UNIVERSITY OF BIRMINGHAM

Cancer diagnostics- opportunities and challenges

Prof Sudha Sundar

Professor of Gynaecological Oncology

University of Birmingham, United Kingdom

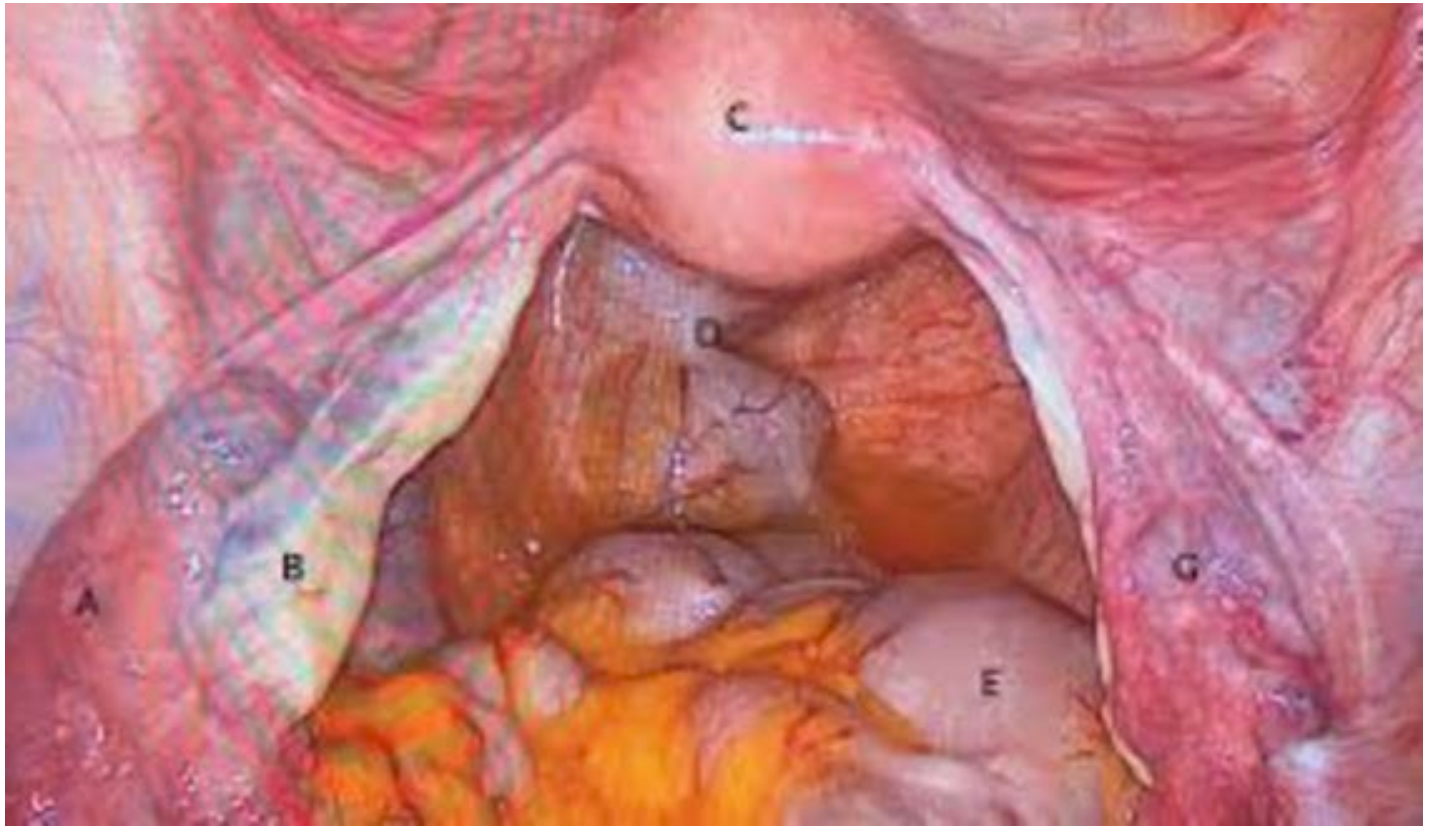
National Ovarian cancer audit lead, UK

BRITISH GYNAECOLOGICAL CANCER SOCIETY

Objectives

- Evaluating a new model/test for clinical practice
- Understand this in the context of cancer diagnosis
- Challenges of conducting such studies
- Example of Ovarian cancer

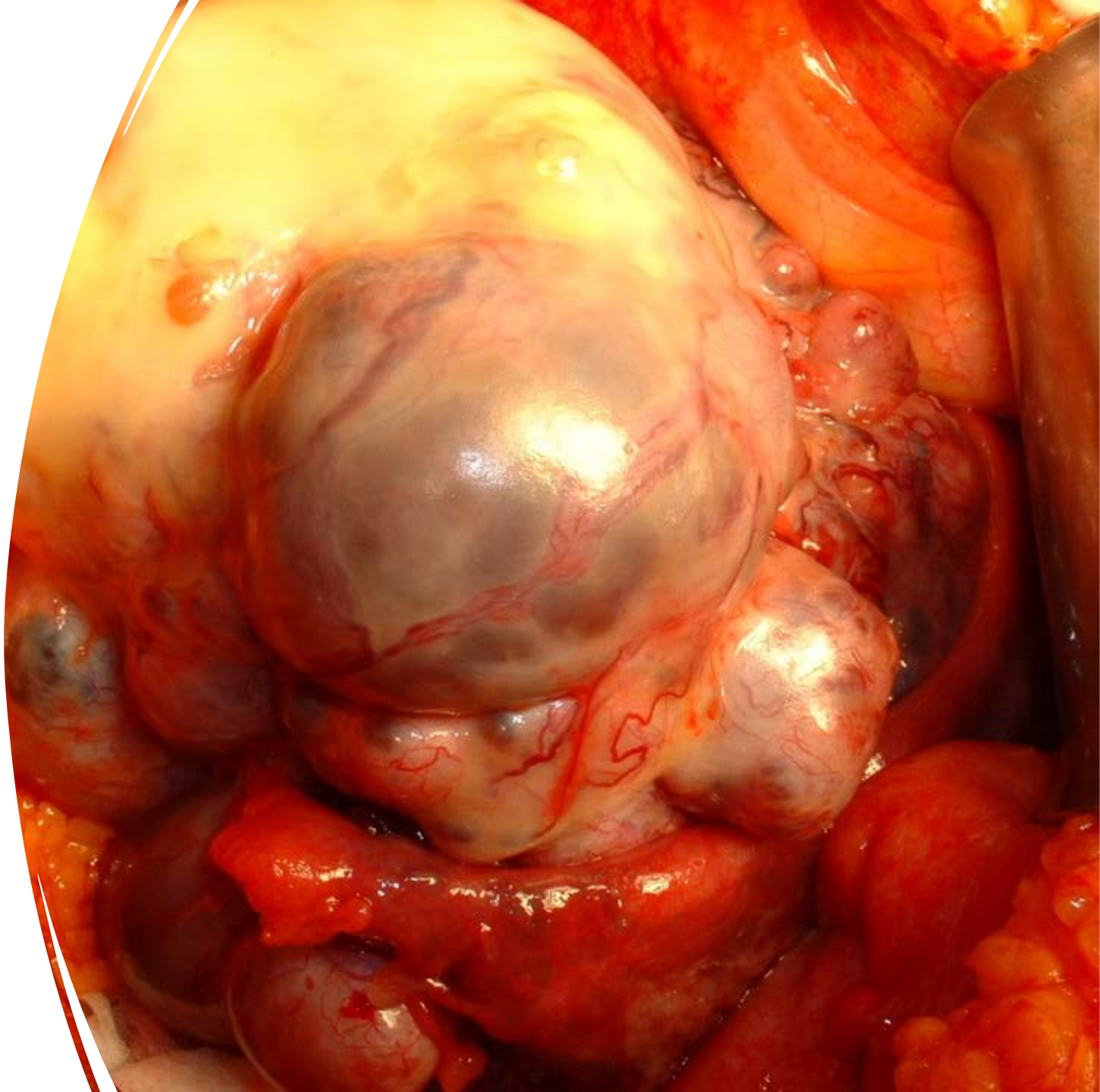
Ovaries



- Key roles in normal life – periods, children
- Situated in the pelvis
- Sisodia and Carmen, NEJM 2022

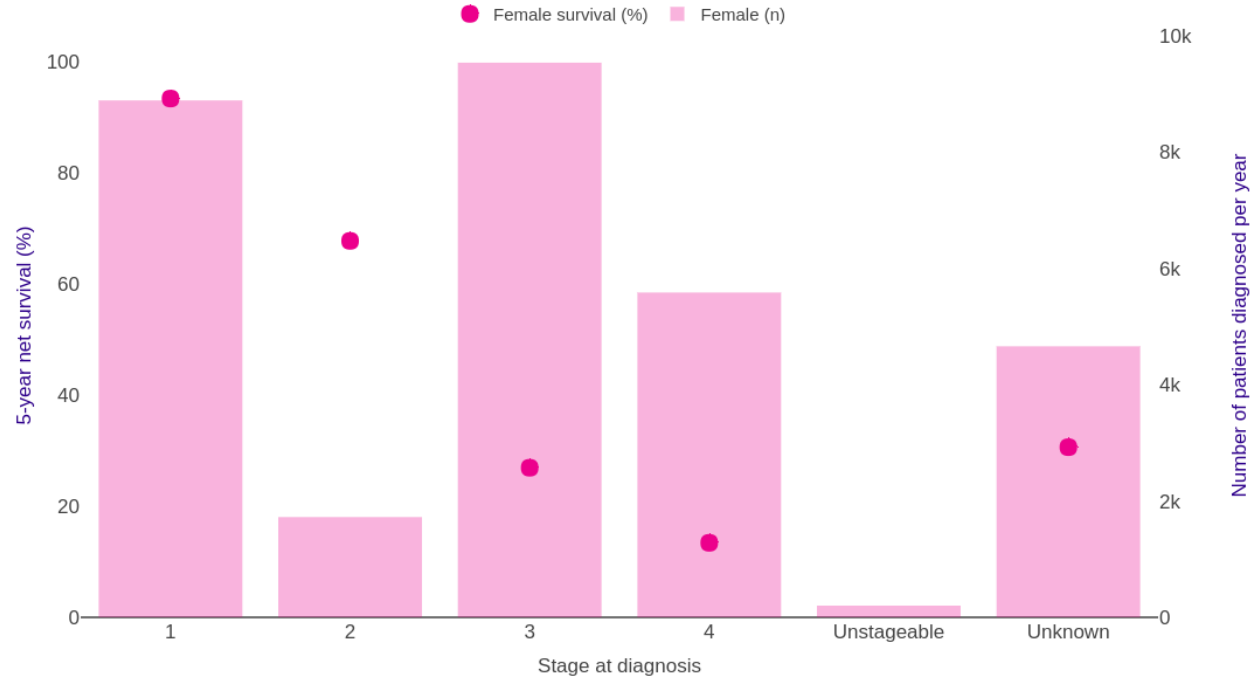
1. Epidemiology

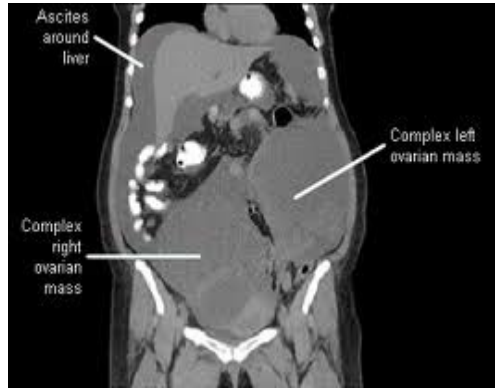
- Incidence – around 160,000 new patients/year worldwide
- Mortality is high – at best 30% 10 year survival

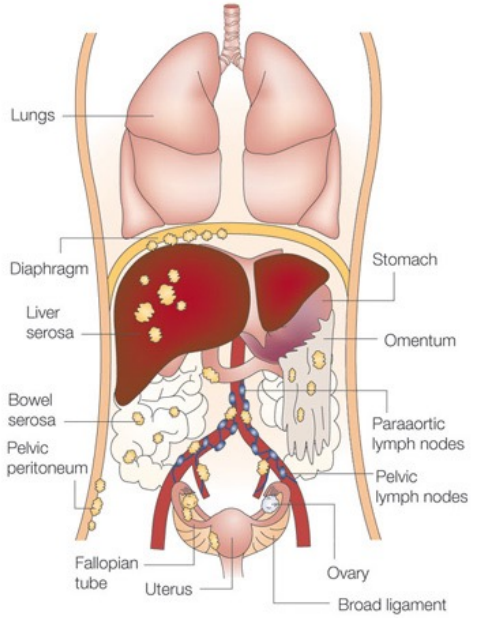


Ovarian Cancer (C56 – C57): data from 2013-2017

Five-Year Survival drops from 93% in Stage 1 to 13% in Stage 4







Nature Reviews | Cancer

CANCER RESEARCH UK

http://info.cancerresearchuk.org/cancerstats/faqs/#How

The National Cancer Registration Service, Eastern Office. Personal communication.

Prepared by Cancer Research UK
Original data source:

http://ecric.org.uk/

2. Challenges in Diagnosis of OC

- Low prevalence
- Not from the ovary!
- Natural history not well understood
- Heterogeneous disease
- No available screening tool

Prevalence of OC in primary care – 1 in 400 pts seen by GP.
Diagnosis of OC in pre-menopausal women poses additional challenges;
only 1 in 1000 symptomatic ovarian cysts are malignant, increasing to 3 in 1000 at age 50 years

SO WHAT DO YOU THINK ABOUT PREVALENCE OF SYMPTOMS IN COMMUNITY?

Improving Diagnosis

- Vague non-specific common symptoms
- Difficult cancer to diagnose in primary care
- Symptoms mimic common benign and physiological sensations
- Lack of patient and clinician awareness regarding symptoms
- Lack of accurate and specific tests to differentiate benign from malignant disease

- Persistent abdominal distension
- Feeling full and/or loss of appetite, pelvic or abdominal pain
- Increased urinary urgency and/or frequency
- Unexplained weight loss
- Fatigue
- Changes in bowel habit
- (or symptoms that suggest irritable bowel syndrome if they are over 50).

Several challenges

- Symptom recognition – presenting to GP
- Current diagnostic pathway – CA125 and ultrasound in primary care Care
- Ca125 nonspecific – misses 50% early-stage cancers
- USG not quality assured
- Most cysts are physiological and will resolve, even in postmenopausal women
- Incidence highest in postmenopausal women, but highest referrals in premenopausal women

Benign caused of raised Ca125

- Menstruation
- Endometriosis
- Any cause of peritoneal irritation
- Pelvic Inflammatory disease
- Diverticular disease

Conditions with increased serum CA-125 concentration (>35 U/mL) and ascites with or without pelvic masses

(Lancet Oncology 2007)

- **Non-malignant disorders**
- *Pelvic-mass associated*
- Multivisceral tuberculosis
- Meigs and pseudo-Meigs syndrome
- Ovarian hyperstimulation syndrome

- *Non-pelvic mass associated*
- Liver cirrhosis
- Tuberculosis peritonitis
- Uremia and renal failure
- Nephrotic syndrome
- Fulminant hepatic failure
- Pancreatitis

- **Malignant disorders**
- *Primary pelvic tumour*
- Ovarian cancer
- Advanced uterine cancer
- Advanced fallopian-tube cancer
- Advanced rectal or bladder cancer

- *Secondary pelvic involvement*
- Lymphoma with peritoneal involvement
- Pancreatic carcinoma
- Breast cancer with peritoneal metastasis
- Gastric cancer with peritoneal metastasis/ Advanced hepatocellular ca

Can We Introduce Better Tests? IOTA ultrasound models? Biomarkers?

IOTA simple rules,
Timmerman et al BMJ
Classify 77% of lesions
Sensitivity 91%, specificity 95%

He4 plus CA125
ROMA

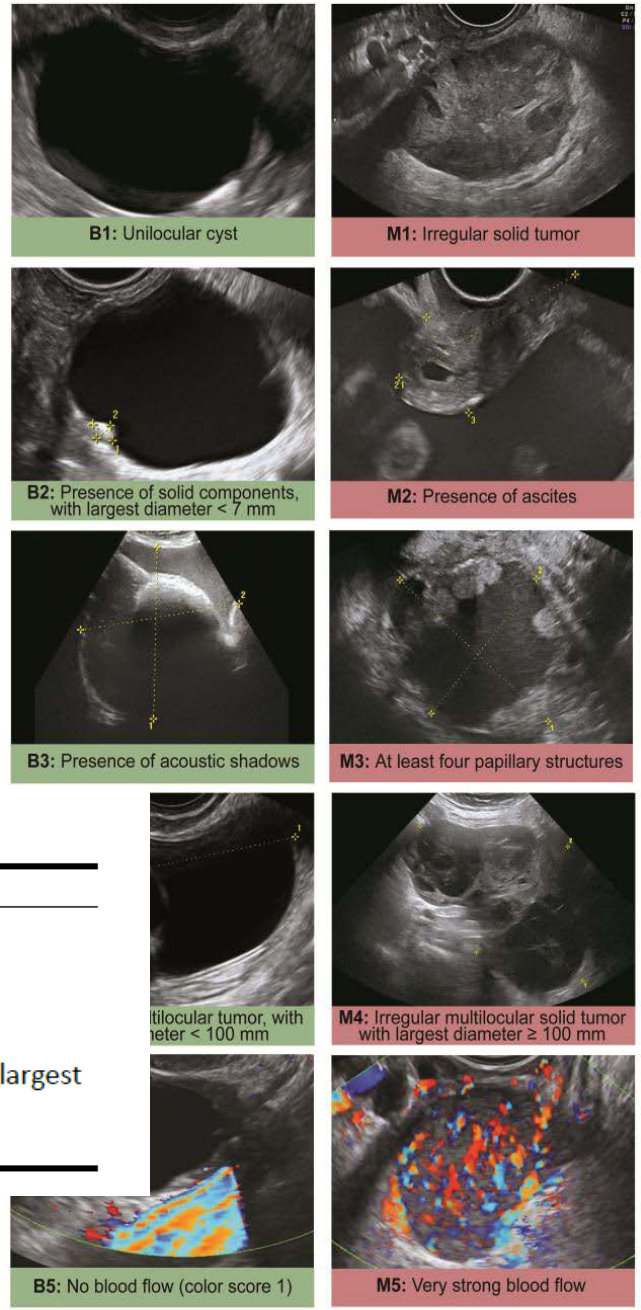


Table 1. IOTA group simple ultrasound rules

B-rules	M-rules
Unilocular cysts	Irregular solid tumour
Presence of solid components where the largest solid component < 7 mm	Ascites
Presence of acoustic shadowing	At least four papillary structures
Smooth multilocular tumour with a largest diameter < 100 mm	Irregular multilocular solid tumour with largest diameter ≥ 100 mm
No blood flow on colour Doppler	Prominent blood flow on colour Doppler

Validation studies

Translational pipeline for cancer diagnostic development and evaluation

Phase 1 Discovery science

- Nanotechnology - chem eng
- Big data - CLOCS

Phase 2 Tech Development

- Genomics - ROCkeTS GEN
- Metabolomics - STEMOVA

Phase 3 Clinical trials

- ROCKETS

Phase 4 Implementation science

- SONATA

Interdisciplinary collaborations – laboratory research – clinical trials units – large scale primary care collaborations

Broad range of methodologies – computation – lab - systematic reviews – clinical trials - cost effectiveness

Interrogating diagnostic pathways starting from the individual- community – primary-secondary through to tertiary care

the eve appeal

AOA Dx

CANCER RESEARCH UK

STAND UP TO CANCER

WELLBEING OF WOMEN

BWCH Global Research Fund

FUNDED BY NIHR | National Institute for Health and Care Research

Shopping can save your life!

JMIR PUBLIC HEALTH AND SURVEILLANCE Brewer et al

Original Paper

Association Between Purchase of Over-the-Counter Medications and Ovarian Cancer Diagnosis in the Cancer Loyalty Card Study (CLOCS): Observational Case-Control Study

Hannah R Brewer[1], BA, PhD; Yasemin Hirst[2], BSc, PhD; Marc Chadeau-Hyam[3], BA, PhD; Eric Johnson[3], BA; Sudha Sundar[4], MPhil, MBBS; James M Flanagan[4], BSc, PhD

[1]Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom
[2]Department of Behavioural Science and Health, University College London, London, United Kingdom
[3]School of Public Health, Faculty of Medicine, Imperial College London, London, United Kingdom
[4]Division of Cancer, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom

There is a difference in purchases of pain and indigestion medications among women with and without ovarian cancer up to 8 months before diagnosis. Facilitating earlier presentation among those who self-care for symptoms using this novel data source could improve ovarian cancer patients' options for treatment and improve survival.

Imperial – PI - Dr James Flanagan, Imperial, Dr Hannah Brewer
UCL/UCLAN – Dr Yasemin Hirst

Brewer HR, …Sundar S, .. Hirst Y.
Cancer Loyalty Card Study (CLOCS): feasibility outcomes for an observational case-control study focusing on the patient interval in ovarian cancer. BMJ Open. 2023 :

Other end of spectrum

Cochrane Database of Systematic Reviews | Review - Diagnostic

Menopausal status, ultrasound and biomarker tests in combination for the diagnosis of ovarian cancer in symptomatic women

Clare Davenport, Nirmala Rai, Pawana Sharma, Jonathan J Deeks, Sarah Berhane, Sue Mallett, Pratyusha Saha, Rita Champaneria, Susan E Bayliss, Kym IE Snell, Sudha Sundar Authors' declarations of interest

Version published: 26 July 2022 Version history

https://doi.org/10.1002/14651858.CD011964.pub2

- 59 studies (32,059 women, 9545 cases of OC)
- In postmenopausal women, both ROMA and IOTA ADneX are more sensitive, ROMA is as specific as RMI
- In premenopausal women, RMI has neither acceptable sensitive nor specific and should not be used.
- Both ROMA or ADneX are more sensitive tests in premenopausal women with a reduction in specificity

- **Conclusion – RMI needs to be changed to better testing strategies (ROMA or ADneX)**

So how should we design a study for a new ML test to diagnose ovarian cancer ?

What questions might we have about the new model/test?

- “If I repeat a new model 10 times, how similar will the results be?”
- “How well does the model differentiate between people with and without OC?”
- “On top of clinical information that I already have (from history and examination and CA125), do these models add much?”
- How similar is the experimental cohort to the model generation cohort?
- “If I introduce these models into clinical practice, will it reduce morbidity and mortality?”

Analytical validity

Diagnostic accuracy

Clinical impact

Stages of the Test Evaluation Pathway

Analytical validity

- Sources of variability
- Reliability (repeatability and reproducibility)
- Measurement accuracy

Clinical / Diagnostic validity

- Test accuracy
- Comparative/incremental test accuracy

Impact

- Change in diagnostic yield
- Change in therapeutic yield
- Change in patient outcomes

What type of study?

- Is there a single study that would answer all of the questions we have about the introduction of the model to improve the detection of Ovarian Cancer or should we undertake a series of studies?

Stages of the Test Evaluation Pathway

Analytical validity

- Sources of variability
- Reliability (repeatability and reproducibility)
- Measurement accuracy

Clinical / Diagnostic validity

- Test accuracy
- Comparative/incremental test accuracy

Impact

- Change in diagnostic yield
- Change in therapeutic yield
- Change in patient outcomes

Analytic validity

Quality assurance

- Subjective versus objective tests
- Inbuilt QA for subjective performance rare in medicine
- Often not transparent
- Cohort conditions can change

EG: Analytic validity factors

Is the test affected by?

- Patient factors

-Time of day, fasting, exercise, tea/coffee, stimulants, other drugs, menstrual cycle, BMI (imaging)

- Sample collection and handling/ Ultrasound Image acquisition & interpretation (personnel competency)

- Sampling technique, delay to testing, temperature, freezing, skill of USS operator.

Reliability and sources of variability

- **Repeatibility**
  - variation in measurements taken by a single person on the same item and under the same conditions
  - test-retest studies
  - Estimates of measurement error and coefficient of variation
- **Reproducibility**
  - agreement between measurements conducted on replicate specimens/ the same images in different healthcare settings by different people.
- **Measurement accuracy**
  - Does the measurement agree with a gold standard quality assured laboratory assay/ USS scoring system?
  - Bland-Altman plots and analyses
- **Any safety analyses**

Stages of the Test Evaluation Pathway

Analytical validity

- Sources of variability
- Reliability (repeatability and reproducibility)
- Measurement accuracy

Clinical /Diagnostic validity

- Test accuracy
- Comparative/incremental test accuracy

Impact

- Change in diagnostic yield
- Change in therapeutic yield
- Change in patient outcomes

What is test/model accuracy?

A comparison between …..

The prevalence of a disease state (target condition) estimated by a test of interest (“the index test”)

&

The best estimate of the true disease state

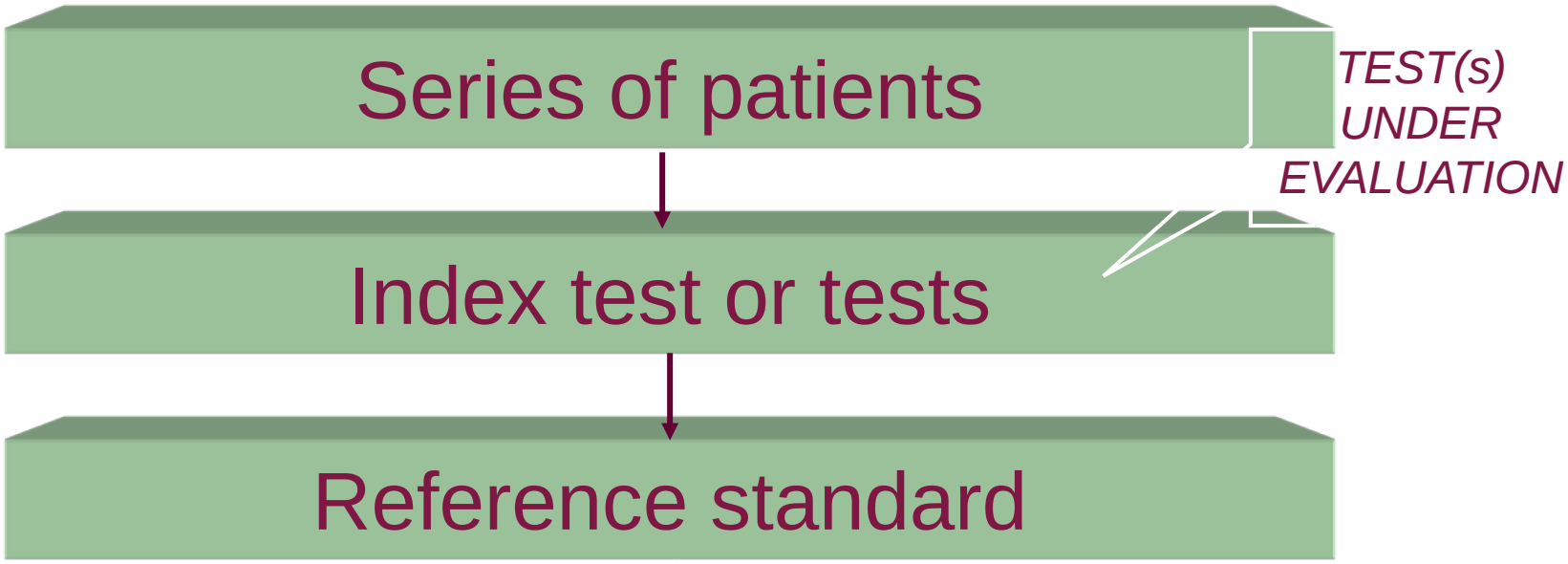
(“the reference standard”)

Against

the best available comparator (standard of care test)

Evaluation of test accuracy is an explicit recognition that most tests make errors even if correctly performed

Test accuracy



VERIFICATON OF INDEX TEST RESULTS WITH A REFERENCE STANDARD : BEST WAY TO DETERMINE PRESENCE OR ABSENCE OF DISEASE

Components of a test accuracy question

Patients:

-presentation: symptomatic / asymptomatic; setting (primary care / hospital)

-prior tests: received prior to the index test being evaluated

Index test: the test or tests being evaluated

Target condition: what is the test trying to detect

Reference standard: used to verify the results of the index test

Outcomes: accuracy (sensitivity, specificity, predictive values…..)

What is test accuracy?

- **Diagnosis**
  - Does this patient have this disease at this point in time?
- **Test accuracy**
  - What proportion of those with the disease does the test detect? (**sensitivity**)
  - What proportion of those without the disease get negative test results? (**specificity**)
  - Requires 2×2 table of test *vs* reference standard

Expressing Test Accuracy Numerically

Test Accuracy Study Design

Study sample

Index Test

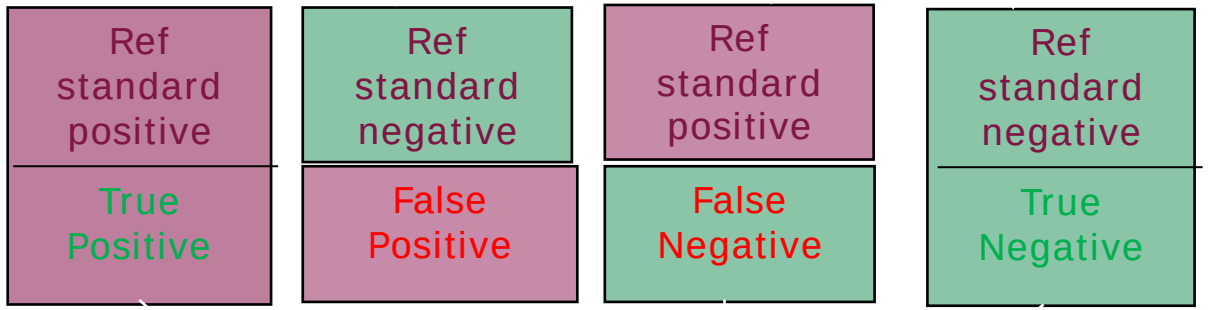
Reference standard applied regardless of index test result

Index Test positive

Index Test negative

Reference standard

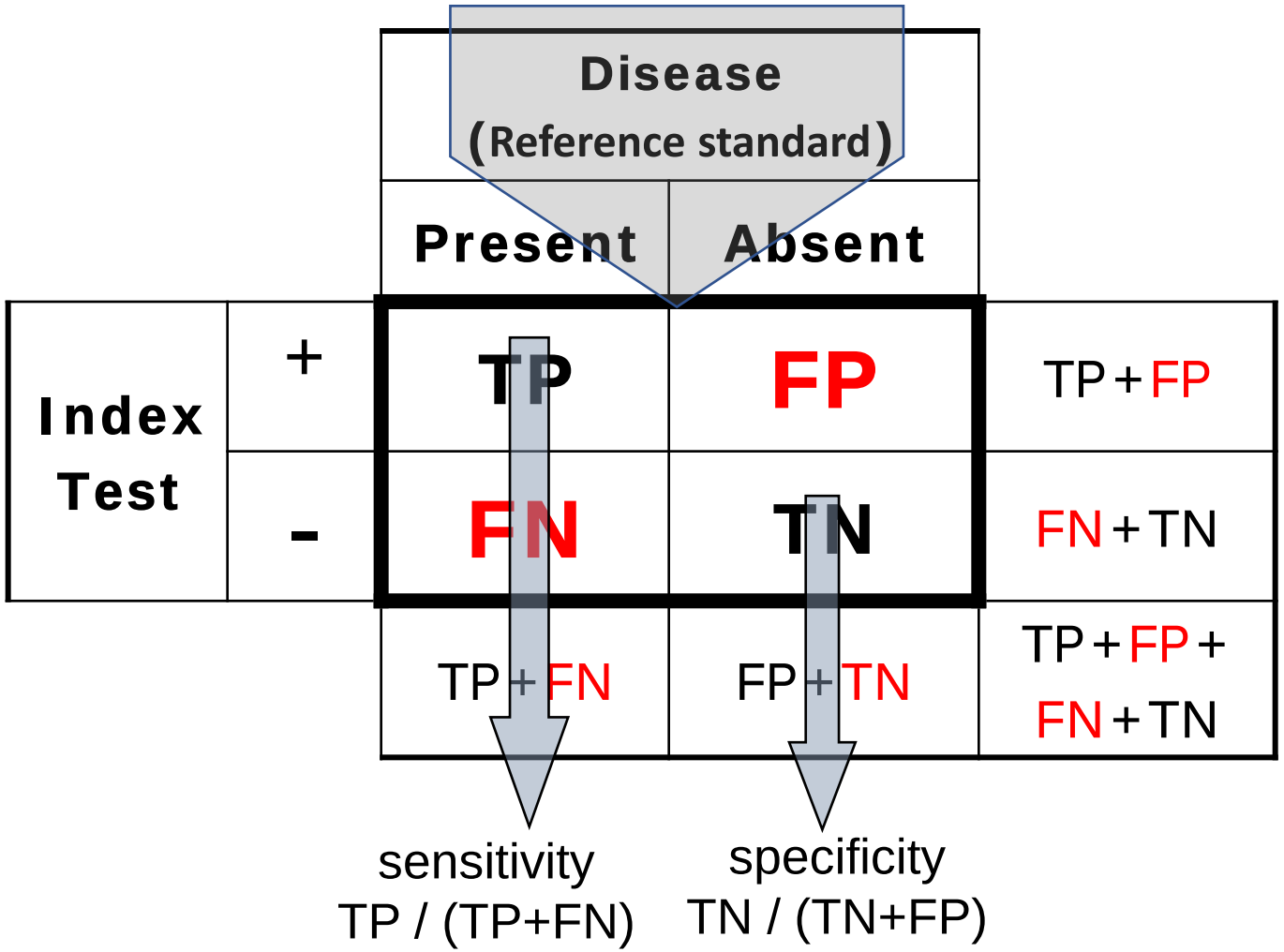
Reference standard



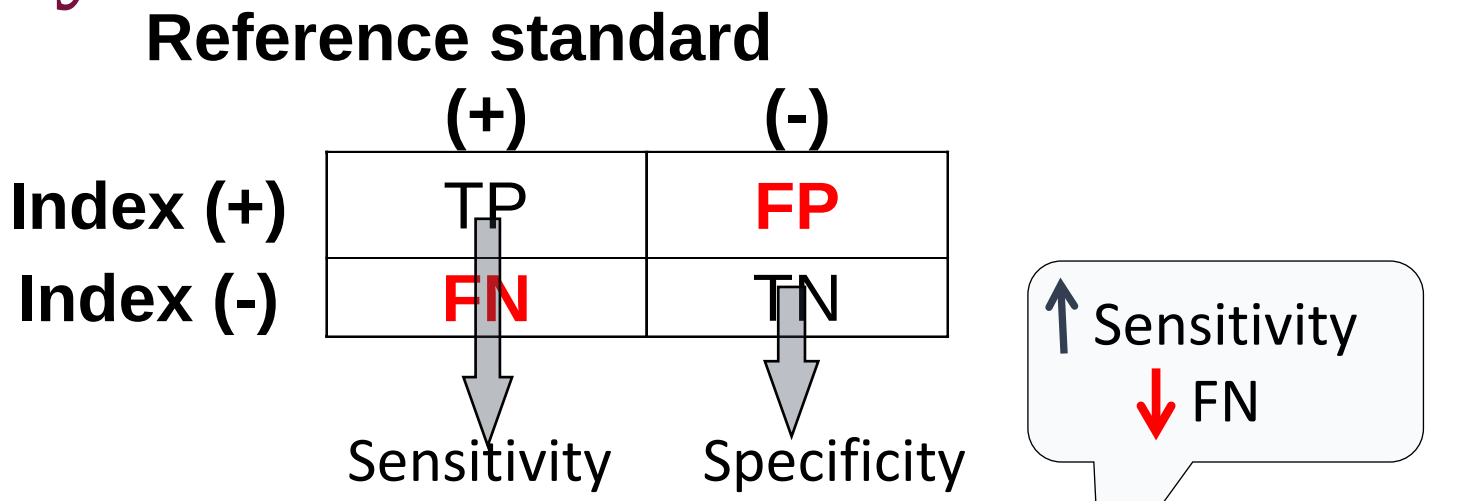
Reference standard and index test results produce 4 subgroups

Estimation of test accuracy
e.g. sensitivity and specificity

Summary test accuracy metrics: calculating sensitivity and specificity



Test accuracy: sensitivity and specificity



- Sensitivity: what proportion of those with the target condition have positive index test results?
- Specificity: what proportion of those without the target condition have negative index test results?
- The nearer either sensitivity or specificity are to 1 or 100%, the better the test

Other measures of diagnostic accuracy

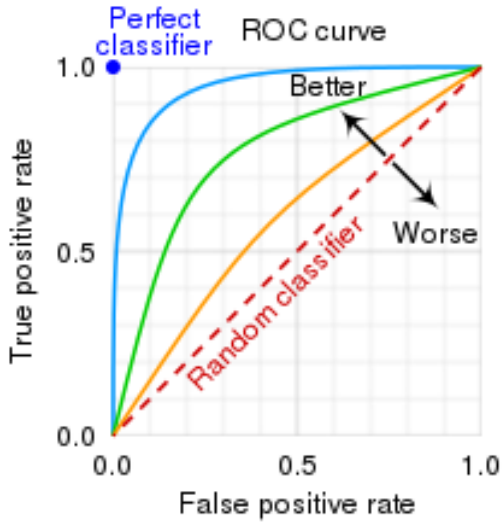
↑PPV
↓ FP

↑NPV
↓ FN

- Positive and negative predictive values – the probability for an individual with a positive test result having the disease (PPV). The probability of an individual with a negative test result not having the disease (NPV)

- PPV and NPV are prevalence dependent

- Area under the ROC curve
- Diagnostic odds ratio



What are the consequences of negative and positive test results?

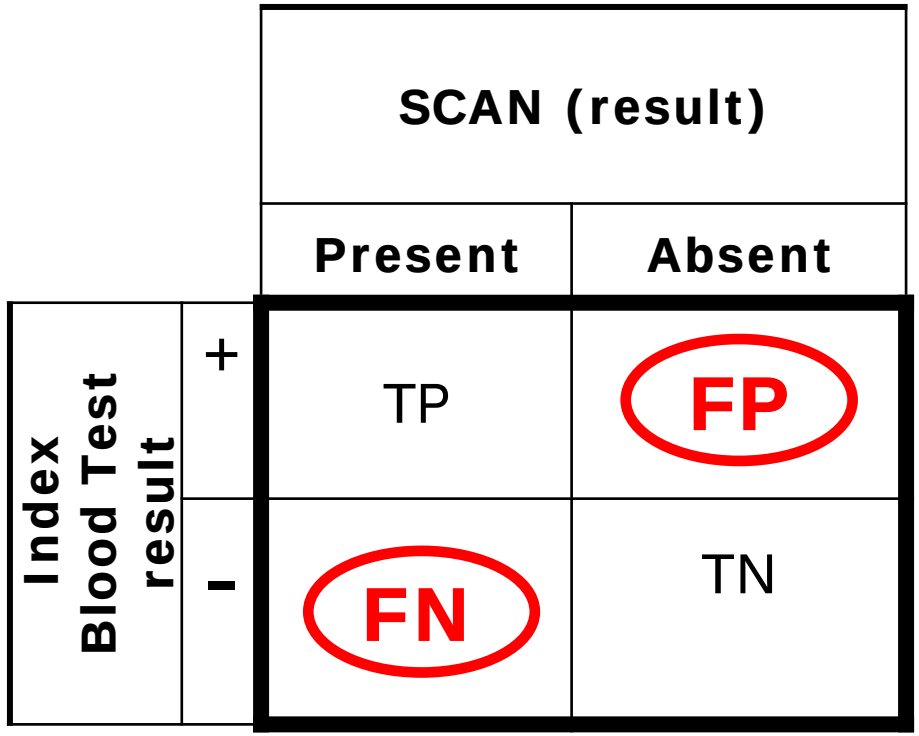
In most testing situations **ONE OR OTHER** of **F**alse **N**egative or **F**alse **P**ositive test errors are more important....

……maximise sensitivity / NPV and minimise false negative test errors

OR

……maximise specificity/PPV and minimise false positive test errors?

Test Errors: model to detect ovarian cancer



FP: combination test positive but no ovarian cancer
CONSEQUENCES: unnecessary further investigation. Patient anxiety. Opportunity costs.

FN: combination test negative but ovarian cancer present
CONSEQUENCES: delayed or missed diagnosis with associated morbidity and mortality

Test Accuracy studies

- What does the test accuracy study tell us?
- What does the test accuracy study not tell us?

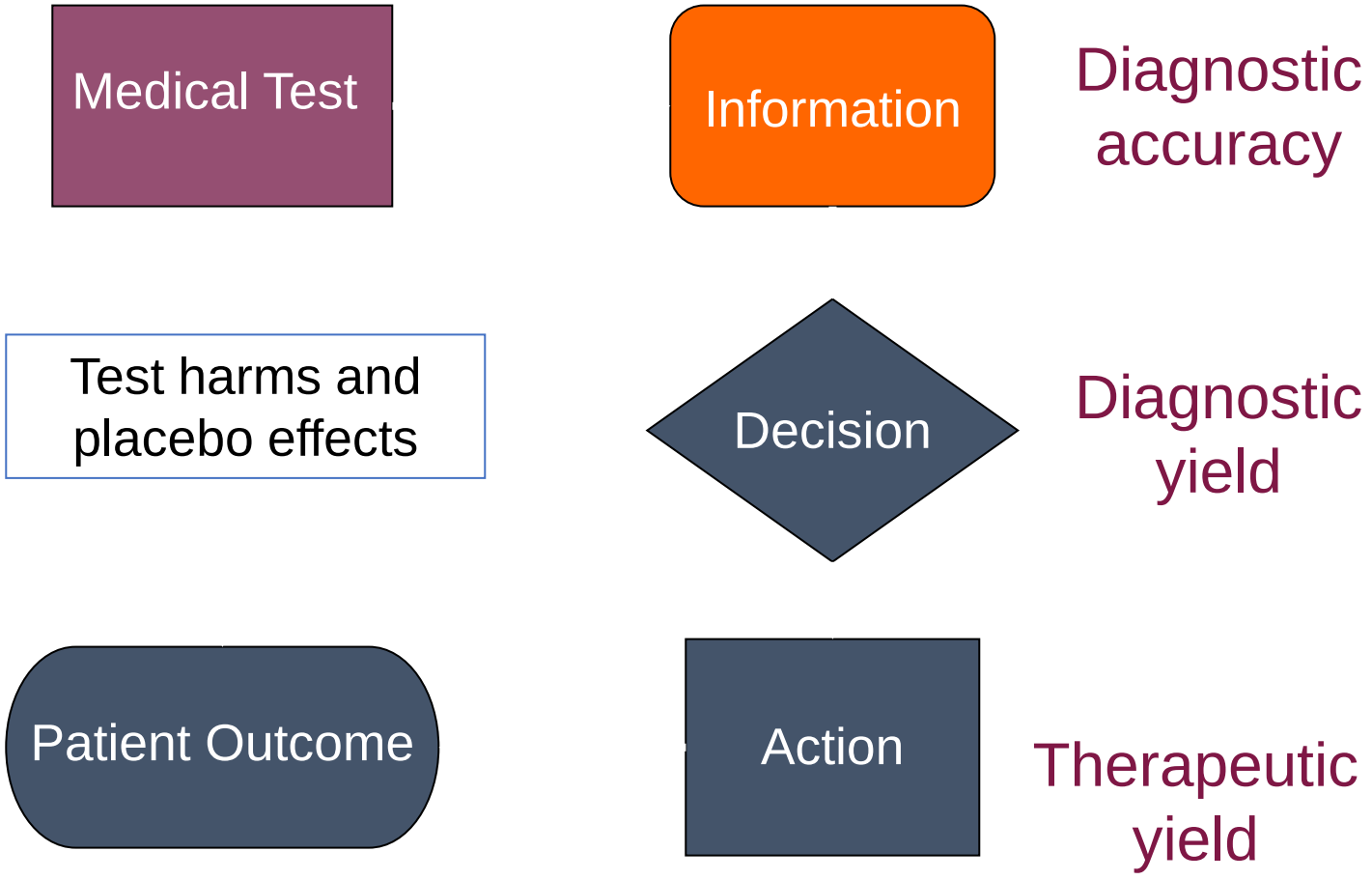
Test Accuracy studies

- What does the test accuracy study tell us?
- What does the test accuracy study not tell us?

Stages of the Test Evaluation Pathway

- **Analytical validity**
  - Sources of variability
  - Reliability (repeatability and reproducibility)
  - Measurement accuracy
- **Clinical / Diagnostic validity**
  - Test accuracy
  - Comparative/incremental test accuracy
- **Impact**
  - Change in diagnostic yield
  - Change in management
  - Change in patient outcomes

Diagnostic Evaluative framework



Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Medical Decision Making 1991;11(2):88-94.

Take home messages

- Test evaluation has multiple stages
- Test accuracy is only one element of test evaluation
- The pace of technological advances in devices pose a challenge to rigorous test evaluation using clinical trials
- Test treat trials (direct evidence) of the clinical impact of a test on patient outcomes are rare
- Decision models can be used in place of test treat trials but their validity depends on the quality and inter-connectedness of existing evidence.

Summary

- Diagnostic accuracy of the model is one component of evaluation
- Placing the model in context of clinical use is critical

Challenges of answering the research question – model evaluation in practice

- Securing funding for the study
- Designing the study – sample size/study design
- Designing the CRF (data collection form)
- Data standardization – doesn’t exist
- Data defined differently by research nurses and patients
- Long trials so algorithms change along the way
- Complex pathways so plenty of ‘loss’ within the study – not real loss but data loss by statisticians
- Clinical sense checking of data

Moral of the story

- For God’s sake talk to the medics
- Patient – public engagement in your research
- DOMAIN specialist is key
- Treat the person giving you the data nicely!

Further reading

- Common pitfalls in statistical analysis: Understanding the properties of diagnostic tests. https://pubmed.ncbi.nlm.nih.gov/29430417/

Further Reading:

- Deeks JJ. Assessing outcomes following tests. In: Price CP and Christenson RH, editors. *Evidence Based Laboratory Medicine*. 2nd Ed. Washington: AACC Press; 2007. p.95-111.
- Ferrante di Ruffano L, et al. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. BMJ 2012;344:e686.
- Hunink MGM, Krestin GP. Study design for concurrent development, assessment and implementation of new diagnostic imaging technology. Radiology 2002; 222: 604-14.
- Lijmer JG, Bossuyt PMM. Various randomized designs can be used to evaluate medical tests. J Clin Epidemiol 2009;62:364-373.
- Lord SJ, et al. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomised trials? Ann Intern Med 2006; 144(11): 850-5
- Lord SJ, et al. Using the principles of randomized controlled trial design to guide test evaluation. Med Dec Mak 2009; 29: E1-12.
- Plevritis SK. Decision analysis and simulation modelling for evaluating diagnostic tests on the basis of patient outcomes. Am J Roentgenol 2005; 185:581-590
- Tatsioni A, et al. Challenges of systematic reviews in diagnostic tests. Ann Intern Med 2005;142:1048-1055.

Thank you!

- s.s.sundar@bham.ac.uk