

Epidemiology basics

Definition, the epidemiological triad, natural history of disease and measures of frequency — with examples from dentistry

Practical Class 8 · Dental Public Health · 2026/2027

Chief Assistant Professor Kostadin Kostadinov, MD, PhD

Department of Social Medicine and Public Health · Medical University of Plovdiv

What you will be able to do after this class

1. Define epidemiology and distinguish **descriptive**, **analytic** and **experimental** epidemiology.
2. Apply the **epidemiological triad** and the **natural history of disease** to caries and periodontitis.
3. Distinguish **proportion, rate and ratio**.
4. Calculate **prevalence**, **cumulative incidence**, **incidence rate** (with person-time), **mortality** and **case-fatality**.
5. Explain why **crude** rates can mislead and what standardisation does.
6. Name the main **data sources** in Bulgaria.

Plan of the class (90 min)

1. What epidemiology is — 15 min
2. Triad and natural history — 12 min
3. Measures of frequency — 33 min
4. Crude and specific rates — 8 min
5. Data sources — 5 min
6. In-class exercises — 15 min

CONTENTS

- 1. What epidemiology is**
2. Epidemiological triad and natural history of disease
3. Measures of disease frequency
4. Crude and specific rates
5. Sources of epidemiological data
6. Exercises

Definition: six key words

*Epidemiology is the study of the **frequency, distribution and determinants** of diseases and other health-related conditions in **human populations**, and the **application** of this study to the promotion of health and to the prevention and control of health problems.*

— Kebede (2004), after Last, A Dictionary of Epidemiology

Key word	Meaning and dental example
population	the unit is the group, not the individual: all 12-year-olds in the region, not the child in the chair
frequency	a quantitative science: how many people, how many new cases, over what time
distribution	person, place, time : e.g. 12-year-olds in villages around Plovdiv have higher DMFT than in the city (Onov & Beltcheva, 2020)
determinants	risk factors (sugars, tobacco) and protective factors (fluoride, sealants)
health-related conditions	not only diseases: injuries, behaviour (toothbrushing), access to dental care
application	the aim is action: prevention programmes, service planning, evaluation of effect

What epidemiology is used for

Community diagnosis

What are the oral health problems of the population and how common are they? (e.g. the share of 12-year-olds with untreated caries)

Finding causes

Which factors raise the risk?
Sugar and caries, tobacco and periodontitis, HPV and oropharyngeal cancer.

Natural history and prognosis

How does disease progress without intervention? How fast does an initial lesion become a cavity?

Evaluating interventions

Does fluoride varnish reduce caries? Is a sealant programme worth it?

Planning services

How many dentists, children's clinics and prosthetic services does an ageing population need?

Surveillance

Ongoing collection, analysis and dissemination of data **for action** — e.g. oral cancer incidence by year.

A short history: from observation to study

Year	Event
5th c. BC	Hippocrates, “Airs, Waters, Places” — environment and way of life influence health
1662	John Graunt analyses the London “Bills of Mortality” — the beginning of health statistics
1747	James Lind compares six treatments for scurvy in sailors — an early controlled trial
1854	John Snow maps cholera around the Broad Street pump — the “father of epidemiology”
1948–1954	Framingham (cohort); Doll and Hill: smoking and lung cancer

The dental story: fluoride

- **1901–1916**: Frederick McKay describes the “Colorado brown stain” — later recognised as **fluorosis**, in which caries is **less common**.
- **1939–1942**: Dean et al. — studies in **21 cities** (8 + 13): a historical finding: as water fluoride rises to about 1 mg/l caries falls, while fluorosis increases with concentration.
- **1945**: Grand Rapids, USA — the first controlled fluoridation of drinking water.

Burt & Eklund, ch. 12, 19, 24, 25 · Dean HT et al. Public Health Rep. 1942;57:1155. doi:10.2307/4584182 · Doll & Hill. BMJ. 1950;2:739. doi:10.1136/bmj.2.4682.739 · Dawber TR et al. AJPH. 1951;41:279. doi:10.2105/AJPH.41.3.279

Descriptive, analytic and experimental epidemiology

Descriptive

Who, where, when?

- person: age, sex, social status
- place: urban/rural, region, country
- time: trends, seasonality

Case report, case series, **cross-sectional survey** (prevalence).
→ **generates** hypotheses

Analytic

Why? Compares groups — exposed and unexposed, or cases and controls:

- ecological
- case-control
- **cohort** (incidence)

→ **tests** hypotheses

Experimental

Does the measure work? The investigator assigns the intervention, if possible at random:

- clinical trial (fluoride varnish)
- community trial (fluoridation)

→ a well-conducted randomised trial is strong evidence of effect

Two axes: **purpose** (describing or testing a hypothesis) and **assignment of exposure** (observational or experimental). In observational studies the investigator does not intervene in the natural course of events.

CONTENTS

1. What epidemiology is
- 2. Epidemiological triad and natural history of disease**
3. Measures of disease frequency
4. Crude and specific rates
5. Sources of epidemiological data
6. Exercises

The epidemiological triad: host, agent, environment

Element	In dental caries
host	tooth (morphology, enamel), saliva (flow, buffering), age, immunity
agent	dysbiotic dental biofilm with acid-producing bacteria (e.g. S. mutans — important, but neither the only nor a sufficient cause)
environment	frequency and amount of sugars, fluoride, hygiene habits, income, access to care
time	duration of exposure — caries is a slow process

Keyes' triad (historical model)

Caries develops where **tooth**, **microflora** and **substrate** (sugars) intersect. Today caries is seen as a dynamic, biofilm-mediated disease strongly driven by diet.

Causal pie model

A sufficient cause is a **complex** of components. Sugar is a **component** cause: not everyone who eats sugar develops caries, but removing it blocks every causal complex that contains it.

Burt & Eklund, 7e (2021), ch. 12: epidemiological triad, Keyes, Rothman

Natural history of disease and levels of prevention

Stage	In dental caries	Level of prevention
1. susceptibility	no disease, but risk factors present: frequent sugars, little fluoride, plaque	primary: health education, fluoride, sealants, less sugar
2. pre-symptomatic (sub-clinical)	early non-cavitated (“white spot”) lesion — painless but clinically detectable; it can arrest or remineralise	secondary: early detection (check-ups, screening), remineralisation
3. clinical	cavitation, pain, pulpitis	secondary/tertiary: restoration, endodontic treatment
4. outcome	without treatment: arrest of the lesion or progression to abscess and tooth loss	tertiary: prosthetics, rehabilitation

Early management of initial caries can avoid more invasive treatment. Absence of pain does **not** mean absence of disease.

The iceberg of disease

Registered cases are the **visible tip**. Below the water remain:

- asymptomatic and early cases;
- people who do not seek care (fear, cost, distance);
- self-treatment;
- inaccurate diagnosis or recording.

In oral diseases the iceberg is large

Periodontitis and early caries do not hurt for a long time; some people see a dentist **only when in pain**. Practice reports show the people **served**, not all the people **affected**; prevalence among attenders may even be higher than in the population (selection).

Consequence

Estimating population prevalence requires **active examinations** of a representative sample — a national oral health survey following the WHO methodology (next class).

CONTENTS

1. What epidemiology is
2. Epidemiological triad and natural history of disease
- 3. Measures of disease frequency**
4. Crude and specific rates
5. Sources of epidemiological data
6. Exercises

Why the number of cases is not enough

Example

In one year town A registers **30** new cases of oral cancer and town B **15**. Where is incidence higher?

Town	New cases	Population	Incidence (annual)
A	30	600,000	$30 / 600,000 \times 100,000 = \mathbf{5 \text{ per } 100,000}$
B	15	150,000	$15 / 150,000 \times 100,000 = \mathbf{10 \text{ per } 100,000}$

The absolute number

tells **how much** resource is needed: examinations, treatments, beds.

The rate

shows the **frequency** of new cases per unit of time — it allows comparison between groups of different size.

The multiplier 10^n (100, 1000, 100,000) is chosen to give a convenient number. Before counting — a **case definition**: diagnostic threshold, dentition, unit (person, tooth, surface).

Proportion, rate, ratio

Measure	Definition	Formula	Dental example
proportion (share)	the numerator is part of the denominator; 0–1 or %	$a / (a + b)$	65 of 100 children have caries experience → 65%
rate	new events per unit of time at risk	events / total time at risk	12 new carious lesions per 100 person-years (one person may contribute several lesions)
ratio	one quantity divided by another; the numerator need not be part of the denominator (a proportion is a special case)	a / b	dentists per 10,000 population; men : women among cases

A ratio from Bulgaria

WHO estimates for 2020: 317 new cases of lip and oral cavity cancer in men and 138 in women → ratio **2.3 : 1**.

Prevalence

Prevalence — the proportion of people in a population who **have** the disease at a given time or during a given period (**new and existing** cases).

$$P = \frac{\text{people meeting the case definition}}{\text{population examined}} \times 10^n$$

- **point** prevalence — at a particular date;
- **period** prevalence — at any time during the interval, each person counted once.

Measured by a **cross-sectional** survey — a “snapshot” of the population.

Prevalence rises with	Prevalence falls with
longer duration of disease	high case-fatality
more effective supportive treatment (fewer deaths)	rapid recovery
higher incidence	lower incidence
better diagnosis	immigration of healthy people
immigration of cases, emigration of healthy people	emigration of cases

Incidence (1): cumulative incidence — risk

Incidence — the frequency of **new** cases during a defined period in the **population at risk**.

$$CI = \frac{\text{new cases during the period}}{\text{people at risk at the start of the period}}$$

- a probability (0–1); always with a **period**: “12% over 3 years”;
- the denominator **excludes** existing cases and those who cannot develop the disease;
- this simple calculation assumes a **closed cohort** and complete ascertainment of outcomes.

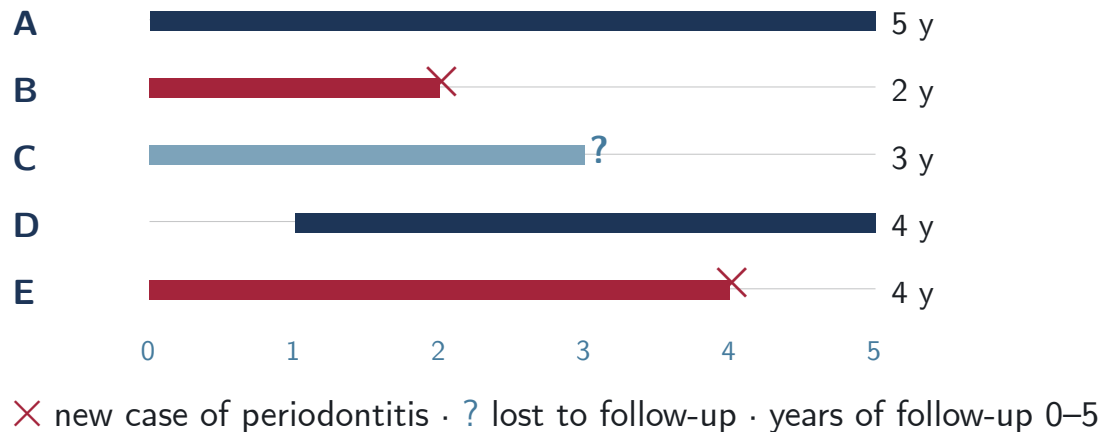
Example

In 2023 a school has 300 six-year-olds; 120 already have caries in the first permanent molars. The remaining **180** children without caries in permanent teeth are followed for 3 years; **45** develop caries.

$$CI = \frac{45}{180} = 0.25 = 25\% \text{ over 3 years}$$

Dividing by 300 would be wrong: 120 children already meet the case definition and cannot become a **new** case (event = first caries in a permanent tooth of the person).

Incidence (2): incidence rate and person-time



$$IR = \frac{\text{new cases}}{\text{total person-time at risk}}$$

Person-years	$5 + 2 + 3 + 4 + 4 = 18$
New cases	2 (B and E)
IR	$2 / 18 = 11.1$ per 100 PY

The rate is **not a probability**: its unit is “per person-year” and it has no upper limit. Person-time handles unequal follow-up, but **informative drop-out** (e.g. sick people leaving more often) can still bias the result.

Prevalence, cumulative incidence, incidence rate

	Prevalence	Cumulative incidence	Incidence rate
numerator	all cases (new + existing)	new cases	new cases
denominator	whole population examined	people at risk at the start	person-time at risk
time	point or period	period must be stated	built into the unit
range	0–1	0–1 (probability)	0 to ∞ , “per PY”
study	cross-sectional	cohort, closed cohort	cohort, dynamic population
answers	how many are ill now?	what is my risk of falling ill?	how fast do new cases appear?

For a rare event and a short period: $CI \approx IR \times t$. E.g. $IR = 2$ per 1000 PY over 5 years $\rightarrow CI \approx 1\%$.

Prevalence and incidence: $P \approx IR \times D$

The bathtub analogy

The tap = incidence (new cases). **The water** = prevalence. **The drain** = recovery or death. The slower it drains (long disease), the more water accumulates.

Conditions: stable population, constant incidence and duration, rare disease; the time units of IR and D must match (in the simple equilibrium model for a more common disease: $\frac{P}{1-P} = IR \times D$).

Example 1 · hypothetical stable population

IR of oral cancer = 3.3 per 100,000 PY; mean duration (from diagnosis to recovery or death) ≈ 4 years
 $\rightarrow P \approx 3.3 \times 4 = 13.2$ per 100,000.

Example 2 · aphthous ulcer

Frequent but short episodes (about 10 days) give a **low** point prevalence.

Hypothetical example; 3.3 is close to the age-standardised incidence for Bulgaria (WHO, 2022), but a standardised rate is not used for this calculation

Mortality and case-fatality

Measure	Formula
crude death rate	deaths in the year / mid-year population $\times$ 1000
cause-specific	deaths from the cause / mid-year population $\times 10^n$
age- (or sex-) specific	deaths in the group / mid-year population of the same group $\times 10^n$
one-year case-fatality proportion	deaths from the disease within 12 months of diagnosis / cases in that diagnosis cohort $\times$ 100
proportional mortality	deaths from the cause / all deaths $\times$ 100

Example (hypothetical town, 250,000 people, 2025)

- 60 new cases of oral cancer;
- each followed **12 months after diagnosis**: 18 died of the cancer;
- in 2025, 24 people in total died of oral cancer (including people diagnosed earlier).

incidence	$60/250,000 = 24$ per 100,000
cause-specific mortality	$24/250,000 = 9.6$ per 100,000
one-year case-fatality	$18/60 = 30\%$

CONTENTS

1. What epidemiology is
2. Epidemiological triad and natural history of disease
3. Measures of disease frequency
- 4. Crude and specific rates**
5. Sources of epidemiological data
6. Exercises

Crude prevalence can mislead: edentulism in two towns

	Town A: population	Town A: edentulous	Town B: population	Town B: edentulous
20–64	8,000	160 (2%)	5,000	100 (2%)
65+	2,000	400 (20%)	5,000	1,000 (20%)
total (crude)	10,000	560 → 5.6%	10,000	1,100 → 11.0%

Age-specific prevalence

Calculated for a subgroup (age, sex). Here it is **identical** in the two towns.

Crude prevalence

Depends on the **structure** of the population. Town B looks twice as “sick” only because it is **older**.

Hypothetical data for adults aged 20+; edentulism = loss of all natural teeth.

Direct age standardisation

Apply the specific rates of **each** town to **the same standard population** (here: 7,000 people aged 20–64 and 3,000 aged 65+).

Age	Expected, town A	Expected, town B
20–64	$7,000 \times 0.02 = 140$	$7,000 \times 0.02 = 140$
65+	$3,000 \times 0.20 = 600$	$3,000 \times 0.20 = 600$
total	$740 / 10,000 = 7.4\%$	$740 / 10,000 = 7.4\%$

With the chosen age standard both towns have the same age-standardised prevalence — **7.4%**.

A standardised measure is a **notional** number — valid for comparison only with the **same** standard, age groups and case definition. Crude numbers describe the real burden and are used for planning (how many dentures are needed).

CONTENTS

1. What epidemiology is
2. Epidemiological triad and natural history of disease
3. Measures of disease frequency
4. Crude and specific rates
- 5. Sources of epidemiological data**
6. Exercises

Data sources

Source	Examples (Bulgaria and international)	Strengths	Limitations
demographic and health statistics	National Statistical Institute (population, deaths by cause); National Centre of Public Health and Analyses — annual health statistics	full coverage, continuity	registered events only; few dental data
registries	National Cancer Registry; NCPHA cancer bulletins; oncology database in the National Health Information System (being built)	detailed follow-up of cases	completeness and timeliness must be assessed
administrative data	National Health Insurance Fund: reported dental activities under the National Framework Contract	large volume, cheap	show services , not disease
surveys	oral health survey following the WHO methodology; European Health Interview Survey (EHIS, NSI)	estimate of population prevalence, social factors	expensive, periodic; non-response; self-report is imprecise
model estimates	GBD (IHME), WHO country profiles	comparable across countries	estimates , not direct measurements

CONTENTS

1. What epidemiology is
2. Epidemiological triad and natural history of disease
3. Measures of disease frequency
4. Crude and specific rates
5. Sources of epidemiological data
- 6. Exercises**

In-class exercises

1 · Prevalence

500 people aged 35–44 are examined; 95 have at least one pocket ≥ 6 mm. What is the prevalence? Is it point or period prevalence?

2 · Cumulative incidence

200 caries-free six-year-olds are followed for 3 years with no losses; 50 develop caries. Calculate the risk.

3 · Person-time

Four people are followed for 3, 5, 2 and 4 years respectively; the second and fourth develop periodontitis at the end of their follow-up. Calculate the IR per 100 PY.

4 · Mortality and case-fatality

Population 400,000; 48 new cases of oral cancer in a year; followed for 12 months after diagnosis, 12 of them died of the cancer; 20 people in total died of this cause during the year. Calculate incidence, cause-specific mortality and case-fatality.

5 · $P \approx IR \times D$

Incidence 0.5 per 1000 PY, mean duration 10 years. What is the expected prevalence?

6 · Crude rate

Aged 20–64: 6,000 people, 1% edentulous; 65+: 4,000 people, 25% edentulous. What is the crude prevalence?

Answers

No.	Solution	Answer
1	95 / 500; on the day of examination	19% point prevalence of the finding “pocket ≥ 6 mm”
2	50 / 200 over 3 years	25% over 3 years
3	$2 / (3 + 5 + 2 + 4) = 2 / 14$	14.3 per 100 PY
4	48 / 400,000; 20 / 400,000; 12 / 48	12 and 5 per 100,000 ; one-year case-fatality 25%
5	$0.0005 \times 10 = 0.005$	5 per 1000
6	$(60 + 1,000) / 10,000$	10.6%

Before you divide, ask: **who is at risk**, **for how long** were they observed, and are we counting **new or all** cases?

Summary

Concepts

1. Epidemiology studies **frequency, distribution and determinants** in populations and applies them to control.
2. Descriptive epidemiology generates hypotheses, analytic tests them, experimental evaluates interventions.
3. The triad and the causal pie explain the multifactorial origin of caries and periodontitis.
4. The natural history defines the levels of prevention.

Measures

1. **Prevalence** — all cases; **incidence** — new cases among people at risk.
2. Cumulative incidence is a **risk**; the rate uses **person-time**.
3. $P \approx IR \times D$.
4. **Mortality** — relative to the population; **case-fatality** — relative to the cases.
5. When populations differ in age structure, compare **specific** or **standardised** measures.

Next class: **Epidemiology of dental diseases** — oral health indices, WHO survey methodology, burden worldwide and in Bulgaria.

Thank you for your attention!