USMLE Step 1 biostatistics formula sheet

The 2×2 table

Test result	Disease present	Disease absent
Positive	TP (a)	FP (b)
Negative	FN (c)	TN (d)

Test measures

Measure	Formula	Starts from
Sensitivity	TP ÷ (TP + FN)	Everyone with disease
Specificity	TN ÷ (TN + FP)	Everyone without disease
False-negative rate	FN ÷ (TP + FN) = 1 sensitivity	Everyone with disease
False-positive rate	FP ÷ (TN + FP) = 1 specificity	Everyone without disease
PPV	TP ÷ (TP + FP)	Everyone testing positive
NPV	TN ÷ (TN + FN)	Everyone testing negative

- SnNout: a negative result on a highly sensitive test helps rule disease out.
- SpPin: a positive result on a highly specific test helps rule disease in.
- Higher prevalence raises PPV and lowers NPV. Sensitivity and specificity stay the same.

Risk measures

Measure	Formula
Absolute risk	People with outcome ÷ people at risk
Relative risk	Risk exposed ÷ risk unexposed
Relative risk reduction	1 RR, or ARR ÷ control risk
Odds ratio	(a × d) ÷ (b × c)
ARR	Control risk treatment risk
ARI	Exposed risk unexposed risk
NNT	1 ÷ ARR
NNH	1 ÷ ARI
Attributable risk percent	(RR 1) ÷ RR × 100
PAR	Population risk unexposed risk

Measure	Formula
PAR%	PAR ÷ population risk × 100

- Use proportions in reciprocals: 1 ÷ 0.04 = 25.
- Round NNT and NNH up to a whole person.
- Always state the outcome and the time frame.
- Use the odds ratio for case-control studies and relative risk for cohorts and trials.

P-values, confidence intervals and errors

- p-value: the probability of a result at least this extreme if there were no real difference.
- A p-value is not the probability that the null hypothesis is true, and it does not measure effect size.
- 95% CI for a ratio (RR, OR): excludes 1 means $p < 0.05$.
- 95% CI for a difference (ARR, means): excludes 0 means $p < 0.05$.
- Type I error (): finding a difference that is not real.
- Type II error (): missing a difference that is real.
- Power = 1 . It rises with sample size, effect size and .

Study designs and bias

Design	Main measure	Watch for
Cross-sectional	Prevalence	Can't show which came first
Case-control	Odds ratio	Recall and selection bias
Cohort	Relative risk	Loss to follow-up
Randomised trial	RR, ARR, NNT	Poor blinding, crossover
Screening data	Survival time	Lead-time, length-time bias

From the Medly Team. Practise more at medlyai.com.