## STUDY QUESTION & DESIGN:
**describe with PECOT**

**P = Participants:**
- Setting
- Eligibility criteria
- Recruitment process
- % of eligibles who participated

**EG = Exposed Group [Reference / Gold Standard positive: RS +ve]**
- Describe RS+ve & how / by whom / when assessed

**CG = Comparison Group [Reference / Gold Standard negative: RS -ve]**
- Describe RS-ve & how / by whom / when assessed

**O = Outcome (Test result)**
- Describe Test & T: how / by whom / when done

---

## STUDY NUMBERS:
**hang on GATE frame**

**Setting**
- Eligibles
  - n = __________
- P
  - n = __________

**Allocated: by measurement**

**REFERENCE STANDARD RESULTS**
- EG
  - (RS +ve)
  - = __________
- CG
  - (RS –ve)
  - = __________

**TEST RESULTS**
- TP
  - a = __________
- FP
  - b = __________
- FN
  - c = __________
- TN
  - d = __________

---

## STUDY ERROR: assess using RAMBOMAN

**Recruitment**
- Setting appropriate to study goals?

**Setting/eligible population appropriate, given study goals?**

**Participants representative of Eligibles?**

**Participants typical of spectrum usually tested?**

**Allocation (± adjustment) to EG & CG done accurately?**
- Was the Reference Standard a valid standard & assessed objectively & blind to Test result?

**Maintenance of EG & CG as allocated sufficient?**
- Proportion of P who had both RS & Test done
- Time period (& any treatment) between RS & Test

**Blind and Objective Measurements of Test?**
- Was Test measured accurately?

---

### STUDY ANALYSES

<table>
<thead>
<tr>
<th>Test</th>
<th>+EGO = a/EG (i.e. sensitivity)</th>
<th>+CGO = b/CG</th>
<th>+ LR = +EGO/CGO ± 95% CI</th>
<th>-EGO = c/EG (i.e. specificity)</th>
<th>-LR = -EGO/CGO ± 95% CI</th>
</tr>
</thead>
</table>

**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:**
- **Non-random error:** amount & direction of bias (RAMBOM)?
- **ANalyses done well?**
- **Random error:** sufficiently low
- **Power/sample size sufficiently high (if no statistically significant effects demonstrated)?**
- **Applicability of findings?** Any important adverse effects?
- **Size of effects sufficient to be meaningful (sensitivity & specificity and LRs)?**
- **Can findings be applied in practice?**
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).
Eligibles: 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 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 & 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: 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 Measurement of Tests: was it done blind to the participants’ RS status? and 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 Reference 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 Reference 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 Reference 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 Reference Standard negative (CG) = 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 Reference 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 Reference 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)?
Random error: there is too much random error if you would 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 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: