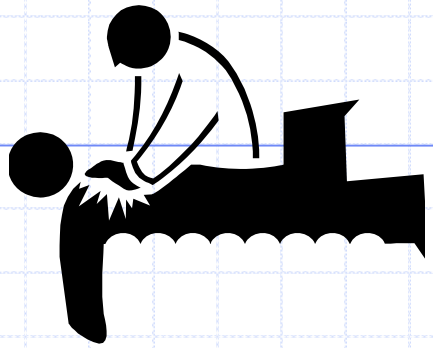
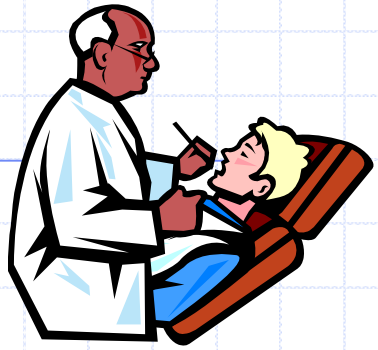
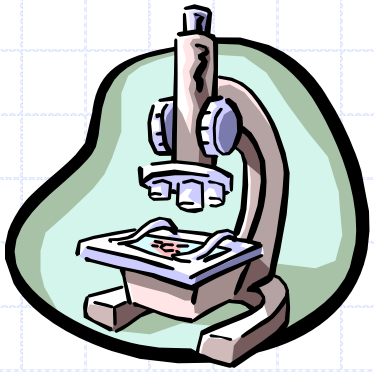
Diagnostic research: incremental value and multivariable approach

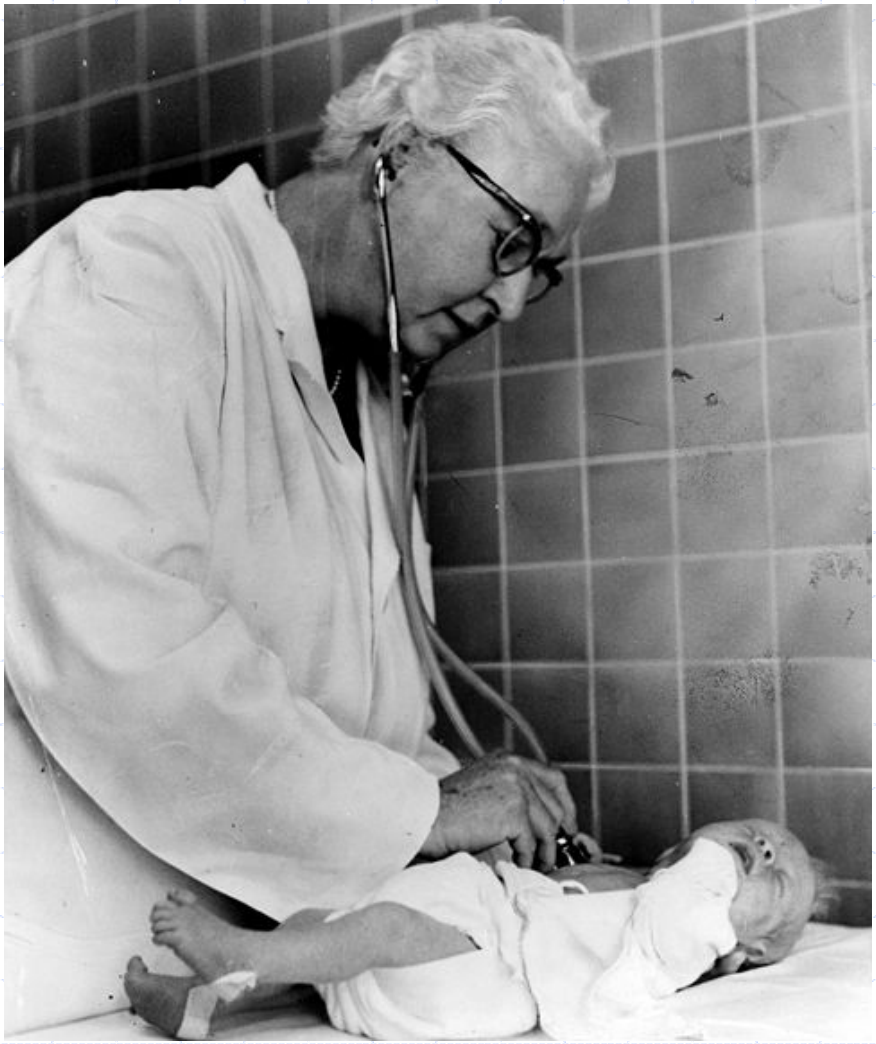




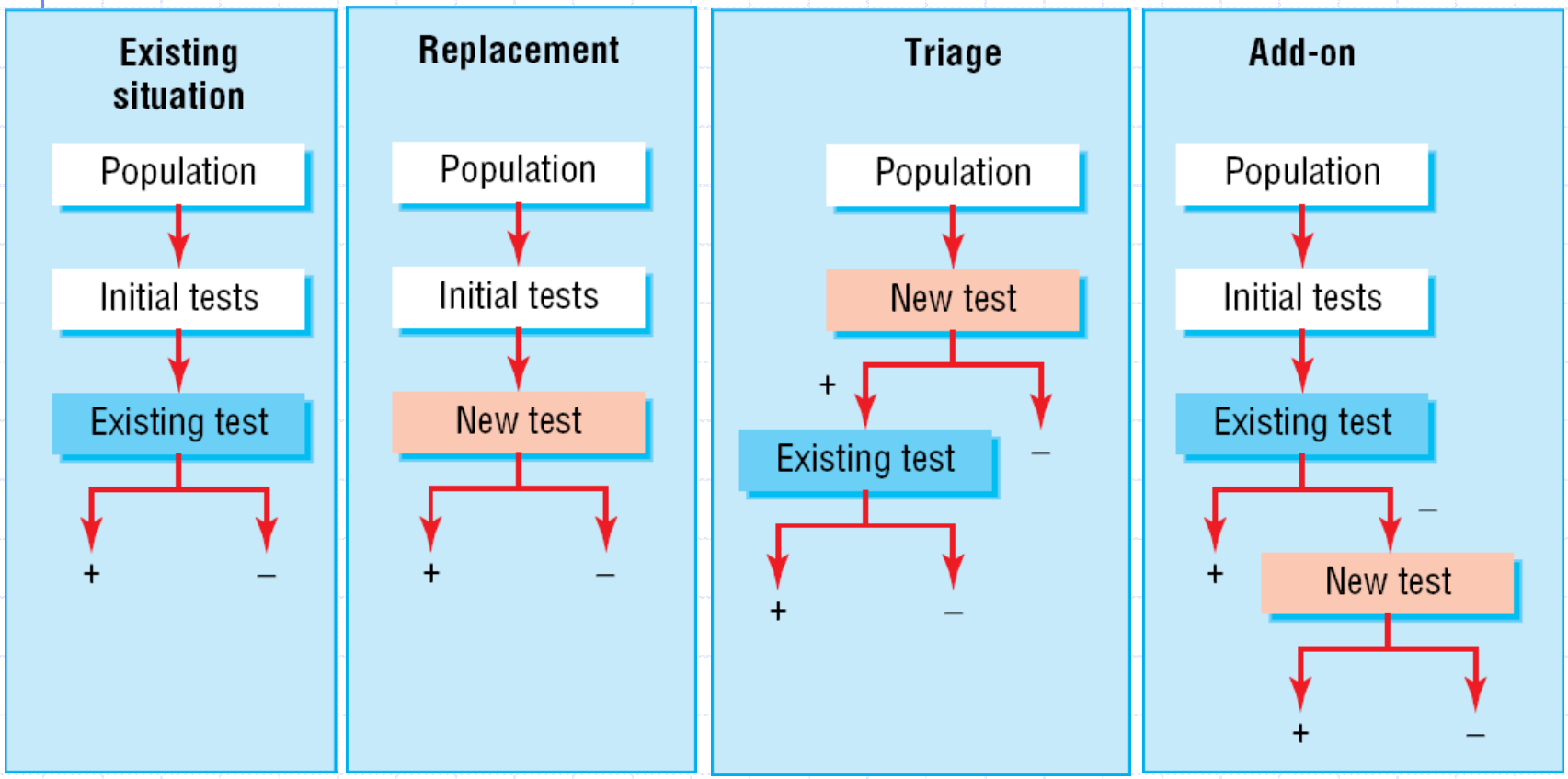


Madhukar Pai, MD, PhD
Associate Professor of Epidemiology, McGill University
Montreal, Canada

Email: madhukar.pai@mcgill.ca



Design is often decided by: what is the real or intended purpose of the test?



Bossuyt, BMJ, 2006

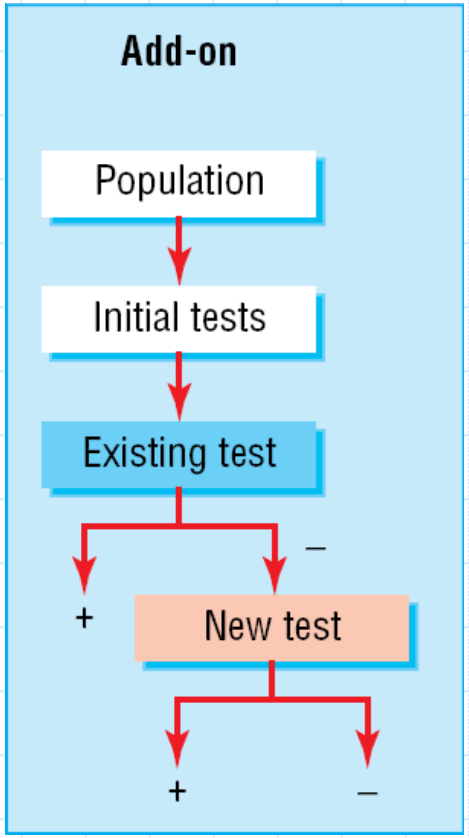
The diagnostic process is probabilistic, multivariable and sequential

1. A diagnosis starts with a patient presenting a complaint (symptom and/or sign) suggestive of a certain disease to be diagnosed.
2. The subsequent work-up is a multivariable process. It involves multiple diagnostic determinants (tests) that are applied in a logical order: from age, gender, medical history, and signs and symptoms, to more complicated, invasive, and costly tests.
3. Setting or ruling out a diagnosis is a probabilistic action in which the probability of the presence or absence of the disease is central. This probability is continuously updated based on subsequent diagnostic test results.
4. The true diagnostic value of a test is determined by the extent to which it provides diagnostic information beyond earlier tests, that is, materially changes the probability estimation of disease presence based on previous test results.
5. The goal of the diagnostic process is to eventually rule in or out the disease with enough confidence to take clinical decisions. This requires precise estimates of the probability of the presence of the target disease(s).

Moons KGM. In: Grobbee & Hoes. Clinical Epidemiology. 2009

Incremental value (added value)

- What does the new test add to the diagnostic process, over and above already existing information?
- Can be answered using a multivariable approach
  - Will also need patient impact studies
- Sensitivity/specificity: each test is treated in isolation (which is not reflective of normal practice)



TB examples

- What is the added value of IGRAs, after initial screening with TST is done, for evaluating patients on TNF-a blockers?
- What is the added value of Xpert MTB/RIF in sputum smear-negative patients with HIV infection?
- What is the added value of the 3rd smear, once two sputum smears are done?

Multivariable process

- *Relate disease probability to test results*
- *Outcome = occurrence of disease (yes/no)*
- *Determinants = diagnostic tests --> dichotomous, continuous, ordinal, nominal*
- *Diagnostic function: P (D+) = f (X_1, X_2.... X_n)*
  - Where X1, X2, etc are various tests

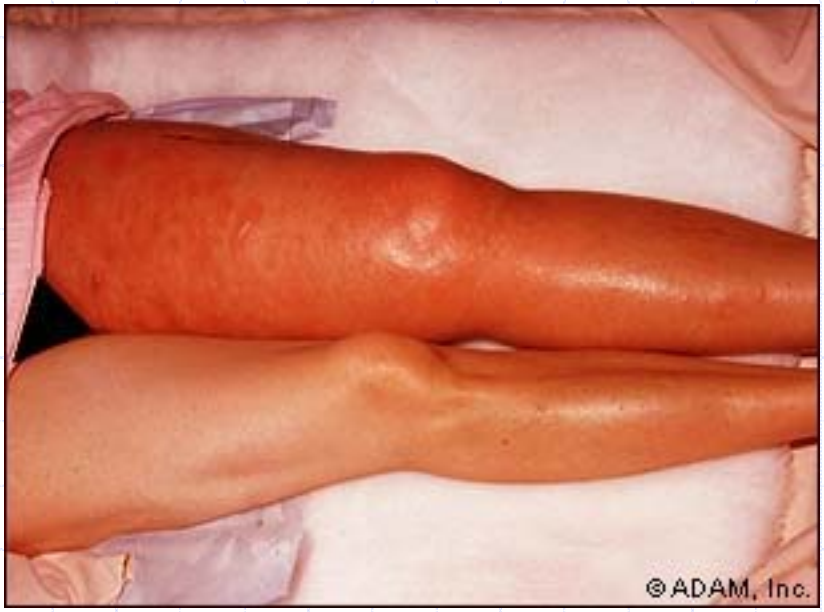
Multivariable process

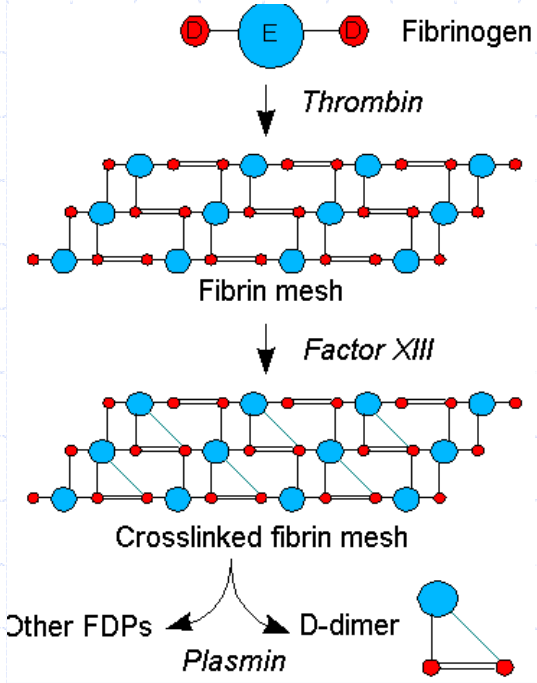
- Logistic regression model:

$$\ln \frac{P(D+|X)}{1-P(D+|X)} = b0 + b1.X1 + b2.X2 + \ldots + bn.Xn$$

$$P(D+|X) = \frac{1}{1+e^{-(b0+b1.X1+\ldots+bn.Xn)}}$$

Multivariable example: does D-dimer add value to ruling out DVT?





Multivariable approach (example)

New Technologies and Diagnostic Tools

Ruling out deep venous thrombosis in primary care

A simple diagnostic algorithm including D-dimer testing

Ruud Oudega, Karel G. M. Moons, Arno W. Hoes

Julius Center for Health Sciences and Primary care, University Medical Center Utrecht, Utrecht, The Netherlands

Summary

In primary care, the physician has to decide which patients have to be referred for further diagnostic work-up. At present, only in 20% to 30% of the referred patients the diagnosis DVT is confirmed. This puts a burden on both patients and health care budgets. The question arises whether the diagnostic work-up and referral of patients suspected of DVT in primary care could be more efficient. A simple diagnostic decision rule developed in primary care is required to safely exclude the presence of DVT in patients suspected of DVT, without the need for referral. In a cross-sectional study, we investigated the data of 1295 consecutive patients consulting their primary care physician with symptoms suggestive of DVT, to develop and validate a simple diagnostic decision rule to safely exclude the presence of DVT. Independent diagnostic indicators of the presence of DVT were male gender, oral contraceptive use, presence of malignancy, recent surgery, absence of leg trauma, vein distension, calf difference and D-dimer test result. Application of this rule could reduce the number of referrals by at least 23% while only 0·7% of the patients with a DVT would not be referred. We conclude that by using eight simple diagnostic indicators from patient history, physical examination and the result of D-dimer testing, it is possible to safely rule out DVT in a large number of patients in primary care, reducing unnecessary patient burden and health care costs.

Oudega et al. Thromb Haemost 2005

Methods

- In a large cross sectional study, 1295 consecutive adult patients (over 18 years) who visited one of the primary care physicians adherent to three non-academic hospitals in The Netherlands, and in whom DVT was suspected by the physician on clinical grounds.
- In accordance with earlier studies, the suspicion of DVT was based on the presence of at least one of the following symptoms or signs of the lower extremities: swelling, redness, and/or pain in the legs

Oudega et al. Thromb Haemost 2005

History and physical

- After informed consent, the primary care physician systematically documented information on the patient's history and physical examination.
- Following history findings were recorded as potential diagnostic determinants: presence of previous DVT, family history of DVT, history of any malignancy (active cancer in the last 6 months), immobilization for more than 3 days, recent surgery (within past 4 weeks), leg trauma (within past 4 weeks), pain when walking, and the presence of duration of the three main symptoms (i.e. a painful, red or swollen leg).
- Physical examination items included the presence of tenderness along the deep vein system in calf or thigh, distension of collateral veins in the symptomatic leg, pitting edema in the symptomatic leg of the calf and thigh, and ≥ 3 cm difference in circumference of the calves.

Oudega et al. Thromb Haemost 2005

Lab tests and reference standard

- After the standardized history taking and physical examination, all patients were referred to the hospital to undergo D-dimer testing.
- After venous blood was drawn, each patient directly underwent real time B-mode compression ultrasonography (CUS) of the lower extremities [Reference standard]

Oudega et al. Thromb Haemost 2005

Data analysis

- After univariate analysis, the authors first quantified which of the 16 history and physical findings independently contributed to the presence or absence of proximal DVT using multivariable logistic regression analysis.
- Starting with the overall model including all history and physical findings, model reduction (stepwise backwards) was performed by excluding variables from the model with a p-value > 0.10 based on the log likelihood ratio test.

Oudega et al. Thromb Haemost 2005

Data analysis

- Subsequently, they added the D-dimer test to this reduced model to quantify its added value, which resulted in the final model.
- The ability of a model to discriminate between patients with and without DVT was estimated using the area under the ROC curve.
- The reliability or calibration of each model was evaluated by comparing the predicted and observed probabilities for deciles of calculated patient risks and tested using the Hosmer-Lemeshow test.

Oudega et al. Thromb Haemost 2005

Results: bivariate analyses

N = 1295 patients

22% had DVT

Diagnostic variables	Total n=1295 %	DVT present n=289 %	DVT absent n=1006 %	OR (95% CI)
Patient history:				
age (years)	60.0 (17.6)[1]	62.0 (16.8)[1]	59.4 (17.8)[1]	1.01 (1.00 – 2.02)[2]
gender + OC use				
males	36	47	33	1.95 (1.47 – 2.57)
females using OC	10	10	10	1.37 (0.87 – 2.17)
females not using OC	54	43	57	-
gender + HRT use				
males	36	47	33	1.86 (1.42 – 2.43)
females using HRT	2	2	2	1.32 (0.48 – 3.63)
females not using HRT	62	51	66	-
previous DVT	24	21	25	0.82 (0.60 – 1.12)
family history of DVT	23	20	24	0.79 (0.57 – 1.09)
presence of malignancy	6	12	5	2.72 (1.71 – 4.32)
immobilization	14	13	14	0.90 (0.61 – 1.33)
recent surgery	14	19	13	1.59 (1.12 – 2.26)
absence of leg trauma	85	89	84	1.58 (1.05 – 2.36)
pain when walking	81	84	80	1.30 (0.92 – 1.84)
days of symptoms	7.9 (7.6)[1]	6.9 (6.7)[1]	8.2 (7.8)[1]	0.98 (0.96 – 0.99)[3]
Physical examination:				
vein distension	20	28	17	1.88 (1.39 – 2.55)
deep vein system tenderness	71	72	71	1.04 (0.78 – 1.39)
swelling whole leg	45	57	42	1.84 (1.41 – 2.39)
calf difference ≥ 3cm	43	67	36	3.63 (2.75 – 4.79)
D-dimer abnormal VIDAS n= 918 Tinaquant n= 377 Combined assays	 78 65 74	 99 98 99	 72 54 66	 38.2 (9.40 – 155.3) 37.3 (9.00 – 154.8) 35.7 (13.3 - 100.0)

DVT = deep vein thrombosis, n = number of patients, OR = Odds Ratio, 95%CI = 95% Confidence Interval; OC=oral contraceptive, HRT=hormonal replacement therapy; -=reference category; D-dimer abnormal for VIDAS ≥ 500 ng/ml and Tinaquant ≥ 400 ng/l; [1]Mean (standard deviation), [2] OR is estimated per year increase or decrease), [3] OR is estimated per day increase or decrease.

Results: multivariable analyses

Table 2: Independent diagnostic indicators of DVT. The final multivariate model, the figures are estimated after model validation and adjustment for over-fitting.

Diagnostic variables	Odds ratio	Regression coefficient*	p-value	Points for the rule
Male gender	1.80 (1.36 – 2.16)	0.59	<0.001	1
Oral contraceptive use	2.12 (1.32 – 3.35)	0.75	0.002	1
Presence of malignancy	1.52 (1.05 – 2.44)	0.42	0.082	1
Recent surgery	1.46 (1.02 – 2.09)	0.38	0.044	1
Absence of leg trauma	1.82 (1.25 – 2.66)	0.60	0.002	1
Vein distension	1.62 (1.19 – 2.20)	0.48	0.002	1
Calf difference ≥ 3 cm	3.10 (2.36 – 4.06)	1.13	<0.001	2
D-dimer abnormal	20.3 (8.25 – 49.9)	3.01	<0.001	6
Constant		–5.47		

DVT= deep vein thrombosis; *=natural logarithm of the odds ratio; D-dimer abnormal for VIDAS ≥ 500 ng/ml and Tinaquant ≥ 400 ng/ml. Probability of DVT as estimated by the final model =1/(1+exp-(-5·47 + 0·59*male gender + 0·75*OC use + 0·42*presence of malignancy + 0·38*recent surgery + 0·60*absence of leg trauma + 0·48*vein distension + 1·13*calf difference ≥ 3cm + 3·01*abnormal D-dimer)).

Oudega et al. Thromb Haemost 2005

Multivariable approach

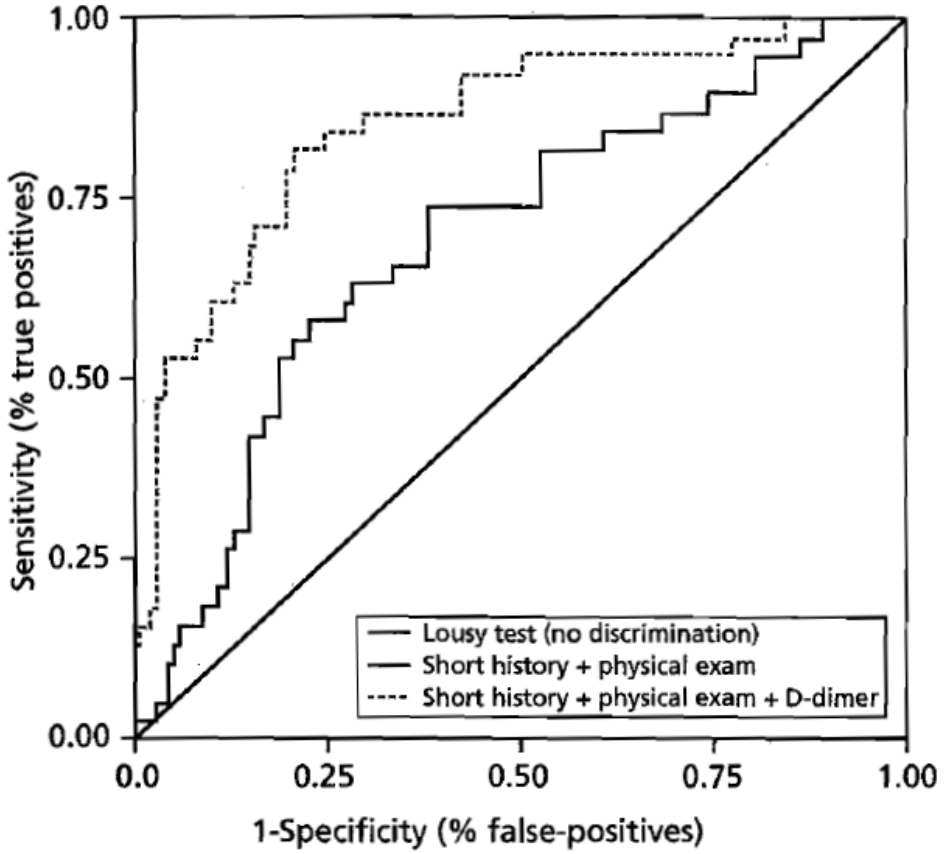


FIGURE 3.3 Example of an ROC curve of the reduced multivariable logistic regression model, including the same six determinants as in Figure 3.2. The ROC area of the "reduced history + physical model" (red) was 0.70 (95% confidence interval [CI], 0.66–0.74) and of the same model added with the D-dimer assay (green) 0.84 (95% CI, 0.80–0.88).

Moons KGM. In: Grobbee & Hoes. Clinical Epidemiology. 2009

Results: scoring system

*1*male gender + 1*OC use + 1*presence of malignancy + 1*recent surgery + 1*absence of trauma + 1*vein distension + 2*calf difference ≥ 3cm +6*abnormal D-dimer test.*

Table 4: Prevalence of DVT across four score (risk) categories.

Probability or risk Category	number of patients n (%)[1]	DVT present n (%)[2]	DVT absent n (%)[3]
Very low (0–3)	293 (23)	2 (0.7)	291 (99.3)
Low (4–5)	66 (5)	3 (4.5)	63 (95.5)
Moderate (7–9)	663 (51)	144 (21.7)	519 (78.3)
High (10–13)	273 (21)	140 (51.3)	133 (48.7)

1=proportion of all (1295) patients; 2=proportion of presence of DVT within risk category; 3=proportion of absence of DVT within risk category.

Oudega et al. Thromb Haemost 2005

TB EXAMPLES

OPEN ACCESS Freely available online

PLoS one

Comparison of a Clinical Prediction Rule and a LAM Antigen-Detection Assay for the Rapid Diagnosis of TBM in a High HIV Prevalence Setting

Vinod B. Patel[1], Ravesh Singh[2], Cathy Connolly[3], Victoria Kasprowicz[2], Allimudin Zumla[4], Thumbi Ndungu[2], Keertan Dheda[4,5,6]*

1 Department of Neurology, University of KwaZulu Natal, Berea, South Africa, **2** HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R. Mandela School of Medicine, University of KwaZulu Natal, Berea, South Africa, **3** Biostatistics Unit, Medical Research Council, Durban, South Africa, **4** Department of Infection, Centre for Infectious Diseases and International Health, University College London, London, United Kingdom, **5** Lung Infection and Immunity Unit, Division of Pulmonology and Department of Medicine, University of Cape Town Lung Institute, University of Cape Town, Cape Town, South Africa, **6** Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa

Abstract

Background/Objective: The diagnosis of tuberculous meningitis (TBM) in resource poor TB endemic environments is challenging. The accuracy of current tools for the rapid diagnosis of TBM is suboptimal. We sought to develop a clinical-prediction rule for the diagnosis of TBM in a high HIV prevalence setting, and to compare performance outcomes to conventional diagnostic modalities and a novel lipoarabinomannan (LAM) antigen detection test (Clearview-TB®) using cerebrospinal fluid (CSF).

Methods: Patients with suspected TBM were classified as definite-TBM (CSF culture or PCR positive), probable-TBM and non-TBM.

Results: Of the 150 patients, 84% were HIV-infected (median [IQR] CD4 count = 132 [54; 241] cells/μl). There were 39, 55 and 54 patients in the definite, probable and non-TBM groups, respectively. The LAM sensitivity and specificity (95%CI) was 31% (17;48) and 94% (85;99), respectively (cut-point ≥0.18). By contrast, smear-microscopy was 100% specific but detected none of the definite-TBM cases. LAM positivity was associated with HIV co-infection and low CD4 T cell count (CD4<200 vs. >200 cells/μl; p = 0.03). The sensitivity and specificity in those with a CD4<100 cells/μl was 50% (27;73) and 95% (74;99), respectively. A clinical-prediction rule ≥6 derived from multivariate analysis had a sensitivity and specificity (95%CI) of 47% (31;64) and 98% (90;100), respectively. When LAM was combined with the clinical-prediction-rule, the sensitivity increased significantly (p<0.001) to 63% (47;68) and specificity remained high at 93% (82;98).

Conclusions: Despite its modest sensitivity the LAM ELISA is an accurate rapid rule-in test for TBM that has incremental value over smear-microscopy. The rule-in value of LAM can be further increased by combination with a clinical-prediction rule, thus enhancing the rapid diagnosis of TBM in HIV-infected persons with advanced immunosuppression.

Citation: Patel VB, Singh R, Connolly C, Kasprowicz V, Zumla A, et al. (2010) Comparison of a Clinical Prediction Rule and a LAM Antigen-Detection Assay for the Rapid Diagnosis of TBM in a High HIV Prevalence Setting. PLoS ONE 5(12): e15664. doi:10.1371/journal.pone.0015664

Editor: Ben J. Marais, University of Stellenbosch, South Africa

Received September 30, 2010; **Accepted** November 22, 2010; **Published** December 22, 2010

Funding: This work was supported by the Columbia University-Southern African Fogarty AIDS International Training and Research Program funded by the Fogarty International Center, National Institutes of Health (grant # D43TW00231; VP), a South African Medical Research Council (SA MRC) grant (VP and KD), the European Union Seventh Framework Programme (TBsusgent; VP and KD), the South African National Research Foundation Research Chairs Initiative (SARChI; TN and KD), a SA MRC Career Development Award (KD), and the European and Developing Countries Clinical Trials Partnership (TESA and TB-NEAT). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: keertan.dheda@uct.ac.za

INT J TUBERC LUNG DIS e-publication ahead of print 28 June 2012

http://dx.doi.org/10.5588/ijtld.11.0743

A model to rule out smear-negative tuberculosis among symptomatic HIV patients using C-reactive protein

G. G. Alvarez,* E. Sabri,* D. Ling,† D. W. Cameron,* G. Maartens,‡ D. Wilson§

*Ottawa Hospital Research Institute, University of Ottawa, Division of Respirology and Infectious Diseases, The Ottawa Hospital, Ottawa, Ontario, †Department of Epidemiology, McGill University, Montréal, Québec, Canada; ‡Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, §Department of Medicine, University of KwaZulu-Natal, Edendale Hospital, Pietermaritzburg, KwaZulu-Natal, South Africa

SUMMARY

SETTING: Improved diagnostic algorithms for sputum smear-negative tuberculosis (SNTB) are needed to address the dramatic increase in SNTB in regions with high human immunodeficiency virus (HIV) prevalence.

OBJECTIVE: To determine whether the addition of C-reactive protein (CRP) to a prediction model using simple clinical criteria improves the diagnosis of SNTB among mostly antiretroviral-naïve adult HIV TB suspects in an out-patient setting.

DESIGN: A multiple logistic regression model was derived from a database of 228 HIV patients to predict the risk of SNTB using data from a previous prospective study.

RESULTS: The derived model demonstrated that male sex, night sweats, fever, low body mass index and anemia increased the probability of having SNTB. CRP improved the accuracy of the model (without CRP, area under the curve [AUC] 0.75, 95%CI 0.68–0.81 vs. model with CRP, AUC 0.81, 95%CI 0.76–0.87, $P = 0.0014$) to predict SNTB. Using reclassification tables, CRP correctly reclassified 27.9% of the patients (net reclassification improvement, $P = 0.0005$) into higher or lower risk categories. The strongest effect was seen in the reclassification improvement among patients with no TB, which was 20.6% ($P = 0.0023$).

CONCLUSION: CRP improved the performance of the prediction model in the diagnosis of SNTB in HIV patients, and may play a role in ruling out SNTB in this population. Prospective validation of this model is needed.

KEY WORDS: biomarkers; tuberculosis diagnostics; clinical prediction rule

Evaluation of Quantitative IFN-γ Response for Risk Stratification of Active Tuberculosis Suspects

John Z. Metcalfe[1], Adithya Cattamanchi[1], Eric Vittinghoff[2], Christine Ho[3,4], Jennifer Grinsdale[3], Philip C. Hopewell[1,3], L. Masae Kawamura[3], and Payam Nahid[1,3]

[1]Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, and [2]Department of Epidemiology and Biostatistics, University of California, San Francisco; [3]Tuberculosis Control Section, Department of Public Health; and [4]Centers for Disease Control and Prevention, San Francisco, California

Rationale: The contribution of interferon-γ release assays (IGRAs) to appropriate risk stratification of active tuberculosis suspects has not been studied.

Objectives: To determine whether the addition of quantitative IGRA results to a prediction model incorporating clinical criteria improves risk stratification of smear-negative–tuberculosis suspects.

Methods: Clinical data from tuberculosis suspects evaluated by the San Francisco Department of Public Health Tuberculosis Control Clinic from March 2005 to February 2008 were reviewed. We excluded tuberculosis suspects who were acid fast–bacilli smear–positive, HIV-infected, or under 10 years of age. We developed a clinical prediction model for culture-positive disease and examined the benefit of adding quantitative interferon (IFN)-γ results measured by QuantiFERON-TB Gold (Cellestis, Carnegie, Australia).

Measurements and Main Results: Of 660 patients meeting eligibility criteria, 65 (10%) had culture-proven tuberculosis. The odds of active tuberculosis increased by 7% (95% confidence interval [CI], 3–11%) for each doubling of IFN-γ level. The addition of quantitative IFN-γ results to objective clinical data significantly improved model performance (c-statistic 0.71 vs. 0.78; $P < 0.001$) and correctly reclassified 32% of tuberculosis suspects (95% CI, 11–52%; $P < 0.001$) into higher-risk or lower-risk categories. However, quantitative IFN-γ results did not significantly improve appropriate risk reclassification beyond that provided by clinician assessment of risk (4%; 95% CI, −7 to +22%; $P = 0.14$).

Conclusions: Higher quantitative IFN-γ results were associated with active tuberculosis, and added clinical value to a prediction model incorporating conventional risk factors. Although this benefit may be attenuated within highly experienced centers, the predictive accuracy of quantitative IFN-γ levels should be evaluated in other settings.

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The role of interferon-γ release assays (IGRAs) in the evaluation of active tuberculosis suspects is controversial. To date, whether IGRAs improve classification of smear negative tuberculosis suspects into clinically relevant risk categories has not been examined.

What This Study Adds to the Field

Quantitative interferon-γ levels measured by QuantiFERON-TB Gold improves risk stratification of smear-negative active tuberculosis suspects when added to objective clinical and demographic risk factors. However, this benefit is attenuated when the judgment of experienced clinicians is also considered.

2) and have better correlation with gradient of *M. tuberculosis* exposure (3–8). In 2005, the Centers for Disease Control and Prevention recommended that QuantiFERON TB-Gold (QFT-G; Cellestis, Carnegie, Australia), the first FDA-approved, commercially available IGRA to experience widespread use, could be used for targeted screening of LTBI in all circumstances in which the tuberculin skin test (TST) is used (9).

Although the advantages of IGRAs in diagnosing LTBI are

AJRCCM 2010

Eur Respir J 2012; 39: 163–171
DOI: 10.1183/09031936.00058911

Interferon release does not add discriminatory value to smear-negative HIV–tuberculosis algorithms

M.X. Rangaka*,#,¶, H.P. Gideon#, K.A. Wilkinson#,+,§, M. Paif, J. Mwansa-Kambafwile#, G. Maartens§,, J.R. Glynn¶, A. Boulle*, K. Fielding¶, R. Goliath#, R. Titus#, S. Mathee## and R.J. Wilkinson#,+,§,¶¶**

ABSTRACT: Clinical algorithms for evaluating HIV-infected individuals for tuberculosis (TB) prior to isoniazid preventive therapy (IPT) perform poorly, and interferon-γ release assays (IGRAs) have moderate accuracy for active TB. It is unclear whether, when used as adjunct tests, IGRAs add any clinical discriminatory value for active TB diagnosis in the pre-IPT assessment.

779 sputum smear-negative HIV-infected persons, established on or about to commence combined antiretroviral therapy (ART), were screened for TB prior to IPT. Stepwise multivariable logistic regression was used to develop clinical prediction models. The discriminatory ability was assessed by receiver operator characteristic area under the curve (AUC). QuantiFERON®-TB Gold in-tube (QFT-GIT) was evaluated.

The prevalence of smear-negative TB by culture was 6.4% (95% CI 4.9–8.4%). Used alone, QFT-GIT and the tuberculin skin test (TST) had comparable performance; the post-test probability of disease based on single negative tests was 3–4%. In a multivariable model, the QFT-GIT test did not improve the ability of a clinical algorithm, which included not taking ART, weight <60 kg, no prior history of TB, any one positive TB symptom/sign (cough ⩾2 weeks) and CD4+ count <250 cells per mm^3, to discriminate smear-negative culture-positive and -negative TB (72% to 74%; AUC comparison p=0.33). The TST marginally improved the discriminatory ability of the clinical model (to 77%, AUC comparison p=0.04).

QFT-GIT does not improve the discriminatory ability of current TB screening clinical algorithms used to evaluate HIV-infected individuals for TB ahead of preventive therapy. Evaluation of new TB diagnostics for clinical relevance should follow a multivariable process that goes beyond test accuracy.

AFFILIATIONS
*Centre for Infectious Disease and Epidemiology Research (CIDER),
#Clinical Infectious Diseases Research Initiative, Institute of Infectious Diseases and Molecular Medicine,
§Dept of Medicine,
**Division of Clinical Pharmacology, University of Cape Town, and
##Provincial Administration of the Western Cape, Cape Town, South Africa.
¶London School of Hygiene and Tropical Medicine,
+MRC National Institute for Medical Research, and
¶¶Division of Medicine, Imperial College London, London, UK.
fMcGill University and Montreal Chest Institute, Montreal, QC, Canada.

CORRESPONDENCE
M.X. Rangaka
Centre for Infectious Disease and

Strengths and limitations of this approach

- Avoids the common problem of sens/spec of single tests in isolation when reality is quite different
- Requires the same data that are collected in a standard test accuracy study, but will require data to be collected in a sequential way, starting from simple to complex
- Multivariable analyses and model building – gives us the ability to account for multiple tests in a natural clinical context
  - Can look at both rule-in and rule-out values at the same time

- Some drawbacks:
  - Blinding is hard to ensure if the same person is collecting all the clinical data in a sequential manner – so, covariates are not necessarily independent
  - Outcomes are not clinically intuitive: e.g. AUC
    - A significantly different AUC might not mean much from a clinical perspective
  - Will require conversion of multivariable model output into scoring systems
  - Scoring systems are not necessarily easy to apply in clinical contexts – will require further knowledge translation work

Newer approaches to assessing added value

STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:157–172
Published online 13 June 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.2929

WILEY InterScience® DISCOVER SOMETHING GREAT

Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond

Michael J. Pencina[1,*,†], Ralph B. D'Agostino Sr[1], Ralph B. D'Agostino Jr[2] and Ramachandran S. Vasan[3]

[1]*Department of Mathematics and Statistics, Framingham Heart Study, Boston University, 111 Cummington St., Boston, MA 02215, U.S.A.*
[2]*Department of Biostatistical Sciences, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, U.S.A.*
[3]*Framingham Heart Study, Boston University School of Medicine, 73 Mount Wayte Avenue, Suite 2, Framingham, MA 01702-5803, U.S.A.*

Net Reclassification Improvement (NRI) Summary Index

GOOD BAD BAD GOOD

$$[P(up|D=1)-P(down|D=1)]-[P(up|D=0)-P(down|D=0)]$$

Correct classification of cases + Correct classification of non-cases

Pencina *Stat Med* 2007

Case study #1

Eur Respir J 2012; 39: 163–171
DOI: 10.1183/09031936.00058911

Interferon release does not add discriminatory value to smear-negative HIV–tuberculosis algorithms

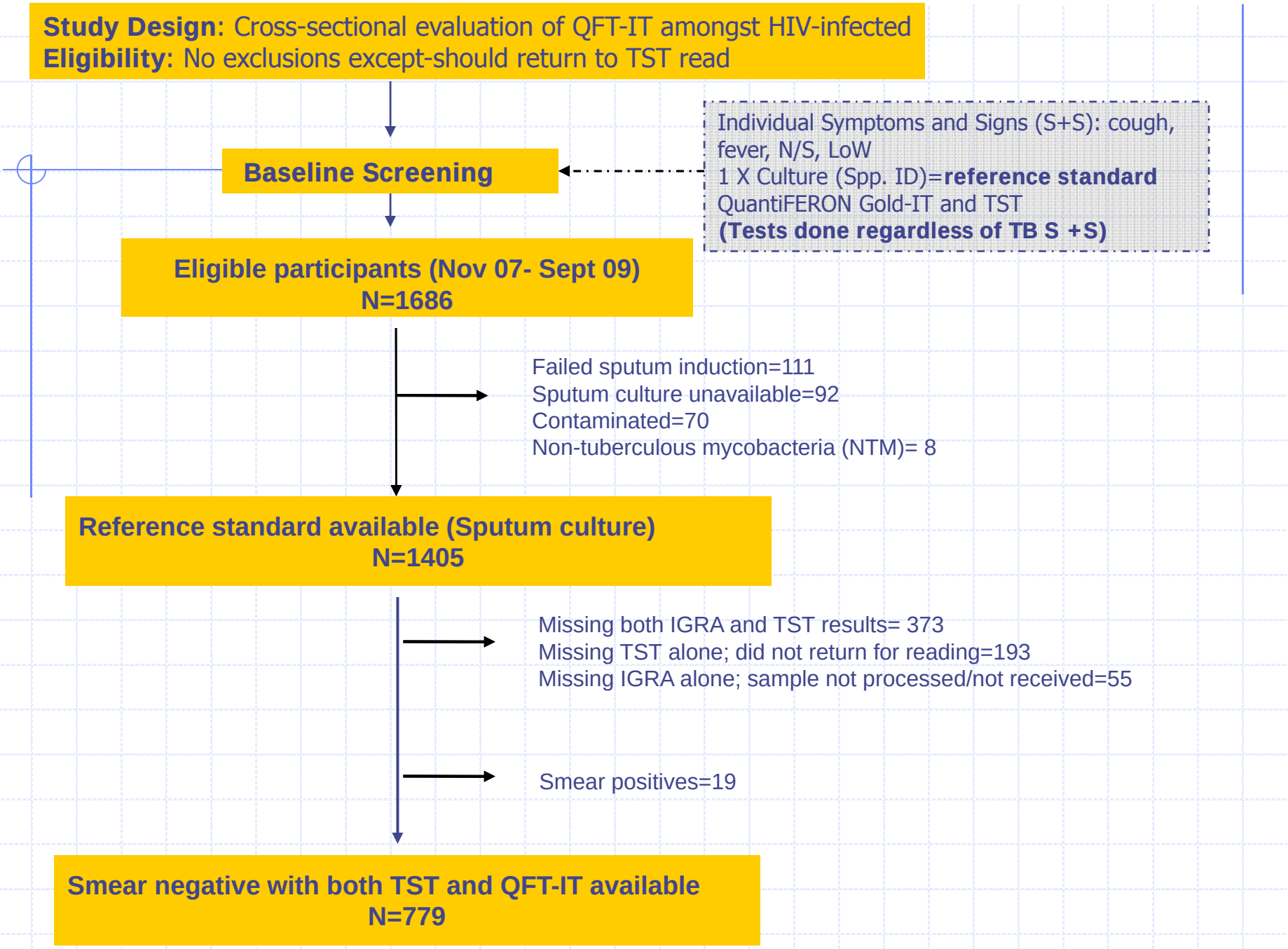
M.X. Rangaka*,#,¶, H.P. Gideon#, K.A. Wilkinson#,+,§, M. Pai[f], J. Mwansa-Kambafwile#, G. Maartens§,, J.R. Glynn¶, A. Boulle*, K. Fielding¶, R. Goliath#, R. Titus#, S. Mathee## and R.J. Wilkinson#,+,§,¶¶**

Courtesy: Lele Rangaka

The Question

Does QuantiFERON Gold In tube add to current clinical algorithms to detect smear negative TB in HIV-infected patients screened for preventive therapy?

Flow into the study (STARD-standards for reporting diagnostic accuracy studies)



Assessing discriminatory value: Model Steps

Potential Predictors (pre-determined)

Clinical: Age, Gender, Weight, ART status, Prior TB, TB Symptoms and signs (any one TB symptom or sign positive), CD4+ count

Simple test of TB infection: TST (5mm cutoff)

More complex: QFT-IT (Standard Manufacturer's cutoffs used)

Model Step 1
Identify best clinical model

- Multivariable logistic regression
- Fit all clinical *a priori* determined predictors; sequentially in the manner collected at the first clinical visit
- Stepwise selection & clinical judgement

Model Step 2
Add simple test

Add TST (5mm cutoff)

Model Step 3
Add complex test

Add QFT-IT (Standard Manufacturer's cutoffs used)

Model Step 4
Both tests added

Add TST & QFT-IT

Discriminatory Ability

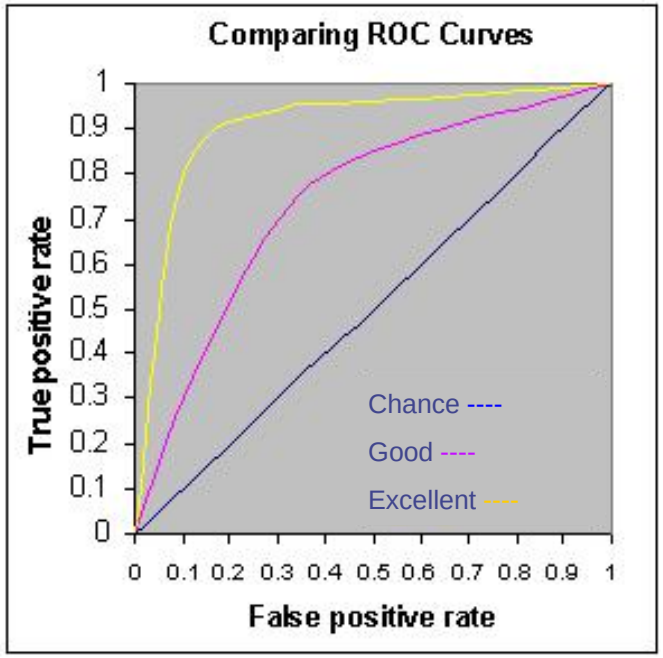
The Area Under the Curve

AUC (c-statistic):

-Interpreted for all diseased and non-diseased pairs: Overall probability that diseased individuals will score higher than non-diseased*

- AUC comparison of model with and without new test is of interest (added value)

- Main criticism: New predictor or marker has to be strongly associated with disease outcome to add value over and beyond existing tools



Reference test Test under assessment	Reference test: Culture	
	(D+)	(D-)
(T+)	True Positive	False Positive
(T-)	False Negative	True Negative

Interpretation:
Overall probability (averaged for all pairs)
AUC ~ 0.50= not better than chance

Description of cohort by *M.tb* culture status: Clinical Observations

Total N=779

(Prevalence of smear negative TB= 6%)

Clinical and Laboratory Features	50 TB culture positive	729 TB culture negative	p-value
Clinical Observations			
Median Age (IQR)	35 (31-40)	36 (31-42)	0.71
Age ≥ 35 y.o	46%	45%	0.92
Male	68%	75%	0.25
No Prior TB	82%	62%	0.004
Median CD4+ count (IQR)	169 (98-239)	198 (136-315)	0.03
CD4 less 250	80%	66% (721)	0.05
Median Weight Kg (IQR)	60 (54-65)	66 (58-76)	<0.001
Weight less than 60kg	52%	33% (722)	0.01
Not on ART at screening	54%	34%	0.004

Description of cohort by *M.tb* culture status: TB symptoms and signs

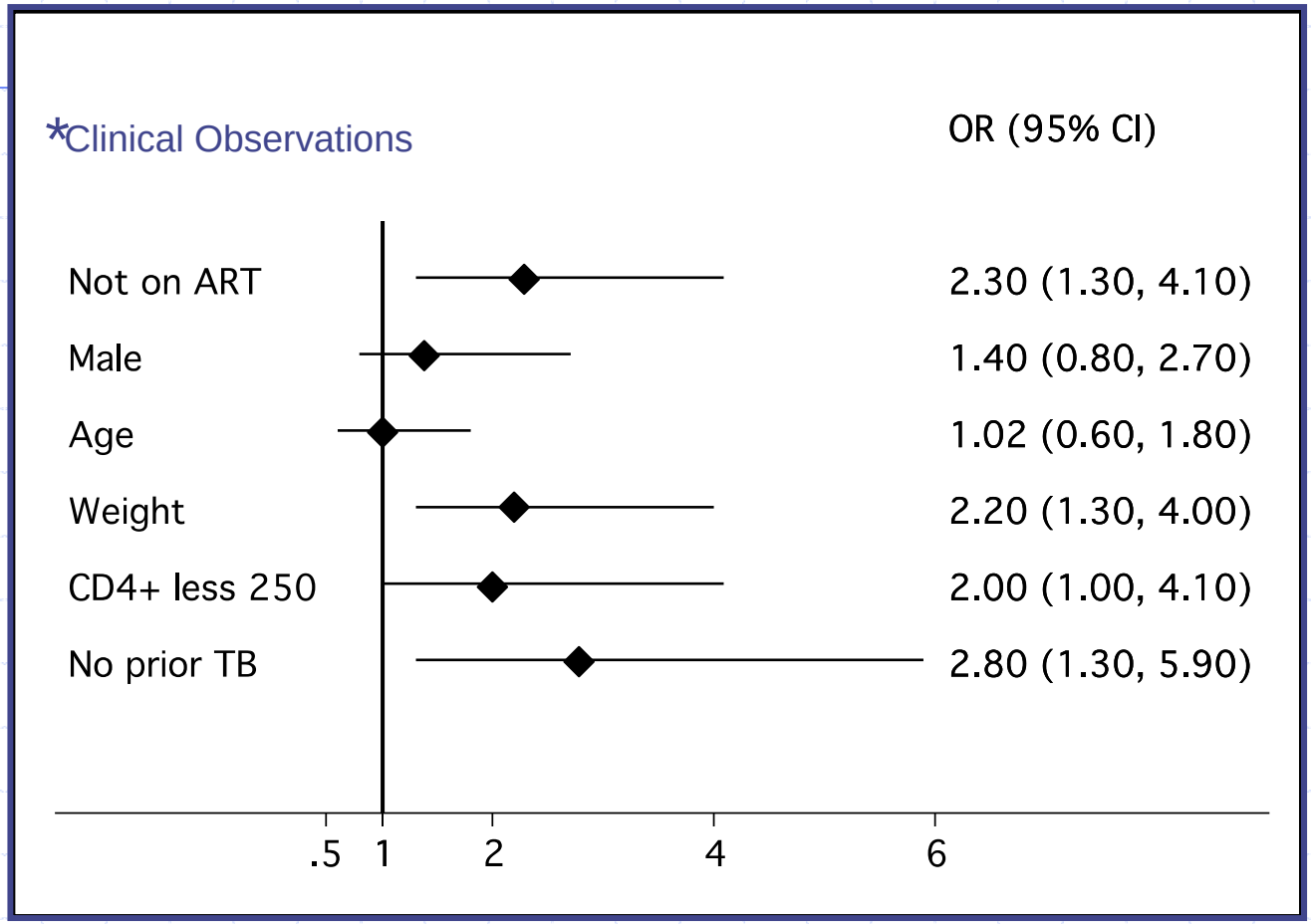
Clinical and Laboratory Features	50 TB culture positive	729 TB culture negative	p-value
Symptoms and signs of TB			
Cough ≥ two weeks	10%	4% (728)	0.05
Night sweats	10%	2% (728)	0.002
Self-reported 'Fever'	1/49	3/727	0.230 (exact)
Nodes on examination	1/49	1/728	0.122 (exact)
Loss of weight	18%	5% (728)	<0.0001
Any one TB symptom or sign positive	26%	8% (728)	<0.0001

*Anyone symptom or sign positive: Cough for >=2 weeks vs. Cough for any duration

Description of cohort by *M.tb* culture status: Tests of TB infection

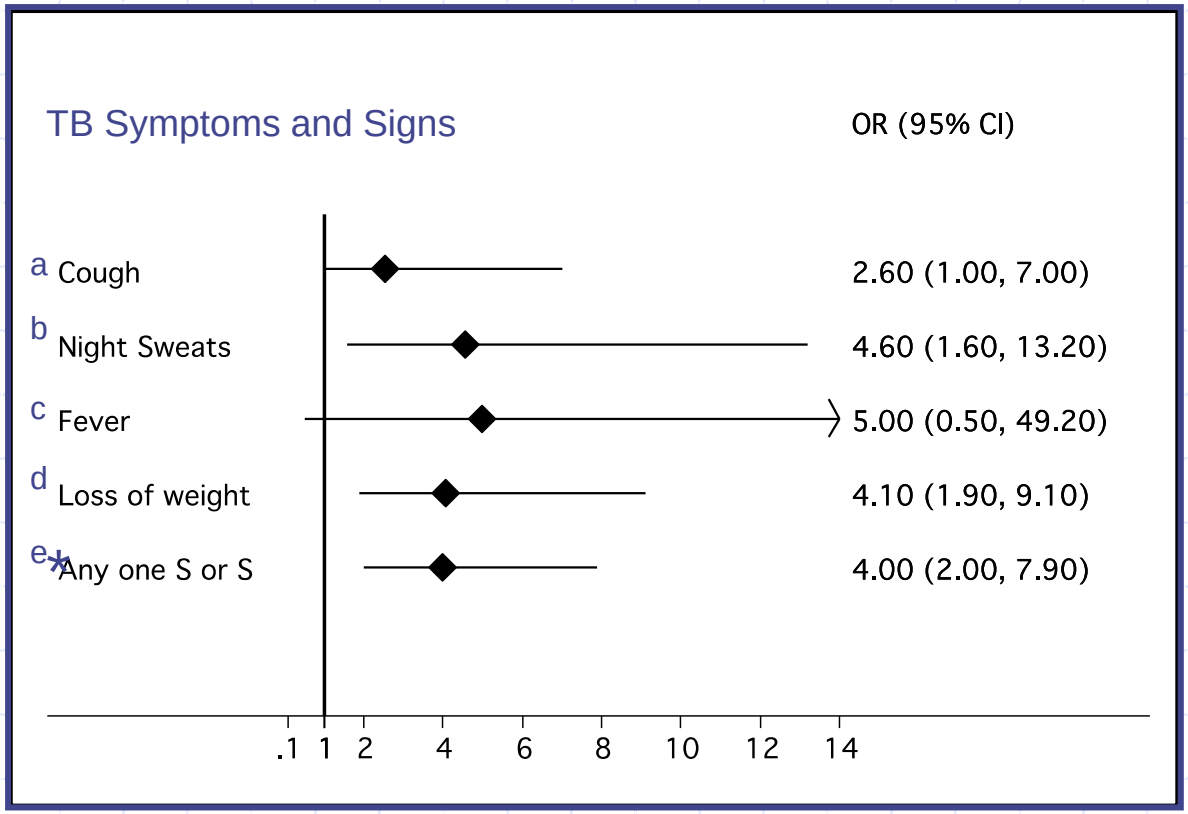
Clinical and Laboratory Features	50 TB culture positive	729 TB culture negative	p-value
Tests of TB infection			
TST positive at 5mm cut-off	68%	41%	<0.0001
TST positive at 10mm cut-off	66%	37%	<0.0001
TST positive at 15mm cut-off	54%	26%	<0.0001
Median TST mm (IQR)	15 (0-20)	0 (0-15)	<0.0001
(Manufacturer's cutoffs) QFT positive	64%	41%	0.004 (exact)
QFT negative	30%	53%	
QFT Indeterminate	6%	7%	
Median QFT quantitative (IQR)	0.5 (0.1-2.6)	0.12 (0-0.85)	0.003
Either TST 5mm/IGRA positive (Indeterminate included with negatives)	80%	56%	0.001
Either TST 5mm/IGRA positive (Indeterminate results excluded)	83% (48)	59% (692)	0.001

Univariable predictors of culture-positive TB disease



*Age>=35y.o, Weight less than 60kg

Univariate predictors of culture-positive TB disease



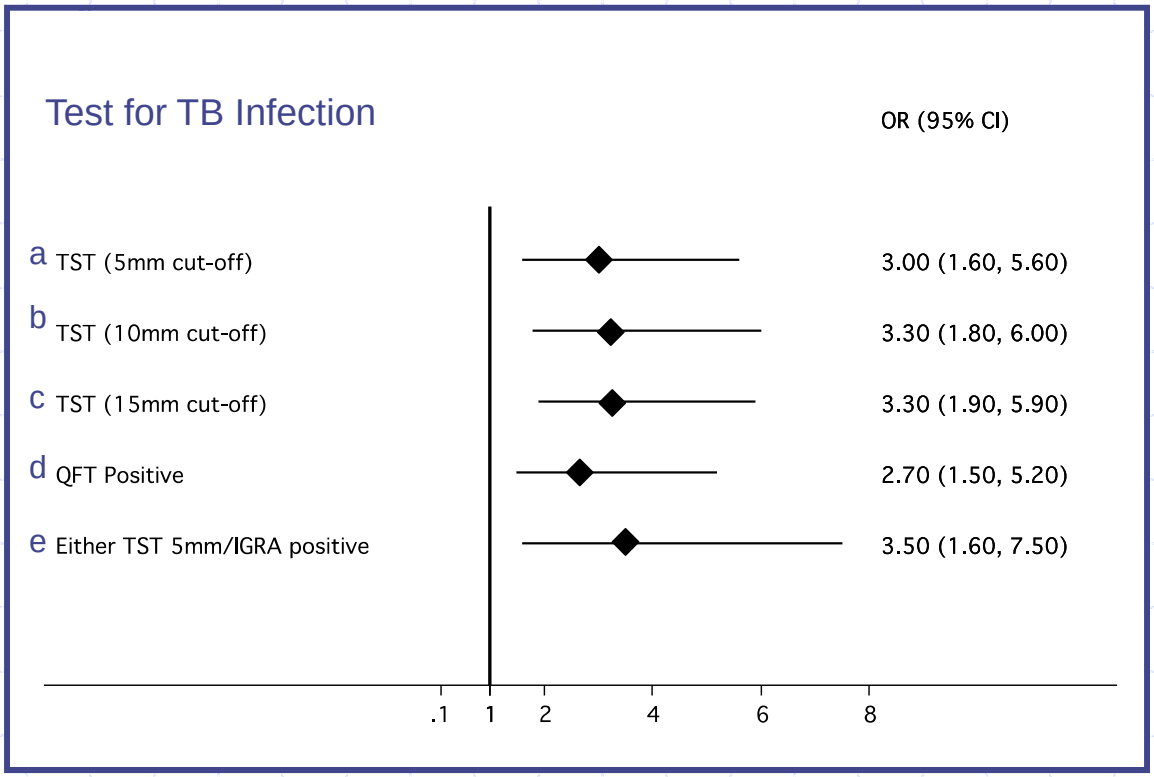
Test Accuracy

	Sensitivity	Specificity	Post-test Pr (Neg)	AUC
a	10	96	5.7	53
b	10	98	5.7	54
c	2	100	5.7	51
d	18	95	5.7	57
e	26	92	4.8	59

Pre-test probability (Prevalence): 6%

*Anyone symptom or sign positive: Includes Nodes and Cough for >=2 weeks

Univariable predictors of culture-positive TB disease

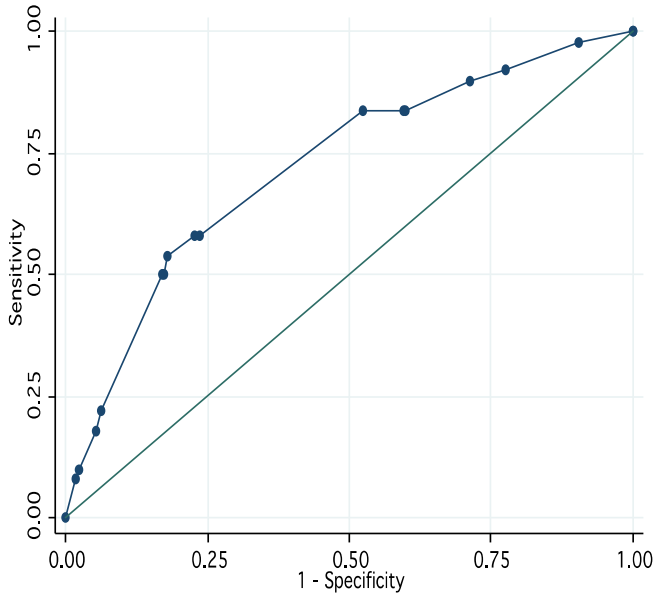


Test Accuracy

	Sensitivity	Specificity	Post-test Pr (Neg)	AUC
a	68	59	3	63
b	66	63	3	64
c	54	74	4	64
d	68	56	4	62
e	83	41	3	62

Pre-test probability (Prevalence): 6%

Discriminatory ability of TB tests (Multivariable Analyses)

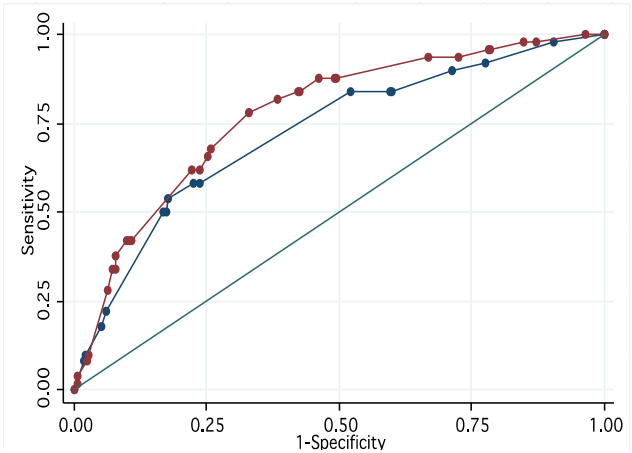


AUC=72%

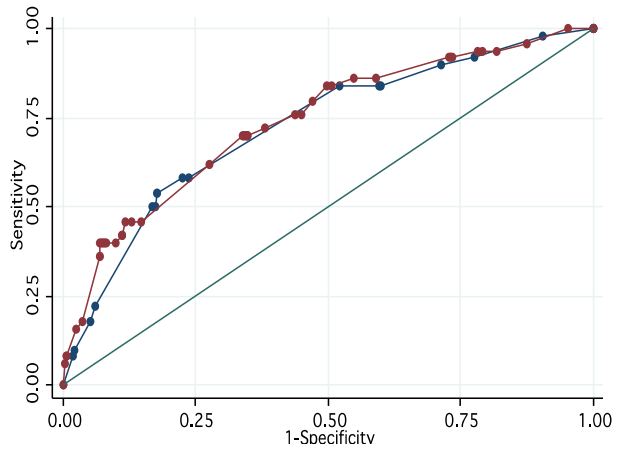
Final clinical model

1. Weight less than 60kg, OR=2
2. No prior TB, OR=3
3. Any one TB S/S positive, OR=3
4. CD4+ less than 250 cells/mm^3 OR=2
5. Not on ART at screening OR=1.2

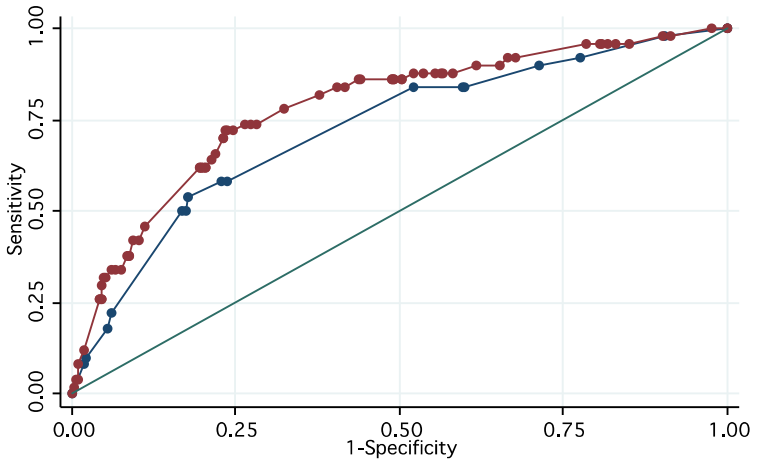
Discriminatory ability of TB tests (Multivariable Analyses)



Clinical **(blue, AUC=72%)** AND TST at 5mm **(red, AUC=77%)**
Comparison p-value=0.03



Clinical **(blue, AUC=72%)** AND QFT **(red, AUC=74%)**
Comparison p-value=0.41



Clinical **(blue, AUC=72%)** and both TST & QFT **(red, AUC=78%)**
Comparison p-value=0.01

Summary

- High prevalence of smear negative culture-positive TB in HIV- infected patients on or starting ART screened for IPT
- Asymptomatic culture-positive TB a concern
- As stand-alone tests, current TB screening tools perform poorly against culture. Best to combine in clinical prediction rule
- QuantiFERON Gold *In Tube*, measuring interferon-gamma, adds little to TB screening tools for evaluating HIV-infected adults for IPT

Case study #2

INT J TUBERC LUNG DIS e-publication ahead of print 28 June 2012

http://dx.doi.org/10.5588/ijtld.11.0743

A model to rule out smear-negative tuberculosis among symptomatic HIV patients using C-reactive protein

G. G. Alvarez,* E. Sabri,* D. Ling,† D. W. Cameron,* G. Maartens,‡ D. Wilson§

*Ottawa Hospital Research Institute, University of Ottawa, Division of Respirology and Infectious Diseases, The Ottawa Hospital, Ottawa, Ontario, †Department of Epidemiology, McGill University, Montréal, Québec, Canada; ‡Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, §Department of Medicine, University of KwaZulu-Natal, Edendale Hospital, Pietermaritzburg, KwaZulu-Natal, South Africa

Derivation of an algorithm to rule out smear negative active TB in HIV patients

Gonzalo G. Alvarez MD, MPH, FRCPC
Assistant Professor at The University of Ottawa at
Associate Scientist at The Ottawa Hospital Research Institute
Divisions of Respirology and Infectious Diseases
Department of Medicine, The Ottawa Hospital

University of Ottawa

The Ottawa Hospital
L'Hôpital d'Ottawa

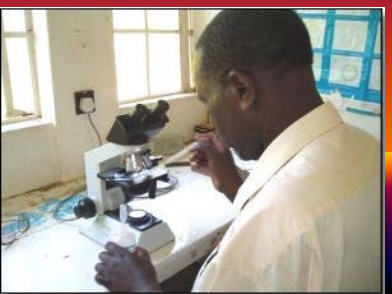
HEALTH
KwaZulu-Natal

Ottawa Health Research Institute
OHRI

Smear microscopy

- The cornerstone of TB diagnosis is smear microscopy in developing countries

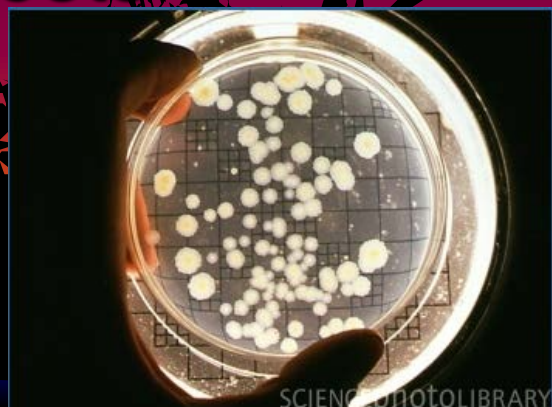
- Method is cheap and can be done in the field *however* many cases of TB can be smear negative and culture positive



TB culture

- Culture is the gold standard test for the diagnosis of TB

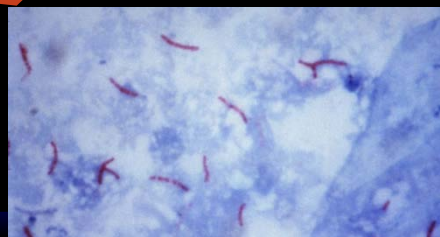
- however in most high burden TB countries it is only used in a fraction of the cases due to associated costs



HIV and SNTB

- HIV positive patients commonly present with smear negative culture positive TB (SNTB) thus making the diagnosis of TB using microscopy difficult

- Smear microscopy in HIV patients has an estimated sensitivity of 35% for active TB

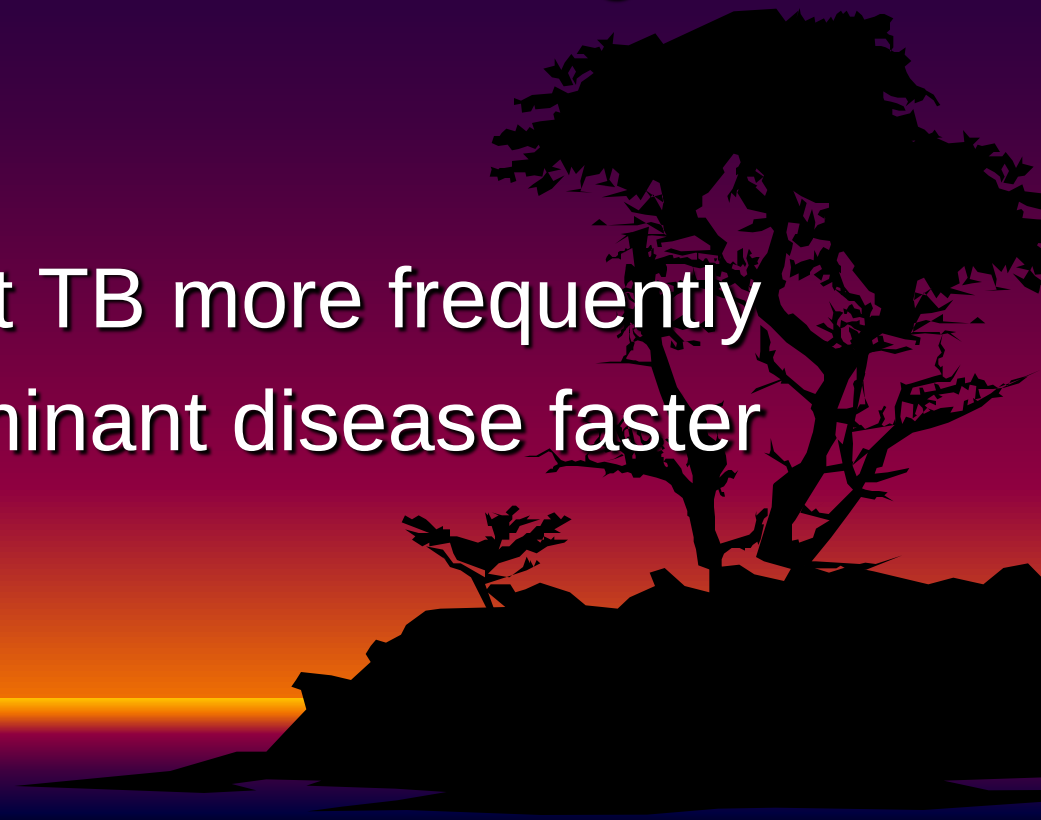


Samb et al1999 IJTLD, Corbett EL 2003 Arch Intern Med

HIV patients & TB

HIV positive patients when compared to HIV negative patients have a higher likelihood of :

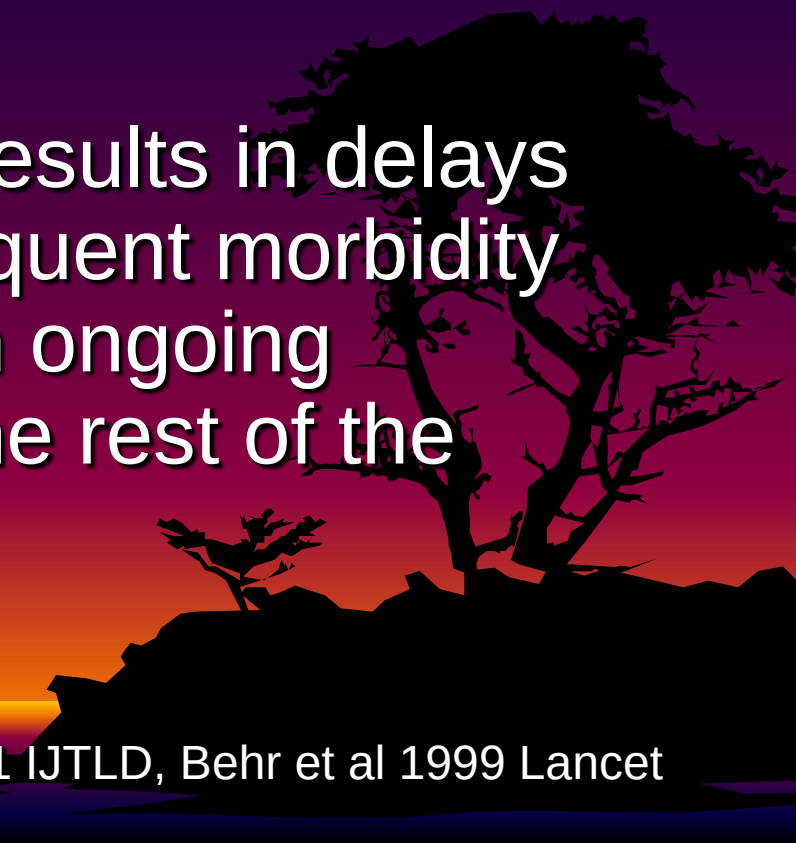
1) contracting TB,
2) reactivate latent TB more frequently
3) progress to fulminant disease faster



Corbett et al Lancet 2003

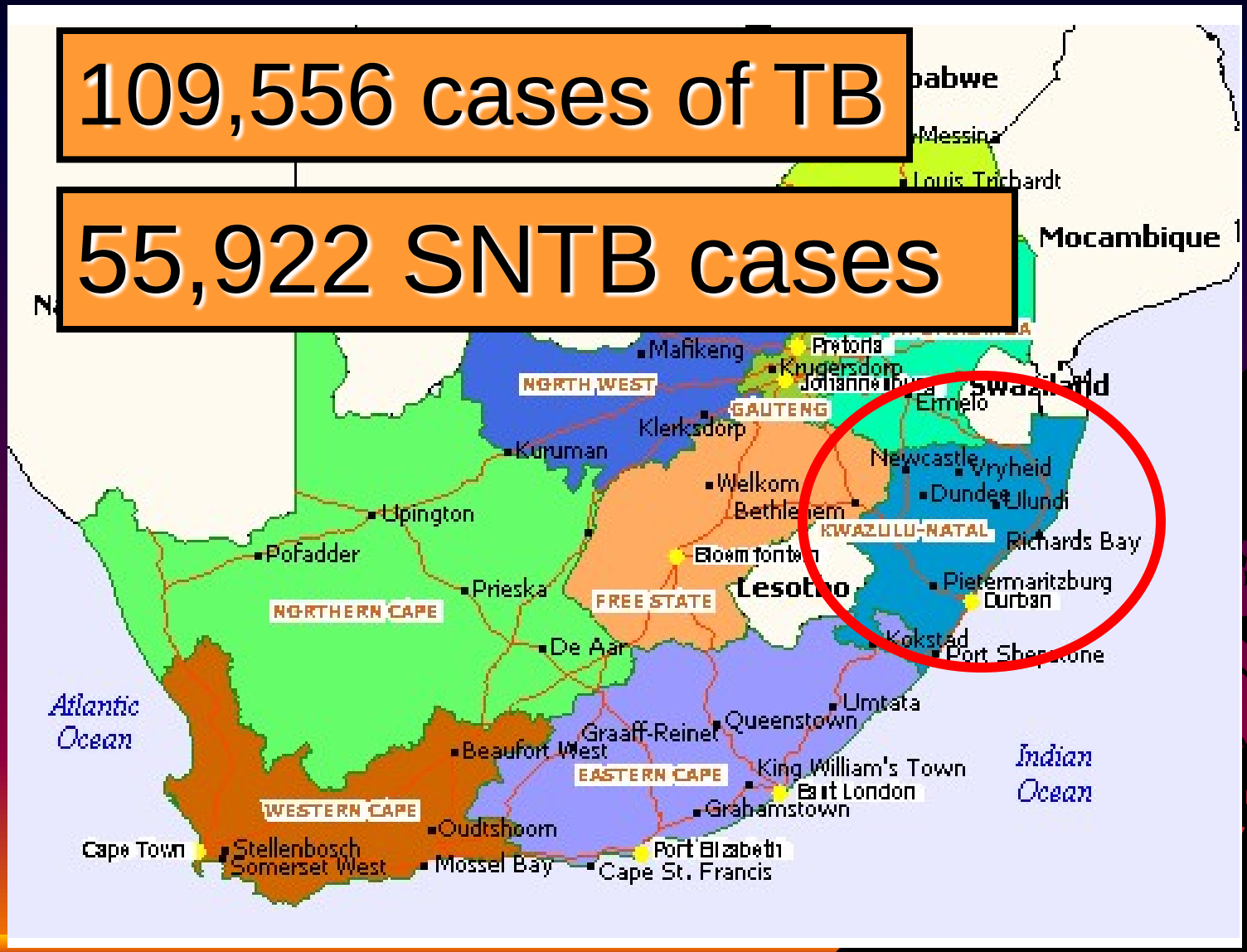
Failure of SNTB diagnosis

- Consequently, many patients are not diagnosed with TB in the community
- Delay in TB diagnosis results in delays in treatment and subsequent morbidity and mortality along with ongoing transmission of TB to the rest of the community



Alvarez GG 2004 IJTLD, Hargreaves NJ 2001 IJTLD, Behr et al 1999 Lancet

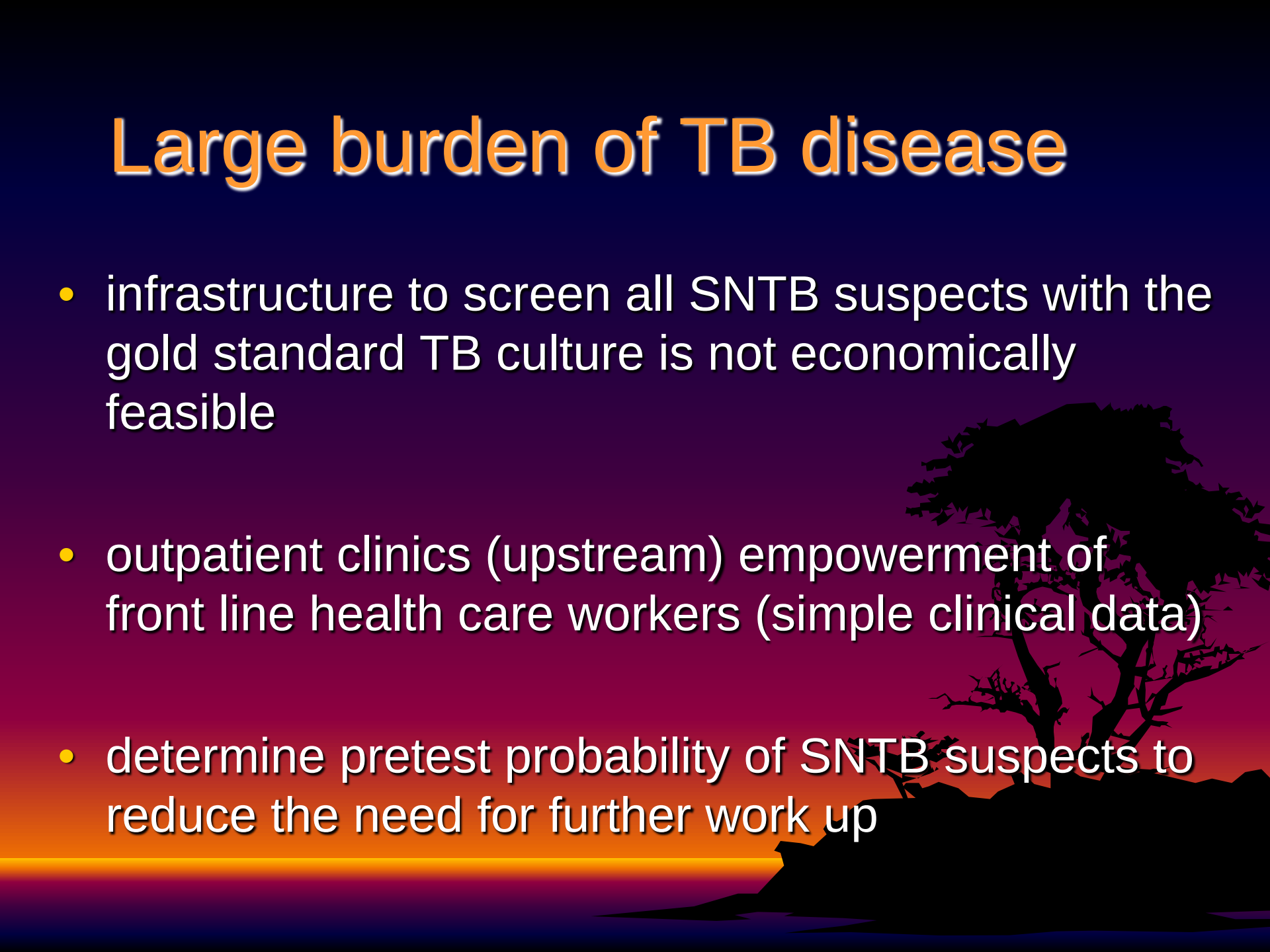
KwaZulu-Natal, South Africa TB stats



HST 2008, South African Health Review, DOH

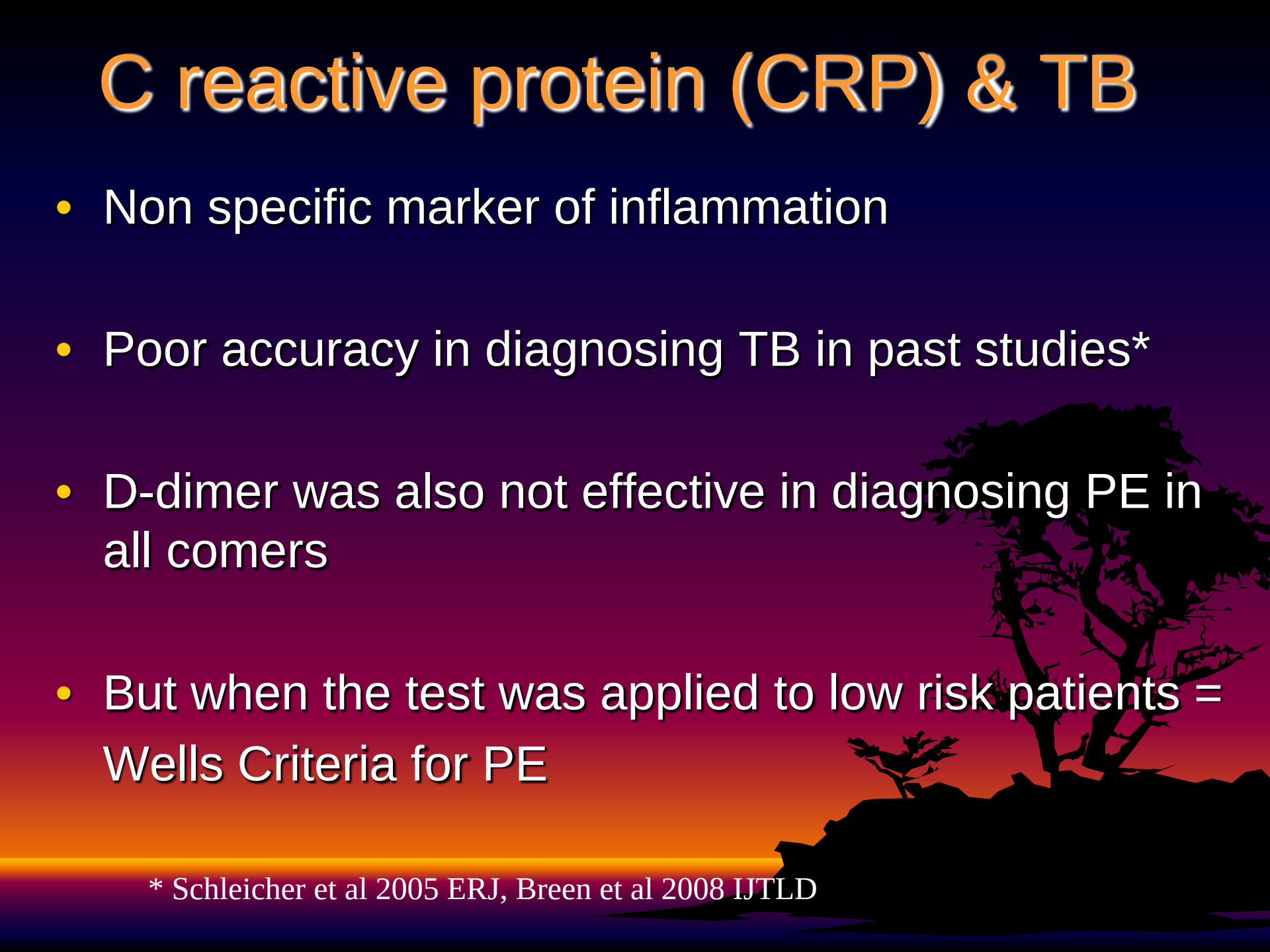
Large burden of TB disease

- infrastructure to screen all SNTB suspects with the gold standard TB culture is not economically feasible
- outpatient clinics (upstream) empowerment of front line health care workers (simple clinical data)
- determine pretest probability of SNTB suspects to reduce the need for further work up



C reactive protein (CRP) & TB

- Non specific marker of inflammation
- Poor accuracy in diagnosing TB in past studies*
- D-dimer was also not effective in diagnosing PE in all comers
- But when the test was applied to low risk patients = Wells Criteria for PE

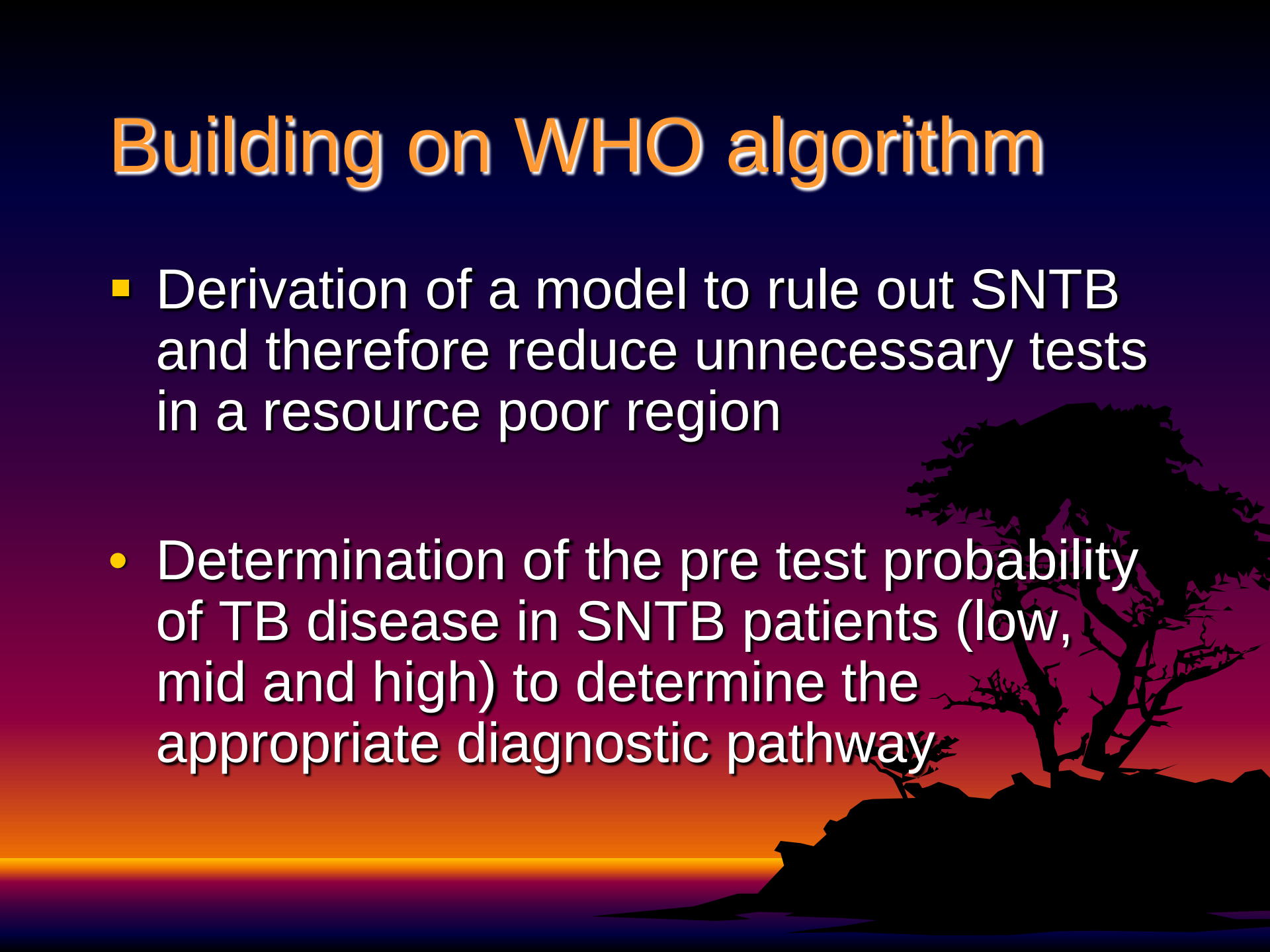


* Schleicher et al 2005 ERJ, Breen et al 2008 IJTLD

Building on WHO algorithm

- Derivation of a model to rule out SNTB and therefore reduce unnecessary tests in a resource poor region

- Determination of the pre test probability of TB disease in SNTB patients (low, mid and high) to determine the appropriate diagnostic pathway



Clinical decision based on pre test probability

Probability of TB	**Clinical decision**
Low	Follow up with community nurse, No CXR or TB culture needed, No further TB investigations ordered
Moderate	CXR and TB Culture should be done
High	PTB treatment

INT J TUBERC LUNG DIS 16(8):1–5

http://dx.doi.org/10.5588/ijtld.11.0743

A model to rule out smear-negative tuberculosis among symptomatic HIV patients using C-reactive protein

G. G. Alvarez,* E. Sabri,* D. Ling,† D. W. Cameron,* G. Maartens,‡ D. Wilson§

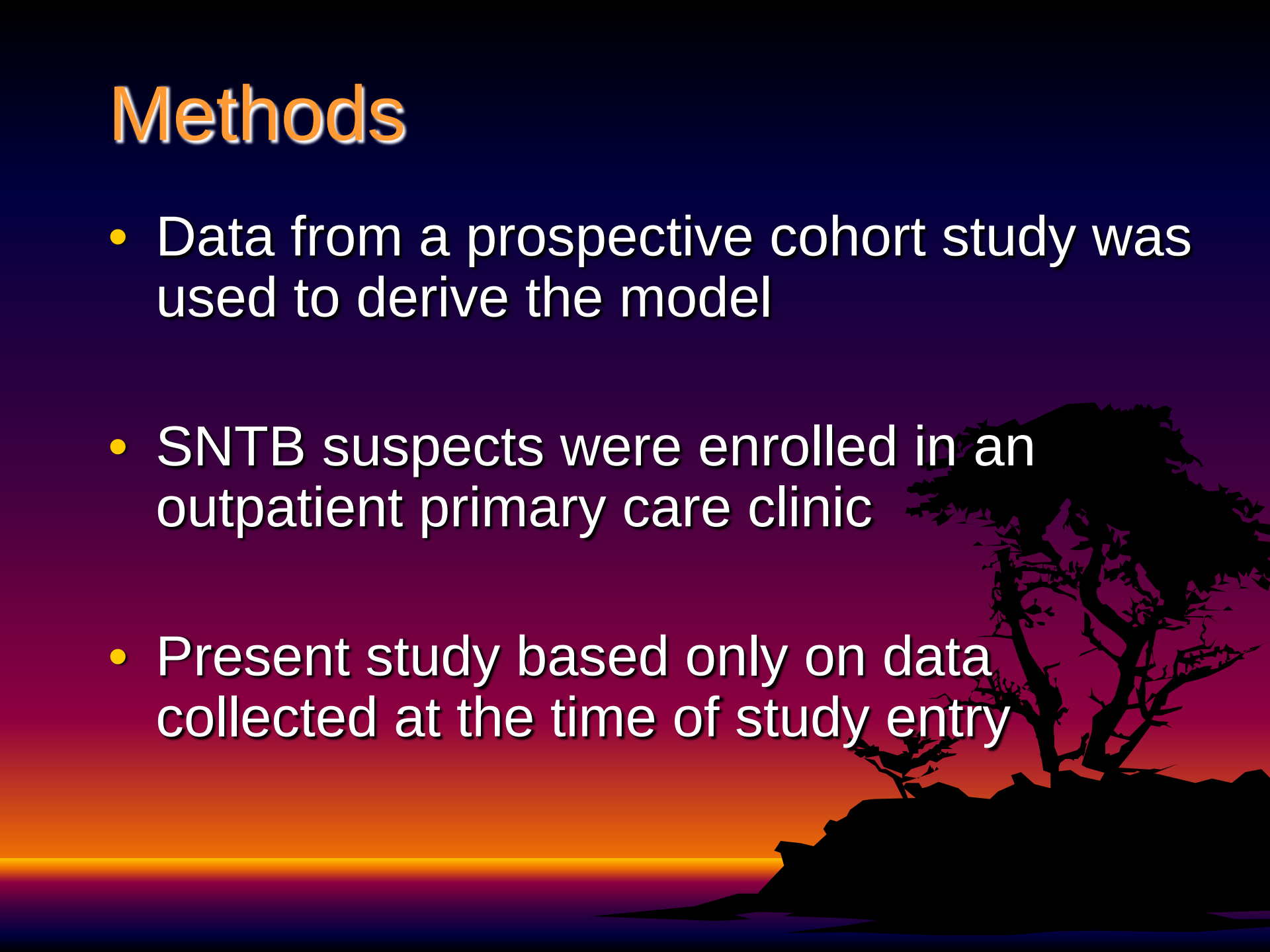
The International Journal of Tuberculosis and Lung Disease

Edendale Hospital, Pietermaritzburg, South Africa

Edendale Hospital catchment area

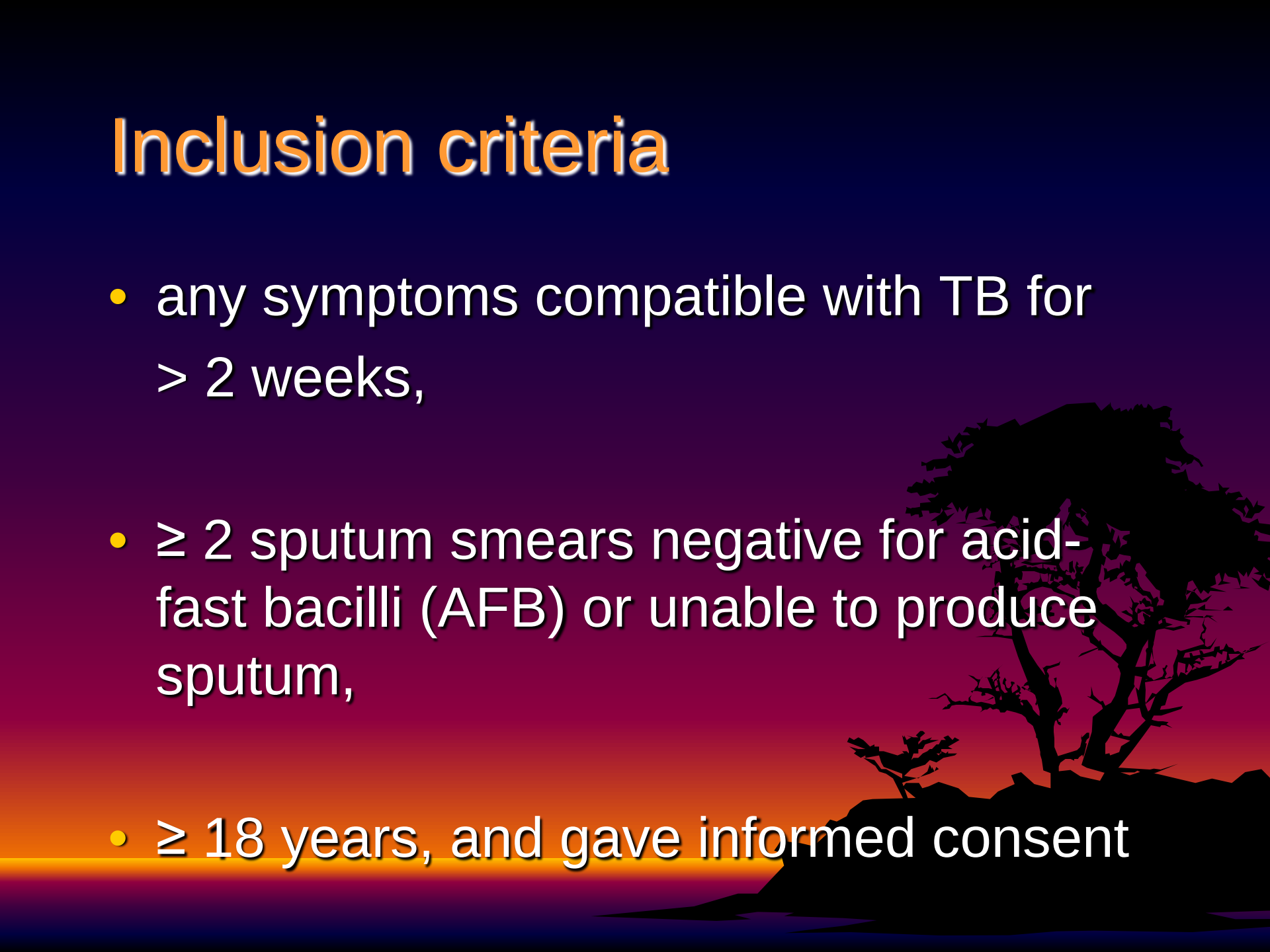
Methods

- Data from a prospective cohort study was used to derive the model
- SNTB suspects were enrolled in an outpatient primary care clinic
- Present study based only on data collected at the time of study entry



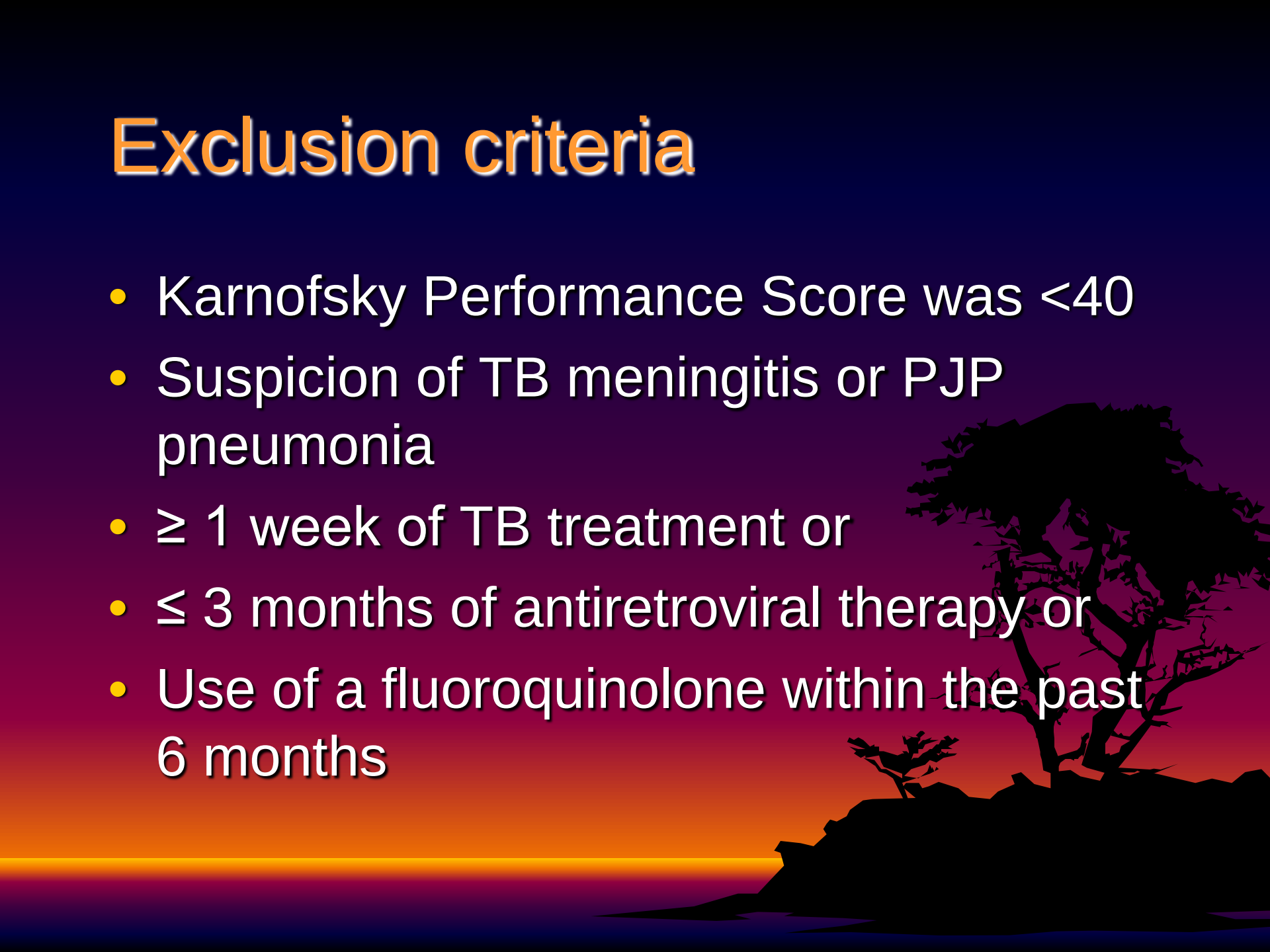
Inclusion criteria

- any symptoms compatible with TB for > 2 weeks,
- ≥ 2 sputum smears negative for acid-fast bacilli (AFB) or unable to produce sputum,
- ≥ 18 years, and gave informed consent



Exclusion criteria

- Karnofsky Performance Score was <40
- Suspicion of TB meningitis or PJP pneumonia
- ≥ 1 week of TB treatment or
- ≤ 3 months of antiretroviral therapy or
- Use of a fluoroquinolone within the past 6 months



Outcomes

- TB present if 1) Culture positive OR 2) acid fast bacilli positive OR 3) histology positive

- TB not present if 1 and 2 and 3 were negative

Results

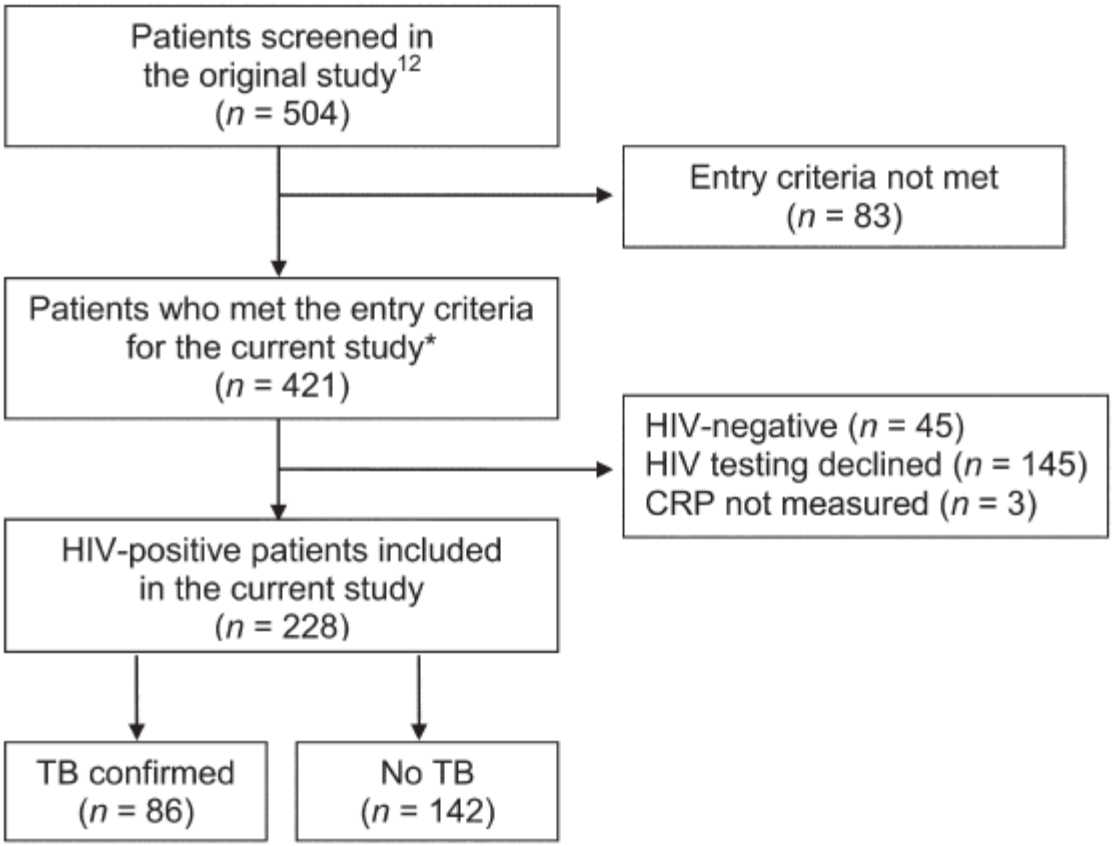


Figure 1 Participant flow chart. *364 patients were included in the original study: 12 persons who died before 8 weeks and 42 persons who did not complete 8 weeks were excluded.[12] The current analysis includes these patients as the data obtained at the time of presentation were available. HIV = human immunodeficiency virus; CRP = C-reactive protein; TB = tuberculosis.

Approach to variable inclusion in the final model

- Initial set of variables were chosen using the literature and clinical experience of our group
- We chose 8 variables to maintain power of the model
- (88 cases of confirmed TB, 1:10 ratio)

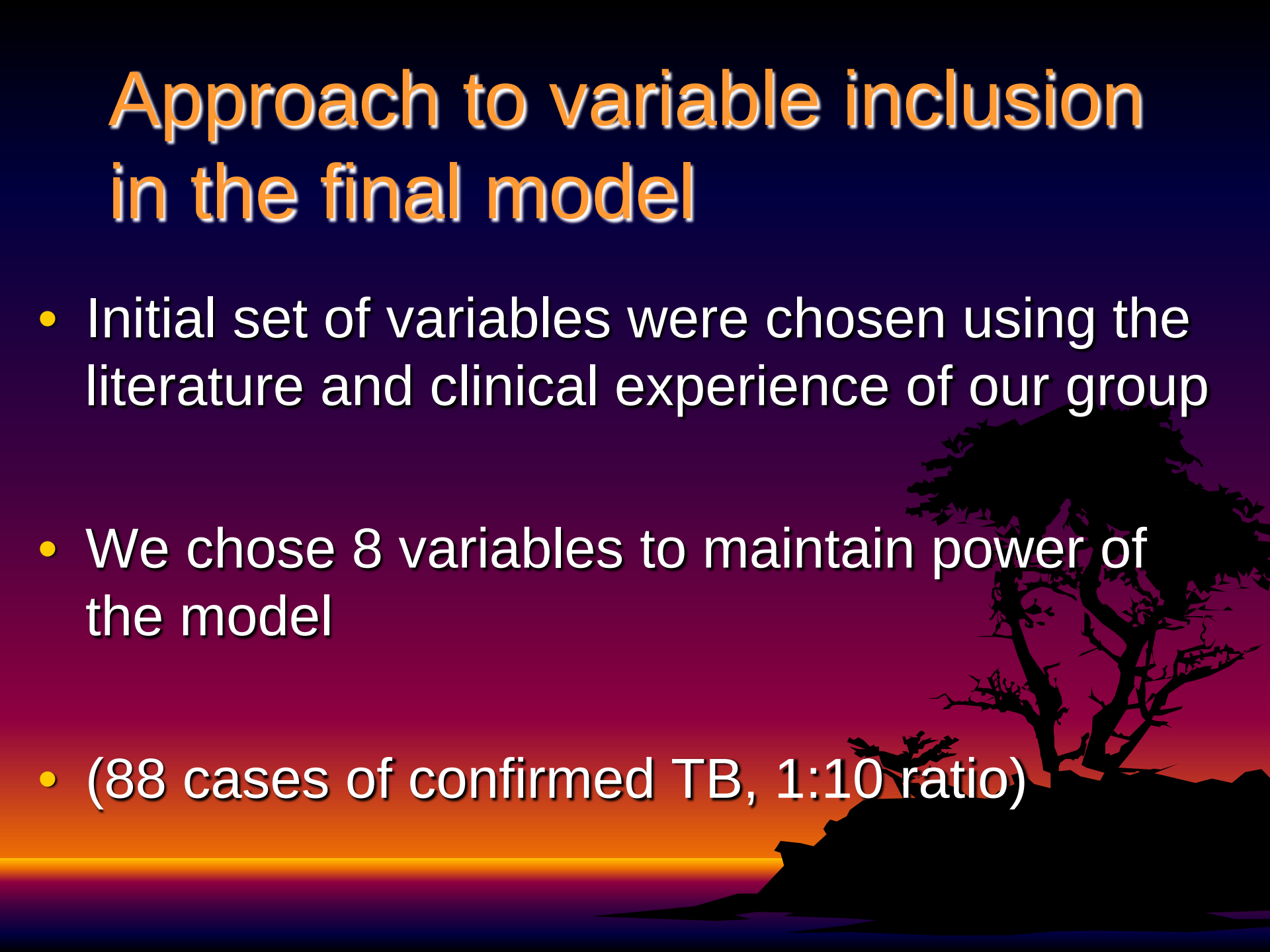


Table 1 Characteristics of human immunodeficiency virus positive patients who were TB suspects with smear-negative microscopy at presentation ($N = 228$)

	Total ($N = 228$) *n* (%)	TB ($n = 86$) *n* (%)	No TB ($n = 142$) *n* (%)	*P* value
Age, years, mean ± SD	34.3 ± 7.5	33.4 ± 6.8	34.8 ± 7.9	0.181
Men	98 (45)	37 (45)	61 (45)	0.99
Cough >2 weeks	210 (92)	78 (91)	132 (93)	0.539
Weight loss	189 (83)	77 (90)	112 (79)	0.038
Drenching night sweats	162 (71)	76 (88)	86 (61)	<0.0001
Haemoglobin, g/dl, mean ± SD	10.4 ± 2.1	9.5 ± 2.1	10.98 ± 1.9	<0.0001
Temperature >38°C	26 (11)	18 (21)	8 (6)	0.0004
BMI, kg/m^2, mean ± SD	21.9 ± 4.6	20.9 ± 4.2	22.5 ± 4.7	0.015

TB = tuberculosis; SD = standard deviation; BMI = body mass index.

Median CD4 count of 137 cells per uL in a sample of 150 patients from the cohort

Bayesian Information Criterion

- Several logistic regression models were fitted with variables selected according to the BIC
- The top 5 models with the highest posterior probabilities were used to pick the variables in the final model
- Any variable that appeared at least once was used in the final model (5 out of 8 variables were retained in final model)

Table 2 Multivariable model derived using variables known to be clinically relevant with CRP and without CRP

	Baseline model without CRP aOR (95%CI)	Baseline model with CRP aOR (95%CI)
Male	1.08 (0.57–2.06)	0.87 (0.43–1.75)
Night sweats	3.95 (1.81–8.65)	4.78 (2.07–11.03)
Temperature >38°C	2.37 (1.12–5.00)	1.27 (0.58–2.77)
BMI ⩽18.5 kg/m^2	1.38 (0.67–2.82)	1.20 (0.55–2.66)
Haemoglobin ⩽12 g/dl	4.84 (1.75–13.75)	1.93 (0.59–6.29)
CRP ⩾11 mg/dl	—	19.32 (5.45–68.9)
BIC	278	248
AUC	0.75 (0.68–0.81)	0.81 (0.76–0.87)*

*Significant difference (P = 0.0014) between the model with CRP and the model without CRP.

CRP = C-reactive protein; aOR = adjusted odds ratio; CI = confidence interval; BMI = body mass index; BIC = Bayesian Information Criterion; AUC = area under the curve.

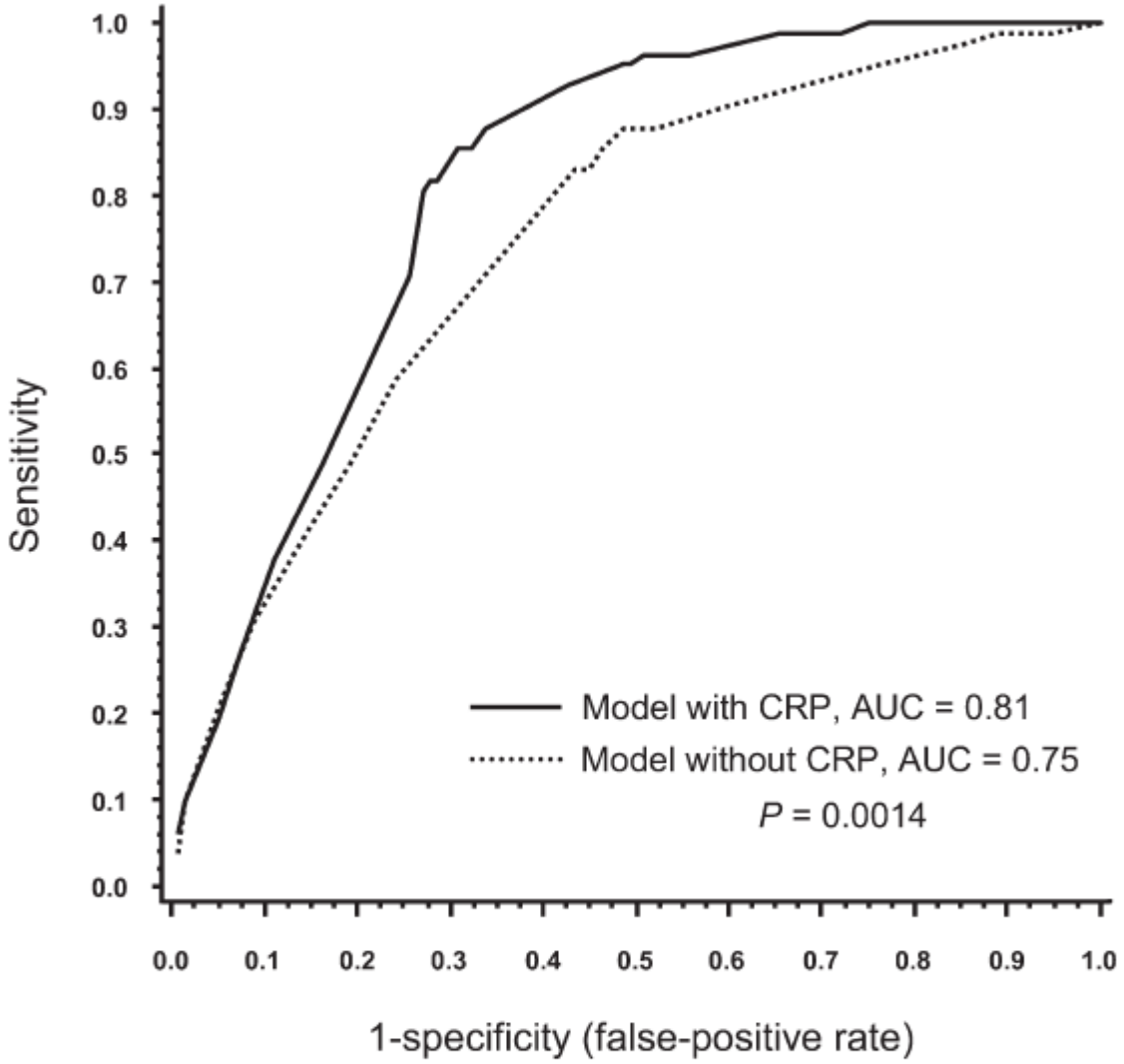


Figure 2 Prediction model ROC curve with and without CRP to predict SNTB among HIV patients. CRP = C-reactive protein; AUC = area under the curve; ROC = receiver operating characteristic; SNTB = smear-negative tuberculosis; HIV = human immunodeficiency virus.

Net reclassification tables

- Quantifies the net movement of patients between pre-specified risk categories
- Net reclassification index is the net proportion of patients that were reclassified correctly by the addition of CRP
- It predicts how well CRP, when added to the model, will shift the patients to lower risk categories for patients who did not have TB

Pencina et al 2007 Statistical Med

Table 3 Risk reclassification following the incorporation of CRP into the baseline model

Model without CRP (baseline model)	Model with CRP <5% risk	5–20% risk	⩾20% risk	Total *n*	Percentage appropriately reclassified*
Patients with no TB (*n* = 136)					
<5% risk	7	5	0	12	−42
5–20% risk	32	2	23	57	16
⩾20% risk	2	22	43	67	36
Total, *n*	41	29	66	136	

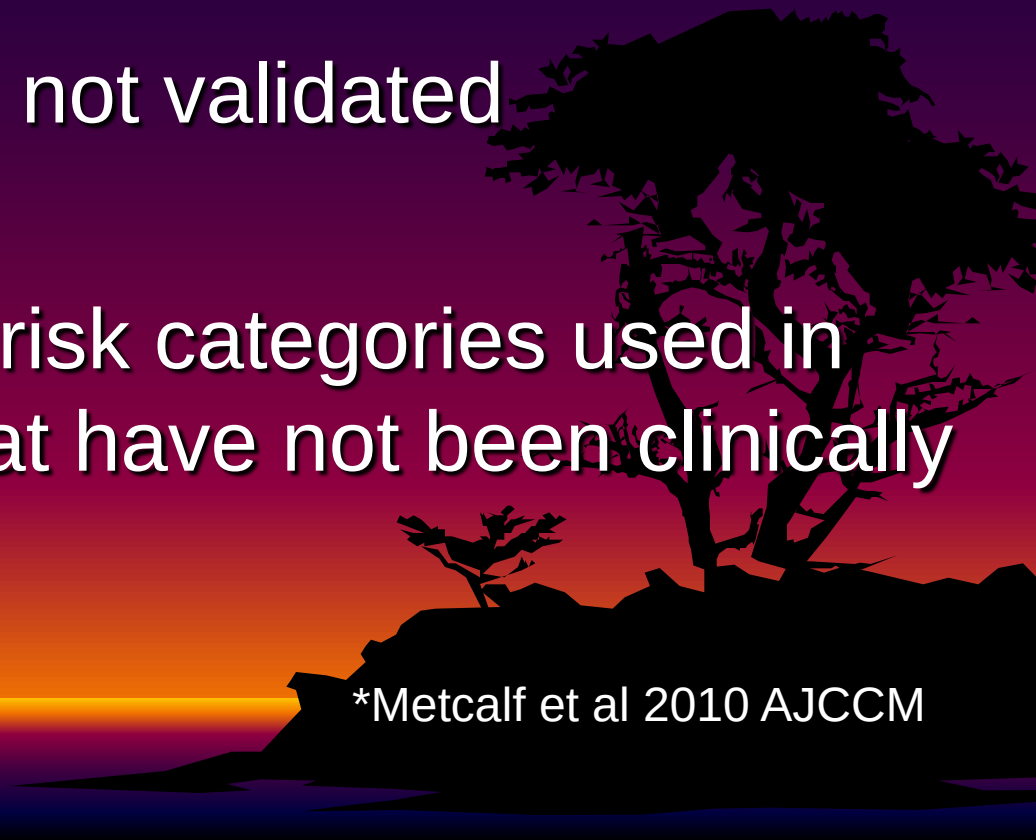
Incorporation of CRP to the model resulted in:

56 patients reclassified to lower risk categories
52 patients had no change in risk categories
28 patients reclassified to higher risk categories

The reclassification improvement among patients with **NO TB** was 20.6% (p=0.0023) when adding CRP to the clinical model (correctly reclassified into a lower risk category)

Limitations

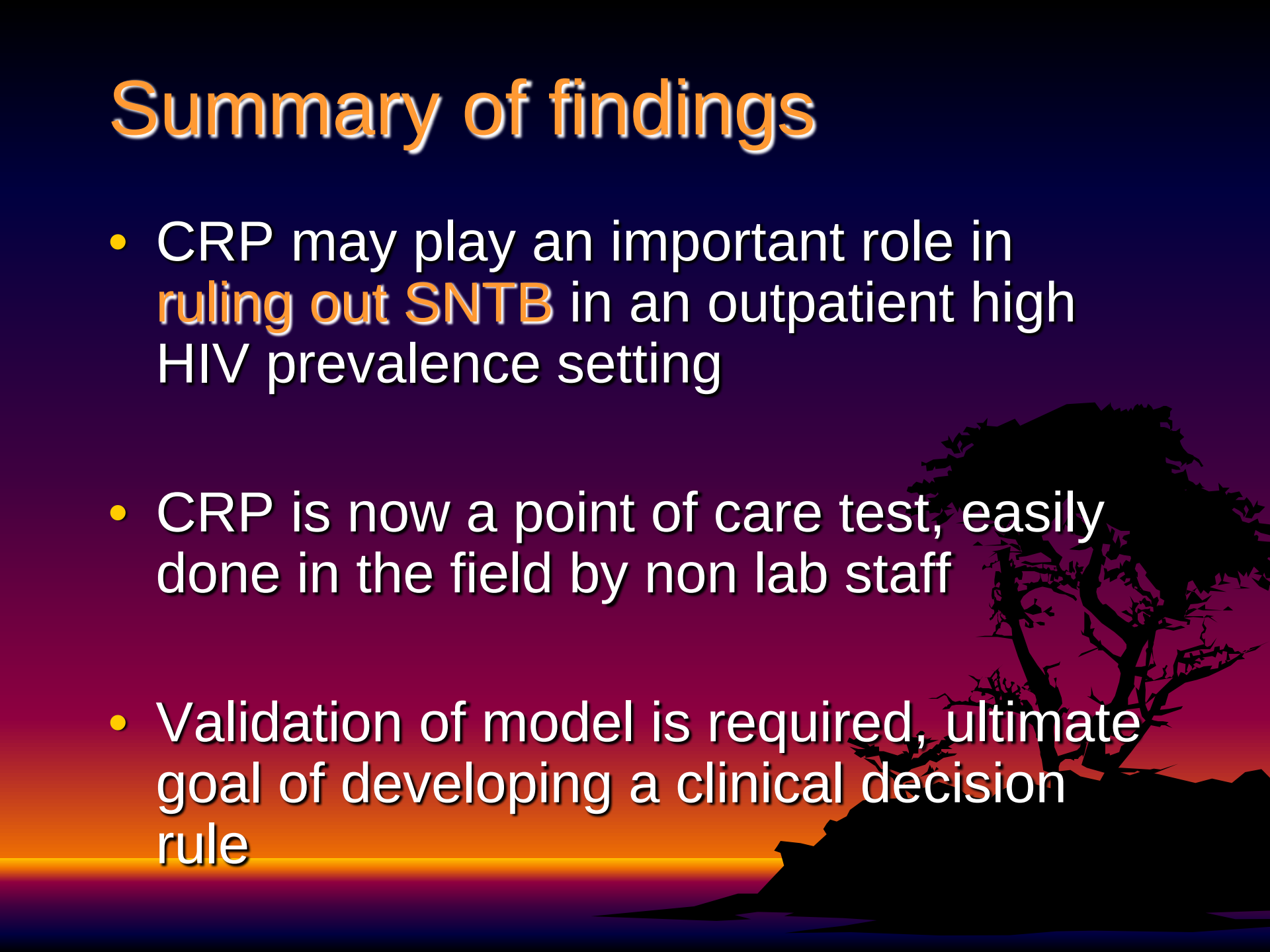
- Retrospective
- Model was derived not validated
- Arbitrary choice of risk categories used in previous paper* that have not been clinically validated



*Metcalf et al 2010 AJCCM

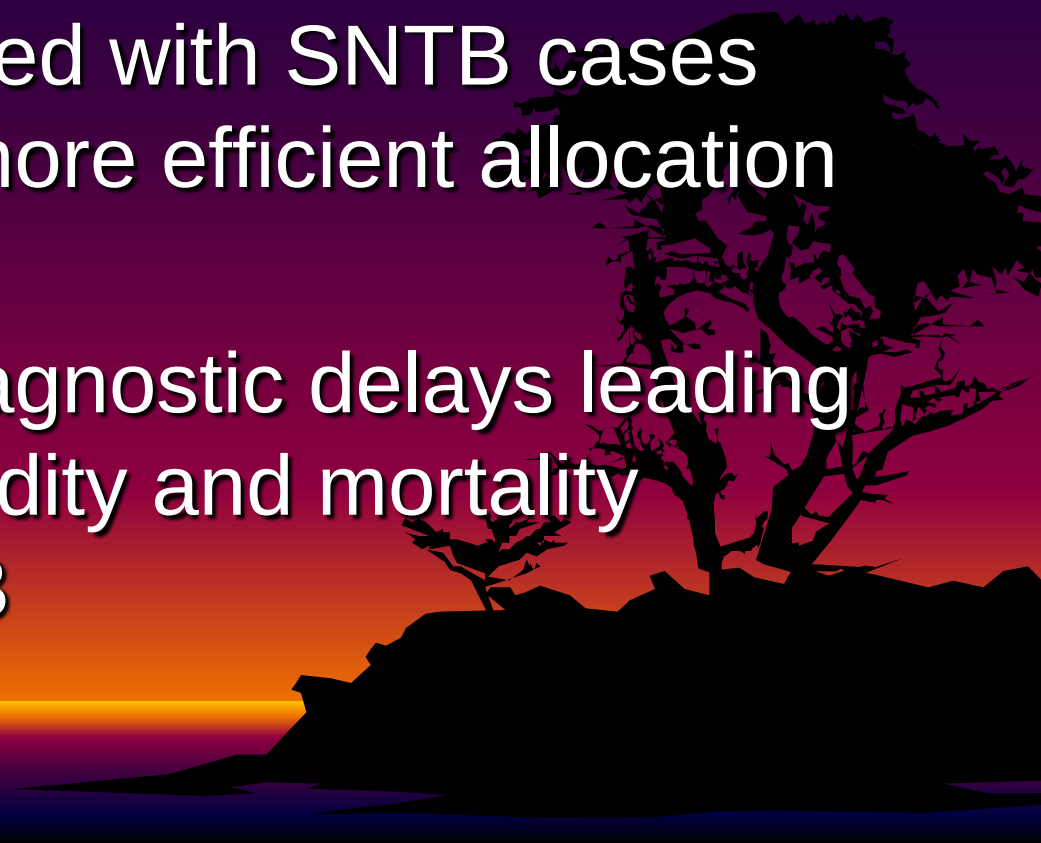
Summary of findings

- CRP may play an important role in ruling out SNTB in an outpatient high HIV prevalence setting
- CRP is now a point of care test, easily done in the field by non lab staff
- Validation of model is required, ultimate goal of developing a clinical decision rule



Implications

- Reducing the number of patients that require the costly and labor intensive work up associated with SNTB cases would allow for more efficient allocation of resources
- A decrease in diagnostic delays leading to reduced morbidity and mortality related to HIV TB



Thank you

- Doug Wilson
- Elham Sabri
- Steve Doucette
- Daphne Ling
- Bill Cameron
- Gary Maartens
- Nurses Zanele Mcabo, MJ Khumalo