Chapter 5

The new clinical diagnostic prediction score for facilitating the possible diagnoses associated with acute lower abdominal pain in women of reproductive age – Considerations for clinical practice

One purpose of any set of clinical prediction rules is to contribute value to clinical practice. The new clinical diagnostic prediction score proposed by this thesis aims to reduce the subjective individual judgment made by many emergency room physicians. Diagnostic processes will be more accurate and more effective if patients are directed to the right specialists. By using this clinical diagnostic prediction score, emergency room physicians can categorize patients into those suffering from appendicitis, obstetric and gynecological conditions (OB-GYNc) or non-specific abdominal pain (NSAP). However, there are some issues to be considered in the clinical usage of the new clinical diagnostic prediction for diagnosis of acute lower abdominal pain in women of reproductive age.

1. How certain can the predicted diagnosis be from the score allocated to each individual patient?

The first question that would be raised by clinicians is 'how much confidence can we have in the score?'. Although the score can identify most cases of appendicitis and OB-GYNc accurately, clinicians also need to know the limitations of the score when it comes to the predicted diagnosis. Interpretation of the likelihood ratios can inform the answer.

Let us review the mathematical basis of likelihood ratios from Baye's theorem that:

Likelihood ratios = post-test odds/pre-test odds

Then,

Post-test odds = likelihood ratio x pre-test odds

We can calculate post-test probability from post-test odds and pre-test probability.

By using data from the study shown in Appendix D[1], we can calculate post-test odds and post-test probabilities from likelihood ratios according to the results of the score predicted diagnosis of individual patients as illustrated in the following table:

Table 5.1 Likelihood ratios and post-test probabilities of appendicitis and Obstetric & Gynecological causes (OB-GYNc) of acute lower abdominal pain

Diagnoses and likelihood ratios (LR)	pre-test odds	pre-test probability	post-test odds	post-test probability
Appendicitis (n=197, total = 302)	197:105 = 1.88	197/302 = 0.65	= pre-test odds x likelihood ratio	= post-test odds/(post-test odds+1)
LR+ = 4.39			8.24	0.89
LR- = 0.10			0.19	0.16
OB-GYNc (n = 63, total = 302)	63:239 = 0.26	63/302 = 0.21		
LR+ = 8.73			2.27	0.69
LR- = 0.29			0.07	0.07

We can estimate that the probability of appendicitis increases from 0.65, without using the score, to 0.89 if the clinical diagnostic prediction score favors a diagnosis of appendicitis; and, the probability of appendicitis decreases to 0.16 if the score predicts otherwise. Similarly, the probability of OB-GYNc changes from 0.21 to 0.69 if the score predicts a diagnosis of OB-GYNc, and to 0.07 if it predicts otherwise.

2. Application of the new clinical diagnostic prediction score to different settings – an issue of external validity

The new clinical diagnostic prediction score that is presented in this thesis has high degrees of accuracy in diagnosing the cause of acute lower abdominal pain in women of reproductive age. However, the score has been developed and validated in the sole setting of a single tertiary-care hospital. We can say that the score is valid internally, but for confident application in other settings it needed to have external validity. External validity depends on two main factors; differences in case-mix and differences in regression coefficients.[2] Differences in case-mix may affect validity in other levels of care where we expect less prevalence of diseases. This does not mean that the score cannot be applied to other settings other than the setting that it had been developed. Alvarado's score, for example, was developed in a single setting but is still widely used in general clinical practice.[3-7]

Differences in case-mix can affect predictive accuracies of the score; however, post-test probabilities can be predicted from pre-test odds and likelihood ratios as shown in table 5.1. If we know pre-test odds (or prevalence of disease) we can estimate post-test probabilities, or positive predictive value and negative predictive value in such settings. One advantage of the score presented in this thesis is that it was developed from a large cohort of patients. Unlike a case-control study where disease prevalence was fixed by case: control ratios, the prevalence (or pre-test odds) of appendicitis and OB-GYNc were the true data sets. This can be an explanation of similar patterns of diagnostic accuracy indices in the derivation study and the validation study.[1,8]

Another advantage of the new clinical diagnostic score is the use of a multiple imputation method in the handling of missing values of the predictors.[9,10] Evidence from simulations suggested that imputation methods are superior to the analysis of complete data only (complete case analysis).[2] Study records with missing values may have some characteristics that differ from the complete data set. If we were using prediction models from complete case analysis only, it would be subjected to some degree of selection bias. This is especially true in the case of this thesis when the records with missing values were usually related to the percentage of neutrophils (26% missing) and the number of white blood cells (23.2% missing). Most of these incomplete records were records of patients referred from rural hospitals.

Differences in regression coefficients in a new setting, on the other hand, are more likely to affect the validity of the diagnostic prediction score. As the score was derived from regression coefficients of the predictor model, different regression coefficients can cause invalidity of diagnostic prediction. Regression coefficients in other settings may be different because of real differences between populations, differences in definitions of predictors and outcomes, and differences in patient selection.[2] A further study for external validation and the updating of the prediction models is needed in order to safely apply this diagnostic prediction score in different settings.

3. Conclusion

To apply a clinical prediction rule to a new setting, the prediction rule needs to be valid. Internal validation of the proposed clinical diagnostic prediction score for acute lower abdominal pain in women of reproductive age is confirmed from the validation study in this

thesis. It is statistically advantageous to apply this score in the clinical practice of the setting where the score was developed. To apply the score in different settings, more external validation studies and a possible updating of the prediction model are necessary.

References

1. 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.

2. Steyerberg EW. *Clinical Prediction Models*. New York: Springer Science+Business Media,LLC; 2009.

3. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* May 1986;15(5):557-564.

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

5. Douglas CD, Macpherson NE, Davidson PM, Gani JS. Randomised controlled trial of ultrasonography in diagnosis of acute appendicitis, incorporating the Alvarado score. *BMJ.* Oct 14 2000;321(7266):919-922.

6. Horzic M, Salamon A, Kopljar M, Skupnjak M, Cupurdija K, Vanjak D. Analysis of scores in diagnosis of acute appendicitis in women. *Coll Antropol.* Jun 2005;29(1):133-138.

7. Ohle R, O'Reilly F, O'Brien KK, Fahey T, Dimitrov BD. The Alvarado score for predicting acute appendicitis: a systematic review. *BMC Med.* 2011;9:139.

8. 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.

9. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med.* Apr 1991;10(4):585-598.

10. Rubin DB. Inference and missing data. *Biometrika.* 1976;63:581-592.