Do doctors understand test results?

By William Kremer
BBC World Service

Are doctors confused by statistics? A new book by one prominent statistician says they are - and that this makes it hard for patients to make informed decisions about treatment.

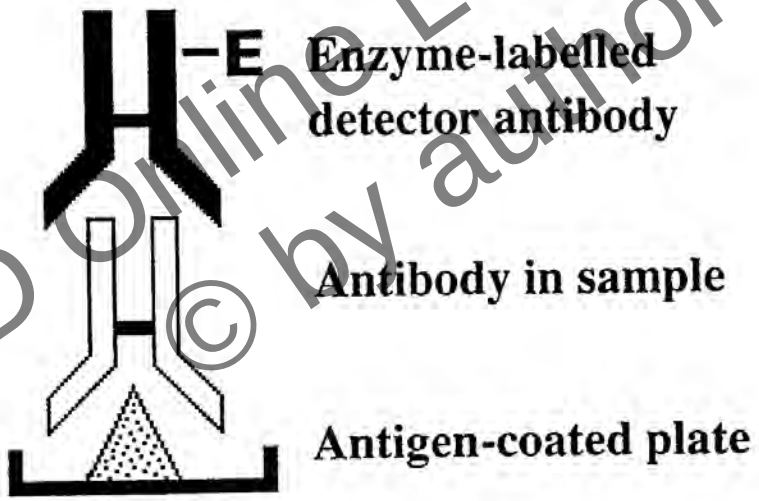
Diagnostic tests

Eskild Petersen

Direct ELISA

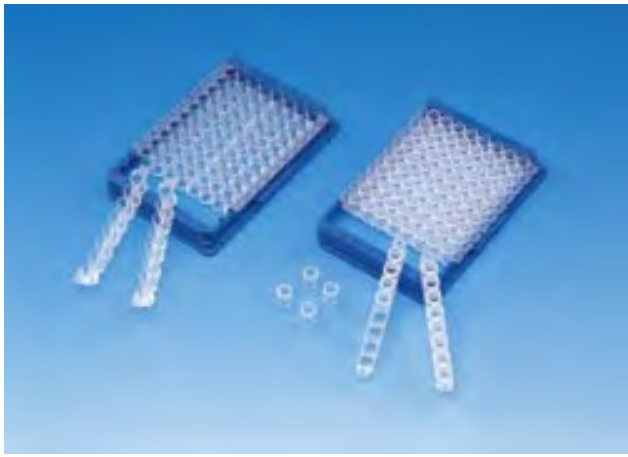
ELISA = Enzyme Linked ImmunoSorbent assay =
EIA = Enzyme ImmunoAssay

Ex. Goat-anti Human-IgG



E = enzym, biotin, flourescin

Ex. IgG in a Blood sample



How to evaluate a diagnostic test

	Disease		
Test	**Yes**	**No**	Total
Positive	True pos. (a)	False pos.	a+b
Negative	False neg.	True neg.	a+d
Total	a+c	b+d	N

Sensitivity and specificity

- Sensitivity is the ability of the test to find patient who has the disease – true positive
- Specificity is the ability of the test to find patiens who do NOT have the disease – true negative

Sensitivity and specificity

	Disease		
Test	**Yes**	**No**	Total
Positive	a	b	a+b
Negative	c	d	c+d
Total	a+c	b+d	N

Sensitivity = a/a+c

Specificity = d/b+d

Ex: Tonsillitis acuta

Clinical Diagnosis	Tons. acuta (bacterial culture) Yes	No	Total
Positive	27	35	62
Negative	10	77	87
Total	37	112	149

Sensitivitet = 27/37 = 73%

Specificitet = 35/112 = 31%

A high sensitivity is good when we …….

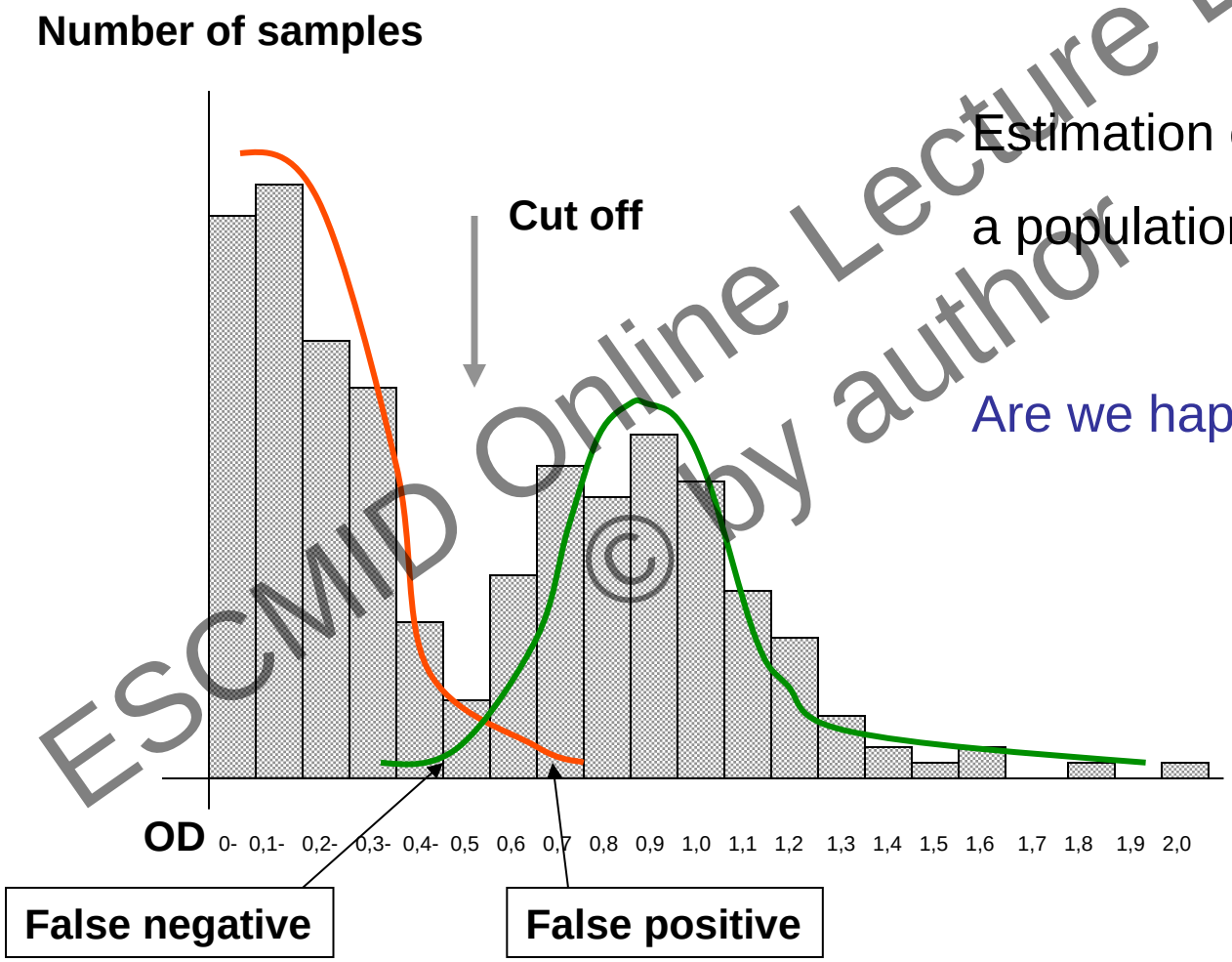
- Need to be sure that the patient do not have the diagnosis
- Ex.:
  - Sensitivity = 99% and specificity = 70%
  - If the test is negative the patient is probably not sick
  - If the test is positive it may be false positive

A high specificity is needed when …….

- The physician want to be sure the patient has a certain diagnosis
- Ex.:
  - Sensitivity = 70% and specificity = 99%
  - If the test is positive the diagnose is correct with a high probability
  - False negative results can not be excluded

Validation of diagnostic tests I

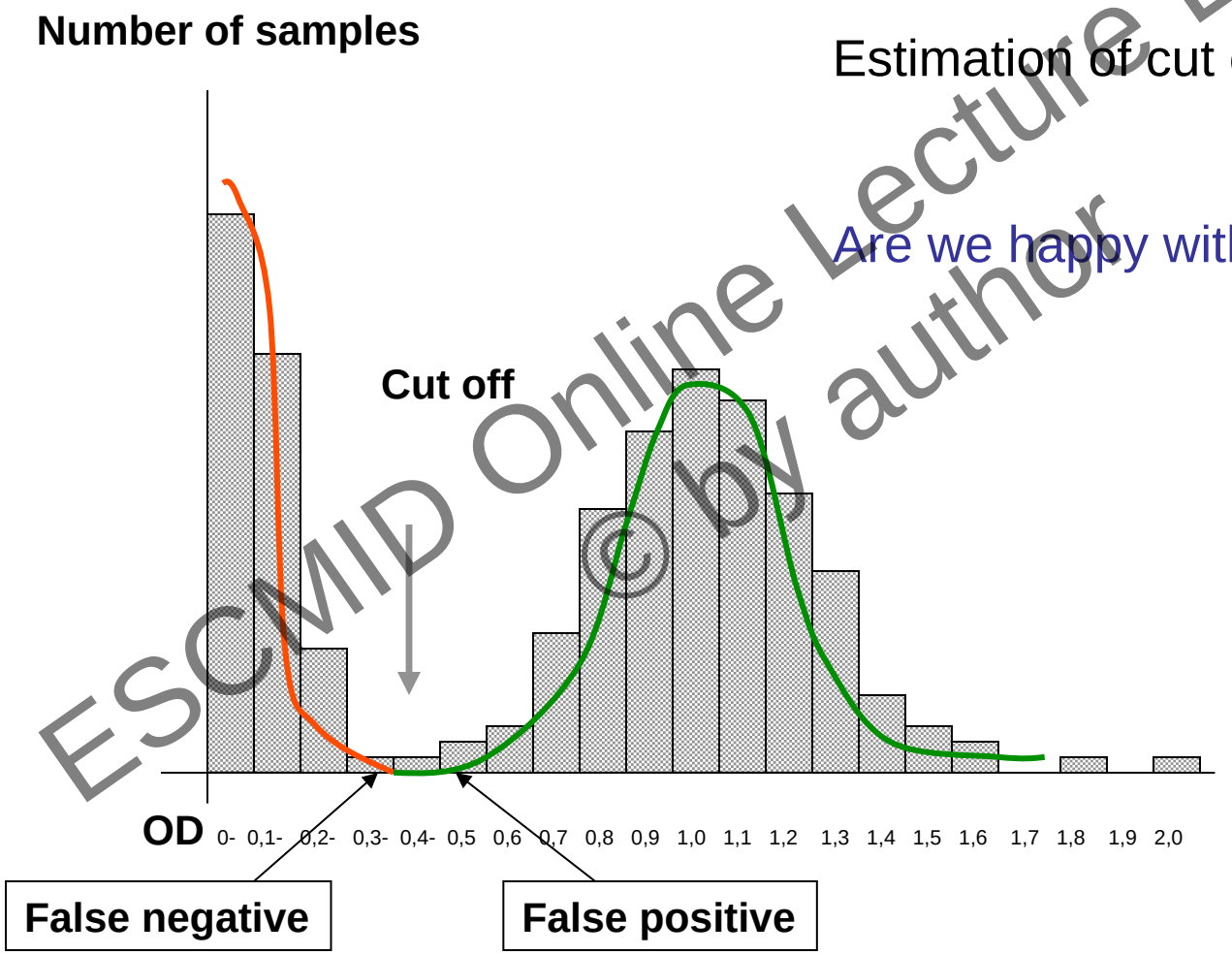
Distribution of population based samples



Estimation of cut offs using a population based sample

Are we happy with this analysis ?

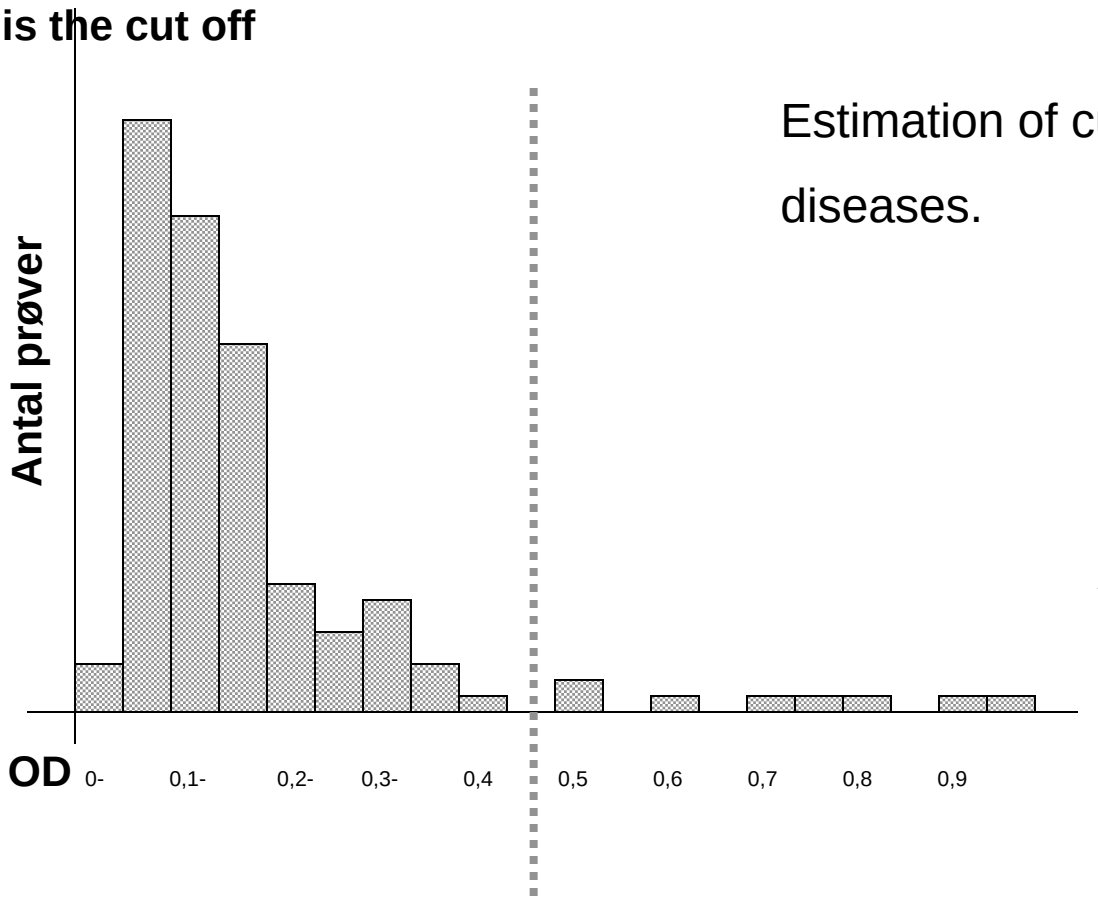
Validation of diagnostic tests II



Percentiles and Standard Deviation

99% of the samples are to the left of the line = 99% percentile is the cut off

1% of the samples are suspected positive



Estimation of cut offs with rare diseases.

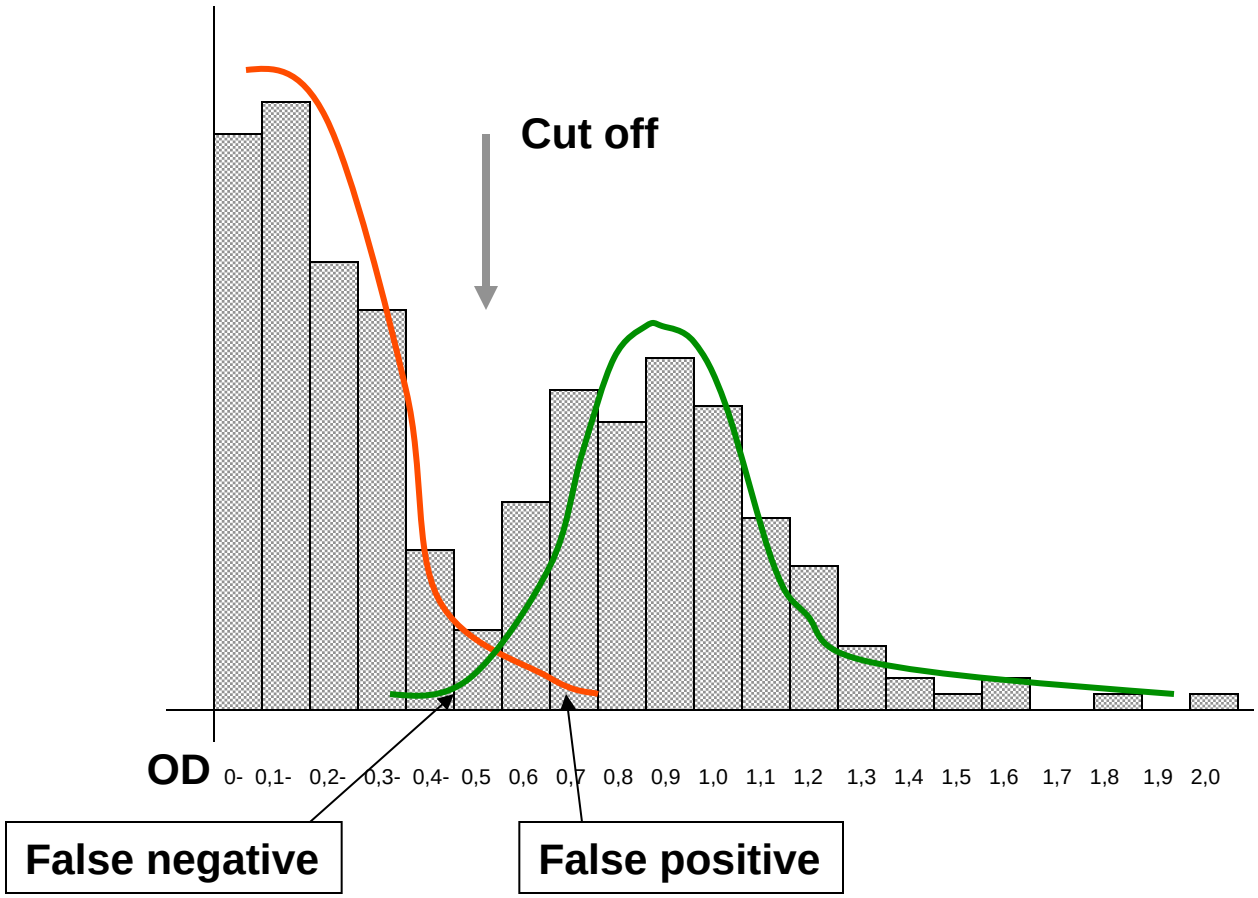
Alternatively the results are expressed as the mean plus

2 SD equivalent to 95%

3 SD equivalent to 97,5%

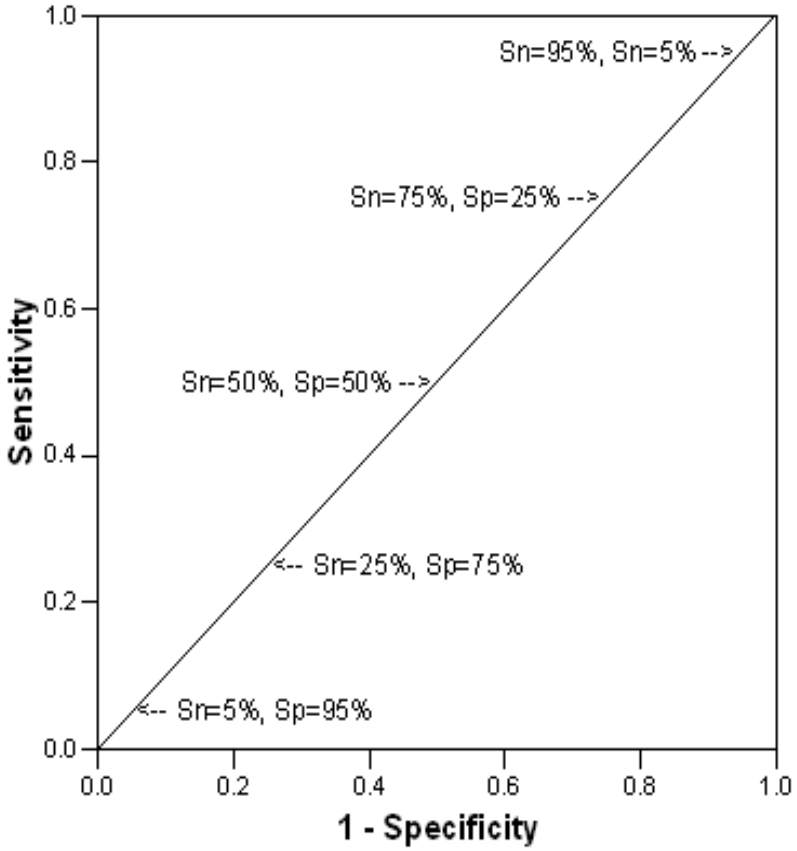
4 SD equivalent to 99%

If the test is validated on samples from patients with and without the diagnosis, samples from patients where the diagnosis is uncertain are removed and the performance will be much improved.



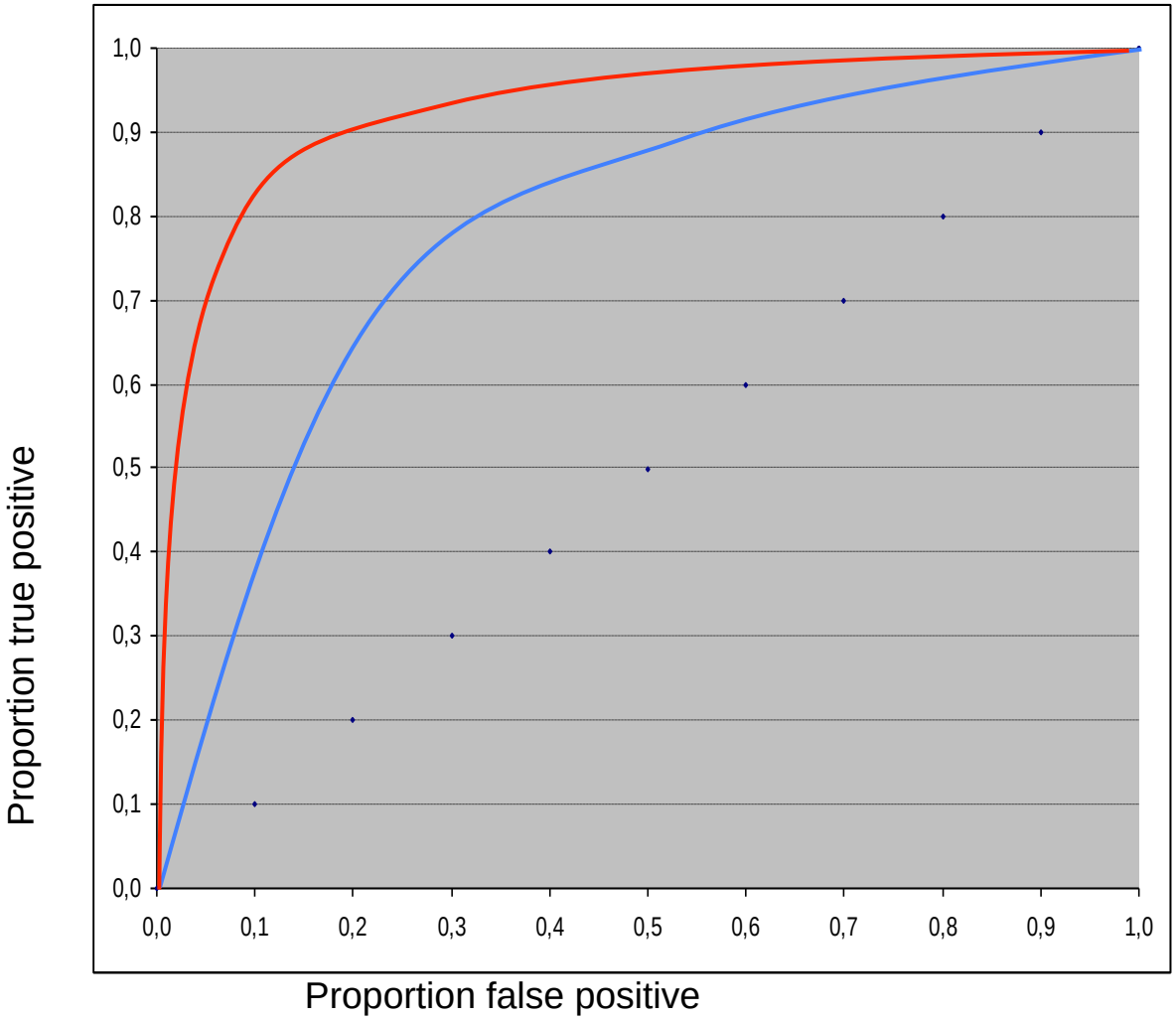
Trade off between sens. and spec.

ROC = Receiver Operating Characteristic Curve



(proportion false positive)

Which test is the preferred one ?



Predivtive values

- Positive predictive value (PPV or +PV)
  - If the test is positive, what is the likelyhood that the patient has the disease ?
- Negative predictive value (NPV or –PV)
  - If the test is negative what is the likelyhood that the patient does not have the disease?

$$+PV = \frac{\textbf{True positive}}{\textbf{True positive + False positive}} = a/a+b$$

$$-PV = \frac{\textbf{True negative}}{\textbf{True negative + False negative}} = d/d+c$$

Prevalence

- How many have the diagnosis in the population
- The relation between prevalence and predictive value

- THIS IS OF PARAMOUNT IMPORTANCE TO UNDERTSTAND THE PERFORMANCE OF A TEST IN REAL LIFE

Ex: PrEdiCtive VALUES

Disease

Test	Yes	No	Total
Positive	9	143	152
Negative	1	47	48
Total	10	190	200

Sens.= 90%, spec.=75%

Prevalence = 5%

+PV = 5,9%

–PV = 97,9%

Disease

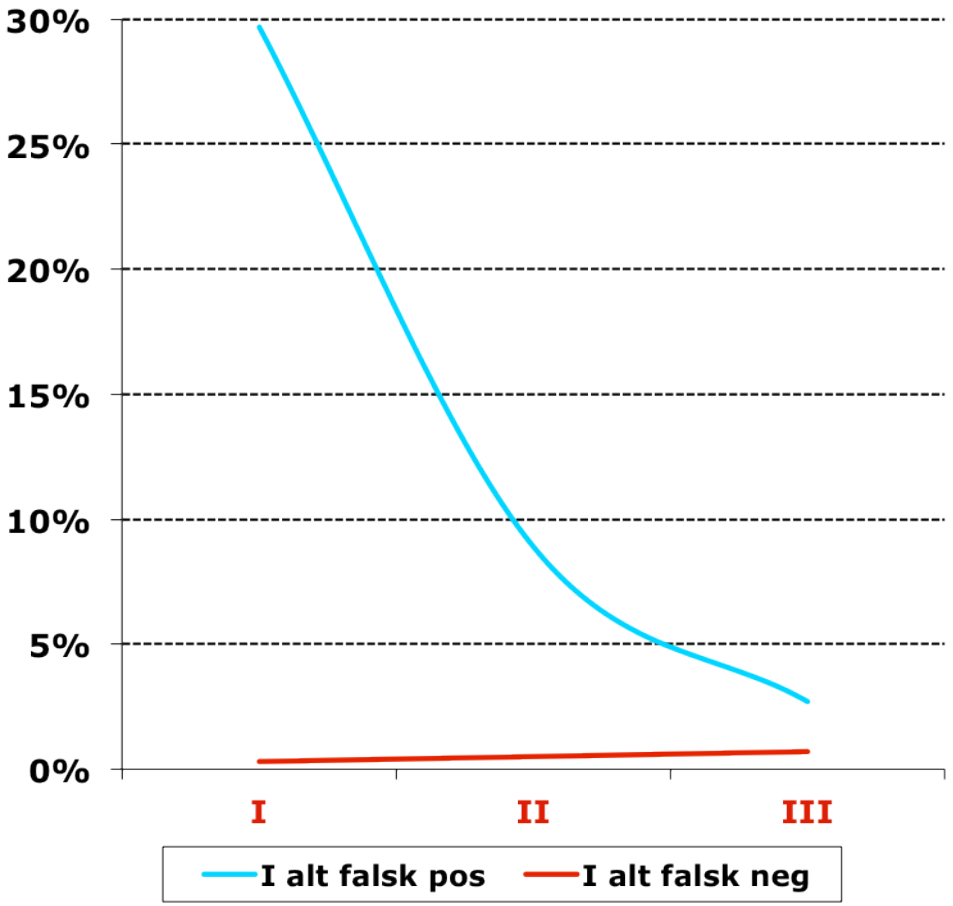
Test	Yes	No	Total
Positive	90	75	165
Negative	10	25	35
Total	100	100	200

Prevalence = 50%

+PV = 54,5%

–PV = 71,4%

Series of tests



Exponential fall in false positive

Additive increase in i false negative

CLINICAL AND VACCINE IMMUNOLOGY, Dec. 2011, p. 2095–2101 Vol. 18, No. 12
1556-6811/11/$12.00 doi:10.1128/CVI.05285-11

Evaluation of Six Commercial Point-of-Care Tests for Diagnosis of Acute Dengue Infections: the Need for Combining NS1 Antigen and IgM/IgG Antibody Detection To Achieve Acceptable Levels of Accuracy[▽]†

Stuart D. Blacksell,[1,2]* Richard G. Jarman,[3] Mark S. Bailey,[4] Ampai Tanganuchitcharnchai,[1] Kemajittra Jenjaroen,[1] Robert V. Gibbons,[3] Daniel H. Paris,[1,2] Ranjan Premaratna,[5] H. Janaka de Silva,[5] David G. Lalloo,[6] and Nicholas P. J. Day[1,2]

Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok 10400, Thailand[1]; Centre for Tropical Medicine, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, England, United Kingdom[2]; Armed Forces Research Institute of Medical Sciences, 315/6 Rajvithi Road, Bangkok 10400, Thailand[3]; Department of Military Medicine, Royal Centre for Defence Medicine, Vincent Drive, Birmingham B15 2SQ, England, United Kingdom[4]; Department of Medicine, Faculty of Medicine, University of Kelaniya, P.O. Box 6, Thalagolla Road, Ragama, Sri Lanka[5]; and Clinical Research Group, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, England, United Kingdom[6]

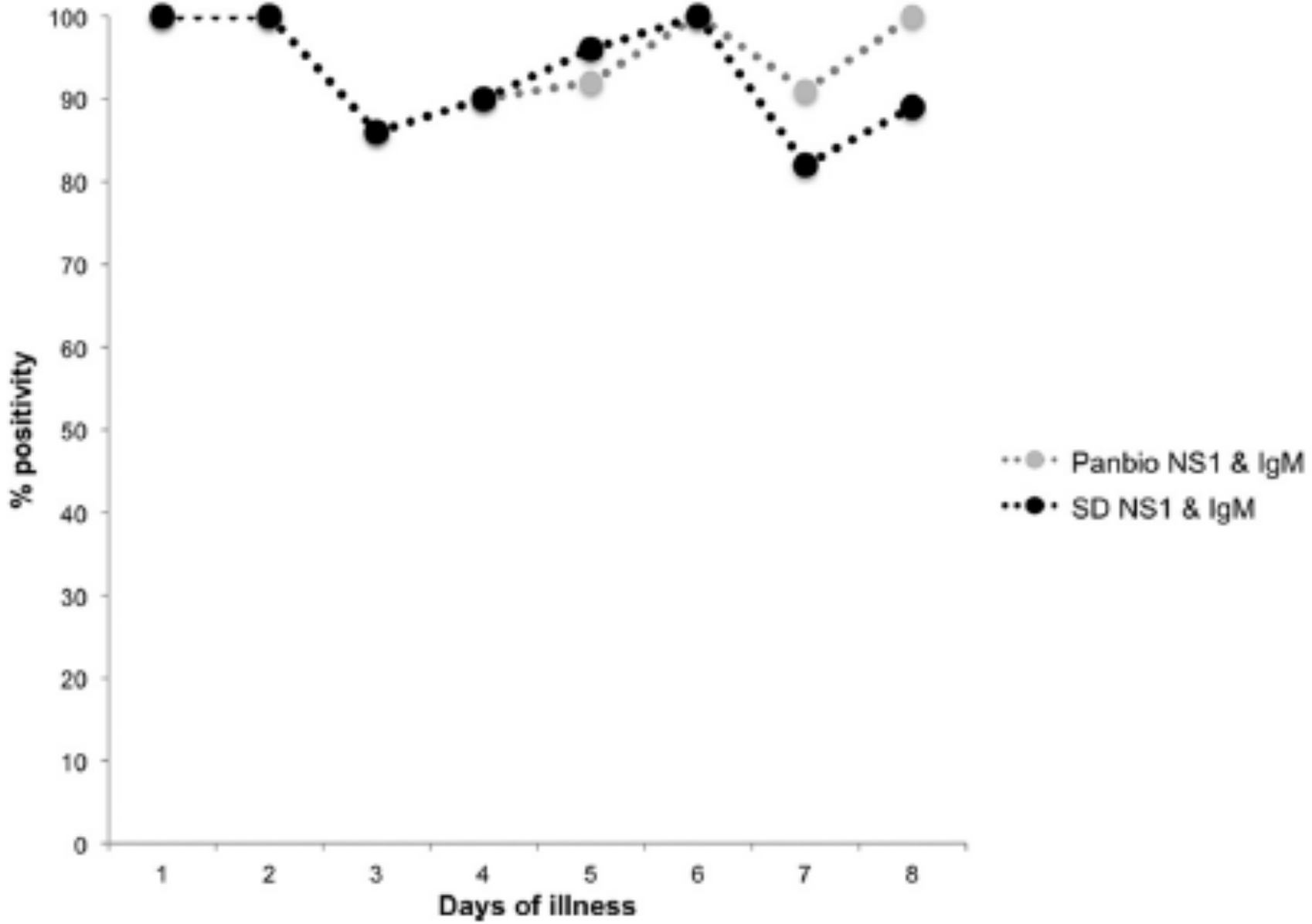
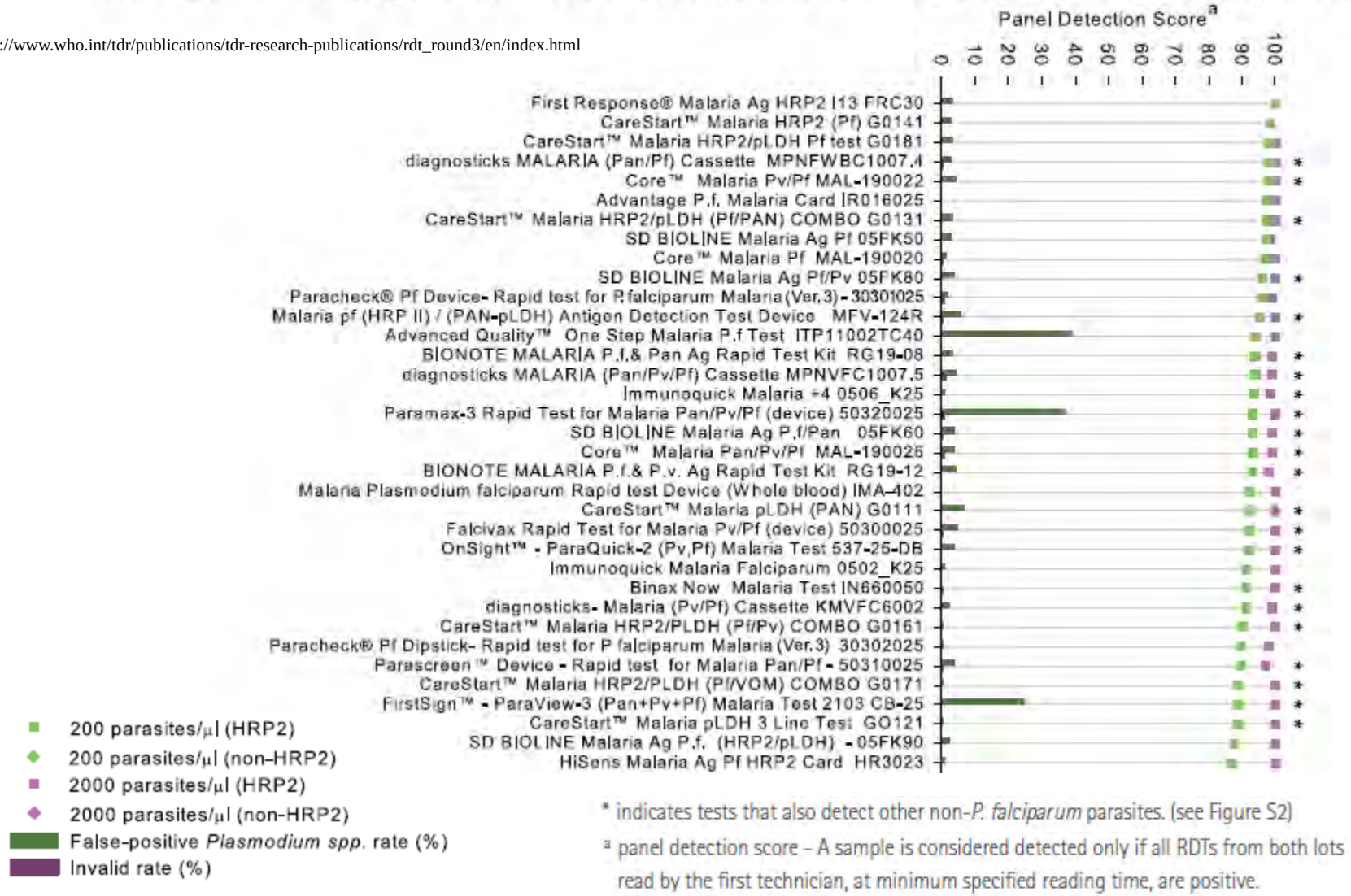

100
90
80
70
60
50
40
30
20
10
0
% positivity
1
2
3
4
5
6
7
8
Days of illness
Panbio NS1 & IgM
SD NS1 & IgM

Figure S1: Malaria RDT performance in Phase 2 of Rounds 1-3 against wild-type (clinical) samples containing *P. falciparum* at low (200) and high (2000 or 5000) parasite densities (parasites/µl) and clean-negative samples

http://www.who.int/tdr/publications/tdr-research-publications/rdt_round3/en/index.html



* indicates tests that also detect other non-*P. falciparum* parasites. (see Figure S2)

[a] panel detection score – A sample is considered detected only if all RDTs from both lots read by the first technician, at minimum specified reading time, are positive.

Use a rapid test which include pan Plasmodial antigen ie. LDH or Aldolase

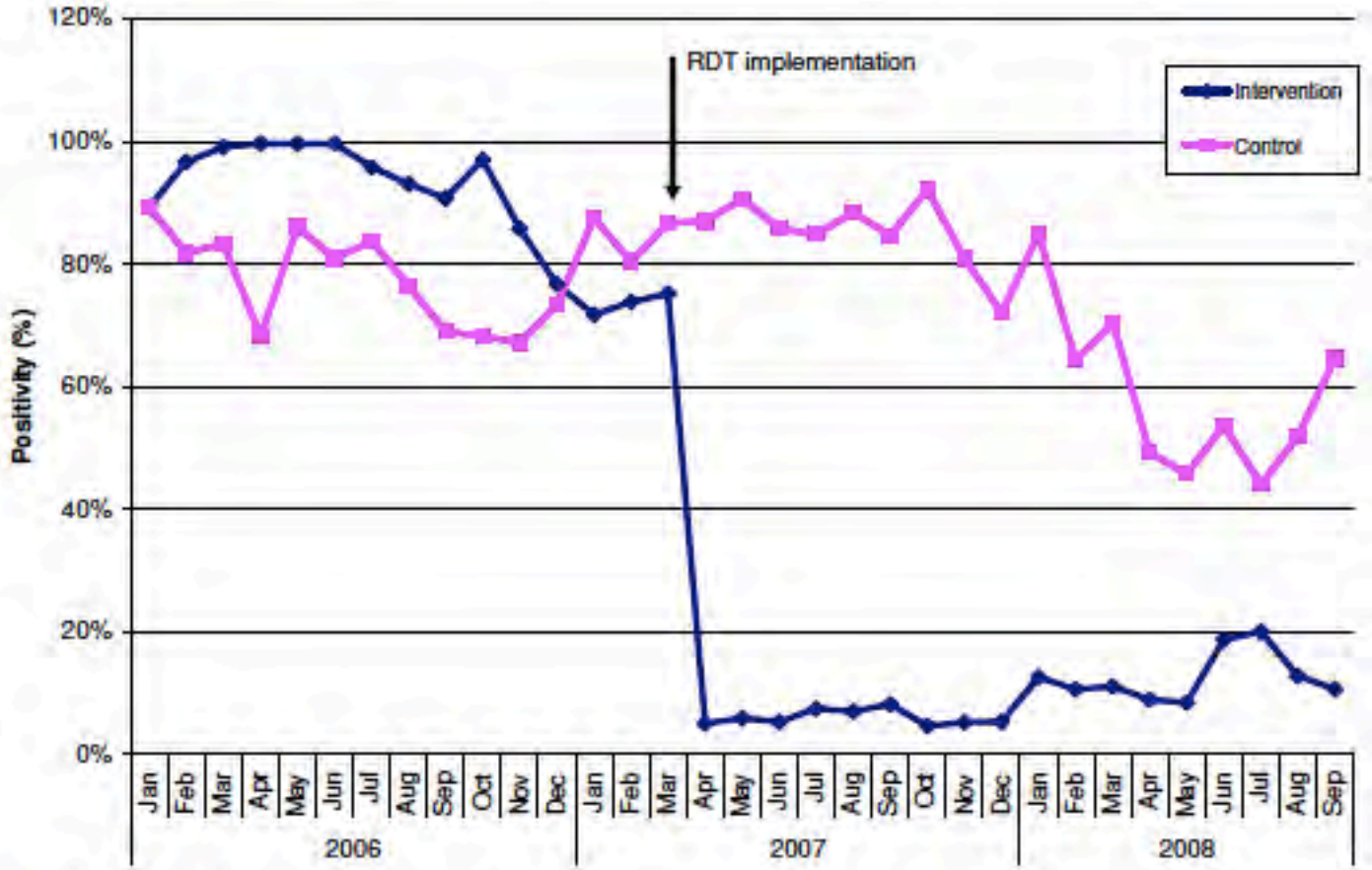


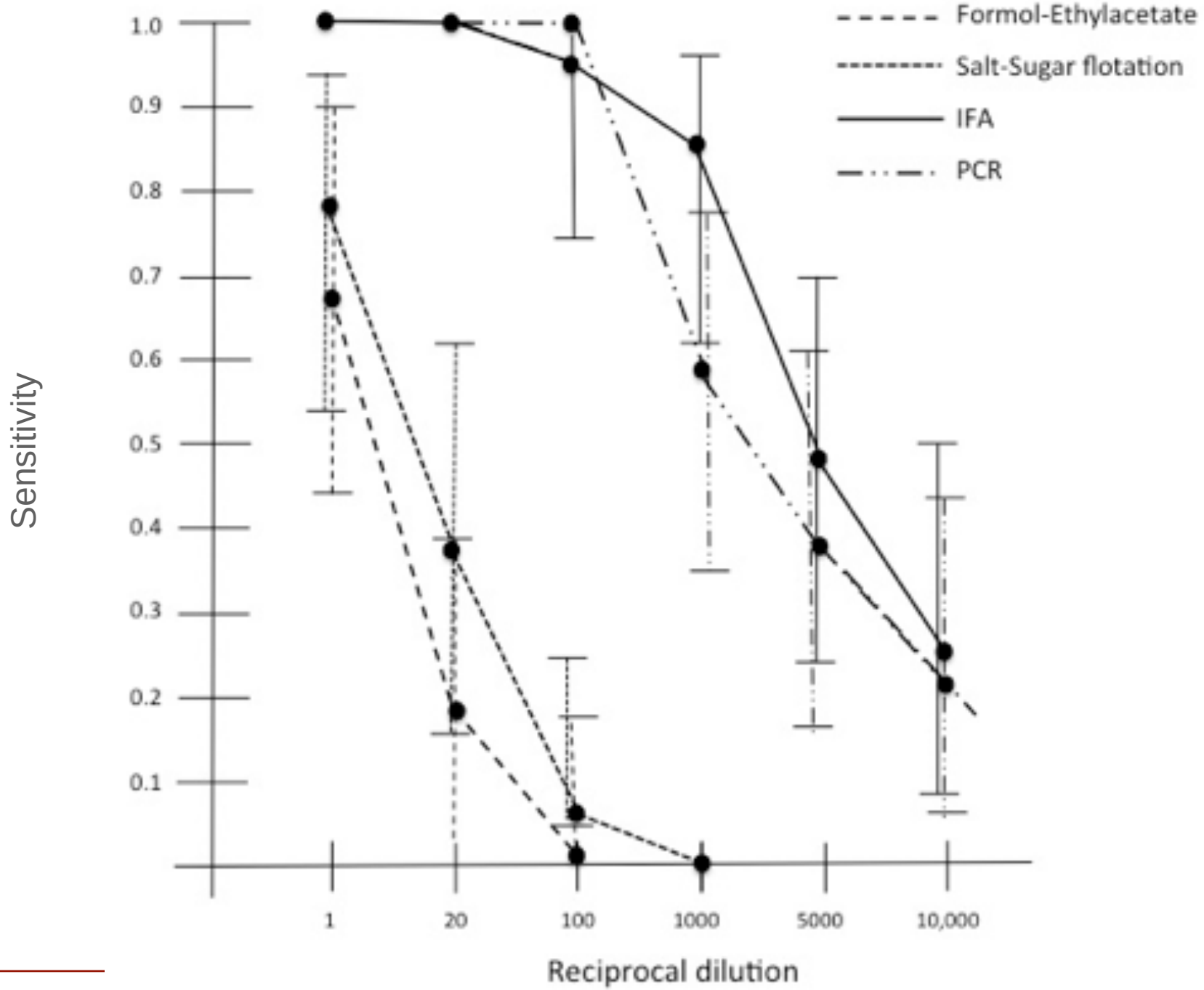
Figure 3 Malaria test positivity rates in one intervention and one matched control health centre before and after RDT introduction (see text for details).

Kahama-Maro et al. Malaria J 2011;10:332

Comparison of sensitivity and specificity of four methods for detection of *Giardia* spp. in feces: immunofluorescence and PCR are superior to microscopy of concentrated iodine stained samples

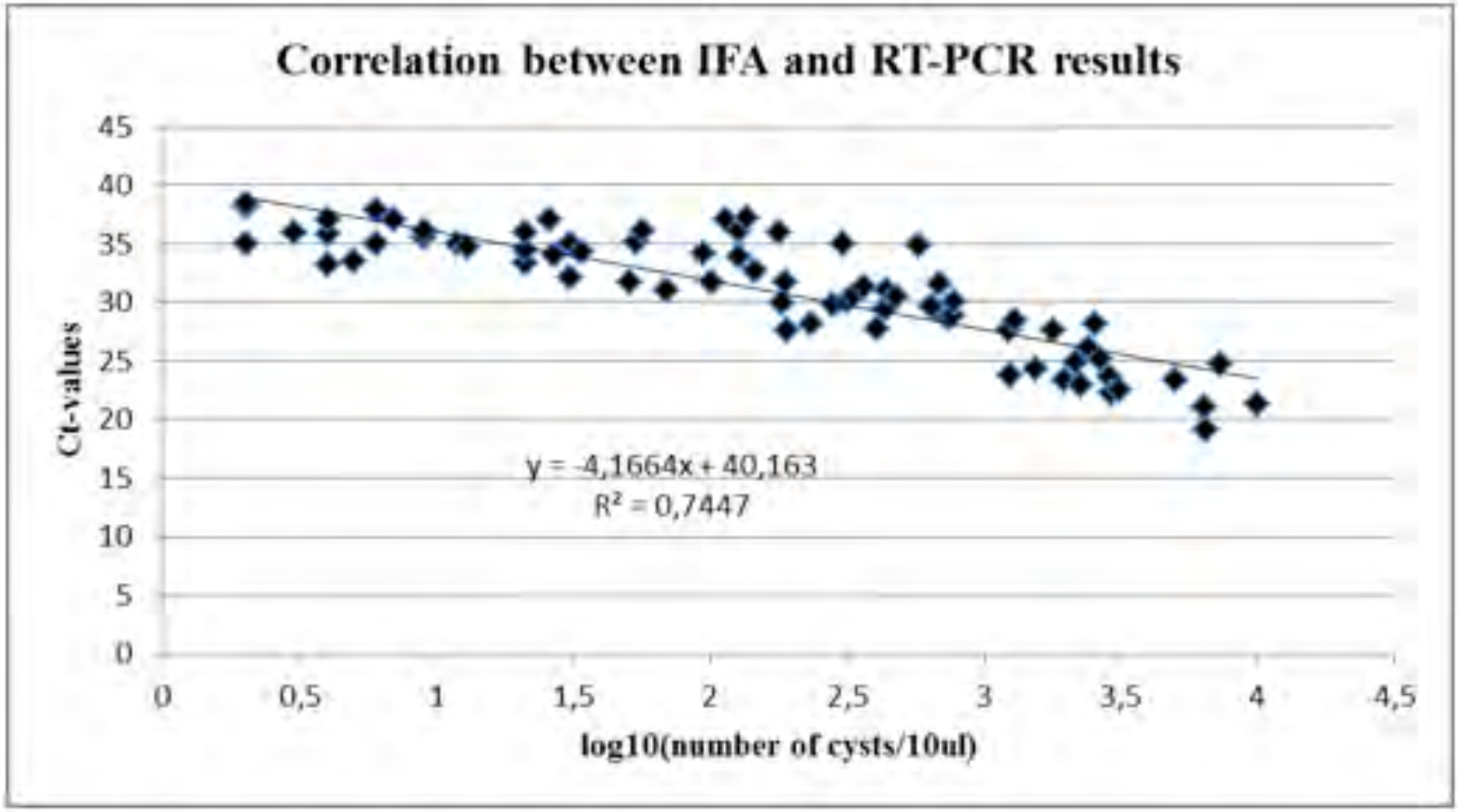
Running title
Compare diagnostic methods to find *Giardia duodenalis*

Helle Gotfred-Rasmussen (1), Marianne Lund (2), Heidi L. Enemark (3), Mogens Erlandsen (4), Eskild Petersen#(1)

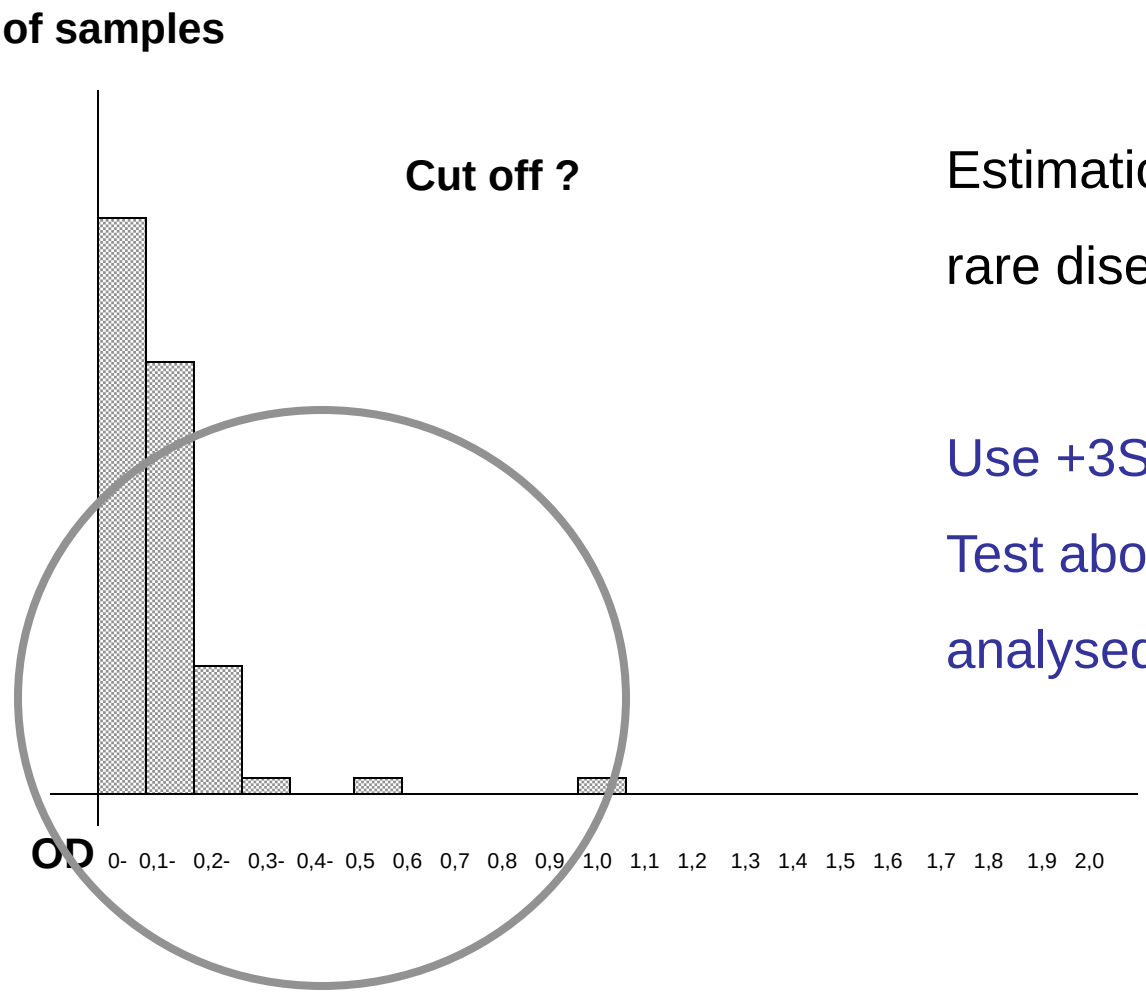

Formol-Ethylacetate
Salt-Sugar flotation
IFA
PCR
Sensitivity
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1
20
100
1000
5000
10,000
Reciprocal dilution


Questions ?

Figure 1


Correlation between IFA and RT-PCR results
Ct-values
45
40
35
30
25
20
15
10
5
0
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
y = -4,1664x + 40,163
R² = 0,7447
log10(number of cysts/10ul)

Estimation of cut off without positive reference samples

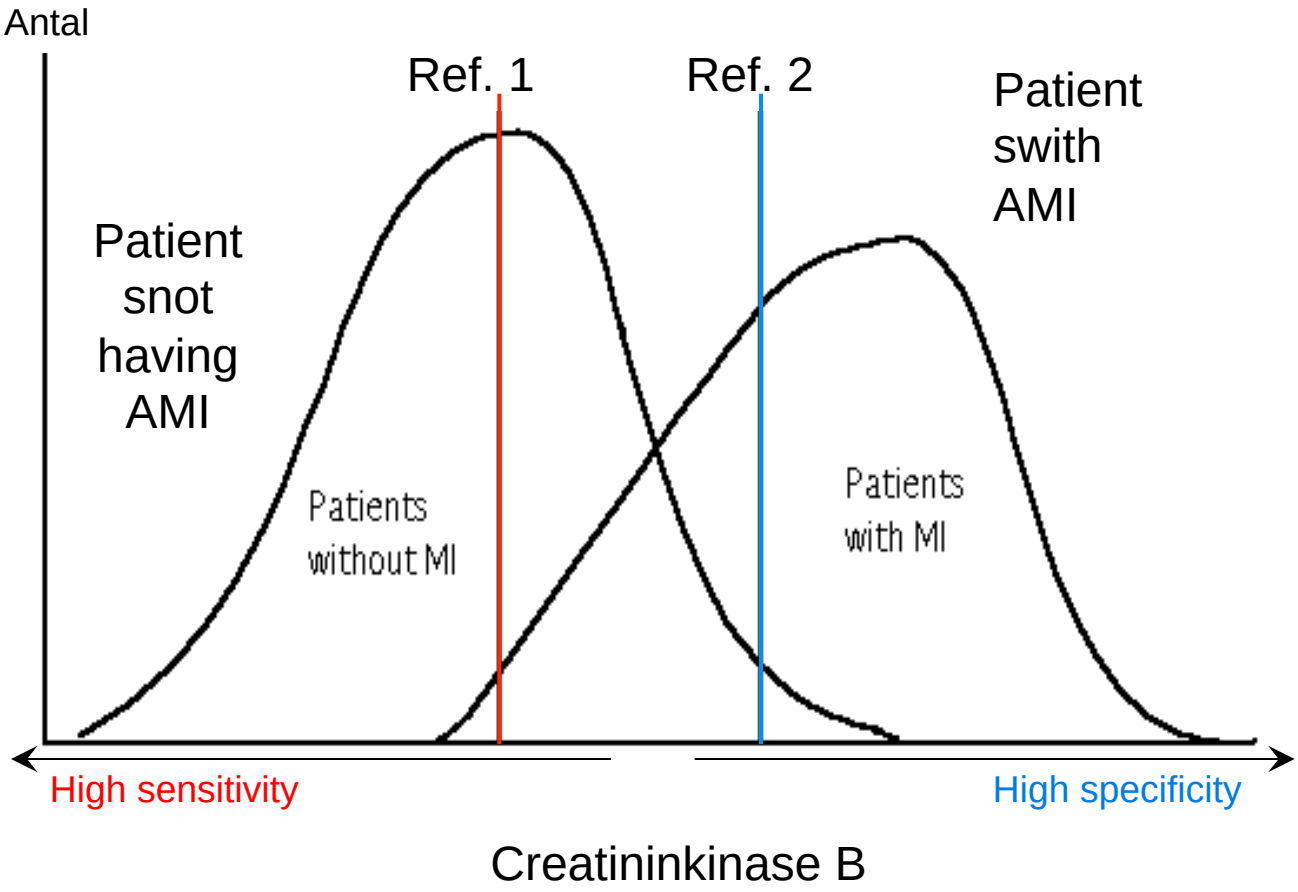


Estimation of cut offs with rare diseases.

Use +3SD or 99% percentile.

Test above should ideally be analysed with a second test

Trade off between sens. og spec.



AMI = Acute Myocardial Infarction

Commercial Dengue Rapid Diagnostic Tests for Point-of-Care Application: Recent Evaluations and Future Needs?

Stuart D. Blacksell[1,2]

[1] *Center for Tropical Medicine, Nuffield Department of Clinical Medicine, Churchill Hospital, University of Oxford, Oxford, UK*
[2] *Mahidol Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok, 10400, Thailand*

Correspondence should be addressed to Stuart D. Blacksell, stuart@tropmedres.ac

Received 12 December 2011; Accepted 11 February 2012

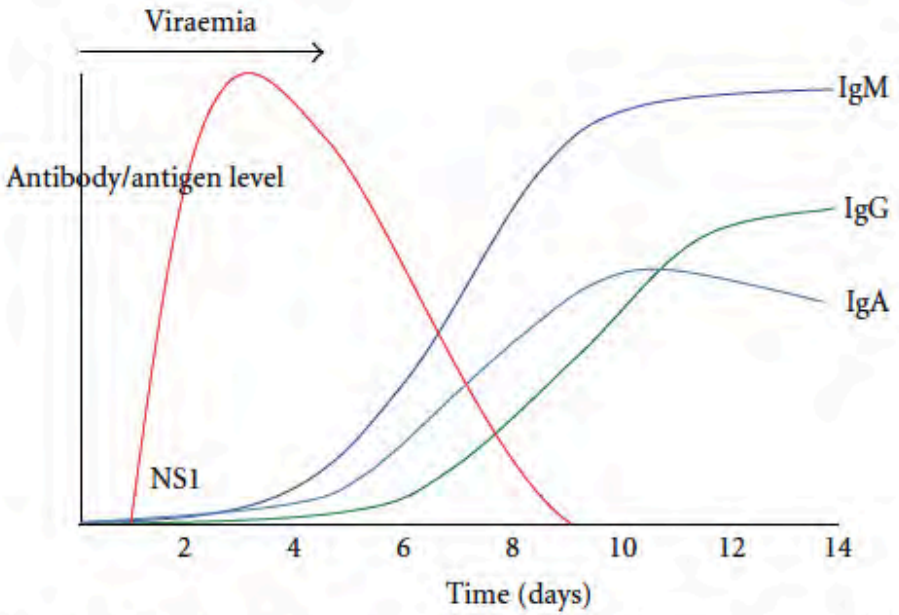


FIGURE 1: Graphical representation of the kinetics of dengue NS1 antigen and IgM, IgG, and IgA antibodies during a primary dengue infection.

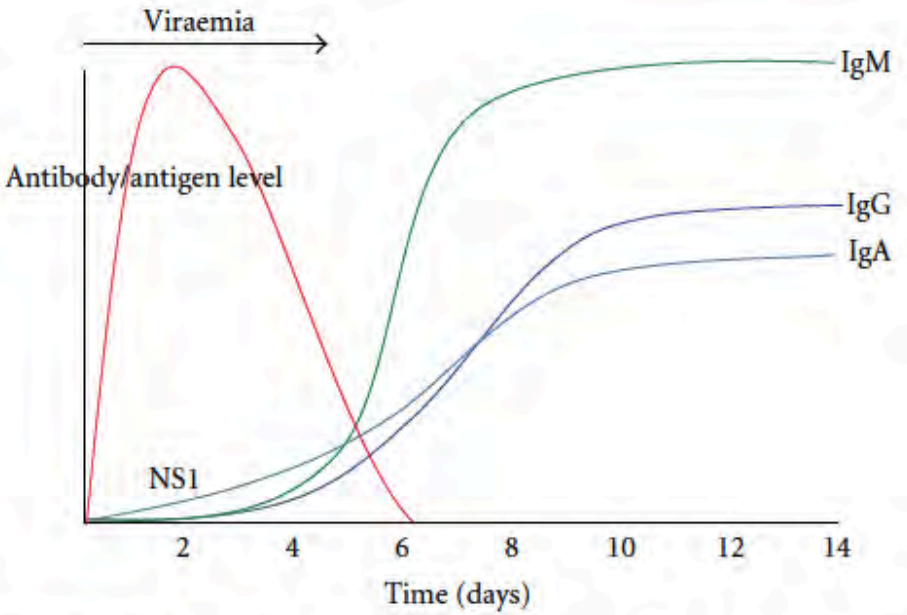


FIGURE 2: Graphical representation of the kinetics of dengue NS1 antigen and IgM, IgG, and IgA antibodies during a secondary dengue infection.

TABLE 4. Overall diagnostic accuracy and sensitivity[a]

Type of antibodies or antigens	Test	Sensitivity (%)	Specificity (%)	PPV[b]	NPV[c]	Kappa value
IgM antibodies	Merlin	72.7 (62.9–81.2)	73.8 (66.2–80.4)	63.2 (53.6–72.0)	81.4 (74.1–87.4)	0.79
	Biosynex	79.8 (70.5–87.2)	46.3 (38.3–54.3)	49.9 (40.1–55.8)	78.7 (69.1–86.5)	0.57
	Standard Diagnostics	79.2 (70.5–87.2)	89.4 (83.5–93.7)	82.3 (73.2–89.3)	87.7 (81.7–92.3)	0.92
	Panbio	70.7 (60.7–79.4)	80.0 (73.0–85.9)	68.6 (58.7–77.5)	81.5 (74.6–87.3)	0.92
NS1 antigen	Standard Diagnostics	48.5 (38.5–58.7)	99.4 (96.6–100)	98.0 (89.1–100)	75.7 (69.3–81.4)	0.96
	Bio-Rad	58.6 (48.2–68.4)	98.8 (95.6–99.9)	96.7 (88.5–99.6)	79.4 (73.1–84.8)	0.94
	Panbio	58.6 (48.2–68.4)	92.5 (87.3–96.1)	82.9 (72.0–90.8)	78.3 (71.7–84.0)	0.95
IgM antibodies and NS1 antigen	Standard Diagnostics	92.9 (83.9–97.1)	88.8 (82.8–93.2)	83.6 (75.4–90.0)	95.4 (90.6–98.1)	Not applicable
	Panbio	89.9 (82.2–95.0)	75.0 (67.6–81.5)	69.0 (60.3–76.8)	92.3 (86.3–96.2)	Not applicable

[a] The 95% confidence intervals are listed in parentheses.
[b] PPV, positive predictive value.
[c] NPV, negative predictive value.

JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2010, p. 2793–2797 Vol. 48, No. 8
0095-1137/10/$12.00 doi:10.1128/JCM.02142-09

Evaluation of a Commercial SD Dengue Virus NS1 Antigen Capture Enzyme-Linked Immunosorbent Assay Kit for Early Diagnosis of Dengue Virus Infection[▽]

Seok Mui Wang and Shamala Devi Sekaran*
Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Malaysia[1]

Received 2 November 2009/Returned for modification 18 April 2010/Accepted 13 June 2010

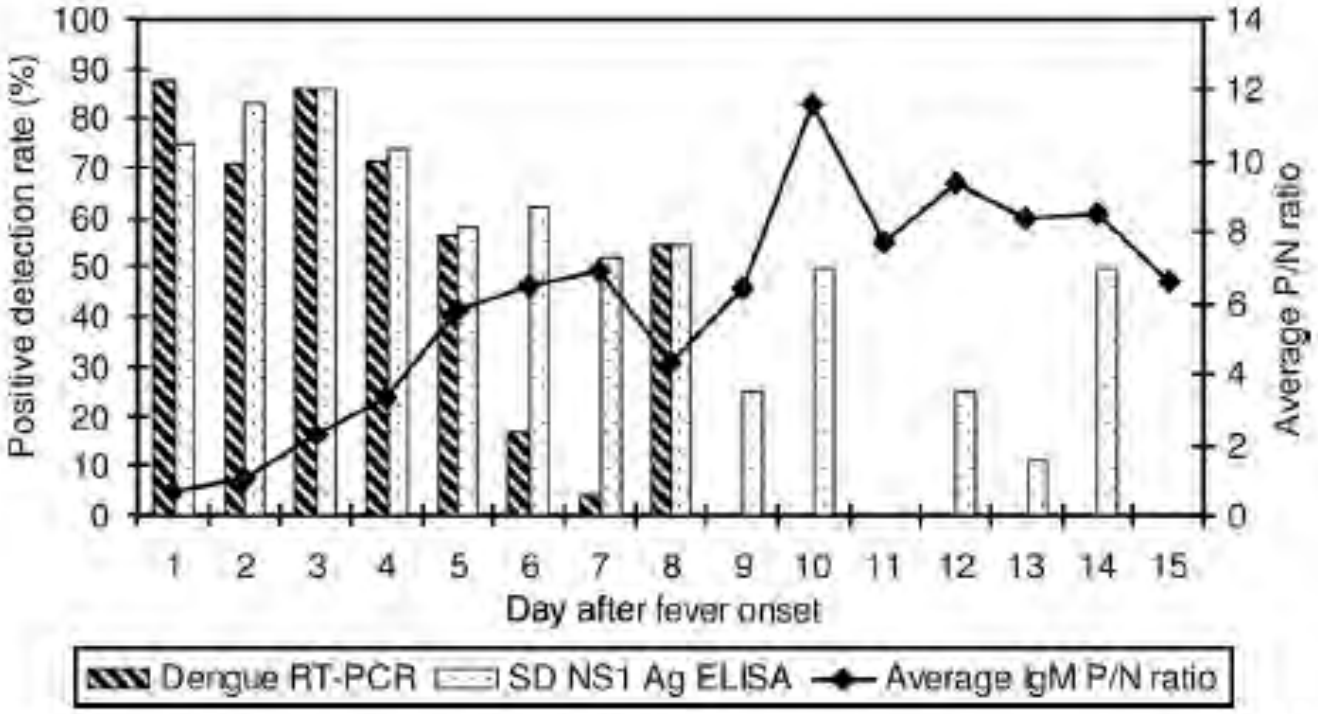
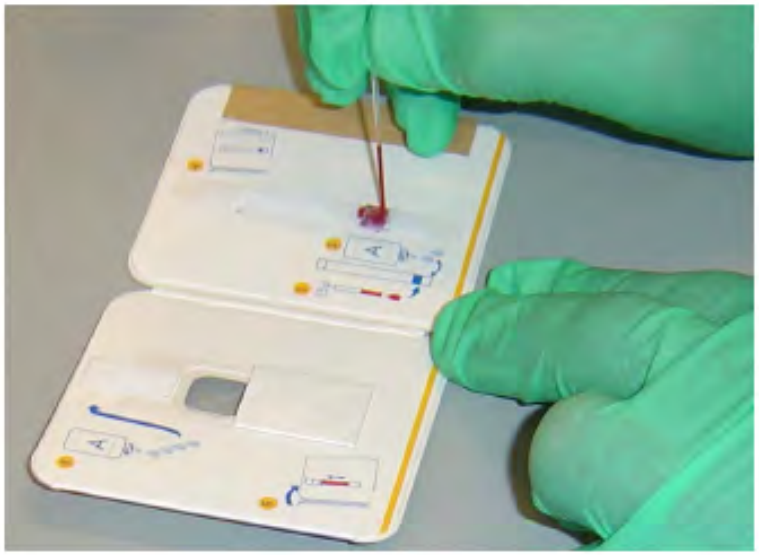
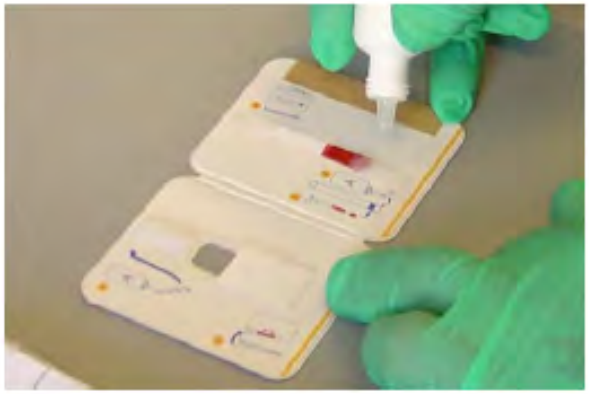


FIG. 2. Comparison of RT-PCR, NS1 antigen detection, and IgM for detection after onset of fever.






Malaria P.f./P.v.
NOW
ICT
C
P.f.
P.v.
(+) (+) (+) (-)
P.f.
P.v.
binax

Limitations of rapid tests

Rapid tests may be false negative in cases with very high *P.falciparum* density (Gillet et al. Malaria J 2009; 8:271)

Prozone effect

(HRP2 deletion in parts of South America)

Variant *P. ovale*

P.knowlesi

(Tordrup et al. Malaria J 2011;10:15)

HRP2 based tests will be negative
Pan-malaria antigens containing tests (LDH, aldolase) will be positive

Malaria Rapid Diagnostic Test Performance – Results of WHO product testing of malaria RDTs: Round 3 (2010-2011)

The evaluation programme is co-sponsored by the Foundation for Innovative New Diagnostics (FIND), the Special Programme for Research and Training in Tropical Diseases (TDR) and the WHO Global Malaria Programme (GMP). Testing is performed at the US Centers for Disease Control and Prevention (CDC)

http://www.who.int/tdr/publications/tdr-research-publications/rdt_round3/en/index.html

Which method should we use ?

Detection levels of different diagnostic methods:

Rapid tests **200 parasites per mm^3**

WHO 2011

Microscopy **5 parasites per mm^3**

(experienced microscopist) Petersen E et al Am J Trop Med Hyg 1996;55:485-489

PCR **0.5 parasites per mm^3**

Gama BE et al. Exp Parasitol. 2007 Mar 2; [Epub]

Table 1. Formol Ethylacetate (FEA) of 19 *Giardia duodenalis* positive clinical samples. The number of positive samples at each dilution and threshold sensitivity expressed as number of cysts per gram feces in the original, undiluted sample.

Dilutions	Positive % (pos/all)	Median (cysts/20µl ± 25%-75% quantile)	Median Cysts in 1 ml of material (± 25%-75% quantile)
1:1	100% (19/19)	1 ± 1 - 3	50 ± 50 - 150
1:20	16% (3/19)	1 ± 1 - 2	2.5 ± 2.5 - 5
1:100	0%	0	0
1:1,000	0%	0	0
1:5,000	0%	0	0
1:10,000	0%	0	0

Table 3. Immunofluorescence assay (IFA). Number of positive samples (n=19) at each dilution and threshold sensitivity expressed as number of cysts per gram feces in the original, undiluted sample.

Dilutions	Positive % (pos/all)	Median number of cysts in 10µl ± 25% - 75% quantile	Median number of cysts in 1ml of material (± 25%-75% quantile)
1:1	100% (19/19)	767 (568 - 1,221)	76,700 (56,800 – 121,100)
1:20	100% (19/19)	433 (211 - 714)	43,300 (21,100 - 71,400)
1:100	95% (18/19)	94 (37 - 169)	9,400 (3,700 - 16,900)
1:1,000	84% (16/19)	7 (4 - 20)	700 (400 - 600)
1:5,000	47% (9/19)	3 (2 - 5)	300 (200 – 500)
1:10,000	26% (5/19)	1 (NA)	100 (NA)

NA = Not applicable