Diagnosing ectopic pregnancy and current concepts in the management of pregnancy of unknown location

E. Kirk1,2*, C. Bottomley3, and T. Bourne2,4

1North Middlesex University Hospital, Sterling Way, London N18 1QX, UK 2Queen Charlotte’s & Chelsea Hospital, Imperial College, Du Cane Road, London W12 0HS, UK 3Chelsea and Westminster Hospital, Fulham Road, London SW10 9NH, UK 4Department of Development and Regeneration, University Hospitals Leuven, Campus Gasthuisberg, KU Leuven, B-3000 Leuven, Belgium

*Correspondence address. E-mail: ejkirk@hotmail.co.uk

Submitted on April 10, 2013; resubmitted on July 31, 2013; accepted on August 15, 2013

TABLE OF CONTENTS

- Introduction
- Methods
  - Tubal ectopic pregnancy
    - History and presentation
    - Ultrasound diagnosis
    - Surgical diagnosis
  - Non-tubal ectopic pregnancy
    - Heterotopic pregnancy
  - Pregnancy of unknown location
    - Diagnosis
    - Clinical outcomes
    - Clinical management
- Future work
- Conclusion

BACKGROUND: A diagnosis of ectopic pregnancy (EP) is primarily achieved using transvaginal ultrasonography (TVS). Pregnancy of unknown location (PUL) is the term used to categorize a pregnancy in a woman with a positive pregnancy test when no pregnancy has been visualized using TVS. This review appraises current tools for the diagnosis of EP, describes the diagnostic criteria for non-tubal EP and reviews the literature on the clinical management of PUL.

METHODS: We performed a targeted search using the PubMed database. All articles published in the English language from January 1984 to March 2013 were screened for eligibility.

RESULTS: Using TVS to diagnose EP is highly sensitive (87–99%) and specific (94–99.9%). Variations exist in the criteria used for ultrasound diagnosis. Studies report that between 5 and 42% of women seen for ultrasound assessment with a positive pregnancy test have a PUL. For PUL, measurements of serum human chorionic gonadotrophin (hCG) and progesterone are used to predict pregnancy viability and therefore give an indication of the risk of an EP. Only 6–20% of PUL are subsequently diagnosed with EP. Non-tubal EPs are relatively uncommon, difficult to diagnose and result in disproportionate morbidity and mortality.

CONCLUSIONS: Access to expertise and equipment for high-quality TVS means the majority of women with EP in developed countries can be diagnosed rapidly and accurately. Identifying PUL, which are low risk and therefore requiring less follow-up, finding better serum markers for EP and safely identifying women who do not require intervention for EP are the current diagnostic challenges.

Key words: ectopic pregnancy / pregnancy of unknown location / transvaginal ultrasonography / prediction models / expectant management

© The Author 2013. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
Introduction

Any pregnancy implanted outside the endometrial cavity is defined as an ectopic pregnancy (EP). Most are located within the Fallopian tube. The incidence of EP varies from 11 to 20 per 1000 live births in developed countries (Centres for Disease Control, 1995; Boufous et al., 2001; Bakken and Skjeldestad, 2003; Lewis 2007), though it is as high as 4% in assisted conception populations (Fernandez and Gervaise, 2004).

The case fatality rate in developed countries has declined over recent years, suggesting that earlier diagnosis and better treatment have made an impact (O’Herlihy, 2011).

Diagnosis based on clinical signs of tubal rupture (hypovolaemic shock) is now uncommon and transvaginal ultrasonography (TVS) is the primary diagnostic tool for clinically stable women with a suspected EP (Jurkovic and Wilkinson, 2011). Data from a specialist early pregnancy unit show that up to 73.9% of women with an EP may be diagnosed on an initial TVS assessment (Kirk et al., 2008) and 94% are diagnosed prior to surgical intervention (Condous et al., 2005a).

Non-tubal EPs although relatively uncommon are associated with higher morbidity and mortality. Pregnancy of unknown location (PUL) is the term used to classify a pregnancy in women with a positive urinary pregnancy test when an initial TVS detects neither an intrauterine nor extrauterine pregnancy. A small proportion of women classified with a PUL have an underlying EP.

The aim of this review is first to detail the diagnostic criteria for both tubal and non-tubal EP. The secondary aim is to define the use of the term ‘PUL’ and describe how to use biochemical markers to triage women to appropriate follow-up when they have a PUL.

Methods

We performed a targeted search using the PubMed database employing combinations of the following search terms: EP, tubal/cornual/interstitial/ Caesarean scar/intramural/ cervical/ abdominal EP, PUL. All articles published in the English language from January 1984 to March 2013 were selected. The restriction to articles after 1984 was due to this being when TVS became commonly used in clinical practice.

All potential papers were screened for eligibility by review of the title and abstract, to assess relevance in terms of focus on either diagnosis of any type of EP or an original or review paper on any aspect of PUL. The full text versions of all papers selected in the screening process were read and assessed for relevance. Reference lists of all primary and review articles were examined for any relevant citations (from any year), which may have been missed in the original key word search.

One main author performed the search for each specific section [tubal EP diagnosis (CB), non-tubal EP diagnosis (EK), PUL (EK)] with all three authors subsequently rechecking the included articles for relevance.

Tubal Ectopic Pregnancy

Most (93–98%) EPs are located within the Fallopian tube. Of these, 13% are isthmic, 75% ampullary and 12% fimbrial (Bouyer et al., 2002; Walker, 2007; Varma and Gupta, 2009). The aetiology remains poorly defined but is likely to be a combination of impaired embryo–tubal transport and alterations in the tubal environment allowing early implantation (Shaw et al., 2010a, b).

History and presentation

Risk factors for tubal EP are well described, with highest risk associated with tubal damage following surgery or infection (especially Chlamydia trachomatis), smoking and IVF (Ankum et al., 1996; Tay et al., 2000; Bouyer et al., 2003; Shaw et al., 2010a, b; Shaw et al., 2011). However although the attributable risk of these factors is 0.76 (Bouyer et al., 2003), many women with EP have no identifiable risk factors and most women with a risk factor do not have an EP.

The triad of pain, vaginal bleeding and amenorrhoea was historically used to suspect a diagnosis of EP (Weckstein et al., 1985). These symptoms, with or without syncope, shoulder tip pain and shock, generally led to surgical intervention. Now, history and physical examination alone rarely leads to the diagnosis or exclusion of an EP, with most diagnosed earlier in the course of the disease. One-third of women with an EP have no clinical signs and up to 10% have no symptoms (Kaplan et al., 1996; Tay et al., 2000). The symptoms of EP are often non-specific and difficult to differentiate from those of other gynaecological, gastrointestinal and urological disorders, including appendicitis, salpingitis, corpus luteum cyst rupture, miscarriage, adnexal torsion or urinary tract infection.

The amount of bleeding associated with EP varies, although classically the patient will complain of ‘spotting’. Heavy bleeding, in the absence of further TVS and hCG assessment, may lead to a clinical misdiagnosis of miscarriage. Abdominal pain is often absent or a late finding, probably due to earlier recognition of pregnancy from the use of highly sensitive commercially available urinary pregnancy tests and to easy access to TVS. The majority of women with abdominal pain in early pregnancy do not have an EP (Bottomley et al., 2009).

Less common features of EP include nausea, vomiting and diarrhoea. In ruptured EP, there may be abdominal distension and tenderness, peritonism and haemorrhagic shock. A diagnosis of EP should be considered in all women of reproductive age with a sudden onset of abdominal pain or gastrointestinal symptoms (Lewis, 2007).

Ultrasound diagnosis

The use of ultrasonography to diagnose EP was first described over 40 years ago (Kobayashi et al., 1969). However until the 1980s, the ultrasound ‘diagnosis’ was based on an inability to visualize an intrauterine pregnancy (IUP) rather than definitive identification of an extrauterine pregnancy. This was considered an indication for laparoscopy, with an inevitable high rate of ‘negative laparoscopy’ (Brown and Doubilet, 1994).

The use of TVS has changed the diagnostic approach to one based on visualizing the ectopic mass (Condous et al., 2005a). TVS has been demonstrated to be superior to transabdominal ultrasound (TAS), with sensitivities for the diagnosis of EP in early studies of 77–80% for TAS and 88–90% for TVS (Cacciarelli et al., 1989; Valenzano et al., 1991). As expertise and equipment has improved, the sensitivity of TVS for the diagnosis of EP has increased further (Condous et al., 2005a; Kirk et al., 2007a).

The accuracy of a diagnosis of tubal EP using ultrasonography

It is important to appreciate that not all EP will be seen using TVS and certainly not on one visit. A prospective study on over 5000 consecutive women including 120 tubal EPs found that 73.9% were visualized on TVS when the patient first attended the clinic (Kirk et al., 2007a). The remaining cases were not seen and were initially classified as PUL. Most EPs were then visualized on subsequent follow-up scans (prior to
surgery) making the overall sensitivity of TVS 98.3%. A number of other studies have also assessed the accuracy of TVS for diagnosing tubal EP. They too report high sensitivities of 87.0–99.0% and specificities of 94.0–99.9% (Braffman et al., 1994; Shalev et al., 1998; atriet al., 2003; Condous et al., 2005a; Kirk et al., 2008). A meta-analysis of 2216 women with an EP concluded that visualizing an adnexal mass separate from the ovary using TVS had a sensitivity of 84.4% and specificity of 98.9% (Brown and Doubilet, 1994).

A study comparing the features of women with an EP who had a diagnosis made on the initial scan compared with those first classified as a PUL found that EPs in the PUL group were too small and probably too early in their natural history to be visualized, rather than being ‘missed’ (Kirk et al., 2008). Suboptimal quality ultrasound equipment, operator inexperience, increased maternal body mass index, uterine fibroids or ovarian pathology may also make visualization of an EP difficult.

Criteria for the ultrasound diagnosis of tubal EP

Endometrium. There is no specific endometrial appearance or thickness that reliably supports a diagnosis of EP (Mehta et al., 1999). In up to 20% cases a collection of fluid may be seen within the endometrial cavity, classically referred to as a ‘pseudosac’ (Marks et al., 1979; Frates and Laing, 1995; Benson et al., 2013). However, an anechoic area in the endometrial cavity is far more likely to be an early IUP (Doubilet and Benson, 2010), and so a presumptive diagnosis of EP should not be made solely on the basis of this finding.

Free pelvic fluid. A small amount of anechoic free fluid in the Pouch of Douglas is commonly found with both intrauterine and ectopic gestations (Nyberg et al., 1991). The presence of echogenic fluid has been reported in 28–56% of women with an EP (Fleischer et al., 1990; Nyberg et al., 1991). This correlates well with the surgical findings of haemoperitoneum (Sickler et al., 1998), but does not confirm tubal rupture, as blood commonly leaks from the fimbrial tube. The amount of echogenic fluid visualized in the pelvis correlates closely with findings at surgery. As a rule of thumb, if fluid reaches the fundus of the uterus or is present in the utero-vesical pouch, it is significant. A further marker of serious intra-abdominal bleeding is the presence of fluid in Morrison’s pouch between the liver and the kidney. This simple examination forms the basis of the focused assessment by sonography for trauma (FAST scan) used in emergency departments (Scalia et al., 1999).

Tubal EP. Tubal EPs are usually positively diagnosed using ultrasound by visualizing an adnexal mass that moves separate to the ovary. The most common finding, in around 60% of cases, is an inhomogeneous or a non-cystic adnexal mass sometimes known as the ‘blob’ sign (Braffman et al., 1994; Shalev et al., 1998; atriet al., 2003; Condous et al., 2005a; Kirk et al., 2008) (Fig. 1). A meta-analysis of 10 studies found that this was the most effective criterion on which to base the diagnosis of a tubal EP (Brown and Doubilet, 1994), with a specificity, positive predictive value, sensitivity and negative predictive value of 98.9%, 96.3%, 84.4% and 94.8%, respectively. The mass is generally spherical, but a more tubular appearance may be seen if bleeding creates a haematosalphinx.

In around 20% of cases it may be possible to visualize an empty extrauterine gestational sac or ‘bagel sign’ (Condous et al., 2005a; Kirk et al., 2008). In another 20%, this sac may contain a yolk sac and/or an embryonic pole that may or may not have cardiac activity (see Fig. 1). The term ‘viable ectopic’ is applied when embryonic cardiac activity is visualized.

Some authors consider that a definitive diagnosis of tubal EP can only be made when an extra-uterine gestation sac containing a yolk sac or embryonic pole is visualized (Barnhart et al., 2011). A recent consensus on nomenclature proposed that the term ‘definite EP’ be used if a yolk sac and/or embryo (with or without cardiac activity) is seen (Barnhart et al., 2011). The term ‘probable EP’ is suggested if an inhomogeneous mass or an extrauterine sac-like structure is visualized. Using this classification, the specificity of ultrasound to detect EP is very high, but the sensitivity is lower as the inhomogeneous and empty sac masses are not counted as definitive EPs.

The rationale for reserving the term ‘definite EP’ to those with embryonic structures is that false-positive diagnoses of EP, though rare, are found even when the criteria for diagnosis are very strict (Krag Moeller et al., 2009), likely due to other pathology such as the presence of pedunculated or broad ligament fibroids, pelvic inflammation or highly exophytic ovarian cysts. This more conservative approach is linked to the need for a very high level of diagnostic certainty in the event that medical treatment is being considered. If methotrexate is used, a false-positive diagnosis of EP may lead to the inadvertent termination of an undetected IUP (Nurmohamed et al., 2011), or severe abnormality in any surviving pregnancy. The American College of Obstetricians and Gynecologists (ACOG) recommend that methotrexate only be administered
to women with a ‘probable ectopic pregnancy’ (as opposed to ‘definitive’) in cases where the hCG has been serially monitored to confirm the change is incompatible with an IUP. The difficult issue is how to define ‘incompatible’. The ACOG defines this as a < 53% rise in hCG over 48 h (Agency for Healthcare Research and Quality, 2008). However, Condous et al. showed that a viable IUP may have an hCG rise of significantly < 50%, so giving methotrexate with an hCG rise of 53% could possibly lead to termination of a wanted pregnancy. An hCG rise of < 35% is now considered a safer definition of non-viability in women with probable EP when methotrexate is being considered (Seeber et al., 2006).

The lower threshold of an initial ultrasound finding of an inhomogeneous mass or empty sac structure is reasonable for the majority of women, when surgical or expectant management is proposed, as the management options do not jeopardize any possible intrauterine gestation. The criteria for ultrasound diagnosis of a tubal EP are summarized in Table I.

Against this background the recent guidelines in the UK published by the National Institute for Clinical Excellence are a major concern. This guidance recommends the use of methotrexate as the first-line treatment for EP; however, no consideration is given to definitively excluding guidance recommends the use of methotrexate as the first-line treatment for EP; however, no consideration is given to definitively excluding the possibility of a viable IUP. The presumption therefore is that a false-negative diagnosis of EP is not possible, which is clearly wrong. If followed, in the authors’ opinions, this guidance will inevitably lead to cases where a viable intrauterine pregnancy is terminated in error (Newbatt et al., 2012; Bourne et al., 2013).

Pulsed Doppler ultrasound has been examined as a tool for diagnosing EP. However, the range of values for vascular indices associated with an EP means that there is significant overlap with those associated with angiogenesis within a corpus luteum (Bourne et al., 1991; Jurkovic et al., 1992; Pellerito et al., 1995). Currently Doppler is not considered to significantly contribute to the diagnosis of EP.

### Surgical diagnosis

Most surgery for EP is now carried out as a therapeutic procedure after an EP has been diagnosed with TVS (Jurkovic and Wilkinson, 2011). Diagnostic surgery is generally reserved for women presenting with signs of an acute abdomen and hypovolaemic shock. A surgical diagnosis may also be made in women with a PUL who become symptomatic. Although laparoscopy is still considered by some to be the ‘gold standard’ for diagnosing EP, this position is now untenable given the high predictive value of ultrasound (Condous et al., 2005a), the widespread use of conservative management strategies and the clinical reality that false-positive and -negative diagnoses of EP occur even at surgery (Li et al., 1991; Atri et al., 1996). If an EP cannot be visualized with TVS performed by an

<table>
<thead>
<tr>
<th>Type of EP</th>
<th>Sonographic criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tubal</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) An inhomogeneous adnexal mass or, an empty extra-uterine gestation sac or a yolk sac or fetal pole + cardiac activity in an extra-uterine sac</td>
</tr>
<tr>
<td>Interstitial</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) Products of conception/gestation sac located in the interstitial (intramyometrial) portion of the tube surrounded by a continuous rim of myometrium</td>
</tr>
<tr>
<td></td>
<td>(3) Interstitial line sign (thin echogenic line extending from a central uterine cavity echo to the periphery of the interstitial sac)</td>
</tr>
<tr>
<td>Cornual</td>
<td>(1) A single interstitial portion of the Fallopian tube in the main uterine body</td>
</tr>
<tr>
<td></td>
<td>(2) Products of conception/gestation sac mobile and separate from the uterus surrounded by the myometrium</td>
</tr>
<tr>
<td></td>
<td>(3) Vascular pedicle joining the gestational sac to the unicorneate uterus</td>
</tr>
<tr>
<td>Cervical</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) Barrel shaped cervix</td>
</tr>
<tr>
<td></td>
<td>(3) Products of conception/gestation sac below the level of the internal cervical os</td>
</tr>
<tr>
<td></td>
<td>(4) Negative sliding organ sign</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td>scar</td>
<td>(2) Products of conception/gestation sac located anteriorly at the level of the internal os covering the presumed site of the previous lower segment Caesarean sectionscar</td>
</tr>
<tr>
<td></td>
<td>(3) Negative sliding organ sign</td>
</tr>
<tr>
<td></td>
<td>(4) Evidence of peritrophoblastic flow on Doppler examination</td>
</tr>
<tr>
<td>Ovarian</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) Cystic structure with a wide echogenic ring on or within the ovary, generally seen surrounded by ovarian cortex and separate from the corpus luteum</td>
</tr>
<tr>
<td>Intramural</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) Products of conception/gestation sac completely surrounded by myometrium and separate from the endometrial cavity</td>
</tr>
<tr>
<td>Abdominal</td>
<td>(1) Empty uterine cavity</td>
</tr>
<tr>
<td></td>
<td>(2) No evidence of a dilated Fallopian tube or complex adnexal mass</td>
</tr>
<tr>
<td></td>
<td>(3) Gestation sac surrounded by loops of bowel and separated by peritoneum</td>
</tr>
<tr>
<td></td>
<td>(4) Wide mobility similar to fluctuation of the sac</td>
</tr>
</tbody>
</table>

Ackerman et al. (1993); timer-Tritsch et al. (1994); Jurkovic et al. (1996); Jurkovic et al. (2003); Gerl et al. (2004); Jermy et al. (2004); Cornstock et al. (2005); Condous et al. (2005a); Laing (2007); Jurkovic and Marvelos (2007).
experienced operator, there is a low chance of finding a visible EP at laparoscopy. In one study 4.5% (2/44) of women were subsequently diagnosed with an EP following a negative laparoscopy (Li et al., 1991). Another study reported a 3–4% false-negative rate and 5% false-positive rate associated with laparoscopy (Atri et al., 1996).

Non-tubal Ectopic Pregnancy

Up to 7% of EP are located outside the Fallopian tube (Bouyer et al., 2002). Although relatively rare, these pregnancies are responsible for significant morbidity and mortality (Lewis, 2007). The diagnostic criteria for non-tubal EP are shown in Table I. Although these criteria exist, there are a paucity of data relating to their diagnostic performance, and care must be taken to avoid misdiagnosis. A common example is of a normal IUP implanted laterally in an arcuate uterus being misdiagnosed as an interstitial pregnancy. Furthermore, new developments in ultrasonography have rendered some of these criteria redundant. Using 3D ultrasound, a coronal view of the uterus can be obtained. This view means that invariably a connection can be seen between the endometrial cavity and interstitial portion of the Fallopian tube (Fig. 2). In the future it seems highly likely that the use of 3D ultrasound in this context will be an integral part of diagnostic guidance (Lawrence and Jurkovic, 1999).

The classification of a Caesarean section scar pregnancy may also be difficult as it is not unusual for a gestation sac to implant over a scar, without extending into the scar (Naji et al., 2013). Given these uncertainties, there is an argument that all such pregnancies over or in a scar should be referred to specialist centres both to confirm the diagnosis and for treatment.

Heterotopic pregnancy

A heterotopic pregnancy occurs when any form of EP is found simultaneously with an IUP (Fig. 3). The incidence of heterotopic pregnancy in natural conceptions was originally estimated to be 1 in 30 000 pregnancies (DeVoe and Pratt, 1948). Clinically they are associated with assisted reproductive techniques, where the incidence is thought to be 1–3 in 100 pregnancies (Molloy et al., 1990; Fernandez and Gervaise, 2004). The risk increases in proportion to the number of embryos transferred with IVF. If more than 4 embryos are transferred the risk has been quoted as 1 in 45 (Dor et al., 1991). In clinical practice, it is important to remember that visualizing an IUP does not exclude the presence of a further pregnancy elsewhere in the pelvis, especially if the pregnancy is the result of IVF.

Figure 2 Interstitial ectopic pregnancy: (a) 2D grey-scale image and (b) 3D image.

Figure 3 Heterotopic pregnancy: (a) 2D grey-scale image and (b) 3D image. The two images are not of the same case.
Pregnancy of unknown location

‘Pregnancy of unknown location’ (PUL) is a descriptive term used to classify a pregnancy when a TVS has shown no signs of either an intrauterine or extraterrestrial pregnancy or retained products of conception. Studies report that 5–42% of women attending for an ultrasound assessment in early pregnancy will be classified as having a PUL (Hahlin et al., 1995; Cacciatori et al., 1996; Banerjee et al., 1999, 2001; Condous et al., 2004a). Reports from specialized early pregnancy units describe lower PUL rates of 8–10% (Kirk et al., 2007b; Cordina et al., 2011). A consensus statement produced by the International Society of Ultrasound in Obstetrics and Gynecology stated that modern units should try to maintain a PUL rate of <15% (Condous et al., 2006a). The PUL rate may be influenced by a number of factors but in particular the quality of the ultrasound performed within a unit is important. Accordingly, it has been proposed as a benchmark for judging the performance of any facility offering early pregnancy ultrasound.

Diagnosis

PUL is not a diagnosis. It is a term used to classify a pregnancy until the final clinical outcome is known. Any woman classified as having a PUL needs follow-up to determine the final clinical outcome, which may be an ongoing viable IUP, a failed pregnancy, an EP or rarely a persisting PUL. The need for appropriate clinical supervision and follow-up when caring for women with a PUL was highlighted in the 2006–2008 Centre for Maternal and Child Enquiries report where a maternal death secondary to a ruptured EP occurred in a woman where a pregnancy described as a PUL was subsequently inappropriately managed (O’Herlihy, 2011). Current thinking has moved away from establishing pregnancy location prior to PUL according to the risk to harm to the patient. Accordingly probable EPs are designated as high risk and probable IUP are classified as low risk alongside PUL that are destined to fail. In this way appropriate follow-up arrangements can be made that reflect the chance of a complication occurring.

Before a pregnancy can be classified as a PUL, it follows that there need to be clear criteria for definitively diagnosing IUP and EP. Unfortunately, these criteria vary worldwide which partly explains why PUL rates and subsequent outcomes differ both between units and in the literature. In a recent review, there was such clinical heterogeneity in the definition of PUL outcome, that only data for the outcome EP could be analyzed (Van Mello et al., 2012). A consensus paper with authors from the USA, UK, Belgium, the Netherlands and Australia also highlighted this problem (Barnhart et al., 2011). The issue relates to when it is considered safe to make a definitive diagnosis of a pregnancy in the uterus or tube. These areas of uncertainty are outlined below.

PUL or early IUP?

When a small empty intrauterine gestational sac-like structure is visualized on TVS, the concern is that it may not be a true gestation sac but a collection of fluid in the endometrial cavity (‘pseudosac’). In the absence of any visible embryonic structures, some operators will classify such a finding as a PUL. To overcome this, Barnhart et al in their consensus paper suggested the following new categorization system for initial ultrasound findings:

(i) **Definite EP:** extraterrestrial gestational sac with yolk sac and/or embryo (with or without cardiac activity).

(ii) **PUL-probable EP:** inhomogeneous adnexal mass or extrauterine sac-like structure.

(iii) ‘True’ PUL: no signs of either an intrauterine or extrauterine pregnancy on TVS.

(iv) **PUL-probable IUP:** intrauterine gestational sac-like structure.

(v) **Definite IUP:** intrauterine gestational sac with yolk sac and/or embryo (with or without cardiac activity).

Generally in the USA and some units in the UK where the broader definition of PUL is used encompassing points 2, 3, and 4 above, a small intrauterine gestational sac without obvious embryonic structures would be classified as a PUL. In other units, the same findings would be regarded as confirmation of an intrauterine gestational sac.

PUL or miscarriage?

When a woman presents with a PUL and a history of heavy bleeding it is tempting to make a diagnosis of complete miscarriage. However, the evidence shows that such cases should be classified as PUL and evaluated further unless a prior ultrasound has confirmed an IUP. In a study of 152 women with such a history, 5.9% were subsequently found to have an underlying EP (Condous et al., 2005b).

Difficulty also arises in women found to have a slightly thickened or irregular endometrium with a possible small amount of retained products of conception. We know that measurements of endometrial thickness or volume on TVS are not good tests for predicting the finding of chorionic villi at curettage, and so are unreliable for diagnosing an incomplete miscarriage (Sawyer et al., 2007), although the use of power Doppler has been reported as being useful for confirming the presence of significant retained products in the cavity (Casikar et al., 2012). There is currently no consensus on the correct management of these cases. Where no definitive retained products of conception are seen, it would seem appropriate to adopt the safer approach and manage these cases as PUL, thus also avoiding the possibility of curettage in the presence of a potentially viable early IUP. It follows that some units with very high PUL rates may be managing a larger proportion of incomplete miscarriages as PUL because of this diagnostic uncertainty.

PUL or EP?

In some protocols, the criteria to definitively diagnose an EP require an extra-uterine gestational sac with a yolk sac or embryo to be visualized (Barnhart et al., 2011). Therefore when an inhomogeneous adnexal mass is seen, an initial classification of PUL will be made. This will lead to more EP being classified as PUL. If a sonographer feels unable to categorically make a diagnosis of EP, it is now recommended that a diagnosis of ‘probable EP’ be made (Barnhart et al., 2011).

These different approaches to the classification of PUL reflect the clinical environment of a given unit. If the quality of ultrasound is inconsistent, patient compliance low, and risk of litigation high, then criteria used to make a definitive diagnosis of an IUP or EP will be more exacting. The result will be that more women will be classified as having a PUL and undergo some sort of supervised follow-up. When ultrasound findings can be relied upon and patients can be expected to return to the clinic in the event of problems, the balance of risk changes and diagnostic criteria may be looser. The stricter criteria for making a definitive diagnosis of EP become particularly relevant when treatment with methotrexate is being considered.
Clinical outcomes

There are four possible clinical outcomes once a pregnancy has been classified as a PUL: an IUP; or failed PUL (low risk of complications) or an EP or persistent PUL (higher risk of complications). The concern for clinicians when a woman is classified with a PUL is that a diagnosis of EP may have been missed. Hence, previously the clinical scenario we now call a PUL would often be categorized as a ‘query ectopic’ or ‘ectopic until proved otherwise’. This is entirely inappropriate as the terms are clearly not synonymous. Evidence shows that from 6% and at most 20% of women with a PUL have an EP (Hahlin et al., 1995; Banerjee et al., 1999, 2001; Condous et al., 2007a, b; Kirk et al., 2007b). Between 30 and 47% of PUL are ultimately diagnosed with an ongoing IUP (Kirk et al., 2006, 2007b; Condous et al., 2007a; Bignardi et al., 2010) with the majority (50–70%) found to be failing pregnancies where the location is never confirmed (Banerjee et al., 1999, 2001; Condous et al., 2006a, 2007a; Kirk et al., 2007b). It is therefore clear that most women with PUL are not at risk of significant complications. Recent management protocols have begun to reflect this, with the emphasis on triage according to risk rather than establishing pregnancy location (Cordina et al., 2011; Van Calster et al., 2013).

A very small proportion of women with a PUL may be given the descriptive term of ‘persisting PUL’. In such cases the hCG level does not spontaneously decline and no intrauterine or EP is identified on follow-up TVS (Condous et al., 2004a). These are likely to be either a small EP that has not been visualized or retained trophoblast in the endometrial cavity. The final classification of a persisting PUL is dependent on the intervention or treatment given and includes: treated persistent PUL, resolved persistent PUL and non-visualized EP. These are explained in more detail in the 2011 consensus statement (Barnhart et al., 2011). A treated persistent PUL is defined as one managed medically with methotrexate without confirmation of the location of the pregnancy by ultrasound, laparoscopy or uterine evacuation. These cases are important as treatment should only be considered when a potentially viable IUP has been definitively excluded (Agency for Healthcare Research and Quality, 2008).

A resolved persistent PUL is defined as resolution of serum hCG to a non-pregnant value after expectant management or after uterine evacuation without evidence of chorionic villi on pathological examination. If the hCG level rises after a uterine evacuation which fails to identify chorionic villi, a diagnosis of a ‘non-visualized EP’ may be made. In the rare event of relatively low and unresolved levels of serum hCG, clinicians should also alert to the possibility of an hCG-secreting tumour being present (Condous et al., 2003; Einenk et al., 2010; Cong et al., 2011). The optimal management of persisting PUL is not known and there is the need for a randomized trial between expectant management, uterine curettage and methotrexate. There is no International Classification of Diseases (ICD-10) code for a final diagnosis of persistent or resolved PUL. A new code has been ascribed to such cases in Denmark and the authors recommend such a classification is adopted internationally.

Serum biochemistry

Human chorionic gonadotrophin. Single measurements of serum hCG are not used to predict the outcome of PUL. A recent meta-analysis confirmed that absolute single levels of hCG perform poorly as a predictor of PUL outcome (Van Mello et al., 2012). There is a common misconception that a single level of serum hCG, for example <1000 IU/l, means an EP is unlikely. This is a false assumption as the majority of EPs detected in modern practice have initial serum hCG values below this level (Condous et al., 2005c). Similarly, a high hCG (for example >2000 IU/l) does not necessarily exclude the presence of a viable IUP, although it makes it less likely (because a miscarriage may be classified as a PUL before the hCG level has declined significantly).

In contrast serial measurement of serum hCG offers good test performance for predicting viability and is certainly better at predicting location than measurements of serum progesterone. The use of the serial hCG level was first described by Kadar and Romero in 1981 who suggested that the minimal rate of increase in serum hCG with a viable IUP was 66% in 48 h. A subsequent study showed the minimum rise in hCG for a viable IUP to be 24% at 24 h and 53% at 48 h (Barnhart et al., 2004a). More recently, a rise of 35% over 48 h was proposed as the minimal rise consistent with a viable IUP (Seeber et al., 2006). In clinical practice many units use a minimum value of between 50 and 66% to indicate an increase in serum hCG compatible with viability (Horne et al., 2011), but this is not sufficiently conservative and may lead to inadvertent termination of a pregnancy should an intervention be carried out. It should also be remembered that some EP also demonstrate a ‘normal’ rise in hCG.

The changes expected in serum hCG levels over 48 h for failing PULs have also been examined. The rate of decline in hCG is dependent on the initial serum hCG level, with a more rapid decline associated with higher starting concentrations (Barnhart et al., 2004b). In a study, the rate of hCG decline ranged from 21 to 35% at 48 h to 60 to 84% at 7 days (Barnhart et al., 2004b). The change in serum hCG over 48 h may be referred to as the hCG ratio (hCG at 48 h/hCG at 0 h). A decrease in hCG of >13% or an hCG ratio of <0.87 has a sensitivity of 92.7% and specificity of 96.7% for predicting a failing PUL (Condous et al., 2006b).

Unfortunately, there is no single method to characterize the pattern of serum hCG change over 48 h in women classified with a PUL who are subsequently diagnosed with an EP (Silva et al., 2006). Up to 15–20% will have serum hCG doubling times similar to that of an IUP whilst around 10% EP will have hCG pattern behaving like a failing PUL (Goldstein, 2006; Silva et al., 2006). The majority though, will have serial serum hCG levels that increase more slowly than would be expected with an IUP (‘suboptimal rise’) or decrease more slowly than would be expected with a failing PUL. The sensitivity of the serum hCG ratio to predict EP ranges from 74 to 100% and the specificity ranges from 28 to 97% (Van Mello et al., 2012).

Progesterone. The serum progesterone level is a good indicator of early pregnancy viability, but a poor predictor of location. Levels of <20 nmol/l have a high positive predictive value for failing pregnancies (Banerjee et al., 2001), whilst levels >25 nmol/l are ‘likely to predict’ and levels >60 nmol/l are ‘strongly associated with’ viable pregnancies (RCOG, 2006). A meta-analysis of 26 studies was performed to assess the accuracy of a single progesterone level for the diagnosis of EP (Mol et al., 1998). This study found that progesterone did not have sufficient discriminative capacity to diagnose EP with certainty.

The utility of progesterone is in fact in selecting women with failing PUL as being at low risk and so needing reduced follow-up. Whether these failing PUL are failing EP or failing IUP is not of clinical importance. A recent systematic review and meta-analysis (5 studies with 1998 participants and cut-off values from 3.2 to 6 ng/ml) has shown that a single
progesterone level predicts a non-viable pregnancy with a pooled sensitivity of 74.6% (95% confidence interval 50.6–89.4%), specificity of 98.4% (90.9–99.7%), positive likelihood ratio of 45 (7.1–289) and negative likelihood ratio of 0.26 (0.12–0.57) (Verhaegen et al., 2012).

Other serum markers. Creatine kinase, cancer antigen 125, activin A, inhibin A, inhibin pro-αC-related immunoreactivity and insulin-like growth factor-binding protein have all been evaluated (Condous et al., 2005c; Kirk et al., 2009; Chetty et al., 2011). Of these, inhibin A may be useful for predicting spontaneous resolution of PUL, but is not as good as progesterone (Kirk et al., 2009; Chetty et al., 2011).

Finding a novel marker to better discriminate EP from the rest of the PUL population nevertheless remains an important area of research. Recent attention has focused on metalloprotease 12 (ADAM 12) (Rausch et al., 2011) and fibronectin (Brown et al., 2013) as potentially promising candidate markers.

Mathematical prediction models
A number of mathematical models have been developed for the prediction of PUL outcome. These include logistic regression models and Bayesian networks, based on variables such as serum hCG and progesterone levels, endometrial thickness and the amount of vaginal bleeding (Banerjee et al., 2001; Condous et al., 2004b, 2007a; Gevaert et al., 2006). These models have been shown to have high sensitivities for the prediction of PUL outcome and can be used to rationalize follow-up (Kirk et al., 2007b). The most widely evaluated model developed is M4 (Condous et al., 2007a), which takes into account the fact that the initial hCG level adds information regarding the likely outcome as well as the hCG ratio.

Rationalizing follow-up
Initially mathematical models were optimized to identify EPs in the PUL population. The main aim for their use now is to identify PUL that are ‘low risk’ (failing PULs and IUPs) in order to safely reduce follow-up for these pregnancies whilst focusing resources towards pregnancies at increased risk of being an EP. In a multi-centre diagnostic accuracy study of 1962 patients using the model M4 to predict outcome, 62–75% of PULs were considered low risk of which 96–98% were confirmed to have either a failed PUL or IUP (Van Calster et al., 2011). In contrast, 80–92% of EPs were considered high risk. In the majority of EP, allocated to the low-risk group, the hCG ratio was declining and so it is likely that these EP would have resolved without the need for intervention.

It is interesting to contrast this approach with the PUL triage study using single measurements of progesterone evaluated by Cordina et al. (2011). In this study a serum progesterone of < 10 nmol/l was used to define a failing pregnancy. Of 676 PUL, 252 were classified as failing PUL with this approach and only 4 of these needed intervention. These results, although not externally validated, are of interest, though with 252 failing PUL in the study, and with 10% of these cases lost to follow-up, it is hard to make a definitive comment on safety. These results contrast with those of Condous et al. (2005a). He defined a low-risk PUL as one with a single hCG of < 25 IU/l or progesterone of < 10 nmol/l. It was found that when this approach was applied retrospectively, 84% of non-ectopic or failing pregnancies could have been discharged from the system, but at a cost of 67% of EP in the study population potentially having no further follow-up. This study suggests a single visit strategy should be approached with caution pending external validation of the approach.

The problem with the use of serum progesterone is that it classifies many viable IUP as high risk, which clearly limits the ability to function as an effective triage tool. The M4 model, on the other hand, classifies both failing pregnancies and viable IUP as low risk, which greatly enhances its utility for triage. Looking forward, it seems likely that a system using both progesterone and hCG as a two-step procedure will be the optimal approach which will enable some failing PUL to be triaged at the initial visit, and the residual pregnancies should be triaged at 48 h after application of the M4 model. The M4-based protocol can be accessed from the following link: http://homes.esat.kuleuven.be/~biomed/M4PUL/M4triage.htm.

Surgical intervention
The combination of the absence of a visualized IUP using TVS and a serum hCG level above a set discriminatory zone historically was considered an indication for diagnostic laparoscopy. This is no longer the case. Laparoscopy is now rarely indicated unless a woman is symptomatic or haemodynamically unstable. This is because many failing PUL and some IUP may have initial hCG levels above, and some clinically significant EP will have hCG levels below, given any discriminatory zone. Condous et al. (2005d) evaluated the diagnostic accuracy of discriminatory zones for hCG levels of > 1000, 1500 and 2000 IU/l for the detection of EP in women with a PUL. It was found that as the discriminatory zone increases, the sensitivity to detect EP decreases, and most EP in the population of women with PUL studied had serum hCG levels below the often-quoted discriminatory zone of 1000 IU/l.

Uterine curettage is used by some clinicians to diagnose pregnancy location with the aim of differentiating between an EP and a non-viable IUP. Advocates of curettage report that a failure to definitively ascertain the location of a pregnancy may have implications for counseling about future risk of EP and lead to unnecessary exposure to methotrexate (Chung et al., 2011). However, there is the risk that a potentially ongoing IUP may be terminated if the preoperative criteria are not stringent enough to exclude a viable IUP before the procedure. Recognized indications for curettage include: (i) no visible IUP on TVS with a serum hCG > 2000 IU/ml; (ii) an abnormal rise in hCG level, defined as < 50% increase in 2 days and (iii) an abnormal fall in hCG level, defined as < 20% decline in 2 days (Chung et al., 2011). In a study of 321 women who underwent uterine curettage with no visible IUP or those with an abnormal hCG trend, 73.2% were ultimately diagnosed with an EP whilst 26.8% were found to have a non-viable IUP (Chung et al., 2011). These and other biochemical criteria that define non-viability in the PUL population have been evaluated to establish if the criteria used could result in inadvertent termination of a pregnancy if curettage was performed (Condous et al., 2006c). Up to 12.3% of women undergoing curettage were deemed at risk of having a viable IUP interrupted (Condous et al., 2006c). Uterine curettage should therefore not be used routinely in the management of PUL but may have a role in diagnosing the location of failing PULs, once the possibility of a potentially viable IUP has been excluded.

Clinical management
The majority of women categorized with a PUL have minimal or absent symptoms and are haemodynamically stable. Expectant outpatient
All compliant women with a PUL should be counselled about the possible outcomes and be given written information detailing the plan for further investigation and when and how to seek urgent medical care during any follow-up period. The majority of women classified with a PUL will be subsequently diagnosed with a failed pregnancy, resolving spontaneously without intervention (Banerjee et al., 1999). There is no consensus on what is an acceptable surgical intervention rate in women classified with a PUL (Condous et al., 2006a), although studies report rates of between 0.5 and 11% (Banerjee et al., 2001; Condous et al. 2003; Kirk et al., 2007b; Cordina et al., 2011). It is evident that the management of PUL needs to concentrate resources and surveillance on women with the highest risk of significant pathology, whilst minimizing intervention and follow-up in women at low risk of complications.

Most women with a PUL are followed up with serum hCG measurements and repeat TVS examination until a final diagnosis is confirmed. Women with PUL should be triaged into low- and high-risk groups that require different levels of intensity for follow-up. By using the M4 model or hCG ratio, PUL may be classified as being ‘low-risk failing’ or ‘low-risk intrauterine’. ‘Low-risk failing’ may be followed up with a urinary pregnancy test after 2 weeks and telephone advice. A woman with a ‘low-risk intrauterine’ may return for an ultrasound scan to assess viability in 2 weeks. In the event of being classified as high risk, further assessment is required either with a further ultrasound scan within 48 h or a further hCG measurement. Figure 4 summarizes suggested management algorithms for PUL.

**Future work**

The ideal diagnostic tool for an EP would be a single serum marker to replace ultrasound and serial biochemistry (Shaw et al., 2010a). While the major focus of work until now has been on diagnosing all EP, reliable tools are needed to identify those women with EP or PUL who do not require active intervention.

The common assumption is that earlier diagnosis of EP means more effective management, because more conservative management options may be employed (Hajenius et al., 2007), although this is not borne out in studies thus far (Fernandez et al., 2013; Van Mello et al., 2013). Until robust evidence suggests that medical treatment leads to lower subsequent EP rates or higher subsequent IUP rates, then the overall benefit from earlier diagnosis is unproved providing the diagnosis is made prior to tubal rupture. It is likely that women with suspected or definite EP that would otherwise spontaneously resolve are currently being over-treated.

---

**Figure 4** Possible PUL management algorithm. *Women must be given information and advice, be deemed compliant with follow-up and have no significant language or other communication barrier. Women should be advised to return for review before the scheduled follow-up visit if they have any severe pain or concerns. **If hCG > 1000 IU/l at 0 h and history not suggestive of complete miscarriage, then repeat TVS as soon as possible. Adapted from Condous et al. (2006b); Kirk et al. (2007b). TVS, transvaginal ultrasound.*
Conclusion

The primary assessment approach for a woman with a possible EP is ultrasonography. This should be carried out transvaginally and all recent data on diagnosing EP, defining PUL, performance of discriminatory zones and diagnosing miscarriage relate to this approach. Ultrasound is operator dependent and requires appropriate training and significant experience for good results to be obtained. We have shown that where there is easy access to expertise and equipment to provide high-quality TVS, the vast majority of women with tubal EPs may be diagnosed rapidly and accurately. The diagnosis of non-tubal EP can also be achieved through careful ultrasound assessment, although diagnosis is more often delayed leading to increased morbidity.

The need for a single serum marker for EP and a method to differentiate between women with PUL and EP who do and do not require intervention are the current diagnostic challenges. The current trend in the management of PUL is thus directed to detecting the women with ‘low-risk’ PUL who can undergo minimal follow-up whilst concentrating resources on the women with ‘high-risk’ PUL in whom the majority of harmful EPs will be found. In the meantime, algorithms such to manage women with PUL safely, consistently and with minimal unnecessary intervention should be adopted.

Authors’ roles

All three authors (E.K., C.B and T.B.) made substantial contributions to the plan of the article, review and interpretation of the literature and drafting and revision of the manuscript. All authors approved the final version to be published.

Funding

T.B. is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the authors and not necessarily reflect those of the NHS, the NHS Trust and Imperial College London. The views expressed are those of the authors and not necessarily reflect those of the NHS, the NIHR or the Department of Health.

References


Baikken IJ, Skjeldstad FE. Reduced number of extrauterine pregnancies—increased fertility of women during the 1990s. Tidsskr Nor Laegeforen 2003;123:3011–3014.


