Published in final edited form as:
Fertil Steril. 2012 January ; 97(1): 101–106.e2. doi:10.1016/j.fertnstert.2011.10.037.

Performance of hCG curves in women at risk for ectopic pregnancy: exceptions to the rules

Christopher B. Morse, BA[a], **Mary D. Sammel, ScD**[b], **Alka Shaunik, MD**[a], **Lynne Allen-Taylor, PhD**[b], **Nicole L. Oberfoell, BA**[a], **Peter Takacs, MD, PhD**[c], **Karine Chung, MD, MSCE**[d], and **Kurt T. Barnhart, MD, MSCE**[a,b]

[a]Department of Obstetrics and Gynecology, University of Pennsylvania School of Medicine, Philadelphia, PA

[b]Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA

[c]Department of Obstetrics and Gynecology, University of Miami School of Medicine, Miami, FL

[d]Department of Obstetrics and Gynecology, University of Southern California, Los Angeles, CA

Abstract

Objective—To investigate the accuracy of serial hCG to predict outcome of a pregnancy of unknown location in an ethnically and geographically diverse setting.

Design—Multi-site cohort study

Setting—University hospital

Patients—Women with a pregnancy of unknown location

Interventions—None

Main outcome measures—Patients were followed until diagnosed with ectopic pregnancy (EP), intrauterine pregnancy (IUP), or miscarriage. To predict outcome, observed hCG was compared to recommended thresholds to assess deviation from defined normal curves. Predicted outcome was compared to standard of care. Sensitivity, specificity, predictive value, and accuracy were calculated, stratified by diagnosis.

Results—The final diagnosis 1,005 patients included 179 EPs, 259 IUPs, and 567 miscarriages. The optimal balance in sensitivity and specificity used the minimal expected two-day rise in hCG of 35%, and the minimal two-day decline in hCG of 36–47% (depending on the level) achieving 83.2% sensitivity, 70.8% specificity to predict EP. However 16.8% of EPs and 7.7% of IUPs would be misclassified solely using serial hCGs. Consideration of a third hCG and early ultrasound decreased IUP misclassification to 2.7%.

Corresponding author: Kurt T. Barnhart, MD, MSCE, Reproductive Research Unit, Department of Obstetrics and Gynecology, University of Pennsylvania Medical Center, 3701 Market Street, Suite 810, Philadelphia, PA 19104, Telephone: (215) 662-2972, Fax: (215) 573-7627, kbarnhart@obgyn.upenn.edu.

Conflicts of interest: None.

Conclusion—Solely using serial hCG values can result in misclassification. Clinical judgment should trump prediction rules and continued surveillance with a third hCG may be prudent, especially when initial values are low or when values are near suggested thresholds.

Introduction

Women presenting with pain, bleeding, and a positive pregnancy test require timely and accurate diagnosis to distinguish non-viable from viable pregnancies. A portion of this population will have a pregnancy of unknown location (PUL), a descriptive term applied to women without evidence of either an intrauterine pregnancy (IUP) or an ectopic pregnancy (EP) on transvaginal sonography (TVS) (1). Although miscarriage is the most common complication of early pregnancy (2), ectopic pregnancy complicates 7–20% of symptomatic first-trimester pregnancies and contributes greatly to maternal morbidity and mortality (3–5).

The diagnosis of EP relies upon the interpretation of serial human chorionic gonadotrophin (hCG) levels in conjunction with TVS and clinical history. TVS is sensitive and specific for distinguishing an IUP from an EP when the presenting hCG is above the discriminatory zone (6, 7). Clinical interpretation of TVS in patients with hCG levels close to, or below the discriminatory zone, is challenging and initial TVS alone cannot detect 26% of EPs (5). Additional factors may impact the diagnostic utility of TVS (8, 9).

Typically, patients with a PUL are reassessed within 48 hours from presentation to determine the change in hCG and to aid in the frequency and method of follow-up. Expected rates of increase for an IUP and decrease for a miscarriage have been reported (10, 11). Values that do not rise or fall as expected raise the suspicion for EP. Based on the findings of serial hCG values, various methods have been proposed to predict final outcome (12–19). Comparing serial hCGs to expected values is accurate and decreases the time and number of visits to diagnosis compared with standard care (16), however this strategy has not been validated in population distinct from which it was developed.

A **strategy** to predict outcome in women with a PUL that validates on an external sample of patients would be valuable. Prediction models may lose excellent test characteristics when applied externally, or may need to be amended or updated to maintain accuracy (20). The purpose of this study is to compare observed hCG curves to expected curves in a diverse set of patients with symptomatic early pregnancy and a PUL in order to predict outcome.

Materials and Methods

Sample Selection

This study was conducted at three sites: the University of Pennsylvania, the University of Miami, and the University of Southern California as part of the Predictors of Ectopic Pregnancy (PEP) Study with Institutional Review Board at each site. Data were collected from October 2007 to June 2009, using a centralized computerized database. Patients were entered into the database when they presented in the first-trimester of pregnancy with pain and/or bleeding.

All included patients did not have signs of an intrauterine or extrauterine gestation on TVS at presentation, had at least two hCG values (at least one day apart), and a documented date of eventual definitive diagnosis. Patients were excluded if they were diagnosed at presentation or never received a definite diagnosis. In order to be comparable to a previous cohort, hCG levels over 10,000 mIU/mL were censored. The analysis was restricted to hCG values prior to, and including, the date of diagnosis.

Clinical Outcomes

Each patient was diagnosed with either an EP, IUP, or miscarriage, matched to recently defined nomenclature (1). EP included visualized and non-visualized EPs, and treated persistent PULs. An IUP was defined as visualization of an intrauterine gestational sac with yolk sac and/or fetal pole on ultrasound. Miscarriage included spontaneously resolved PUL (the resolution of serum hCG, two decreasing hCGs with the final level below 25 mIU/mL, or three declining hCGs with the final level below 500 mIU/mL), resolved persistent PUL, and histological IUP. ***It is not possible to report the assay variability for all three sites over a 3 year period.*** HCG concentration were performed at the clinical laboratory of each participating center and all clinical laboratories were CAP certified. ***CAP certification standards are coefficients of variation for hCG labs to be 8% or less (but does not require reporting of each laboratory on a yearly basis).***

Model Based Classification

Prediction rules for model based classification were applied retrospectively and did not impact actual clinical care. This study did not dictate when or how often serial hCG values were obtained. The timing and frequency of hCG values was at the desertion of treating clinician. The prediction of clinical outcome has been previously described (16). Briefly, the trend between initial hCG values was determined to be rising or falling. For those with an initial rise, the rate of increase, or slope, was computed and compared to the minimum expected slope for an IUP (11). For patients with an initial decline, the observed decrease was compared to an expected decrease, calculated using previously defined bounds (10). If the observed change in serial hCG concentration was less than the minimum defined by the lower bound (either a rise or a fall), the patient would be classified as a suspected EP by the model. This logic is depicted in supplemental figure 1.

If the observed hCG did not cross the threshold for diagnosis (i.e. it was increasing or decreasing above a specified cutoff), the process was repeated based on comparison of the next hCG with the previous value. If, on subsequent values, the slope now failed to increase (or decrease) as expected, or the slope switched directions (i.e. a rise followed by a fall or vice versa), the patient was classified by the model as a suspected EP. If the change in hCG did not deviate from ‘normal’ during the course of evaluation, diagnosis was made based on ultrasound findings, clinical symptoms, or resolution of hCG from the serum.

Confidence interval (CI) bounds representing the expected rise based on the 95%, 99% and 99.9% CIs for the slope of rising IUPs were used (11). The lower limits of these bounds correspond to a two-day expected rise of 71%, 53%, and 35%, respectively.

The expected decline was determined for the 90% and 95% CI based upon decreasing miscarriage curves (10). The exact decrease was contingent upon the value of hCG at presentation. For example, the expected two-day decline in hCG for the 90% CI ranged from 36–47%, while the expected decline for the 95% CI ranged from 21–35%. Thus, the expected decline for the 90% CI was “faster” whiles the decline for the 95% CI “slower”.

The use of three serial hCG values was also explored in patients with an initial rise to minimize theoretical misclassification of a viable IUP with a falsely slow rise in hCG. The hCG concentration immediately following model classification of an ‘abnormal’ rise (based on two values), was evaluated before a classification of ‘abnormal’ was made. This simulated the clinical practice of obtaining one additional hCG value before intervention. If the subsequent rise (based on three values) was within the expected range, the subject was reclassified as an IUP.

Data Analysis

Sensitivity, specificity, positive and negative predictive value (PPV and NPV), and accuracy were calculated for each combination of bounds. To calculate these values, 2×2 tables were created with respect to each of the three outcomes. For each table, disease positive was defined as having the outcome of interest (i.e. an EP) while disease negative was defined by combining both of the other two outcomes (i.e. an IUP or a miscarriage). Days and visits to clinical diagnosis and days to diagnosis based on prediction rules using hCG curves were determined. Categorical demographic variables were evaluated by chi-square test of association. Accuracy was calculated as the percentage of model-based classifications correctly matched to clinical outcome.

The statistical software package SAS 9.2 (SAS Institute Inc., Cary, NC) was used for all data analyses.

Results

Of 1,180 patients that met inclusion criteria, 167 were lost to follow-up and 8 were excluded due to missing data, resulting in 1,005 patients for analysis. Demographic characteristics are presented in supplemental Table 1. An initial rise in hCG was observed in 453 (45.1%) patients while 552 (54.9%) had an initial fall.

First, the sensitivity and specificity of the model to predict the outcome of EP was examined (Table 1). The combination that resulted in the fewest misclassified EPs and IUPs was the 35% expected rise for an IUP and the “faster” expected hCG decline (based upon the 90% CI). The combination with the best overall accuracy to predict an EP also used an expected rise of 35% but instead used the “slower”, expected hCG decline (based on the 95% CI).

Next, model performance to predict the outcome of IUP was evaluated (Table 2). The most sensitive method for correctly identifying an IUP used an expected rise of 35%. This was very specific and accurate with high PPV and NPV.

We also examined various combinations of expected hCG rise and decline to predict miscarriages (data not shown). Using the “slower” expected decline achieved a sensitivity of 67.5%., specificity of 96.6%, a PPV of 96.2%, NPV of 69.7%, and overall accuracy of 80.2%. A “faster” expected decline of hCG lowered sensitivity, NPV, and accuracy, while minimally improving specificity and PPV.

The number of days and visits saved is presented in comparison to the results obtained from the original development of this rule (the prior cohort) in Table 3. The sensitivity for detecting EPs was comparable between two cohorts, but the sensitivity for detecting IUPs was slightly lower. The mean number of days and visits saved was similar. With respect to IUPs, the model saved 9.2 days and 0.98 visits (data not shown).

The number of misclassifications based solely on the use of hCG curves was quantified and presented alongside results from the prior cohort in Table 4. The combination that resulted in the fewest total misclassified EPs and IUPs used a 35% expected rise in hCG and the “faster” expected decline (36–47% over two days). Compared to prior results, the percentage of misclassified EPs was lower, while the percentage of misclassified IUPs was moderately increased.

The clinical course of those misclassified suing solely hCG values was evaluated. 30 women with an EP were classified as IUP or miscarriage. Of these, 24 had a rise greater than 35% over two days and 6 had a fall consistent with of miscarriage. Six of the 30 misclassified EPs (20.0%) were diagnosed clinically with rupture. Three of these six patients had

confirmed hemoperitoneum at laparoscopy. The remaining 24 cases were diagnosed by visualization of EP by TVS, or after dilation and curettage.

Twenty women with an IUP were classified as a nonviable gestation (EP or miscarriage). Of these women, 18 had a rise in hCG less than the minimal expected rise, 35% in two days, and two had a change in the direction of their hCG curve, but with a subsequent rise and ultrasound confirming a viable gestation.

Incorporation of a third data point correctly reclassified six IUPs because the third hCG value was rising 'normally,' even when the second was too slow. However elevating the third hCG value also resulted in misclassification of nine EPs, and two miscarriages to an IUP. Overall, the three data point model resulted in modest changes in model performance.

To perform this analysis we included anyone that had more than one hCG value. This could include women who had a value on day 0, and day 1 (as well as subsequent values). However very few women (2%) had only two values 24 hrs apart (i.e., diagnosis was made 1 day later), the vast majority of women had values 48 hrs apart (or in some cases longer than 48 hrs apart). Most women (85%) who had a value 24 hours after presentation had subsequent values as well."

Discussion

Diagnosis of symptomatic first-trimester pregnancy of unknown location is challenging. A simple strategy to identify women at high risk for EP would help clinicians optimally monitor certain patients and intervene, if necessary. Comparing the serial hCG curve of the actual patient to that of what is expected for a growing pregnancy or a resolving pregnancy is widely advocated (4, 7, 10 –16). However, such a method should minimize misclassification. The important clinical question when evaluating serial hCG values is, how much of a deviation from 'normal' is acceptable before deciding to intervene and definitively diagnose a woman at risk for EP? The goal of this study was to validate prior prediction rules on a new, racially and geographically diverse group of patients. These prediction rules were applied independently and did not affect clinical care.

These data demonstrate and validate that it is possible to compare the curve generated from serial hCGs as a patient is followed to 'normal' curves of a growing IUP or a resolving miscarriage to identify and distinguish patients that can continue outpatient surveillance from patients that need closer monitoring or intervention. The current results are similar to application of hCG curves to a previous cohort (16). Of note, the minimal rise for an IUP of 35% achieve the optimal balance of misclassification is slower than 53% previously published that characterized a rise in hCG for a viable gestation (11).

Optimal classification balances the sensitivity to diagnose an EP versus the specificity of not intervening when a gestation does not need intervention. Thus, our proposed cutoffs are not symmetric. A minimum expected rise of 35% over two days was selected (representing the theoretical rise in hCG of 99.9% of 'normally' growing IUPs) so as to not interrupt an IUP, an error with great consequence, while we chose relatively "fast" hCG decline (a minimum decline of 36–47% over two days representing 90% of women with a miscarriage) as a cutoff to better identify EPs. This combination resulted in the fewest number of misclassified IUPs and EPs.

This "optimal" combination was not the combination with the highest accuracy overall for all three outcomes. The combination with the highest accuracy used a slower decline in hCG rise as a cut off. This reduced misclassification of miscarriages (increasing overall accuracy) but misclassified nine more EPs. The difference between a misclassified EP and a

misclassified miscarriage is significant because the former carries the potential for significant morbidity and mortality. We suggest that it is better to 'over treat' some miscarriages in order to more accurately diagnosis an EP.

It is expected that solely using serial hCG concentrations will misclassify some EPs. Approximately 21% of EPs have a rise in hCG similar to an IUP and 8% have a fall similar to a miscarriage (21). Of the 30 misclassified EPs in this cohort, 24/30 (80.0%) had an initial rise in hCG consistent with a growing IUP whereas only 6/30 (20.0%) declined like a miscarriage. This reflects the selection of a slow rise of hCG as "normal". Selecting a cutoff of 35% for hCG rise prioritized correct classification of IUPs over EPs.

The clinical consequence of misclassified EPs in this context was minimal. While six cases of "missed" cases of EP intervened upon due to symptoms consistent with rupture, only three had intra-operative confirmation of hemoperitoneum. This represents 3/1005 (0.3%) of all patients in the cohort of women with a PUL and 3/179 (1.7%) of all EPs. This is an acceptable rate given that tubal rupture is an unpredictable outcome.

A severe error would be to misclassify a desired IUP. Therefore, diagnosis using hCG values should be made in conjunction with symptoms and ultrasound findings. For example, all women with an abnormal rise in hCG (and thus misclassified by the model) would not have been interrupted based on clinical criteria. Of the 20 cases of false negative classification of a woman with an IUP based on hCG alone, 13/20 (65%) had ultrasound findings suggestive of an IUP and were appropriately followed without intervention. Of those women without early ultrasound findings, the abnormality in hCG curves presented very early, typically around 500 mIU/mL. In these cases further values were obtained that were within the expected range for an IUP.

Using three hCG values (instead of two) correctly reclassified a small percentage of IUPs, most of which presented at levels well below the discriminatory zone. However, the benefit of correctly reclassifying some IUPs came with a cost, as it also incorrectly reclassified some EPs to that of a growing IUP.

Other studies have developed models to predict outcome when a patient presents with a PUL. Some include logistic regression models based progesterone level, serum hCG, endometrial thickness, and bleeding (13); hCG ratio alone (14), the hCG ratio, hCG average, and its quadratic effect (17, 18), or Bayesian networks (15). The models vary regarding varying complexity and results and have often not been validated in a separate population. Previous application of a sample of our population to a prediction model from the United Kingdom decreased the sensitivity for EP from 80% to 49% (20). It is likely that these models are limited because more than two hCG values are necessary to obtain an accurate clinical diagnosis.

The strengths of the model comparing expected hCG values to observed until deviation for that of expect for "normal" included flexibility, validation, and rationality. Importantly, this strategy allows for change over time greater than 48 hour. Second, we present data demonstrating validation well when applied to a broad cross-section of race, ethnicity, and geography.

In summary, these data validate that standardized prediction rules based on serial hCG concentration can accurately be applied to subjects with a pregnancy of unknown location and at risk for EP, saving time and visits to diagnosis. It should be noted that the use of serial hCG curves will incorrectly classify some women with an IUP or an EP, perhaps more than expected. Serial hCG values should not be used alone to diagnosis a woman, but should be used in combination with clinical judgment, evaluation of symptoms and repeat

ultrasound (as needed). Whereas the majority of cases have hCG curves that are well within the limits of 'normal', it is prudent to reassess a patient before intervention when the rise or fall in hCG is close to the extremes. To minimize misclassification of an IUP, the minimal expected rise for normal should be used (35% in two days) and a third value is recommended, especially when values are low. These rules should not replace clinical judgment. Further insight into the natural history of hCG curves, along with integration of data like symptoms, suspicious ultrasound findings, and perhaps race and ethnicity, will likely improve prediction rules in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: Supported by grant numbers R01-HD036455 (KB, MDS), K24HD060687 (KB) and the Doris Duke Clinical Research Fellowship (CBM).

References

1. Barnhart K, van Mello NM, Bourne T, Kirk E, Van Calster B, Bottomley C, et al. Pregnancy of unknown location: a consensus statement of nomenclature, definitions, and outcome. Fertil Steril. 2010
2. Wilcox AJ, Weinberg CR, O'Connor JF, Baird DD, Schlatterer JP, Canfield RE, et al. Incidence of early loss of pregnancy. N Engl J Med. 1988; 28;319(4):189–94.
3. Chang J, Elam-Evans LD, Berg CJ, Herndon J, Flowers L, Seed KA, et al. Pregnancy-related mortality surveillance--United States, 1991--1999. MMWR Surveill Summ. 2003; 21;52(2):1–8.
4. Barnhart KT. Clinical practice. Ectopic pregnancy. N Engl J Med. 2009; 23;361(4):379–87.
5. Kirk E, Papageorghiou AT, Condous G, Tan L, Bora S, Bourne T. The diagnostic effectiveness of an initial transvaginal scan in detecting ectopic pregnancy. Hum Reprod. 2007; 22(11):2824–8. [PubMed: 17855406]
6. Barnhart KT, Simhan H, Kamelle SA. Diagnostic accuracy of ultrasound above and below the beta-hCG discriminatory zone. Obstet Gynecol. 1999; 94(4):583–7. [PubMed: 10511363]
7. Cacciatore B, Stenman UH, Ylostalo P. Diagnosis of ectopic pregnancy by vaginal ultrasonography in combination with a discriminatory serum hCG level of 1000 IU/l (IRP). Br J Obstet Gynaecol. 1990; 97(10):904–8. [PubMed: 2223681]
8. Barnhart KT, Fay CA, Suescum M, Sammel MD, Appleby D, Shaunik A, et al. Clinical factors affecting the accuracy of ultrasonography in symptomatic first-trimester pregnancy. Obstet Gynecol. 2011; 117(2 Pt 1):299–306. [PubMed: 21252743]
9. Kirk E, Bourne T. Diagnosis of ectopic pregnancy with ultrasound. Best Pract Res Clin Obstet Gynaecol. 2009; 23(4):501–8. [PubMed: 19356985]
10. Barnhart K, Sammel MD, Chung K, Zhou L, Hummel AC, Guo W. Decline of serum human chorionic gonadotropin and spontaneous complete abortion: defining the normal curve. Obstet Gynecol. 2004; 104(5 Pt 1):975–81. [PubMed: 15516387]
11. Barnhart KT, Sammel MD, Rinaudo PF, Zhou L, Hummel AC, Guo W. Symptomatic patients with an early viable intrauterine pregnancy: HCG curves redefined. Obstet Gynecol. 2004; 104(1):50–5. [PubMed: 15229000]
12. Romero R, Kadar N, Jeanty P, Copel JA, Chervenak FA, DeCherney A, et al. Diagnosis of ectopic pregnancy: value of the discriminatory human chorionic gonadotropin zone. Obstet Gynecol. 1985; 66(3):357–60. [PubMed: 3895079]
13. 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(2):158–63. [PubMed: 11236115]

14. Condous G, Okaro E, Khalid A, Timmerman D, Lu C, Zhou Y, et al. The use of a new logistic regression model for predicting the outcome of pregnancies of unknown location. Hum Reprod. 2004; 19(8):1900–10. [PubMed: 15205400]

15. Gevaert O, De Smet F, Kirk E, Van Calster B, Bourne T, Van Huffel S, et al. Predicting the outcome of pregnancies of unknown location: Bayesian networks with expert prior information compared to logistic regression. Hum Reprod. 2006; 21(7):1824–31. [PubMed: 16601010]

16. Seeber BE, Sammel MD, Guo W, Zhou L, Hummel A, Barnhart KT. Application of redefined human chorionic gonadotropin curves for the diagnosis of women at risk for ectopic pregnancy. Fertil Steril. 2006; 86(2):454–9. [PubMed: 16753158]

17. Condous G, Van Calster B, Kirk E, Haider Z, Timmerman D, Van Huffel S, et al. Clinical information does not improve the performance of mathematical models in predicting the outcome of pregnancies of unknown location. Fertil Steril. 2007; 88(3):572–80. [PubMed: 17499248]

18. Condous G, Van Calster B, Kirk E, Haider Z, Timmerman D, Van Huffel S, et al. Prediction of ectopic pregnancy in women with a pregnancy of unknown location. Ultrasound Obstet Gynecol. 2007; 29(6):680–7. [PubMed: 17486691]

19. Casanova BC, Sammel MD, Chittams J, Timbers K, Kulp JL, Barnhart KT. Prediction of outcome in women with symptomatic first-trimester pregnancy: focus on intrauterine rather than ectopic gestation. J Womens Health (Larchmt). 2009; 18(2):195–200. [PubMed: 18991513]

20. Barnhart KT, Sammel MD, Appleby D, Rausch M, Molinaro T, Van Calster B, et al. Does a prediction model for pregnancy of unknown location developed in the UK validate on a US population? Hum Reprod. 2010; 25(10):2434–40. [PubMed: 20716562]

21. Silva C, Sammel MD, Zhou L, Gracia C, Hummel AC, Barnhart K. Human chorionic gonadotropin profile for women with ectopic pregnancy. Obstet Gynecol. 2006; 107(3):605–10. [PubMed: 16507931]

Table 1

Test Performance of Various hCG Cutoffs to Predict the Outcome of Ectopic Pregnancy

Expected Two-Day Rise for an IUP	Sensitivity (%) [95%CI]	Specificity (%) [95%CI]	PPV (%) [95%CI]	NPV (%) [95%CI]	Accuracy (%) [95%CI]
35% Rise in hCG	83.2 [77.7–88.8]	70.8 [67.7–73.9]	38.2 [33.4–43.0]	95.1 [93.4–96.8]	73.0 [70.3–75.8]
53% Rise in hCG	91.1 [86.8–95.3]	66.6 [63.4–69.8]	37.1 [32.6–41.7]	97.2 [95.8–98.5]	70.9 [68.1–73.8]
71% Rise in hCG	92.2 [88.2–96.2]	62.8 [59.5–66.1]	35.0 [30.6–39.3]	97.4 [96.0–98.7]	68.1 [65.2–70.9]

hCG, human chorionic gonadotropin; PPV, positive predictive value; NPV, negative predictive value; IUP, intrauterine pregnancy. Expected two-day rise of 35%, 53%, and 71% reflects the 99.9%, 99%, and 95% confidence interval bounds, respectively, for an IUP. Test performance criteria utilize the 90% confidence interval bounds for the expected two-day decline for a spontaneous miscarriage, corresponding to a decline of 36–47% (depending on level)..

Table 2

Test Performance of Various hCG Cutoffs to Predict the Outcome of Intrauterine Pregnancy

Expected Two-Day Rise for an IUP	Sensitivity (%) [95%CI]	Specificity (%) [95%CI]	PPV (%) [95%CI]	NPV (%) [95%CI]	Accuracy (%) [95%CI]
35% Rise in hCG	92.3 [89.0–95.6]	94.0 [92.3–95.7]	84.2 [79.9–88.4]	97.2 [96.0–98.4]	93.5 [92.0–95.1]
53% Rise in hCG	82.6 [78.0–87.3]	97.2 [96.0–98.4]	91.1 [87.4–94.7]	94.2 [92.5–95.8]	93.4 [91.9–95.0]
71% Rise in hCG	72.6 [67.1–78.1]	98.1 [97.1–99.1]	93.1 [89.5–96.6]	91.2 [89.2–93.1]	91.5 [89.8–93.3]

hCG, human chorionic gonadotropin; PPV, positive predictive value; NPV, negative predictive value; IUP, intrauterine pregnancy. Expected two-day rise of 35%, 53%, and 71% reflects the 99.9%, 99%, and 95% confidence interval bounds, respectively, for an IUP. Test performance criteria utilize the 90% confidence interval bounds for the expected two-day decline for a spontaneous miscarriage, corresponding to a decline of 36–47% (depending on level).

Table 3

Performance in Validation Cohort versus Original Cohort for Various hCG Cutoffs to Predict the Outcome in Pregnancy of Unknown Location

Expected Two-Day Rise for an IUP	Sensitivity for EP (%)		Sensitivity for IUP (%)		Mean number of days saved (range)[#]		Mean number of visits saved (range)[#]	
	Validation	Original[*]	Validation	Original	Validation	Original	Validation	Original
35% Rise in hCG	83	83	92	95	2.87 (0–35)	2.64 (0–34)	0.92 (0–7)	1.22 (0–9)
53% Rise in hCG	91	88	83	90	3.27 (0–35)	2.85 (0–34)	1.07 (0–7)	1.30 (0–9)
71% Rise in hCG	92	91	73	78	3.44 (0–37)	2.94 (0–34)	1.12 (0–7)	1.35 (0–9)

hCG, human chorionic gonadotropin; EP, ectopic pregnancy; IUP, intrauterine pregnancy. Expected two-day rise of 35%, 53%, and 71% reflects the 99.9%, 99%, and 95% confidence interval bounds, respectively, for an IUP. Test performance criteria utilize the 90% confidence interval bounds for the expected two-day decline for a spontaneous miscarriage, corresponding to a decline of 36–47% (depending on level).

[*] Original cohort described by Seeber BE, Sammel MD, Guo W, Zhou L, Hummel A, Barnhart KT. Application of redefined human chorionic gonadotropin curves for the diagnosis of women at risk for ectopic pregnancy. Fertil Steril 2006;86(2):454–9.

[#] For patients with outcome of ectopic pregnancy..

Table 4

Performance in Validation Cohort versus Original Cohort for Various hCG Cutoffs to Predict the Outcome in Pregnancy of Unknown Location

Expected Two-Day Rise for an IUP	Number of misclassified EPs (%)		Number of misclassified IUPs (%)		Number of misclassified miscarriages (%)	
	Validation	Original*	Validation	Original	Validation	Original
35% Rise in hCG	30 (16.8)	34 (17.3)	20 (7.7)	12 (4.6)	221 (39.0)	222 (28.0)
53% Rise in hCG	16 (8.9)	24 (12.2)	45 (17.4)	26 (10.0)	231 (40.7)	224 (28.2)
71% Rise in hCG	14 (7.8)	18 (9.2)	71 (27.4)	58 (22.2)	236 (41.6)	225 (28.4)

hCG, human chorionic gonadotropin; EP, ectopic pregnancy; IUP, intrauterine pregnancy. Expected two-day rise of 35%, 53%, and 71% reflects the 99.9%, 99%, and 95% confidence interval bounds, respectively, for an IUP. Test performance criteria utilize the 90% confidence interval bounds for the expected two-day decline for a spontaneous miscarriage, corresponding to a decline of 36–47% (depending on level).

*Original cohort described by Seeber BE, Sammel MD, Guo W, Zhou L, Hummel A, Barnhart KT. Application of redefined human chorionic gonadotropin curves for the diagnosis of women at risk for ectopic pregnancy. Fertil Steril 2006;86(2):454–9.