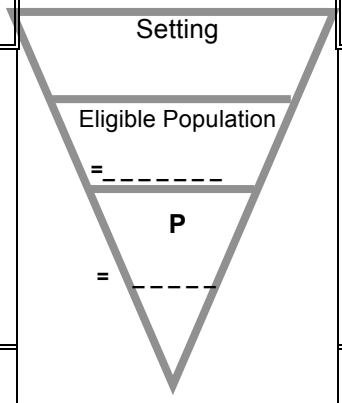
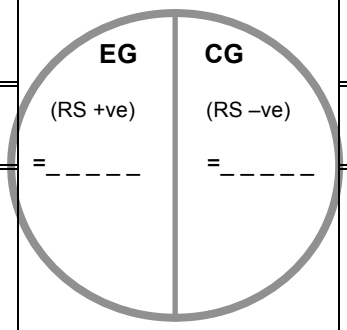
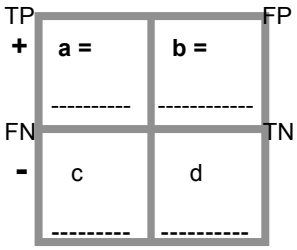
GATE-lite for Diagnostic Test Accuracy Studies 2013

Study details:

STUDY QUESTION & DESIGN - describe with **PECOT**	**STUDY NUMBERS -** hang on **GATE frame**	**STUDY ERROR -** assess using **RAMBOMAN**
P = Participants:		**Recruitment** appropriate to study goals/ able to define who findings applicable to?
Briefly describe – Setting: Eligibility criteria: Recruitment process: % of invited eligibles who participated:		Setting/eligible population appropriate to goals & well described? Participants likely to be similar to Eligibles? Participants typical of spectrum usually tested?
EG = **Exposed** Group [Reference / Gold Standard positive: RS +ve]	**Allocated by measurement**	**Allocation** to EG & CG done accurately?
Describe RS+ve & how / by whom / when assessed:	**REFERENCE STANDARD RESULTS**	Was the Reference Standard a valid standard & assessed objectively & blind to Test result?
CG = **Comparison** Group [Reference / Gold Standard negative: RS -ve]		**Maintenance** in allocated groups during study sufficient?
Describe RS-ve & how / by whom / when assessed:	**TEST RESULTS**	Proportion of P who had both RS & Test done Time period (& any treatment) between RS & Test
O = Outcome (Test result) **T = Time** when test done cf Ref Standard →		
Describe Test & T - how / by whom / when done:		**Blind** and **Objective Measurements** of Test? Was Test measured accurately?

STUDY ANALYSES	**Test**	**+EGO** = a/EG (i.e. sensitivity) **+CGO** = b/CG	**+ LR** = +EGO/+CGO ± 95% CI	**-EGO** = c/EG **-CGO** = d/CG (i.e. specificity)	**- LR** = -EGO/-CGO ± 95% CI

ANalyses: Did cross-tabulation of RS x Test include indeterminate/missing results?____ Adjusted if EG & CG different? (seldom done) _____**95% CIs** or p-values given?_____

Summary:

1. Non-random error sufficiently low? (AMBOM: amount & direction of bias)

2. Analytical error sufficiently low? (AN: analyses of missing results/adjusted analyses)

3. Random error sufficiently low? (95% CIs: and if no statistically significant effects demonstrated was study power/sample size sufficiently high)

4. Size of effects sufficient to be meaningful? (sensitivity & specificity and LRs)?

5. If 1-4 ok, are findings applicable in practice? (R)

GLOSSARY

Use this form for questions about Diagnostic Test Accuracy Studies

Hang the study on the GATE Frame

STUDY QUESTIONS/DESIGN: use PECOT to define study question & describe study design

Setting of study: Timing & locations in which Eligibles identified (e.g. country/urban/hospital).

Eligible population: those from study Setting who meet eligibility (i.e. inclusion / exclusion) criteria.

How were Eligibles identified from the study setting? e.g. based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?

P: Participants: recruited from Eligibles & allocated to EG & CG. How recruited from Eligibles (e.g. consecutive patients)?

EG: Exposure Group: participants allocated to the reference (or gold) standard positive group (i.e. those considered to have the disease/condition being tested for).

CG: Comparison Group: participants allocated to alternative reference standard negative group

Outcome: Test being investigated. If several cut-offs for same Test, use additional GATE-lites.

Time: when were the test measurements done in relation to Reference standard assessment.

STUDY VALIDITY (non random error or bias): use RAMBOM to identify possible non random errors

Recruitment (mainly about external validity): were setting & eligibles appropriate given the study aims &/or the reviewer's interests? Were participants similar to all Eligibles? Are the results applicable to relevant populations Was there a wide enough range (spectrum) of participants with and without the condition/disease & similar to those usually tested in practice?

Allocation: how accurately were participants allocated to Reference Standard (RS) positive and negative groups? Was the RS a valid measure of condition? Were measurements of RS: done **blind** to knowledge of the Test result, & done **objectively** (e.g. automated lab tests, radiography)? Were measurement methods well described? Was RS measurement replicable based on data provided or referenced?

Maintenance: Completeness of follow-up: what proportion of eligible Participants had both Test & RS? How much time was there between the RS and Test measurements? Co-intervention: did participants remain untreated in their initially allocated groups (RS positive or negative OR Test positive or negative)?

Blind Measurement of Tests**:** was it done blind to the participants' RS status? **and/or** were Tests **Objectively Measured**? e.g. biopsies; x-rays, validated questionnaires.

STUDY ANALYSES (estimates of sensitivity & specificity [EGO & CGO], effect sizes [LRs] & random error [95%CI])

Adjusted analyses (for confounders): Were factors that could effect the Test measurements (e.g. age) distributed similarly in the Reference Standard positive and negative groups? If not, were analyses stratified or adjusted.

+EGO: The positive Exposure Group Occurrence is the likelihood of a positive test (a) in those who are Ref Standard positive (EG) = **sensitivity** of test. +EGO = a/EG or a/a+c. ('a' are the True Positives [TP])

+CGO: The positive Comparison Group Occurrence is the likelihood of a positive test (b) in those who are Ref Standard negative.. +CGO = b/CG or b/b+d. ('b' are the False Positives [FP])

-EGO: The negative Exposure Group Occurrence is the likelihood of a negative test (c) in those who are Ref Standard positive (EG). - EGO = c/EG or c/a+c. ('c' are the False Negatives [FN])

-CGO: The negative Comparison Group Occurrence is the likelihood of a negative test (d) in those who are Ref Stan negative = **specificity** of test). -CGO = d/CG or d/b+d. ('d' are the True Negatives [TN])

Effect Estimates (measures comparing EGO & CGO): Likelihood Ratio (LR) = EGO/CGO; the Likelihood Ratio (LR) in a diagnostic test accuracy study is the equivalent of the Risk Ratio in a RCT or cohort study. However there are two LRs: a positive LR = + EGO / + CGO; and a negative LR = - EGO / - CGO

PPV*: the Positive Predictive Value is the probability of being Ref Standard positive (i.e. having the condition or disease) if the Test result is positive. PPV = a/a+b.

NPV*: the Negative Predictive Value is the probability of being Ref Standard negative (i.e. no disease) if the Test result is negative. NPV = d/c+d.

***** the PPV and NPV calculated from a study are only meaningful if the patients these values are applied to have a similar prevalence/severity of the disease/condition as the Participants (P) in the study. In contrast, +ve & -ve LRs (& sensitivity & specificity) are more generalisable from studies to a range of patient populations.

Random error in estimates of EGO, CGO, LR, etc is assessed by width of confidence interval (CI). A wide CI (i.e. big gap between upper & lower confidence limits (CL) = more random error = less precision.

STUDY SUMMARY

Non-random error (bias): what was the likely amount & direction of bias: is bias likely to substantially increase or decrease the observed difference between EGO & CGO (and therefore the effect sizes)?

Analytical error: were analyses done appropriately? Were missing data accounted for? Were adjusted analyses done if important differences between EG & CG?

Random error: would you make a different decision if the real effect was close to upper CL rather than the lower CL? Power: if the effect sizes were not statistically significant, was study just too small to show meaningful effects?

Effect sizes: was the magnitude of LRs, sensitivity & specificity sufficient for test to be meaningful/useful in practice?

Applicability: if effect sizes meaningful & errors small, are the findings likely to be applicable in practice?

REFERENCE:

Jackson et al. The GATE frame: critical appraisal with pictures. In: Evidence-Based Medicine. 2006;11;35-38. Also in: Evidence-Based Nursing 2006; 9: 68-71, and in ACP Journal Club 2006; 144: A8-A11.

LIKELIHOOD RATIO NOMOGRAM:

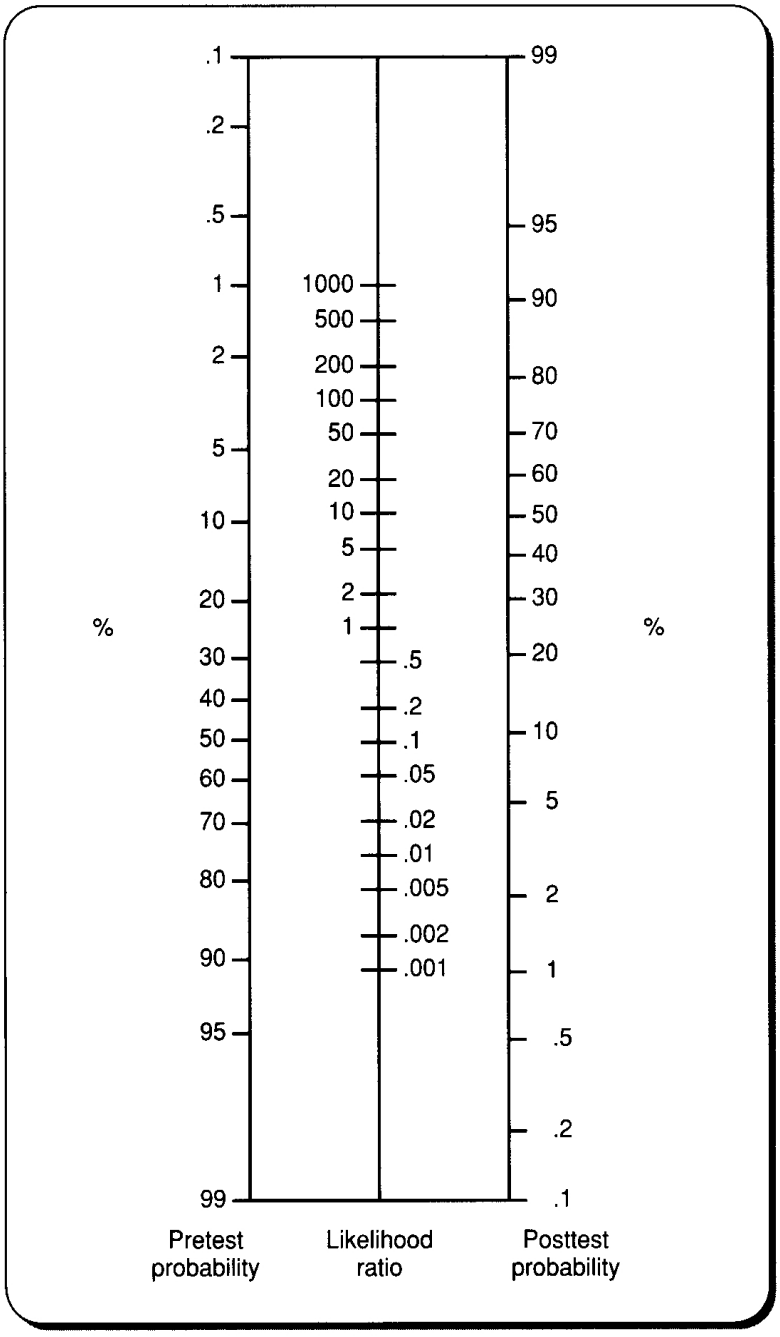
Draw straight lines from the pre-test probability (left hand scale) through the LRs (middlle scale – there is a different LR for positive and negative results) and continue the line to the right hand scale which is the estimated post-test probability of the condition being tested for.

Pre-test probability =- …… +LR = …….. Post-test probability =…….....
–LR = ……Post-test probability - ……

Pre-test Probability (Pre-TP) = proportion of people similar to the person about to be tested who have the condition

= prevalence

= EG/(EG & CG)



Post-test Probability (Post-TP) = proportion of people with +ve or –ve test result who have condtion

Note 1: If the test is positive, use the +ve LR to estimate the Post-TP = Positive Predictive Value (PPV). When the test is negative, use the -ve LR to estimate the Post-TP = 1 – Negative Predictive Value (1-NPV).

Note 2: the probability that a person has the target condition after a test has been done depends on two factors: 1. the probability of the condition in the population being tested and 2. the accuracy of the test. As the probability of the target condition will vary between different populations (e.g. young versus old or primary care patients versus hospitalised patients), the same test result will predict different post-test probability in different populations.