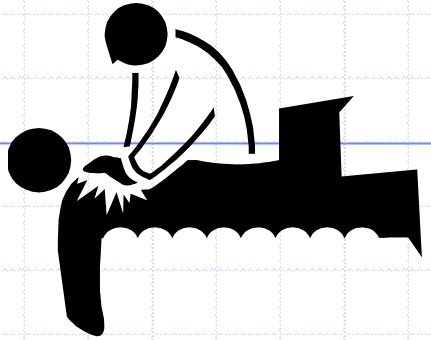
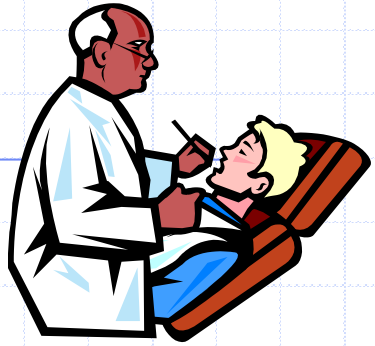
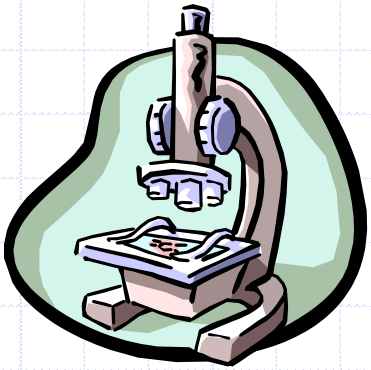
Diagnostic research: incremental value and multivariable approach







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

Email: madhukar.pai@mcgill.ca

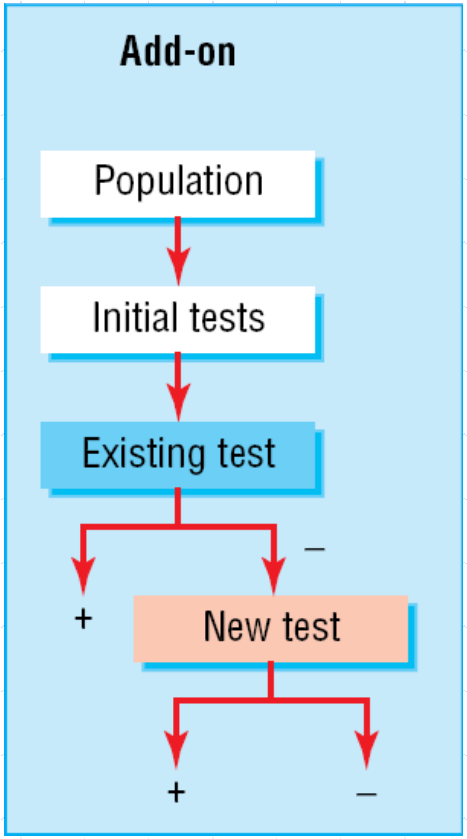
The diagnostic process is probabilistic, multivariable and sequential

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 household contacts?
- What is the added value of Xpert MTB/RIF in sputum smear-negative patients with HIV infection?
- What is the added value of IGRAs in diagnosing active TB, beyond smear microscopy?

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

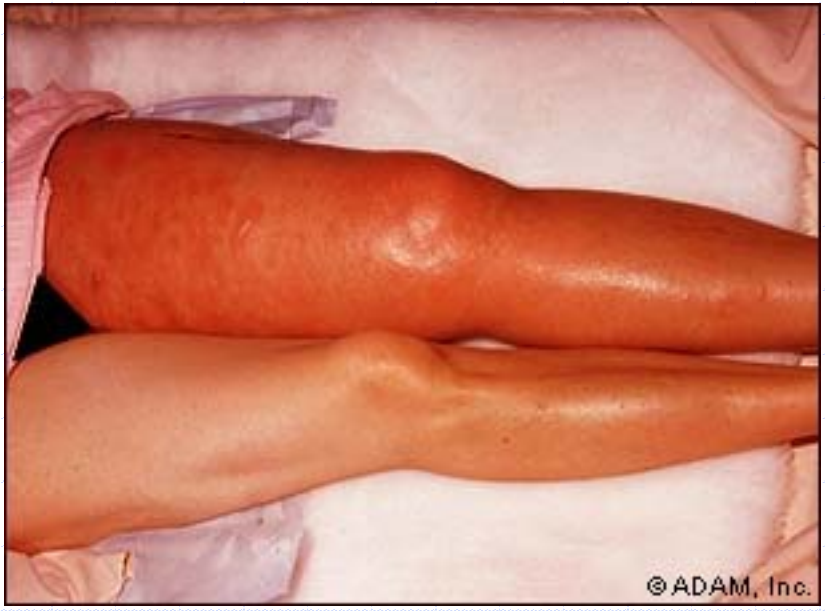
Multivariable process

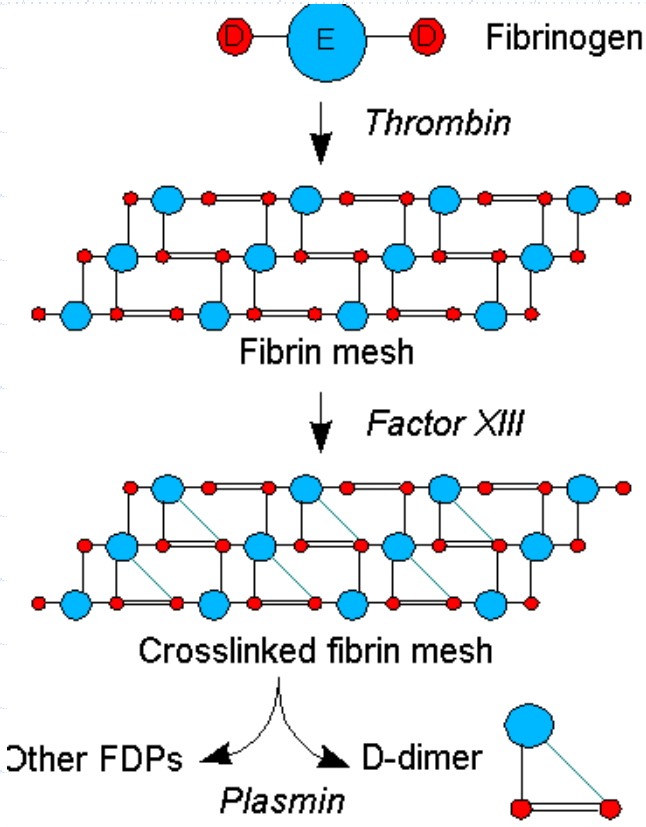
- Logistic regression model:

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

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

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

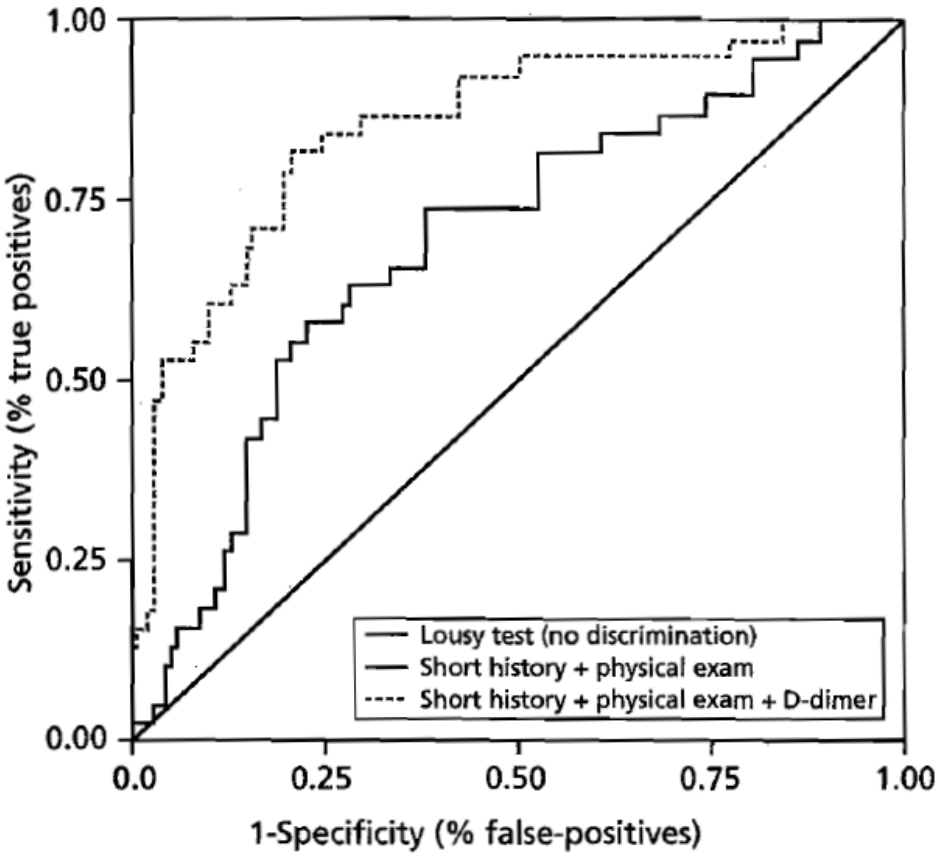


FIGURE 3.3 Example of an ROC curve of the reduced multivariable logistic regression model, including the same six determinants as in Figure 3.2. The ROC area of the "reduced history + physical model" (red) was 0.70 (95% confidence interval [CI], 0.66–0.74) and of the same model added with the D-dimer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

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

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.*

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

Limited added value of T-SPOT.*TB* blood test in diagnosing active TB: A prospective bayesian analysis

Aileen E. Boyd [a,*], Anthony Ashcroft [b], Marc Lipman [c], Graham H. Bothamley [d]

[a] *Department of Medicine, Royal Free Hospital, Pond Street, London NW3 2QG, UK*
[b] *Department of Microbiology, Homerton University Hospital, Homerton Row, London E9 6SR, UK*
[c] *Department of Respiratory Medicine, Royal Free Hospital, Pond Street, London NW3 2QG, UK*
[d] *Department of Respiratory Medicine, Homerton University Hospital, Homerton Row, London E9 6SR, UK*

Accepted 10 April 2011
Available online 15 April 2011

KEYWORDS
Bayes' theorem;
Diagnosis;
Interferon-gamma release assay;
Likelihood ratio;
Tuberculosis

Summary *Objectives:* To determine the diagnostic value of a blood interferon-gamma release assay in suspected active tuberculosis (TB).
Methods: 136 subjects with suspected pulmonary TB (pTB) at a single London centre with intermediate TB incidence, were clinically graded into low (<25%), medium, or high (>75%) likelihood of active pTB and then tested by T-SPOT®.TB assay. The diagnosis was confirmed by culture ($n = 33$), treatment response ($n = 13$) or a firm alternative diagnosis ($n = 90$).
Results: Overall, the T-SPOT.*TB* sensitivity was 74% (95% confidence intervals 60—84%), positive predictive value (PPV) 56% (43—68%), negative predictive value (NPV) 83% (71—90%), positive likelihood ratio (PLR) 1.75 and negative likelihood ratio (NLR) 0.45. Results for high pTB likelihood subjects: PPV 100%, NPV 25% (7—60%), PLR >69, NLR 0.31. Results for intermediate pTB likelihood subjects: PPV 67% (41—85%), NPV 88% (65—96%), PLR 2.39, NLR 0.26. Results for low pTB likelihood subjects: PPV 15% (6—34%), NPV 92% (79—97%), PLR 1.23, NLR 0.80. False negatives occurred in 24% of cases of active tuberculosis (4 smear and culture-positive, 3 smear negative and culture-positive, and 4 culture negative).
Conclusions: The predictive values and likelihood ratios show the T-SPOT.*TB* test does not assist in confidently confirming or excluding active TB, regardless of the pre-test probability of disease.

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

ERJ Express. Published on June 30, 2011 as doi: 10.1183/09031936.00058911

No discriminatory value of interferon release added to smear negative HIV-tuberculosis algorithms

Molebogeng X Rangaka[1,2,3], Hannah P Gideon[2], Katalin A Wilkinson[4,2,5], Madhukar Pai[6], Judith Mwansa-Kambafilwe[2], Gary Maartens[7,5], Judith R Glynn[3], Andrew Boulle[1], Katherine Fielding[3], Rene Goliath[2], Raylene Titus[2], Shahied Mathee[8], Robert J Wilkinson[9,2,4,5]

[1]Centre for Infectious Disease and Epidemiology Research (CIDER), [2]Clinical Infectious Diseases Research Initiative, Institute of Infectious Diseases and Molecular Medicine,[5] Department of Medicine, [7] Division of Clinical Pharmacology, University of Cape Town, Observatory 7925, South Africa

[3] London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

[4] MRC National Institute for Medical Research, Mill Hill, London, NW7 1AA, UK

[6] McGill University and Montreal Chest Institute, Montreal, Quebec, Canada

[8] Provincial Administration of the Western Cape, Khayelitsha 7784, South Africa

[9] Division of Medicine, Imperial College London, Norfolk Place, W2 1PG, UK

Correspondence and requests for reprints should be addressed to Molebogeng Rangaka, Centre for Infectious Disease and Epidemiology Research (CIDER), School of Health Sciences University of Cape Town, Observatory, 7925 South Africa Tel: +27 (0)21 406 6389 Fax: +27 (0)21 406 6796 Email: mxrangaka@yahoo.co.uk

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 or NRI
    - 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