Cochrane Deutschland

LUNCH & LEARN

Systematic Reviews of Diagnostic Test Accuracy

Speaker: Mariska Leeflang, PhD, Assoc Prof.
(University of Amsterdam)

Host: Dr. Angelika Eisele-Metzger
(Cochrane Germany)

Gefördert durch:

Bundesministerium für Gesundheit

aufgrund eines Beschlusses des Deutschen Bundestages

Diagnostic test accuracy systematic reviews

Mariska Leeflang, PhD

Epidemiology and Data Science

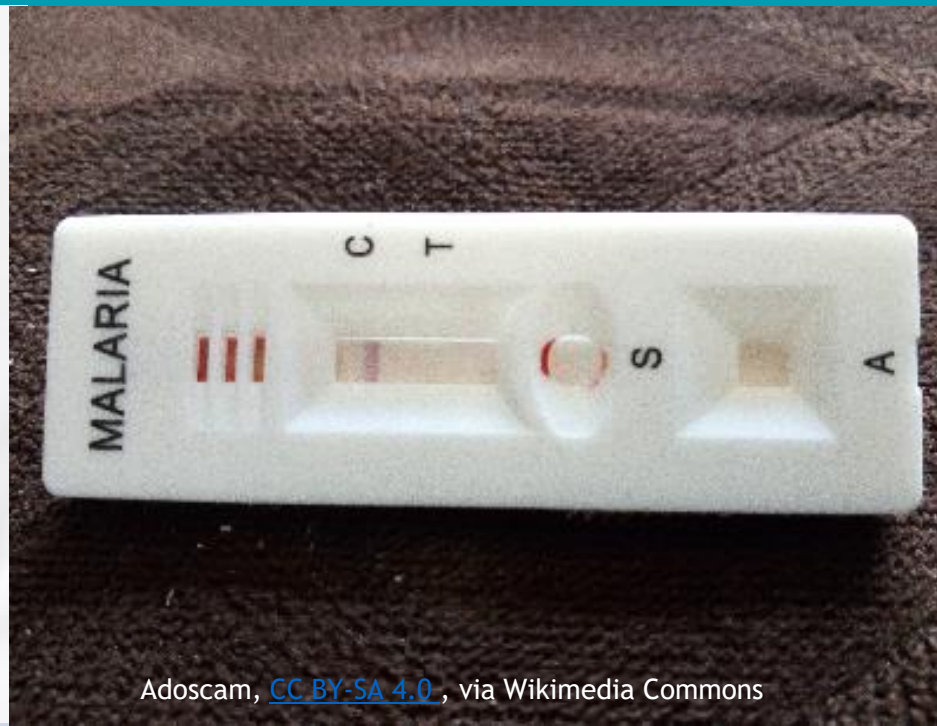
Amsterdam UMC, University of Amsterdam

m.m.leeflang@amsterdamumc.nl

X: @leeflang_m

Prabhu, Faqi, Khalifa and Awad, CC BY 4.0, via Wikimedia Commons

GOsites, CC BY-SA 3.0, via Wikimedia Commons



Adoscam, CC BY-SA 4.0 , via Wikimedia Commons

Amsterdam UMC
Universitair Medische Centra

Potential conflicts

Published:
Search strategy evaluations
QUADAS-2
QUADAS-C
QUAPAS

Systema ×

ng.cochrane.org/handbook-diagnostic-test-accuracy

Contact | Cochrane

Cochrane Training
Trusted evidence.
Informed decisions.
Better health.

Search...

Online learning | Learning events | Guides and handbooks | Trainers' Hub

Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy

Version 2.0, 2022

The *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* is the official guide that describes in detail the
and maintaining systematic reviews of test accuracy for Cochrane. The *Handbook* has been produced by the Cochrane Sc
Test Methods Group. It is a step-by-step guide for those conducting systematic reviews of test accuracy and a reference fo
authors.

Photo by M. Leeflang

Diagnostic tests (e.g. in EU IVD-R)

Technical or analytical validation	Clinical performance	Clinical utility

Diagnostic tests

Diagnostic test accuracy

	Target condition present	Target condition absent
Index test positive	TP	FP
Index test negative	FN	TN

Sensitivity

Specificity

Predictive Values

Likelihood Ratios

Steps in a systematic review

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Question formulation

Does the index test sufficiently well discriminate between people with and without the target condition?

Question formulation

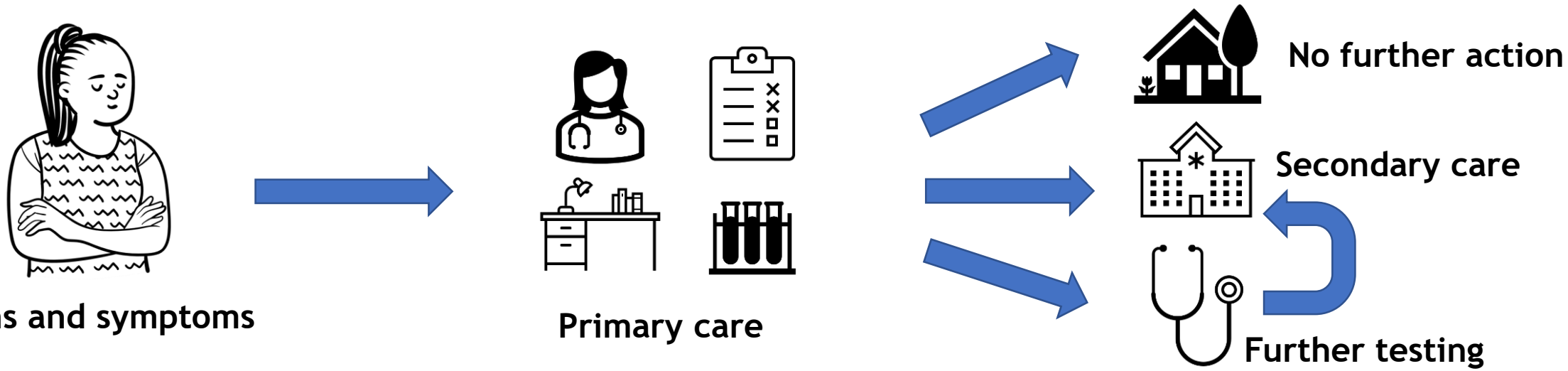
Does the index test sufficiently well discriminate between people with and without the target condition?

- “sufficient” depends on intended role of test
- Is this test fit for purpose?
- Which test is better fit for purpose?
- Wat ***is*** the purpose?

Question formulation: PICO → PIT

Patients	→	Who will be tested in practice? In which health care setting will test be used? What are previous/following steps?
Index tests	→	What is the relevant comparison? Variations of a test
Target Condition	→	What is the target condition? How is the target condition established?

Question formulation: Patients / Setting



Question formulation: examples

- Is serology sufficiently specific to serve as a triage test for Lyme disease in general practice?

- Is a modified two-tiered test strategy for Lyme disease sufficiently accurate to replace standard two-tiered testing in secondary care?

Question formulation: P I T

Is serology sufficiently specific to serve as a triage test for Lyme disease in general practice?

Is a modified two-tiered test strategy for Lyme disease sufficiently accurate to replace standard two-tiered testing in secondary care?

Steps in a systematic review

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Searching for diagnostic accuracy studies

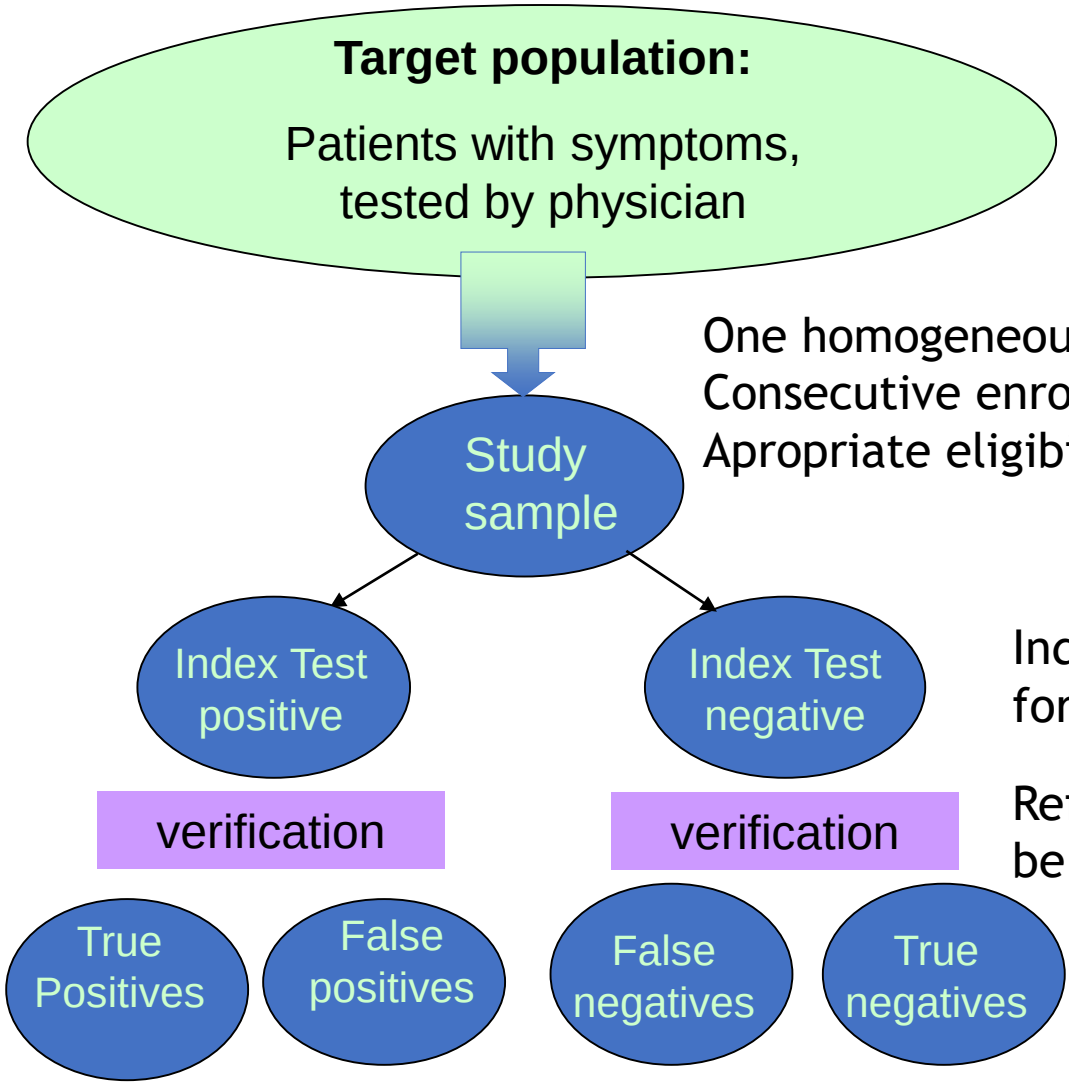
- Multiple synonyms and terms for target condition
- Multiple synonyms and terms for index test(s)

- If necessary / appropriate: terms for population or setting

- Be careful when using search filters: the more sensitive filters do not filter well

Expect about 2000-5000 titles and abstracts to screen!

Selection of studies



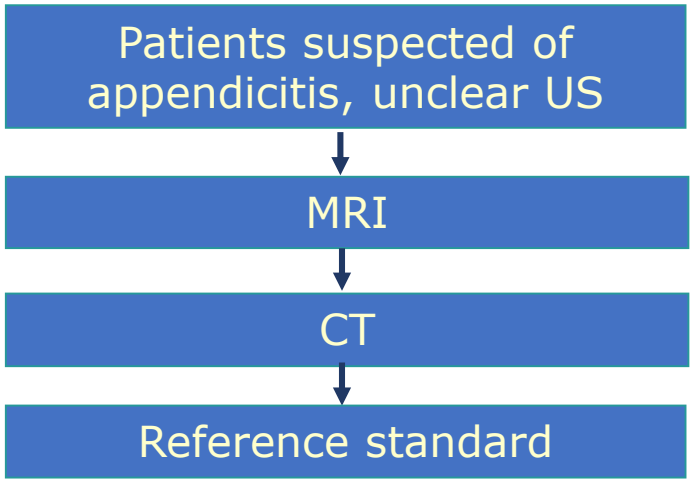
One homogeneous group
Consecutive enrollment
Apropriate eligibility criteria

Index test should be relevant for review question

Reference standard should be sufficiently reliable

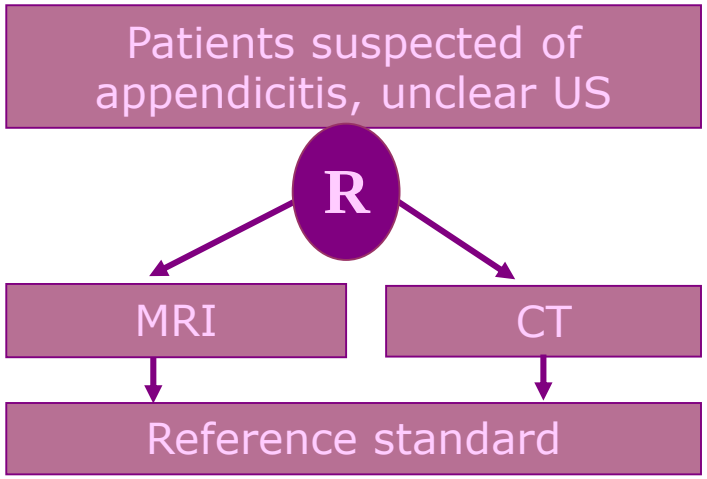
Selection of comparative accuracy studies

Fully paired design



Assumptions:
- MRI will not influence performance of CT
- Patients are representative to those tested in practice

Randomized design



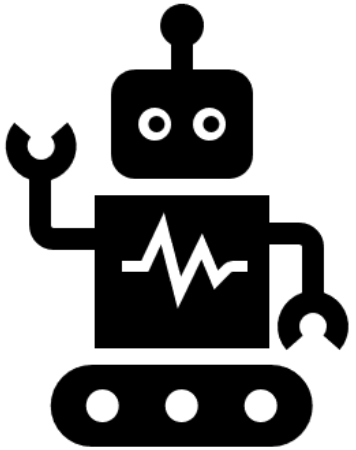
Randomized design:
- No pairwise comparisons possible
- Loss of power

Automated Search and Selection

- Several supporting systems rank titles and abstracts according to probability of being relevant

- Examples are:
  - EPPI Reviewer
  - AS Review
  - Rayyan

- The systems are of little use for small one-off reviews
- They may be useful for large and living reviews



Steps in a systematic review

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Quality assessment: QUADAS-2 and QUADAS-C

	Patient Selection	Index Test(s)	Reference Standard	Flow and Timing
Risk of Bias	Consecutive enrollment No case control Appropriate eligibility	Blinding Pre-specified cut-off	Likely to correctly classify the disease Blinding	Appropriate time interval Partial Verification Differential verification Missings
Concerns regarding applicability	Representative sample?	Index test same as in practice?	Appropriate target condition?	N/A

Signaling questions

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection:

- Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
- Was a case-control design avoided? Yes/No/Unclear
- Did the study avoid inappropriate exclusions? Yes/No/Unclear

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients (prior testing, presentation, intended use of index test and setting):

Is there concern that the included patients do not match the review question? **CONCERN: LOW/HIGH/UNCLEAR**

Steps in a systematic review

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Focus on sensitivity and specificity

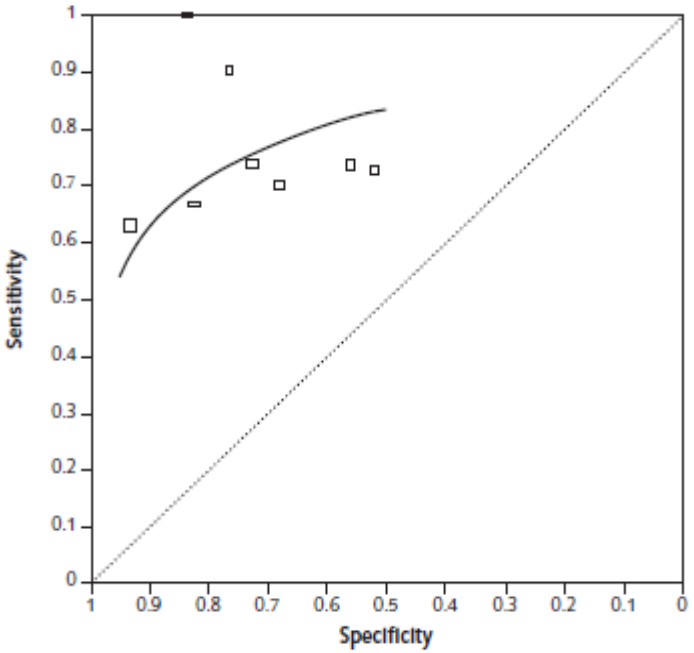
- Sensitivity and specificity are negatively correlated
- Their value depends on the chosen positivity threshold
- Sensitivity and specificity are typically heterogeneous

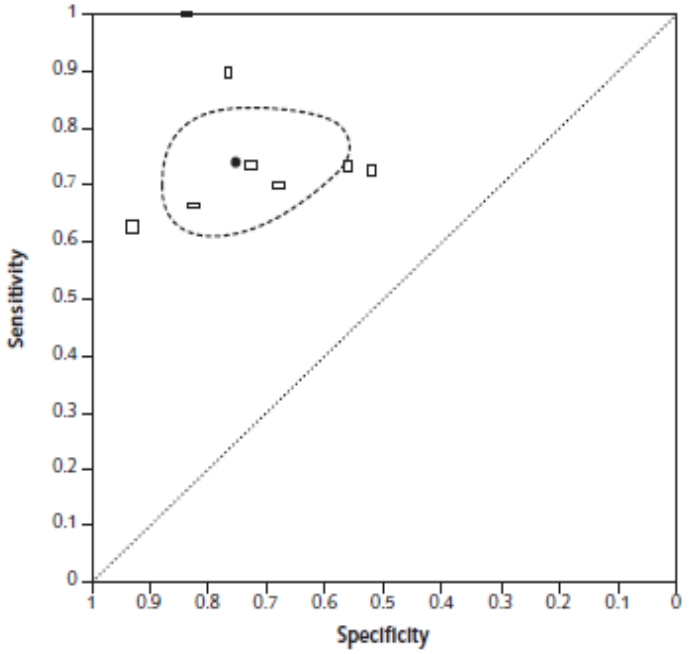
These factors should all be taken into account!

Hierarchical regression models needed

- Hierarchical / multi-level
  - allows for both within and between study variability
  - allows for correlation between diseased and non-diseased
- Logistic
  - correctly models sampling uncertainty
  - no zero cell adjustments needed
- Random effects
  - allows for heterogeneity between studies

Roughly two ways to summarize sensitivity and specificity



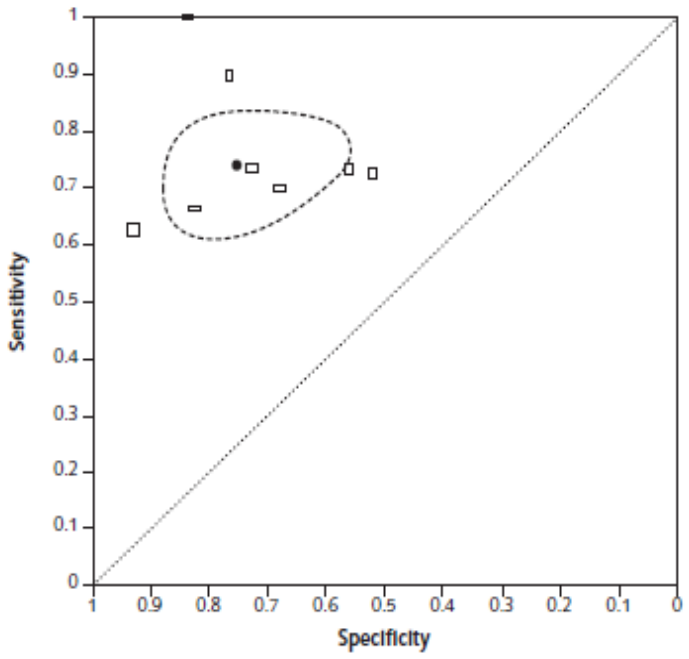


Figures made by M. Leeflang

Summary ROC curve

Summary sensitivity and specificity

Summary point estimates - bivariate model

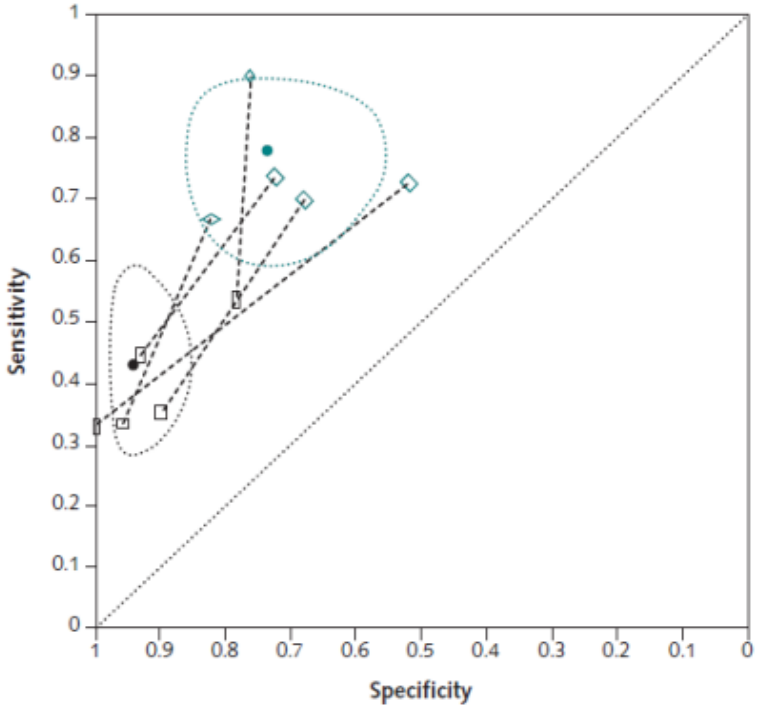


Bivariate

- Mean logit sens
- Variance logit sens

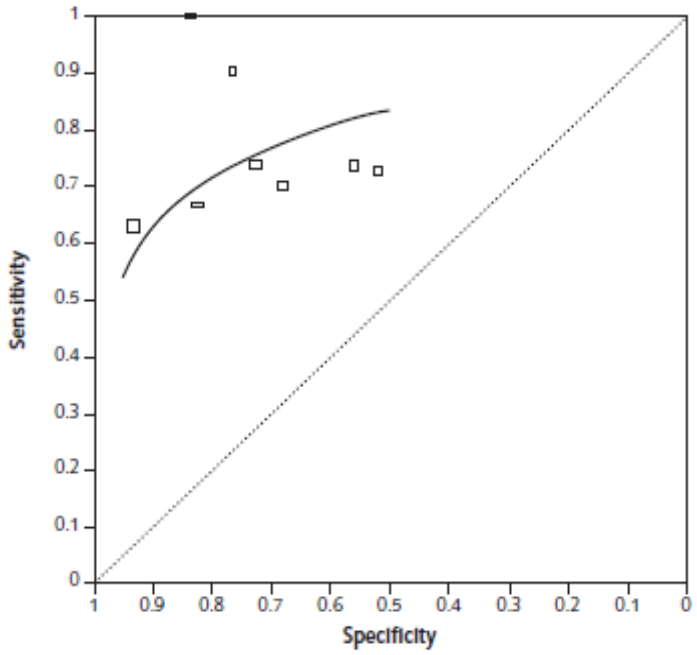
- Mean logit spec
- Variance logit spec

- Correlation between sensitivity and specificity



Figures made by M. Leeflang

Summary curve - HSROC model

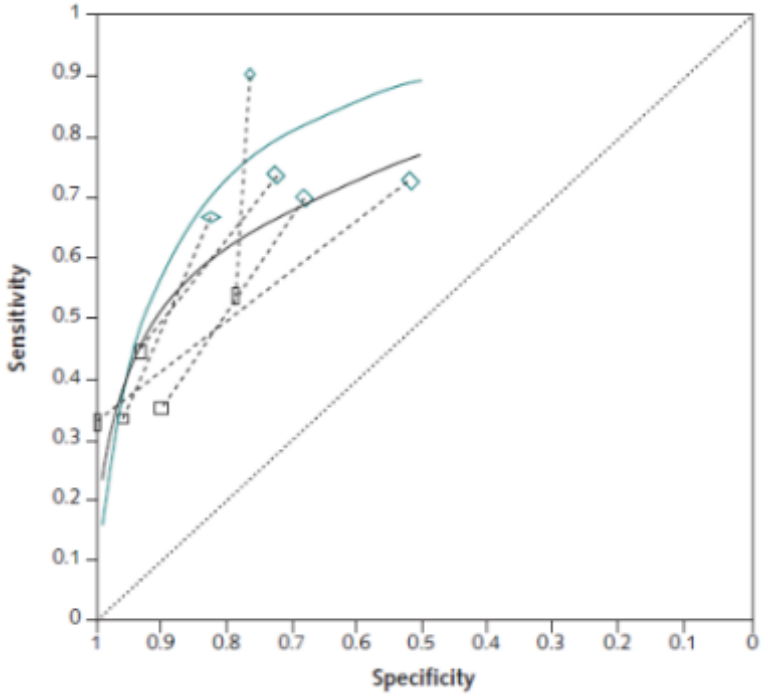


HSROC

- Mean lnDOR
- Variance lnDOR

- Mean threshold
- Variance threshold

- Shape of ROC



Figures made by M. Leeflang

Vast heterogeneity

- Differences in setting and population?
- Differences in test conduct?
- Differences in reference standard?
- Risk of bias?

Poor reporting complicates investigations

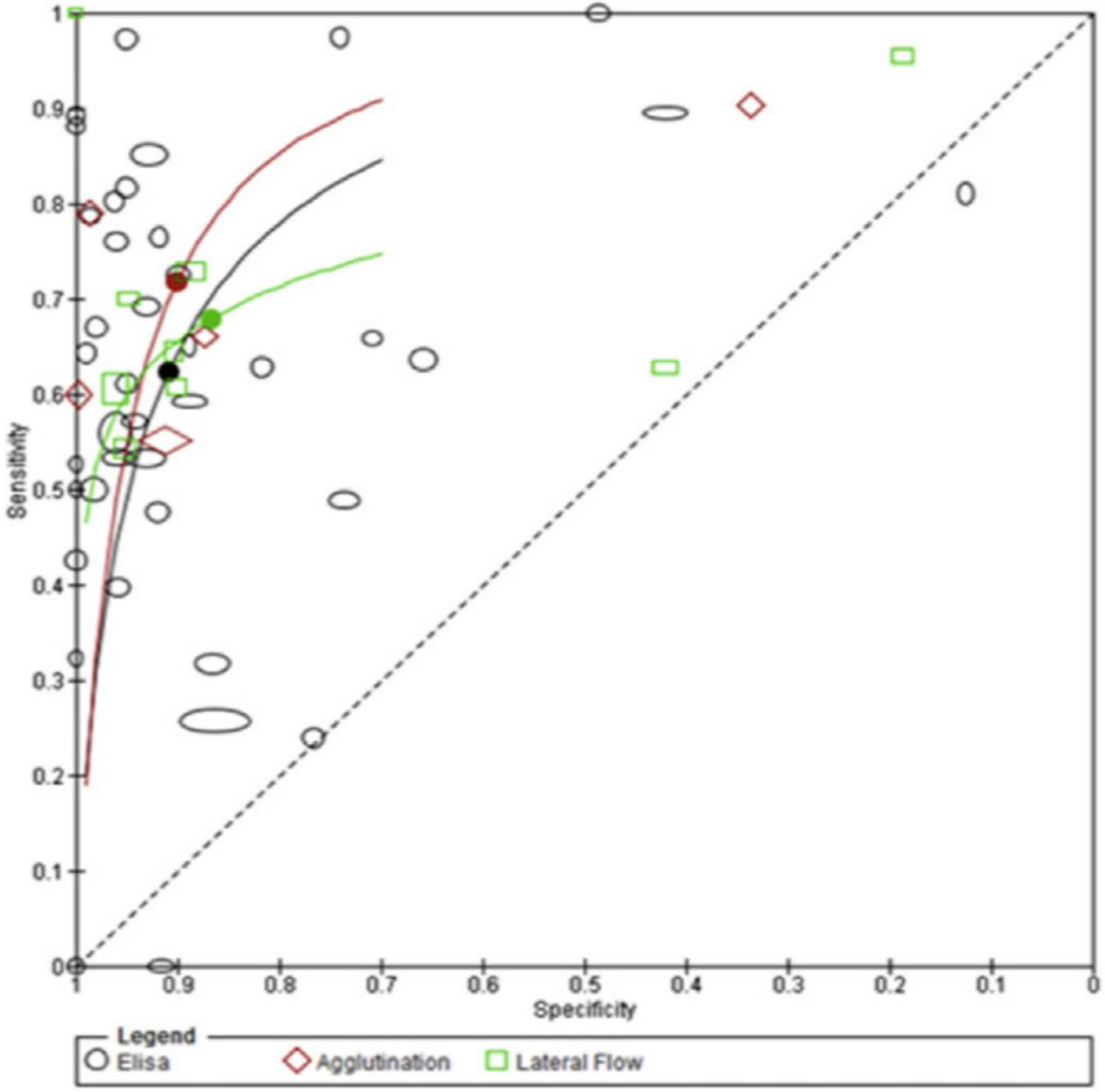


Figure made by M. Leeflang

Steps in a systematic review

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Interpretation and conclusions

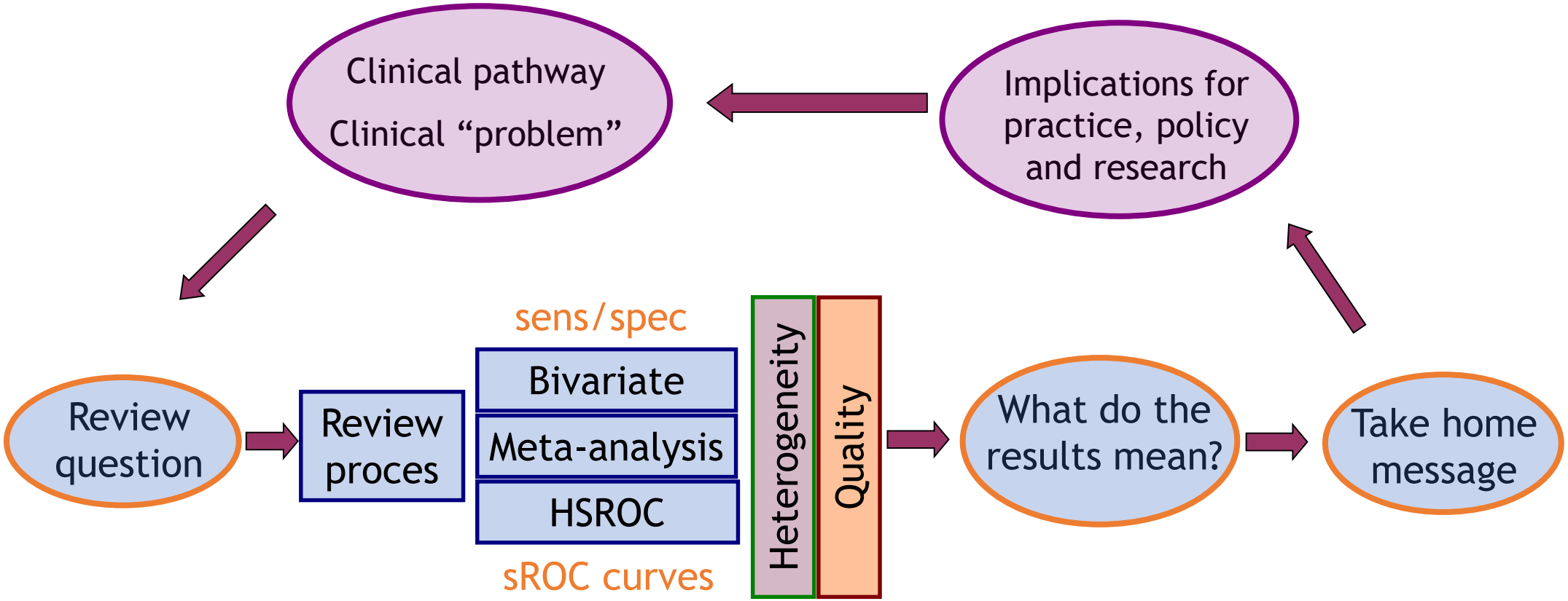


Figure made by M. Leeflang

Steps in a systematic review - questions?

1. Question formulation
2. Search and selection
3. Data-extraction and Quality assessment
4. Meta-analysis
5. Interpretation and Conclusions

Cochrane Deutschland

Lunch & Learn

HANDOUT

RECORDING

Further handouts and recordings of our past webinars:
https://www.cochrane.de

Next Lunch & Learn:

11. September 2024: Cochranes neues Review-Format *(in German)*

Cochrane Deutschland

Translations

Cochrane Kompakt

Social Media

Workshops & Events

Newsletter

Blog

WISSEN WAS WIRKT

Ressources

Cochrane Training

Cochrane Deutschland

https://www.cochrane.de/interessentenliste-workshops

Systematische Übersichtsarbeiten erstellen (online)

19.-20.09. & 24.-25.09.24

Systematische Reviews kritisch bewerten AMSTAR 2 (online)

08.11.2024

GRADE - Grundlagen (online)

GRADE

26.-28.11.24

RCTs kritisch bewerten RoB 2 (online)

04.12.2024

https://www.cochrane.de/veranstaltungen

Cochrane Deutschland

Evaluation

https://survey.lamapoll.de/Lunch_and_Learn_DTA_reviews

workshops@cochrane.de

Zuverlässige Evidenz.
Informierte Entscheidungen.
Bessere Gesundheit.

Gefördert durch:
Bundesministerium für Gesundheit
aufgrund eines Beschlusses des Deutschen Bundestages