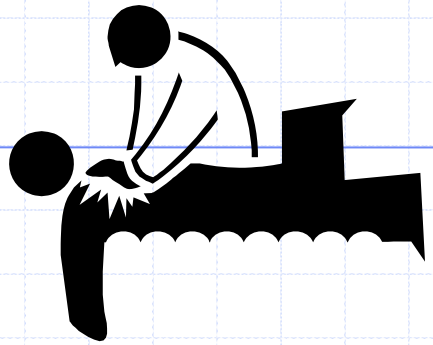
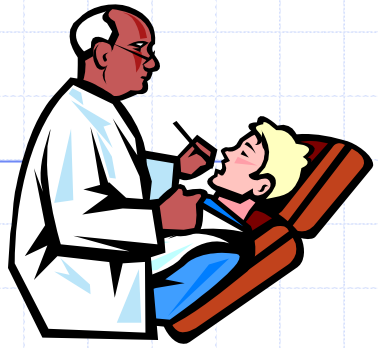
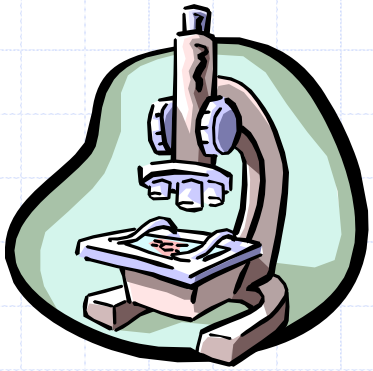
Diagnostic research: multivariable approach







Madhukar Pai, MD, PhD
Assistant Professor of Epidemiology, McGill University
Montreal, Canada
Professor Extraordinary, Stellenbosch University, S Africa

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

Sensitivity and specificity are a good starting point, but not enough

Table 1. **Hierarchy of Diagnostic Evaluation and the Number of Studies Available for Different Levels of Diagnostic Test in a Technology Assessment of Magnetic Resonance Spectroscopy for Brain Tumors***

Level	Description	Examples of Study Purpose or Measures	Studies Available, *n*	Patients, *n*
1	Technical feasibility and optimization	Ability to produce consistent spectra	85	2434
2	Diagnostic accuracy	Sensitivity and specificity	8	461
3	Diagnostic thinking impact	Percentage of times clinicians' subjective assessment of diagnostic probabilities changed after the test	2	32
4	Therapeutic choice impact	Percentage of times therapy planned before MRS changed after the test	2	105
5	Patient outcome impact	Percentage of patients who improved with MRS diagnosis compared with those without MRS (e.g., survival, quality of life)	0	0
6	Societal impact	Cost-effectiveness analysis (e.g., use to detect tumor in asymptomatic population)	0	0

* MRS = magnetic resonance spectroscopy.

Tatsioni. Annals

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies

The GRADE system can be used to grade the quality of evidence and strength of recommendations for diagnostic tests or strategies. This article explains how patient-important outcomes are taken into account in this process

SUMMARY POINTS

As for other interventions, the GRADE approach to grading the quality of evidence and strength of recommendations for diagnostic tests or strategies provides a comprehensive and transparent approach for developing recommendations

Cross sectional or cohort studies can provide high quality evidence of test accuracy

However, test accuracy is a surrogate for patient-important outcomes, so such studies often provide low quality evidence for recommendations about diagnostic tests, even when the studies do not have serious limitations

Inferring from data on accuracy that a diagnostic test or strategy improves patient-important outcomes will require the availability of effective treatment, reduction of test related adverse effects or anxiety, or improvement of patients' wellbeing from prognostic information

Judgments are thus needed to assess the directness of test results in relation to consequences of diagnostic recommendations that are important to patients

BMJ 2008

Redundancy of Single Diagnostic Test Evaluation

Karel G.M. Moons,[1,2,3] Gerri-Anne van Es,[4] Bowine C. Michel,[5] Harry R. Büller,[6] J. Dik F. Habbema,[3] and Diederick E. Grobbee[1]

Moons et al. Epidemiology 1999

Diagnostic research

Diagnostic studies as multivariable, prediction research

K G M Moons, D E Grobbee

Patient outcomes in diagnostic research

Moons et al. JECH 2002

Opinion

Test Research versus Diagnostic Research

Moons et al. Clin Chem 2004

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

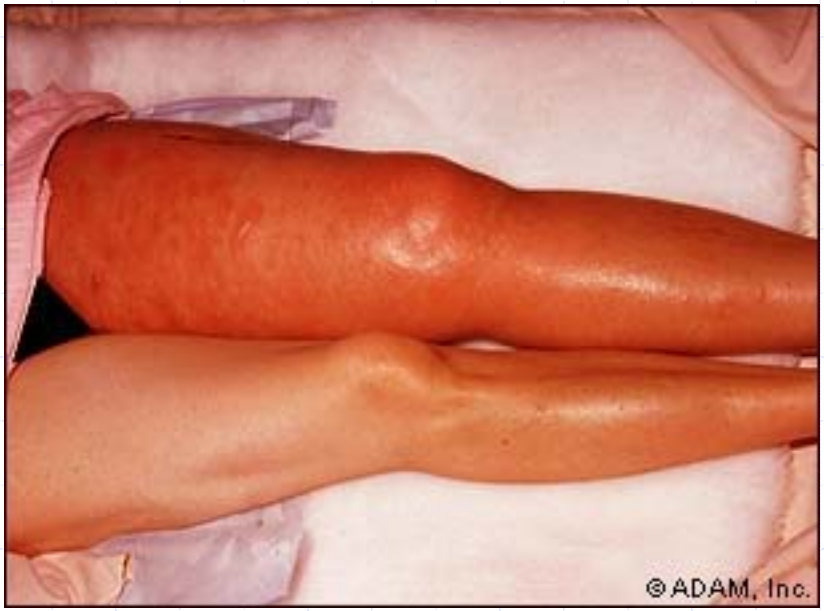
Multivariable process

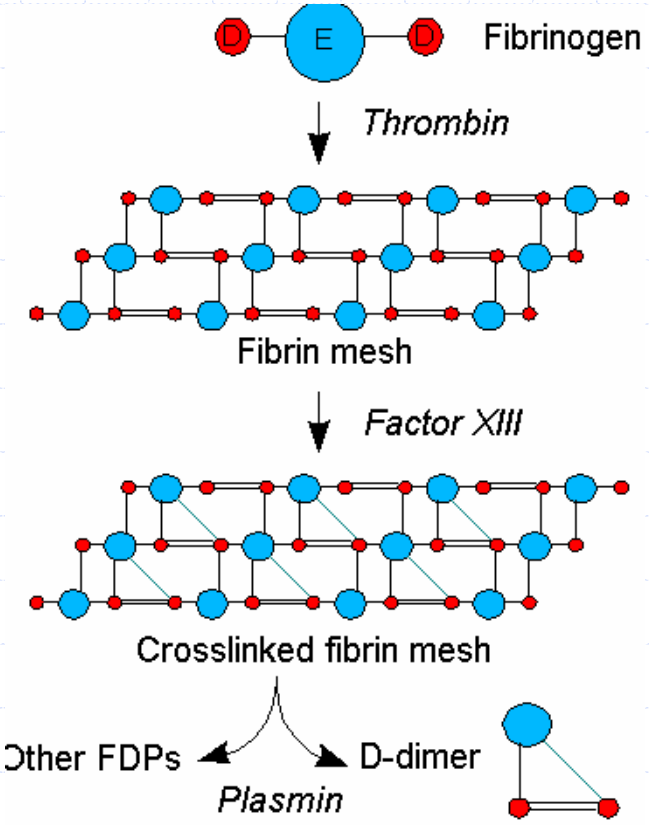
- Logistic regression model:

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

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

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 we identified 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 $\geq$ 3 cm difference in circumference of the calves.
- For women two additional predictors were documented, i.e. the use of oral hormonal contraception and of estrogen replacement therapy.

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, we 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, we 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

Data analysis

- Bootstrapping techniques, repeating the entire modelling process were used to validate the final model and to adjust the estimated performance and regression coefficients (odds ratios) for over-fitting [shrinkage process]
- To construct an easily applicable diagnostic rule, the regression coefficients of the variables were transformed to integers according to their relative contributions to the risk estimation.
- Finally, after estimating the score for each patient, we estimated the absolute percentages of correctly diagnosed patients across score categories.

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	78	99	72	38.2 (9.40 – 155.3)
Tinaquant n= 377	65	98	54	37.3 (9.00 – 154.8)
Combined assays	74	99	66	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

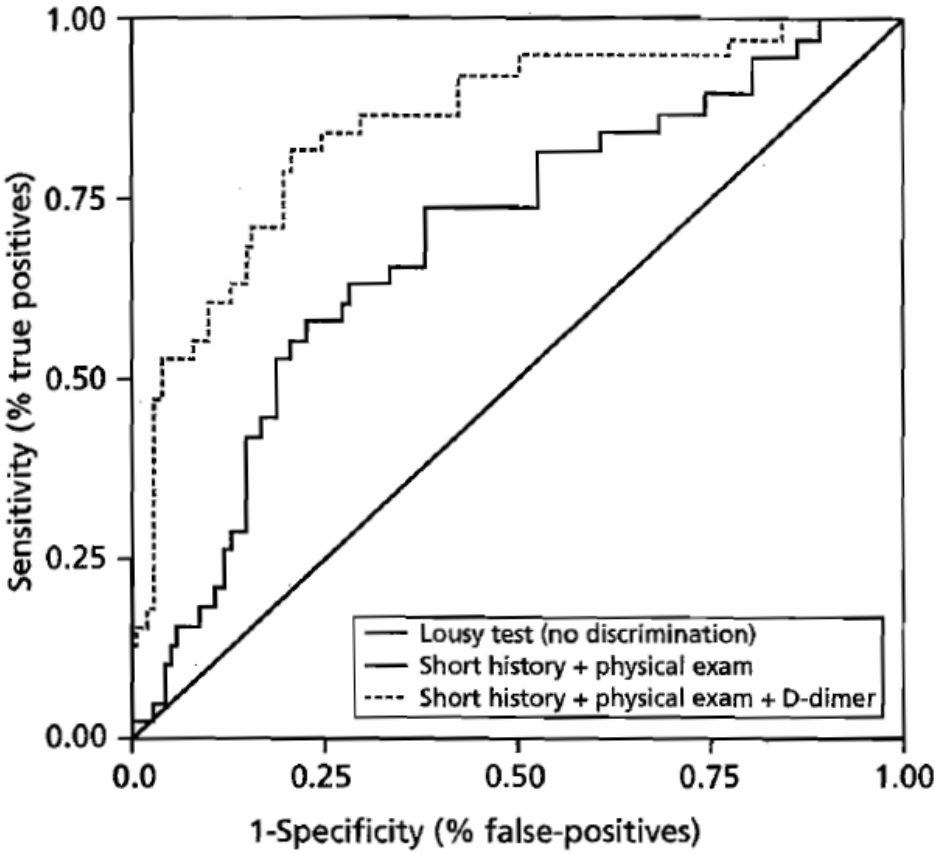


FIGURE 3.3 Example of an ROC curve of the reduced multivariable logistic regression model, including the same six determinants as in Figure 3.2. The ROC area of the "reduced history + physical model" (red) was 0.70 (95% confidence interval [CI], 0.66–0.74) and of the same model added with the D-dimer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

Another example

Acta Pædiatr 90: 611–617. 2001

Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures

R Oostenbrink[1,2], KGM Moons[1,2], ART Donders[2,3], DE Grobbee[2] and HA Moll[1]

Outpatient Department of Paediatrics[1], Sophia Children's Hospital, Rotterdam; Julius Centre for Patient Oriented Research[2], University Medical Centre, Utrecht; Centre for Biostatistics[3], University of Utrecht, The Netherlands

Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Pædiatr 2001; 90: 611–617. Stockholm. ISSN 0803-5253

Physicians often have to perform a lumbar puncture to ascertain the diagnosis in patients with meningeal signs, because of the serious consequences of missing bacterial meningitis. The aim of this study was to derive and validate a clinical rule to predict bacterial meningitis in children with meningeal signs, to guide decisions on the performance of lumbar punctures. Information was collected from records of patients (aged 1 mo to 15 y) consulting the emergency department of the Sophia Children's Hospital between 1988 and 1998 with meningeal signs. Bacterial meningitis was defined as cerebrospinal fluid (CSF) leucocyte count >5 cells μl^{-1} with a positive bacterial culture of CSF or blood. The diagnostic value of predictors was judged using multivariate logistic modelling and area under the receiver operating characteristic curves (ROC area). In the derivation set (286 patients, years 1988–1995) the duration of the main complaint, vomiting, meningeal irritation, cyanosis, petechiae and disturbed consciousness were independent clinical predictors of bacterial meningitis. The ROC area of this model was 0.92. The only independent predictor from subsequent laboratory tests was the serum C-reactive protein concentration, increasing the ROC area to 0.95. Without missing a single case, this final model identified 99 patients (35%) without bacterial meningitis. Validation on 74 consecutive patients in 3 subsequent years (1996–1998) yielded similar results.

Conclusion: This prediction rule identifies about 35% of the patients with meningeal signs in whom a lumbar puncture can be withheld without missing a single case of bacterial meningitis. For the individual patient this prediction rule is valuable in deciding whether or not to perform a lumbar puncture.

Table 3. Independent predictors for bacterial meningitis

Variable	Clinical evaluation model OR (95% CI)	Clinical evaluation + laboratory model OR (95% CI)	Risk score
Patient history			
Duration of the main complaint (per day)[a]	1.5 (1.2–1.9)	1.5 (1.2–1.9)	1
Vomiting	2.4 (1.0–5.4)	2.3 (0.9–5.5)	2
Physical examination			
Meningeal irritation	25.0 (3.2–197.5)	21.1 (2.6–172.4)	7.5
Cyanosis	24.0 (2.0–289.4)	13.0 (1.1–151.3)	6.5
Petechiae or ecchymoses	7.5 (2.2–25.6)	4.9 (1.4–17.9)	4
Disturbed consciousness	22.2 (9.4–52.4)	21.8 (8.6–55.2)	8
Laboratory tests			
Serum CRP (per 10 mg l^{-1})[b]		1.1 (1.0–1.1)	0.1
ROC area (95% CI) in derivation set	0.92 (0.89–0.95)	0.95 (0.92–0.97)	0.94 (0.91–0.97)
ROC area (95% CI) in validation set	0.92 (0.86–0.98)	0.92 (0.86–0.98)	0.92 (0.86–0.98)

[a] Duration of the main complaint rounded off to half days, with a maximum of 7 points.
[b] Points assigned to serum CRP: 0.1 point per 10 mg l^{-1} increase, thus 0–9 mg l^{-1}: 0 points; 10–19 mg l^{-1}: 0.1 points; etc., with a maximum of 2 points.
OR: odds ratio; CI: confidence interval; CRP: C-reactive protein; ROC: receiver operating characteristic.

$$\text{Total score} = 1 \times \text{duration main complaint (d)} + 2 \times \text{vomiting} + 7.5 \times \text{meningeal irritation} + 6.5 \times \text{cyanosis} + 4 \times \text{petechiae} + 8 \times \text{disturbed consciousness} + 0.1 \times \text{serum CRP (per 10 mg } l^{-1})$$

Table 4. Frequency of bacterial meningitis (BM) related to the risk score

	Derivation set ($n = 286$)		Validation set ($n = 74$)	
Risk score (points)	BM present	BM absent	BM present	BM absent
0–4.9	0	64 (100%)	0	20 (100%)
5.0–9.4	0	35 (100%)	0	14 (100%)
9.5–14.9	17 (16%)	88 (84%)	3 (15%)	17 (85%)
15.0–19.9	24 (63%)	14 (37%)	4 (44%)	5 (56%)
≥20.0	43 (98%)	1 (2%)	8 (73%)	3 (27%)

Bacterial meningitis was absent in all patients with a score <9.5 and present in almost all patients with a score >=20. The threshold value <9.5 identified 99 patients without bacterial meningitis (35%; 95% CI 29–40%), without missing a single case of bacterial meningitis. In patients with meningeal signs, a lumbar puncture can be withheld in 35% of cases without missing a single case of bacterial meningitis.

Demo:

D-dimer testing to rule out pulmonary venous embolism

Data from: Moons KGM et al (used with permission)