Implementation of probabilistic decision rules improves the predictive values of algorithms in the diagnostic management of ectopic pregnancy

Ben W.J.Mol1,2,3, Fulco van der Veen2 and Patrick M.M. Bossuyt1

1Department of Clinical Epidemiology and Biostatistics and 2Centre for Reproductive Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
3To whom correspondence should be addressed: Department of Clinical Epidemiology and Biostatistics, Academic Medical Centre, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands

Current algorithms for the diagnosis of ectopic pregnancy do not take into account the heterogeneity in patient profiles. Such heterogeneity can lead to differences in the pre-test probability of ectopic pregnancy. In patients with clinical symptoms, for example, the probability of presence of an ectopic pregnancy is higher than in symptom-free patients. Any additional tests should then be interpreted differently, depending on the pre-test probability. We present a diagnostic algorithm that uses probabilistic decision rules for the evaluation of women with suspected ectopic pregnancy with flexible cut-off levels for test positivity. We compare it with a general algorithm that uses fixed cut-off levels. Fictitious cohorts, varying in prevalence of ectopic pregnancy were put together, using data obtained from a cohort of >800 women with suspected ectopic pregnancy. In the inflexible algorithm, ectopic pregnancy was diagnosed whenever it could be visualized at transvaginal sonography, or where serum human chorionic gonadotrophin (HCG) exceeded a rigid cut-off level; ectopic pregnancy was rejected if an intrauterine pregnancy was seen or when serum HCG decreased. In the flexible algorithm, post-test probability was obtained after each test, using pre-test probabilities and test-based likelihood ratios. Ectopic pregnancy was diagnosed whenever the post-test probability for ectopic pregnancy exceeded 95%, whereas this diagnosis was rejected if the calculated post-test probability fell below 1%. For both algorithms, sensitivity and specificity as well as predictive values were calculated. At each prevalence, the inflexible algorithm was associated with a sensitivity of 93% and a specificity of 97%. In contrast, the sensitivity and specificity of the flexible, individualized algorithm depended on the prevalence of ectopic pregnancy. Consequently, predictive values varied strongly when the inflexible algorithm was used, whereas they were much more stable after using the flexible algorithm. For five possible valuations of false positive and false negative diagnoses, the flexible algorithm reduced the expected disutility, compared with the inflexible algorithm. It is concluded that clinicians should incorporate probabilistic decision rules in algorithms used for the diagnosis of ectopic pregnancy.

Key words: algorithms/ectopic pregnancy/human chorionic gonadotrophin/probabilistic decision rules

Introduction

Until a decade ago, suspected ectopic pregnancy was a difficult clinical problem, due to the absence of non-invasive diagnostic tests, laparoscopy being the only reliable way to establish a final diagnosis. The introduction of accurate urine pregnancy tests, of transvaginal sonography and of serum human chorionic gonadotrophin (HCG) measurement have reduced the importance of laparoscopy (Ankum et al., 1993; Carson and Buster, 1993; Mishalani et al., 1994). A questionnaire completed by Dutch gynaecologists in 1995 showed that transvaginal sonography and serum HCG were implemented in the vast majority of the hospitals in The Netherlands (Mol et al., 1996). Apparently, these new tools have answered a need in clinical practice.

Since the variety of non-invasive clinical tools offers numerous possibilities for developing diagnostic strategies, there was a need for clinical guidelines in the diagnosis of ectopic pregnancy. To support the clinician in daily practice, several authors have developed algorithms for the management of suspected ectopic pregnancy, incorporating transvaginal sonography and serum HCG measurement, sometimes preceded by serum progesterone measurement (Ankum et al., 1993; Carson and Buster, 1993; Stovall and Ling, 1993). These algorithms have shown excellent sensitivity and specificity in heterogeneous populations of women with suspected ectopic pregnancy, thereby reducing the number of unnecessary laparoscopies.

As a consequence of the general availability of non-invasive diagnostic tools, and since these tools are a minor burden for the patients, the clinical context of the patient with suspected ectopic pregnancy has changed dramatically. Nowadays, the classical clinical case, a woman with severe abdominal pain and a possible pregnancy, is rarely seen, since ectopic pregnancy is often already suspected in early pregnancy, sometimes even before the onset of clinical symptoms (Cacciatore et al., 1994; Mol et al., 1997). In these patients, the prevalence of ectopic pregnancy is lower than in those with the classical clinical picture, thus increasing the hazard of false positive diagnoses (Sackett et al., 1985).

These differences in the strength of suspicion of ectopic pregnancy should have consequences for diagnostic management. In women with a clinical profile corresponding to a low
degree of suspicion, test results corresponding with higher likelihood ratios are required for a diagnosis ectopic pregnancy. In women with a clinical profile corresponding to a high degree of suspicion of ectopic pregnancy, the probability of an ectopic pregnancy need only be increased a little further in order to establish the diagnosis and to warrant therapeutic action.

This heterogeneity in the degree of suspicion of ectopic pregnancy has not been incorporated in the diagnostic algorithms available in the medical literature. Since these algorithms all use rigid cut-off levels for clinical decisions, they do not take into account differences in pre-test probability. For example, the likelihood ratio of a serum HCG level of 3000 IU/l is 15 where sonography shows neither an intrauterine pregnancy nor adnexal abnormalities (Mol et al., 1998). A pre-test probability for ectopic pregnancy of 5%, well suited for a symptom-free patient with a previous ectopic pregnancy, results in this patient having a post-test probability for ectopic pregnancy of 44% (Mol et al., 1997). The same serum HCG value in a patient with abdominal pain, with a pre-test probability of 40%, results in a post-test probability of 92%. Such differences could have clinical implications, since confirmatory laparoscopy or treatment with systemic methotrexate is only justified when the probability of ectopic pregnancy is high. The use of rigid cut-off levels for the diagnosis of ectopic pregnancy could lead to unwarranted therapeutic action and harm to the patient.

In this study, we present a more flexible individualized diagnostic algorithm that uses probabilistic decision rules for the evaluation of suspected ectopic pregnancy. We have compared its diagnostic performance with that of a general algorithm based on rigid cut-off levels.

**Materials and methods**

The backbone of both the inflexible, general algorithm and the flexible, individualized diagnostic algorithm is the combination of transvaginal sonography and serum HCG measurement, as presented previously (Ankum et al., 1993; Mol et al., 1998).

In the inflexible algorithm, an intrauterine pregnancy is diagnosed when an intrauterine gestational sac is visualized at transvaginal sonography. If an intrauterine gestational sac cannot be visualized, both adnexal regions are scanned; ectopic pregnancy is diagnosed if a yolk sac, a fetal pole, fetal cardiac activity or an ectopic mass and fluid in the pouch of Douglas is seen. If transvaginal sonography does not lead to a diagnosis, serum HCG measurement is performed. Taking into account the ‘discretorium zone principle’, as defined previously (Kadar et al., 1981), ectopic pregnancy is diagnosed for serum HCG concentration >2000 IU/l, whereas in patients in whom sonography showed an adnexal mass or fluid in the pouch of Douglas is seen. If transvaginal sonography shows an adnexal mass or fluid in the pouch of Douglas, a serum HCG concentration of 1500 IU/l or higher suffices (Kadar et al., 1981; Mol et al., 1998). When the serum HCG concentration is <1500 IU/l, transvaginal sonography and serum HCG are repeated at 2 day intervals until a diagnosis is made. The cut-off level for diagnosis of ectopic pregnancy at repeat serum HCG measurement is 1000 IU/l. If the second repeat HCG measurement does not reach this threshold, ectopic pregnancy is diagnosed when serum HCG is increasing, whereas decreasing serum HCG concentrations are supposed to rule out ectopic pregnancy.

The flexible, individualized algorithm, which uses probabilistic decision rules, is shown in Figure 1. Based on observations in our own hospital, we let the pre-test probability of an ectopic pregnancy in a patient depend on three findings of the medical history: abdominal pain, vaginal bleeding and previous exposure to a risk indicator. In the absence of risk indicators, the pre-test probability for ectopic pregnancy is 34% in a patient with abdominal pain, 18% in a patient with vaginal bleeding, and 39% when a patient presents with both symptoms. If one or more risk indicators are present, these probabilities are 42, 34 and 54% respectively. The pre-test probability of ectopic pregnancy in the absence of abdominal pain and vaginal bleeding in a woman without a risk indicator is <1%. In symptom-free women with at least one risk indicator this pre-test probability is 6%. Risk indicators are a history of ectopic pregnancy, tubal surgery or pelvic inflammatory disease; tubal disease detected by hysterosalpingography or laparoscopy; in-uterro exposure to diethyl stilboestrol (DES), sterilization, and an intrauterine contraceptive device (IUCD) in situ at the moment of conception (Mol et al., 1995; Ankum et al., 1996).

When additional tests are performed these pre-test probabilities are modified through Bayes’ theorem, using likelihood ratios. A likelihood ratio of a particular test result is defined as the ratio of the relative frequency of that particular test result in patients with an ectopic pregnancy and the relative frequency of that particular test result in patients without an ectopic pregnancy. A test result with a likelihood ratio >1 increases the probability of ectopic pregnancy as compared to the probability of ectopic pregnancy before the performance of a test. The higher the likelihood ratio >1, the more likely becomes the probability of an ectopic pregnancy. A test result with a likelihood ratio <1 decreases the probability of ectopic pregnancy as compared to the probability of ectopic pregnancy before the performance of a test. The lower the likelihood ratio below 1, the less likely becomes the probability of an ectopic pregnancy. Each possible result is associated with a single, fixed likelihood ratio. The magnitude of these likelihood ratios is based on previous research (Mol et al., 1998). This way, women with identical findings at additional testing but with different pre-test probabilities will have different post-test probabilities; the higher the pre-test probability, the higher the post-test probability will be.

The first test is scanning of the intrauterine cavity. The presence of an intrauterine gestational sac is associated with a likelihood ratio of 0.07; absence of an intrauterine gestational sac has a likelihood ratio of 2.2 (Mol et al., 1998). If the calculated post-test probability drops below a threshold that we presume to be low enough to rule out ectopic pregnancy, say 1%, the diagnosis of ectopic pregnancy is rejected and further diagnostic tests can be delayed. Alternatively, if the post-test probability exceeds a threshold that we presume to be high enough to justify confirmatory laparoscopy, or the start of medical treatment, say 95%, the diagnosis ectopic pregnancy can be made, and further diagnostic tests can also be delayed. For example, a pregnant patient with abdominal pain and a previous ectopic pregnancy has a pre-test probability of 42%. When transvaginal sonography in this patient does not show an intrauterine pregnancy, one can derive from the Appendix that the probability of ectopic pregnancy increases to 60%. This post-test probability does not reach the threshold of 95% that was required for the diagnosis ectopic pregnancy. Consequently, we need a second test, i.e. scanning of the adnexal region.

Scanning of the adnexal region can show ectopic cardiac activity, an ectopic gestational sac, an ectopic mass and/or fluid in the pouch of Douglas. Ectopic cardiac activity proves the presence of an ectopic pregnancy (likelihood ratio of infinity), whereas the other features are associated with likelihood ratios of 23, 3.6 and 4.4 respectively (Mol et al., 1998). Presence of an ectopic mass with fluid in the pouch of Douglas has a likelihood ratio of 9.9. A completely normal adnexal region has a likelihood ratio of 0.55.
Results

Table I shows the results of the analysis. The inflexible algorithm showed stable sensitivities and specificities for different prevalences, 93 and 97%, respectively. In contrast, the sensitivity and specificity of the flexible algorithm depended on the prevalence. When the prevalence of ectopic pregnancy was 5%, the sensitivity of the algorithm was 90% with a specificity of 99%. If the prevalence increased, the sensitivity also increased with a slowly decreasing specificity; at a prevalence of ectopic pregnancy of 50%, the sensitivity of the flexible algorithm was 95% with a specificity of 94%.

Table I and Figure 1 also show the predictive values generated by the algorithms for different prevalences. In the inflexible algorithm the predictive values were strongly dependent on the prevalence. At a prevalence of 5%, the positive predictive value was only 58% with a negative predictive value of 0.4%. When the prevalence of ectopic pregnancy increased, the positive predictive value also strongly increased, with a slightly increasing negative predictive value.

At a prevalence of 5%, the flexible algorithm had a positive predictive value of 78%, 20% higher than the inflexible algorithm, with a negative predictive value that was only slightly higher (0.5%). If the prevalence of ectopic pregnancy increased, the positive and negative predictive values also increased, to become 96 and 5.4% for a prevalence of 50%. In the inflexible algorithm, the diagnosis of ectopic pregnancy was delayed for 2 days in 8% of the patients and for 4 days in 27% of the patients at a prevalence of 5%. In the flexible algorithm, these rates were 15 and 39% at a prevalence of 5%, and 15 and 28% at a prevalence of 50%.

Figure 2 shows two ROC curves, the first evaluating the impact of different thresholds for the probability required to diagnose ectopic pregnancy, the second evaluating the impact of different thresholds for the probability required to rule out ectopic pregnancy. Both ROC curves were constructed assuming a pre-test prevalence of 10%. At a threshold probability to diagnose ectopic pregnancy of 95% the algorithm had a sensitivity of 92% and a specificity of 98%. Above this threshold, the sensitivity dropped slightly without an increase in specificity. If the threshold decreased below 95%, the specificity decreased strongly without an increase in sensitivity. At a threshold probability of 1% to rule out ectopic pregnancy sensitivity and specificity were 92 and 98% respectively. An increase of this threshold to 5% resulted in a strong decrease in sensitivity for a very small increase in specificity.

Figure 3 shows the expected disutility reduction by using the flexible algorithm as compared with the inflexible algorithm. For virtually all valuations of false positive and false negative results, the flexible algorithm performed better than the inflexible algorithm. Only at a prevalence of over 35% and a valuation of a false negative diagnosis only 0.10 times worse than a false positive diagnosis did the inflexible algorithm perform slightly better.