



0 • Exam Map

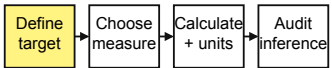
COURSE.JSON · BRIEF

Quiz 20	Oral 20	Assign 20	Exam 40
ASSESSMENT		WEIGHT · TIMING	
Online quiz (60 mins)		20% · Week 6	
Tutorial Oral Assessment and Participation		20% · Weeks 4, 8, 12	
Written Assignment: Written short answer assignment – 1500 words		20% · Week 10	
Written Exam: Written assignment – 2hours		40% · examination period	

“This is an open book examination.” (Written Exam Instructions, Epidemiology 1, 2026)

1 • Answer Spine

COURSE OVERVIEW



Source figure redrawn: population/time/denominator → measure → interpretation → bias/transport.

- Name population, exposure, outcome and time.
- Separate observed inputs from derived quantities; preserve units and signs.
- State the scale: proportion, rate, ratio, difference or weighted mean.
- Interpret for the actual comparison; then test measurement, selection, confounding, causation and transportability.

CRUDE RATE · WHERE IT BITES
 $R = (D/N) \times k$
D events; N stated population; k scale; attach the period.

TRAP A count cannot distinguish burden from detection when case definition, coverage or denominator changes.

REPRESENTATION CHECK BEFORE ARITHMETIC

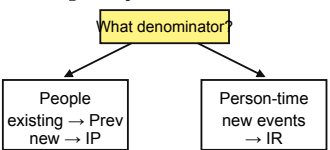
WRITE FIRST	WHAT IT PREVENTS
Observed inputs vs derived quantities	Smuggling an answer into the setup
Symbol, value and unit	Population/person-time or scale mismatch
Expected direction	Accepting a sign or boundary error
Decision consequence	Repeating a result without answering the prompt

After the baseline answer, change the input most tied to the mechanism, hold the rest fixed and recompute only affected steps. Explain whether the movement is plausible. A strong final sentence states **what can be concluded, for whom and over what period**, then attaches the uncertainty or limitation that has power to change it.

FOUR KINDS OF STATEMENT

2 • Frequency

C2



Adapted from Disease Frequency figures: the denominator determines the question answered.

MEASURE	NUMERATOR / DENOMINATOR	USE
Prevalence	Existing cases / population	Point or period burden
Incidence proportion	New cases / initially at-risk people	Fixed cohort + period
Incidence rate	New events / person-time at risk	Unequal follow-up

INCIDENCE RATE
 $IR = D / \Sigma t_i$
Units: events per person-time, never a bare percentage.

People and person-time are not interchangeable. Incidence proportion needs a defensible fixed at-risk cohort and period; incidence rate accommodates varying observed time.

EXAM VERB “Calculate and distinguish” means formula + denominator + time window + one sentence explaining why the alternatives answer different questions.

SET-UP TESTS

PROMPT DETAIL	CONSEQUENCE
Existing cases at a point/period	Prevalence burden; not disease onset
Initially disease-free cohort	Eligible denominator for incidence proportion
Staggered entry or unequal follow-up	Accumulate each person's time at risk
Outcome can recur	Clarify whether numerator is people or events
Reporting/case definition changed	Observed trend may be surveillance artefact

Interpretation stems: “Over [period], [x]% of the initially at-risk population developed [outcome]” for incidence proportion. “There were [x] events per [unit] person-time” for incidence rate. “At [time], [x]% of [population] had [condition]” for point prevalence.

Scaling by 1,000 or another k changes presentation, not the underlying rate. Show the unscaled division before multiplying, and carry the time unit through the final line.

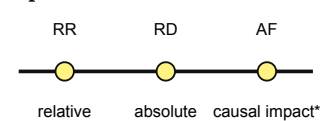
Always expose the denominator: it is the quickest way to distinguish burden, risk and rate.

TRANSMISSION LIMIT

Frequency does not by itself encode contact or susceptibility. Infectious-disease transmission measures add those assumptions.

3 • Association & Impact

C3 · C8 · C9



Adapted from Exposure–Outcome Association: report scale, baseline and causal qualification.

	OUTCOME +	OUTCOME –
Exposed	a	b
Unexposed	c	d

COHORT / TRIAL
 $Risk_1 = a/(a + b)$ · $Risk_0 = c/(c + d)$
 $RR = Risk_1/Risk_0$ · $RD = Risk_1 - Risk_0$
CASE-CONTROL
 $OR = ad/bc$
Exposure odds among cases ÷ exposure odds among controls.

DESIGN	WHAT CAN BE ESTIMATED?	MAIN AUDIT
RCT	Risk, RR, RD by assigned arm	Allocation, adherence, attrition
Cohort	Risk/rate; RR/RD	Follow-up, confounding
Case–control	OR, not sampled risk	Source population/control selection

Interpretation: identify exposed and reference groups, outcome and period. RR is relative; RD is percentage-point excess when risks are percentages. Attributable-fraction language requires **exchangeability and valid measurement**.

TRAP Case/control sampling fractions do not estimate population disease risk.

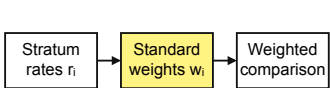
2×2 EXECUTION ORDER

- Label exposure rows and outcome columns before placing any count.
- For cohort/trial data, compute each risk inside its own exposure denominator.
- Choose ratio for relative contrast or difference for absolute contrast.
- For case–control data, use the cross-product OR and state the sampling qualification.
- Only then audit selection, measurement and confounding.

ESTIMATE	NULL	SAFE INTERPRETATION
RR	1	Risk in exposed relative to unexposed
RD	0	Absolute exposed-minus-unexposed risk
OR	1	Exposure odds among cases relative to controls

4 • Standardisation

C4



Adapted from Disease Distribution: common weights remove age-composition mixing.

DIRECT STANDARDISED RATE
 $R_0 = \Sigma_i w_i r_i / \Sigma_i w_i$
If weights sum to 1: $R_0 = \Sigma_i w_i r_i$.

- Describe pattern by **person, place and time**.
- Ask whether age composition is mixed with the comparison.
- Choose and name one common standard population.
- Multiply each population's stratum-specific rates by the same weights; sum.
- Compare standardised values; keep crude burden separate.

QUANTITY	MEANING
Crude rate	Observed burden in that population's actual composition
Age-specific rate	Observed rate within an age stratum
Standardised rate	Comparison construct under chosen common weights

TRAP Changing the standard changes the comparison statistic. Always state the standard; never present R_0 as an observed crude burden.

DIRECTION CHECK

If older strata carry higher rates, a population with more older people can have a higher crude rate even when age-specific rates are comparable. Standardisation targets that composition difference.

PERSON–PLACE–TIME FIRST

DIMENSION	ASK
Person	Which compositional factors differ?
Place	Are populations and ascertainment comparable?
Time	Did disease, population or measurement change?

Before standardising, inspect age-specific rates: a crude difference can reflect composition, genuine within-stratum rate differences, or both. Direct standardisation requires stratum-specific rates and a common set of weights. It improves comparability for the chosen stratifier; it does not eliminate selection, measurement or other confounding.

Report three items together: standard used, weighted rate and scale (for example, per 1,000). If weights do not sum to one, divide by Σw_i . If they do, the result is simply the weighted sum.

5 • Screening 2×2

C5

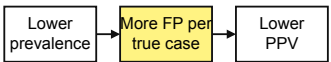
	Disease +	Disease –
Test +	TP	FP
Test –	FN	TN

columns → validity · rows → prediction
Redrawn from Measurement Validity: orient the table before calculating.

VALIDITY
Sensitivity = $TP/(TP+FN)$
Specificity = $TN/(TN+FP)$

PREDICTIVE VALUE
 $PPV = TP/(TP+FP)$
Interpret in the tested population.

QUESTION	READ	MEASURE
Among diseased, positive?	Disease + column	Sensitivity
Among non-diseased, negative?	Disease – column	Specificity
Among positives, diseased?	Test + row	PPV



Same test, different population: predictive value moves with prevalence.

AUDIT Verification procedures can also change predictive values. Do not call sensitivity, specificity and PPV interchangeable “accuracy”.

BUILD FROM TOTALS

- Fix disease status columns and test-result rows.
- Split diseased people into TP and FN; split non-diseased into FP and TN.
- Check all row and column totals before dividing.
- Read the conditioning phrase: “among diseased” selects a column; “among positives” selects a row.
- Interpret the percentage as a conditional probability, not as a causal effect.

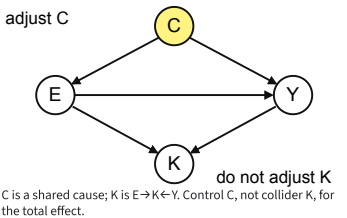
IF PREVALENCE FALLS WHILE SE/SP FIXED	EXPECTED MECHANISM
Fewer true cases	Fewer true positives
Large non-diseased pool remains	False positives can dominate positives
Positive row composition changes	PPV falls

A validity estimate describes performance against the condition classification used as reference. A predictive-value estimate describes what a result means in the tested population. State which probability you calculated and why that denominator answers the question.

6 • Causation & DAGs

C6

CAUSAL TARGET
 $ACE = E[Y(1) - Y(0)]$
Mean potential outcome under exposure minus non-exposure.

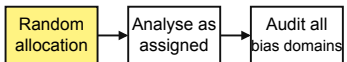


ROLE	GRAPH CUE	ACTION FOR TOTAL EFFECT
Confounder	Common cause of E and Y	Block backdoor path
Mediator	$E \rightarrow M \rightarrow Y$	Do not adjust
Collider	$E \rightarrow K \leftarrow Y$	Do not condition; it can open bias

Exchangeability means compared groups represent relevant counterfactual outcomes. A DAG records assumed causal relations; it does not prove them.

7 • RCT Rule Set

C7



Adapted from RCT figures: randomisation → ITT → risk-of-bias judgement.

TRIAL RISK RATIO
 $RR_{ITT} = (Y_1/N_1)/(Y_0/N_0)$
Outcomes and denominators follow original assignment.

ITT retains participants in assigned groups regardless of adherence. Distinguish assignment, treatment received and adherence.

CHECK	QUESTION
Allocation	Was assignment genuinely random?
Adherence	Were outcomes analysed by assignment?
Outcome	Could measurement differ by arm?
Attrition/reporting	Could missingness or selective reporting shift effect?

Limit Randomisation does not repair attrition, measurement or selective-reporting bias.

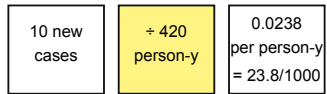
FROM QUESTION TO CAUSAL CLAIM



8 • Drill 1: Frequency

C2 WORKED

Given: 10 cases among 500 people with unequal follow-up totalling **420 person-years**.



Worked route: numerator → person-time denominator → rate with units.

CALCULATION
 $IR = 10/420 = 0.0238$ per person-year = 23.8 per 1,000 person-years.

- Identify **new events**.
- Use total observed person-time because follow-up is unequal.
- Attach person-year units at both scales.
- Do not substitute 10/500 as an incidence proportion unless a defensible fixed cohort and period are specified.

MARKER LINE “The measures are not interchangeable because their denominators encode different time experience.”

FREQUENCY ERROR GRID

ERROR	REPAIR
Uses 500 automatically	Ask whether every person contributes the same complete period
Reports 0.0238%	0.0238 per person-year is a rate; ×1,000 gives 23.8
Drops time from answer	Write “per person-year” beside the denominator and result
Calls rate “risk”	Risk is a proportion over a stated period

Quick plausibility test: 10 events divided by 420 person-years must be a little above 0.02 per person-year. Rescaling to 1,000 multiplies the numerical value by 1,000, not the underlying frequency.

Write the population and time window beside every estimate.

9 • Drill 2: RR + RD

C3 WORKED

Given: risk 12% exposed; 4% unexposed.

$RR = 0.12/0.04 = 3.0$
 $RD = 0.12 - 0.04 = 0.08 = 8 \text{ percentage points}$

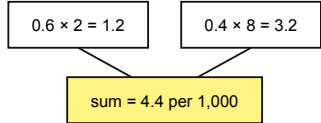
Report both: exposed people have three times the risk; the absolute excess is eight percentage points. Causal impact language still requires exchangeability and valid measurement.

DO	DO NOT
Name both groups and outcome	Say only “three times higher”
Call 0.08 eight percentage points	Call it an 8% relative increase

10 • Drill 3: Standardise

C4 WORKED

Given: standard = 60% younger, 40% older; study rates = 2 and 8 per 1,000.



Apply the stated standard weights to age-specific rates, then sum.

CALCULATION
 $R_s = 0.6 \times 2 + 0.4 \times 8 = 1.2 + 3.2 = 4.4 \text{ per } 1,000$.

- Use standard—not study—weights.
- Preserve the rate unit.
- Name the standard population.
- Do not substitute 10/500 as an incidence proportion unless a defensible fixed cohort and period are specified.

BOUNDARY CHECK A weighted average with positive weights must lie between 2 and 8; 4.4 passes.

STANDARDISATION ERROR GRID

ERROR	WHY IT FAILS
Uses each study’s own age shares	Leaves composition mixed with comparison
Combines crude rates	Skips required stratum-specific information
Omits the standard	Hides a defining part of the estimate
Calls 4.4 observed	R_s is a comparison construct

When the standard changes, the weighted result can change even though age-specific rates do not. That is expected: the target distribution is part of the estimand.

11 • Drill 4: Screening

C5 WORKED

Given: among 100 diseased, 90 test positive; among 900 non-diseased, 90 test positive.

	DISEASE +	DISEASE −	TOTAL
Test +	90 TP	90 FP	180
Test −	10 FN	810 TN	820
Total	100	900	1,000

Sensitivity = $90/100 = 90\%$
PPV = $90/180 = 50\%$

Half of positive results are false positives in this population. Sensitivity and PPV answer different conditional questions.

Checks: FN = 100−90 = 10; TN = 900−90 = 810; positive total = 90+90 = 180. Sensitivity conditions on disease (100), while PPV conditions on positive tests (180).

CHANGED CONDITION	RECOMPUTE
Lower prevalence; same Se/Sp	Rebuild counts; PPV usually falls
Partial verification	Question whether disease status is comparably known

12 • Drill 5: Case–Control

C9 WORKED

Given: cases 80 exposed, 40 unexposed; controls 60 exposed, 120 unexposed.

	Exposed	Unexposed
Cases	80	40
Controls	60	120

OR = diagonal product ratio

Arrange outcome by exposure, then cross-multiply.

$OR = (80 \times 120) / (40 \times 60) = 9,600/2,400 = 4.0$.

Sampled cases have four times the exposure odds of controls under the study’s selection process.

- Name the source population.
- Ask whether controls represent exposure odds in that source population.
- Check exposure measurement and recall.
- Interpret as an **odds ratio**, not risk.

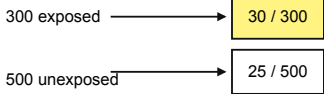
TRAP Do not calculate disease risks from fixed case/control totals.

SELECTION LOGIC

CASES	CONTROLS
Meet the same outcome definition	Could have arisen as cases in the same source population
Exposure measured comparably	Exposure distribution represents that source population
Ascertainment not driven by exposure	Participation not jointly driven by exposure and study entry

13 • Cohort Audit

C8



Forward follow-up establishes temporality; validity still depends on follow-up and confounding.

A cohort follows exposed and unexposed participants to subsequent outcomes, supporting temporal order. Before interpreting its contrast, check whether eligibility came from a common study base, exposure preceded outcome, measurement was comparable, and outcome observation remained complete.

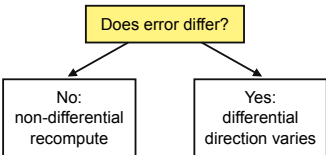
THREAT	TRACE THE PATHWAY
Loss to follow-up	Does remaining observed membership depend on exposure and prognosis?
Confounding	Did a common cause generate exposed/unexposed risk differences?
Measurement	Could ascertainment differ after exposure classification?

Temporality is necessary, not sufficient, for causation.

Selection, confounding and missing outcomes can distort a temporally ordered association.

14 • Misclassification

C10



Adapted from Measurement Error: classify the mechanism before discussing bias direction.

OBSERVED EXPOSURE PROBABILITY
 $P(E^*) = Se \cdot P(E) + (1 - Sp)[1 - P(E)]$
= correctly classified exposed + false positives.

TYPE	QUESTION	SAFE CONCLUSION
Random error	Imprecision without systematic group difference?	Repeat measures may vary
Systematic error	Measurement shifted consistently?	Can create information bias
Non-differential	Same classification process across comparison?	Do not memorise a universal direction
Differential	Error differs by group/outcome?	Rebuild cells; direction may vary

RECLASSIFICATION STEPS

- Find true exposed and unexposed totals in each outcome group.
- Apply sensitivity to true exposed.
- Apply false-positive rate 1−Sp to true unexposed.
- Add to obtain observed exposed; rebuild table and recompute association.

RULE Bias direction is a calculation or causal argument, not a slogan.

MECHANISM-FIRST ANSWER

TERM	COMPACT DEFINITION
Information bias	Systematic error from measurement or classification
Differential	Error differs across comparison groups or outcome status
Non-differential	Error does not vary by the comparison dimension considered

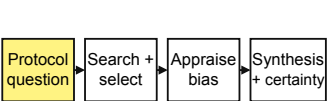
With sensitivity 80% and specificity 90% in both outcome groups, an observed exposed count combines 0.80 of true exposed people with 0.10 of true unexposed people. Apply that operation separately within each outcome group before rebuilding the association.

SURVEILLANCE ECHO

A changed reporting system can raise notifications through detection alone. Check stable definitions, coverage, denominators and corroborating sources before claiming disease increased.

15 • Evidence Synthesis

C11



Adapted from Systematic Reviews: a pooled number is the end, not the review method.

INVERSE-VARIANCE MEAN
 $\hat{\theta} = \sum w_i \hat{\theta}_i / \sum w_i$
Only combine compatible estimates under a stated model.

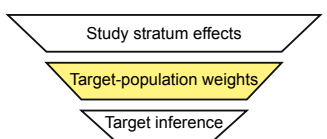
STAGE	EXAM AUDIT
Protocol	Question and methods specified in advance?
Search/select	Could dissemination or eligibility depend on result?
Appraise	What biases sit inside each study?
Synthesise	Are questions, measures and models compatible?
Interpret	Heterogeneity, publication bias, certainty?

Publication bias occurs when dissemination depends on results or direction. Removing unpublished null studies can shift the pooled result.

LIMIT A pooled estimate cannot repair biased studies or incompatible questions.

16 • Transportability

C12



Adapted from External Validity: target weights turn study strata into a target estimate.

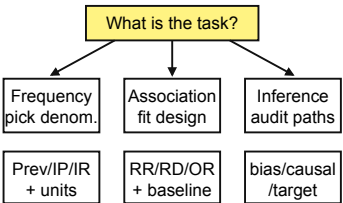
$E_T[Y^*] = \sum_Z E[Y^* \mid Z=z, S=1] P_T(Z=z)$

Name the target first. Ask whether stratum-specific effects transport and whether selected strata capture relevant effect modification. Representativeness alone does not establish causal transport.

TRANSPORT QUESTION	EVIDENCE NEEDED
Who is the target?	Population and setting named explicitly
What differs?	Distribution of modifiers Z in study vs target
Can strata travel?	Effect within Z is applicable from S=1 to target
Are key modifiers captured?	Chosen Z supports, rather than merely assumes, transport

17 • Method Chooser

C1–C12



Fast route: denominator → design-compatible measure → inference audit.

IF PROMPT ASKS...	START WITH...
Current burden	Prevalence; specified population/time
New risk over fixed period	Incidence proportion
Events with unequal follow-up	Incidence rate + person-time
Relative + absolute impact	RR and RD; baseline risk
Sampled cases/controls	OR; source-population selection
Compare age-mixed populations	Direct standardisation
Test performance	Orient diagnostic 2×2
Causal adjustment	DAG roles and exchangeability
Trial effect	Assignment → ITT → bias domains
Pool studies	Compatibility before weighting
Apply elsewhere	Name target; modifiers; triangulate

18 • Final Error Sweep

ALL CHAPTERS

- Denominator:** at-risk people, population or person-time?
- Units:** proportion, per person-time, ratio, percentage points or units?
- Design:** does the sampling permit the measure?
- Measurement:** which cells change under error?
- Selection:** participation/loss related to exposure and outcome?
- Causal role:** confounder, mediator or collider?
- Target:** study or target population weights?
- Claim:** association, causal effect or public-health decision?

FINISH Answer the question in one sentence; name the limitation most able to reverse the conclusion.

THIRTY-SECOND ARITHMETIC SWEEP

- Does a proportion lie between 0 and 1?
- Does a positive weighted mean lie between its component values?
- Does a ratio use two like quantities and have the correct null of 1?
- Does a difference preserve sign and have the null of 0?
- Did rescaling change only the displayed unit?