

Protocols for prediction of ectopic pregnancy in pregnancy of unknown location:a systematic review and meta-analysis

Zhuo Chen¹, Siyu Yang¹, Botao Sun¹, Wenqing Yang¹, and Yu Zhang¹

¹Xiangya Hospital Central South University

August 7, 2021

Abstract

Background: The ectopic pregnancy(EP) patients requires the closely monitor. However, there is no international consensus through which method to select EP patients from the pregnancy of unknown location (PUL) patients. **Objective:** To summarize and review the protocols of screening patients with EP when being diagnosed with PUL. **Search strategy:** We searched MEDLINE, web of science and Embase from inception to May 2020. All the articles were dual-reviewed based on predetermined selection criteria. **Selection criteria:** Studies exploring PUL outcomes can be included. **Data collection and analysis:** Prediction results, final diagnosis, and expense depending on patient's visits and examination were analyzed by R. version 3.6.3 and Revman version 5.4. **Results:** 29 studies were included. M6 model had the areas under the curve(AUC) of 0.944, the progesterone cut-offs an AUC of 0.725, and the M4 model an AUC of 0.871 respectively. When the sum of visits and examinations of the protocols increased from 3 to 5, the rate of lost to follow-up patients increased from 11.19% to 18.63%. The average production utility of progesterone cut-offs is 0.242, the M4 is 0.174, and the M6 is 0.157. **Conclusions:** The M6 model had the best performance to predict EP among the PUL patients. The progesterone cut-offs is the most cost-effective method to predict the final outcome of EP. **Fundings:** Project supported by National Natural Science Foundation of China (82073323) and the Joint Funds of the National Natural Science Foundation of China (U20A20368). **Keywords:** ectopic pregnancy, pregnancy of unknown location, predictive protocol,hCG,progesterone.

Introduction

Ectopic pregnancy (EP) occurs when fertilized ovum locate outside the uterine cavity, and its incidence rate is about 1.5% to 2% in pregnancy¹. The death toll of ectopic pregnancy can account for more than 6% of maternal mortality, which is the leading cause of maternal death in early pregnancy^{2, 3}. According to a research in Scotland⁴, nearly half of the patients had to visit the clinicians more than three times to be diagnosed for ectopic pregnancy. Moreover, it is estimated that 136, 400 pounds sterling has been spend yearly on the diagnosis of EP in pregnancy of unknown location(PUL). PUL refers to pregnancy test positive, but the first transvaginal ultrasound (TVS) has not found evidence of pregnancy in both intrauterine and extrauterine, accounting for about 5% to 42% of all early pregnant women⁵. Among them, ectopic pregnancy, accounted for 6% to 20% of all PUL patients,required surgery or chemotherapy and had relatively higher mortality comparing with the rest of PUL patients.

The elevated serum hCG and progesterone are the most widely used indications for early pregnancy. Some protocols use a single biomarker to distinguish patients with different outcomes by setting certain threshold, while others use established mathematical formulas. In European countries and regions, the logistic regression model M4, established by George Condous et al⁶, is now the most widely used formula for predicting the outcomes of PUL. But the sensitivity for EP was too low for the model to be used in clinical practice in the US⁷. This suggests that different definitions of pathology and disease management processes may lead to distinct accuracy of prediction protocols in specific regions. Thus, there is no international consensus on how to manage women with a pregnancy of PUL. In this study, we will plot the SROC curve and calculate

the area under the curve(AUC)by meta-analysis of studies, jointly contrasting the sensitivity and specificity of individual protocols. We will introduce the concepts of average production utility and marginal benefit assessing the feasibility of individual protocols, and attempt to describe the impact of the complexity of the protocols.

Method

Protocol and registration

The search strategies, inclusion and exclusion criteria, outcomes measures, quality assessment and strategy for data synthesis were specified in advance and documented in the study protocol, and registered in PROSPERO (<https://www.crd.york.ac.uk/prospero/>, with registration number CRD42021262155). The systematic review was carried out in accordance with the PRISMA checklist.

Inclusion and exclusion criteria of studies

We used “pregnancy of unknown location[Title]” as free text search strings to search all studies diagnosed as PUL in MEDLINE, web of science and Embase at the first visit. All studies exploring PUL outcomes can be included. Other studies related to pregnancy diagnosis were excluded. The eligibility criteria were as follows: (a) the study population was adult patients diagnosed as PUL; (b) the diagnosis of ectopic pregnancy was confirmed by ultrasound, non-visualized treatment (increase in hCG after uterine evacuation) or surgery; and (c) data regarding the true-positive (TP), false-positive (FP), false-negative (FN) and true-negative (TN) rates can be found either directly or indirectly in all studies. All prospective or retrospective randomized controlled trials, case-control studies, cohort studies and observational studies met the inclusion. All reviews, meta-analyses, poster or oral meeting presentation, case reports, letters, responses or comments, and non-English written studies were excluded. The studies before 1984 was excluded, because transvaginal ultrasound had not been a routine test for early pregnancy evaluation before that⁸.

Selection of primary studies and data extraction

After retrieving the references, the titles and abstracts were screened in duplicate by two reviewers. Once it is determined that relevant studies may be included, the full text will be screened. If there was any doubt about eligibility, a third reviewer will be the arbitrator. Upon completion of the above work, two reviewers used a tested data extraction form to abstract data independently. This included the characteristics of the references, the year of publication, first author, study period, definition of PUL and the outcomes, experimental design, study methodology, management protocol, and final clinical results.

Risk of bias assessment

All evaluations of the studies were reflected in the results of the QUADAS-2⁹ by two reviewers. When no consensus can be reached in the assessment of the risk of bias in particular studies, the third reviewer will be introduced as the arbitrator.

Outcomes of interest

In the present study, the main outcome was the final diagnosis of the EP and persistent pregnancy of unknown location. In the context of PUL, the outcome can either be: [1] ectopic pregnancy (EP), [2] failed PUL (FPUL) or [3] an intrauterine pregnancy (IUP)¹⁰. Persistent PUL (PPUL) is considered as a high-risk outcome, counted together with the outcome of ectopic pregnancy. To assess the feasibility of the tests, the secondary outcome was the expense depending on the numbers of patient’s visits and the examination.

Statistical analysis

Paired forest plots of sensitivity and specificity with corresponding 95% confidence intervals were produced for each protocol measuring the main outcome of EP. Considering the potential threshold effect, we decided to use summary receiver operating characteristic(SROC) curves to evaluate the accuracy of the model. The area under the ROC curve(AUC) was calculated by numerical integration of the SROC curve area. All analyses were performed using R. version 3.6.3 and Revman version 5.4.

Results

Study selection

By searching MEDLINE, SCIE and EMBASE, 219 articles were found after eliminating duplicate citations. After reading the full text, ninety-one citations were retained. Next, the following citations were excluded: thirty-nine were not intended to describe and predict the outcome of PUL, twelve had insufficient data integrity to support the analysis, nine did not have high-risk category (i.e., EP or PPUL), and two did not propose specific prediction protocols. The flow of paper retrieval is summarized in figure 1. As a result, twenty nine studies were included.^{7, 10-37} Among them, three used M1 model, four used single progesterone threshold protocol, two used M6 model, nine used M4 model, one used P1 model, five used hCG ratio (hCG 48 hours / hCG at 0 hours) protocol, two used the two steps method, the combination of progesterone and hCG threshold protocol, seven used single hCG cut-off protocol, one used FFN protocol, one used ultrasound imaging results and hCG, one used ultrasound imaging results and one used Activin-A protocol.

Characteristics of Included Studies

Table S1 summarizes twenty-nine studies. PPUL included in most studies included was not considered as a separate classification, so our meta-analysis didn't separate it as a specific type. Since we did not analyze the performance of predicting the low-risk outcomes (IUP and FPUL), we assumed that the heterogeneity of their definition do not affect the final outcomes of the analysis.

Risk of bias within studies

QUADAS-2 criteria was used for assessing the risk of bias (figure 2). The reference standard refers to the final clinical outcomes of all patients. Reference standard always remains high risk in risk of bias, for it has to be included in differential verification bias. For example, the reference standard of EP could be ultrasound, non-visualized treatment or surgery, whereas that for an persistent pregnancy of unknown location(PPUL) could be non-declining hCG and a non-confirmed location of the pregnancy that were confirmed neither by TVS nor by laparoscopy between four consecutive measurements, or classified at D7 if the hCG was plateauing by D7 with no visible pregnancy on repeat TVS.

Diagnostic performance of prediction protocols for EP in PUL

A summary of the data reported in the twenty-nine studies and paired forest plots of sensitivity and specificity with corresponding 95% confidence intervals is presented in table S2. The summary of sensitivity and specificity of each study for the high-risk outcome is represented in Table 1. Figure 3 shows the summary receiver operating characteristic(SROC) curves for the protocol performances of M1, single hCG cut-offs, progesterone cut-offs, M4, M6 and hCG ratio respectively. As summarized in figure 3, M1 had an AUC of 0.907, the progesterone cut-offs an AUC of 0.725, the hCG ratio an AUC of 0.824, the hCG cut-offs an AUC of 0.586, the M4 an AUC of 0.871, and the M6 an AUC of 0.944 respectively. Pooled sensitivity, specificity, positive prediction value (PPV) and negative prediction value(NPV) data for the prediction protocols are also summarized in table 1.

Feasibility of prediction protocols for EP in PUL

In order to evaluate the impact of the complexity of each protocol on clinical work efficiency, we propose to introduce the concept of “the rate of lost”, which has never been put forward before. The rate of lost refers to the ratio of the number of patients who have not received all of the diagnostic approaches included in protocols before follow-up, management or final diagnosis to the number of patients with complete data. Table 2 shows the number of examination items and visits required by each protocol. When it comes to the relationship between the sum of the number of visits and the number of inspection items and the rate of lost, table 2 shows that the average rate of lost in each category. The average rate of lost is 11.19% (95%CI 4.67-17.72) for the sum of 3, 18.63% (95%CI 9.67-17.71) for the sum of 5 and 9.29% (95%CI cannot be counted) for the sum of 6.

We also introduce average production utility, an economic concept, to quantify the cost of improving prediction accuracy. It was used to evaluate the effectiveness of each protocol under the expenditure of unit cost. Considering the different economic development level, medical insurance policies and cultural concepts of different countries and regions, the sum of the number of visits and the number of examinations required by the protocols were used to define the total cost, and AUC to define the total benefit. Table 3 shows the average production utility of the protocols. Among them, M1 is 0.181, progesterone cut-off is 0.242, M4 is 0.174, hCG ratio is 0.165, hCG cut-off is 0.195, and M6 is 0.157.

Discussion

Main findings

Laparoscopy was the gold standard for the diagnosis of ectopic pregnancy for a long period of time since 1993³⁸. However, it is reported that 3.0% ~ 4.5% of EP patients failed to diagnose EP at the time of initial laparoscopic examination^{39, 40}. As a result, laparoscopy is no longer considered as the gold standard for the diagnosis of ectopic pregnancy according to the 2016 Royal College of Obstetricians and Gynaecologists and Association of Early Pregnancy Units (RCOG/AEPU) joint guidelines on diagnosis and management of ectopic pregnancy⁴¹. Currently, there is a lack of effective means for early diagnosis of ectopic pregnancy, because it was difficult to distinguish from threatened or inevitable abortion, intrauterine pregnancy with corpus luteum rupture and intrauterine and extrauterine pregnancy. The American College of Obstetricians and Gynaecologists (ACOG) and RCOG/AEPU joint guidelines recommend that transvaginal ultrasound is the first choice for the diagnosis of EP^{41, 42}. However, ultrasound examination is also affected by equipment, pathway, physician's operation, pregnant women's obesity, combined with uterine fibroids or ovarian tumors. 8% to 31% of the early pregnant site cannot be diagnosed at the first ultrasound examination¹⁹.

PUL was a temporary status with various outcomes. EP is the high risk outcome of PUL, because it would cause internal hemorrhage and endanger patients' lives¹. The high-risk outcomes as EP in PUL patients need to be screened out before the occurrence of rupture and internal hemorrhage⁴³. Precise prediction of PUL outcome can provide not only timely and correct management protocols for EP, but also reduce medical burden and unnecessary medical intervention for IUP and FPUL patients.

The M4 model is currently one of the most widely used prediction models, especially in European countries like Britain. In 2016, Ben Van Calster et al²⁶ proposed the M6 model, which introduced both progesterone and hCG values(0h and 48h)as variables in anticipation of greater predictive accuracy, because the presence of low serum progesterone concentrations in patients with EP has been known since the late 1970s⁴⁴. A meta-analysis by S Bobdiwala et al. in 2019 showed that the areas under the curves (95% CI) of hCG cut-offs, hCG ratio (0/48h), progesterone cut-offs and M4 model were 0.42 (0.00-0.99), 0.69 (0.57-0.78), 0.69 (0.54-0.81) and 0.87 (0.83-0.91), respectively. The prediction accuracy of the model with hCG at 0 h and 48 h was higher than that with single hCG value, and the model with single progesterone value had higher accuracy than that with single HCG value, which is the same trend as our statistics. Moreover, through our systematic meta-analysis of all published prediction protocols of ectopic pregnancy outcome in PUL, we found that, consistent with previous studies⁴⁵, M4 model has better prediction accuracy. As for the latest protocol, M6 also showed a trend of higher prediction accuracy, reaching the AUC of 0.94. It is worth noting that hCG ratio and progesterone cut-offs also showed good sensitivity and specificity in the scheme using a single biomarker or single biomarker detection point, with AUC of 0.82 and 0.72, respectively. In view of the short time required, few testing items, fast guidance for clinicians (especially in developing countries and regions), and reduced cost for patients, they still have certain practical value. Most studies focusing on PUL outcome prediction models have been conducted in European and North American countries and regions with abundant medical resources. There often were well-established protocols for early pregnancy diagnosis, but fewer studies had been conducted on PUL outcome prediction in low-income countries and areas. In addition, predicting EP in PUL patients by M4 in the United Kingdom and United States, Barnhart K. T. et al.⁷ found that even after adjusting the diagnostic criteria to a consistent level in the two countries, the sensitivity of the model differed. The study revealed that the PUL outcome prediction protocols were related to the database used, or the different medical guidelines and the levels of healthcare organization.

In the past, the acceptability of each prediction protocol has hardly been evaluated, so we propose to use average production utility to evaluate the cost performance of the protocols. Due to the differences of medical charges, medical insurance policies and health policies in different countries and regions, and the differences of medical development level, we use the sum of the number of visits and the number of inspection items to replace the medical cost, and the number of inspection items is defined as the minimum number of inspections that can be used to predict the outcome of the protocols. It is reasonable to assume the numbers of tests the protocols required can reflect the medical cost. In addition, more visits and examination items, and more complex prediction protocols could likely lead to follow-up losing. Therefore, we believe that the data loss caused by the above reasons reflects the acceptability of each protocol to a certain extent. We propose a new evaluation method. Table 2 shows the relationship between the sum of the number of visits and the number of inspection items and the rate of lost. Previously, we expected that as the protocol took longer and the number of examinations and visits increased, more patients might not be able to use the protocols because of medical related payment pressure and severe clinical symptoms, which may lead to the loss of follow-up and the lack of timely diagnosis and treatment of EP patients. However, different from our expectation, the rate of lost did not increase with the number of examinations and visits required by the protocols. When the time required for the protocols was extended from one day to two days, the average rate of lost increased from 11.56% (95% CI 6.96% - 16.16%) to 17.46% (95% CI 11.46% - 23.46%). Although there was a certain growth trend, there was no statistical significance. This may be due to the following reasons: First, the earliest time of all 29 studies can be traced back to 1991, and the latest time is 2018. Over the past 20 years, great changes may have taken place in medical policy, popularization of medical science knowledge and national economic level, which may affect the willingness of patients to follow up and the ability of medical institutions to track and manage patients; Second, some protocols (such as P1, M1, etc.) have not been studied extensively, which may cause large bias. In fact, after excluding the protocols with less than or equal to 2 studies, the average rate of lost increased from 11.19% (95% CI 4.67-17.72) to 18.63% (95% CI 9.67-17.71) when the sum of visits and examinations changed from 3 to 5. Although there was no statistical significance, the trend was the same as our prediction.

Besides that, table 3 shows that the complicated protocols to improve the accuracy of prediction also needs higher cost. The average production utility of M4 model, which requires at least 2 visits and 3 examinations, is higher than that of M6 model requiring at least 3 visits and 3 examinations, and lower than the hCG cutoffs model and progesterone-cutoffs model which only need one visit and one inspection. When this trend is reflected in clinical work, it seems that more complex prediction schemes may bring higher costs. Simple prediction protocols still have certain application value in low-income countries and regions.

Strengths and Limitations

For the first time, we evaluated the acceptability of protocols for predicting the outcome of ectopic pregnancy in patients with PUL, which has implications for the application of these protocols to practical clinical work. We evaluated for the first time the accuracy of the M6 model in predicting ectopic pregnancy. However, the studies were inconsistent in terms of the definition of EP, the definition of PUL and how to present the results of data analysis, and the threshold of protocols were not identical, which may result in heterogeneity between studies.

Interpretations

All PUL prediction protocols are not diagnostic and may misclassify patients. They can guide clinicians to increase the follow-up frequency of high-risk patients, but should not be used to choose interventions. Unnecessary intervention may result in harm to an intrauterine pregnancy. In addition, the prediction accuracy of some protocols are unstable in different regions.⁷ Based on this, it is suggested that clinicians be cautious about the prediction results of any protocol, and comprehensively evaluate the patients in combination with clinical manifestations (such as abdominal pain or signs and symptoms of hemodynamic compromise).⁵

Conclusion

this systematic review and meta-analysis summarized the current outcome-predictive protocols for PUL. It suggests that M6 has considerable accuracy in diagnosis of EP in PUL. However, there are few studies of M6, and the negative impact of more complex prediction steps on clinical work is unknown. Using a single biomarker or a single type of biomarker, such as progesterone threshold method, has a higher average production utility, which indicates that it is more cost-effective.

Contribution to authorship

ZC: Methodology, validation, data curation, writing-original draft. SY: Formal analysis, data curation, software. BS: Software, visualization. WY: Conceptualization, validation, writing-review and editing, supervision. YZ: Conceptualization, supervision, project administration, funding acquisition. All the authors approved to publish the final version of this paper.

Conflict of Interest

The authors declare that they have no conflict of interest.

Supporting Information

Additional supporting information may be found online.

Table S1. Details of the included studies.

Table S2. Ectopic pregnancy data reported in studies and paired forest plots of sensitivity and specificity with corresponding 95% confidence intervals.

Fundings

This project was supported by National Natural Science Foundation of China (82073323) and the Joint Funds of the National Natural Science Foundation of China (U20A20368).

Reference

1. Chang, J.; Elam-Evans, L. D.; Berg, C. J.; Herndon, J.; Flowers, L.; Seed, K. A.; Syverson, C. J., Pregnancy-related mortality surveillance—United States, 1991–1999. *MMWR Surveill Summ* **2003**, *52* (2), 1-8.
2. Farquhar, C. M., Ectopic pregnancy. *Lancet* **2005**, *366* (9485), 583-91.
3. Alkatout, I.; Honemeyer, U.; Strauss, A.; Tinelli, A.; Malvasi, A.; Jonat, W.; Mettler, L.; Schollmeyer, T., Clinical diagnosis and treatment of ectopic pregnancy. *Obstet Gynecol Surv* **2013**, *68* (8), 571-81.
4. Wedderburn, C. J.; Warner, P.; Graham, B.; Duncan, W. C.; Critchley, H. O.; Horne, A. W., Economic evaluation of diagnosing and excluding ectopic pregnancy. *Hum Reprod* **2010**, *25* (2), 328-33.
5. Kirk, E.; Bottomley, C.; Bourne, T., Diagnosing ectopic pregnancy and current concepts in the management of pregnancy of unknown location. *Hum Reprod Update* **2014**, *20* (2), 250-61.
6. Condous, G.; Van Calster, B.; Kirk, E.; Haider, Z.; Timmerman, D.; Van Huffel, S.; Bourne, T., Prediction of ectopic pregnancy in women with a pregnancy of unknown location. *Ultrasound Obstet Gynecol* **2007**, *29* (6), 680-7.
7. Barnhart, K. T.; Sammel, M. D.; Appleby, D.; Rausch, M.; Molinaro, T.; Van Calster, B.; Kirk, E.; Condous, G.; Van Huffel, S.; Timmerman, D.; Bourne, T., Does a prediction model for pregnancy of unknown location developed in the UK validate on a US population? *Human Reproduction* **2010**, *25* (10), 2434-2440.
8. Schwimer, S. R.; Lebovic, J., Transvaginal pelvic ultrasonography. *J Ultrasound Med* **1984**, *3* (8), 381-3.

9. Whiting, P. F.; Rutjes, A. W.; Westwood, M. E.; Mallett, S.; Deeks, J. J.; Reitsma, J. B.; Leeflang, M. M.; Sterne, J. A.; Bossuyt, P. M., QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* **2011**, *155* (8), 529-36.
10. Bignardi, T.; Condous, G.; Alhamdan, D.; Kirk, E.; Van Calster, B.; Van Huffel, S.; Timmerman, D.; Bourne, T., The hCG ratio can predict the ultimate viability of the intrauterine pregnancies of uncertain viability in the pregnancy of unknown location population. *Human Reproduction* **2008**, *23* (9), 1964-1967.
11. Bobdiwala, S.; Christodoulou, E.; Farren, J.; Mitchell-Jones, N.; Kyriacou, C.; Al-Memar, M.; Ayim, F.; Chohan, B.; Kirk, E.; Abughazza, O.; Guruwadahyarhalli, B.; Guha, S.; Vathanan, V.; Bottomley, C.; Gould, D.; Stalder, C.; Timmerman, D.; van Calster, B.; Bourne, T., Triaging women with pregnancy of unknown location using two-step protocol including M6 model: clinical implementation study. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* **2020**, *55* (1), 105-114.
12. Bobdiwala, S.; Guha, S.; Van Calster, B.; Ayim, F.; Mitchell-Jones, N.; Al-Memar, M.; Mitchell, H.; Stalder, C.; Bottomley, C.; Kothari, A.; Timmerman, D.; Bourne, T., The clinical performance of the M4 decision support model to triage women with a pregnancy of unknown location as at low or high risk of complications. *Human Reproduction* **2016**, *31* (7), 1425-1435.
13. Condous, G.; Kirk, E.; Lu, C.; Van Huffel, S.; Gevaert, O.; De Moor, B.; De Smet, F.; Timmerman, D.; Bourne, T., Diagnostic accuracy of varying discriminatory zones for the prediction of ectopic pregnancy in women with a pregnancy of unknown location. *Ultrasound in Obstetrics and Gynecology* **2005**, *26* (7), 770-775.
14. 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. *Human Reproduction* **2005**, *20* (5), 1398-1403.
15. 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. *Human Reproduction* **2004**, *19* (8), 1900-1910.
16. Condous, G.; Van Calster, B.; Kirk, E.; Haider, Z.; Timmerman, D.; Van Huffel, S.; Bourne, T., Clinical information does not improve the performance of mathematical models in predicting the outcome of pregnancies of unknown location. *Fertility and Sterility* **2007**, *88* (3), 572-580.
17. Condous, G.; Van Calster, B.; Kirk, E.; Haider, Z.; Timmerman, D.; Van Huffel, S.; Bourne, T., Prediction of ectopic pregnancy in women with a pregnancy of unknown location. *Ultrasound in Obstetrics and Gynecology* **2007**, *29* (6), 680-687.
18. Day, A.; Sawyer, E.; Mavrelou, D.; Taylor, A.; Helmy, S.; Jurkovic, D., Use of serum progesterone measurements to reduce need for follow-up in women with pregnancies of unknown location. *Ultrasound in Obstetrics and Gynecology* **2009**, *33* (6), 704-710.
19. Fistouris, J.; Bergh, C.; Strandell, A., Classification of pregnancies of unknown location according to four different hCG-based protocols. *Human Reproduction* **2016**, *31* (10), 2203-2211.
20. Guha, S.; Ayim, F.; Ludlow, J.; Sayasneh, A.; Condous, G.; Kirk, E.; Stalder, C.; Timmerman, D.; Bourne, T.; Van Calster, B., Triaging pregnancies of unknown location: The performance of protocols based on single serum progesterone or repeated serum hCG levels. *Human Reproduction* **2014**, *29* (5), 938-945.
21. Kirk, E.; Condous, G.; Haider, Z.; Lu, C.; Van Huffel, S.; Timmerman, D.; Bourne, T., The practical application of a mathematical model to predict the outcome of pregnancies of unknown location. *Ultrasound in Obstetrics and Gynecology* **2006**, *27*(3), 311-315.
22. Kirk, E.; Condous, G.; Van Calster, B.; Van Huffel, S.; Timmerman, D.; Bourne, T., Rationalizing the follow-up of pregnancies of unknown location. *Human Reproduction* **2007**, *22* (6), 1744-1750.

23. Morse, C. B.; Sammel, M. D.; Shaunik, A.; Allen-Taylor, L.; Oberfoell, N. L.; Takacs, P.; Chung, K.; Barnhart, K. T., Performance of human chorionic gonadotropin curves in women at risk for ectopic pregnancy: exceptions to the rules. *Fertil Steril* **2012**, *97* (1), 101-6.e2.
24. Nadim, B.; Leonardi, M.; Infante, F.; Lattouf, I.; Reid, S.; Condous, G., Rationalizing the management of pregnancies of unknown location: Diagnostic accuracy of human chorionic gonadotropin ratio-based decision tree compared with the risk prediction model M4. *Acta Obstetrica et Gynecologica Scandinavica* **2020**, *99* (3), 381-390.
25. Van Calster, B.; Abdallah, Y.; Guha, S.; Kirk, E.; Van Hoorde, K.; Condous, G.; Preisler, J.; Hoo, W.; Stalder, C.; Bottomley, C.; Timmerman, D.; Bourne, T., Rationalizing the management of pregnancies of unknown location: Temporal and external validation of a risk prediction model on 1962 pregnancies. *Human Reproduction* **2013**, *28* (3), 609-616.
26. Van Calster, B.; Bobdiwala, S.; Guha, S.; Van Hoorde, K.; Al-Memar, M.; Harvey, R.; Farren, J.; Kirk, E.; Condous, G.; Sur, S.; Stalder, C.; Timmerman, D.; Bourne, T., Managing pregnancy of unknown location based on initial serum progesterone and serial serum hCG levels: development and validation of a two-step triage protocol. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* **2016**, *48* (5), 642-649.
27. Zee, J.; Sammel, M. D.; Chung, K.; Takacs, P.; Bourne, T.; Barnhart, K. T., Ectopic pregnancy prediction in women with a pregnancy of unknown location: Data beyond 48 h are necessary. *Human Reproduction* **2014**, *29* (3), 441-447.
28. Dart, R.; Howard, K., Subclassification of indeterminate pelvic ultrasonograms: stratifying the risk of ectopic pregnancy. *Acad Emerg Med* **1998**, *5* (4), 313-9.
29. Dart, R.; Kaplan, B.; Ortiz, L.; Cloherty, J.; Lavoie, T., Normal intrauterine pregnancy is unlikely in emergency department patients with either menstrual days > 38 days or beta-hCG > 3,000 mIU/mL, but without a gestational sac on ultrasonography. *Acad Emerg Med* **1997**, *4* (10), 967-71.
30. Dart, R. G.; Burke, G.; Dart, L., Subclassification of indeterminate pelvic ultrasonography: prospective evaluation of the risk of ectopic pregnancy. *Ann Emerg Med* **2002**, *39* (4), 382-8.
31. Dart, R. G.; Dart, L.; Mitchell, P.; Berty, C., The predictive value of endometrial stripe thickness in patients with suspected ectopic pregnancy who have an empty uterus at ultrasonography. *Acad Emerg Med* **1999**, *6* (6), 602-8.
32. Dart, R. G.; Mitterando, J.; Dart, L. M., Rate of change of serial beta-human chorionic gonadotropin values as a predictor of ectopic pregnancy in patients with indeterminate transvaginal ultrasound findings. *Ann Emerg Med* **1999**, *34* (6), 703-10.
33. El Bishry, G.; Ganta, S., The role of single serum progesterone measurement in conjunction with beta hCG in the management of suspected ectopic pregnancy. *J Obstet Gynaecol* **2008**, *28*(4), 413-7.
34. Florio, P.; Severi, F. M.; Bocchi, C.; Luisi, S.; Mazzini, M.; Danero, S.; Torricelli, M.; Petraglia, F., Single serum activin a testing to predict ectopic pregnancy. *J Clin Endocrinol Metab* **2007**, *92* (5), 1748-53.
35. Kaplan, B. C.; Dart, R. G.; Moskos, M.; Kuligowska, E.; Chun, B.; Adel Hamid, M.; Northern, K.; Schmidt, J.; Kharwadkar, A., Ectopic pregnancy: prospective study with improved diagnostic accuracy. *Ann Emerg Med* **1996**, *28* (1), 10-7.
36. Mol, B. W.; Hajenius, P. J.; Engelsbel, S.; Ankum, W. M.; Van der Veen, F.; Hemrika, D. J.; Bossuyt, P. M., Serum human chorionic gonadotropin measurement in the diagnosis of ectopic pregnancy when transvaginal sonography is inconclusive. *Fertil Steril* **1998**, *70* (5), 972-81.
37. Nowacek, G. E.; Meyer, W. R.; McMahon, M. J.; Thorp, J. R.; Wells, S. R., Diagnostic value of cervical fetal fibronectin in detecting extrauterine pregnancy. *Fertil Steril* **1999**, *72*(2), 302-4.

38. Ankum, W. M.; Van der Veen, F.; Hamerlynck, J. V.; Lammes, F. B., Laparoscopy: a dispensable tool in the diagnosis of ectopic pregnancy? *Hum Reprod* **1993**, *8* (8), 1301-6.

39. Li, T. C.; Tristram, A.; Hill, A. S.; Cooke, I. D., A review of 254 ectopic pregnancies in a teaching hospital in the Trent Region, 1977-1990. *Hum Reprod* **1991**, *6* (7), 1002-7.

40. Atri, M.; Leduc, C.; Gillett, P.; Bret, P. M.; Reinhold, C.; Kintzen, G.; Aldis, A. E.; Thibodeau, M., Role of endovaginal sonography in the diagnosis and management of ectopic pregnancy. *Radiographics* **1996**, *16* (4), 755-74; discussion 775.

41. Diagnosis and Management of Ectopic Pregnancy: Green-top Guideline No. 21. *Bjog* **2016**, *123* (13), e15-e55.

42. ACOG Practice Bulletin No. 94: Medical management of ectopic pregnancy. *Obstet Gynecol* **2008**, *111* (6), 1479-85.

43. Bobdiwala, S.; Al-Memmar, M.; Farren, J.; Bourne, T., Factors to consider in pregnancy of unknown location. *Womens Health (Lond)* **2017**, *13* (2), 27-33.

44. Milwidsky, A.; Adoni, A.; Segal, S.; Palti, Z., Chorionic gonadotropin and progesterone levels in ectopic pregnancy. *Obstet Gynecol* **1977**, *50* (2), 145-7.

45. Bobdiwala, S.; Saso, S.; Verbakel, J. Y.; Al-Memmar, M.; Van Calster, B.; Timmerman, D.; Bourne, T., Diagnostic protocols for the management of pregnancy of unknown location: a systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology* **2019**, *126* (2), 190-198.

Table 1. Pooled sensitivity, specificity data for the prediction protocols of PUL assessed in the systematic review and meta-analysis.

Protocol	Se (95% CI)	Sp (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC
M1	0.78[0.66,0.86]	0.90[0.88,0.92]	0.42[0.34,0.51]	0.98[0.96,0.99]	0.907
Progesterone cut-off	0.80[0.76,0.83]	0.51[0.49,0.52]	0.16[0.15,0.17]	0.96[0.95,0.96]	0.725
M6	0.94[0.91,0.95]	0.78[0.77,0.79]	0.36[0.33,0.38]	0.99[0.99,0.99]	0.944
M4	0.78[0.76,0.81]	0.79[0.78,0.80]	0.34[0.32,0.36]	0.96[0.96,0.97]	0.871
P1	0.69[0.53,0.82]	0.73[0.68,0.78]	0.24[0.17,0.33]	0.95[0.91,0.97]	/
hCG ratio	0.74[0.70,0.77]	0.81[0.80,0.83]	0.41[0.38,0.44]	0.95[0.94,0.95]	0.824
Progesterone+hCG	0.59[0.47,0.70]	0.74[0.70,0.77]	0.77[0.70,0.83]	0.93[0.90,0.95]	/
hCG cut-off	0.52[0.47,0.57]	0.78[0.76,0.80]	0.32[0.28,0.35]	0.89[0.88,0.91]	0.586
Ultrasound+hCG ratio	0.82[0.64,0.92]	0.73[0.67,0.78]	0.26[0.18,0.36]	0.97[0.93,0.99]	/
Ultrasound	0.78[0.63,0.89]	0.62[0.58,0.66]	0.14[0.10,0.19]	0.97[0.95,0.99]	/
FFN	0.42[0.16,0.71]	0.65[0.46,0.80]	0.29[0.11,0.56]	0.76[0.56,0.89]	/
Activin A	1.00[0.94,1.00]	1.00[0.98,1.00]	0.97[0.90,1.00]	1.00[0.99,1.00]	/

Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve.

Table 2. The average rate of lost of each protocol.

No. of visits	No. of examinations	No. of examinations	Sum of the Number*	Sum of the Number*	protocol
1	1	1	1	2	Ultrasound
1	1	2	2	3	hCG cut-off Progesterone Activin A FFN

No. of visits	No. of examinations	No. of examinations	Sum of the Number*	Sum of the Number*	protocol
1	1	3	3	4	Progesterone+h
2	2	3	3	5	M4 hCG ratio M1 Ultrasound+hC
2	2	4	4	6	M6
3	3	3/4	3/4	6/7	P1

*Sum of the number is defined as the sum of the number of visits and the number of examinations required by the protocols.

**NaN: Not a Number

Table 3. The average production utility of protocols.

Protocols	Average production utility*
M4	0.174
hCG ratio	0.165
M1	0.181
hCG cut-offs	0.195
Progesterone cut-offs	0.242
M6	0.157

*In the definition of average production utility, revenue is defined as AUC, and cost is defined as the sum of the number of visits and the number of examinations required.

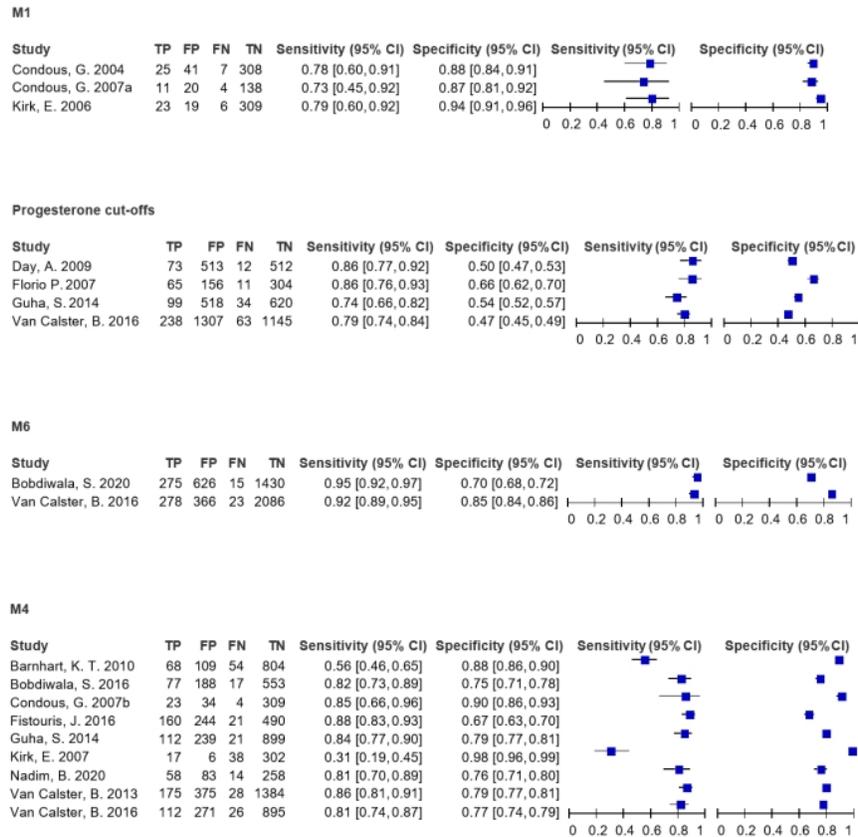
Table S1. Details of the included studies.

Author (year)	Study period	Protocol	Definition of EP in protocol
Nadim, B. 2020	2011.08-2018.09	P1	hCG ratio (48/0h) 0.8-0.99, and D7 hCG rises; hCG ratio (48/0h) \geq 1 EP [?]5%
Bobdiwala, S. 2020	2015.01-2017.01	M4 M6	Step 1 progesterone \geq 2nmol/L and Step 2 M6 model [?]5%
Van Calster, B. 2016	2003.07-2013.12	M6	Step 1 progesterone \geq 2nmol/L and Step 2 M6 model [?]5%
Fistouris, J. 2016	2011.01.01-2013.12.31	M4 Progesterone cut-offs(10nmol/L)	[?]5% EP \geq 10nmol/L high risk

Author (year)	Study period	Protocol	Definition of EP in protocol
		hCG ratio	hCG ratio (48/0h) changes between 0.87-1.66
Bobdiwala, S. 2016	2012.08-2013.12	M4	[?]5% EP high risk
Van Calster, B. 2013	2003.07-2011	M4	[?]5% EP high risk
Barnhart, K. T. 2010	2003.02.01-2007.09.30(US) 2003.07.18-2004.10.09(UK)	M4	[?]5% EP high risk
Condous, G. 2007b	2002.03.04-2003.07.17	M4 M5	[?]5% EP high risk combined clinical symptoms as predictive factors with M4
Bignardi, T. 2008	2001.06.25-2004.10.09	hCG ratio (hCG 48/0 h)	hCG ratio (48/0h) changes between 0.87-1.66
Kirk, E. 2007	2005.02.21-2006.02.20	M4	[?]5% EP high risk
Kirk, E. 2006	2003.08.07-2004.12.18	M1	>0.21
Condous, G. 2005b	2001.6.25-2003.4.14	hCG cut-offs(1000) hCGcut-offs(1500) hCGcut-offs(2000)	>1000 IU/L >1500 IU/L >2000 IU/L
Condous, G. 2005a	2001.6.25-2003.4.14	Progesterone&hCG cut-offs	Progesterone 10-50nmol/l and hCG>25u/l
Condous, G. 2004	2001.06-2002.12	M1	>0.21
Guha, S. 2014	2005.02-2012.08	M4 Progesterone cut-offs(10nmol/l)	[?]5% >10nmol/L
Kaplan, B. C. 1996	1991.08-1992.09	hCG ratio(0.87-1.66)	0.87-1.66
Dart, R. G. 1999b	1991.08-1998.08	hCG cut-offs(1000mIU/mL)	[?]1000mIU/mL
Dart, R. G. 1999a	1991.08-1998.08	hCG ratio+Ultrasonography	hCG ratio 0.5-1.66 and an empty uterus at ultrasound
Dart, R. G. 2002	1995.01-2000.08	Ultrasonography	Empty endometrial cavity with or without a thickened endometrium.
Dart, R. G. 1998	1991.08-1994.12	hCG cut-offs(1000mIU/mL)	[?]1000mIU/mL
Day, A. 2009	2004.01-2006.12	Progesterone cut-offs	>10nmol/L
Mol, B. W. 1998	1993.09-1996.04	hCG cut-offs	>1000IU/L
		hCG cut-offs	>1500IU/L
		hCG cut-offs	>2000IU/L
Dart, R. G. 1999a	1991.08-1997.08	hCG cut-offs	[?]1000mIU/mL
Condous, G. 2007a	2002.03-2003.07	M1	>0.21
Dart, R. G. 1996	1991.08-1994.12	hCG cut-off	>3000
El Bishry, G. 2008	-	hCG+progesterone cut-offs	16[?]P;80 and hCG>25 P>80 and hCG >1000
Nowacek GE. 1999	-	FFN	FFN>50ng/mL

Author (year)	Study period	Protocol	Definition of EP in protocol
Morse, C. B. 2012	2007.10-2009.06	hCG ratio	0.53~0.64-1.35
Florio P. 2007	2004.11-2005.09	Progesterone cut-offs hCG cut-offs	5.01 ng/ml 658IU/L
Zee, J. 2014	2007-2009	Activin A hCG ratio	0.37ng/ml j1.66

Table S2. Ectopic pregnancy data reported in studies and paired forest plots of sensitivity and specificity with corresponding 95% confidence intervals.



P1

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Nadim, B. 2020	29	92	13	249	0.69 [0.53, 0.82]	0.73 [0.68, 0.78]		

hCG ratio

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bignardi, T. 2008	51	83	22	784	0.70 [0.58, 0.80]	0.90 [0.88, 0.92]		
Fistouris, J. 2016	123	129	58	605	0.68 [0.61, 0.75]	0.82 [0.79, 0.85]		
Guha, S. 2014	85	129	48	1009	0.64 [0.55, 0.72]	0.89 [0.87, 0.90]		
Morse, C. B. 2012	149	241	30	585	0.83 [0.77, 0.88]	0.71 [0.68, 0.74]		
Zee, J. 2014	118	177	28	323	0.81 [0.73, 0.87]	0.65 [0.60, 0.69]		

Progesterone+hCG

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Condous, G. 2005a	23	89	14	362	0.62 [0.45, 0.78]	0.80 [0.76, 0.84]		
El Bishry, G. 2008	20	53	16	37	0.56 [0.38, 0.72]	0.41 [0.31, 0.52]		

hCG cut-offs

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Condous, G. 2005b	10	61	36	420	0.22 [0.11, 0.36]	0.87 [0.84, 0.90]		
Dart, R. G. 1996	7	78	20	89	0.26 [0.11, 0.46]	0.53 [0.45, 0.61]		
Dart, R. G. 1998	13	47	16	144	0.45 [0.26, 0.64]	0.75 [0.69, 0.81]		
Dart, R. G. 1999a	29	94	16	70	0.64 [0.49, 0.78]	0.43 [0.35, 0.51]		
Florio P. 2007	57	110	19	350	0.75 [0.64, 0.84]	0.76 [0.72, 0.80]		
Kaplan, B. C. 1996	21	37	35	346	0.38 [0.25, 0.51]	0.90 [0.87, 0.93]		
Mol, B. W. 1998	75	32	54	193	0.58 [0.49, 0.67]	0.86 [0.81, 0.90]		

FFN

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Nowacek GE. 1999	5	12	7	22	0.42 [0.15, 0.72]	0.65 [0.46, 0.80]		

Ultrasound+hCG ratio

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dart, R. G. 1999b	27	75	6	199	0.82 [0.65, 0.93]	0.73 [0.67, 0.78]		

Ultrasound

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dart, R. G. 2002	36	223	10	366	0.78 [0.64, 0.89]	0.62 [0.58, 0.66]		

Actin A

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Florio P. 2007	76	2	0	458	1.00 [0.95, 1.00]	1.00 [0.98, 1.00]		

TP: true positive; FP: false positive; FN: false negative; TN: true negative.

Hosted file

F1_PRISMA flowchart .docx available at <https://authorea.com/users/429508/articles/533193-protocols-for-prediction-of-ectopic-pregnancy-in-pregnancy-of-unknown-location-a-systematic-review-and-meta-analysis>

Hosted file

F2_QUADAS2.docx available at <https://authorea.com/users/429508/articles/533193-protocols-for-prediction-of-ectopic-pregnancy-in-pregnancy-of-unknown-location-a-systematic-review-and-meta-analysis>

Hosted file

F3_SR0C.docx available at <https://authorea.com/users/429508/articles/533193-protocols-for-prediction-of-ectopic-pregnancy-in-pregnancy-of-unknown-location-a-systematic-review-and-meta-analysis>