

GATE-LITE for Diagnostic Test Accuracy Studies (23/11/10)

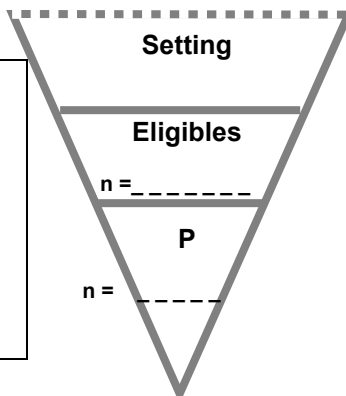
STUDY DESIGN: describe with **PECOT**

STUDY NUMBERS: hang on **GATE** frame

STUDY BIAS: assess using **RAMBO**

P = Participants:

Describe:
 - Setting.
 - Eligibility criteria.
 - How Eligibles identified from Setting.
 - How Participants recruited from Eligibles.



Recruitment appropriate to study goals?

Eligible population appropriate given study goals?
 Participants representative of Eligibles?
 Participants = usual patient spectrum to be tested?

EG = Exposure Group [Reference / Gold Standard positive: RS +ve]

Describe RS+ve & how / by whom / when assessed

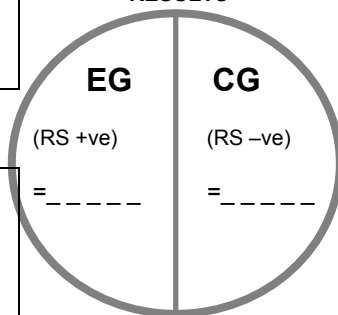
Allocated
REFERENCE STANDARD RESULTS

Allocation to EG & CG done accurately?

Was the Reference Standard a valid standard & assessed objectively & blind to test result?

CG = Comparison Group [Reference / Gold Standard negative: RS -ve]

Describe RS-ve & how / by whom / when assessed



Maintenance of EG & CG as allocated?

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 of Ref Standard

Describe Test & how / by whom / when done:

TEST RESULTS

TP	+	a	b	FP
	=	-----	-----	
FP	-	c	d	TN
	=	-----	-----	

Blind or Objective assessment of Test?

Was Test measured accurately?

STUDY ANALYSES	Test	+EGO = a/EG = sensitivity (TTP)	+CGO = b/CG (FPP)	+ LR = +EGO/+CGO	-EGO = c/EG (FNP)	-CGO = d/CG= specificity (TNP)	- LR = -EGO/-CGO
	95% Conf interval						
	95% Conf interval						

Analyses: Did cross-tabulation of RS x Test include indeterminate/missing results? ____ Were there measures of Test accuracy/reproducibility? ____ 95% CI or p-values given? ____

Summary: Magnitude of sensitivity & specificity and LRs?
 Amount & direction of bias (RAMBO)?
 Amount of random error (width of CI)?
 Power/sample size sufficient (if 95% CIs for LRs include 1.0)?
 Any adverse effects?
 Applicability?

GLOSSARY

*Hang the study on the GATE Frame***STUDY QUESTIONS/DESIGN: use PECOT to define study question & describe study design**

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

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

How was the EP 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 EP & allocated to EG & CG. How recruited from EP (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 multiple Tests or different cut-offs for same Test, use another GATE-lite.

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

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

Recruitment: was the setting / EP appropriate given the study aims &/or the reviewer's interests? Were participants representative of the EP? Could the results be generalised to relevant populations Was there a wide enough range (spectrum) of participants with and without the condition/disease?

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

Maintenance: 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)? Completeness of follow-up: what proportion of eligible Participants had both Test & RS?

Blind assessment of Tests: was it done blind to the participants' RS status? **or** were Tests measured

Objectively? 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 distributed similarly in the Reference Standard positive and negative groups? If not, were analyses stratified.

+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 or true positive proportion (TPP). $+EGO = a/EG$ or $a/a+c$.

+CGO: The positive Comparison Group Occurrence is the likelihood of a positive test (b) in those who are Ref Standard negative. This is the false positive proportion (FPP). $+CGO = b/CG$ or $b/b+d$.

-EGO: The negative Exposure Group Occurrence is the likelihood of a negative test (c) in those who are Ref Standard positive (EG). This is the false negative Proportion (FNP) - $EGO = c/EG$ or $c/a+c$.

-CGO: The negative Comparison Group Occurrence is the likelihood of a negative test (d) in those who are Ref Stan negative = **specificity** of test or true negative proportion (TNP) $+CGO = d/CG$ or $d/b+d$.

Effect Estimates (measures comparing EGO & CGO): Likelihood Ratio (LR) = EGO/CGO; in diagnostic test accuracy studies the Likelihood Ratio (LR) is the equivalent of the Risk Ratio in a RCT or cohort study. There are two LR's – 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 patient group these values are applied to have the same prevalence/severity of the disease/condition as the Participants (P) in the study. In contrast, +ve & -ve LR's (& sensitivity & specificity) are more generalisable from study to study.

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

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

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)?

Random error: would you make a different decision if the real effect was closer to upper CL than lower CL?

Power: if the effect sizes were not statistically significant, was study just too small to show a real difference?

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

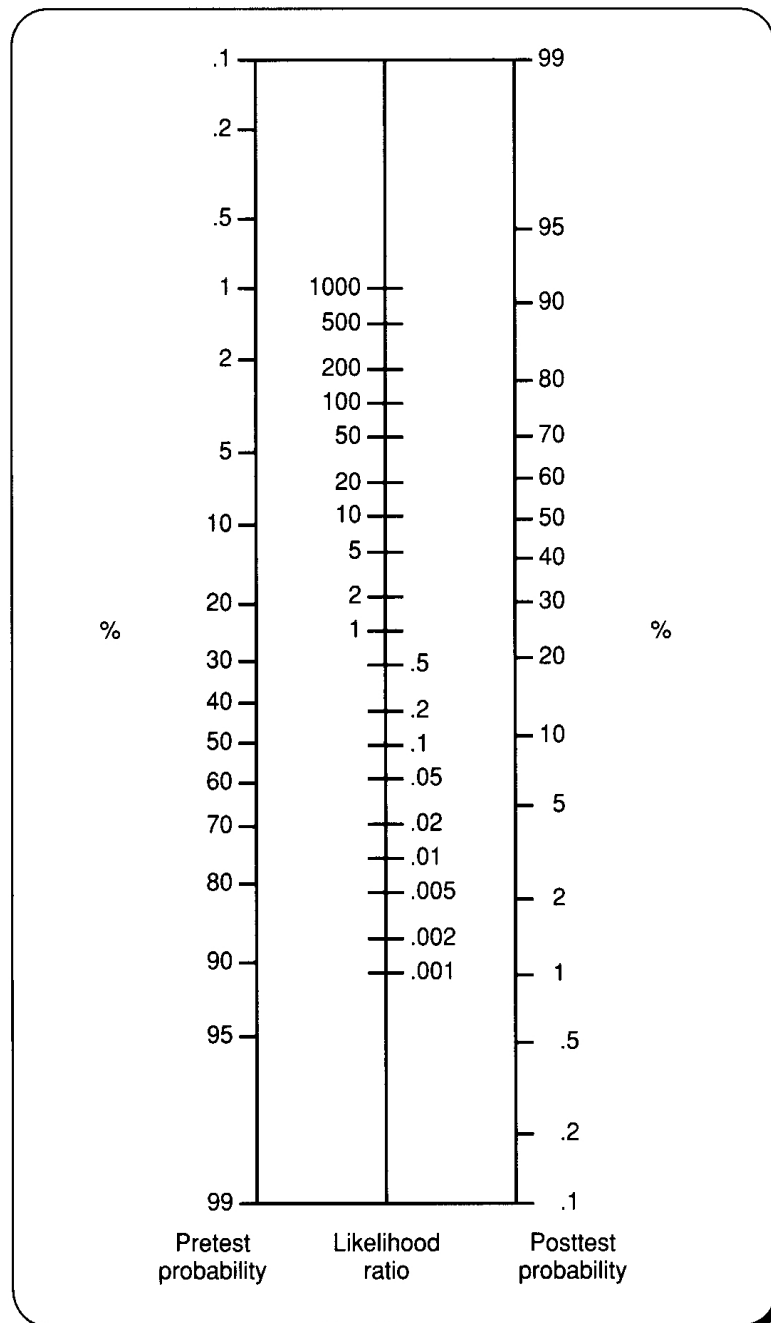
REFERENCE:

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 a line through pre-test probability and LRs to estimate post-test probability of the target condition (e.g. disease)

Pre-test Probability (Pre-TP):
 proportion of population who have the condition
 = prevalence
 = $EG / (EG + CG)$
 (Note: the validity of the Post-TP depends on the Pre-TP (prevalence) of the condition of interest, in the population you are applying the test. This may differ from the prevalence in the study producing the LRs)



Post-test Probability (Post-TP):
 proportion of people with +ve or -ve test who have condition
 When the test is positive, use +ve LR to estimate the Post-TP = Positive Predictive Value (PPV)
 When the test is negative, use -ve LR to estimate the Post-TP = 1 - Negative Predictive Value (1-NPV)

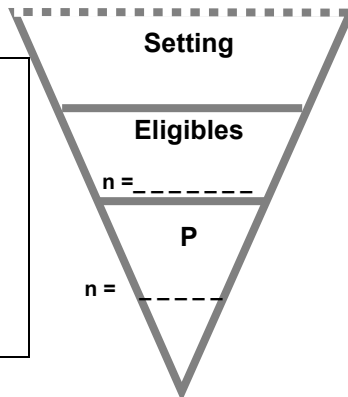
GATE-very LITE for Diagnostic Test Accuracy Studies (23/11/10)

STUDY DESIGN: describe with **PECOT**

STUDY NUMBERS: hang on **GATE** frame

STUDY BIAS: assess using **RAMBO**

P = Participants:



Recruitment appropriate to study goals?

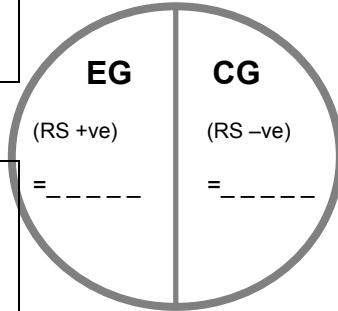
EG = Exposure Group [Reference / Gold Standard positive: RS +ve]

Allocated

Allocation to EG & CG done accurately?

CG = Comparison Group [Reference / Gold Standard negative: RS -ve]

REFERENCE STANDARD RESULTS

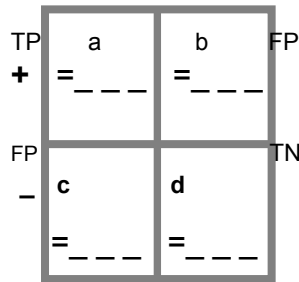


Maintenance of EG & CG as allocated?

TEST RESULTS

O = Outcome (Test result)

T = Time when test done of Ref Standard →



Blind or **Objective** assessment of Test?

STUDY ANALYSES	Test	+EGO = a/EG = sensitivity (TTP)	+CGO = b/CG (FPP)	+ LR = +EGO/+CGO	-EGO = c/EG (FNP)	-CGO = d/CG = specificity (TNP)	- LR = -EGO/-CGO
	95% Conf interval						
	95% Conf interval						

Analyses: Did cross-tabulation of RS x Test include indeterminate/missing results? ____ Were there measures of Test accuracy/reproducibility? ____ 95% CI or p-values given? ____

Summary: Magnitude of sensitivity & specificity and LRs?
 Amount & direction of bias (RAMBO)?
 Amount of random error (width of CI)?
 Power/sample size sufficient (if 95% CIs for LRs include 1.0)?
 Any adverse effects?
 Applicability?