Chapter 4

Clinical prediction rules and other diagnostic modalities

Publication short communicated in this chapter:

Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Clinical Scoring for Diagnosis of Acute Lower Abdominal Pain in Female of Reproductive Age. Emergency Medicine International. 2013;2013:6.

Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Validation of the Diagnostic Score for Acute Lower Abdominal Pain in Women of Reproductive Age. Emergency Medicine International. 2014;2014:6

1. Clinical prediction rules

A Clinical prediction rule has been described as the process by which a combination of clinical findings that have statistically demonstrated meaningful predictability in determining a condition or prognosis of a patient who has been provided with a specific treatment.[1, 2]

The need for a clinical prediction rule arises in: 1) areas of crucial clinical uncertainty; 2) clinical scenarios where improper intervention may result in unfavorable outcomes, and 3) patient screening for under-diagnosed conditions.[3] With these criteria and their applications, a clinical prediction rule is valuable in the specific diagnostic possibilities associated with acute lower abdominal pain in women of reproductive age.

The most well-known and widely accepted clinical prediction rule for acute lower abdominal pain is Alvarado's score for appendicitis.[4] It was developed from a review of the medical records of 305 hospitalized patients who suffered with abdominal pain. The clinical findings were evaluated for their positive and negative predictive values, and the diagnostic weight was calculated from them. Item scores were assigned for clinical findings with high diagnostic weight. Alvarado's score has been validated by several studies with a proven track record of being a reliable risk stratifier for appendicitis.[5-8] The score and its interpretation are show in table 4.1 and 4.2.

Table 4.1 Alvarado' score

Characteristics	Score
M= Migratory pain to RLQ	1
A = Anorexia	1
N = Nausea and vomiting	1
T = Tenderness at RLQ	2
R = Rebound tenderness	1
E = Elevated temperature > 37.3° C	1
L = Leukocytosis (wbc > 10,000)	2
S = Shift of wbc to the left (N>75%)	1
Total	10

Table 4.2 Interpretation of Alvarado's score

Levels	Interpretations
Score 5 – 6	Compatible with appendicitis
Score 7 – 8	Probable appendicitis
Score 9 – 10	Very probable appendicitis

(**Table 4.1** & **Table 4.2** are adapted from *Alvarado A: Practical score of the early diagnosis of acute appendicitis* [4]*. Ann Emerg 1986;15:557-564*)

The value of a clinical prediction rule in the diagnosis of acute abdominal pain has been evaluated. In one before-and-after trial[9] which incorporated a clinical prediction rule to improve the accuracy of the diagnosis of appendicitis, the results were disappointing. The diagnostic success of the final examiner decreased when they used the score (specificity, 86% vs 78%; positive predictive value, 67% vs 50%; and accuracy, 88% vs 81%). Therefore, it is not advised to use the clinical prediction rule alone for the diagnosis of appendicitis.

Despite the discouraging results when using a clinical prediction rule alone in attempts to improve accuracy of diagnosis, clinical prediction rules have their value in risk stratification for informing further investigation and clarifying criteria for the admission of patients. Studies have shown improved diagnostic yield when Alvarado's score has been used for admission criteria particularly when it has been used in conjunction with an ultrasound or CT scan.[5, 8, 10]

There are few clinical prediction rules available for the diagnosis of gynecological conditions.[11-13] Diagnoses of gynecological conditions are mostly established by clinical evaluation, ultrasound or laparoscopy.

2. A diagnostic prediction rule for acute lower abdominal pain in women of reproductive age

The development of a clinical diagnostic prediction rule for acute lower abdominal pain in women of reproductive age is more complicated than other conditions because more possibilities of diseases or condition need to be diagnosed. It is difficult to obtain a prediction model for this patient group from simple logistic regression which can handle only binary outcomes. Polytomous logistic regression may be used more on multiple outcomes in the case of diagnostic research.[14, 15] In the development of a diagnostic score for acute lower abdominal pain in women of reproductive age, the score was developed for a diagnostic predictive model using polytomous logistic regression to predict whether the patient would be suffering from appendicitis, obstetric and gynecological conditions (OB-GYNc) or non-specific abdominal pain. This diagnostic prediction rule has been designed for use by emergency room physicians and doctors in rural hospitals to inform the selective management of patients.

The derivative phase of the study to identify the diagnostic score for acute lower abdominal pain in women of reproductive age was conducted at Nakornping Hospital, a tertiary care hospital in Chiang Mai, Thailand.[16] Data were extracted from the medical records of 542 women of 15 to 50 years old who were admitted to the hospital between January and December 2008 with acute lower abdominal pain and were diagnosed on discharge with either appendicitis, OB-GYNc or non-specific abdominal pain. Data were analyzed with a stepwise polytomous logistic regression to identify significant diagnostic indicators.

The regression coefficients from the final stepwise polytomous regression model for appendicitis and OB-GYNc are shown in Table 4.3. Predictors that were significant for the diagnosis of appendicitis or OB-GYNc were: guarding or rebound tenderness, pregnancy, leukocytosis, neutrophil levels ≥ 75%, locations of tenderness and presence of diarrhea. A significant constant needed to be included in the model because it was necessary to enable the comparison of the probabilities of the two diseases in the final model. Regression coefficients were known as linear predictors because they can be summed-up for the prediction of the probability of the disease. For the development of the diagnostic prediction score, regression coefficients were rounded into itemised scores to make the prediction score more straightforward to calculate. Item scores have positive and negative values according to their increase or decrease of the probabilities of the diseases, for example, pregnancy decreases the probability of appendicitis (item score -1.7 for appendicitis) while it increases the probability of OB-GYNc (item score +2.4 for OB-GYN). Based on the concept of relative probabilities of the polytomous logistic regression, criteria for the diagnostic prediction of appendicitis or OB-GYNc were developed. The criteria for diagnostic prediction are basically comparing the sum of item scores for appendicitis with the sum of item scores for OB-GYNc in each patient (Table 4.4).

The diagnostic performance of the diagnostic prediction score was evaluated with an apparent validation (see Appendix C). The score correctly predicted 249 out of 279 patients with appendicitis (87.37%), 46 out of 69 OB-GYNc patients (73.02%), the overall accuracy in correct diagnosis being 79.20%. In comparison with Alvarado's score for appendicitis, the new diagnostic prediction score had a slightly better discriminative ability (Figure 4.1). The area under the receiver operating characteristic curve (ROC), which indicates discriminative ability, was 0.8257 (discrimination ability of 82.57 %) for the new diagnostic prediction score, while the ROC of Alvarado's score was 0.8095 (discrimination ability of 80.95 %).

A temporal type of external validation study was performed retrospectively using patients who were admitted between January and July 2009.[17] Patients' characteristics in the validation cohort were similar to those of the derivative cohort. The researchers compared a predicted diagnosis from the diagnostic prediction score with the final diagnosis in the medical record of each patient in the validation cohort. The study found that the diagnostic prediction score had sensitivities of 91.9% and 73.0%, specificities of 79.0% and 91.6% and areas under ROC of 0.85 and 0.82, in diagnosis for appendicitis and OB-GYNc, respectively. The overall correct classification was 83.1%. If we consider that appendicitis is more serious condition, triaging patients with this score would results in 'under-diagnosis' in 8.1% (16/197). Therefore; close follow-up of a patient is still needed in NSAP and OB-GYNc groups. More details of the diagnostic indices scores were shown in Table 4.5.

The diagnostic prediction score can be separately applied to each diagnosis, based on the appendicitis score and OB-GYN score (see table 4.3& 4.4). As the appendicitis section and OB-GYN section of the score were related to the diagnostic probabilities of the respective diagnosis, the appendicitis score and OB-GYN score can therefore be applied separately. The predicted probability of each score level can be plotted as shown in Figure 4.2. The scores predicted as probability curves in Figure 4.2 implied that an OB-GYNc can be fairly confidently be ruled out if the patient's OB-GYN score approached 0. Similarly, the diagnosis of appendicitis is very unlikely if the patient's appendicitis score is less than -2. In cases of intermediate probability levels (appendicitis score 0-1 and OB-GYN score 2-3), physicians can advise further investigations such as ultrasonography or CT scans prior to making a final diagnosis and recommending treatment.

Table 4.3 Coefficients (95% confidence interval: CI) and assigned item scores of selected predictors for diagnosis of appendicitis, or common obstetric and gynecological conditions (OB-GYNc), from polynomial logistic regression analysis.

Predictors	Appendicitis Coefficients (95%CI)	p-value	OB-GYNc Coefficients (95%CI)	p-value	Assigned score: Appendicitis score	Assigned score: OB-GYN score
Guarding/ rebound tenderness	1.85 (1.12, 2.59)	<0.001	0.40 (-0.54, 1.34)	0.407	1.9	0
Pregnancy	-1.70 (-3.28, -0.12)	0.035	2.39 (1.05, 3.73)	<0.001	-1.7	2.4
Leukocytosis	1.53 (0.78, 2.29)	<0.001	-0.13 (-1.11, 0.84)	0.787	1.5	0
Neutrophil ≥75%	1.25 (0.35, 2.15)	0.007	1.61 (0.49, 2.73)	0.005	1.3	1.6
RLQ tenderness	1.52 (0.40, 2.64)	0.008	-0.42 (-1.46, 0.62)	0.429	1.5	0
LLQ tenderness	-1.11 (-2.26, 0.05)	0.062	1.93 (0.87, 2.98)	<0.001	0	1.9
Diarrhea	-1.44 (-2.41, -0.48)	0.003	-2.26 (-3.79, -0.74)	0.004	-1.4	-2.3
Constant	-1.45 (-2.61, -0.30)	0.014	-0.57 (-1.63, 0.49)	0.290	-1.5	0

Table 4.4 The scoring scheme for appendicitis and obstetric-gynecological conditions (OB-GYNc), and the criteria used to diagnose appendicitis, abdominal pain caused by obstetric and gynecological conditions (OB-GYNc) or non-specific abdominal pain (NSAP).

Predictors	Assigned score		Suggested diagnoses	Criteria
	Appendicitis score	OB-GYN score		
Guarding or rebound tenderness	1.9	0	Appendicitis	Sum of appendicitis score > sum of OB-GYN score <u>and</u> appendicitis score >0
Pregnancy	-1.7	2.4		
Leukocytosis (WBC ≥10,000/μL)	1.5	0	OB-GYNc	Sum of OB-GYN score ≥ sum of appendicitis score <u>and</u> sum of OB-GYN score >0
Neutrophil ≥75%	1.3	1.6		
RLQ tenderness	1.5	0	NSAP	Sum of appendicitis score ≤0 <u>and</u> sum of OB-GYN score ≤0
LLQ tenderness	0	1.9		
Diarrhea	-1.4	-2.3		
Constant	-1.5	0		

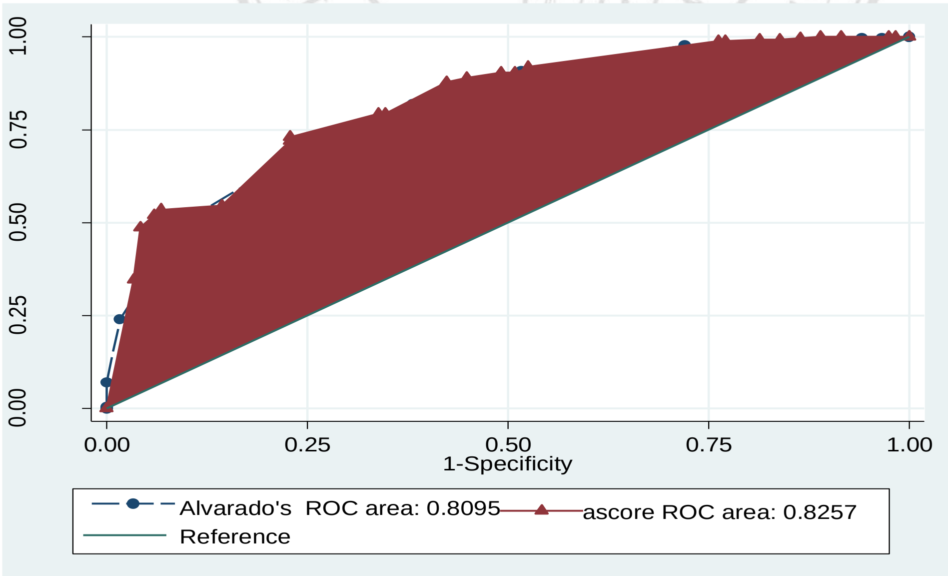

Figure 4.1 Receiver operating characteristic (ROC), curves of Alvarado's score (dash line) and appendicitis score (ascore, solid line) for diagnosis of appendicitis

(**Table 4.3** and **figure 4.1** were adapted from Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Clinical Scoring for Diagnosis of Acute Lower Abdominal Pain in Females of Reproductive Age. *Emergency Medicine International.* 2013;2013:6.)

Table 4.5 Diagnostic indices (and 95% confidence interval) of the scoring system for appendicitis (versus non-appendicitis), and OB-GYNc (versus non-OB- GYNc) in the validation data.

	Appendicitis (vs non-appendicitis)	**OB-GYNc (vs non-OB-GYNc)**
Sensitivity (%)	91.9 (87.1-95.3)	73.0 (60.3-83.4)
Specificity (%)	79.0 (70.0-86.4)	91.6 (87.4-94.8)
Receiver operating characteristic area	0.85 (0.81-0.90)	0.82 (0.77-0.88)
Likelihood ratio of positive	4.39 (3.02-6.37)	8.73 (5.59-13.62)
Likelihood ratio of negative	0.10 (0.06-0.17)	0.29 (0.20-0.44)
Positive predictive value (%)	89.2 (84.1-93.1)	69.7 (57.1-80.4)
Negative predictive value (%)	83.8 (75.1-90.5)	92.8 (88.7-95.7)

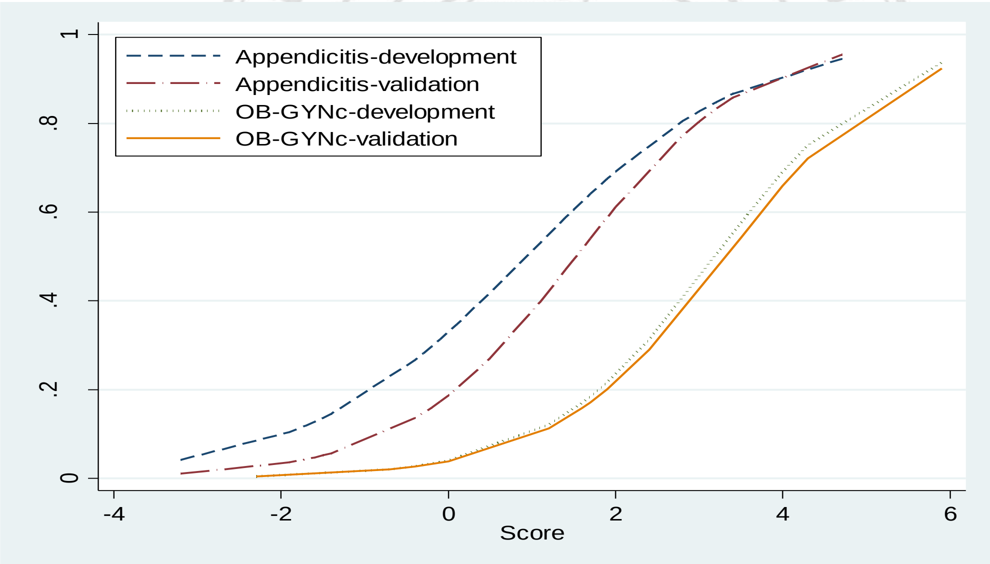

Figure 4.2 Score of predicted probabilities of appendicitis (dash and dash-dot lines) and OB-GYNc (dot and solid lines) in the development data set and the validation data set.

(**Table 4.4, 4.5** and **Figure 4.2** are from Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Validation of the Diagnostic Score for Acute Lower Abdominal Pain in Women of Reproductive Age. *Emergency Medicine International.* 2014;2014:6.)

3. Magnetic Resonance Imaging (MRI)

With no radiation exposure and a high level of accuracy, MRI can have a major role in the diagnosis of the cause of acute lower abdominal pain in pregnant women and children. Although it is not as straightforward as ultrasound the primary imaging modality, it can quickly establish a diagnosis of the cause of acute lower abdominal pain when a CT scan is unwarranted.[18] In pregnant and post partum women, an emergency MRI without oral contrast media has a sensitivity of 100% and 86%, and a positive predictive value of 100% and 100% for appendicitis and ovarian torsion, respectively.[19] It can be used in adjunct with ultrasound to evaluate acute lower abdominal pain in pregnant women facilitating a high degree of accuracy in diagnosis.[20]

4. Computer aided diagnosis and clinical decision tools

Since 1971, de Dombal et. al have demonstrated a higher diagnostic accuracy with computer assisted diagnosis of abdominal pain when compared with clinical evaluation alone.[21] In subsequent studies,[22, 23] they found that during the period of trial, diagnostic performance of doctors was better than the period prior; however, it reverted to previous levels shortly after the trial finished. When computer aided diagnosis techniques are used properly, they can reduce the mismanagement rate from 0.9% to 0.2%.[24]

The impact of evidence-based medicine enables the diagnosis of patients' conditions to be drawn into algorithms. These algorithms can be used as clinical decision tools, and they can be used in association with computer software to enable the diagnosis of several conditions. Widespread computer usage in medicine results in greater development of computerized clinical decision tools. A systematic review reported that clinical decision tools have a lower false-positive rate and a higher false-negative rate than when doctors are working unaided; and they are extremely useful in the diagnosis of appendicitis.[25] However, paper clinical decision tools, including clinical prediction rules, are more cost-effective than computerized clinical decision tools.

References

1. Randolph AG, Guyatt GH, Calvin JE, Doig G, Richardson WS. Understanding articles describing clinical prediction tools. Evidence Based Medicine in Critical Care Group. Crit Care Med. 1998;26(9):1603-12.

2. Cook CE. Potential pitfalls of clinical prediction rules. J Man Manip Ther. 2008;16(2):69-71.

3. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? Aust J Physiother. 2006;52(3):157-63.

4. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15(5):557-64.

5. Chan MY, Tan C, Chiu MT, Ng YY. Alvarado score: an admission criterion in patients with right iliac fossa pain. Surgeon. 2003;1(1):39-41.

6. Ohmann C, Yang Q, Franke C. Diagnostic scores for acute appendicitis. Abdominal Pain Study Group. Eur J Surg. 1995;161(4):273-81.

7. Pouget-Baudry Y, Mucci S, Eyssartier E, Guesdon-Portes A, Lada P, Casa C, Arnaud JP, Hamy A. The use of the Alvarado score in the management of right lower quadrant abdominal pain in the adult. J Visc Surg. 2010;147(2):e40-4.

8. McKay R, Shepherd J. The use of the clinical scoring system by Alvarado in the decision to perform computed tomography for acute appendicitis in the ED. Am J Emerg Med. 2007;25(5):489-93.

9. Ohmann C, Franke C, Yang Q. Clinical benefit of a diagnostic score for appendicitis: results of a prospective interventional study. German Study Group of Acute Abdominal Pain. Arch Surg. 1999;134(9):993-6.

10. Douglas CD, Macpherson NE, Davidson PM, Gani JS. Randomised controlled trial of ultrasonography in diagnosis of acute appendicitis, incorporating the Alvarado score. BMJ. 2000;321(7266):919-22.

11. Peipert JF, Ness RB, Blume J, Soper DE, Holley R, Randall H, Sweet RL, Sondheimer SJ, Hendrix SL, Amortegui A, Trucco G, Bass DC. Clinical predictors of endometritis in women with symptoms and signs of pelvic inflammatory disease. Am J Obstet Gynecol. 2001;184(5):856-63; discussion 63-4.

12. Morishita K, Gushimiyagi M, Hashiguchi M, Stein GH, Tokuda Y. Clinical prediction rule to distinguish pelvic inflammatory disease from acute appendicitis in women of childbearing age. Am J Emerg Med. 2007;25(2):152-7.

13. Huchon C, Staraci S, Fauconnier A. Adnexal torsion: a predictive score for pre-operative diagnosis. Hum Reprod. 2010;25(9):2276-80.

14. Biesheuvel CJ, Vergouwe Y, Steyerberg EW, Grobbee DE, Moons KG. Polytomous logistic regression analysis could be applied more often in diagnostic research. J Clin Epidemiol. 2008;61(2):125-34.

15. Van Calster B, Valentin L, Van Holsbeke C, Testa AC, Bourne T, Van Huffel S, Timmerman D. Polytomous diagnosis of ovarian tumors as benign, borderline, primary invasive or metastatic: development and validation of standard and kernel-based risk prediction models. BMC Med Res Methodol. 2010;10:96.

16. Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Clinical Scoring for Diagnosis of Acute Lower Abdominal Pain in Female of Reproductive Age. Emergency Medicine International. 2013;2013:6.

17. Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Validation of the Diagnostic Score for Acute Lower Abdominal Pain in Women of Reproductive Age. Emergency Medicine International. 2014;2014:6.

18. Heverhagen JT, Klose KJ. MR imaging for acute lower abdominal and pelvic pain. Radiographics. 2009;29(6):1781-96.

19. Singh AK, Desai H, Novelline RA. Emergency MRI of acute pelvic pain: MR protocol with no oral contrast. Emerg Radiol. 2009;16(2):133-41.

20. Masselli G, Brunelli R, Casciani E, Polettini E, Bertini L, Laghi F, Anceschi M, Gualdi G. Acute abdominal and pelvic pain in pregnancy: MR imaging as a valuable adjunct to ultrasound? Abdom Imaging. 2010.

21. De Dombal FT, Horrocks JC, Staniland JR, Gill PW. Simulation of clinical diagnosis: a comparative study. Br Med J. 1971;2(5761):575-7.

22. de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-aided diagnosis of acute abdominal pain. Br Med J. 1972;2(5804):9-13.

23. De Dombal FT, Leaper DJ, Horrocks JC, Staniland JR, McCann AP. Human and computer-aided diagnosis of abdominal pain: further report with emphasis on performance of clinicians. Br Med J. 1974;1(5904):376-80.

24. Adams ID, Chan M, Clifford PC, Cooke WM, Dallos V, de Dombal FT, Edwards MH, Hancock DM, Hewett DJ, McIntyre N, et al. Computer aided diagnosis of acute abdominal pain: a multicentre study. Br Med J (Clin Res Ed). 1986;293(6550):800-4.

25. Liu JL, Wyatt JC, Deeks JJ, Clamp S, Keen J, Verde P, Ohmann C, Wellwood J, Dawes M, Altman DG. Systematic reviews of clinical decision tools for acute abdominal pain. Health Technol Assess. 2006;10(47):1-167, iii-iv.