Ultrasound Obstet Gynecol 2006; **27**: 311–315
Published online 2 February 2006 in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/uog.2702

The practical application of a mathematical model to predict the outcome of pregnancies of unknown location

E. KIRK*, G. CONDOUS*, Z. HAIDER*, C. LU†, S. VAN HUFFEL†, D. TIMMERMAN† and T. BOURNE*

**Early Pregnancy, Gynaecological Ultrasound and MAS Unit, St George's Hospital, London, UK and †Department of Obstetrics and Gynaecology, University Hospital Gasthuisberg and Department of Electrical Engineering (ESAT), KU Leuven, Belgium*

KEYWORDS: ectopic pregnancy; hCG ratio; mathematical model; PUL

ABSTRACT

Objective *A logistic regression model has been developed previously to predict which pregnancies of unknown location (PULs) become ectopics. This model was based on the human chorionic gonadotropin (hCG) ratio (hCG 48 h/hCG 0 h). The aim of this study was to evaluate the model in an early pregnancy clinical setting.*

Methods *This was a prospective non-interventional observational study. Women classified as having a PUL had serum hCG and progesterone levels taken at presentation and 48 h later. At 48 h, nurse practitioners in the early pregnancy unit gave their subjective assessment of the likely pregnancy outcome based upon existing strategies to predict the outcome of PULs and their own previous experience. This was used as the basis for the clinical management of the women. The hormonal data were then entered into the model, which was available on the clinic computer in the form of a Microsoft Excel package, and the predicted outcome according to this model was recorded for each case. The model was weighted in favor of detecting an ectopic pregnancy at the expense of failing PULs and intrauterine pregnancies. All women were followed up until the final true diagnosis was established: a failing PUL, an intrauterine pregnancy or an ectopic pregnancy. The predicted outcome according to the model was compared with that of the subjective assessment of the expert operator and the true outcome after follow-up.*

Results *Data were obtained from 357 women classified as having PULs. The final clinical outcomes were: 162 (45.4%) failing PULs, 166 (46.5%) intrauterine pregnancies and 29 (8.1%) ectopic pregnancies. Subjective assessment utilizing current strategies gave sensitivities of 79.3%, 92.8% and 87.0% for the detection of ectopic pregnancy, intrauterine pregnancy and failing PUL, respectively. The model detected ectopic pregnancy, intrauterine pregnancy and failing PUL with sensitivities of 82.8%, 86.8% and 73.5%, respectively. There were no adverse outcomes.*

Conclusion *This logistic regression model, based on the hCG ratio, can be used to predict the outcome of PULs, especially the ectopic pregnancies, with a high degree of certainty. It compares favorably with subjective assessment by experienced operators using current strategies to predict the outcome of PULs. The model can be used by those with limited knowledge or understanding of the behavior of serum biochemistry in the first trimester and in turn aid in the classification of PULs into those at low and those at high risk of ectopic pregnancy at 48 h. Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.*

INTRODUCTION

In the early pregnancy unit (EPU), the location of a pregnancy may be confirmed on the basis of the initial transvaginal ultrasound examination[1,2] in 90–92% of women who present with a positive urinary pregnancy test. Diagnoses include a viable or failed intrauterine pregnancy, or an ectopic pregnancy. The remaining 8–10% have an empty uterus with no signs of an intra- or extrauterine pregnancy, and these are classified as pregnancies of unknown location (PULs). Whilst the majority will be either intrauterine pregnancies too early to visualize on ultrasound or failing PULs, 10.8–14% of PULs are ectopic pregnancies that are too early to visualize or are missed on ultrasound[1,2]. The failing PUL group includes both intra- and extrauterine pregnancies, as the

Correspondence to: Dr E. Kirk, Early Pregnancy, Gynaecological Ultrasound and MAS Unit, St George's Hospital School, Cranmer Terrace, London, SW17 0RE, UK (e-mail: ejkirk@hotmail.co.uk)

Accepted: 27 July 2005

Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.

location of the pregnancy is never determined. A small group of women will have a persisting PUL, defined as one which behaves biochemically like an ectopic pregnancy: the serum human chorionic gonadotropin (hCG) level fails to decline but no evidence of the pregnancy is ever identified by transvaginal ultrasound or laparoscopy[2]. A 'wait and see' approach has been demonstrated to be safe in the management of PULs[1,3,4].

Currently, women with PULs are followed up with serial hormone measurements, repeat transvaginal ultrasound examinations and possible laparoscopy until a diagnosis is confirmed. The established hormonal criteria, based on serial hCG and progesterone levels, predict the viability of a PUL, but not its location[2]. Women often end up undergoing a lengthy follow-up before a diagnosis is made. For the majority this is often unnecessary, as there is a low risk for ectopic pregnancy[5]. We have previously developed a logistic regression model to predict the outcome of PULs based on the results of two serum hCG levels taken 48 h apart[6]. This model was found to have a sensitivity of 92%, a specificity of 84% and a positive predictive value of 28% for the detection of ectopic pregnancy. The aim of the present study was to evaluate the performance of this model in comparison with clinical management decisions made by experienced nurse practitioners in the clinic.

METHODS

This was a prospective non-interventional observational study. From 7th August 2003 to 18th September 2004, women presenting to the EPU at St George's Hospital, London with a positive urinary pregnancy test underwent a transvaginal ultrasound examination for various reasons. Indications included lower abdominal pain, vaginal bleeding, maternal anxiety or confirmation of gestational age.

Women were classified as having a PUL if there was no evidence of an intra- or extrauterine pregnancy on transvaginal sonography. Exclusion criteria were: 1) visualization of an intrauterine gestational sac with or without a fetal pole, 2) identification of an adnexal mass thought to be an ectopic pregnancy, 3) heterogeneous tissue within the endometrial cavity thought to represent retained products of conception, and 4) maternal hemodynamic instability or pain, or demonstration of a hemoperitoneum on transvaginal sonography. Women classified with a PUL had blood taken to measure their serum hCG (World Health Organization, Third International Reference 75/537) and progesterone (Roche Elecsys 2010 Progesterone II test, Roche Diagnostics, Lewes, UK) level using automated electrochemiluminescence immunoassays (ECLIAs) at presentation (0 h) and 48 h later.

Nurse practitioners, whose clinical and scanning experience in the EPU varied from 5 to 10 years, gave their subjective assessment of the predicted outcome of the PUL at 48 h when the result of the woman's second hCG level became available. This was based on existing strategies to predict the outcome of PULs and their own previous experience. If the serum hCG level increased by $> 66\%$ in 48 h they generally classified PULs as probable intrauterine pregnancies[6]. If the hCG level decreased over 48 h or the initial progesterone level was < 20 nmol/L they generally classified a pregnancy as a failing PUL[1]. Possible ectopic pregnancies were classified on the basis of a either a discriminatory zone of ≥ 1500 U/L hCG or a suboptimal rise of $< 66\%$ over 48 h. These predicted outcomes were recorded at 48 h and compared with the final outcomes as well as with the predicted outcomes of the logistic regression model. The nurses were blinded to the outcomes predicted by the model.

The logistic regression model based on the hCG ratio was used to predict the probability of an ectopic pregnancy (P_{ectopic}), a failing PUL (P_{failing}) or an intrauterine pregnancy (P_{IUP}). Models have been developed using age and progesterone level as variables[6]. However, they did not perform as well as did the hCG ratio model, so for the purposes of this study they were not used. The model calculated the probability for each class directly:

$$P_{\text{ectopic}} = \frac{e^{5.79 - 4.21\ \text{hCG ratio}}}{1 + e^{5.79 - 4.21\ \text{hCG ratio}} + e^{9.92 - 7.66\ \text{hCG ratio}}}$$

$$P_{\text{failing}} = \frac{e^{9.92 - 7.66\ \text{hCG ratio}}}{1 + e^{5.79 - 4.21\ \text{hCG ratio}} + e^{9.92 - 7.66\ \text{hCG ratio}}}$$

$$P_{\text{IUP}} = \frac{1}{1 + e^{5.79 - 4.21\ \text{hCG ratio}} + e^{9.92 - 7.66\ \text{hCG ratio}}}$$

Predictions were made using two probability cut-offs that had been calculated previously to maximize the sensitivity and specificity of the model when it was being developed. If the predicted probability of a PUL to be an ectopic pregnancy was > 0.21, then it was classified as an ectopic pregnancy; otherwise it was classified as a non-ectopic PUL. If a pregnancy was classified as non-ectopic, and the predicted probability for it to be a failing PUL was > 0.72, it was classified as a failing PUL; otherwise it was classified as an intrauterine pregnancy.

The costs of misdiagnosing an ectopic pregnancy are greater than those of misdiagnosing a failing PUL or an intrauterine pregnancy. We therefore used a weighting for misclassification of a failing PUL, intrauterine pregnancy and ectopic pregnancy of 1, 1 and 4, as previously published[6]. The optimal weighting was then determined retrospectively.

The model was incorporated into a Microsoft Excel (Microsoft Corp., Redmond, WA, USA) package and installed on the desktop of the central EPU computer so as to provide a 'user-friendly interface' (Figure 1). The serum hCG results at 0 h and 48 h were entered onto the Excel spreadsheet, giving not only an hCG ratio in real time but also probabilities for the three possible outcomes: 1) a failing PUL, 2) an intrauterine pregnancy and 3) an ectopic pregnancy. These predicted outcomes were recorded but not used for clinical management.

Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.

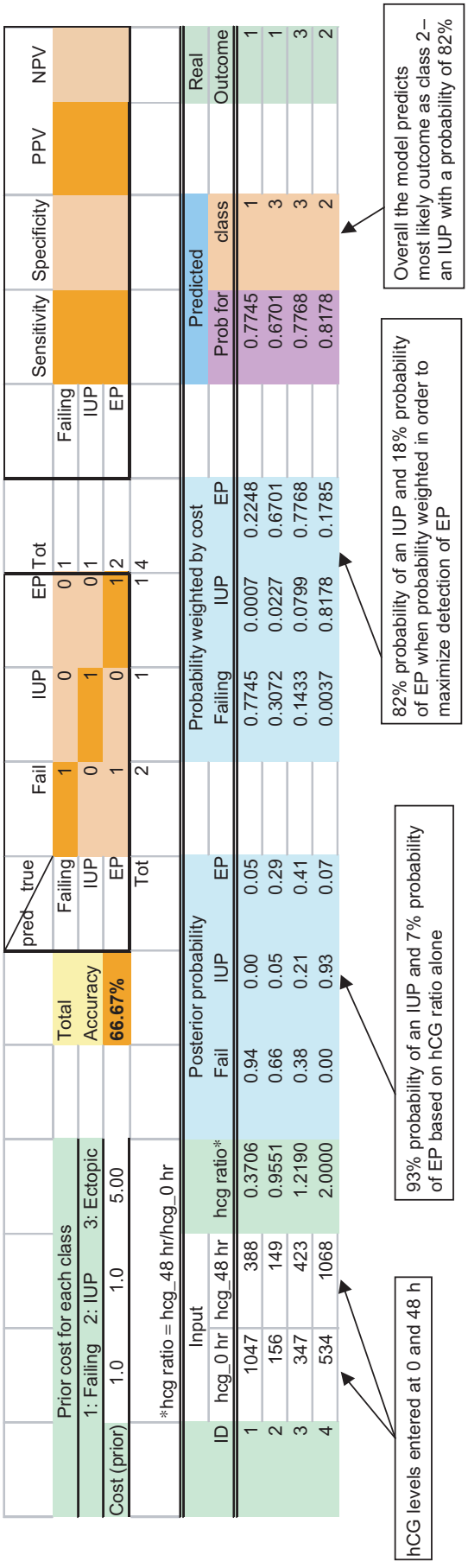


Figure 1 Example of weighted model as it appears on a computer database in the clinical setting. EP, ectopic pregnancy; hCG, human chorionic gonadotropin; IUP, intrauterine pregnancy.

Women were followed up until the final true clinical outcome was known. Follow-up was in the form of repeat serum hCG and progesterone level measurements and/or transvaginal sonography as necessary with or without laparoscopy. The final diagnoses were failing PUL, intrauterine pregnancy or ectopic pregnancy.

Management options were as follows: if the serum hCG level had increased by $\geq 66\%$ in 48 h, the women were classified as having an intrauterine pregnancy and were rescanned in 2 weeks to confirm the diagnosis[6]. Diagnosis of an intrauterine pregnancy was confirmed by visualization on transvaginal ultrasound of a hyperechoic ring (gestational sac) placed eccentrically in the endometrial cavity with or without a yolk sac or fetal pole, or heterogeneous tissue within the endometrial cavity thought to represent retained products of conception. If the initial serum progesterone level was < 20 nmol/L, women were classified as having a failing PUL[1] and serum hCG level was repeated within 7 days to confirm the diagnosis. The diagnosis of a failing PUL was confirmed by a spontaneous decrease in the hCG level to < 5 U/L, with the disappearance of pregnancy symptoms. Women not falling into either category had their hCG level repeated every 48 h and also had further scans (depending upon the hCG level and clinical symptoms) until a diagnosis was made.

Diagnosis of an ectopic pregnancy was made either from the operative findings at the time of surgery with confirmatory histopathology or on transvaginal sonography. Ectopic pregnancy was diagnosed on transvaginal sonography if one of the following criteria was fulfilled: 1) an inhomogeneous adnexal mass ('blob sign')[7], 2) an empty extrauterine gestational sac with a hyperechoic ring ('bagel sign')[8], or 3) a fetal pole with or without cardiac activity in an extrauterine sac. The visualization of an adnexal mass has been shown to have a sensitivity of over 90% for the detection of ectopic pregnancy[7,9]. Women who had a low hCG level that plateaued with no pregnancy ever visualized on transvaginal sonography were classified as having a persistent PUL. These almost certainly represented missed ectopic pregnancies and behaved biochemically like the ectopic pregnancy group. They were therefore analyzed as ectopic pregnancies.

The predicted outcomes from both subjective assessment and use of the model were compared with the true pregnancy outcomes for each woman. Those defining the true pregnancy outcomes were blinded to the outcomes predicted by the nurse practitioners and the model. The performance of the model and subjective assessment were evaluated in terms of sensitivity, specificity and positive and negative predictive values (PPV and NPV) for each of the pregnancy outcomes (failing PUL, intrauterine pregnancy and ectopic pregnancy). The overall diagnostic accuracy of the model was also calculated. Statistical analyses were conducted with SAS (version 8.2 for Windows, SAS Institute Inc., Cary, NC, USA).

RESULTS

A total of 5350 women presented to the EPU during the study period. Of these, 91.1% (4876/5350) were

Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.

Table 1 Predicted probabilities according to subjective assessment based on existing strategies

	True			
Predicted	*Failing*	*Intrauterine*	*Ectopic*	*Total*
Failing	141	7	4	152
Intrauterine	7	154	2	163
Ectopic	14	5	23	42
Total	162	166	29	357
	Sensitivity (%)	*Specificity (%)*	*PPV (%)*	*NPV (%)*
Failing	87.0	91.7	92.8	86.3
Intrauterine	92.8	85.9	94.5	84.5
Ectopic	79.3	89.9	54.8	93.7

NPV, negative predictive value; PPV, positive predictive value.

diagnosed as having either an intra- or an extrauterine pregnancy based only on the transvaginal ultrasound findings at their initial scan; 97.7% (4765/4876) were intrauterine and 2.3% (111/4876) were extrauterine pregnancies. The remaining 8.9% (474/5350) were classified as PULs. There were complete data on 357 PULs.

After follow-up, the final true clinical outcomes were 162 failing PULs (45.4%), 166 intrauterine pregnancies (46.5%) and 29 ectopic pregnancies (8.1%). Included in the ectopic pregnancy group were two women who had persistent PULs. The mean time from presentation until diagnosis of an ectopic pregnancy was 7.2 (range, 2–17) days. On average, the diagnosis of an intrauterine pregnancy was confirmed 2 weeks after presentation, at the time of follow-up ultrasound examination.

The nurse practitioners predicted that there were 152 failing PULs (42.6%), 163 intrauterine pregnancies (45.6%), and 42 ectopic pregnancies (11.8%). Their sensitivity for the detection of ectopic pregnancy was 79.3% (Table 1).

The model, when weighted 1:1:4, predicted 133 failing PULs (37.2%), 152 intrauterine pregnancies (42.6%), and 72 ectopic pregnancies (20.2%). The sensitivity for the detection of ectopic pregnancy, intrauterine pregnancy and failing PUL was 75.9%, 88.6% and 78.4%, respectively. When the cost of misclassification was changed to 1:1:5, the sensitivity for the detection of ectopic pregnancy, intrauterine pregnancy and failing PUL was 82.8%, 86.8% and 73.5%, respectively. Increasing the misclassification cost to 1:1:10 only slightly improved the sensitivity for the detection of ectopic pregnancy (86.2%), but significantly reduced the sensitivity for the detection of intrauterine pregnancy (77.1%) and failing PUL (58.0%) (Table 2).

Table 2 Performance of the model based on weighted probabilities, with different cost values for misclassifying a failing pregnancy of unknown location, an intrauterine pregnancy and an ectopic pregnancy

	True pregnancies														
	Cost values 1:1:4				*Cost values 1:1:5*				*Cost values 1:1:10*						
Predicted pregnancies	*Failing*	*Intrauterine*	*Ectopic*	*Total*	*Failing*	*Intrauterine*	*Ectopic*	*Total*	*Failing*	*Intrauterine*	*Ectopic*	*Total*			
Failing	127	1	5	133	119	1	3	123	94	1	2	97			
Intrauterine	3	147	2	152	3	144	2	149	2	128	2	132			
Ectopic	32	18	22	72	40	21	24	85	66	37	25	128			
Total	162	166	29	357	162	166	29	357	162	166	29	357			
	Sens. (%)	*Spec. (%)*	*PPV (%)*	*NPV (%)*	*Sens. (%)*	*Spec. (%)*	*PPV (%)*	*NPV (%)*	*Sens. (%)*	*Spec. (%)*	*PPV (%)*	*NPV (%)*			
Failing	78.4	86.7	95.5	75.5	73.5	86.2	98.4	71.8	58.0	78.5	96.9	58.9			
Intrauterine	88.6	78.1	96.7	72.6	86.8	74.9	96.6	68.8	77.1	62.3	97.0	52.9			
Ectopic	75.9	83.5	30.6	96.1	82.8	80.2	28.2	97.7	86.2	67.7	19.5	96.9			

NPV, negative predictive value; PPV, positive predictive value; Sens., sensitivity; Spec., specificity.

DISCUSSION

We have shown that a mathematical model which incorporates serum biochemistry in the first trimester can function as well as can experienced clinicians. We

Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.

acknowledge, however, that this was a non-interventional study and the model was used in parallel with current management strategies.

In detecting ectopic pregnancies, the model appeared to out perform the expert operators, who had a sensitivity of 79.3%. Using the original version of the model, with a misclassification cost of 1:1:4, the sensitivity for detection of ectopic pregnancy was 75.9%. However, when we analyzed retrospectively the data, the highest sensitivity was achieved with a misclassification cost of 1:1:5 (82.8%). We have implemented this in a new prospective clinical trial, in which the model is currently being used in real time and its effectiveness will be evaluated.

Current strategies to predict the outcome of PULs include: a progesterone level < 20 nmol/L to predict a failing pregnancy[1], a suboptimal rise in hCG of $< 66\%$ over 48 h and a discriminatory zone of 1500 U/L to predict an ectopic pregnancy, and an hCG rise of $> 66\%$ to predict an intrauterine pregnancy[6]. The advantage of using this model over current strategies is that no prior knowledge of the behavior of serum hCG and progesterone in early pregnancy is required. Practitioners of varying clinical experience and levels of training can therefore use it in a multidisciplinary setting. This is particularly useful in the ectopic pregnancy PUL group, whose heterogeneous biochemical behavior can often result in misclassification, some ectopic pregnancies mimicking failing PULs and others mimicking ongoing intrauterine pregnancies. At present the follow-up of women with a PUL is often protracted, especially in those women whose biochemistry does not conform to either a failing PUL or an intrauterine pregnancy. This could now be addressed, as the model classifies PULs at 48 h. Although it may take more than a week to confirm the location of the pregnancy, use of the model should be able to classify women into those at high and those at low risk of ectopic pregnancy. Further follow-up can then be targeted only at those at high risk of ectopic pregnancy.

The model was designed to predict ectopic pregnancies at the expense of intrauterine pregnancies and failing PULs, and so it is not surprising that the model overdiagnosed ectopic pregnancy. It missed 17.2% (5/29) of ectopic pregnancies: three were classified as failing PULs and two as intrauterine pregnancies. The majority (65.6%, 40/61) of women incorrectly predicted to have an ectopic pregnancy were subsequently diagnosed with a failing PUL. Interestingly, increasing the cost for the misclassification of an ectopic pregnancy to 10 only resulted in a slight increase in the sensitivity for detecting ectopic pregnancy (86.2%), but reduced the sensitivity for the detection of both intrauterine pregnancies and failing PULs. As a consequence of increasing the sensitivity, the specificity for the detection of ectopic pregnancy was reduced. This may result in more falsely diagnosed ectopic pregnancies undergoing laparoscopy or treatment with methotrexate.

The prevalence of ectopic pregnancies in this study was lower (8.1%) than in previously published studies; consequently, the PPV was low. As the number of ectopic pregnancies in a PUL population increases, so should the PPV of the model. In the UK, many EPUs still use transabdominal ultrasound examination. Consequently, more ectopic pregnancies and intrauterine pregnancies will be missed and so more pregnancies will be classified as PULs. It is reported that, using transabdominal scans, intrauterine pregnancies and therefore ectopic pregnancies can still be missed up to an hCG level of 6500 U/L[9]. Therefore, one would expect the model to perform well in such circumstances due to the likely increased prevalence of ectopic pregnancy in such PUL populations. However, this model is no substitute for the provision of a high-quality ultrasound service. As the model relies on the objective measurement of serum hCG level alone, one would expect these results to be reproducible in other units. However, this needs to be validated in a multicenter trial on different populations.

We believe that it was important to perform our validation study for this new model. The use of the model is independent of experience and does not require an understanding of serum biochemistry. We have demonstrated that the model performs as well as do senior clinicians in terms of classifying the likely outcome for a PUL and so selecting women for appropriate management. If the model is found to apply as well in other populations, it could be used by anyone in the multidisciplinary team to safely manage women with PULs and as a basis on which to justify management decisions.

REFERENCES

1. Banerjee S, Aslam N, Woelfer B, Lawrence A, Elson J, Jurkovic D. Expectant management of early pregnancies of unknown location: a prospective evaluation of methods to predict spontaneous resolution of pregnancy. *BJOG* 2001; **108**: 158–163.
2. Condous G, Lu C, Van Huffel S, Timmerman D, Bourne T. Human chorionic gonadotrophin and progesterone levels for the investigation of pregnancies of unknown location. *Int J Gynecol Obstet* 2004; **86**: 351–357.
3. Hahlin M, Thorburn J, Bryman I. The expectant management of early pregnancies of uncertain site. *Hum Reprod* 1995; **10**: 1223–1227.
4. Mol BW, van der Veen F. Letter to the Editor re. expectant management of women with early pregnancy of unknown location. *Ultrasound Obstet Gynecol* 2000; **15**: 265.
5. Condous G, Okaro E, Khalid A, Lu C, Van Huffel S, Timmerman D, Bourne T. A prospective evaluation of a single visit strategy to manage pregnancies of unknown location. *Hum Reprod* 2005; **20**: 1398–1403.
6. Condous G, Okaro E, Khalid A, Timmerman D, Lu C, Zhou Y, Van Huffel S, Bourne T. The use of a new logistic regression model for predicting the outcome of pregnancies of unknown location. *Hum Reprod* 2004; **19**: 1900–1910.
7. Condous G, Okaro E, Khalid A, Lu C, Van Huffel S, Bourne T. The accuracy of transvaginal ultrasonography for the diagnosis of ectopic pregnancy prior to surgery. *Hum Reprod* 2005; **20**: 1404–1409.
8. Goldstein S, Timor-Tritsch IE. *Ultrasound in Gynecology*. Churchill Livingstone: New York, 1995; Chapter 15, p 228.
9. Romero R, Kadar N, Jeanty P, Copel JA, Chervenak FA, DeCherney A, Hobbins JC. Diagnosis of ectopic pregnancy: value of the discriminatory human chorionic gonadotropin zone. *Obstet Gynecol* 1985; **66**: 357–360.

Copyright © 2006 ISUOG. Published by John Wiley & Sons, Ltd.