Expectant management of early pregnancies of unknown location: a prospective evaluation of methods to predict spontaneous resolution of pregnancy

Saikat Banerjee, Naaila Aslam, Brigitte Woelfer, Alexandra Lawrence, Janine Elson, Davor Jurkovic *

Objective To assess prospectively the ability of two multiparameter diagnostic models and their individual components to predict the outcome of early pregnancies which could not be identified on transvaginal ultrasound scan.

Design Prospective observational study.

Setting Dedicated early pregnancy unit in an inner city teaching hospital.

Population Women with a positive urine pregnancy test and clinical suspicion of early pregnancy complications.

Methods A full medical history, clinical examination and transvaginal ultrasound scan were carried out at the initial visit. When the location of the pregnancy could not be ascertained by ultrasound, serum beta-human chorionic gonadotrophin (\(b\)-hCG) and progesterone levels were measured. All women were managed expectantly until either a normal pregnancy was visualised on scan; the pregnancy resolved spontaneously or intervention was required due to a worsening of clinical symptoms or non-declining \(b\)-hCG levels.

Main outcome measures Spontaneous resolution of pregnancy (i.e. cessation of symptoms and decline in serum \(b\)-hCG level to \(< 20 \text{iu/L}\) without need for any active intervention.

Results Of the 104 women recruited, 72 (69%) pregnancies resolved spontaneously. Both multiparameter diagnostic models identified resolving pregnancies with positive predictive values \(\geq 95\%\). Their performances were not significantly better compared with individual progesterone levels which achieved a positive predictive value of 97\% using a cutoff level of 20 nmol/L.

Conclusion Serum progesterone measurement alone is as accurate as more complex diagnostic models for the prediction of successful expectant management in pregnancies of unknown location.

INTRODUCTION

Women with suspected early pregnancy problems, such as miscarriage or ectopic pregnancy, can be initially assessed using either biochemical tests \(^1\text{-}^3\) or by ultrasound \(^4\). Biochemical assessment is usually based on serial measurement of serum levels of human chorionic gonadotrophin (\(b\)-hCG) and progesterone \(^3\). However, biochemical assessment is often difficult to implement in busy units with a large workload. It requires expensive laboratory support and a need for good follow up on a large scale. With the use of transvaginal ultrasound, a normal intrauterine pregnancy can be identified at four weeks and three days in a woman with a regular 28 day cycle \(^5\). This compares favourably with transabdominal scanning, where the pregnancy cannot be routinely visualised before the sixth week of gestation \(^6\). Transvaginal sonography also enables conclusive diagnosis of miscarriages and ectopic pregnancies to be made in the majority of cases and has therefore been increasingly used as the method of choice for the initial assessment of women with suspected early pregnancy abnormalities \(^4\).

However, in 8\% to 31\% of women with suspected abnormal early pregnancies who are referred for assessment, the diagnosis cannot be made by ultrasound at the initial visit \(^7\text{-}^9\). In these women either surgical intervention or biochemical assessment are used to reach the correct diagnosis. There is preliminary evidence that expectant management is safe and successful in these cases with the majority of pregnancies resolving spontaneously \(^7\text{-}^9\). However, up to 29\% of these women will eventually require intervention because of worsening clinical symptoms or non-resolution of the pregnancy \(^7\).

Recently two diagnostic models were proposed to identify those pregnancies, which are likely to resolve spontaneously with expectant management. The first model, which used a combination of serum progesterone and serial \(b\)-hCG measurements, was described by Hahlín...
et al.⁷. In their study, the model predicted spontaneous resolution of pregnancy with a sensitivity of 73% and specificity of 97%. The second approach used a combination of clinical, ultrasound and two biochemical parameters in a logistic regression model⁹. At the initial visit the diagnosis of spontaneously resolving pregnancy was made with a sensitivity and specificity of 92%.

In this study we tested both diagnostic models prospectively on a representative sample of women with early pregnancies of unknown location. In addition we compared the performance of the models with their individual components for the prediction of spontaneous pregnancy resolution.

**METHODS**

This was a prospective observational study of pregnant women with suspected early pregnancy complications, who were referred for assessment by either their general practitioners or hospital consultants. Our dedicated early pregnancy assessment unit serves a mixed inner city population with a high level of socio-economic deprivation. All women had a positive urine pregnancy test (Clearview HCG II⁸, Unipath, Bedford, UK), a monoclonal antibody test, which according to the manufacturer’s specifications has a sensitivity of 99% at a urine β-hCG level > 25 iu/L. All women included in this study conceived spontaneously and were not taking exogenous progestogens.

A full medical history was documented and clinical examination carried out by the attending physician. A transvaginal ultrasound scan was then performed using a 5 MHz probe (Aloka SSD 550–2000, Aloka Co. Ltd., Tokyo, Japan). The diagnosis of a pregnancy of unknown location was made at the initial visit in all cases where there was no clear ultrasound evidence of an intrauterine pregnancy, retained products of conception or an ectopic pregnancy. In agreement with previous studies, this diagnosis was calculated using the logistic regression model⁹:

\[
\text{Probability of spontaneous resolution} = \frac{1}{1 + e^{-z}}
\]

where:

\[
z = -2.20 - 0.15 \times \text{progesterone (nmol/L)} + 3.36 \times \text{bleeding score} - 0.0013 \times \text{serum β-hCG (iu/L)} + 0.45 \times \text{endometrial thickness (mm)}
\]

The endometrial thickness was measured by ultrasound in the longitudinal plane at the point of maximum thickness. The absence or presence of bleeding was expressed as a bleeding score of 0 or 1. A probability score of > 70% was taken to be a positive prediction of a spontaneously resolving pregnancy, as it gave the best model performance in the original study⁹.

On the first follow up visit the relative daily β-hCG change ratio was calculated using the Hahlin formula⁷:

\[
\text{first β-hCG/sample interval in days}/\text{first β-hCG} \times 100
\]

According to Hahlin et al.⁷ a spontaneous resolution of pregnancy was diagnostic in cases with a β-hCG change ratio of < -5% and the initial serum progesterone level of < 20 nmol/L.

We calculated that 98 women needed to be recruited in order to detect a 20% difference between the sensitivities of both models with 80% power at the 5% level of certainty. The Student’s *t* test, Mann-Whitney *U* test, one way analysis of variance, Kruskal-Wallis test and $\chi^2$ test were used for statistical comparisons. A *P* value of $< 0.05$ was considered significant. A comparison of the predictive value of the individual parameters and the two models were made using receiver operating characteristic (ROC) curves using GraphROC for Windows™. This allowed the non-parametric statistical testing of ROC graphs using the method of comparing the areas under the curves¹⁰.

**RESULTS**

During a ten-month period we examined 2114 women. A normal intrauterine pregnancy was diagnosed in 1531 (73%) women, 410 (19%) had a miscarriage and 60 (3%) had an ectopic pregnancy. In 113 women (5%) the pregnancy was not identified at the initial visit. Data sets were incomplete in nine cases, which were excluded from further analysis. The indications for assessment in the remaining 104 women were: vaginal bleeding with or without abdominal pain in 79 (76%), abdominal pain with no bleeding in 15 (14%), high risk of ectopic pregnancy¹¹ in six (6%), pregnancy dating in three (3%) and previous history of miscarriage in one (1%).

At follow up 23 (22%) women were found to have a normal intrauterine pregnancy. Pregnancy resolution without the need for any intervention occurred in 72 (69%) cases. Therapeutic intervention was indicated in the remaining nine (9%) women. In eight cases serum $\beta$-hCG measurements were rising and one woman requested surgery because of worsening abdominal pain. At operation two women were found to have a miscarriage and seven had an ectopic pregnancy.

The average length of follow up was 9.0 days (range 2 to 57 days). A diagnosis of a normal pregnancy was made on average after 4.4 days (range: 2 to 10 days). In both women with miscarriages surgery was carried out after six days and in those with an ectopic pregnancy after a mean of 6.4 days (range 2 to 14 days). In cases of pregnancies that resolved spontaneously, complete resolution was reached after 10.8 days (range 2 to 57 days). None of the women suffered from excessive vaginal bleeding or required blood transfusion. There were no cases of clinically detectable infection and none of the ectopic pregnancies resulted in rupture.

Table 1. The distribution of measured parameters in terms of final outcome: *data distributed normally with values given as the mean and distribution to include 95% normal range [1.96 SD]; †data distributed non-parametrically with values given as the median [25th to the 75th interquartile range]; ‡discrete data given as median [range]; §discrete data given as number with the variable present and (%) with the feature for each final outcome. SR = spontaneous resolution; SVD = spontaneous vaginal delivery; LSCS = lower segment Caesarean section; PID = pelvic inflammatory disease.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR</th>
<th>Active treatment</th>
<th>Normal pregnancies</th>
<th>Significance of differences between outcome groups: $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age [yrs] *</td>
<td>31.0</td>
<td>27.5</td>
<td>31.1</td>
<td>24.7</td>
</tr>
<tr>
<td>Amnorrhea (days) *</td>
<td>53.4</td>
<td>34.0</td>
<td>39.3</td>
<td>39.8</td>
</tr>
<tr>
<td>Gravida *</td>
<td>3.2 [1-9]</td>
<td>2.0 [1-3]</td>
<td>3.6 [1-10]</td>
<td>2.8 [1.7]</td>
</tr>
<tr>
<td>Previous SVDs *</td>
<td>1 [0-6]</td>
<td>0 [0-0]</td>
<td>1.4 [0-5]</td>
<td>0.4 [0-2]</td>
</tr>
<tr>
<td>Previous abortions *</td>
<td>0.6 [0-6]</td>
<td>0 [0-0]</td>
<td>0.7 [0-3]</td>
<td>0.6 [0-4]</td>
</tr>
<tr>
<td>Previous miscarriages *</td>
<td>0.4 [0 - 3]</td>
<td>0.5 [0-1]</td>
<td>0.1 [0-1]</td>
<td>0.7 [0-3]</td>
</tr>
<tr>
<td>Previous LSCS *</td>
<td>0.1 [0-3]</td>
<td>0 [0-0]</td>
<td>0.1 [0-1]</td>
<td>0.1 [0-1]</td>
</tr>
<tr>
<td>Previous ectopics *</td>
<td>0.3 [0-1]</td>
<td>0.5 [0-1]</td>
<td>0.1 [0-1]</td>
<td>0.1 [0-1]</td>
</tr>
<tr>
<td>Pain **</td>
<td>50 (69)</td>
<td>2 (100)</td>
<td>2 (29)</td>
<td>15 (65)</td>
</tr>
<tr>
<td>Bleeding **</td>
<td>68 (94)</td>
<td>1 (50)</td>
<td>5 (71)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Endometrial thickness (mm) *</td>
<td>9.4 [1.2-17.6]</td>
<td>15.5 [14.4-16.5]</td>
<td>11.5 [1.0-18.4]</td>
<td>11.5 [5.2-17.8]</td>
</tr>
<tr>
<td>$\beta$-hCG (iu/L)**</td>
<td>320 [93-847]</td>
<td>139</td>
<td>811 [542-1025]</td>
<td>385 [297-582]</td>
</tr>
<tr>
<td>Progesterone [nmol/l]**</td>
<td>7 [5.0-12]</td>
<td>79</td>
<td>34 [9-61]</td>
<td>77 [66-98]</td>
</tr>
</tbody>
</table>

serum progesterone at a cut off value of 20 nmol/L achieved a positive predictive value > 95%. This was not significantly different from the performance of either model (Table 2).

DISCUSSION

In this study a transvaginal ultrasound scan performed at the initial visit enabled localisation of early pregnancy in 95% of cases. This confirms the efficacy of this method in the management of women with suspected early pregnancy abnormalities. We also confirmed that women with pregnancies of unknown location, who are clinically stable, could be safely managed expectantly with the majority of pregnancies resolving spontaneously. Only 9% of women eventually required surgical intervention which is significantly less than the 23% to 29% previously reported7,8. The reduction in the number of interventions may be partially explained by our increased experience in managing these cases. Furthermore, the investigators were not blinded to the results of diagnostic tests, which could have influenced the final outcomes.

The sensitivity and specificity of the logistic model in the detection of spontaneously resolving pregnancies were less than in our original study9. This is often the case when logistic models are tested prospectively. The design of the model is determined by the exact characteristics of the original data set, which will never be identical to any subsequent data set. There were three women with false positive diagnoses of a spontaneously resolving pregnancy, two of whom required intervention for an ectopic pregnancy. One false positive finding was a result of the increased endometrial thickness and the second was caused by a low progesterone level of 5 nmol/L. One woman, who required an intervention for a miscarriage, was incorrectly classified as having a spontaneously resolving pregnancy due to the presence of bleeding, low β-hCG and thick endometrium, despite high progesterone levels of 40 nmol/L.

In contrast, the sensitivity of the Hahlin’s model was better than originally reported7. When reviewing their original data, if the initial serum progesterone level < 20 nmol/L alone was taken to be diagnostic of a spontaneously resolving pregnancy, then the sensitivity would increase from 73% to 82%. However, there would be a consequent fall in specificity due to the presence of patients with low serum progesterone and rising β-hCG and thus requiring intervention. Few such patients were seen in this prospective study. Furthermore, in this prospective study only nine (9%) women required an active intervention, compared with 23 (29%) in Hahlin’s original publication. These differences in study populations may have contributed to the improved performance of Hahlin’s model on our data set.

In this prospective study both models performed

Table 2. Univariate analysis of prognostic predictors of spontaneous resolution and comparison with Hahlin’s and logistic regression model. SE = standard error; PPV = positive predictive values; NPV = negative predictive value.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Area under ROC curve</th>
<th>SE</th>
<th>P</th>
<th>Cut-off value</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bleeding</td>
<td>0.83</td>
<td>0.06</td>
<td>&lt; 0.01</td>
<td>0/1</td>
<td>0.94 (0.88 - 0.98)</td>
<td>0.72 (0.56 - 0.84)</td>
<td>0.88 (0.82 - 0.94)</td>
<td>0.85 (0.78 - 0.92)</td>
</tr>
<tr>
<td>Endometrial thickness</td>
<td>0.67</td>
<td>0.06</td>
<td>0.01</td>
<td>11.0</td>
<td>0.61 (0.51 - 0.71)</td>
<td>0.66 (0.50 - 0.79)</td>
<td>0.80 (0.73 - 0.88)</td>
<td>0.43 (0.34 - 0.53)</td>
</tr>
<tr>
<td>β-hCG</td>
<td>0.47</td>
<td>0.06</td>
<td>0.48</td>
<td>295 iu/l</td>
<td>0.41 (0.32 - 0.52)</td>
<td>0.63 (0.46 - 0.77)</td>
<td>0.71 (0.67 - 0.75)</td>
<td>0.32 (0.28 - 0.36)</td>
</tr>
<tr>
<td>Progesterone</td>
<td>0.95</td>
<td>0.04</td>
<td>&lt; 0.01</td>
<td>20 nmol/l</td>
<td>0.93 (0.86 - 0.97)</td>
<td>0.94 (0.81 - 0.99)</td>
<td>0.97 (0.94 - 0.97)</td>
<td>0.86 (0.85 - 0.91)</td>
</tr>
<tr>
<td>Logistic model</td>
<td>0.93</td>
<td>0.03</td>
<td>&lt; 0.01</td>
<td>70%</td>
<td>0.82 (0.72 - 0.88)</td>
<td>0.91 (0.77 - 0.97)</td>
<td>0.95 (0.93 - 0.97)</td>
<td>0.69 (0.65 - 0.73)</td>
</tr>
<tr>
<td>Hahlin’s model</td>
<td>0.92</td>
<td>0.04</td>
<td>&lt; 0.01</td>
<td>hCG change &lt; -5% progesterone &lt;20 nmol/L</td>
<td>0.88 (0.79 - 0.93)</td>
<td>0.97 (0.85 - 1.00)</td>
<td>0.98 (0.96 - 0.99)</td>
<td>0.78 (0.74 - 0.82)</td>
</tr>
</tbody>
</table>

equally well by comparing the areas under their ROC curves. The prediction of spontaneous pregnancy resolution based on initial serum progesterone alone was as accurate as prediction based on more complex multiparameter models (Fig. 1). Using a cutoff level of 20 nmol/L, serum progesterone measurements achieved a sensitivity of 93% and specificity of 94%. The progesterone level was a component of both diagnostic models, and this result suggests that the other variables used did not significantly add to their accuracy when tested prospectively. A low serum progesterone alone may therefore be used to predict which pregnancies of unknown location will resolve without the need for an intervention. As the test specificity is not 100% it would still be necessary to follow up this group of women and observe a quantitative fall in the levels of β-hCG. This could be done by measuring serum β-hCG at a single additional visit seven days later.

It has already been proposed that a low initial serum progesterone level (<15.9 nmol/L, i.e. 5 ng/L) identifies non-viable pregnancies and enables the use of diagnostic curettage to try and distinguish between failed intrauterine and extrauterine pregnancies. Although we agree that low progesterone can be used to exclude a viable intrauterine pregnancy, it is not a useful test to differentiate between miscarriage and ectopic pregnancy. If we applied the concept of diagnostic curettage for women with a progesterone of <15.9 nmol/L in our group of women with pregnancies of unknown location, only two out of the seven ectopic pregnancies would be identified (sensitivity 29%). In 58 out of 72 (81%) pregnancies which resolved spontaneously, curettage would have been performed unnecessarily. The positive predictive value of this protocol for the diagnosis of ectopic pregnancy would thus be only 3%.

Another problem lies in the assumption that a positive diagnosis of an ectopic pregnancy can be made in all cases with absent chorionic villi at the curettage. In a study of 272 women with suspected miscarriages who underwent diagnostic curettage chorionic villi were not identified in 20% of cases. In this group of women only one-third were found to have an ectopic pregnancy on subsequent laparoscopy. These findings demonstrate the limited value of curettage for the diagnosis of ectopic pregnancy.

Serum progesterone has also been used to differentiate between normal intrauterine and abnormal pregnancies. In our study all normal intrauterine pregnancies but one had an initial serum progesterone level > 60 nmol/L. Such a high level can be used to predict the probability of a normal pregnancy with a sensitivity of 96%, specificity of 95% and positive predictive value of 85% which is consistent with previous reports. A high progesterone level reflects a normally functioning corpus luteum and placenta of a viable pregnancy. However, Shepherd and colleagues have shown that 10% of ectopic pregnancies may be viable and in such cases the serum progesterone levels may also be high. Two out of seven (29%) of ectopic pregnancies in our study presented with a serum progesterone level > 60 nmol/L.

Cacciatore et al. first suggested that at a serum β-hCG level of above 1000 iu/L a normal intrauterine pregnancy should always be visualised by transvaginal scan. Therefore in a pregnancy of unknown location with a serum β-hCG > 1000 iu/L, a normal intrauterine pregnancy may be excluded. In addition, if the serum progesterone is high (>60 nmol/L), a spontaneously resolving pregnancy is unlikely and therefore an intervention should be considered. If the serum β-hCG is <1000 iu/L and the serum progesterone level is high the optimum time to exclude an ectopic pregnancy is to repeat the ultrasound scan when the serum β-hCG is expected to reach the level of 1000 iu/L. This can be calculated based on a normal serum β-hCG doubling time of 48 hours.

In this study nine women (9%) with a pregnancy of unknown location had initial serum progesterone between 20 nmol/L to 60 nmol/L. Three women in this group (33%) had an ectopic pregnancy which required surgery. In comparison, only 3% of women with progesterone <20 nmol/L and 8% of those with progesterone > 60 nmol/L had an ectopic pregnancy. Therefore in our study mid-range progesterone levels indicated a high risk of ectopic and careful follow with quantitative serial hCG levels at 48-hour intervals is required in these cases.

It should be stressed, however, that serum progesterone measurements may be confounded by the presence of multiple corpora lutea or the use of exogenous progesterogens which may reduce its value in the setting of an assisted conception unit. Further work is required to define appropriate management strategies in women undergoing fertility treatment.

CONCLUSION

The result of this prospective study indicates that the recently developed complex multiparameter models are not better than single serum progesterone measurement for the assessment of women with pregnancies of unknown location. Serum progesterone measurement may be used to identify spontaneously resolving pregnancies and formulate an effective and safe management strategy at the initial visit. This aids patient counselling, helps to reduce the number of follow up visits and reduce the need for surgical interventions, with no serious adverse outcomes.

References

2. Stovall TG, Ling FW, Cope BJ, Buster JE. Preventing ruptured ectopic


Accepted 12 May 2000