GATE: a **G**raphic **A**pproach **T**o **E**vidence based practice

GATE CAT – Diagnostic Test Accuracy Studies

THE UNIVERSITY OF AUCKLAND
FACULTY OF MEDICAL AND HEALTH SCIENCES

updates from previous version in red

Critically Appraised Topic (CAT): Applying the 5 steps of Evidence Based Practice
Using evidence from Diagnostic test accuracy studies

Assessed by:	**Date:**

Problem

Describe the problem that led you to seek an answer from the literature about diagnostic accuracy.

Step 1: Ask a focused 5-part question using PECOT framework (EITHER 'your question' OR 'the study's question')

Population / patient / client	Describe relevant patient/client/population group (be specific about: symptoms, signs, medical condition, age group, sex, etc.) that you are considering testing
Exposure (Target disorder)	Describe the Target disorder (disease/condition) to be diagnosed. Is it relevant to consider levels/categories of severity/stage?
Comparison (no Target disorder)	Describe the typical health status of those without the target disorder who would also receive the test. Are they likely to be disease free or have other co-morbidities?
Outcome (Test)	Describe the test, including levels/categories if relevant, that you are considering doing (note the 'outcome' in a diagnostic test accuracy study is the test result.
Time	Time is not usually considered explicitly in a diagnostic test accuracy question

Step 2: Access (Search) for the best evidence using the PECOT framework

PECOT item	Primary Search Term		Synonym 1		Synonym 2	
Population / Participants / patients / clients	Enter your key search terms for P, E & O. C & T seldom useful for searching. Add mesh terms (e.g. sensitivity & specificity) +/or diagnostic filter to refine. Use MESH terms (from PubMed) if available, then text words.	OR	Include relevant synonym	OR	Include relevant synonym	AND
Exposure (Target disorder)	As above	OR	As above	OR	As above	AND
Comparison (no Target disorder)	As above	OR	As above	OR	As above	AND
Outcomes (Test)	As above	OR	As above	OR	As above	AND
(Time)	As above	AND	As above	AND	As above	
Limits & Filters	PubMed has **Limits** (eg age, English language, years) & PubMed Clinical Queries has **Filters** (e.g. study type) to help focus your search. List those used.					

Databases searched:

Database	Cochrane SRs	Other Secondary Sources	PubMed / Ovid Medline	Other
Number of publications (Hits)	Enter number of hits from Cochrane database search for Systematic Reviews (SR).	Enter number of hits from other secondary sources (specify source)	Enter number of hits from PubMed /Ovid/etc (specify database)	Enter number of hits from other sources (e.g. Google scholar, Google)

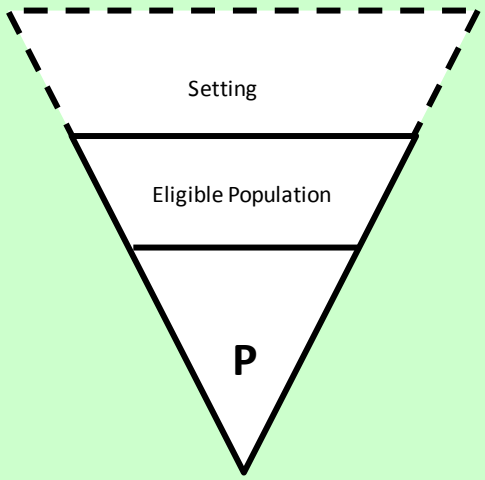
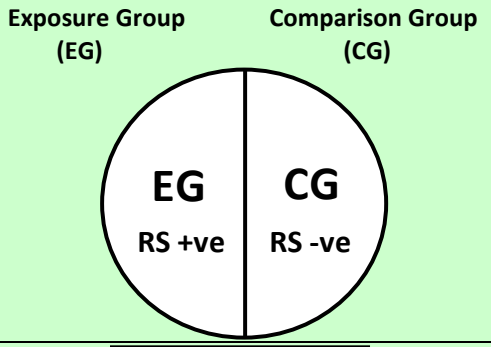
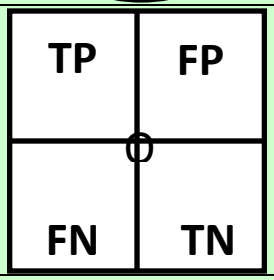
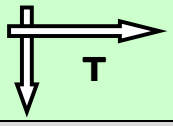
Evidence Selected

Enter the full citation of the publication you have selected to evaluate.

Justification for selection

State the main objectives of the study.
Explain why you chose this publication for evaluation.

Diagnostic test accuracy studies
Step 3: Appraise Study
3a. Describe study by hanging it on the GATE frame (also enter study numbers into the separate excel GATE calculator)

Population		Study Setting	Describe when & from where participants recruited (e.g. what year(s), which country, urban / rural / hospital / community)
		Eligible population Recruitment process	Define eligible population / main eligibility (inclusion and exclusion) criteria (e.g. was eligibility based on presenting symptoms / signs, results of previous tests, or participants who had received the test or reference standard? Describe recruitment process (e.g. were eligibles recruited from hospital admissions / electoral / birth register, etc). How they were recruited (e.g. consecutive eligibles)?
		Participants	What percentage of the invited eligibles participated? What reasons were given for non-participation among those otherwise eligible?
Exposure & Comparison		Allocation method	Allocated by measurement of Target disorder into those with disorder (Ref standard +ve) & those without disorder (Ref standard -ve)
		Exposure (Ref Std. +ve)	Describe reference standard positive disorder: what, how defined, how measured, when, by whom (level of expertise?). Include description of categories if more than yes/no
		Comparison (Ref Std. –ve)	Describe reference standard negative disorder (as above)
Outcomes		Outcome(s) (Test)	Describe the diagnostic test: what, how defined, how measured, when, by whom (level of expertise?). Include description of categories if more than yes/no
Time		Time	State when test was done in relation to when the reference standard was done.

Reported Results	**Enter the main reported results ➔**	Outcome	Risk estimate	Confidence Interval
		Sensitivity		
		Specificity		
		+ve LR		
		-ve LR		
		PPV		
		NPV		

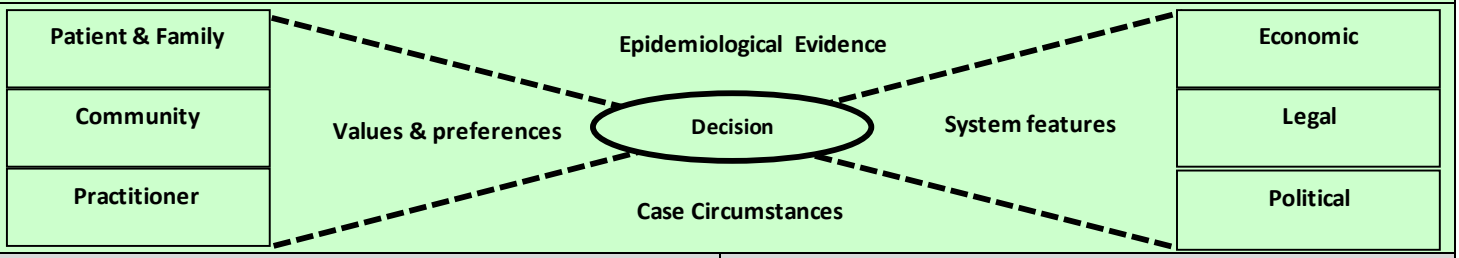
Complete the Numbers on the separate GATE Calculator for Diagnostic Studies

Diagnostic test accuracy studies Step 3: Appraise Study 3b. Assess risk of errors using RAMboMAN			
Appraisal questions (RAMboMAN)		**Risk of errors +, x, ?, na**	Notes
Recruitment/Applicability '**errors'**: questions on risks to application of results in practice are in blue boxes			
Internal study design **errors**: questions on risk of errors within study (design & conduct) are in pink boxes			
Analyses **errors**: questions on errors in analyses are in orange boxes			
Random **error**: questions on risk of errors due to chance are in the green box			
Key for scoring risk of errors: + = low; x = of concern; ? = unclear; na = not applicable			
Participant Population	**Recruitment** - are the findings based on these recruited participants applicable in practice?		
	Study Setting relevant to practice?	Score risk of error as: +, x, ? or na (see key above)	Is the study setting (e.g. what year(s), which country, urban / rural, hospital / community) likely to influence the applicability of the study results?
	Study planned before reference standard and tests done?		Was the study done prospectively or was it a retrospective use of available data? If retrospective was the participant population chosen primarily because of available test data or target disorder data?
	Eligible population relevant to practice?		Was the eligible population from which participants were identified relevant to the study objective and to practice? Were inclusion & exclusion criteria well defined & applied similarly to all potential eligibles?
	Participants similar to all eligibles?		Did the recruitment process identify participants likely to be similar to all eligibles? Was sufficient information given about eligibles who did not participate?
	Key personal (risk/prognostic) characteristics of participants reported? Appropriate spectrum of participants?		Was there sufficient information about baseline characteristics of participants to determine the applicability of the study results? Was any important information missing? Was there an appropriate spectrum of people similar to those in whom the test would be used in practice?
Exposures & Comparisons	**Allocation** to EG & CG done well?		
	Reference standard sufficiently well defined and well measured so participants allocated to correct Target disorder groups?		Were reference standard definitions described in sufficient detail for the measurements to be replicated? Were the measurements done accurately? Were criteria / cut-off levels of categories well justified)
	Reference standard measured prior to Test? If not, was it measured blind to Test result?		Was reference standard administered whatever the test result and interpreted without knowledge of the test result? If not, was it likely to cause bias?
	Prevalence (pre-test probability) of Target disorder typical of usual practice?		Note: If prevalence (pre-test probability) of target disorder similar to usual practice, these data can be used to help determine post-test probabilities in practice (also need LRs)
	Maintenance in allocated groups and throughout study sufficient?		
	Proportion of intended participants receiving both Test and Reference Standard sufficiently high?		Was there a particular subgroup of the eligible participants not given either the Test or the Reference Standard? Was this sufficient to cause important errors?

	Change in Target disorder/Test status in period between Test and Reference Standard being administered		If there was a considerable delay between Test and Reference Standard then study could some new events have occurred or treatment may have been started that could influence the results of the Test/Ref Standard? If so, was this sufficient to cause important bias?
Outcomes	**blind or objective Measurement** of Outcomes: were they done accurately?		
	Test measured blind to Reference Standard status?		Were Testers aware of whether participants were Reference Standard positive or negative? If yes, was this likely to lead to biased measurement?
	Test measured objectively?		How objective was the Test measurements (e.g. automatic test, strict criteria)? Where significant judgment was required, were independent adjudicators used? Was reliability of measures relevant (inter-rater & intra-rater), & if so, reported?
	Test safe, available, affordable & acceptable in usual practice?		Would it be practical to implement this Test in usual practice? How safe, available, affordable & acceptable might it be?
Results	**ANalyses:** were they done appropriately?		
	If Ref Standard +ve & -ve groups not similar at baseline was this adjusted for in the analyses?		Some factors that differ between those with & without the target disorder could influence test accuracy (e.g. age, obesity, co-morbidities), although these are typically not reported.
	Estimates of Test sensitivity/specificity etc given or calculable? Were they calculated correctly?		Were raw data reported in enough detail to allow 2x2 tables to be constructed (i.e. TP, FP, FN & TN) in GATE frame & to calculate estimates of test specificity and sensitivity if entered into GATE calculator? Were GATE results similar to reported results?
	Measures of the amount of random error in estimates given or calculable? Were they calculated correctly?		Were confidence intervals &/or p-values for study results given or possible to calculate? If they could be entered into GATE calculator, were GATE results similar to reported results?
	Summary of Study Appraisal		
	Study design & conduct: was risk of error low (i.e. results reasonably unbiased)?		Use responses to questions in pink boxes above
	Study analyses: was risk of error low (i.e. results reasonably unbiased)?		Use responses from the orange boxes above
	Random error in estimates of intervention effects: were CIs sufficiently narrow for results to be meaningful?		Use responses to questions in green box above. Would you make a different decision if the true effect was close to the upper confidence limit rather than close to the lower confidence limit?
	Applicability: are these findings applicable in practice?		Use responses to questions in blue boxes above

Diagnostic test accuracy studies
Step 4: Apply. Consider/weigh up all factors & make (shared) decision to act

The X-factor



Epidemiological evidence: summarise the quality of the study appraised, the magnitude and precision of the measure(s) estimated and the applicability of the evidence. Also summarise its consistency with other studies (ideally systematic reviews) relevant to the decision.	Case circumstances: what circumstances (e.g. disease process/ co-morbidities [mechanistic evidence], social situation) specifically related to the problem you are investigating may impact on the decision?
System features: were there any system constraint or enablers that may impact on the decision?	What values & preferences may need to be considered in making the decision?

Decision: Taking into account all the factors above what is the best decision in this case?

Step 5: Audit usual practice (For Quality Improvement)

Is there likely to be a gap between your usual practice and best practice for the problem?