caries por sexo a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías (OMS). ¿Cómo interpreta los resultados?
- c) Calcule la extensión de caries según la edad a partir del criterio diagnóstico de lesiones iniciales y de lesiones tardías. ¿Cómo interpreta los resultados?
- d) Construya una tabla y un gráfico donde se describa la situación de caries (utilizando $c_{3eo/d}$) y cómo influye el sexo en ella.

Cáncer bucal

Breve descripción clínico-epidemiológica

El diagnóstico de cáncer de boca, en la mayor parte de los casos, es tardío. Cerca del 60 % de los pacientes llegan a los consultorios odontológicos o médicos con lesiones avanzadas (Seoane-Romero *et al.*, 2012). Hay dolor mínimo en la fase inicial, lo que puede explicar en parte la demora en buscar ayuda profesional. La falta de estrategias de despistaje (*screening*) impide el diagnóstico precoz.

El cáncer bucal aún tiene un pronóstico generalmente negativo, con valores de sobrevivencia de menos de 50 % a los cinco años para cáncer de lengua, cavidad bucal y oro-faringe, y valores notablemente mayores (90 %) para el cáncer de labio (Warnakulasuriya, 2010).

De acuerdo con la séptima clasificación de la Unión Internacional Contra el Cáncer (UICC), revisada junto a la American Joint Committee on Cancer (AJCC), el carcinoma espinocelular (CEC) de cavidad bucal se clasifica en cuatro estadios clínicos. Esta clasificación se basa en el sistema TNM que se considera el tamaño (T), las metástasis regionales en nódulos linfáticos (N) y las metástasis a distancia (M). Este sistema es mundialmente reconocido para describir la real extensión del tumor, su pronóstico y estrategias terapéuticas (Edge *et al.*, 2010).

Tabla 12.1 Descripción de cada clase del sistema de clasificación TNM

Tumor primario (T)	
TX	Tumor primario no puede ser evaluado.
T0	Sin evidencia de tumor primario.
Tis	Carcinoma <i>in situ</i> .
T1	Tumor de 2 cm o menos en su mayor dimensión.
T2	Tumor de entre 2 y 4 cm en su mayor dimensión.
T3	Tumor mayor a 4 cm en su mayor dimensión.
T4a	El tumor primario tiene una ligera invasión local de estructuras adyacentes.*
T4b	El tumor primario tiene una gran invasión de estructuras adyacentes.**
Metástasis regionales (N)	
NX	Nódulos linfáticos regionales no pueden ser evaluados.
N0	Sin evidencia de nódulos linfáticos regionales.
N1	Un linfonodo homolateral menor a 3 cm en su mayor dimensión.
N2a	Un linfonodo homolateral entre 3 y 6 cm en su mayor dimensión.
N2b	Múltiples linfonodos homolaterales de no más de 6 cm.
N2c	Linfonodos bilaterales o contralaterales de no más de 6 cm.
N3	Linfonodos de más de 6 cm en su mayor dimensión.
Metástasis a distancia (M)	
M0	No hay metástasis a distancia.
M1	Hay metástasis a distancia.

* Hueso cortical maxilar o mandibular, musculatura de lengua, seno maxilar, piel facial.

** Espacio masticatorio, placas pterigoideas, base de cráneo o carótida interna.

Tabla 12.2 Etapas anatómicas y grupos de pronóstico según el sistema de clasificación TNM

Estadio clínico	Clasificación TNM		
Estadio 0	Tis	N0	M0
Estadio I	T1	N0	M0
Estadio II	T2	N0	M0
Estadio III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Estadio IVA	T4a	N0	M0
	T4a	N1	M0
	T1	N2	M0
	T2	N2	M0
	T3	N2	M0
Estadio IVB	T4a	N2	M0
	Cualquier T	N3	M0
Estadio IVC	T4b	Cualquier N	M0
	Cualquier T	Cualquier N	M1

Indicadores de cáncer bucal

A pesar de que el cáncer es una enfermedad de baja frecuencia, presenta una gran mortalidad en todo el mundo; es por esto que las tasas de prevalencia de esta patología se expresan cada 100.000 personas. Las tasas que se utilizan para realizar las comparaciones entre países deben estandarizarse por edad, ya que la estructura etaria de las poblaciones generalmente difiere. Se utilizan las tasas estandarizadas por edad para expresar la prevalencia y la incidencia de la enfermedad (se muestra un ejemplo en la tabla 12.3).

En 2012 se reportaron cerca de 14,1 millones de nuevos casos y 8,2 millones de muertes por esta enfermedad. Se prevé que el número de nuevos casos aumente en aproximadamente un 70 % en los próximos veinte años (IARC/OMS, 2014) siendo, por lo tanto, necesaria una mejora en los abordajes diagnósticos y terapéuticos de esta enfermedad (Ferlay *et al.*, 2010).

El cáncer de cabeza y cuello (CCC) es la sexta neoplasia maligna más común y está asociada a gran morbilidad y mortalidad. El carcinoma espinocelular es más común dentro de las neoplasias malignas de cabeza y cuello (CECC) (Warnakulasuriya, 2009). En el año 2012 en el mundo hubo más de medio millón de casos de CECC, 274.000 de ellos en boca (WHO, 2008). Dentro de las CECC, tres de cada cuatro casos nuevos ocurren en países en desarrollo, en ellos esta neoplasia representa cerca del 3 % del total de cáncer (Jemal *et al.*, 2007). En 2012 se estimaron 325.000 muertes por CECC en el mundo (Winn *et al.*, 2015).

En Uruguay, en 2014, la Comisión Honoraria de Lucha Contra el Cáncer (CHLCC) publicó un relevamiento de cáncer en el país en el período 2007-2011. Ese estudio muestra que el cáncer oro-faríngeo posee mayor mortalidad en hombres ($n = 1079$ casos), ocupando el noveno lugar, con una mortalidad de 6,5 casos cada 100.000 individuos. Con respecto a la incidencia anual, en el período 2007-2011 se observó una tasa de incidencia de 10,33 casos por 100.000 habitantes. En mujeres, en el mismo período, fueron identificados 404 casos, con una tasa de incidencia de 2,76 casos por 100.000 habitantes (Barrios *et al.*, 2014). En la mujer los cánceres buco-faríngeos no entran en los primeros diez puestos de frecuencia.

Tabla 12.3 Incidencia, mortalidad y prevalencia a los 5 años de cáncer de labio y de cavidad oral

Region		Incidence		Mortality		5-year prevalence	
		N (%)	ASR (W)	N (%)	ASR (W)	N (%)	Prop
World	Men	198975 (2.7)	5.5	97940 (2.1)	2.7	467157 (3.1)	18.0
	Women	101398 (1.5)	2.5	47413 (1.3)	1.2	234992 (1.4)	9.1
	Both sexes	300373 (2.1)	4.0	145353 (1.8)	1.9	702149 (2.2)	13.5
WHO African Region	Men	80509 (3.0)	3.4	5026 (2.4)	2.2	18446 (3.9)	7.3
	Women	5475 (1.4)	2.0	3504 (1.4)	1.4	12766 (1.4)	4.9
	Both sexes	13484 (2.1)	2.7	8530 (1.9)	1.8	31212 (2.3)	6.1
WHO American Region	Men	31898 (2.2)	5.9	8532 (1.3)	1.5	94953 (2.5)	26.9
	Women	17302 (1.2)	2.6	4271 (0.7)	0.6	48526 (1.2)	13.1
	Both sexes	49200 (1.7)	4.2	12803 (1.0)	1.0	143479 (1.8)	19.9
WHO East Mediterranean Region	Men	11601 (4.4)	5.1	6185 (3.2)	2.8	27236 (5.9)	12.9
	Women	9080 (3.1)	4.1	4812 (2.7)	2.3	21570 (2.9)	10.7
	Both sexes	20681 (3.7)	4.6	10997 (3.0)	2.5	48806 (4.1)	11.8
WHO Europe Region	Men	45567 (2.3)	7.1	18642 (1.7)	2.9	118151 (2.5)	33.2
	Women	20366 (1.2)	2.4	6560 (0.8)	0.7	51933 (1.1)	13.4
	Both sexes	65933 (1.8)	4.6	25202 (1.3)	1.7	170084 (1.8)	22.8
WHO South East Asian Region	Men	70816 (8.7)	8.9	45247 (7.3)	5.7	122976 (9.9)	18.4
	Women	32648 (3.6)	4.0	20487 (3.7)	2.5	58034 (2.8)	8.9
	Both sexes	103464 (6.0)	6.4	65734 (5.6)	4.1	181010 (5.5)	13.7
WHO Western Pacific Region	Men	31013 (1.2)	2.7	14292 (0.8)	1.2	85233 (1.9)	11.4
	Women	16511 (0.9)	1.3	7776 (0.7)	0.6	42123 (0.9)	5.8
	Both sexes	47524 (1.0)	2.0	22068 (0.7)	0.9	127356	8.6

Fuente: Torre *et al.*, 2015.

Trastornos témporo mandibulares (TTM)

Breve descripción clínico-epidemiológica de los TTM y el bruxismo

McNeill (1993) define con la terminología genérica *trastornos témporo mandibulares* (TTM) a un grupo heterogéneo de desórdenes psicofisiológicos con características comunes de dolor oro-facial, disfunción masticatoria o ambas.

El bruxismo ha sido definido como una actividad muscular mandibular repetitiva caracterizada por el apretamiento o frotamiento de los dientes o por empuje mandibular (Lobbezoo *et al.*, 2013).

Si bien son entidades patológicas diferentes, el bruxismo puede ser considerado un factor predisponente, iniciador o perpetuante de los TTM (Okeson, 2013). El rol causal del bruxismo para TTM ha sido demostrado, y su abordaje incluye terapias conductuales, dispositivos intraorales, medicación, higiene del sueño y técnicas de relajación (Luiz de Barreto Aranha *et al.*, 2018).

Los estudios epidemiológicos de bruxismo y trastornos temporomandibulares a nivel mundial no difieren de los de nuestro país. Los valores pueden variar según la metodología utilizada en los diferentes estudios, pero siempre se refleja una alta prevalencia.

Según un estudio realizado en Uruguay (Riva *et al.*, 2011), padecer o haber padecido bruxismo tiene una prevalencia de un 71,95% en Montevideo y de un 62,17% en el interior. La prevalencia de «tener la sensación al despertar de haber dormido apretando los dientes» fue de 30,72% en Montevideo y 23,19% en el interior. Para complementar estos datos estadísticos, los autores obtuvieron información acerca de la autopercepción de estrés con la pregunta «¿Se considera usted una persona nerviosa, estresada?», se halló que más del 50% de los encuestados se consideraban nerviosos o estresados (Riva *et al.*, 2011).

Los TTM son afecciones que presentan alta prevalencia con respecto a la población general. La prevalencia de signos en una población europea es de un 34,5% aproximadamente (Jussila *et al.*, 2017). Otro estudio realizado en Australia mostró

una prevalencia de síntomas de TTM del 77,2% en una cohorte de estudiantes de odontología (Lung *et al.*, 2018).

Otro estudio de prevalencia muestra valores del 29,1%, donde el signo más prevalente fue el ruido articular (17,4%), seguido de movimiento anormal de mandíbula (12,3%). En este mismo estudio se encontró diferencia significativa entre géneros, siendo más prevalente en las mujeres (Xie *et al.*, 2018).

En una revisión publicada en 2018, con artículos donde la muestra se componía de niños y adolescentes de entre 10 y 19 años, utilizando el Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), se encontraron resultados que fluctúan entre el 7,3% y el 30,4%, donde el diagnóstico más prevalente fue el dolor miofascial, seguido del desplazamiento discal con reducción (Christidis *et al.*, 2019).

En Uruguay, el 55% de la población presenta por lo menos un síntoma de TTM y el 44% de la población presenta por lo menos un signo (Riva *et al.*, 2011). Estos estudios coinciden con los mencionados en señalar una mayor prevalencia en mujeres (Riva *et al.*, 2011).

Principales indicadores epidemiológicos

Indicadores de bruxismo

El polisomnograma ha sido tomado como patrón oro en investigación para diagnóstico de bruxismo del sueño, aunque se podría considerar un examen de difícil lectura para el odontólogo general y poco accesible para los pacientes.

El autoreporte de bruxismo extraído de la anamnesis a través de las siguientes preguntas ha sido asociado también a presencia de TTM del tipo *dolorosos*:

- ¿Se despierta con sensación de haber dormido apretando los dientes?
- ¿Se despierta con sensación de cansancio en sus músculos masticadores?
- ¿Le comentan en su casa que hace ruido con los dientes mientras duerme?

Las facetas de desgaste dentarias han sido tradicionalmente asociadas a bruxismo, sin embargo su etiología es multifactorial y no se asocian al diagnóstico ni a la severidad de este. Su hallazgo establece la presencia de bruxismo en un determinado momento de la vida del paciente, pero no determina la presencia de un bruxismo activo.

Para el relevamiento de la población uruguaya (en adultos) se utilizó la presencia de facetas de desgaste en cúspides de caninos superiores como indicador de padecer o haber padecido bruxismo.

La respuesta positiva a la pregunta «¿se despierta con sensación de haber dormido apretando los dientes?» se utilizó como indicador de la presencia de bruxismo activo y del sueño.

Indicadores de TTM

Teniendo en cuenta el DC/TMD, se deben considerar el dolor y la disfunción como indicadores de presencia de TTM (Schiffman *et al.*, 2014).

Los indicadores que se presentan a continuación se registran como variables categóricas dicotómicas.

Dolor

En TTM el síntoma *dolor* puede llevar a confusión entre su origen muscular o articular, especialmente por la cercanía de las estructuras anatómicas. Para el diagnóstico diferencial debe identificarse el polo lateral de las articulaciones temporomandibulares (ATM) del masetero profundo. La identificación se realiza solicitando al paciente que realice movimientos de apertura y protrusión, sintiendo su traslado en el pulpejo del dedo índice. Si el dolor es de origen articular, el paciente relatará el dolor que le es familiar.

El dolor a nivel muscular puede ser *mialgia* o *dolor miofascial*. A nivel articular, *artralgia*. Para identificarlos se tienen en cuenta los siguientes criterios/variables en el diagnóstico:

Dolor muscular:

- Mialgia:
 - Confirmación del dolor en una localización correspondiente a una estructura muscular.
 - Dolor a la palpación en temporal o masetero.
 - Dolor a la máxima apertura no asistida o asistida.
 - Dolor que le es familiar a la apertura o a la palpación.
- Dolor miofascial:

- Dolor sostenido e identificación de patrones referidos.

Dolor articular:

- Artralgia:
 - Dolor a la palpación en el polo lateral (1 kg de presión).
 - Dolor que le es familiar al mover o palpar.

Disfunción

Las *disfunciones* son alteraciones del rango de movimiento mandibular. Para diagnosticarlas se debe identificar uno de los siguientes síntomas:

- Registro de ruido en las articulaciones en el último mes.
- Examen clínico:
 1. Desplazamiento discal con reducción: chasquido al menos en 1 de 3 palpaciones articulares laterales.
 2. Desplazamiento discal sin reducción con apertura limitada: apertura asistida menor a 40 mm.
 3. Desplazamiento discal sin reducción ni limitación de la apertura: apertura asistida mayor o igual a 40 mm. Deberá estar asociada a otras variables de origen muscular y datos de resonancia magnética.
 4. Enfermedad articular degenerativa:
 - Crepitación a la palpación.
 - Crepitación reportada al movimiento.

- Agerholm, D. M., & Ashley, F. P. (1996, February). Clinical assessment of periodontitis in young adults – Evaluation of probing depth and partial recording methods. *Community Dent Oral Epidemiol*, 24(1), 56-61.
- Ainamo, J., & Bay, I. (1975). Problems and proposals for recording gingivitis and plaque. *Int Dent J*, 25(4), 229-235.
- Alves, L., Moura, M., Groisman, S., & Maltz, M. (2016). Epidemiologia da cárie dentária: Arte. En: M. Maltz, L. M. Tenuta, S. Groisman, J. A. Cury, & A. Médicas (Ed.), *Cariologia: conceitos básicos, diagnóstico e tratamento não restaurador*. São Paulo: Série Abeno.
- Antunes, J. L. F. & Peres, M. A. (2013). *Epidemiologia da saúde bucal*. Santos.
- Arana, A., Bernabé, E., & Salazar, F. (2005). Diagnóstico de caries dental. En: G. Henostroza, *Diagnóstico de caries dental* (113-120). Universidad Peruana Cayetano Heredia.
- Argimon Pallas, J., & Jiménez Villa, J. (2013). Anexo 1. Medidas de frecuencia. En: J. M. Argimon Pallas, & J. Jiménez Villa, *Métodos de investigación clínica y epidemiológica* (328-330). Elsevier.
- Armitage, G. C. (1999). Development of a classification system for periodontal diseases and conditions. *Ann Periodontol*, 4 (1), 1-6.
- Baelum, V., Sheiham, A., & Burt, B. A. (2008). Caries control for populations. En: O. Fejerskov & E. S. Kidd, *Dental Caries: The Disease and its Clinical Management* (506-523). Blackwell Munksgaard.
- Barrios, E., Garau, M., & Musetti, C. (2014). *IV Atlas de incidencia del cáncer en el Uruguay 2007-2011*. Comisión Honoraria de Lucha Contra el Cáncer, Montevideo. Recuperado el 27 de junio de 2019, de: http://www.comisioncancer.org.uy/uc_394_1.html.
- Bastos, J. L., Boing, A. F., Perez, K. G., Antunes, J. L., & Peres, M. A. (2011). Periodontal outcomes and social, racial and gender inequalities in Brazil: a systematic review of the literature between 1999 and 2008. *Cad Saude Publica*, 27(Suppl 2), S141-S153.

- Beltran-Aguilar, E. D., Eke, P. I., Thornton-Evans, G., & Peterson, P. E. (2012). Recording and surveillance systems for periodontal diseases. *Periodontol 2000* 2012;60(1):40-53. *Periodontol 2000*, 60(1), 40-53.
- Benigeri, M., Payette, M., & Brodeur, J. M. (1998). Comparison between the DMF indices and two alternative composite indicators of dental health. *Community Dentistry and Oral Epidemiology*, 26(5), 303-309. DOI: [10.1111/J.1600-0528.1998.TB01965.X](https://doi.org/10.1111/J.1600-0528.1998.TB01965.X)
- Borrell, L. N., & Crawford, N. D. (2012). Socioeconomic position indicators and periodontitis: examining the evidence. *Periodontol 2000*, 58(1), 69-83.
- Borrell, L. N., Burt, B. A., & Neighbors, H. W. (2006). The role of individual and neighborhood social factors on periodontitis: the third National Health and Nutrition Examination Survey. *J Periodontol*, 77(3), 444-453.
- Braga, M. M., Mendes, F. M., Martignon, S., Ricketts, D. N. J., & Ekstrand, K. R. (2009). In vitro comparison of Nyvad's system and ICDAS-II with Lesion Activity Assessment for evaluation of severity and activity of occlusal caries lesions in primary teeth. *Caries Research*, 43(5), 405-412. DOI: [10.1159/000239755](https://doi.org/10.1159/000239755)
- Brathall, D. (2000). Introducing the Significant Caries Index together with a proposal for new global oral health goal for 12-year-olds. *Int Dent J*, 50(6), 378-384.
- Breilh, J. (2010). Parte 3: La reproducción social y el proceso epidemiológico (o breve lectura epidemiológica del capital). En: J. Breilh, *Epidemiología: economía política y salud. Bases estructurales de la determinación social de la salud*. Universidad Andina Simón Bolívar – Corporación Editora Nacional.
- Breilh, J. (2013). La determinación social de la salud como herramienta de transformación hacia una nueva salud pública (salud colectiva). *Rev. Fac. Nac. Salud Pública*, 31(supl1), S13-S27.
- Burt, B., Baelum, V., & Fejerskow, O. (2008). The epidemiology of dental caries. En: O. Fejerskow, & E. Kidd, *Dental Caries: The Disease and its Clinical Management* (pp.124-145). Blackwell Munksgaard.

- Chávarry, N.G., Vettore, M.V., Sansone, C., & Sheiham, A. (2009). The relationship between diabetes mellitus and destructive periodontal disease: a meta-analysis. *Oral Health Prev Dent*, 7, 107-127.
- Chaves, M. (1962). Odontología sanitaria. En: M. Chaves, *Problemas. Teoría de la odontología sanitaria*. OPS-OMS.
- Christidis, N., Lindström Ndanshau, E., Sandberg, A., & Tsilingaridis, G. (2019). Prevalence and treatment strategies regarding temporomandibular disorders in children and adolescents. A systematic review. *J Oral Rehabil*, 46(3), 291-301.
- Cortés, F.J. (1999). Medición de la enfermedad en odontología comunitaria. En: E. Cuenca, C. Manau & L. Serra, *Odontología preventiva y comunitaria* (303-325). Masson.
- Dahlgren, G. & Whitehead, M. (2007). A framework for assessing health systems from the public's perspective: the ALPS approach. *Int J Health Serv*, 37(2), 363-378. DOI: [10.2190/U814-6X80-N787-807J](https://doi.org/10.2190/U814-6X80-N787-807J)
- Edge, S., Byrd, D., Compton, C., Fritz, A., Greene, F., & Trotti, A. (2010). *The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual* (7.^{ma} ed.). Springer.
- Eke, P.I., & Genco, R.J. (2007). CDC Periodontal Disease Surveillance Project: background, objectives, and progress report. *J Periodontol*, 78(7 Suppl), 1366-1371.
- Eke, P., Page, R., Wei, L., Thornton-Evans, G., & Genco, R.J. (2012). Update of the case definitions for population-based surveillance of periodontitis. *J Periodontol*, 83(12), 1449-1454.
- Facultad de Odontología. (2014). *Plan de estudios 2011. Carrera de Doctor en Odontología* [plan de estudios].
- Fejerskov, O. (2004, Jun). Changing paradigms in concepts on dental caries: consequences for oral healthcare. *Caries Res*, 38(3), 182-91.
- Fejerskov, O., & Manji, F. (1990). Risk assessment in dental caries. En: J. D. Bader (Ed.), *Risk Assessment in Dentistry*. University of North Carolina Dental Ecology.
- Ferlay, J., Parkin, D., & Steliarova-Foucher, E. (2010). Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer*, 46(4), 765-781.

- Genco, R.J., & Borgnakke, W.S. (2013). Risk factors for periodontal disease. *Periodontol 2000*, 62(1), 59-64.
- Graziani, F., Gennai, S., Solini, A., & Petrini, M. (2018). A systematic review and meta-analysis of epidemiologic observational evidence on the effect of periodontitis on diabetes: An update of the EFP-AAP review. *Journal of Clinical Periodontology*, 45(2), 167-187.
- IARC/OMS. (2016). Prevención y control del cáncer en el contexto de un enfoque integrado (Consejo Ejecutivo EB140/31). OMS. Recuperado el 30 de marzo de 2017, de: https://apps.who.int/gb/ebwha/pdf_files/EB140/B140_31-sp.pdf
- Ismail, A. I., Sohn, W., Tellez, M., Amaya, A., Sen, A., Hasson, H., & Pitts, N. B. (2007). The International Caries Detection and Assessment System (ICDAS): an integrated system for measuring dental caries. *Community Dentistry and Oral Epidemiology*, 35(3), 170-178. DOI: [10.1111/J.1600-0528.2007.00347.X](https://doi.org/10.1111/J.1600-0528.2007.00347.X)
- Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., & Thun, M. (2007). Cancer Statistics. *CA Cancer J Clin*, 57(1), 43-66.
- Jussila P, Kiviahde, H., Näpänkangas, R., Päckilä, J., Pesonen, P., Sipilä, K., Pirttiniemi, P., & Raustia, A. (2017, Spring). Prevalence of Temporomandibular Disorders in the Northern Finland Birth Cohort 1966. *J Oral Facial Pain Headache*, 31(2), 159-164.
- Khader, Y. S., Dauod, A. S., El-Qaderi, S. S., Alkafajei, A., & Batayha, W. Q. (2006). Periodontal status of diabetics compared with nondiabetics: A meta-analysis. *J Diabetes Complications*, 20, 59-68.
- Kidd, E., & Fejerskov, O. (2004). What constitutes dental caries? Histopathology of carious enamel and dentin related to the action of cariogenic biofilms. *J Dent Res*, 83(Spec n.ºC), C35-38.
- Kim, J. S., & Dailey, R. J. (2008a). Measures of Central Tendency, dispersion and skewedness. En: J. S. Kim & R. J. Dailey, *Biostatistic for Oral Healthcare* (pp. 27-53). Wiley-Blackwell.
- Kim, J. S., & Dailey, R. J. (2008b). Probability. En: J. S. Kim & R. J. Dailey, *Biostatistics for oral healthcare* (pp. 55-79). Wiley-Blackwell.

- Kingman, A., & Albandar, J. M. (2002). Methodological aspects of epidemiological studies of periodontal diseases. *Periodontol 2000*, 29(1), 11-30.
- Lencova, E., Broukal, Z., & Duskova, J. (2006). Psychosocial, behavioural and oral health indicators – review of the literature. *Prague Med Rep*, 107(3), 305-316.
- Lobbezoo, F., Ahlberg, J., Glaros, A. G., Kato, T., Lavigne, G., de Leeuw, R., Manfredini, D., Svensson, P., & Winocur, E. (2013). Bruxism defined and graded: an international consensus. *J Oral Rehabil*, 40(1), 2-4.
- Lockhart, P. B., Bolger, A. F., Papapanou, P. N., Osinbowale, O., Trevisan, M., & Levison, M. E. (2012). Periodontal Disease and Atherosclerotic Vascular Disease: Does the Evidence Support an Independent Association?: A Scientific Statement From the American Heart Association. *Circulation*, 125, 2520-2544.
- Luiz de Barreto Aranha, R., Nogueira Guimarães de Abreu, M., Serra-Negra, J., & Martins, R. (2018, Nov). Evidence-Based Support for Sleep Bruxism Treatment Other Than Oral Appliances Remains Insufficient. *J Evid Based Dent Pract*, 18(2), 159-161. DOI: [10.1016/j.jebdp.2018.03.003](https://doi.org/10.1016/j.jebdp.2018.03.003)
- Lung, J., Molly, H., Coming, S., & Ariyawardana, A. (2018). Prevalence of temporomandibular disorders among a cohort of university undergraduates in Australia. *Journal of Investigative and Clinical Dentistry*, 9(3), e12341.
- Maltz, M., Jardim, J., & Alves, L. (2010). Health promotion and caries dental caries. *Braz Oral Res*, 24(Spec. Iss 1), 18-25.
- McNeill, C. (1993). Introduction. En: C. McNeill, *Temporomandibular Disorders: Guidelines for Classification, Assessment, and Management* (pp. 11-13). Quintessence Publishing Co.
- Mehata, A. (2012). Comprehensive review of caries assessment systems developed over the last decade. *RSBO*, 9(3), 316-321.
- Meyle, J., & Chapple, I. (2015). Molecular aspects of the pathogenesis of periodontitis. *Periodontol 2000*, 69(1), 7-17.

Ministério da Saúde (2009). *SB Brasil 2010: Pesquisa Nacional de Saúde Bucal: Manual da equipe de campo*. Recuperado de: https://edisciplinas.usp.br/pluginfile.php/7465690/mod_resource/content/1/REFER%C3%80NCIA%20-%20SBBrazil2010_Manual_Equipe_Campo.pdf

Nyvad, B., Machiulskie, V., & Baelum, V. (1999). Reliability of new caries diagnostic system differentiating active and inactive caries lesions. *Caries Res*, *33*(4), 252-260.

O'Leary, T. (1967). The periodontal screening examination. *J Periodontol*, *38*(6), Suppl: 617-624.

O'Leary, T. (1972). The Plaque Control Record. *J Periodontol*, *43*(1), 38.

Okeson, J. (2013). *Tratamiento de oclusión y afecciones temporomandibulares* (7.^a ed.). Elsevier SL.

Oppermann, R. V., Haas, A. N., Rosing, C. K., & Susin, C. (2015). Epidemiology of periodontal diseases in adults from Latin America. *Periodontal 2000*, *67*(1), 13-33.

Schiffman, E., Ohrbach, R., Truelove, E., Look, J., Anderson, G., Goulet, J.-P., List, T., Svensson, P., Gonzalez, Y., Lobbezoo, F. Michelotti, A., Brooks, S. L., Ceusters, W., Drangsholt, M., Ettlin, D., Gaul, C., Goldberg, L. J., Haythornthwaite, J. A., Hollender, L., Jensen, R., John, M. T., De Laat, A., De Leeuw, R., Maixner, W., Van Der Meulen, M., Murray, G. M., Nixdorf, D. R., Palla, S., Petersson, A., Pionchon, P., Smith, B., Visscher, C. M., Zakrzewska, J., & Dworkin, S. F. (2014). Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: Recommendations of the International RDC/TMD Consortium Network* and Orofacial Pain Special Interest Group†. *Journal of Oral & Facial Pain and Headache*. *28*(1); 6-27.

Page, R. C., Offenbacher, S., Schoeder, H. E., Seymour, G. J., & Kornman, K. S. (1997). Advances in the pathogenesis of periodontitis: summary of developments, clinical implications and future directions. *Periodontol 2000*, *14*, 216-248.

- Page, R., & Eke, P. (2007). Case definitions for use in population-based surveillance of periodontitis. *J Periodontol*, 1387-1399.
- Peres, M. A., Peres, K. G., Cascaes, A. M., Correa, M. B., Demarco, F. F., Hallal, P. C., Horta, B. L., Gigante, D. P., Menezes, A. B (2012). Validity of partial protocols to assess the prevalence of periodontal outcomes and associated sociodemographic and behavior factors in adolescents and young adults. *J Periodontol*, 83(3), 369-378. DOI: [10.1902/jop.2011.110250](https://doi.org/10.1902/jop.2011.110250)
- Petersen, P. E. (2003, Dec). The World Oral Health Report 2003: continuous improvement of oral health in the 21st century—the approach of the WHO Global Oral Health Programme. *Community Dent Oral Epidemiol*, 31(Suppl. 1), 3-23.
- Petersen, P., & Beahni, P. C. (2012). Periodontal health and global public health. *Periodontol 2000*, 60(1), 7-14.
- Petersen, P., Erik, B., Ramon J., & World Health Organization. (2013). Oral health surveys: basic methods, 5th ed. World Health Organization. <https://iris.who.int/handle/10665/97035>
- Pitts, N. (1993). Current methods and criteria for diagnosis in Europe. *J Dent Educ*, 57(6), 409-414.
- Pitts, N. (2004). «ICDAS»—an international system for caries detection and assessment being developed to facilitate caries epidemiology, research and appropriate clinical management. *Community Dental Health*, 21(3), 193-198.
- Pitts, N., & Fyffe, H. (1988). The effect of varying diagnostic thresholds upon clinical caries data for a low prevalence group. *Journal of Dental Research*, 67(3), 592-596. DOI: [10.1177/00220345880670031401](https://doi.org/10.1177/00220345880670031401)
- Research, Science and Therapy Committee. (2003). Position Paper: Diagnosis of Periodontal Diseases. (1237-1247, Ed.) *J Periodontol*, 74(8), 1237-1247.
- Riva, R., Sanguinetti, M., Rodríguez, A., & Guzzetti, L. (2011). Prevalencia de trastornos témporo mandibulares y bruxismo en Uruguay PARTE I. *Odontoestomatología*, 13(17), 54-71.

- Rouquayrol, M. Z., Goldbaum, M., & de Pinho Santana, E. W. (2014). Epidemiología, historia natural e prevención de doenças. En: M. Z. Rouquayrol & M. Gurgel, *Epidemiologia & Saúde* (pp. 11-24). MedBook.
- Savage, A., Eaton, K., & Needleman, I. (2009). A systematic review of definitions of periodontitis and methods that have been used to identify this disease. *J Clin Periodontol*, *36*, 458-467.
- Schuller, A. A., & Holst, D. (2001). Oral status indicators DMFT and FS-T: reflections on index selection. *Eur J Oral Sci*, *109*(3), 155-159. DOI: [10.1034/j.1600-0722.2001.00016.x](https://doi.org/10.1034/j.1600-0722.2001.00016.x)
- Selwitz, R. H., Ismail, A. I., Pitts, N. B. (2007). Dental caries. *Lancet*, *369*(9555), 51-9. DOI: [10.1016/S0140-6736\(07\)60031-2](https://doi.org/10.1016/S0140-6736(07)60031-2)
- Seoane-Romero, J., Vázquez-Mahía, I., Seoane, J., Varela-Centelles, P., Tomás, I., & López-Cedrún, J. (2012). Factors related to late stage diagnosis of oral squamous cell carcinoma. *Med Oral Patol Oral Cir Bucal*, *17*(1), e35-40.
- Sheiham, A., & Nicolau, B. (2005). Evaluation of social and psychological factors in periodontal disease. *Periodontol 2000*, *39*, 118-131.
- Sheiham, A., Maizels, J., & Maizels, A. (1987). New composite indicators of dental health. *Community Dent Health*, *4*(4), 407-414.
- Silness, J., & Løe, H. (1964). Periodontal disease in pregnancy. II. Correlation between oral hygiene and periodontal condition. *Acta Odont Scand*, *22*, 112-135.
- Stevens, S. S. (1946, junio). On the theory of scales of measurement. *Science. New Series*, *103*(2684), 677-680. Recuperado de: <http://www.jstor.org/stable/1671815>
- Susin, C., Kingman, A., & Albandar, J. M. (2005). Effect of partial recording protocols on estimates of prevalence of periodontal disease. *J Periodontol*, *76*(2), 262-267.
- Suvan, J., D'Aiuto, F., Moles, D. R., Petrie, A., & Donos, N. (2011, mayo). Association between overweight/obesity and periodontitis in adults. A systematic review. *Obes Rev Off J Int Assoc Study Obes*, *12*(5), e381-404.
- Thomson, W. N., Sheiham, A., & Spencer, A. J. (2012). Sociobehavioral aspects of periodontal disease. *Periodontol 2000*, *60*(1), 54-63.

- Thornton, E., Eke, P., Wei, L., Palmer, A., Moeti, R., Hutchins, R., Borrell, L. N., & Centers for Control and Prevention (CDC). (2013). Periodontitis among adults aged ≥ 30 years – United States, 2009-2010. *MMWR*, *62*(Suppl 3), 129-135.
- Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J. & Jemal, A. (2015), Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, *65*, 87-108. DOI: [10.3322/caac.21262](https://doi.org/10.3322/caac.21262)
- Verga, F., & Cáceres, R. (2004). Descripción de estadística univariada. En: C. Ketzoian, *Estadística médica* (pp.83-123). Oficina del Libro FEFMUR.
- Vettore, M., Lllamarca, G., Leao, A., Sheiham, A., & Leal, M. (2007). Partial recording protocols for periodontal disease assesment in epidemiological surveys. *Cad Saúde Pública*, *23*(1), 33-42.
- Walsh, J. (1970). International patterns of oral health care—the example of New Zealand. *NZ Dental J.*, *66*(304), 143-152.
- Warnakulasuriya, S. (2009). Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol*, *45*(4-5), 309-316.
- Warnakulasuriya, S. (2010). Living with oral cancer: epidemiology with particular reference to prevalence and life-style changes that influence survival. *Oral Oncol*, *46*(6), 407-410.
- WHO (2008). *World Cancer Report 2008*. Recuperado el 27 de junio de 2019 de: <https://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2008>
- WHO (2013). *Oral health surveys: basic methods* (5th ed.). World Health Organization. Recuperado de: <http://www.who.int/iris/handle/10665/97035>
- WHO. (1979). *A guide to oral health epidemiological investigations*. World Health Organization.
- WHO. (1997). *Oral health surveys: basic methods* (4th ed ed.). World Health Organization. Recuperado de: <http://www.who.int/iris/handle/10665/41905>

Esta publicación cuenta con el apoyo de la Comisión Sectorial de Enseñanza de la Universidad de la República. Forma parte de la serie Manuales Didácticos, que tiene como objetivo mejorar las condiciones de aprendizaje de los estudiantes y, al mismo tiempo, propiciar la autoformación docente mediante la reflexión sobre sus prácticas y sobre el estado del arte en su disciplina. Secundariamente, esta publicación pretende colaborar en la constitución de tradiciones disciplinares y culturas educativas nacionales.

ISBN: 978-9974-0-2059-7

