The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

# MINIMISING DIAGNOSTIC UNCERTAINTIES IN EARLY PREGNANCY

Alison Richardson
MBChB, MA, MRCOG

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

August 2017

‘Medicine is a science of uncertainty and

an art of probability’

*Sir William Osler*

# Abstract


## Introduction

Approximately one in five women experience abdominal pain and/or vaginal bleeding in early pregnancy. This usually prompts referral to an Early Pregnancy Assessment Unit where an ultrasound scan will be performed. Following the ultrasound, either a certain or uncertain diagnosis will be made. Certain diagnoses may be positive, i.e. a viable intrauterine pregnancy, or negative, i.e. a non-viable or ectopic pregnancy. Uncertain diagnoses occur when there is ambiguity regarding either the location or the viability of the pregnancy. Up to 25% of women attending an Early Pregnancy Assessment Unit are given such a diagnosis at their initial visit.

All women with a diagnosis of either a pregnancy of unknown location or uncertain viability need to be followed-up until a definitive diagnosis can be made. At present this is haphazard and protracted, commonly taking up to two weeks to resolve and requiring multiple visits for various different investigations. This takes a considerable amount of time and costs a not insignificant amount of money. Furthermore, in the time taken to make a definitive diagnosis, a stable woman with an unknown miscarriage or ectopic pregnancy may become unstable and require immediate resuscitation, life-saving blood transfusion and/or emergency surgery.

## Aims

The aim of this PhD was to develop methods to minimise the number of women given uncertain diagnoses in early pregnancy, or to at least minimise the duration of uncertainty if the diagnosis is unavoidable. Several different studies were undertaken in an attempt to accomplish this.

## Studies

We initially undertook a prospective cohort study to determine the levels of anxiety generated by uncertain diagnoses in early pregnancy was undertaken. Women with uncertain diagnoses were found to be significantly more anxious (as measured using the standardized short form of Spielberger's

state-trait anxiety inventory) than their counterparts given certain diagnoses (23±0.79 versus 14±6.6), even if these certain diagnoses were not associated with an ongoing pregnancy. This study served to further justify the main objective.

We then performed a systematic review and meta-analysis to identify and determine the diagnostic accuracy of various different ultrasonographic features to predict (a) an intrauterine pregnancy prior to visualization of embryonic contents and (b) a tubal ectopic pregnancy in the absence of an obvious extra-uterine embryo. This study identified the double decidual sac sign as a potential marker to be able to accurately differentiate a true gestation sac from a pseudosac with a sensitivity of 82% (95% CI, 68-90%), specificity of 97% (95% Ci, 76-100%), positive likelihood ratio of 30 (95% CI, 2.8-331) and negative likelihood ratio of 0.19 (95% CI, 0.10-0.35). The quality of the studies included in the meta-analysis however precluded the use of the double decidual sac sign in clinical practice without further validation

As a consequence, we carried out a prospective study following STARD guidelines to determine the diagnostic accuracy of the double decidual sac sign to predict an intrauterine pregnancy prior to visualization of embryonic contents using modern, high-resolution transvaginal ultrasound. This study found that the double decidual sac sign predicted an intrauterine pregnancy with a sensitivity of 94% (95% Ci, 85-98%), specificity of 100% (95% CI, 16-100%) and overall diagnostic accuracy of 94% (95% CI, 88-100%). The positive and negative predictive values are 100% (95% CI, 94-100%) and 33% (95% CI, 4.3-78%) respectively whilst the positive likelihood ratio was infinite and the negative likelihood ratio was 0.06 (95% CI, 0.02-0.16). These results suggest that the meta-analysis under-estimated the ability of the double decidual sac sign to differentiate between a true gestation sac and a pseudosac.

Subsequently, we conducted a study incorporating off-line analysis of ultrasonographic images to determine the inter- and intra-observer reliability of the double decidual sac sign to predict an intrauterine pregnancy prior to ultrasonographic visualization of embryonic contents. This involved fifteen observers from around the United Kingdom remotely assessing a selection of two-dimensional images from 25 of the diagnostic accuracy study participants. There was significant ($p<0.01$) agreement amongst the observers but the level of agreement was only 'fair', reflected by kappa statistics of 0.25,

0.33 and 0.21. Following a period of focused training, the inter-observer reliability significantly increased demonstrated by kappa statistics of 0.70, 0.63 and 0.53. The intra-observer reliability ranged from 'substantial' (K=0.65) to 'almost perfect' (K=0.92). These results demonstrate that the double decidual sac sign has the potential, after training, to be both reliable and precise, making it a very useful ultrasonographic sign in clinical practice.

Finally, we undertook a prognostic research study, following REMARK recommendations, investigating the ability of five serum biomarkers to predict pregnancy outcome in women with pregnancies of uncertain viability. Candidate biomarkers included Angiopoietin-1 (Ang-1), Angiopoietin-2 (Ang-2), soluble FMS-like Tyrosine Kinase-1 (Flt-1), serum TNF-Related Apoptosis Inducing Ligand and Interleukin-15. Serum concentrations of Ang-2 and Flt-1 were significantly lower in pregnancies of uncertain viability that were subsequently proven to be viable than those that were subsequently proven to be non-viable (Ang-2 1510pg/ml versus 2365pg/ml and Flt-1 103pg/ml versus 202pg/ml). Furthermore, there were statistically significant ($p<0.01$), linear (p-value for trend <0.01) associations between Ang-2 and Flt-1 concentrations and subsequent pregnancy viability such that women with a pregnancy of uncertain viability and Ang-2 concentrations greater than or equal to 2666pg/ml were 64% less likely to have a viable pregnancy than those with Ang-2 concentrations less than or equal to 1382pg/ml and women with a pregnancy of uncertain viability and Flt-1 concentrations greater than or equal to 142pg/ml were 50% less likely to have a viable pregnancy than those with Flt-1 concentrations less than or equal to 87pg/ml. These findings suggest that Ang-2 and Flt-1 may be useful in the prediction of pregnancy viability in cases of uncertainty.

# Discussion

One of the biggest challenges in early pregnancy ultrasonography is accurate differentiation between a true gestation sac and a pseudosac. Pseudosacs, although rare, are strongly suggestive of an ectopic pregnancy, hence it is an important distinction to make, ideally as soon as possible. Both appear initially as intrauterine fluid collections or 'empty sacs'. Whilst experts may claim that it is not difficult to differentiate between the two structures, in clinical practice, many of the individuals undertaking the scans in early pregnancy do not claim to be experts. Traditional teaching has always been to wait until either a yolk

sac or fetal pole are visualized within the sac before confirming a definite intrauterine pregnancy. Although safe, inherent with this approach is that an intrauterine fluid collection is visible using transvaginal ultrasound from around day 28 but a yolk sac is not visible until at least day 35. If an ultrasound is undertaken during this time, an 'empty sac' will be seen and uncertainty will ensue.

Application of the results from the studies described above could potentially revolutionize the care of women with diagnostic uncertainties in early pregnancy. Firstly, the confirmation that uncertain diagnoses in early pregnancy are highly anxiogenic, means that Early Pregnancy Assessment Units can now justify the allocation of resources to help alleviate this distress. This is crucial if we are to improve the holistic nature of the care provided to women with complications of early pregnancy.

Furthermore, the discovery that the double decidual sac sign can accurately predict an intrauterine pregnancy prior to visualization of embryonic contents (and therefore effectively exclude an ectopic pregnancy) means that we can rationalise follow-up, improve consistency and minimise error in the management of women with ultrasonographic evidence of an empty sac in early pregnancy.

Although it could be argued that utilization of the double decidual sac sign does not minimise the number of women given uncertain diagnoses in early pregnancy, merely swap concerns regarding location to ones regarding viability, in clinical practice it is the potential consequences of pregnancies of unknown location that are most hazardous, both because of the immediate threat to health and the future threats to fertility. Furthermore, if the findings from our prognosis study are confirmed, and appropriate threshold levels for our biomarkers determined, it may be possible to minimise the duration of uncertainty for women with pregnancies of uncertain viability to hours rather than weeks.

Using a combination of approaches therefore, we have achieved the overall aim of this thesis in minimising diagnostic uncertainties in early pregnancy, the clinical benefits of which are multifold. Not only does it reduce anxiety for women, but also prevents unnecessary investigations from being performed in those with an intrauterine pregnancy and minimise morbidity and mortality,

permit earlier, potentially less invasive intervention and possibly preserve future fertility for women with an ectopic pregnancy.

# Publications

Richardson, A., Deb, S., Campbell, B. and Raine-Fenning, N. *Serum concentrations of Ang-2 and Flt-1 may be predictive of pregnancy outcome in women with pregnancies of uncertain viability: a phase I exploratory prognostic factor study* J Obstet Gynecol (in press)

Richardson, A., Raine-Fenning, N., Deb, S., Campbell, B. and Vedhara, K. *Anxiety associated with diagnostic uncertainties in early pregnancy* Ultrasound Obstet Gynecol; 2017; **50**(2):247-254

Richardson, A., Hopkisson, J., Campbell, B. and Raine-Fenning, N. *Use of the double decidual sac sign to confirm intra-uterine pregnancy location prior to ultrasonographic visualization of embryonic contents: a diagnostic accuracy study* Ultrasound Obstet Gynecol; 2017; **49**:643-648

Richardson, A., Gallos, I., Dobson, S., Campbell, B.K., Coomarasamy, A. and Raine-Fenning, N. *Accuracy of first trimester ultrasound in the diagnosis of an ectopic pregnancy in the absence of an extra-uterine embryo: a systematic review and meta-analysis* Ultrasound Obstet Gynecol; 2016; **47**:28-37

Richardson, A., Gallos, I., Dobson, S., Campbell, B.K., Coomarasamy, A. and Raine-Fenning, N. *Accuracy of first trimester ultrasound in the diagnosis of an intra-uterine pregnancy prior to the development of a yolk sac: a systematic review and meta-analysis* Ultrasound Obstet Gynecol; 2015; **46**:142-149

# Presentations

## Fertility 2017: 10th Joint Conference of UK Fertility Societies (BFS, ACE and SRF): Edinburgh, UK, January 2017

An ectopic pregnancy cannot be excluded… Or can it? (Oral)

Serum concentrations of Ang-2 and Flt-1 are predictive of outcome in women with pregnancies of uncertain viability (Poster)

## Association of Early Pregnancy Assessment Units Annual Meeting: Cardiff, UK, November 2016

Anxiety associated with diagnostic uncertainties in early pregnancy (Oral)*

Serum concentrations of Ang-2 and Flt-1 are predictive of outcome in women with pregnancies of uncertain viability (Oral)

An ectopic pregnancy cannot be excluded… Or can it? (Oral)

## Midlands Academy of Medical Sciences Research Festival: Leicester, UK, April 2016

An ectopic pregnancy cannot be excluded… Or can it? (Poster)

## Royal College of Obstetricians and Gynaecologists Annual Academic Meeting: London, UK, March 2016

Is it possible to minimise the duration of uncertainty for women with a PUV? (Poster)

An ectopic pregnancy cannot be excluded… Or can it? (Poster)

## University of Nottingham Research Showcase Forum: Nottingham, UK, June 2015

Minimising diagnostic uncertainties in early pregnancy (Poster)

## University of Nottingham, School of Medicine Postgraduate Research Forum, Nottingham, UK, May 2015

Minimising diagnostic uncertainties in early pregnancy (Oral)**

## International Society of Ultrasound in Obstetrics and Gynecology Annual Meeting: Barcelona, Spain, September 2014

Accuracy of early ultrasound in the diagnosis of intrauterine pregnancy prior to development of a yolk sac: a systematic review and meta-analysis (Oral)

Accuracy of early ultrasound in the diagnosis of ectopic pregnancy in the absence of an extra-uterine embryo: a systematic review and meta-analysis (OraL)

# Prizes

*The Durbin PLC Annual Award for the Best Original Research Oral Presentation at the Annual Meeting of the Association of Early Pregnancy Units, Cardiff, November 2016

**The Sue Watson Prize for the best oral presentation at the University of Nottingham, School of Medicine Postgraduate Research Forum, May 2015

# Abbreviations

All abbreviations or acronyms are detailed here as well as in first use within the thesis

| | |
|---|---|
| AM | Adnexal Mass |
| Ang-1 | Angiopoietin-1 |
| Ang-2 | Angiopoietin-2 |
| AUC | Area Under Curve |
| BAI | Beck Anxiety Inventory |
| CA125 | Cancer Antigen 125 |
| CI | Confidence Interval |
| CONSORT | CONsolidation Standards Of Reporting Trials |
| CRS | Chorionic Rim Sign |
| CT | Computerized Tomography |
| DDSS | Double Decidual Sac Sign |
| EGA | Estimated Gestational Age |
| ELISA | Enzyme-Linked ImmunoSorbent Assay |
| EP | Ectopic Pregnancy |
| EPAU | Early Pregnancy Assessment Unit |
| ERPC | Evacuation of Retained Products of Conception |
| EU | Empty Uterus |
| FF | Free Fluid |
| FLDA | Functional Linear Discriminant Analysis |
| FN | False Negative |
| FP | False Positive |
| GS | Gestation Sac |
| GSD | Gestation Sac Diameter |
| GnRH | Gonadotropin Releasing Hormone |
| hCG | human Chorionic Gonadotropin |
| IDS | Intradecidual Sign |
| IL-15 | Interleukin-15 |
| IUFC | Intrauterine Fluid Collection |
| IUP | Intrauterine Pregnancy |
| LMP | Last Menstrual Period |
| LR- | Negative Likelihood Ratio |
| LR+ | Positive Likelihood Ratio |

| | |
|---|---|
| M | Mean difference |
| MRI | Magnetic Resonance Imaging |
| MSD | Mean Sac Diameter |
| NPV | Negative Predictive Value |
| p-value | Probability value |
| PPV | Positive Predictive Value |
| PS | Pseudosac |
| PUL | Pregnancy of Unknown Location |
| PUV | Pregnancy of Uncertain Viability |
| PVB | Per Vaginal Bleeding |
| QMC | Queen's Medical Centre |
| QUADAS | Quality Assessment of Diagnostic Accuracy Studies |
| RCOG | Royal College of Obstetricians and Gynaecologists |
| REC | Research Ethics Committee |
| REMARK | REporting recommendations for tumour MARKer prognostic studies |
| ROC | Receiver Operating Characteristic |
| SE | Standard Error |
| sFlt-1 | serum FMS-like tyrosine kinase-1 |
| SSF | Standardized Short Form |
| STAI | State Trait Anxiety Inventory |
| STARD | STAndards for reporting of Diagnostic accuracy studies |
| TAS | Transabdominal Ultrasound |
| TN | True Negative |
| TP | True Positive |
| TRAIL | TNF-Related Apoptosis Inducing Ligand |
| TVS | Transvaginal Ultrasound |
| UPT | Urinary Pregnancy Test |
| US | Ultrasound |
| YS | Yolk Sac |

# Acknowledgements

I would like to thank several key individuals, without whom, this work would not have been possible: my supervisors, Dr Nick Raine-Fenning and Professor Bruce Campbell for providing me with this opportunity and for their support, advice and encouragement; Professor Kavita Vedhara, for her expertise, guidance and friendship; Dr Shilpa Deb for her co-operation; Professor Arri Coomarasamy and Ioannis Gallos from the University of Birmingham, and Dr Sam Dobson for their invaluable help with the systematic reviews; Catherine Pincott-Allen for her assistance with the biomarker assays; Andrea Venn for her input with the statistical analyses; and all the staff and patients at Nurture Fertility and the Early Pregnancy Assessment Unit at the Queen's Medical Centre, Nottingham.

# Table of Contents

# List of Figures

# List of Tables

# List of Appendices

# 1. Normal Early Pregnancy Development

## 1.1 Introduction

It is important to have an understanding of both normal and abnormal early pregnancy development before attempting to understand diagnostic uncertainties in early pregnancy, which is the focus of this thesis. This chapter will first discuss the normal development of an embryo from fertilization to implantation and then subsequent development of the amniotic cavity, yolk sac and chorionic cavity and formation of the trilaminar embryonic disc. It will then go on to describe the normal ultrasonographic development of an early pregnancy including the sequence of events in which structures such as the gestation sac, yolk sac, fetal pole and fetal heart become visible. Finally, the two principal hormones involved in normal early pregnancy development and their role in clinical practice will be discussed.

*NB; In section 1.2 the term 'days' refers to the number of days following fertilization. In all other sections of this thesis, the term days refers to the menstrual age i.e. the number of days since the first day of the last menstrual period. Fertilization is conventionally said to occur at a menstrual age of fourteen days.*

## 1.2 Embryology

Embryogenesis is a complex and carefully coordinated process that begins following the fusion of definitive male and female gametes.

### Fertilization

Fertilization occurs in the ampulla of the fallopian tube when a viable spermatozoon, having penetrated its way through the cumulus, is able to bind to the zona pellucida surrounding an oocyte. Upon binding, the acrosome in the head of the sperm releases digestive enzymes that enable the sperm to

penetrate the zona pellucida. Following penetration, the cell membranes of the spermatozoon and oocyte fuse, which not only initiates a sequence of events that ultimately results in the zona becoming impenetrable by additional spermatozoa but also enables the oocyte to complete meiosis and develop into a definitive oocyte. Within the fertilized oocyte, the nuclei of the sperm and oocyte swell to form male and female pronuclei. Their nuclear membranes subsequently disappear as both maternal and paternal chromosomes are replicated in preparation for mitosis.

## Cleavage

Following fertilization, a rapid series of mitotic cell divisions called cleavage is initiated within the fertilized oocyte (zygote). During cleavage, the embryo remains enclosed within the zona pellucida and does not increase in size at all as cell growth does not accompany the cell division at this stage. The first cleavage division therefore splits the zygote into two smaller daughter cells called blastomeres. The second division, which is complete within 40 hours of fertilization, produces four blastomeres, which are smaller still. By day three, the embryo consists of six to twelve cells and by day four, the morula as it is now known, consists of sixteen to 32 cells.

## Compaction

During the process of compaction, which occurs simultaneously with cleavage, some blastomeres segregate to the centre of the morula and others to the outside. The centrally placed blastomeres constitute the inner cell mass, which ultimately gives rise to the embryo, whereas the blastomeres at the periphery form the trophoblast, which develops into the fetal component of the placenta.

## Blastocyst Formation

From day four onwards, the morula begins to absorb fluid. This is possible because the developing trophoblast cells express a membrane ion channel which transports sodium into and potassium out of the morula. Water subsequently follows the sodium by osmosis. As the hydrostatic pressure increases, a large fluid cavity develops within the morula, which is now

referred to as a blastocyst. The inner cell mass or embryoblast then forms a compact mass at one side of this cavity and the trophoblast organizes into a thin, single-layered epithelium.

## Hatching

Within three or four days of fertilization, the morula, still surrounded by the zona pellucida, enters the uterus. By day five, the blastocyst hatches from the zona pellucida by enzymatically boring a hole in it and squeezing out. The blastocyst is now relieved of all its original investments and is able to directly interact with the endometrium.

## Implantation

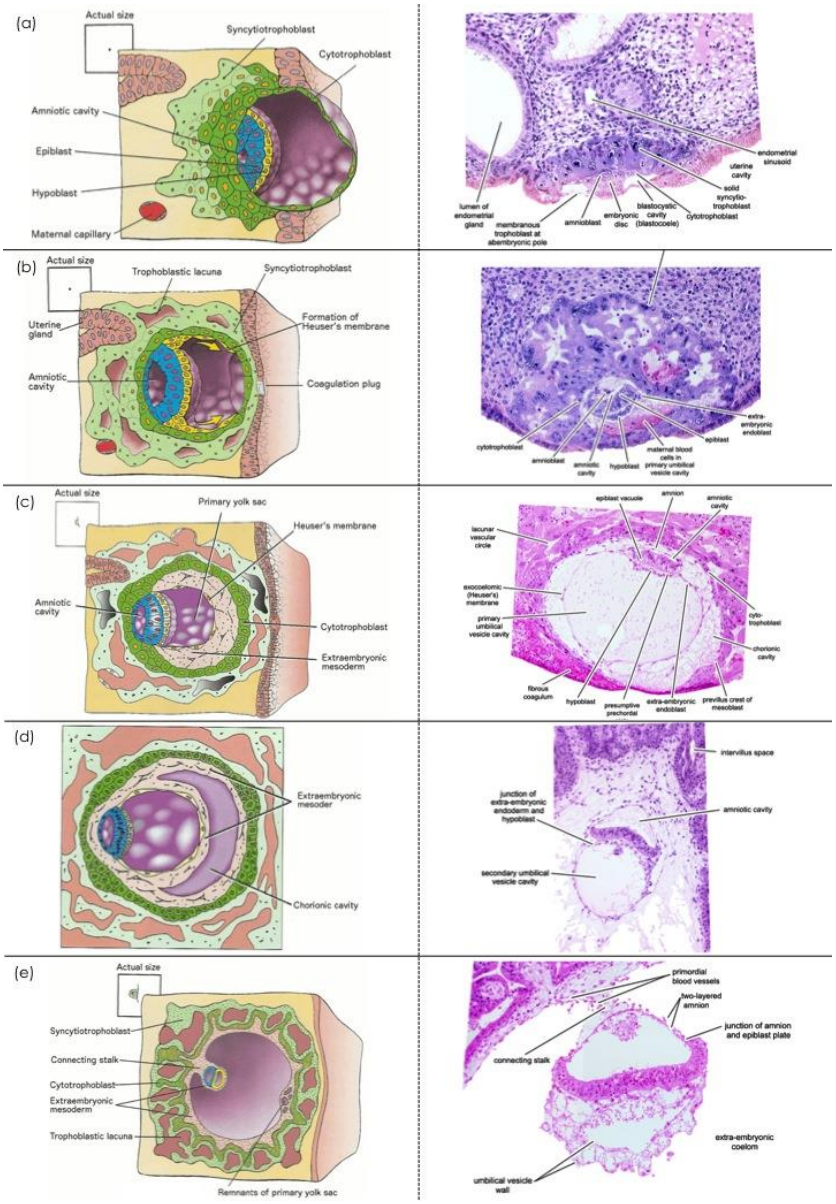
Upon entering the uterus, the blastocyst adheres to the lining of the womb. The endometrial stromal cells respond to both the presence of the blastocyst and the progesterone secreted by the corpus luteum by differentiating into secretory decidual cells. This response is called the decidual reaction. In addition to this, nearby endometrial glands enlarge and the local uterine wall becomes more highly vascularized and oedematous. The decidual cells and endometrial glands secrete growth factors and metabolites that support the development of the implanting embryo. Once the embryo has implanted, the trophoblast cells produce the hormone human chorionic gonadotropin (hCG), which supports the corpus luteum and thus maintains the supply of progesterone. The corpus luteum continues to secrete progesterone for up to twelve weeks, after which time the placenta itself takes over progesterone production and the corpus luteum slowly involutes, becoming a corpus albicans.

Contact with the endometrium induces the trophoblast at the embryonic pole (the side of the blastocyst containing the inner cell mass) to proliferate. Some of these proliferating cells lose their cell membranes and coalesce to form a mass of cytoplasm containing numerous dispersed nuclei and this is called the syncytiotrophoblast.

In contrast, the cells of the trophoblast that line the wall of the blastocyst, retain their cell membranes and constitute the cytotrophoblast. The

syncytiotrophoblast increases in volume throughout the second week as cells detach from the proliferating cytotrophoblast at the embryonic pole and fuse with the syncytium (Figure 1.1a).

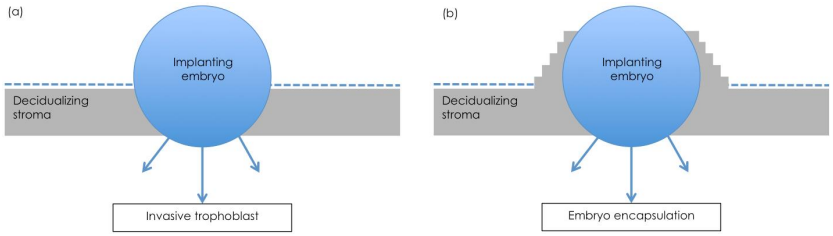
***Figure 1.1: Appearances (a) seven, (b) nine, (c) eleven, (d) thirteen and (e) fourteen days after fertilization (Images on the left taken from Larsen's Human Embryology [1] and images on the right courtesy of The Virtual Human Embryo project (www.prenatalorigins.org/virtual-human-embryo/)***



Nine days after fertilization, the embryo is fully implanted within the endometrium (Figure 1.1b). Active finger like projections extend from the syncytiotrophoblast and penetrate between the endometrial cells further pulling the embryo into the endometrium. As implantation progresses further, the expanding syncytiotrophoblast envelops the entire blastocyst excluding a small region at the abembryonic pole (the side of the blastocyst opposite the inner cell mass). A coagulation plug, seals this small hole where the blastocyst implanted, temporarily marking this point in the endometrial epithelium.

This stepwise process of implantation described above, involving apposition and adherence of a blastocyst to the endometrium followed by breaching of the luminal epithelium and finally invasion of maternal tissues, has been conceptualized largely on the basis of animal models [2]. This process is analogous to an invading cancer, with the embryo driving the destruction of endometrial epithelial cells, the enzymatic digestion of the stromal matrix, and finally, the invasion of the maternal decidua and inner myometrium (Figure 1.2a) [3].

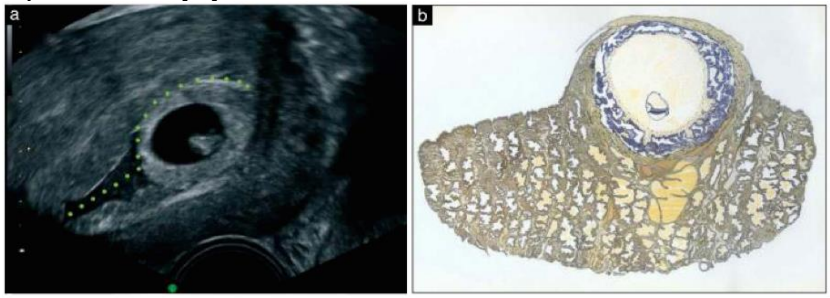
***Figure 1.2: The (a) traditional and (b) emerging view of implantation (adapted from [3])***




Notwithstanding the complexity at a cellular level, implantation in the conventional paradigm outlined above seems to require little more than a receptive endometrium and an invasive embryo capable of evading maternal immune detection. This may be true for certain species, for example the mouse [2], but recent studies [4-10] indicate that implantation of a human embryo is much more dynamically controlled by the endometrium than previously appreciated: it is not simply a case of having an invasive embryo, capable of evading maternal immune detection, and a receptive endometrium. Emerging concepts suggest that human embryos do not

embed in the endometrium randomly but rather at receptive sites where stromal cells are poised to encapsulate the conceptus and create a micro-environment tailored to the individual embryo (Figure 1.2b) [3], a concept that fits well with the appearance of early human implantation sites on high-resolution ultrasound (Figure 1.3a) or histological analysis (Figure 1.3b).

***Figure 1.3: Embryo invasion or decidual encapsulation? (a) High-resolution ultrasound image of an early implantation site. The abnormal amount of free fluid in the uterine cavity (dotted line) allowed clear visualization of the gestational sac bulging into the lumen cavity (b) Drawing of an early human implantation site [11].***




## Formation of a Bilaminar Embryonic Disc

During the process of implantation, the inner cell mass, or embryoblast, begins to differentiate into two epithelial layers: an external upper (dorsal) layer of columnar cells called the epiblast, and an internal lower (ventral) layer of cuboidal cells called the hypoblast (or primitive endoderm).

## Development of the Amniotic Cavity

The amniotic cavity appears on day eight as fluid begins to collect between the cells of the epiblast and the overlying trophoblast. A layer of epiblast cells migrates towards the embryonic pole forming a thin membrane that separates the newly developed amniotic cavity from the overlying cytotrophoblast cells. This membrane is the lining of the amnion. Although the amniotic cavity is initially much smaller than the blastocyst cavity, it expands steadily. By the eighth week, the amnion encloses the entire embryo.

## Development of the Yolk Sac and Chorionic Cavity

The hypoblast cells also begin to proliferate and migrate from day eight onwards extending into the blastocyst cavity forming the primary yolk sac (Figure 1.1c). At the same time, the extraembryonic mesoderm develops, filling the remainder of the blastocyst cavity with loosely arranged cells. A new space develops within the extraembryonic mesoderm and this is the beginning of the chorionic cavity (Figure 1.1d), which separates the embryo with its dorsal amnion and ventral yolk sac from the outer wall of the blastocyst, now called the chorion.

By day twelve, the primary yolk sac is displaced as a subsequent wave of migrating hypoblast cells form the secondary or definitive yolk sac. By day thirteen, the bilaminar embryonic disc is suspended in the chorionic cavity by a thick connecting stalk of extraembryonic mesoderm (Figure 1.1e). The yolk sac remains a major structure associated with the developing embryo through the fourth week and performs important early functions. After the fourth week, the yolk sac is rapidly overgrown by the developing embryonic disc.

## Gastrulation and Formation of a Trilaminar Embryonic Disc

On day fifteen, a thickening, containing a midline groove develops along the mid-sagittal plane of the now oval bilaminar embryonic disc. Over the next 24 hours, this primitive streak elongates to occupy half the length of the embryonic disc, and the primitive groove becomes deeper and more defined. The cranial end of the primitive streak expands to form the primitive node, which contains a depression called the primitive pit. This continues caudally with the primitive groove. Formation of the primitive streak therefore defines the cranial-caudal, medial-lateral, left-right and dorsal-ventral axes of the developing embryo.

From day sixteen onwards, epiblast cells lateral to the primitive streak begin to move into the primitive streak and migrate into the space between the epiblast and hypoblast. This process is known as gastrulation. Initially the migrating epiblast cells invade and displace the hypoblast forming a new layer of cells known as definitive endoderm which gives rise to the lining of the future gut and gut derivatives. Subsequently, the epiblast cells migrate in the

space between the epiblast and definitive endoderm to form a third germ layer known as the intraembryonic mesoderm. These cells migrate bilaterally from the primitive streak and organize themselves into four main subdivisions: cardiogenic mesoderm, paraxial mesoderm, intermediate mesoderm and lateral plate mesoderm. Additionally, a fifth population of mesodermal cells migrates cranially from the primitive node in the midline to form a thick walled midline tube called the notochordal process.

During the third week of development, two faint depressions form in the ectoderm and the ectoderm in these areas fuse directly with the endoderm below (excluding any interspersed mesoderm) and two bilaminar membranes, which later become the blind ends of the gut tube, are formed. One of these is the oropharyngeal membrane located at the cranial end of the embryo overlying the prechordal plate and the other is the cloacal membrane, located caudally, behind the primitive streak. The oropharyngeal membrane disintegrates in week four to become the opening into the oral cavity and the cloacal membrane breaks down in week seven to form the openings of the anus and urogenital tract.

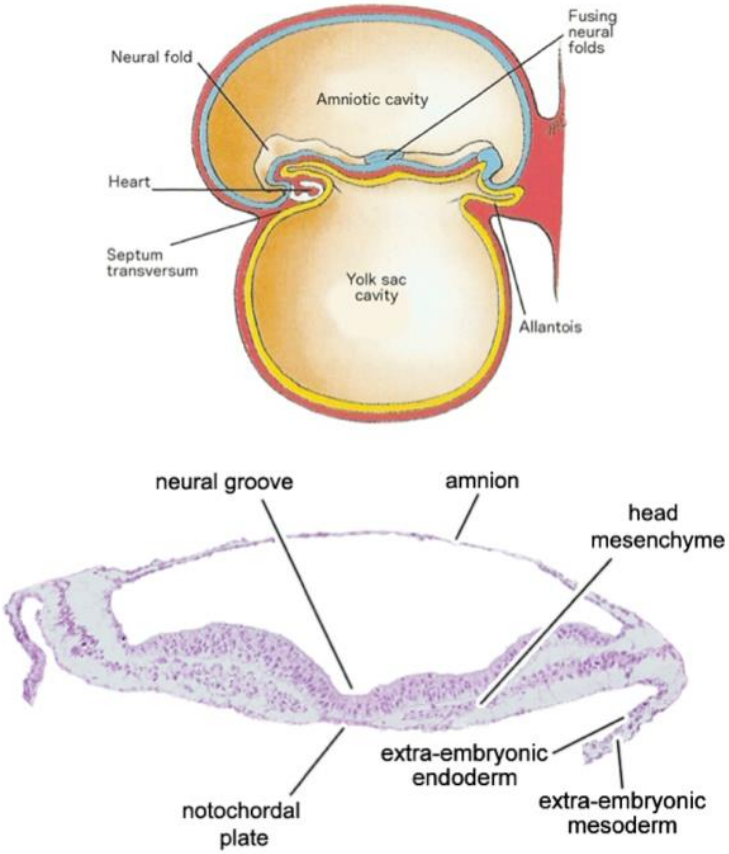
Once formation of the definitive endoderm and intraembryonic mesoderm is complete, epiblast cells no longer migrate towards the primitive streak and the remaining epiblast cells now constitute ectoderm, which quickly differentiates into the central neural plate and peripheral surface ectoderm. The process of gastrulation is then complete and the previously bilaminar embryonic disc is now trilaminar, with all three germ layers being derived from the epiblast.

## Forming the Embryo

At the end of the third week, the embryo is a flat, ovoid, trilaminar disc. During the fourth week it grows rapidly, particularly in length, and undergoes a process of folding (Figure 1.4). Although some active remodeling of tissue layers takes place, the main force responsible for embryonic folding is differential growth of various embryonic structures. During the fourth week, the embryonic disc and amnion grow vigorously, whereas the yolk sac hardly grows at all. As the outer rim of the embryonic endoderm is attached to the yolk sac, the expanding disc bulges into a convex shape. Folding commences in the cranial and lateral regions of the embryo on day 22 and in

the caudal region on day 23. As a result of folding, the cranial, lateral and caudal edges of the embryonic disc are brought together along the ventral midline. The endodermal, mesodermal and ectodermal layers of the embryonic disc each fuse to the corresponding layer on the opposite side, thus creating a tubular, three dimensional body form.

***Figure 1.4: Development of the three-dimensional 'tube-within-a-tube' body plan. Top image taken from Larsen's Human Embryology [1] and image below courtesy of The Virtual Human Embryo project (www.prenatalorigins.org/virtual-human-embryo/)***




## Forming the Heart

The heart is the first functioning organ in humans. It begins beating rhythmically as early as day 22 and pumps blood by day 25. Much of the development of the heart therefore occurs whilst the heart is pumping blood.

This is essential to provide nutrients and oxygen and dispose of waste during embryonic and fetal development. Morphologically, the embryonic heart is first identifiable as a single heart tube composed of contractile myocardium surrounding an inner endocardial tube. Between weeks four and eight, this primitive heart tube undergoes a process of looping, remodeling, realignment and septation that transforms its single lumen into the four chambers of the definitive heart.

## 1.3 Ultrasonography

Ultrasound is one of the most useful diagnostic tools in the field of obstetrics and gynaecology. Its main advantage over other imaging modalities being that it does not involve the use of the ionizing radiation. In addition to this, there are other more subjective benefits such as that it is relatively non-invasive, inexpensive, portable, quick, accurate and widely acceptable. Over the last few decades, ultrasound has been used extensively in the evaluation of early pregnancy enabling the development and growth of the fetus to be monitored, an ability which has been significantly enhanced over the years by dramatic improvements in ultrasound technology.

Transabdominal ultrasonography utilizes lower frequencies with poorer axial resolution than its transvaginal counterpart and nowadays it is therefore used predominantly in the second and third trimesters of pregnancy. Its use in the first trimester is relatively limited and mostly diagnostic in nature but this has not always been the case. Much of the information obtained from the original early pregnancy studies was collected using transabdominal ultrasound. The relatively recent introduction of higher frequency transvaginal ultrasound probes that can be placed closer to the pelvic organs providing images with better resolution has revolutionized the ultrasonographic study of very early pregnancy. For example, a gestation sac can be visualized from day 28 [12] using transvaginal ultrasound but often not until day 42 [13] using transabdominal ultrasound. A recent study comparing transabdominal and transvaginal ultrasonography in the same group of 50 women has concluded that transvaginal ultrasound reliably identifies normal and abnormal pregnancies at an earlier stage than transabdominal ultrasound [14].

Furthermore, the ability to visualize structures in early pregnancy is critically dependent not only the imaging route (transabdominal versus transvaginal) but also on the transducer frequency. In a study of 39 pregnancies, women were initially imaged with a 5-MHz transvaginal transducer and immediately afterwards were re-imaged with a 9MHz transvaginal transducer. Threshold values and discriminatory sizes used to distinguish normal and abnormal pregnancies were found to be smaller on higher frequency than on lower frequency imaging [15, 16]. It is therefore of paramount importance that thresholds be determined for both the specific imaging route and transducer frequency.

With that caveat in mind, we shall now go on to discuss the ultrasonographic development of a normal very early pregnancy.

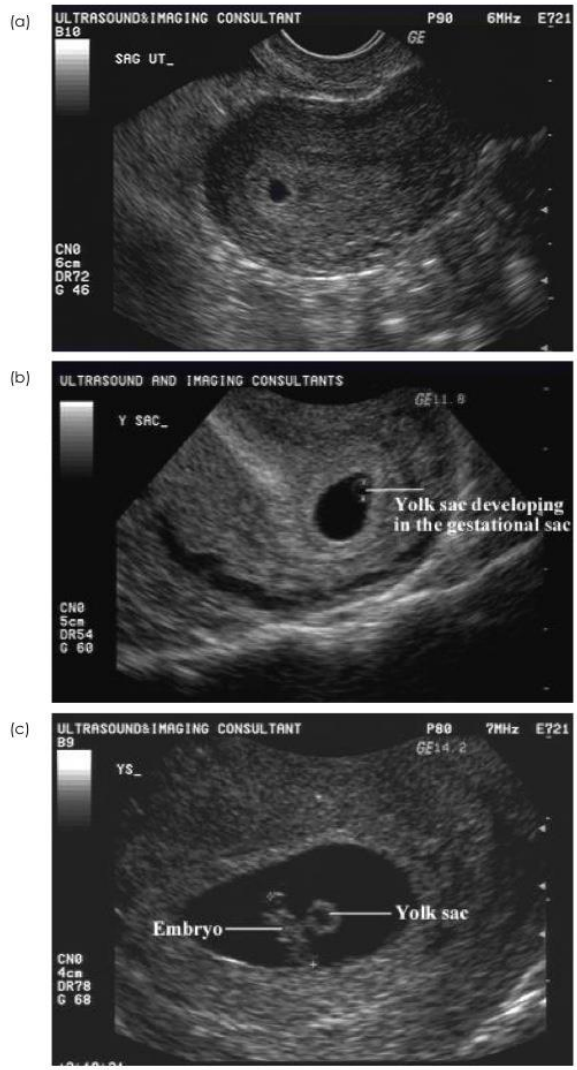
## Endometrial Thickness

The first visible ultrasonographic finding suggestive of a potential pregnancy is a thickened hyperechogenic homogenous endometrium (known as the decidual reaction).

## Gestation Sac

The first ultrasonographic sign of an intrauterine pregnancy is the appearance of a gestation sac (Figure 1.5a). It appears as a uniformly round, hypoechoic structure with an echogenic rim and is situated asymmetrically within the decidua, at or near the uterine fundus. A gestation sac is first visible with transvaginal ultrasound from day 28-31 onwards [12] when it measures approximately 2-3mm in diameter. Thereafter it grows at a rate of approximately 1mm/day (rendering it a good predictor of gestational age) and eventually it acquires a more elliptical outline. The gestation sac should always be present by 37 days when it measures approximately 5mm [17]. Initially the gestation sac does not contain any internal echoes and at this stage can be mistaken for a 'pseudosac', that is, an endometrial fluid collection that occurs in up to 15% of ectopic pregnancies [18]. Although this finding is relatively uncommon with an ectopic pregnancy, it is clinically important not to mistake the two structures.

***Figure 1.5: Ultrasonogram showing a (a) gestation sac (b) yolk sac and (c) fetal pole***




*Images taken from: www.fetalultrasound.com*

## Yolk Sac

The identification of the (secondary) yolk sac therefore is especially important because it is, at present, the first incontrovertible sign of an intrauterine pregnancy. The yolk sac is the first structure to become visible within the gestation sac and, if detected, is a reliable indicator of a true gestation sac and therefore an intrauterine pregnancy with a positive predictive value of 100% [19]. Identification of a yolk sac therefore excludes the possibility of an

ectopic pregnancy (unless in the very rare circumstance of a heterotopic pregnancy). It is usually visible with transvaginal ultrasound from about day 35 (four to seven days after the appearance of the gestation sac). It may be seen in gestation sacs as small as 5-6mm and should always be seen transvaginally when the gestation sac measures more than 10mm. The yolk sac appears as a spherical, hyperechoic ring and is situated eccentrically within the gestation sac (Figure 1.5b). It normally grows slowly during the first trimester, ranging from 2mm at five weeks gestation, to 6mm at ten weeks gestation. Thereafter it starts to regress and has usually disappeared completely by approximately twelve weeks gestation [20].

## Fetal Pole

The next structure to appear within the gestation sac is the fetal pole. The earliest reported sighting of this structure is day 35 but it may not be seen within the gestation sac until six weeks gestation even in normal pregnancies. It initially appears as a small linear echogenic structure adjacent to the yolk sac, on the side closest to the gestational sac (Figure 1.5c), giving the appearance of a 'signet ring'. The fetus is approximately 1-2mm when first detected ultrasonographically and increases in size by approximately 1mm/day. As it grows, it develops into a 'kidney bean' shape and, in doing so, gradually becomes positioned further away from the yolk sac.

## Fetal Heart

A fetal heart pulsation is typically appreciated as soon as, if not before, a fetal pole is detected. It has been documented as early as 37 days gestation in normal pregnancies [21] which is when the fetal heart tube starts to beat [22]. Cardiac activity should be evident when the fetal pole measures 2mm or more [23] but it is however not abnormal for it to be absent until the fetal pole is larger: 5-10% of viable embryos measuring between 2 and 4mm do not have visible cardiac activity [24, 25]. Once detected, the initial fetal heart rates are relatively slow, increasing from 100-110bpm at six weeks gestation to 150-170bpm at eight weeks gestation.

# 1.4 Serology

The two principal hormones in early pregnancy development are human chorionic gonadotropin (hCG) and progesterone. These may be used alone or in combination to give information regarding pregnancy gestation (hCG), viability (serial hCG and progesterone) and to a lesser extent, location (serial hCG).

## Human Chorionic Gonadotropin

Human chorionic gonadotropin is secreted by the syncytiotrophoblast in response to gonadotropin-releasing hormone (GnRH) production by the adjacent cytotrophoblast cells. hCG is a glycoprotein composed of 237 amino acids and like many other pituitary hormones, it consists of two subunits: an α-subunit which is identical to that of luteinizing hormone, follicle-stimulating hormone and thyroid stimulating hormone and a β-subunit that is unique to hCG.

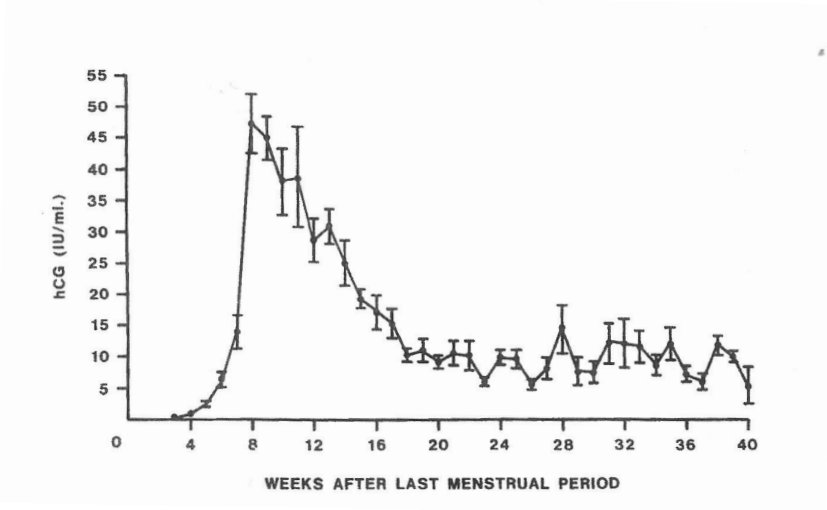
hCG acts to maintain the function of the corpus luteum that would otherwise degenerate in the absence of a pregnancy. During a menstrual cycle without conception, progesterone concentrations in the serum increase for the first 6-7 days of the luteal phase, followed by a 3-4 day plateau and then a decrease resulting in shedding of the endometrial lining. After conception and implantation, the corpus luteum continues to secrete progesterone and 17-hydroxyprogesterone for another 4-6 weeks. The maternal serum concentrations of progesterone and 17-hydroxyprogesterone then decrease, indicating a marked diminution in corpus luteum function. The fall in 17-hydroxyprogesterone continues but the drop in progesterone levels is only transient. This marks the transition from dependence on ovarian progesterone production to placental progesterone secretion and the corpus luteum subsequently regresses. This occurs towards the end of the first trimester of pregnancy.

In male fetuses, hCG also stimulates the early secretion of testosterone by the Leydig cells, an action that is critical to masculine genital tract differentiation. The very high hCG levels have enough structural overlap with thyroid stimulating hormone to stimulate increased maternal thyroid activity in early

pregnancy and may also contribute to the development of hyperemesis gravidarum.

If fertilization has occurred, hCG is detectable in the maternal circulation approximately ten days after ovulation. As the pregnancy develops, so too does the syncytiotrophoblast and hence the concentration of hCG in the maternal blood increases with advancing gestation. Serum hCG levels increase at an exponential rate, reaching a peak at between ten and twelve weeks gestation. The concentration then declines to a stable plateau for the remainder of the pregnancy (Figure 1.6)[26].

***Figure 1.6: Mean (± standard error) maternal serum hCG levels throughout normal pregnancy [26]***




Urinary and serum pregnancy tests work by detecting the presence of hCG in either the urine or blood respectively. Urinary pregnancy tests detect hCG levels as low as 20iu/l depending on the brand of the test. Older tests tend to only detect the presence or absence of the hormone in the urine but newer tests are increasingly analyzing the level of hormone present in the urine and translating this into an estimated gestation, which appears as a number in a digital display screen.

Serum hCG analysis can detect levels as low as 5iu/l and have the advantage of being a quantitative measure. A single serum hCG analysis therefore has the ability to confirm the diagnosis of pregnancy and may also give some indication as to gestation with higher hCG levels generally implying a more

advanced gestation. However, this is not a precise indicator of gestation as the range of normal hCG levels seen at any gestation is wide with considerable overlap [27, 28] (see Table 1.1). Furthermore, higher hCG levels are seen with multiple and molar pregnancies.

***Table 1.1: Correlation between gestation mean sac diameter (MSD), serum hCG [27, 28] and menstrual age [28]***

| Gestation Sac MSD (mm) | Menstrual Age (weeks$^{+days}$) | Predicted hCG (95% CI) (iu/l) |
|---|---|---|
| 3 | $4^{+4}$ | 1710 (1050-2800) |
| 4 | $4^{+5}$ | 2320 (1440-3760) |
| 5 | $4^{+6}$ | 3100 (1940-4980) |
| 6 | $5^{+1}$ | 4090 (2580-6530) |
| 7 | $5^{+2}$ | 5340 (3400-8450) |
| 8 | $5^{+3}$ | 6880 (4420-10810) |
| 9 | $5^{+4}$ | 8770 (5680-13660) |
| 10 | $5^{+5}$ | 11040 (7220-17050) |
| 11 | $5^{+6}$ | 13730 (9050-210408) |
| 12 | $6^{+0}$ | 16870 (11230-25640) |
| 13 | $6^{+1}$ | 20480 (13750-30880) |
| 14 | $6^{+2}$ | 24560 (16650-36750) |
| 15 | $6^{+3}$ | 29110 (19910-43220) |
| 16 | $6^{+4}$ | 34100 (23530-50210) |
| 17 | $6^{+5}$ | 39460 (27470-57640) |
| 18 | $6^{+6}$ | 45120 (31700-65380) |
| 19 | $7^{+0}$ | 50970 (40700-81150) |
| 20 | $7^{+1}$ | 56900 (40700-81150) |
| 21 | $7^{+2}$ | 62760 (45300-88790) |
| 22 | $7^{+3}$ | 68390 (49810-95990) |
| 23 | $7^{+4}$ | 73640 (54120-102540) |
| 24 | $7^{+5}$ | 78350 (58100-108230) |
| 25 | $7^{+6}$ | 82370 (61640-112870) |
| 26 | $8^{+0}$ | 85560 (64600-116310) |
| 27 | $8^{+1}$ | 87820 (66900-118420) |
| 28 | $8^{+2}$ | 89050 (68460-119130) |
| 29 | $8^{+3}$ | 89230 (69220-118420) |
| 30 | $8^{+4}$ | 88340 (69150-116310) |

Serum hCG analysis can detect levels as low as 5iu/l and have the advantage of being a quantitative measure. A single serum hCG analysis therefore has the ability to confirm the diagnosis of pregnancy and may also give some indication as to gestation with higher hCG levels generally implying a more advanced gestation. However, this is not a precise indicator of gestation as the range of normal hCG levels seen at any gestation is wide with considerable overlap [27, 28] (see Table 1.1). Furthermore, higher hCG levels are seen with multiple and molar pregnancies.

The main clinical benefit of a single serum hCG is that it can be used to rationalise the use of ultrasound in women presenting with symptoms of abdominal pain and/or vaginal bleeding in early pregnancy and help interpret sonograms when there is uncertainty. If the serum hCG is below a certain level or 'discriminatory zone' , it is unlikely that ultrasound will detect a gestation sac. If however the serum hCG is above this discriminatory zone, a gestation sac should be visible on ultrasound with a sensitivity approaching 100% [29]. With the use of high resolution transvaginal ultrasound, the discriminatory level has been reported to be approximately 1000iu/l [30] although the American Fertility Society recommends a more conservative level of 2400iu/l. In reality, the discriminatory zone may vary in different units depending on the specific hCG assay utilized, the quality of the ultrasound equipment available and the experience of the individual sonographer. In women with a hCG result above the discriminatory level but with no gestation sac visible on ultrasound, there is a high possibility of an ectopic pregnancy (PPV 18.2%) [31].

As described therefore, a single serum hCG can be used to confirm pregnancy, give an approximate estimate of gestation and help rationalise and interpret ultrasonography but unfortunately, a single serum hCG on its own, provides little or no information regarding pregnancy location or viability. Ectopic pregnancies secrete hCG as do failing pregnancies and serum hCG levels can remain elevated for several weeks following a complete miscarriage [32]. Serial serum hCG levels may therefore be of more benefit than solitary measures in providing information on pregnancy viability. A doubling of hCG is often expected every 48 hours, although this tends to vary with gestation: as pregnancy progresses, the doubling time lengthens [33]. In 1981, the concept of a minimal rise in serum hCG of 66% in 48 hours to predict a viable intrauterine pregnancy was first described in 20 women using an 85%

confidence interval [34]. It predicts an intrauterine pregnancy with a positive predictive value of 96.5%. Intervention for an hCG rise of less than 66% over two days, a practice supported by previous data, might however potentially result in the interruption of a viable pregnancy [35, 36], which is unacceptable. Recent studies have therefore attempted to redefine the hCG values in an effort to reduce the unintentional termination of a viable pregnancy. In a cohort of 287 women who presented with vaginal bleeding and/or pain, the minimum rise for a potentially viable pregnancy was 53% at two days [37]. However, in clinical practice, based on observations from 1249 women, a more conservative minimal rise in serum hCG of 35% over 48 hours has been suggested to minimise the potential risk of interrupting an on-going viable pregnancy [38]. It is important to remember however that up to 15% of normal pregnancies will have an abnormal doubling time [34].

## Progesterone

Progesterone is initially secreted solely by the corpus luteum. From approximately seven weeks' gestation, the placenta begins to synthesize progesterone and by twelve weeks it produces enough to replace the corpus luteum source. Cholesterol extracted from maternal plasma serves as the major precursor for placental progesterone. The synthetic pathway is identical to that of the adrenal gland and the ovary. By the end of pregnancy, placental progesterone production reaches a level that is ten times greater than peak production by the corpus luteum.

Progesterone has many functions during early pregnancy. It not only stimulates endometrial glands to secrete nutrients on which the developing embryo depends but also maintains the decidual lining of the uterus, where it induces prolactin synthesis. Prolactin helps inhibit the maternal immune responses to fetal antigens of paternal origin, thus helping to prevent rejection of the fetus. Progesterone also promotes uterine quiescence by inhibiting prostaglandin production and desensitizing the uterus to oxytocin thereby encouraging prolongation of the pregnancy.

Serum progesterone levels are therefore elevated during pregnancy. Levels change little during the first 8-10 weeks of gestation unless the pregnancy fails. Several studies [39-41], including a meta-analysis [42], have shown levels

>25nmol/l are 'likely to indicate', and levels >60nmol/l are 'strongly associated with', pregnancies subsequently shown to be viable. However, a small proportion (0.3%) of viable pregnancies have been reported with initial levels <15.9nmol/l [33]. Unfortunately however, whilst progesterone levels seem to be good at predicting pregnancy viability, they are poor at predicting pregnancy location [43].

## 1.5 Summary

This chapter has focused on normal early pregnancy development including the development of an embryo from fertilization to implantation, including the development of the amniotic cavity, yolk sac, chorionic cavity, formation of the trilaminar embryonic disc, and fetal heart. It has also described the ultrasonographic findings of a normal early pregnancy including the sequence of events in which structures such as the gestation sac, yolk sac, fetal pole and fetal heart become visible. Finally, the two principal hormones involved in normal early pregnancy development and their role in clinical practice, including their limitations, has been discussed.

# 2. Abnormal Early Pregnancy Development

## 2.1 Introduction

Having considered normal early pregnancy development, we shall now move on to discuss abnormal early pregnancy development, for example miscarriage and ectopic pregnancy, in more detail.

Complications arise more frequently during the first trimester of pregnancy than at any other stage of pregnancy. Most present with abdominal pain and/or vaginal bleeding, both of which cause considerable anxiety for the couple. Unfortunately, in the vast majority of cases, no intervention alters the outcome. The main aim of clinical management is a prompt and accurate diagnosis (Table 2.1) with reassurance if the pregnancy is progressing appropriately, or suitable intervention if it is not.

***Table 2.1: Differential diagnosis of abdominal pain and/or vaginal bleeding in the first trimester of pregnancy [44]***

| **Related to Pregnancy** | **Unrelated to Pregnancy** | |
|---|---|---|
| | **Gynaecological** | **Non-gynaecological** |
| Miscarriage | Ovarian cyst accident | Urinary tract infection |
| Ectopic pregnancy | Torsion/degeneration of a pedunculated fibroid | Renal colic |
| Hydatidiform mole | Bleeding from cervical malignancy | Bowel obstruction |
| Cervical ectropion | Dysfunctional uterine bleeding | Cholecystitis |
| | Pelvic inflammatory disease | Appendicitis |
| | Endometriosis | |

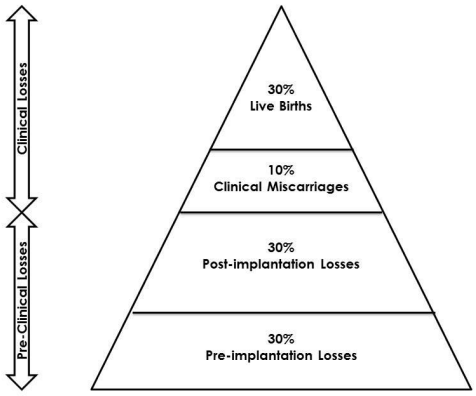
This chapter first introduces the clinical problems of miscarriage and ectopic pregnancy and then goes on to describe the ultrasonographic appearance of abnormal early pregnancies including miscarriages and various different

types of ectopic pregnancies. The final section discusses the role of hCG and progesterone in diagnosing abnormal early pregnancy development.

## Miscarriage

Spontaneous miscarriage refers to the natural loss of a pregnancy before independent viability of the fetus. Viability implies the ability of the fetus to survive extra-uterine life but the gestational age at which the fetus is capable of independent existence is controversial. The definition of miscarriage according to the World Health Organization is 'the expulsion from its mother of an embryo or fetus weighing 500g or less' (500g is approximately the 50$^{th}$ centile for 20 weeks' gestation). In the UK, any pregnancy loss before 24 weeks is regarded as a miscarriage.

***Figure 2.1: The pregnancy loss iceberg [45]***




Miscarriage occurs in approximately 25% of recognized pregnancies. It is known however that a far greater unrecognized pregnancy loss is present in the background, with more pregnancies being lost before the pregnancy has been suspected, recognized or confirmed, giving rise to the term 'pregnancy loss iceberg' (Figure 2.1). Preliminary investigations using hCG to study early pregnancy demonstrated great variations in the rate of unrecognized pregnancy loss ranging from 3-34% [45]. This variation reflects the limitations of the hCG assays used, the different patient populations involved and methodological problems relating to the timing of ovulation. The detection of 'background hCG' in non-pregnant women may have further confused the issue. Many of these problems were addressed by a landmark study in which

daily urine samples from 221 women, collected during six months of attempted conception, were analyzed for hCG [46]. The degree of sensitivity of the hCG assay employed meant that background hCG produced by the endometrium of non-pregnant women could be detected. Therefore a control group of women who had undergone sterilization by tubal ligation were also studied, and cut off values for identifying pregnancy were determined. Of the 198 pregnancies that were detected, 31% were subsequently lost. 22% of all pregnancy losses were occult, occurring before the woman could have been aware of the pregnancy. In a similar study in which 200 women collected daily urine specimens for hCG analysis over three menstrual cycles, an overall pregnancy loss rate of 31% was observed but only 13% of these were occult [47]. The discrepancy may be due to the lower sensitivity of the assay used or the lack of a control group. However, when taken together, data from the published studies point to a rate of pregnancy loss prior to implantation of 30%, a further 30% following implantation but prior to the missed period and 10% as clinical miscarriages.

***Table 2.2: Different types of miscarriage [48]***

| **Type of Miscarriage** | **Description** |
|---|---|
| Threatened | Vaginal bleeding and an ongoing pregnancy |
| Inevitable | The cervix begins to dilate |
| Incomplete | Passage of some, but not all, products of conception |
| Complete | All products of conception have been passed from the uterus |
| Missed/silent | Where the fetus has died in utero but has not been expelled |
| Anembryonic | A type of missed miscarriage in which embryonic development fails at a very early stage in the pregnancy; the sac continues to develop but there are no fetal parts evident on ultrasound scan |
| Septic | A complication of an incomplete miscarriage when intrauterine infection occurs |
| Recurrent | The somewhat arbitrary definition of three or more consecutive miscarriages |

*NB; Whilst accepting that a significant proportion of pregnancy losses occur prior to implantation, for the purposes of this research, a miscarriage was defined as that which occurs after a positive urinary pregnancy test has been*

*obtained. The delineation was made at this point because whilst perhaps not clinically recognizable (based on a history of a missed menstrual period or ultrasonographically identifiable), the pregnancy had been recognized by the woman.*

***Table 2.3: Terminology for classifying pregnancy failure prior to viability [49]***

| **Term** | **Description of pregnancy loss and clinical or ultrasound findings** |
|---|---|
| Pregnancy loss | Spontaneous pregnancy demise |
| Early pregnancy loss | Spontaneous pregnancy demise before 10 weeks of gestational age (before 8$^{th}$ developmental week) |
| Non-visualized pregnancy loss | Spontaneous pregnancy demise based on decreasing serum or urinary hCG levels and non-localization on ultrasound , if performed |
| Biochemical pregnancy loss | Spontaneous pregnancy demise based on decreasing serum or urinary hCG levels without an ultrasound evaluation |
| Resolved pregnancy loss of unknown location | Pregnancy demise not visualized on transvaginal ultrasound with resolution of serum hCG after expectant management or after uterine evacuation without chorionic villi on histology |
| Treated pregnancy loss of unknown location | Pregnancy demise not visualized on transvaginal ultrasound with resolution of serum hCG after medical management |
| Miscarriage | Intrauterine pregnancy demise confirmed by histology or ultrasound |
| Early miscarriage | Intrauterine pregnancy loss <10 weeks' size on ultrasound |
| Anembyonic (empty sac) miscarriage | Intrauterine pregnancy loss with a gestational sac but without a yolk sac or an embryo on ultrasound |
| Yolk sac miscarriage | Intrauterine pregnancy loss with a gestational sac and yolk sac, without an embryo on ultrasound |
| Embryonic miscarriage | Intrauterine pregnancy loss with an embryo without cardiac activity on ultrasound |
| Fetal miscarriage | Pregnancy loss ≥ 10 weeks' size with a fetus (≥33mm) on ultrasound |

Not all pregnancy losses are symptomatic occurring following presentation with abdominal pain and/or vaginal bleeding. Pregnancy losses may also be asymptomatic, identified only at the time of a routine scan. Equally, not all women that present with abdominal pain and/or vaginal bleeding go on to miscarry. In one study, only 12% of pregnancies in which bleeding occurred went on to miscarry. Miscarriages have traditionally been classified in a clinical way (Table 2.2). More recently, recommendations for pregnancy terminology and definitions for adverse pregnancy outcomes before viability have been proposed primarily to provide clear, consistent and widely applicable terminology for early pregnancy research (Table 2.3) [49].

The maternal risks of any type of miscarriage include haemorrhage (which may be significant), infection and the psychological effects of the loss of the pregnancy (which may be severe and prolonged) [50]. In the latest confidential enquiry into maternal deaths in the UK [51], nine women died as a consequence of complications of spontaneous miscarriage, an increase compared with each of the five previous triennia. Four of these deaths occurred secondary to infection and the remaining five women died as a result of haemorrhage.

Management may be conservative, medical or surgical, the choice of which depends on the clinical situation and patient preference. Medical management involves the use of misoprostol with or without prior treatment with mifepristone. Mifepristone is an anti-progestogenic steroid, which sensitizes the myometrium to prostaglandin-induced contractions and ripens the cervix. Misoprostol is a synthetic prostaglandin analogue that can be administered orally or vaginally to induce medical miscarriage or to ripen the cervix prior to surgical evacuation. A dose of 600-800µg is recommended. Surgical management may involve manual vacuum aspiration under local anaesthetic in an outpatient or clinic setting or evacuation of retained products of conception in theatre under general anaesthetic. Several studies exist that demonstrate that in selected women, both conservative and medical management compare favourably with surgical evacuation of the uterus with no increase in the risk of infection or severe haemorrhage [52].

## Ectopic Pregnancy

An ectopic pregnancy is one that occurs in a site outside of the uterine cavity, but usually in an adjacent site. The most common site for an ectopic pregnancy is within the fallopian tube but a pregnancy can implant anywhere (Figure 2.2), for example, on the ovary or within the abdominal cavity [53]. Ectopic pregnancies within the appendix [54], liver [55], spleen [56] and omentum [57] have all been reported in the literature. Pregnancies that implant in the uterine cornua or cervix, or within a previous caesarean section scar, are also referred to as ectopic pregnancies.

***Figure 2.2: Common sites of ectopic pregnancies***

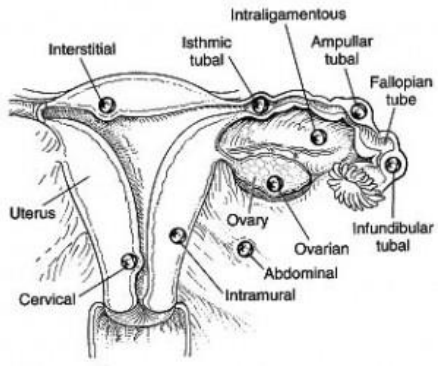



*Image taken from: www.ectopic.org.uk*

Approximately 1 in 100 spontaneous pregnancies are ectopic in nature and the incidence of is rising. This is largely due to the increased prevalence of sexually transmitted diseases and availability of assisted reproduction techniques, both of which are major risk factors for the development of the condition. Although the recognition of high risk individuals has improved, and prompt diagnosis using high resolution transvaginal ultrasound and serum hCG levels is possible, the number of deaths from ectopic pregnancy has remained relatively constant over the years although the latest confidential enquiry into maternal deaths in the UK [51] shows the case fatality rate of ectopic pregnancies to be the lowest since these figures were first estimated in 1988 (Figure 2.3).

Maternal risks include haemorrhage and its consequences, implications for future reproductive performance and the psychological effects of the loss of the pregnancy. The choice of management depends largely on the clinical situation and to a lesser extent patient preference. It may be conservative, medical using methotrexate, or surgical involving either a salpingectomy or salpingotomy performed via laparoscopy or rarely laparotomy. Whilst laparoscopy is associated with shorter operating times, smaller intraoperative blood losses, shorter hospital stays, lower costs, decreased analgesia requirements and less adhesion formation than laparotomy [58-60], evidence suggests that there is no difference in the rates of subsequent successful pregnancy between the two approaches. Similarly, there does not appear to be a significant difference in subsequent ongoing pregnancies (56.2% versus 60.7%) or repeat ectopic pregnancies (5% versus 8%) in tubal ectopic pregnancies managed via salpingectomy or salpingotomy in women without a history of fertility-reducing factors [61].

***Figure 2.3: Numbers of deaths from ectopic pregnancies and rates per 100, 000 estimated ectopic pregnancies [51]***

| Triennium | Total estimated pregnancies | Total estimated ectopic pregnancies* | Ectopic pregnancies per 1000 pregnancies | | Deaths from ectopic pregnancies | Death rate per 100 000 estimated ectopic pregnancies | |
|---|---|---|---|---|---|---|---|
| | *n* | *n* | Rate | 95% CI | *n* | Rate | 95% CI |
| **England and Wales** | | | | | | | |
| 1988–90 | 2 880 814 | 24 775 | 8.6 | 8.5–8.7 | 15 | 60.5 | 36.5–100.4 |
| **United Kingdom** | | | | | | | |
| 1991–93 | 3 141 667 | 30 160 | 9.6 | 9.5–9.7 | 9 | 29.8 | 15.5–57.4 |
| 1994–96 | 2,917 391 | 33 550 | 11.5 | 11.4–11.6 | 12 | 35.8 | 20.3–63.0 |
| 1997–99 | 2,878 018 | 31 946 | 11.1 | 11.0–11.2 | 13 | 40.7 | 23.6–70.1 |
| 2000–02 | 2,736 364 | 30 100 | 11.0 | 10.9–11.1 | 11 | 36.5 | 20.2–66.0 |
| 2003–05 | 2,891 892 | 32 100 | 11.1 | 10.9–11.1 | 10 | 31.2 | 16.8–57.9 |
| 2006–08 | 3,139 315 | 35 495 | 11.3 | 11.2–11.4 | 6 | 16.9 | 7.6–37.6 |

## 2.2 Ultrasonography

Not only is ultrasound a sensitive method for dating early pregnancies, but it can also be used to localize and assess viability in women who present with abdominal pain and vaginal bleeding in early pregnancy.

An understanding of the ultrasonographic development of a normal viable early intrauterine pregnancy is a pre-requisite for being able to diagnose abnormal early pregnancy development, for example non-viable (miscarriage) and extra-uterine (ectopic) pregnancies.

## The Use of Ultrasound to Diagnose Miscarriage

Whilst defining fetal viability is relatively easy (the presence of a fetal pole with a fetal heart beat), the definition of non-viability is often less straightforward.

The diagnosis of miscarriage by ultrasound was first described in the 1960s [62]. It can be diagnosed with confidence relatively easily when an ultrasound scan detects an intrauterine pregnancy without a visible fetal heartbeat when previously it had been visible and/or when there is failure of the embryo or fetus to increase in size over a period of at least one week.

Unfortunately however not all miscarriage diagnoses are as clear-cut as this. As described in chapter one, there are landmarks in normally developing pregnancies for when structures such as the gestation sac, yolk sac, fetal pole and fetal heart should be visible with ultrasound. However, it cannot be assumed that the absence of certain landmarks by a certain time implies a failed or failing pregnancy because, except in cases of assisted conception, the timing of ovulation and fertilization cannot be known with certainty. Furthermore, not all pregnancies exhibit uniform growth in the first trimester [63]. The key question is, when can the absence of a feature be used to diagnose miscarriage with absolute certainty?

To complicate matters further there is considerable geographic variation in the criteria used to confirm non-viability and therefore diagnose miscarriage. Until relatively recently, guidelines (based on expert committee reports or opinions and/or clinical experience of respected authorities) produced by the Royal College of Obstetricians and Gynaecologists and the Royal College of Radiologists stated that pregnancies with a fetal pole (or crown rump length) measuring 6mm or more with no visible fetal heart activity or a mean gestational sac diameter measuring 20mm or more without an identifiable yolk sac or fetal were non-viable [64]. However, the American College of Obstetricians and Gynaecologists and the Hong Kong College of Obstetricians and Gynaecologists use a crown rump length threshold of 5mm and a mean sac diameter of 16mm for the diagnosis of a miscarriage [65, 66] whilst the Society of Gynaecology of Canada suggest a mean sac diameter exceeding 8mm without a yolk sac is sufficient for the diagnosis of a miscarriage [67]. This disparity reflects the limited evidence available to define viability. Furthermore

the guidelines are based on a small number of poor quality studies performed in the 1980s and 1990s that used ultrasound technology unquestionably inferior to that available today. Many of these studies were underpowered with sample sizes ranging between 55 and 211 and did not take into consideration the reliability of the measurements made. A recent systematic review [68] showed that an empty gestation sac with mean sac diameter of 25mm or more or an absent yolk sac with a mean gestation sac diameter of 20mm or more were associated with the highest and most precise estimates of specificity for diagnosing early embryonic demise. Whilst these thresholds were associated with an estimated specificity of 1.00 their confidence intervals (95% CI 0.96-1.00) showed even these values could lead to a false positive diagnosis in 4% of cases.

The false negative rate of diagnosing a miscarriage must however be zero, otherwise termination of a potentially viable pregnancy may ensue. Abdallah *et al* [69] defined the false negative rate for the diagnosis of miscarriage using different thresholds for both crown rump length with absent cardiac activity and mean sac diameter with an empty gestation sac in a cross-sectional observational study of 1060 women with intrauterine pregnancies of uncertain viability. Using the American and Hong Kong guidelines, an empty gestation sac with a mean sac diameter of 16mm and a crown rump length threshold of 5mm were associated with false negative rates of 4.4% and 8.3% respectively. An empty gestation sac with a mean sac diameter of 20mm was associated with a false negative rate of 0.5%. The false negative rate was zero only at a mean sac diameter of greater than 21mm. A crown rump length threshold of 6mm was also associated with a false negative rate of zero. This would imply that a diagnosis of a miscarriage can only be made if there is a gestation sac with a mean sac diameter measuring at least 21mm with no identifiable yolk sac or fetal pole or there is a crown rump length of at least 6mm with no visible fetal heart activity.

However, one must take into consideration the inter- and intra-observer reliability of the mean gestation sac diameter and crown rump length measurements and these have been shown to be poor in pregnancies between six and nine weeks of gestation [70]. The limits of agreement between observers for mean sac diameter and crown rump length measurements are ±19% and ±14% respectively. This means that in order to take into account the variation that can occur when measuring early

pregnancies, the threshold for the safe and accurate diagnosis of miscarriage should be increased to a gestation sac with a mean sac diameter measuring 25mm or more or a crown rump length of at least 7mm with no visible fetal heart activity. This fact was incorporated into an addendum on the 2006 RCOG Guideline on the 'Management of Early Pregnancy Loss' in October 2011 [64].

## The Use of Ultrasound to Diagnose Ectopic Pregnancy

### Tubal Ectopic Pregnancy

The majority of tubal ectopic pregnancies should be visualized on transvaginal ultrasound, which has reported sensitivities and specificities of 87.0-99.0% and 94.0-99.9% respectively [71-75]. The presence of a yolk sac and/or fetal pole with or without cardiac activity in a sac outside the uterus is the only pathognomonic sign of a tubal ectopic pregnancy (Figure 2.4a) but unfortunately this is reported in as few as 8-26% of ectopic pregnancies detected on transvaginal ultrasound [72, 76-78]. Other non-specific ultrasonographic findings suggestive but not diagnostic of tubal ectopic pregnancy are discussed in chapter six.

### Cornual Ectopic Pregnancy

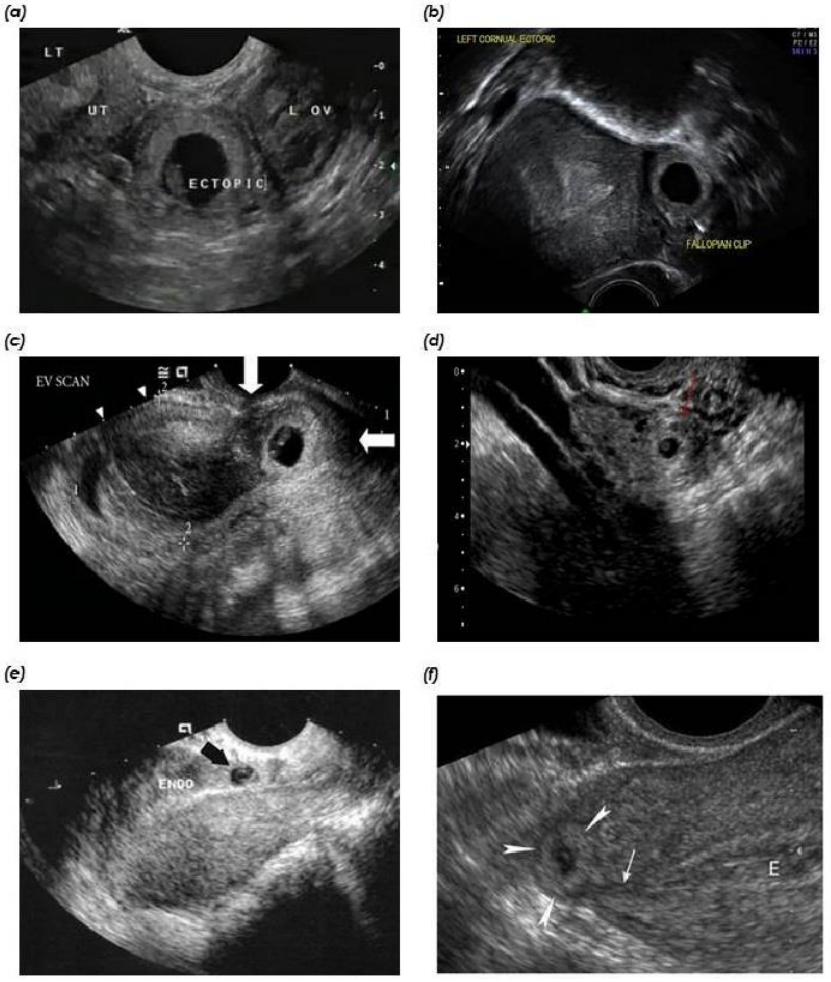
Cornual pregnancies are the rarest form of ectopic pregnancy with a reported incidence of 1 in 76000 pregnancies [79]. A cornual ectopic pregnancy can be diagnosed using ultrasound following visualization of (1) a single interstitial portion of fallopian tube in the main uterine body, (2) a mobile gestation sac/products of conception separate from the uterus and completely surrounded by myometrium and (3) a vascular pedicle adjoining the gestation sac to the unicornuate uterus [80] (Figure 2.4b). However it may not be directly recognizable because there is often no unique discernable abnormal feature and/or clotted blood in the cul-de-sac [81].

### Cervical Ectopic Pregnancy

Cervical pregnancies are rare, accounting for less than 1% of all ectopic pregnancies [82]. The ultrasonographic diagnosis of a cervical pregnancy (Figure 2.4c) is facilitated by the proximity of the probe to the area of interest

[81]. The following criteria is used to diagnose a cervical ectopic pregnancy (1) an empty uterus (2) a barrel shaped cervix (3) a gestational sac present below the level of the internal cervical os (4) the absence of the 'sliding sign' (when pressure is applied to the cervix using the probe, in a miscarriage the gestation sac slides against the endocervical canal but in an implanted cervical pregnancy it does not) and (5) blood flow around the gestation sac using colour Doppler [83, 84].

***Figure 2.4: Ultrasonogram showing a (a) tubal (b) cornual, (c) cervical, (d) ovarian, (e) Caesarean scar and (f) interstitial ectopic pregnancy***




*Images taken from: www.fetalultrasound.com*

## Ovarian Ectopic Pregnancy

There are no specific agreed criteria for the ultrasonographic diagnosis of an ovarian ectopic pregnancy, despite there being strict surgical criteria. Ultrasound findings have been described in individual case reports. A study on six cases of ovarian pregnancies reported that the pregnancies commonly appear as a cyst, on or within the ovary, with a wide echogenic outside ring (Figure 2.4d). A yolk sac or fetal pole is not commonly seen [85, 86]. It is not possible to separate the cystic structure or gestation sac from the ovary on gentle palpation (negative sliding organ sign). The corpus luteum should be identified separate from the suspected ovarian pregnancy. Colour Doppler may aid detection of a fetal heart pulsation within the ovary. Diagnosis of an ovarian ectopic pregnancy is usually confirmed surgically and histologically as it is difficult to differentiate them from corpus luteal cysts, adherent tubal ectopic pregnancies, ovarian germ cell tumours and other ovarian pathologies.

## Caesarean Scar Ectopic Pregnancy

Caesarean scar pregnancy is defined as implantation into the myometrial defect occurring at the site of previous uterine incision. The prevalence of Caesarean scar pregnancy is estimated to be approximately 1 in 2000 pregnancies [87] although the true prevalence is likely to be somewhat higher than estimated in the literature as some cases will end in the first trimester either by spontaneous miscarriage or termination, and go unreported and undiagnosed. Ultrasound is the primary modality for diagnosing a Caesarean scar ectopic pregnancy. A transvaginal approach may be supplemented by transabdominal imaging if required. The diagnosis of a Caesarean scar ectopic pregnancy (Figure 2.4e) can be made using the following criteria: (1) an empty uterus, (2) a gestation sac located anteriorly at the level of the internal os covering the site (visible or presumed) of the previous lower segment Caesarean section incision, (3) thin or absent layer of myometrium between the gestation sac and the bladder, (4) evidence of functional trophoblastic/placental circulation on Doppler examination and (5) an empty endocervical canal/absent 'sliding sign' [88-94]. Magnetic resonance imaging (MRI) can be used as a second-line investigation if the diagnosis is equivocal and there is local expertise available. It is important to note that

the diagnostic criteria described above have not been subject to validation and are derived from descriptive case series.

## Abdominal Ectopic Pregnancy

The following ultrasound criteria have been suggested as being diagnostic of an early abdominal pregnancy: (1) absence of an intrauterine gestation sac, (2) absence of both an evident dilated tube and a complex adnexal mass, (3) a gestational cavity surrounded by loops of bowel and separated from them by peritoneum and (4) a wide mobility similar to fluctuation of the sac, particularly evident with pressure of the transvaginal probe toward the posterior cul-de-sac [95]. MRI can be a useful diagnostic adjunct in advanced abdominal pregnancy [96].

## Interstitial Ectopic Pregnancy

Interstitial pregnancy occurs when the ectopic pregnancy implants in the interstitial part of the fallopian tube. The reported incidence varies between 1.0% and 6.3% of ectopic pregnancies [97-99]. The interstitial part of the fallopian tube is about 1-2cm in length and traverses the muscular myometrium of the uterine wall, opening via the tubal ostium into the uterine cavity [100].

The following ultrasound criteria may be used for the diagnosis of interstitial pregnancy: (1) an empty uterus, (2) products of conception/gestation sac located laterally in the interstitial (intramural) part of the tube and surrounded by less than 5mm of myometrium in all imaging planes, and (3) presence of the 'interstitial line sign' i.e. a thin echogenic line extending from the central uterine cavity echo to the periphery of the interstitial sac (see the arrow in Figure 2.4f). The interstitial line sign has been shown to have a sensitivity of 80% and specificity of 98% for the diagnosis of interstitial ectopic pregnancy [101]. Three-dimensional ultrasound, if available, can be used to confirm two-dimensional ultrasound findings as can MRI. In the three-dimensional coronal view of the uterus, a connection between the endometrial cavity and the interstitial part of the tube can be visualized [102, 103] and on MRI, a gestational sac-like structure is seen lateral to the cornua surrounded by the myometrium. The presence of the intact junctional zone (endomyometrial

junction) between the uterine cavity and the gestational-sac like structure also supports the diagnosis [104, 105].

## 2.3 Serology

The role of serum hCG and progesterone and their limitations in normal early pregnancy development has been discussed in the previous chapter. These hormones can also aid in the diagnosis of abnormal early pregnancy development for example in miscarriage and ectopic pregnancy. Unfortunately however, as will become apparent in this section, these two hormones are unable to diagnose abnormal early pregnancy development with absolute certainty and hence numerous other biomarkers have been investigated in an attempt to definitively discriminate viable intrauterine pregnancies from non-viable and/or extra-uterine pregnancies and the more promising candidates will be discussed in more detail in chapters three and nine.

### Human Chorionic Gonadotropin (hCG) and Progesterone

hCG is secreted by the syncytiotrophoblast in response to GnRH production by the adjacent cytotrophoblast cells. As a pregnancy develops it increases in size and hence the syncytiotrophoblast produces more hCG. An increase in hCG of at least 66% in forty eight hours is associated with an intrauterine pregnancy with a positive predictive value of 96.5% [34]. A decrease in hCG of at least 15% in a forty eight hour period is most usually associated with a failing pregnancy [33] although a similar decline is seen in approximately 8% of ectopic pregnancies [38, 106].

If the serum hCG increases by less than 66% or decreases by less than 15% in a forty eight hour period, the most likely outcome is an ectopic pregnancy [32, 38, 106]. Approximately 71% of ectopic pregnancies have a suboptimal increase or decrease in serum hCG over a forty eight hour period but this is not diagnostic of an ectopic pregnancy as 13-21% of ectopic pregnancies have a hCG pattern which mimics that of a viable intrauterine pregnancy [34, 106]. A suboptimal rise in serum hCG over a forty eight hour period predicts an ectopic pregnancy with a positive predictive value of 43.5% [31].

As mentioned in the previous chapter, the hCG can be used in combination with ultrasonographic findings to guide the management of women when uncertainties in early pregnancy exist. For example, if a woman has an empty uterus on ultrasound and a serum hCG below the discriminatory zone, the pregnancy is most likely too small to be visible on scan and a very early intrauterine pregnancy is still a possibility (although ectopic cannot be excluded). In this instance, as long as the woman is in a stable condition, management can be more conservative. If a woman has an empty uterus on ultrasound and the serum hCG is above the discriminatory zone, there is a much higher possibility that the pregnancy is ectopic and management may be more aggressive.

Serum progesterone levels are elevated during pregnancy and change little during the first 8-10 weeks of gestation unless the pregnancy fails [33]. Several studies [39-41], including a meta-analysis [42], have shown serum progesterone levels less than 25nmol/l to be associated with non-viability although a small proportion (0.3%) of viable pregnancies have been reported with initial levels <15.9nmol/l [107] . Levels less than 20nmol/l have a positive predictive value of more than 95% at predicting a failing pregnancy, and this compares favourably with complex multi-parameter diagnostic models [40]. Progesterone is therefore a useful marker of pregnancy viability.

Unfortunately however, progesterone is a less useful indicator of pregnancy location. Although a low progesterone concentration has been associated with ectopic pregnancy since the late 1970s, there is currently no well-established cut-off to discriminate between intra and extra-uterine pregnancies. Furthermore, a small proportion (2.6%) of ectopic pregnancies have been reported with a serum progesterone concentration of more than 60nmol/l [42]. A meta-analysis [42] of twenty six studies assessing the accuracy of a single serum progesterone measurement in the diagnosis of an ectopic pregnancy concluded that a single serum progesterone measurement can identify patients at risk of ectopic pregnancy who need further evaluation, but its discriminative capacity is insufficient to diagnose ectopic pregnancy with certainty.

In summary so far then, although a suboptimal increase or decrease in serum hCG over a 48-hour period may be indicative of abnormal early pregnancy

development, it is unfortunately not diagnostic of either a miscarriage or an ectopic pregnancy. Furthermore, whilst a low serum progesterone level may be suggestive of abnormal early pregnancy development with regards to viability, unfortunately its ability to predict a failing pregnancy is not 100% specific and therefore it cannot be solely relied upon in clinical practice.

## 2.4 Summary

This chapter has introduced the common clinical problems of miscarriage and ectopic pregnancy. It has described the ultrasonographic appearances of abnormal early pregnancies including miscarriage and various different types of ectopic pregnancies. The role of hCG and progesterone in diagnosing abnormal early pregnancy development has also been discussed. Having briefly described both normal and abnormal early pregnancy development, we can now move on to discuss diagnostic uncertainties in early pregnancy, which is the focus of the remainder of this thesis.

# 3. Diagnostic Uncertainties In Early Pregnancy

## 3.1 Introduction

Abdominal pain and/or vaginal bleeding in early pregnancy is common and nowadays women are able to find out they are pregnant before they even miss a menstrual period. This means that an increasing number of women are presenting to Emergency Departments and Early Pregnancy Assessment Units in the very early stages of pregnancy.

If a woman presents during very early pregnancy therefore and an ultrasound scan is performed, it is possible that the findings will be inconclusive and a definitive diagnosis cannot be provided at that time. Depending on the exact findings during the ultrasound examination, the woman will be given one of two possible interim 'diagnoses', both of which reflect an 'uncertainty' in early pregnancy, either in location or viability.

These uncertainties are hazardous in many ways. Firstly they are likely to cause a considerable amount of anxiety for the pregnant woman and her partner. Secondly they contribute a significant part of the workload for Early Pregnancy Assessment Units which has associated cost and time implications (serial blood tests, follow-up ultrasound scans, counseling and occasionally admission to hospital and surgery) and finally, perhaps most importantly, during the time it takes to make a definite diagnosis, a haemodynamically stable woman with an unknown miscarriage or ectopic pregnancy, may deteriorate and become unstable and need immediate resuscitation, blood transfusion and/or emergency life-saving surgical intervention as opposed to more conservative forms of management which could have been utilized if the diagnosis had been made whilst in a stable condition.

## 3.2 Pregnancies of Unknown Location

A pregnancy of unknown location is the descriptive term used to describe the situation when a woman has a positive pregnancy test but there are no signs

of an intra- or extra-uterine pregnancy or retained products of conception on ultrasound scan [108].

There are four potential reasons why a pregnant woman may have an empty uterus on transvaginal ultrasound scan. Firstly the woman may not be pregnant. It is therefore important that whenever an empty uterus is identified during an ultrasound that a pregnancy test is performed. Other conditions, for example germ cell tumours, can lead to elevated hCG levels and although these are extremely rare, they must be excluded. Secondly, the woman may be pregnant and the pregnancy may be intrauterine but it may be too early in the pregnancy to identify any pregnancy related structures. A gestation sac is not visible on ultrasound until at least day 28-31 [12] and so if a woman presents before this time she will be given an interim diagnosis of a pregnancy of unknown location. Thirdly, the woman may have experienced a complete miscarriage. This will usually be indicated by a history of heavy vaginal bleeding with or without crampy abdominal pain. The serum hCG remains elevated often for several weeks after a miscarriage. However, as 6% of women with such a history are subsequently diagnosed with an ectopic pregnancy [109], it is important not to assume, even with such a history, that a miscarriage has occurred and hence the woman should be given an interim diagnosis of a pregnancy of unknown location and followed up accordingly (unless an intrauterine pregnancy has been visualized previously). The fourth possibility is that uterus is empty because the pregnancy is ectopic and is either too small to be visualized at that time, the adnexa have not been assessed thoroughly or the ectopic is outside the pelvis.

Even with expert use of transvaginal ultrasound using agreed criteria, it may not be possible to confirm if a pregnancy is intra- or extra-uterine in 8-31% of cases at the first visit [43]. However, in specialized scanning units, the overall incidence of pregnancies of unknown location is as low as 8-10%. There is a consensus that modern early pregnancy units should strive to maintain a pregnancy of unknown location rate of 15% or less [110].

Of women given an initial diagnosis of a pregnancy of unknown location, 50-70% will have a final diagnosis of a resolving pregnancy/failing pregnancy of unknown location, 30-50% will be diagnosed with an intrauterine pregnancy on subsequent ultrasound scan, 7-20% of women will be diagnosed with an ectopic pregnancy and less than 5% of women will have a persistent

pregnancy of unknown location [39, 111-113]. A persistent pregnancy of unknown location is when the serum hCG levels fail to decline and no evidence of a pregnancy is ever identified on transvaginal ultrasound, diagnostic laparoscopy or uterine curettage [109].

All women with a pregnancy of unknown location are rigorously followed-up until a definite diagnosis has been made. This at least requires one additional serum hCG and transvaginal ultrasound scan. As intervention is only required in the minority of women with a pregnancy of unknown location (those with an ectopic pregnancy or a persistent pregnancy of unknown location), limited resources are expended and significant anxiety inflicted upon a large proportion of women unnecessarily. It would therefore be preferable if the proportion of women being given a diagnosis of a pregnancy of unknown location was minimised or, at the very least, if there was some way of differentiating women with a pregnancy of unknown location who are at high risk of an ectopic pregnancy from those who are at low risk of an ectopic pregnancy and rationalising follow-up appropriately.

## Minimising the Diagnosis of Pregnancy of Unknown Location

The ideal would be to minimise the diagnosis of a pregnancy of unknown location in the first instance as this would minimise anxiety for women and would also enable limited hospital resources to be more appropriately utilized. There appear to be two main issues to consider:

### Ultrasound Technology

The incidence of pregnancy of unknown location in some early pregnancy units is as high as 31% but in other units it may be as low as 8% [43]. The relative proportions of pregnancy of unknown location are determined by many factors, but the quality of ultrasound examination is probably the most important one. As already discussed, the diagnostic capability of ultrasound in early pregnancy increased dramatically following the introduction of transvaginal ultrasound and it continues to increase with the development of higher frequency transvaginal transducers [15]. Threshold values and discriminatory sizes used to distinguish normal and abnormal pregnancies are smaller on higher frequency than lower frequency imaging.

## Ultrasonographer Expertise

Even if using the latest in ultrasound technology, the ability of the machine to produce good quality images is critically dependent on the capability of the ultrasonographer performing the examination.

Obviously if a woman has no evidence of an intra- or extra-uterine pregnancy on ultrasound scan and definitely has a positive pregnancy test then she has, until proven otherwise, a pregnancy of unknown location. This is uncontroversial. However, anecdotally many women are given the diagnosis of a pregnancy of unknown location when there is ultrasonographic evidence of an intrauterine fluid collection within the uterus and it is these 'pregnancies of unknown location' that have the potential to become pregnancies of uncertain viability.

As already discussed, the yolk sac is currently the first incontrovertible ultrasonographic sign of an intrauterine pregnancy. Many ultrasonographers therefore choose to wait until a yolk sac is visible before confirming that the intrauterine fluid collection that they have detected on ultrasound is indeed a true gestational sac and not a pseudosac. A pseudosac is seen in up to 15% of ectopic pregnancies and caused by the presence of an intrauterine fluid collection surrounded by a thick decidual reaction or a detached decidual reaction (decidual cast) containing fluid. This may appear to be a sensible approach as it is obviously important not to misdiagnose an ectopic pregnancy as an intrauterine pregnancy and vice versa, but, according to experts, using transvaginal ultrasound it is not difficult to differentiate a true gestation sac, which is eccentrically located and surrounded by two layers, from a pseudosac, which is centrally located and only surrounded by one layer [114].

Understandably ultrasonographers may not want or feel able to comment on the origin of the intrauterine fluid collection that they have visualized. They may not be fully aware of the clinical history of the patient and even if they are, they may not be able to interpret the significance of such information. Ultrasonographers therefore need to comment in their report not only on the size of the intrauterine fluid collection that they have seen, and whether or not it has any contents, for example a yolk sac or a fetal pole, but also on its precise location within the cavity and the number of surrounding layers, so

that clinicians can interpret these findings in combination with the clinical picture and are then better equipped to make the distinction between a true gestation sac of an intrauterine pregnancy and a pseudosac of an ectopic pregnancy. If this can be successfully accomplished then the number of women that are given the diagnosis of a pregnancy of unknown location will be significantly reduced.

## Rationalising Follow-up of Pregnancies of Unknown Location

When a diagnosis of a pregnancy of unknown location is made, it is essential to follow women up until a definite diagnosis has been made and although follow-up strategies vary between and within units, most involve the use of hCG with or without progesterone, repeat transvaginal ultrasound examinations and possibly even occasional surgical intervention. The timing of further serum hCG estimations and repeat ultrasound examinations and possible surgery is very ad-hoc and clinician dependent and women can therefore experience a very protracted period of follow-up with multiple clinic visits and possible unnecessary investigations. This would perhaps be acceptable if the majority of women with a pregnancy of unknown location had an adverse outcome but, as already discussed, they do not and it would therefore be preferable to differentiate those women with a pregnancy of unknown location at high risk of having an ectopic pregnancy from those women that are low risk so that follow-up, and limited resources, can be more appropriately rationed. Several different methods have been proposed in an attempt to accomplish this, some incorporating different blood tests, other utilizing mathematical models and a small proportion involving surgical intervention. These shall now be discussed in more detail.

### Serology

#### Human Chorionic Gonadotropin (hCG)

A single serum hCG in the management of pregnancies of unknown location is of limited value. If the serum hCG is above the discriminatory level (usually around 1500iu/l but this can vary) and no intrauterine pregnancy is visualized on ultrasound, some may assume that the woman has an ectopic pregnancy and a diagnostic laparoscopy is subsequently performed. However, this

practice can lead to the performance of unnecessary surgical procedures (in women subsequently found to have a failing pregnancy of unknown location or an intrauterine pregnancy) and may fail to diagnose ectopic pregnancies, as studies have shown that more than 50% of ectopic pregnancies in the pregnancy of unknown location population will have initial serum hCG levels below the discriminatory level [109]. Furthermore, even if these women do have ectopic pregnancies, as their serum hCG is low, they may be suitable candidates for non-surgical management and adopting a wait and see approach and monitoring their hCG rather than immediately resorting to surgery may have its benefits. It is essential however that women with a pregnancy of unknown location are not given methotrexate (even if the hCG is above the discriminatory zone and no intrauterine pregnancy has been seen on scan) unless a viable intrauterine pregnancy has been confidently excluded.

Serial serum hCG measurements have also been used to help predict the outcome in women with pregnancies of unknown location. In a study of 389 women with a pregnancy of unknown location, a hCG ratio (serum hCG at 48hrs/serum hCG at 0hrs) of less than 0.87 predicted a failing pregnancy with a sensitivity and specificity of 93.1% (95% CI 85.9-97.0&) and 90.8% (95% CI 90.0-99.1%) respectively [115]. Whilst the hCG ratio is of more use than a single serum hCG level in the management of pregnancies of unknown location, even that is prone to error and hence caution is required. Solely using serial hCG values to predict the likelihood of ectopic pregnancy in women with a pregnancy of unknown location can result in misdiagnosis as approximately 20% of ectopic pregnancies have a doubling time similar to that of intrauterine pregnancies [34, 106] and 10% will have decreasing hCG levels similar to spontaneous miscarriages [38, 106]. Clinical judgement 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 [116].

**Progesterone**

Serum progesterone can help predict the outcome of pregnancies of unknown location although it is of more use in predicting viability than location. However, a low serum progesterone suggestive of pregnancy failure could be used to help rationalise the follow-up of women with pregnancies of

unknown location. In a study of 1110 women with a pregnancy of unknown location, women with a progesterone of less than or equal to 10nmol/L at presentation were at low (2.1%) risk of requiring medical intervention. If the hCG was also less than 450iu/l, the intervention rate was even lower (1.3%). Such women may not benefit from attending routine follow-up visits [117].

**Cancer Antigen 125**

Cancer Antigen 125 (CA125) is thought to originate from decidual cells affected by chorionic invasion or placental separation. Levels peak during the first trimester of pregnancy and drop to non-pregnant levels in the second and third trimester. Although there is evidence to suggest that CA125 levels do not predict spontaneous miscarriage in the first trimester of pregnancy, there is conflicting evidence regarding the use of CA125 to differentiate between intrauterine pregnancies and ectopic pregnancies. Serum CA125 levels have been reported to be significantly lower in ectopic pregnancies compared to intrauterine pregnancies [118] but in a larger, more recent study [119], single serum measurements of CA125 failed to accurately differentiate between a viable intrauterine pregnancy, spontaneous miscarriage or ectopic pregnancy. The CA125 ratio however (CA125 at 48hrs/CA125 at 0hrs) can distinguish a failing pregnancy of unknown location from an intrauterine pregnancy but unfortunately is not able to detect an ectopic pregnancy, rendering it of limited use in the clinical setting [120].

**Creatine Kinase**

Creatine Kinase is a key enzyme for energy metabolism of contraction and relaxation in both striated and smooth muscle. It has been found in all smooth muscles studied to date including the fallopian tube. One theory is that damage to the fallopian tube in an ectopic pregnancy is sufficient to cause an increase in serum creatine kinase. Although some studies [121, 122] have demonstrated that maternal creatine kinase levels can be an important biochemical marker for the diagnosis of an ectopic pregnancy, the general consensus is that maternal serum creatine kinase levels do not reliably predict ectopic pregnancy [123-126].

**Activin A and Inhibin A**

Activin A and inhibin A are glycoproteins secreted by the feto-placental unit. Their levels have been used to predict miscarriage in women with a history of recurrent miscarriage [43] and to predict the outcome in women with a pregnancy of unknown location [127]. In a study of 141 women with a pregnancy of unknown location, serum concentrations of activin A and inhibin A were measured at 0 and 48 hours. Activin A levels at 0 and 48hrs were found to be unhelpful in predicting the outcome of pregnancies of unknown location as levels were not significantly different between the various outcome groups. Serum inhibin A levels however were significantly lower in the failing pregnancy of unknown location group but they did not perform as well as serial serum hCG levels and are therefore of little use in this clinical context.

Serum hCG, progesterone, inhibin A, inhibin pro-αC-related immunoreactivity, and insulin like growth factor binding protein 1 (IGFBP-1) were measured in 109 women with a pregnancy of unknown location [128]. Progesterone and inhibin A were significantly lower and IGFBP-1 significantly higher in failing pregnancies than in those that required further intervention. In decision tree analysis, the novel markers were less useful than progesterone and hCG in predicting failing pregnancies of unknown location. Inhibin pro-αC-RI and IFGBP-1 were not at all useful in the prediction of failing pregnancies of unknown location. Inhibin A on the other hand was found to be more predictive than hCG alone, but serum progesterone was still considered to be the best single marker and progesterone and hCG combined continue to be the best way of predicting failing pregnancies of unknown location.

**Disintegrin and Metalloprotease Protein-12**

When measured in isolation, disintegrin and metalloproteases protein-12 levels had limited value as a diagnostic biomarker for ectopic pregnancy in 120 women with a pregnancy of unknown location [129]

Unfortunately therefore, at present, no single factor has been identified to accurately predict the outcome of a pregnancy of unknown location.

## Mathematical Models

A number of mathematical models have consequently been developed in an attempt to predict the outcome of pregnancies of unknown location. The main advantage of mathematical models in general is that they do not require any understanding of clinical biochemistry in early pregnancy and their use is independent of clinical experience.

In 2004, three logistic regression models to predict the outcome of pregnancies of unknown location were created from simple demographic and hormonal data and subsequently tested on 185 women with a pregnancy of unknown location [130]. The first model, M1 involved the hCG ratio (hCG at 48hrs/hCG at 0hrs). The second model, M2, incorporated the average progesterone level ((progesterone at 0hrs + progesterone at 48hrs)/2) and the final model, M3, simply utilized the patient's age. When tested prospectively, M1 out-performed M2 and M3. It had an area under the curve (AUC) of 0.975, 0.966 and 0.885 for a failing pregnancy of unknown location, intrauterine pregnancy and ectopic pregnancy respectively. Furthermore, M1 had a sensitivity of 91.7%, specificity of 84.2%, positive likelihood ratio of 5.8, positive predictive value of 27.5% and negative predictive value of 99.4% for the detection of ectopic pregnancy. Although promising, multi-centre trials are needed to test the reproducibility and validity of this model before it is adopted in the clinical setting.

In 2006, Gevaert attempted to develop Bayesian networks to predict ectopic pregnancies in the pregnancy of unknown location population [131]. Variables such as age, abdominal pain, vaginal bleeding, gestational age, endometrial thickness, midline echo, free fluid, hCG ratio, progesterone levels at 0 and 48 hours and the clinical outcome of the pregnancy of unknown location were investigated. The best Bayesian network used the gestational age, the hCG ratio and the progesterone level at 48 hours and had an AUC of 0.88 for predicting ectopic pregnancies when tested prospectively. However these complex models need to be validated, ideally using prospective multi-centre studies, before being relied upon in clinical practice.

In 2007, Condous devised and tested three different models in an attempt to predict pregnancy failure in the pregnancy of unknown location population. The first, M1, was a logistic regression model incorporating vaginal bleeding,

endometrial thickness, initial serum progesterone and hCG levels. The second, M2, utilized the serum progesterone at presentation only and the third model, M3 involved the serum hCG ratio (i.e. serum hCG at 48hrs/serum hCG at 0hrs) [132]. The hCG ratio was found to be the optimal test for the prediction of pregnancy failure in the pregnancy of unknown location population as it had an AUC of 0.980 (SE 0.004) compared to M1 which had an AUC of 0.907 (SE 0.015) and M2 which had an AUC of 0.952 (SE 0.010). The difference between M3 and M1 and M2 was statistically significant ($p<0.0001$ and $p=0.0076$ respectively).

In an attempt to improve on the performance of the M3 model, a new model, M4, was devised and tested against the hCG ratio [133]. This new model involved a multinomial logistic regression model containing the log of the hCG average, the hCG ratio and its quadratic effect. In the prediction of ectopic pregnancy, this new model, M4, gave an AUC of 0.9 compared to M3, which had an AUC of 0.842. Although this difference was found to be statistically significant ($p=0.0303$), in clinical terms, it did not result in substantially more pregnancies being classified correctly as developing ectopic pregnancies. Furthermore, when tested in a cohort in the United States, the M4 model performed less well [134] highlighting that caution is required when applying algorithms from one centre to another, where the definitions of pathology may differ.

The effectiveness of these different models are difficult to compare as they are not all aiming to predict the same outcome. Some are interested in predicting pregnancy failure [132] presumably so that these women can be followed-up less closely or even discharged, whilst others are interested in identifying women at high risk of an ectopic pregnancy that need closer surveillance [131, 133]. Only one looks at predicting all pregnancy of unknown location outcomes [130]. A further disadvantage of all of these models is that unfortunately they require multiple visits in order to be able to collect the data that is needed to be input into the calculation.

## Single Visit Strategies

The possibility of whether women with a pregnancy of unknown location could be safely excluded from potentially unnecessary multiple clinic visits has also

been assessed [135]. A single visit protocol was developed based on data from 200 women. Pregnancies of unknown location were divided into groups according to their probable risk of an ectopic pregnancy. Women were considered to be at low risk if they were thought to have a failing pregnancy of unknown location (i.e. they had an initial serum progesterone of ≤10nmol/l or an initial serum hCG of ≤25iu/l) or a probable intrauterine pregnancy (i.e. an initial serum progesterone of >50nmol/l regardless of the serum hCG). Women with a pregnancy of unknown location and an initial serum progesterone of between 10 and 50nmol/l or an initial serum hCG of >25iu/l were considered to be at high risk of an ectopic pregnancy. This protocol was then tested on 318 consecutive women with an ultrasonographic diagnosis of a pregnancy of unknown location. The sensitivity, specificity, positive predictive value and negative predictive value of this single visit strategy to detect women with a pregnancy of unknown location at low risk of an ectopic pregnancy was 84%, 33%, 96% and 9.4% respectively. In conclusion therefore, although this strategy eliminates 84% of non-ectopic pregnancies correctly from the system, 67% of ectopic pregnancies are discharged without adequate follow-up rendering this single visit strategy inappropriate for the management of women with pregnancies of unknown location.

More recently, the efficacy and safety of a modified clinical protocol using serum progesterone for the management of women with a pregnancy of unknown location has been prospectively evaluated [136]. Two hundred and fifty two women with a pregnancy of unknown location and a serum progesterone of ≤10nmol/l were discharged after their initial visit and followed-up via telephone four weeks later. These women accounted for 37% of the total pregnancy of unknown location population. Follow-up was complete in 227 of the 252 women (90%). The pregnancy resolved without any complications in 212 of the 227 women (93.4%). Fifteen women (6.6%) re-attended with persistent or worsening symptoms. Five women (2.2%) were diagnosed with an ectopic pregnancy following a subsequent ultrasound examination. Of these, two had surgery and three were managed conservatively. Another two women had surgery. One had a negative laparoscopy and an evacuation of retained products of conception but no products of conception were found and another was diagnosed with an incomplete miscarriage and had an evacuation of retained products of conception in which histology confirmed an intrauterine pregnancy. Therefore only four women (1.8%) needed surgical intervention. The positive

predictive value of a serum progesterone of ≤10nmol/l for diagnosing a spontaneously resolving pregnancy was therefore found to be 98.2% (95% CI 96.8-99.7%) and it was concluded that a clinical protocol based on a single serum progesterone measurement is effective for triaging women with pregnancies of unknown location as it reduces the need for follow-up of these women without compromising safety.

It therefore appears possible to triage women with a pregnancy of unknown location at low risk of an ectopic pregnancy so that they can receive less intensive or even no follow-up but the clinically important identification of women with a pregnancy of unknown location at high risk of ectopic pregnancy remains elusive following a single visit strategy.

## Surgery

Uterine curettage and diagnostic laparoscopy are occasionally used in the management of women with a pregnancy of unknown location. One commonly used algorithm to diagnose ectopic pregnancy is based on the use of transvaginal ultrasound to demonstrate the absence of an intrauterine gestation sac followed by uterine curettage if a viable intrauterine pregnancy has first been excluded either on the basis of a low progesterone level or a suboptimal rise in hCG over 48 hours. Following curettage, if the hCG decreases by less than 15% or increases, a diagnosis of ectopic pregnancy is assumed. This protocol could however result in the inadvertent termination of viable intrauterine pregnancies as we have already explained the diagnostic limitations of using progesterone and/or hCG to determine pregnancy location. Uterine curettage should therefore not have a routine place in the management of women with pregnancies of unknown location although it may have a role in identifying the location of the pregnancy after the possibility of a viable pregnancy has been excluded. This may be important because it can help advise on future pregnancies.

Similarly, diagnostic laparoscopy should not be used in the routine management of women with a pregnancy of unknown location. In some units, as already mentioned, it is performed when there is no intrauterine pregnancy visible on ultrasound and the hCG is above the discriminatory level. This may not only lead to the performance of unnecessary surgery (in women subsequently found to have a failing pregnancy of unknown location

or an intrauterine pregnancy) but it may also falsely reassure clinicians as many women with ectopic pregnancies have serum hCG levels below even the lowest used discriminatory levels. Similarly, some women with failing pregnancies of unknown location and early intrauterine pregnancies have very high serum hCG levels and may therefore undergo unnecessary surgery with its associated risks.

## 3.3 Pregnancies of Uncertain Viability

As eluded to in chapter two, the definition of a pregnancy of uncertain viability is when there is an intrauterine gestation sac less of than 25mm mean diameter with no obvious yolk sac or fetal pole or a fetal pole of less than 7mm with no obvious fetal heart activity.

The visualization of a small intrauterine gestation sac without a fetal pole, or a small fetal pole without fetal heart pulsation is not an uncommon finding during a transvaginal ultrasound scan in very early pregnancy [137]. In cases of known intrauterine location, viability will be uncertain in approximately 10% of women at their first visit to the Early Pregnancy Assessment Unit ([138]. It may represent a normal early pregnancy of between four and six weeks' gestation or it may be a failed or failing pregnancy with arrested development, which is destined to miscarry.

It is therefore not difficult to conceive that erroneous diagnoses of early fetal demise can occur in this situation, which may subsequently lead to the unintentional termination of a viable pregnancy. Reports of such diagnostic errors prompted a national enquiry in the UK in the mid-1990s [139]. The diagnostic guidelines, which were issued as a result, state that no ultrasound diagnosis of early fetal demise should be made at the initial visit. In order to confirm or refute viability, a repeat scan at a minimal interval of one week is recommended [140] and in cases of an empty sac of less than 15mm in diameter, the guideline requires that a follow-up scan be organized two weeks later. This policy, whilst minimising the risk of diagnostic errors, contributes greatly to the workload of often already under-resourced Early Pregnancy Assessment Units, not to mention causing a tremendous amount of anxiety for women and their partners. Some authors therefore, whilst acknowledging the value of time and serial observation, do not agree with

these diagnostic guidelines, preferring to spare women from unnecessary follow-up examinations when a failed pregnancy has been reliably diagnosed [141].

The prediction of outcome in cases of pregnancies of uncertain viability is challenging [139, 141]. Failure to identify certain landmarks ultrasonographically by a certain gestation should not be used to reliably diagnose pregnancy failure because true gestational age cannot be confirmed in the majority of cases as the reported date of the last menstrual period is notoriously unreliable [142-144]. Furthermore, growth in the first trimester is variable [145-148] and not, as was previously assumed, uniform [149].

Several studies have attempted to establish methods of differentiating viable from non-viable pregnancies in cases of pregnancies of uncertain viability.

## Predicting Outcome

### Serology

The measurement of serum hCG and progesterone have been used in the past in an attempt to discriminate between various pregnancy complications. Although there is some evidence that measuring serum hCG is helpful in the diagnosis of ectopic pregnancy, most studies have concluded that it cannot reliably discriminate between viable and non-viable pregnancies [150]. For example, in a study of 200 women with a pregnancy of uncertain viability, serum hCG levels were not significantly different between the women that subsequently miscarried (n=82) and those that did not (3556iu/L versus 3974iu/L) [151].

Having said that, an increase in hCG of at least 66% in forty eight hours is associated with an intrauterine pregnancy with a positive predictive value of 96.5% [34] and a decrease in hCG of at least 15% in a forty eight hour period is most usually associated with a failing pregnancy [33] although a similar decline is seen in approximately 8% of ectopic pregnancies [38, 106].

The inability of hCG to differentiate viable from non-viable pregnancies with certainty is in part due to the wide range of hCG levels recorded in normal pregnancy and the long half-life in blood following fetal demise.

Serum progesterone levels are elevated during pregnancy and change little during the first 8-10 weeks of gestation unless the pregnancy fails [33]. Several studies [39-41], including a meta-analysis [42], have shown serum progesterone levels less than 25nmol/l to be associated with non-viability although a small proportion (0.3%) of viable pregnancies have been reported with initial levels <15.9nmol/l [152]. Levels less than 20nmol/l have a positive predictive value of more than 95% at predicting a failing pregnancy, and this compares favourably with complex multi-parameter diagnostic models [40]. Progesterone, whilst useful, is not diagnostic and therefore cannot be relied upon in clinical practice.

## Ultrasonography

Various authors have attempted to predict the outcome of pregnancies of unknown viability following a single ultrasound scan. The benefit of this approach is that it minimises additional workload for Early Pregnancy Assessment Units and does not cause undue anxiety for the couple.

### Single Visit Strategies

A study by Tongsong et al [153] concluded that it was possible using transvaginal ultrasound to distinguish viable and non-viable empty gestational sacs at a single examination in the majority (73%) of cases. They found that in a study of 211 women with threatened miscarriage and an empty gestation sac on ultrasound, a mean sac diameter was the most useful criterion for determining non-viability. A gestational sac with a mean diameter of 17mm or more without an embryo or a gestational sac with a mean diameter of 13mm or more without a yolk sac were reliable predictors of non-viability with both a specificity and positive predictive value of 100%. A gestational sac with a mean diameter of 13mm or more without a yolk sac had the greatest sensitivity.

In another study the relationship between maternal age, menstrual age, mean gestation sac diameter, presence or absence of a yolk sac and/or sub-

chorionic haematoma and serum hCG levels and pregnancy outcome was investigated in 50 women with vaginal bleeding in the first trimester and an ultrasound scan showing a gestation sac with a mean sac diameter of 16mm or less and no visible fetal pole [147]. 64% of women subsequently miscarried prior to 22 weeks gestation. A yolk sac was visualized in 72.2% of the viable pregnancies but only 40.6% of the non-viable pregnancies suggesting that failure to visualize the yolk is a risk factor for pregnancy failure. Sub-chorionic haematomas were not found to be of any prognostic significance. Maternal age, menstrual age, mean gestation sac diameter standard deviation score (defined as the number of standard deviations from the expected mean) and serum hCG level were significantly correlated with subsequent miscarriage with AUCs of 0.73, 0.84, 0.91 and 0.66 respectively. In particular, all women over 35 years of age, or with a serum hCG of less than 1200mUI/l or a gestational age of more than 7 weeks or a mean gestation sac diameter more than 1.64 standard deviations below the mean underwent a miscarriage. However when these variables were entered into a multiple logistic regression analysis, only the mean gestation sac standard deviation score was found to be independently associated with miscarriage.

Other features in addition to mean gestation sac diameter have been proposed in an attempt to predict the subsequent outcome of pregnancies with initially uncertain viability. In the aforementioned study by Tongsong, deformed gestational sac shape, a low position within the uterine cavity and a thin decidual reaction were all shown to be strongly suggestive of non-viability but were not 100% accurate [153] and therefore these features should not be relied upon in clinical practice.

Another study involving thirty five women with either an empty gestational sac or a gestational sac containing a yolk sac but no fetal pole looked at the size and shape of the gestation sac, the prominence of the trophoblastic reaction and the continuity of the reaction around the sac in an attempt to determine if any of these features could be used to help predict outcome in pregnancies of uncertain viability [154]. Fifteen pregnancies (42%) were subsequently proven to be viable, whilst the remainder (n=20) were non-viable. The average diameter of the gestational sac in the viable pregnancies was 13mm compared to 18mm in the non-viable pregnancies. There was however considerable overlap in the sac size between the two groups to enable accurate differentiation between the two. Furthermore, although 80% of non-

viable pregnancies had a pointed sac, 53% of viable pregnancies also exhibited this peculiarity rendering sac shape an inaccurate predictor of pregnancy outcome. With regards to the quality of the trophoblastic reaction, although 100% of the viable pregnancies had a good or fair reaction, 63% of the failing pregnancies also demonstrated similar findings. 87% of the viable pregnancies had a continuous trophoblastic reaction around the gestational sac compared to only 50% of the non-viable pregnancies but again there was considerable overlap between the two groups rendering this sign of little use clinically. However, 67% of the viable pregnancies had a good ring of continuous trophoblast compared to only 12.5% of the non-viable pregnancies. Thus this combination of these two features has a higher degree of accuracy in predicting pregnancy outcome than either factor alone.

In addition to the gestation sac, the appearance and volume of the yolk sac has been extensively assessed but its relationship with pregnancy outcome remains controversial.

In a study of 250 asymptomatic first trimester pregnancies, an abnormal yolk sac shape was found in 29% (n=31) of pregnancies that subsequently miscarried compared to only 4.6% (n=219) in pregnancies that continued beyond the first trimester [155]. An abnormal yolk sac shape was found to predict adverse pregnancy outcome with a sensitivity of 29%, specificity of 95%, positive predictive value of 47% and negative predictive value of 90.5%.

In the same study, the diameter of the yolk sac was also assessed. It was found to strongly correlate with both menstrual age (R=0.958) and crown rump length (R=0.943). Pregnancies that miscarried within the first trimester were more likely to have a yolk sac diameter greater than two standard deviations from the mean than pregnancies that continued beyond the first trimester (64.5% versus 3.7%). Using two standard deviations as the cut off, abnormal pregnancy outcome was predicted with a sensitivity, specificity, positive predictive value and negative predictive value of 65%, 97%, 71% and 95% respectively [155]. Hence, yolk sac diameter was shown, in this study at least, to be a better predictor of adverse pregnancy outcome than yolk sac shape.

Conversely, a much more recent study [156] involving 62 early first trimester pregnancies found the yolk sac volume to be poorly correlated with both

gestational age (R=0.188) and crown rump length (R=0.203). This may be due to differences in caliper placement and measurement technique between the various studies however.

The relative development of the yolk sac to the gestation sac has also been investigated as a prognostic marker for pregnancy outcome [157]. In a study of 49 pregnancies conceived following IVF, three-dimensional measurements of the gestation sac, yolk sac and crown rump length were used to form nomograms in the pregnancies where there were normal outcomes. Measurements from the abnormal pregnancies were then compared with these nomograms. The embryo volumes and gestation sac volumes were shown to have a good predictive value for adverse pregnancy outcome ($p<0.05$) but the yolk sac volumes were not found to differ significantly between viable and non-viable pregnancies. The yolk sac volume:gestation sac volume ratio was however found to have a good predictive value but this is likely to be a function of the effect of gestation sac volume alone.

In another study however [158], the mean diameter and volume of the yolk sac and gestational sac, fetal heart rate, maternal age, gestational age and presence of a subchorionic haematoma were recorded in 125 asymptomatic women with a singleton pregnancy. Twenty-five women were included at each gestational age between six and ten weeks and nomograms were constructed for volumes, mean diameters and fetal heart rate. The main outcome measure was miscarriage before 20 weeks gestation. Regression analysis demonstrated that the only variables that were significantly associated with spontaneous miscarriage were maternal age more than 34 years, gestational sac mean diameter less than the 5$^{th}$ percentile and a fetal heart rate less than the 5$^{th}$ or more than the 95$^{th}$ percentiles. The authors therefore concluded that new three-dimensional parameters were of no additional clinical benefit in predicting miscarriage in asymptomatic first trimester pregnancies.

Some pregnancies of uncertain viability however have a visible fetal pole but no visible fetal heart pulsation. There are multiple references in the literature to the ultrasonographic detection of early embryos before the onset of cardiac contractility [23, 25, 159-161]. Therefore it is possible that early embryos without visible heart pulsations are actually alive but their hearts have simply not yet

begun to beat. Equally it is possible that such embryos represent failing pregnancies.

The 'yolk stalk sign' has been proposed to help differentiate between the two possible diagnoses [162]. In early embryonic development, the yolk stalk has not yet developed. Thus the embryo is detected immediately adjacent to the yolk sac, producing the so-called 'signet ring' appearance (Figure 1.3c). If the embryo is separated from the yolk sac, the separation is due to the development of the yolk stalk. As the yolk stalk lengthens, the distance between the embryo and the yolk sac increases. For small embryos (less than 5mm), there should be no separation between the embryo and the yolk sac because the yolk stalk has not yet developed. The depiction of separation of an embryo with a crown rump length of 5mm or less from the yolk sac indicates development of a yolk stalk and thus a more advanced stage of gestation than would have been deduced from the crown rump length alone.

In a retrospective study of 159 pregnancies with a crown rump length of 5.4mm or less with no visible heartbeat, 21 (13.2%) demonstrated this yolk stalk sign. All of these cases were subsequently proven to be failed pregnancies, thus the positive predictive value of this sign in determining early pregnancy failure was 100% in this cohort. As it was a retrospective study, no effort was made during the ultrasound examination to demonstrate this finding, hence the sensitivity of the yolk stalk sign for predicting miscarriage was not calculated [162].

In addition to the size and appearance of the gestation sac and yolk sac, single measurements of fetal size below expected for gestational age can also help predict subsequent miscarriage. In 1991, Koornstra et al [163] suggested a link between first trimester growth restriction and subsequent miscarriage. Using customized first trimester growth charts, the authors calculated the difference between the observed crown rump length of 403 singleton pregnancies and the expected crown rump length based on the duration of amenorrhoea. When the observed crown rump length was 7 or more days less than the expected crown rump length, the risk of miscarriage was 16%, compared to only 5% when the difference was less than 7 days ($p<0.001$).

Similarly, in a study of 292 women with certain menstrual dates and a single viable fetus, the initial crown rump length of the 14% who subsequently

miscarried was below the expected mean for gestational age. In 61% of those that miscarried, the crown rump length was at least two standard deviations below expected [145].

**Multiple Visit Strategies**

Other studies, involving multiple visits and serial ultrasound scans, have looked at the growth rate of various structures in an attempt to establish cut off values that could be used to definitively confirm or refute viability. Such information would be extremely useful clinically. Although it is recommended that when a pregnancy of uncertain viability is diagnosed a repeat ultrasound scan at a minimum period of one week is required, few, if any protocols define how the ultrasound findings are expected to change over that period of time. This lack of clarity leads to both unnecessary protracted follow-up and potentially inappropriate intervention.

In a study of 83 women [164], those that had a viable pregnancy (n=53) had a mean gestation sac growth velocity of 1.13mm/day (range 0.71-1.75mm) compared to 0.70mm/day (range 0.14-1.71mm/day) which was observed in women that were subsequently proven to have a non-viable pregnancy (n=30). The authors concluded that a mean gestational sac growth of less than 0.6mm per day is associated with an abnormal outcome. However, as there is considerable overlap between the ranges in the two groups, this should be interpreted with caution.

In 2007, the crown rump length growth rate in viable pregnancies that continue beyond the first trimester was compared to that of viable pregnancies that subsequently miscarry in the first trimester [165]. Using functional linear discriminant analysis (FLDA), the crown rump length growth rate was found to be significantly lower ($p<0.001$) in the group that miscarried. FLDA was found to differentiate between normal and abnormal growth and thus predict miscarriage with a much higher specificity than a single measure of crown rump length (93.1% versus 72.2%).

A multi-centre study looking at changes in both mean gestational sac diameter and embryonic crown rump length in women initially diagnosed with a pregnancy of unknown viability and correlating these with subsequent pregnancy outcome found that both the mean gestational sac diameter

growth and crown-rump length growth was significantly higher in viable pregnancies compared to non-viable pregnancies (1.003 versus 0.503mm/day and 0.673 versus 0.148mm/day respectively) [166]. However, there was an overlap in mean sac diameter growth rate between ultimately viable and non-viable pregnancies of uncertain viability and no cut off was found below which a viable pregnancy could be confidently excluded making it of little use in clinical practice. However a crown rump length growth rate of 0.2mm/day or less was always associated with a miscarriage and the finding of an empty gestational sac on two scans more than a week apart was highly suggestive of miscarriage, irrespective of sac growth.

**Mathematical Models**

As we have seen, it is difficult using solitary features to be able to predict with certainty the eventual outcome of pregnancies of uncertain viability. Various authors have therefore investigated the use of mathematical models in an attempt to identify any clinical, ultrasonographic or biochemical parameters that may be of clinical benefit.

A prospective observational study of 200 pregnant women with an initial transvaginal ultrasound finding of a gestation sac of less than 20mm and no obvious fetal pole was undertaken [151]. Information was collected including maternal age, gestational age, symptomatology, mean gestational sac diameter and serum progesterone and hCG levels. A repeat ultrasound was performed one to two weeks after the initial scan. Of the 200 women included, miscarriage was confirmed in 82 (41%) at the time of the repeat ultrasound examination. Women that miscarried were significantly older (32.3 years versus 29.3 years), more likely to have experienced vaginal bleeding (76.8% versus 34.7%), have a larger mean gestation sac diameter (10.7mm versus 6.8mm) and a lower serum progesterone level (31nmol/L versus 84nmol/L) than women who did not miscarry.

Using a regression equation derived from the forward stepwise selection of variables, maternal age, mean gestation sac diameter and serum progesterone were found to be independent with statistically significant coefficients and were therefore included in a logistic regression model. Using this model, at a cut-off value of 10% probability, the diagnosis of a viable intrauterine pregnancy could be made with 99.2% (95% CI 95.8-99.97)

sensitivity and 70.7% (95% CI 61.3-78.9) specificity. A comparison of receiver operating characteristic (ROC) curves showed that for the logistic regression model, the AUC was 0.9693, which was significantly better than all the other parameters (gestational age, mean gestation sac diameter, maternal age and serum hCG) except serum progesterone, which had an AUC of 0.9493.

However, in order to ensure no cases of viable pregnancy were wrongly classified as non-viable, the cut-off value of probability had to be decreased to 1%. At this level the sensitivity is 100% (95% CI 97.5-100) but the specificity decreases to 43.9% (95% CI 34.6-53.6). This makes the model less useful clinically, especially since an almost identical result could be achieved using a serum progesterone cut-off level of 25nmol/L, in which viable pregnancies could be diagnosed with 100% sensitivity (95% CI 96.8-100) and 40.2% specificity (95% CI 31.1-50.0). In addition, the model is complex to use and therefore necessitates the use of an electronic calculator or computer, which may or may not be readily available. It does however give a numerical probability of the pregnancy being viable, which improves counseling of women. The use of the logistic regression model expresses the degree of diagnostic uncertainty and gives the opportunity to patients and physicians to decide at what level of probability the treatment for the presumed medical condition may be initiated.

Similarly, in a prospective study of 493 women, an attempt was made to define the incidence and outcome of pregnancies of uncertain viability and identify any maternal demographic, ultrasound and symptom variables which could help to predict first trimester outcome, using a mathematical model or scoring system [167]. Demographic details, obstetric history and pregnancy symptoms were sought prior to undertaking a transvaginal ultrasound scan in which the gestation sac, yolk sac and fetal pole were measured (if present). The presence or absence of a subchorionic haematoma was also noted. A repeat ultrasound was performed one to two weeks later and the outcome recorded. If the outcome at that stage was a viable pregnancy, a further ultrasound scan was performed at the end of the first trimester to confirm ongoing viability.

Two hundred and ninety eight (60.4%) women had an empty gestational sac during the first ultrasound scan. Of these, 178 (59.7%) had a viable pregnancy confirmed at the follow-up scan one to two weeks later. 165 (33.5%) women

had a gestational sac containing a yolk sac but no fetal pole at the initial scan and of these, 127 (95.2%) had a viable pregnancy confirmed during the follow-up scan. The remaining women (n=30) had a small fetal pole on the initial ultrasound scan but no visible fetal heartbeat. Of these, only two (6.7%) had a viable pregnancy confirmed during the subsequent scan. In this cohort therefore, 37.7% of women given an initial diagnosis of a pregnancy of uncertain viability were confirmed to have miscarried within two weeks.

Of the 307 pregnancies that were deemed viable after one to two weeks, first trimester outcome data was available for 258 (84.0%). Of these, 225 (87.2%) remained viable and the remainder miscarried by the end of the first trimester. In the group with no embryonic structures visible on initial scan, 44.4% had a pregnancy which progressed into the second trimester. In the group with a yolk sac but no fetal pole visible initially, 71.0% continued beyond the first trimester and in the group with a small fetal pole on the initial ultrasound scan but no visible fetal heartbeat, all that were viable at the initial follow-up ultrasound scan progressed into the second trimester. Therefore, although almost two thirds of women with a pregnancy of uncertain viability may have a viable pregnancy confirmed at the initial follow-up scan, ultimately only half will have pregnancies that progress into the second trimester and beyond.

Univariable analysis showed that pain, previous miscarriage, parity and the presence of a sub-chorionic haematoma were not associated with miscarriage. Advanced maternal age, gestational age, heavy vaginal bleeding, increasing deviation of the mean gestation sac diameter above or below 7mm, absence of a yolk sac and mean yolk sac diameter deviation above 4mm on the other hand were all associated with miscarriage at one to two weeks and at the end of the first trimester. In the multivariable model, only maternal age, gestational age, bleeding score, increasing mean gestation sac diameter deviation from 7mm and presence of a yolk sac remained significant. Logistic regression models provided the estimated probability of viability at the initial one to two weeks follow-up scan and at the end of the first trimester for a woman with a pregnancy of uncertain viability and gave AUCs of 0.821 (95% CI 0.756-0.885) and 0.774 (95% CI 0.701-0.848) respectively.

This study concludes that in women with a pregnancy of uncertain viability, whilst definitive prediction of pregnancy outcome with 100% accuracy might not be possible, it could be predicted with reasonable accuracy using

mathematical models. This information could be used in order to prepare couples for the predicted likely outcome and time follow-up scans accordingly. There is evidence to suggest that women would benefit psychologically from this type of information, even if it is not absolute [168].

## 3.4 The Optimal Timing of an Ultrasound Scan

Whilst performing an ultrasound scan too early in pregnancy may lead to the diagnosis of a pregnancy of unknown location or a pregnancy of uncertain viability which has its associated hazards, deferring an ultrasound scan based on arbitrary limits for gestation (especially when as many as 45% of pregnant women have a suspect menstrual history [169]) may also be associated with physical and psychological morbidity or mortality due to a delay in the diagnosis of a miscarriage [170] or ectopic pregnancy [171].

Based on a study of 1442 women [172], the commonest transvaginal ultrasound finding prior to 35 days gestation was a pregnancy of unknown location and from 42 days gestation it was a viable intrauterine pregnancy although the chance of confirming viability increased rapidly per day of gestation until 49 days and thereafter plateaued. A miscarriage could not be diagnosed on initial transvaginal ultrasound prior to 35 days gestation. Between 35 and 41 days gestation the commonest transvaginal ultrasound finding was a pregnancy of uncertain viability. It was concluded therefore that in asymptomatic women, with no previous ectopic pregnancy, transvaginal ultrasound could be delayed until 49 days gestation, which would decrease the number of inconclusive ultrasound scans performed, without an associated increase in morbidity from missed ectopic pregnancy. Symptomatic women however should not have an ultrasound scan deferred, regardless of gestation, due to the risk of serious morbidity associated with a missed diagnosis of an ectopic pregnancy.

## 3.5 Summary

This chapter has introduced the concept of diagnostic uncertainties in early pregnancy, namely pregnancies of uncertain location and pregnancies of unknown viability, and the clinical difficulties associated with them. It has outlined strategies to minimise the diagnosis of pregnancies of unknown

location (including advancements in ultrasound technology and improvements in ultrasonographer training) and methods to rationalise the follow-up if a pregnancy of unknown location diagnosis cannot be avoided (including multiple serological markers, mathematical models, single visit strategies and surgical intervention). This chapter has also discussed methods of predicting pregnancy outcome when a pregnancy of uncertain viability diagnosis has been made (including serological markers, ultrasonographic findings and mathematical models) and the optimal timing of an ultrasound scan in asymptomatic women to avoid diagnostic uncertainties in early pregnancy.

## 3.6 Rationale for Research

As we have discussed, diagnostic uncertainties in early pregnancy are a relatively common phenomenon. At present, all methods to differentiate viable from non-viable pregnancies in pregnancies of uncertain viability, and intrauterine from ectopic pregnancies in pregnancies of unknown location, lack the required levels of diagnostic accuracy.

In addition to this, diagnostic uncertainties in early pregnancy are associated with many hazards. Firstly, it is likely that they generate considerable anxiety for women and their partners. Secondly, the follow-up of women with uncertain diagnoses contributes a significant part of the workload of Early Pregnancy Assessment Units, taking precious time and utilizing valuable and limited resources. Delayed diagnosis of early pregnancy complications such as non-viable and ectopic pregnancies can have catastrophic consequences, not only for a woman's health, but also her future fertility. Whilst waiting for a definitive diagnosis, a stable woman with an unknown miscarriage or ectopic pregnancy, may deteriorate and become unstable and require immediate resuscitation, life-saving blood transfusion and/or emergency surgery. Therefore, minimising the number of women given these uncertain diagnoses in early pregnancy, or at the very least, minimising the duration of uncertainty when the diagnosis is unavoidable, would be preferable. This rationale forms the premise for the research projects undertaken as part of this thesis.

# 4. Anxiety Associated With Diagnostic Uncertainties In Early Pregnancy

*NB: An abridged version of this chapter has been published in the journal, Ultrasound in Obstetrics and Gynaecology [173]*

## 4.1 Introduction

Approximately one in five women experience abdominal pain and/or vaginal bleeding in early pregnancy. This usually prompts referral to an Early Pregnancy Assessment Unit. Following a pelvic ultrasound, women may be given one of five possible diagnoses, which may be certain or uncertain. Certain diagnoses may be positive i.e. a viable intrauterine pregnancy or negative i.e. a non-viable intrauterine pregnancy or an ectopic pregnancy. Uncertain diagnoses include pregnancies of unknown location or uncertain viability. Pregnancies of uncertain viability are diagnosed in approximately 10% of women attending an Early Pregnancy Assessment Unit and a further 8-31% of women are diagnosed with a pregnancy of unknown location [43, 109]. All women with uncertain diagnoses need to be followed up until a definitive diagnosis is made. Follow-up of these uncertainties utilizes limited resources and is often haphazard and protracted and women may deteriorate clinically during the process.

In medicine, diagnostic tests, such as ultrasound, inherently harbor uncertainty. Uncertainty, defined as a cognitive state created when an event cannot be adequately structured or categorized because sufficient cues are lacking [174], occurs when the decision-maker is unable to assign definite values to objects and events and/or is unable to accurately predict outcomes [175]. It is generated by events characterized as vague, ambiguous, unpredictable, unfamiliar, inconsistent, or lacking information [176]. Uncertainty in medicine causes stress [177] and has been linked to poor coping with health-related issues, as well as poor adaptation and recovery [176].

Enduring negative affect styles such as stress, anxiety and depression have been found to be associated with greater morbidity and mortality in a range of clinical contexts including chronic heart disease [178], autoimmune disease [179], multiple sclerosis [180], upper respiratory tract infection [181], human immunodeficiency virus [182] and cancer [183]. In early pregnancy specifically, women who feel stressed, anxious, depressed, out of control and/or overwhelmed in the first trimester have a higher risk of miscarriage than those who feel happy, relaxed and in control (OR 2.47; 95% CI, 2.02-3.02) [184].

The Transactional Model of Stress suggests that stress is the result of an interaction between a person and their environment [185]. When an individual finds the environmental demands taxing and/or threatening, and simultaneously feels insufficiently equipped to cope with them due to a lack (actual or perceived) of personal or environmental resources, stress is experienced. Biological, psychological and/or behavioral stress responses then ensue depending on the level of perceived threat. This implies therefore that different people may experience the same event in a completely different way based on their individual appraisal of the situation. In the early pregnancy setting therefore, a woman who has had several miscarriages previously or who took a long time to conceive might perceive the experience of abdominal pain and/or vaginal bleeding to be a greater threat than a woman who is contemplating termination of pregnancy.

Whilst there is an abundance of evidence in the literature regarding the psychological sequelae following miscarriage and, to a lesser extent, ectopic pregnancy [186-193], very little is known about how uncertain diagnoses in early pregnancy affect women.

## 4.2 Aims

The aim of this study therefore was to determine anxiety levels of women presenting to Early Pregnancy Assessment Units with abdominal pain and/or vaginal bleeding and assess how these change over time and according to ultrasonographic diagnosis (viable and non-viable intrauterine pregnancies, ectopic pregnancies and pregnancies of unknown location and uncertain viability).

## 4.3 Hypotheses

- Amongst women with viable and non-viable intrauterine pregnancies, ectopic pregnancies and pregnancies of unknown location or uncertain viability, there would be no significant difference in the level of trait anxiety
- Amongst women with viable and non-viable intrauterine pregnancies, ectopic pregnancies and pregnancies of unknown location or uncertain viability, there would be no significant difference in the level of state anxiety prior to the ultrasound scan
- Immediately after the ultrasound scan, women with a certain positive diagnosis i.e. a viable intrauterine pregnancy, would have the lowest levels of state anxiety compared to any other group
- Women with a certain diagnosis would have lower levels of anxiety than women with an uncertain diagnosis 48-72 hours after the ultrasound scan
- Anxiety levels would decrease over time for women with certain diagnoses, especially if they were also positive, and increase over time for women with uncertain diagnoses

## 4.4 Methods

### Ethical Approval

Ethical approval for this study was obtained from Nottingham 1 Research Ethics Committee (13-EM-0081) (Appendix 1).

### Assessment of Anxiety

Anxiety, defined as a negative emotional state characterized by subjective feelings of tension, apprehension and nervousness, may occur as a transitory state, existing at a given moment in time or in response to a particular situation i.e. state anxiety, or as a stable disposition i.e. trait anxiety.

Two different measures of anxiety were considered for the assessment of anxiety in this study: Spielberger's State-Trait Anxiety Inventory (STAI) [194] and the Beck Anxiety Inventory (BAI) [195]. Although there are numerous measures of anxiety in existence, these two were selected initially because they fulfilled

certain pre-defined criteria, specifically they were both considered to be: measures of general anxiety (rather than specific anxiety disorders such as panic disorder, obsessive-compulsive disorder and post-traumatic stress disorder) and the severity of anxiety symptoms; administered by self-report; and supported by adequate psychometric data.

The purpose of the STAI is to measure the presence and severity of current symptoms of anxiety (state anxiety) and a generalized propensity to be anxious (trait anxiety). As the STAI provides operational measures of both state and trait anxiety, it consists of two self-report questionnaires, one for assessing state anxiety and the other for determining trait-anxiety. The different anxieties are deliberately separated so that both questionnaires are reliable in their own right.

The state anxiety scale evaluates the current state of anxiety, asking how a respondent feels 'right now' using items that measure subjective feelings of apprehension, tension, nervousness, worry and activation/arousal of the autonomic nervous system. Responses for the state anxiety scale assess the intensity of feelings 'at this moment' and include: (1) not at all (2) somewhat (3) moderately so and (4) very much so.

The trait anxiety scale evaluates relatively stable aspects of 'anxiety proneness' including general states of calmness, confidence and security. The questionnaire consists of 20 questions, including anxiety-present and anxiety-absent questions and are rated using a 4-point Likert scale. Responses assess the frequency of feelings 'in general' and include: (1) almost never (2) sometimes (3) often and (4) almost always.

For each scale, the item scores are added. Scoring is reversed for anxiety absent items. Scores range from 20 to 80 and higher scores represent greater levels of anxiety. A cut-off of 39-40 has been suggested to detect clinically significant symptoms of state anxiety [196, 197], although a higher cut-off of 54-55 has been suggested for older adults [198].

The BAI is a brief measure of anxiety that focuses on the somatic and, to a lesser extent, cognitive symptoms of anxiety. The cognitive subscale provides a measure of fearful thoughts and impaired cognitive functioning, for example, fear of losing control, of the worst happening or of dying, and the

somatic subscale measures the symptoms of physiological arousal such as numbness, light-headedness and palpitations. It has a total of 21 items in which respondents indicate how much they have been bothered by each symptom over the past week. Responses are rated on a 4-point Likert scale and include: (0) not at all (1) mildly, it did not bother me much (2) moderately, it wasn't pleasant at times and (3) severely, it bothered me a lot.

Scoring is easily accomplished by summing the scores of each item. Scores range from 0 to 63. The following guidelines are recommended for the interpretation of scores: 0-7, normal or no anxiety; 8-15, mild to moderate anxiety; 16-25, moderate to severe anxiety; and 26-63, severe anxiety.

Both the STAI and the BAI have been shown to have good reliability. Test-retest coefficients range from 0.31 to 0.86 (with intervals ranging from one hour to 104 days) for the STAI and 0.62 to 0.93 (with intervals ranging from one to seven weeks) for the BAI [194, 199-201]. Similarly, internal consistency alpha coefficients were also high for both inventories, ranging from 0.86 to 0.95 and 0.90 to 0.94 for the STAI and BAI respectively [194, 199-201]. The validity of both the STAI and the BAI is considered to be moderate although both are somewhat limited in being able to discriminate anxiety from depression [202-205].

The BAI has been demonstrated to be responsive to change over time [206, 207] as has the state anxiety subscale of the STAI. The intention of the trait anxiety subscale of the STAI is to characterize anxiety proneness as a longstanding trait or characteristic and as such it is less responsive to change.

Both the STAI and the BAI are brief to administer (taking approximately 5-10 minutes to complete) and simple to interpret. They have both been translated into multiple different languages and have both been used in a variety of clinical and research settings.

The STAI was ultimately selected for use in this study primarily because, in addition to the twenty item subscale for assessing state anxiety, there also exists a standardized short form (SSF) of the inventory [208], which consists of only six questions and produces scores ranging from six to 24. This shortened version of the state anxiety subscale is therefore even less burdensome for the respondent to complete. The STAI-SSF has acceptable reliability and produces

scores that are similar to those produced with the full-form across subject groups manifesting normal and raised levels of anxiety. Two of the three anxiety-absent items it contains are those identified by Spielberger to be particularly sensitive to low stressors whilst all three of the anxiety-present items included he reported as being particularly sensitive to high stressors. When compared with the full-form of the inventory, the six-item short form offers a briefer and more acceptable scale for subjects whilst maintaining results that are comparable to those obtained from the full-form and remains sensitive to different degrees of anxiety. Further benefits of the abridged inventory are that it is likely to maximize response rates and minimise the number of response errors and unanswered items thus improving the validity and generalizability of any findings [208].

Furthermore, as the BAI measures symptoms experienced in the preceding seven days, it is more a measure of prolonged state anxiety rather than current state or background trait anxiety. It was felt that an assessment of prolonged state anxiety might be less sensitive to change over a relatively brief encounter than the state subscale of the STAI. There was also the concern that the BAI might be interpreted in different ways by different individuals, for example, should the highest rating on the scale be marked by the respondent even if the symptom was only bothersome very transiently or should a general state or indeed an average state be recorded? The assessment of feelings 'right now' and 'at this moment,' as in the state subscale of the STAI, was therefore considered to be less ambiguous.

Another limitation of the BAI is that it focuses primarily on the somatic symptoms of anxiety. With fifteen of the 21 items included measuring physiological symptoms, many of which overlap with some of the typical physical symptoms experienced in normal early pregnancy, for example, palpitations, dizziness, indigestion and facial flushing, scores may be increased for reasons other than anxiety in some individuals.

## Study Design

All women presenting to the Early Pregnancy Assessment Unit at the Queen's Medical Centre, Nottingham with abdominal pain and/or vaginal bleeding between 6th February 2015 and 30th April 2015 were eligible to take part in the

study. Upon arrival to the Unit, a brief history was taken using a standardized clerking proforma by a nurse specialist. A pelvic ultrasound scan, performed either transabdominally or transvaginally depending on the estimated gestational age, was then conducted by a trained ultrasonographer. Following the ultrasound, a diagnosis of a viable intrauterine pregnancy, non-viable intrauterine pregnancy, ectopic pregnancy, pregnancy of unknown location or pregnancy of uncertain viability was made. A viable intrauterine pregnancy was confirmed by the presence of an intrauterine gestation sac with a fetal pole of any length with demonstrable fetal heart pulsations. A non-viable intrauterine pregnancy was diagnosed when either there was an empty intrauterine gestation sac with mean sac diameter greater than 25mm or an intrauterine gestation sac containing a fetal pole with crown rump length greater than 7mm with no demonstrable fetal heart pulsations or, in the absence of a viable embryo, there was no significant growth of the gestation sac or fetal pole on two ultrasound scans performed more than seven days apart. Ultrasonographic appearances strongly suggestive of an ectopic pregnancy included an empty endometrial cavity with either an inhomogeneous adnexal mass or an empty extra-uterine sac or a yolk sac or fetal pole with or without cardiac activity in an extra-uterine sac. A pregnancy of unknown location was reported when, in the presence of a positive urinary pregnancy test, there was no ultrasonographic evidence of an intra- or extra-uterine pregnancy or retained products of conception. A pregnancy of uncertain viability was defined as the presence of an intrauterine gestation sac of less than 25mm mean diameter with no obvious yolk sac or fetal pole or an intrauterine gestation sac containing a fetal pole of less than 7mm with no obvious fetal heart pulsations.

Following the ultrasound, women were reviewed by a nurse specialist or a gynaecology doctor, and, if necessary, a plan for further management made according to departmental protocols: women with viable intrauterine pregnancies were reassured and discharged back to the care of their general practitioner; women with non-viable or ectopic pregnancies were given the options of conservative, medical or surgical management, which was usually carried out within the next 24-72 hours; women with pregnancies of uncertain viability were given an appointment for a repeat ultrasound scan after an interval of no less than seven days; and women with pregnancies of unknown location had their serum hCG concentrations measured at 0 and 48 hours after the scan and were then reviewed back in the Early Pregnancy

Assessment Unit with the results. At that time, if the serial serum hCG concentration had increased by at least 66%, the anticipation was of a viable intrauterine pregnancy and hence a repeat ultrasound scan was performed after an interval of at least seven days; if the serial serum hCG concentration had decreased by at least 15%, the pregnancy of unknown location was considered to be resolving and weekly serum hCG measurements were recommended until the concentration was less than 25iu/L; if the serial serum hCG concentration had decreased by less than 15% or increased by less than 66%, an ectopic pregnancy was suspected and either surgery, in the form of a laparoscopy with or without definitive treatment of an ectopic pregnancy if identified, or a repeat ultrasound scan was performed, depending on the specific combination of clinical, ultrasonographic and serological findings (Appendix 2).

Women were excluded from the study if they had attended the Unit previously in the same pregnancy; if they were unable to comprehend the questions asked of them; if they failed to answer all of the questions on all of the questionnaires: or if they were unable or unwilling to give consent.

## Timing of Assessments of Anxiety

Women were asked to complete the trait anxiety questionnaire (Figure 4.1) and the first of three STAI-SSF questionnaires (Figure 4.2) upon arrival to the Early Pregnancy Assessment Unit (State 1). The second STAI-SSF questionnaire was completed immediately after the ultrasound scan or as soon after as was considered practical (but always prior to leaving the unit) (State 2). At this point the questionnaires were returned to a member of the Early Pregnancy Assessment Unit staff and the ultrasonographic diagnosis recorded. Women were contacted by one of the investigators 48-72 hours later via telephone or e-mail using contact information provided exclusively for the purpose by the woman to complete the third and final STAI-SSF questionnaire (State 3) (Appendix 3).

Questionnaires were confidential but not completely anonymous although identifying information was kept to an absolute minimum (appointment date and time) and only used for the purpose of correlating the correct ultrasonographic diagnosis to each woman.

***Figure 4.1: Spielberger's trait anxiety inventory***

A number of statements which people have used to describe themselves are given below. Read each statement and then write the number in the blank to the right of the statement that indicates ***how you generally feel***. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you *generally* feel. Please make sure you answer all the questions.

**1 = Almost never 2 = Sometimes 3 = Often 4 = Almost always**

1. I feel pleasant ____
2. I feel nervous and restless ____
3. I feel satisfied with myself ____
4. I wish I could be as happy as others seem to be ____
5. I feel like a failure ____
6. I feel rested ____
7. I am 'cool, calm and collected' ____
8. I feel that difficulties are piling up so that I cannot overcome them ____
9. I worry too much over something that really doesn't matter ____
10. I am happy ____
11. I have disturbing thoughts ____
12. I lack self-confidence ____
13. I feel secure ____
14. I make decisions easily ____
15. I feel inadequate ____
16. I am content ____
17. Some unimportant thought runs through my mind and bothers me ____
18. I take disappointments so keenly that I can't put them out of my mind ____
19. I am a steady person ____
20. I get in a state of tension or turmoil as I think over my recent concerns and interests ____

***Figure 4.2: The standardized short form of Spielberger's state anxiety inventory***

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the number to the right of the statement that indicates ***how you feel right now***. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best. Please make sure you answer all the questions.

| | | Not at all | Somewhat | Moderately | Very much |
|---|---|---|---|---|---|
| 1. | I feel calm | 1 | 2 | 3 | 4 |
| 2. | I am tense | 1 | 2 | 3 | 4 |
| 3. | I feel upset | 1 | 2 | 3 | 4 |
| 4. | I am relaxed | 1 | 2 | 3 | 4 |
| 5. | I feel content | 1 | 2 | 3 | 4 |
| 6. | I am worried | 1 | 2 | 3 | 4 |

## Statistical Analysis

The reliability of the questionnaires was assessed using Cronbach's α. This is a measure of the internal consistency used to determine how much the items on a scale are measuring the same construct [209]. It is commonly used when there are multiple Likert questions in a survey/questionnaire that form a scale or subscale, and it is necessary to determine if the scale is reliable. It is expressed as a number between 0 and 1. There are different reports about the acceptable values of α, ranging from 0.70 to 0.95 [210-212].

The Cronbach's α and the corresponding levels of internal consistency for the trait anxiety questionnaire and the three state anxiety questionnaires are illustrated in Table 4.1. In the state 1 questionnaire, questions 1 (I feel calm), 4 (I am relaxed) and 5 (I feel content) had zero variance and were removed from the scale resulting in an unacceptable level of internal consistency. This is because α is affected by the length of the test and if this is too short, the value is reduced. Additionally, α is grounded in the 'tau equivalent model', which assumes that each test item measures the same latent trait on the same scale. If the number of test items is too small it will also violate the assumption of tau-equivalence and will underestimate reliability [213].

***Table 4.1: Internal consistency associated with each of the questionnaires***

| | **Cronbach's α** | **No. of Items** | **Internal Consistency** |
|---|---|---|---|
| Trait | 0.94 | 20 | Excellent |
| State 1 | 0.46 | 3 | Unacceptable |
| State 2 | 0.96 | 6 | Excellent |
| State 3 | 0.94 | 6 | Excellent |

To determine if trait anxiety levels were statistically significantly different amongst the different groups, a one-way ANOVA was conducted after assessment of outliers, normality and homogeneity of variances.

To determine if state anxiety levels were statistically significantly different amongst the different groups at the three different time points, a mixed ANOVA was conducted after assessment of outliers, normality, homogeneity of variances and sphericity.

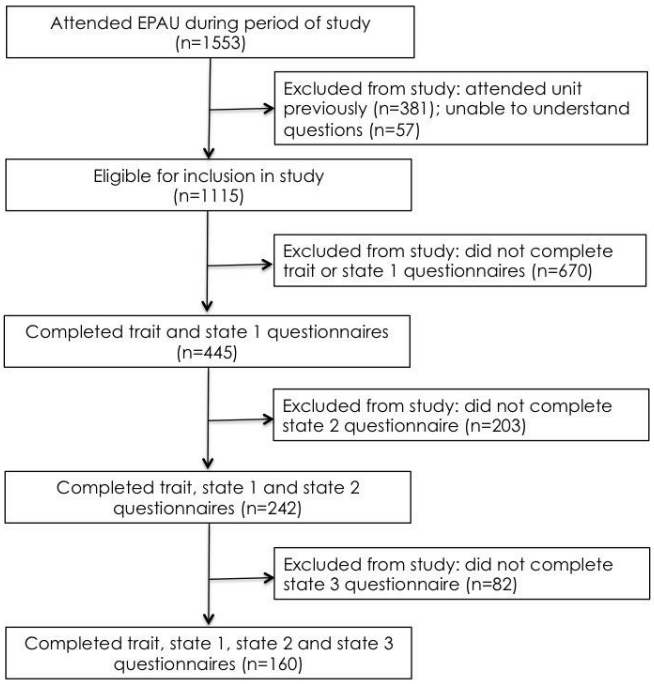
A post-hoc Tukey-Kramer test was carried out for any statistically significant findings from the ANOVAs. Data are presented as mean ± standard deviation unless otherwise indicated. p-values of $<0.05$ were considered to be statistically significant for all tests.

## 4.5 Results

Between 1st February 2015 and 30th April 2015, 1553 women attended the Early Pregnancy Assessment Unit at the Queen's Medical Centre, Nottingham. 381

(25%) women were excluded because they had visited the Unit previously during the same pregnancy and a further 57 (3.7%) were excluded because they were unable to understand the questions. Of the 1115 women who were eligible to take part in the study: 670 (60%) did not return any completed questionnaires; 203 (18%) only completed the trait and state 1 questionnaire; 82 (7.4%) only completed the trait, state 1 and state 2 questionnaires. One hundred and sixty women (14%) completed all four questionnaires and formed our study sample (Figure 4.3).

***Figure 4.3: A flow chart to demonstrate movement of participants through the different phases of the study***




Of these 160 women, 64 (40%) had a viable intrauterine pregnancy, 48 (30%) a non-viable intrauterine pregnancy, 16 (10%) an ectopic pregnancy, 13 (8.1%) a pregnancy of unknown location and 19 (12%) a pregnancy of uncertain viability diagnosed following the ultrasound. One hundred and twenty-eight (80%) women therefore had a certain diagnosis (viable intrauterine, non-viable

intrauterine or ectopic pregnancy) and 32 (20%) women had an uncertain diagnosis (pregnancy of unknown location or uncertain viability). Of those with a certain diagnosis, 64 (50%) women had a positive diagnosis (viable intrauterine pregnancy) and a further 64 women had a negative diagnosis (non-viable intrauterine or ectopic pregnancy). The relative frequencies of the different diagnoses were not significantly different amongst women that completed all four questionnaires (i.e. those that formed the final study sample) and women that only completed two (trait and state 1) or three (trait, state 1 and state 2) questionnaires (Table 4.2).

***Table 4.2: Relative frequencies of the different diagnoses according to the number of questionnaires completed***

| Diagnosis | n(%) | | | p-value$ |
|---|---|---|---|---|
| | Excluded | | Included | |
| | 2* complete (n=203) | 3¥ complete (n=82) | 4§ complete (n=160) | |
| Viable IUP | 71 (35) | 36 (44) | 64 (40) | 0.33 |
| Non-viable IUP | 71 (35) | 19 (23) | 48 (30) | 0.14 |
| EP | 17 (8.4) | 4 (4.9) | 16 (10) | 0.39 |
| PUL | 18 (8.9) | 10 (12) | 13 (8.1) | 0.63 |
| PUV | 26 (13) | 13 (16) | 19 (12) | 0.68 |
| Certain | 159 (78) | 59 (72) | 128 (80) | 0.35 |
| Uncertain | 44 (22) | 23 (28) | 32 (20) | 0.35 |
| Positive | 71 (35) | 36 (44) | 64 (40) | 0.33 |
| Negative | 88 (43) | 23 (28) | 64 (40) | 0.56 |

**Trait and State 1; ¥Trait, State 1 and State 2; §Trait, State 1, State 2 and State 3; $Chi-squared test*

Similarly, irrespective of the diagnosis, there were no significant differences in the levels of trait or state anxiety amongst those that were included in the final study sample and those that were excluded due to failure to complete all four questionnaires (Table 4.3).

## Trait Anxiety Levels in Women Attending The Early Pregnancy Assessment Unit With Abdominal Pain And/Or Vaginal Bleeding

We first examined whether there were any differences between the groups in the level of trait anxiety. This was accomplished by conducting a one-way

***Table 4.3: Anxiety levels according to diagnosis, timing of assessment and number of questionnaires completed***

| Diagnosis | Timing of Assessment | Level of Anxiety (mean±SD) | | | p-value[$] |
|---|---|---|---|---|---|
| | | Excluded | | Included | |
| | | 2* complete | 3[¥] complete | 4[§] complete | |
| Overall | Trait | 33±7.1 | 34±3.9 | 33±7.1 | 0.88 |
| | State 1 | 22±1.2 | 22±1.0 | 22±1.1 | 0.90 |
| | State 2 | | 16±7.1 | 16±6.8 | 0.72 |
| Viable IUP | Trait | 34±6.8 | 33±7.5 | 34±8.3 | 0.79 |
| | State 1 | 22±1.1 | 22±1.1 | 22±1.1 | 0.76 |
| | State 2 | | 7.7±1.3 | 7.8±1.1 | 0.91 |
| Non-viable IUP | Trait | 34±7.7 | 33±5.9 | 33±7.8 | 0.89 |
| | State 1 | 22±1.1 | 22±1.2 | 22±1.1 | 0.66 |
| | State 2 | | 20±1.0 | 20±1.1 | 0.56 |
| EP | Trait | 32±6.2 | 31±7.9 | 35±9.2 | 0.56 |
| | State 1 | 22±1.3 | 22±0.50 | 22±1.2 | 0.96 |
| | State 2 | | 22±0.50 | 23±0.7 | 0.53 |
| PUL | Trait | 32±7.8 | 35±5.3 | 36±8.7 | 0.52 |
| | State 1 | 22±1.3 | 22±1.0 | 22±1.1 | 0.53 |
| | State 2 | | 23±0.67 | 23±0.58 | 1.0 |
| PUV | Trait | 34±6.1 | 37±7.3 | 35±8.7 | 0.43 |
| | State 1 | 22±1.2 | 22±0.86 | 22±1.1 | 0.97 |
| | State 2 | | 23±0.80 | 23±0.92 | 0.85 |
| Certain | Trait | 33±7.2 | 33±7.0 | 34±8.2 | 0.89 |
| | State 1 | 22±1.1 | 22±1.1 | 22±1.1 | 0.62 |
| | State 2 | | 13±6.4 | 14±6.55 | 0.16 |
| Uncertain | Trait | 33±6.8 | 36±6.4 | 36±8.5 | 0.30 |
| | State 1 | 22±1.3 | 22±0.90 | 22±1.09 | 0.74 |
| | State 2 | | 23±0.73 | 223±0.79 | 0.85 |
| Positive | Trait | 33±7.1 | 33±7.5 | 34±8.3 | 0.79 |
| | State 1 | 22±1.2 | 22±1.1 | 22±1.1 | 0.80 |
| | State 2 | | 7.7±1.3 | 7.8±1.1 | 0.91 |
| Negative | Trait | 33±7.4 | 32±6.1 | 34±8.1 | 0.92 |
| | State 1 | 22±1.1 | 22±1.1 | 22±1.1 | 0.63 |
| | State 2 | | 20±1.3 | 21±1.5 | 0.80 |

**Trait and State 1; [¥]Trait, State 1 and State 2; [§]Trait, State 1, State 2 and State 3; [$]one-way ANOVA.*

ANOVA. There were no significant outliers, as assessed by boxplot; data was approximately normally distributed as assessed by the Shapiro-Wilk test (p>0.05) when the sample size was small and Normal Q-Q plots when the sample size was large; and there was homogeneity of variances, as assessed by Levene's test (p>0.05).

Trait anxiety levels amongst women presenting to the Early Pregnancy Assessment Unit with abdominal pain and/or vaginal bleeding were generally fairly low (33±7.1). Trait anxiety levels increased from 32±6.2 in women with non-viable intrauterine pregnancies, to 33±7.2 in women with viable intrauterine pregnancies, to 34±7.5 in women with pregnancies of uncertain viability, to 35±8.6 in women with ectopic pregnancies, to 35±8.0 in women with pregnancies of unknown location, but the differences in trait anxiety levels between the groups were not statistically significant, $F(4, 155)=0.74$, $p=0.57$.

## State Anxiety Levels According To Certainty Of Ultrasonographic Diagnosis And Timing Of Assessment

There was a statistically significant interaction between the level of certainty of the diagnosis (i.e. certain and uncertain) and time on the level of state anxiety, $F(2, 316)=102$, $p<0.0005$, partial $\eta^2=0.39$.

### Simple Main Effect for Group

Prior to the ultrasound scan (state 1), the mean level of state anxiety was 22±1.1. State anxiety levels were 22±1.1 in women with a certain and uncertain diagnosis. Immediately after the ultrasound scan (state 2), the mean level of state anxiety was 16±6.8. State anxiety levels ranged from 14±6.6 in women with a certain diagnosis to 23±0.79 in women with an uncertain diagnosis. 48-72 hours after the ultrasound scan (state 3), the mean level of state anxiety was 13±6.3. State anxiety levels ranged from 11±4.3 in women with a certain diagnosis to 23±1.7 in women with an uncertain diagnosis (Table 4.4 and Figure 4.4).

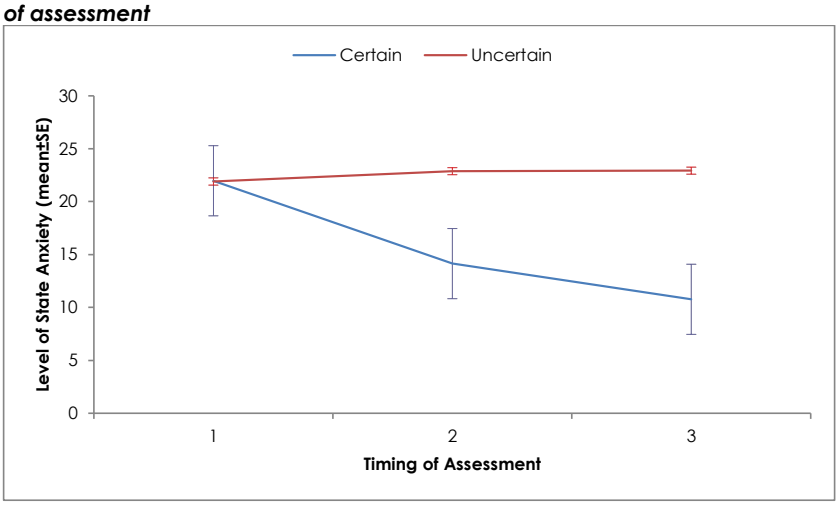
There was no statistically significant difference in the level of state anxiety between the two groups (certain and uncertain) prior to the ultrasound scan being undertaken (state 1), $F(1, 158)=0.080$, $p=0.78$, partial $\eta^2=0.001$ but there

was a statistically significant difference in the level of state anxiety between the two groups both immediately after (state 2), $F(1, 158)=56$, $p<0.005$, partial $\eta^2=0.26$ and 48-72 hours after (state 3), $F(1, 158)=246$, $p<0.005$, partial $\eta^2=0.61$ the ultrasound scan (Table 4.4 and Figure 4.4).

***Table 4.4: State anxiety levels according to certainty of diagnosis and timing of assessment***

| | State Anxiety Level (mean±SD) | | p-value (group) |
|---|---|---|---|
| | Certain (n=128) | Uncertain (n=32) | |
| **State 1** | 22±1.1 | 22±1.1 | ns |
| **State 2** | 14±6.6 | 23±0.79 | <0.005 |
| **State 3** | 11±4.3 | 23±1.7 | <0.005 |
| **p-value (time)** | <0.0005 | <0.001 | |

***Figure 4.4: State anxiety levels according to certainty of diagnosis and timing of assessment***




## Simple Main Effect for Time

In women with a certain diagnosis, state anxiety levels were highest (22±0.10) before the ultrasound scan (state 1) and lowest (11±0.35) 48-72 hours after the ultrasound scan (state 3). Immediately after the ultrasound scan (state 2), state anxiety levels were 14±0.52. In women with an uncertain diagnosis, state anxiety levels were highest (23±0.69) 48-72 hours after the ultrasound (state 3) and lowest (22±0.20) before the ultrasound scan (state 1). Immediately after

the ultrasound scan (state 2), state anxiety levels were 23±1.0) (Table 4.4 and Figure 4.4).

There was a statistically significant effect of time on the level of state anxiety for women with both a certain, $F(2, 254)=351$, $p<0.0005$, partial $\eta^2=0.73$ and uncertain, $F(2, 62)=7.3$, $p=0.001$, partial $\eta^2=0.19$ diagnosis (Table 4.4 and Figure 4.4).

For women with a certain diagnosis, the level of state anxiety was statistically significantly lower ($M$=-7.8, $SE$=0.59, $p<0.001$) immediately following the ultrasound scan (state 2) compared to prior to the ultrasound scan (state 1). Similarly, the level of state anxiety was statistically significantly lower ($M$=-3.4, $SE$=0.22, $p<0.001$) 48-72 hours following the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

For women with an uncertain diagnosis, the level of state anxiety was statistically significantly higher ($M$=0.97, $SE$=0.25, $p=0.002$) immediately following the ultrasound scan (state 2) compared to prior to the ultrasound scan (state 1). The level of state anxiety was slightly, but not statistically significantly, higher ($M$=0.063, $SE$=0.32, $p>0.05$) 48-72hrs after the ultrasound scan (state 3) than it had been immediately after the ultrasound scan (state 2).

## State Anxiety Levels According To Type Of Ultrasonographic Diagnosis And Timing Of Assessment

There was a statistically significant interaction between the type of diagnosis (i.e. positive, negative or uncertain) and time on the level of state anxiety, $F(4,314)=916$, $p<0.0001$, partial $\eta^2= 0.92$.

### Simple Main Effect for Group

Prior to the ultrasound scan (state 1), the mean level of state anxiety was 22±1.1. State anxiety levels ranged from 22±1.1 in women with an uncertain diagnosis to 22±1.1 in women with a positive diagnosis. State anxiety levels in women with a negative diagnosis were 22±1.1. Immediately after the ultrasound scan (state 2), the mean level of state anxiety was 16±6.8. At this time, state anxiety levels were lowest (7.8±1.1) in women with a positive

diagnosis and highest (23±0.79) in women with an uncertain diagnosis. State anxiety levels in women with a negative diagnosis were 21±1.5. 48-72 hours after the ultrasound scan (state 3) the mean level of state anxiety was 13±6.3. At this time, state anxiety levels ranged from 6.7±0.59 in women with a positive diagnosis to 23±1.7 in women with an uncertain diagnosis. State anxiety levels in women with a negative diagnosis were 15±1.7 (Table 4.5 and Figure 4.5).

There was no statistically significant difference in the level of state anxiety between the different types of diagnosis prior to the ultrasound scan being performed (state 1) $F(2, 157)=0.24$, $p>0.05$, partial $\eta^2=0.003$ (Table 4.5 and Figure 4.5).

Immediately after the ultrasound scan (state 2) there was a statistically significant difference in the level of state anxiety between the different types of diagnosis, $F(2, 157)=2278$, $p<0.001$, partial $\eta^2=0.97$ (Table 4.5 and Figure 4.5). At this time, state anxiety levels were significantly higher ($M=2.3$, $SE=0.27$, $p<0.001$) in women with an uncertain diagnosis compared to women with a negative diagnosis. Similarly, state anxiety levels were significantly higher ($M=13$, $SE=0.22$, $p<0.001$) in women with a negative diagnosis compared to women with a positive diagnosis.

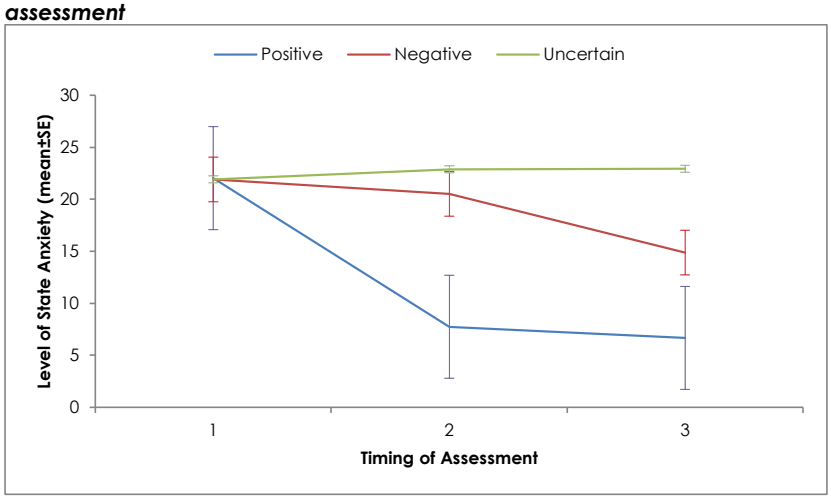
***Table 4.5: State anxiety levels according to type of diagnosis and timing of assessment***

| | **State Anxiety Level (mean±SD)** | | | **p-value (group)** |
|---|---|---|---|---|
| | **Certain** | | **Uncertain (n=32)** | |
| | **Positive (n=64)** | **Negative (n=64)** | | |
| **State 1** | 22±1.1 | 22±1.1 | 22±1.1 | ns |
| **State 2** | 7.8±1.1 | 20.53±1.53 | 23±0.79 | <0.001 |
| **State 3** | 6.7±0.59 | 15±1.7 | 23±1.7 | <0.001 |
| **p-value (time)** | <0.0005 | <0.0005 | <0.001 | |

Similarly there was a statistically significant difference in the level of state anxiety between the different types of diagnosis 48-72hrs after the ultrasound scan (state 3), $F(2, 157)=1660$, $p<0.001$, partial $\eta^2=0.96$ (Table 4.5 and Figure 4.5). At this time, state anxiety levels were significantly higher ($M=8.1$, $SE=0.29$, $p<0.001$) in women with an uncertain diagnosis compared to women with a

negative diagnosis. Similarly, state anxiety levels were significantly higher (*M*=8.2, *SE*=0.24, *p*<0.001) in women with a negative diagnosis compared to women with a positive diagnosis.

***Figure 4.5: State anxiety levels according to type of diagnosis and timing of assessment***




## Simple Main Effect for Time

In women with a positive diagnosis, state anxiety levels were highest (22±0.14) immediately before the ultrasound scan (state 1) and lowest (6.7±0.17) 48-72 hours later (state 3). State anxiety levels were 7.8±0.16 immediately after the ultrasound scan (state 2). Similarly, in women with a negative diagnosis, state anxiety levels were highest (22±0.14) immediately before the ultrasound scan (state 1) and lowest (15±0.17) 48-72 hours later (state 3). State anxiety levels were 21±0.16 immediately after the ultrasound scan (state 2). Conversely, in women with an uncertain diagnosis, state anxiety levels were lowest (22±0.20) immediately before the ultrasound scan (state 1) and highest (23±0.24) 48-72 hours later (state 3). State anxiety levels were 23±0.22 immediately after the ultrasound scan (state 2) (Table 4.5 and Figure 4.5).

There was a statistically significant effect of time on state anxiety levels for women with a positive diagnosis, $F(2, 126)=6084$, $p<0.0005$, partial $\eta^2=0.99$ (Table 4.5 and Figure 4.5). In women with a positive diagnosis, state anxiety levels were statistically significantly lower (*M*=-14, *SE*=0.17, *p*<0.001) immediately after the ultrasound scan (state 2) compared to state anxiety

levels prior to the ultrasound scan (state 1). State anxiety levels were also statistically significantly lower (*M*=-1.1, *SE*=0.14, *p*<0.001) 48-72 hours after the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

There was a statistically significant effect of time on state anxiety levels for women with a negative diagnosis, *F*(2, 126)=508, *p*<0.0005, partial $\eta^2$=0.89 (Table 4.5 and Figure 4.5). In women with a negative diagnosis, state anxiety levels were statistically significantly lower (*M*=-1.4, *SE*=0.26, *p*<0.001) immediately after the ultrasound scan (state 2) compared to state anxiety levels prior to the ultrasound scan (state 1). State anxiety levels were also statistically significantly lower (*M*=-5.7, *SE*=0.12, *p*<0.001) 48-72 hours after the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

There was a statistically significant effect of time on state anxiety levels for women with an uncertain diagnosis, *F*(2, 62)=7.3, *p*=0.001, partial $\eta^2$=0.19 (Table 4.5 and Figure 4.5). In women with an uncertain diagnosis, state anxiety levels were statistically significantly higher (*M*=0.97, *SE*=0.25, *p*=0.002) immediately after (state 2) the ultrasound scan compared to state anxiety levels prior to the ultrasound scan (state 1). State anxiety levels were slightly, but not statistically significantly, higher (*M*=0.063, *SE*=0.32, *p*>0.05) 48-72 hours after the ultrasound scan (state 3) than they were immediately after the ultrasound scan (state 2).

## State Anxiety Levels According To Specific Ultrasonographic Diagnosis And Timing Of Assessment

There was a statistically significant interaction between the specific ultrasonographic diagnosis (viable IUP, non-viable IUP, EP, PUL and PUV) and time on the level of state anxiety, *F*(8,310)=605, p<0.0005, partial $\eta^2$ = 0.94.

### Simple Main Effect For Group

Prior to the ultrasound scan (state 1), the mean level of state anxiety was 22±1.1. State anxiety levels ranged from 22±1.1 in women with a pregnancy of uncertain viability to 22±1.1 in women with a viable intrauterine pregnancy. State anxiety levels in women with a non-viable intrauterine pregnancy,

ectopic pregnancy and pregnancy of unknown location were 21±1.1, 22±1.1 and 22±1.1 respectively. Immediately after the ultrasound scan (state 2), the mean level of state anxiety was 16±6.8. At this time, state anxiety levels were lowest (7.8±1.1) in women with a viable intrauterine pregnancy and highest (23±0.58) in women with a pregnancy of unknown location. State anxiety levels in women with a non-viable intrauterine pregnancy were 20±1.1, ectopic pregnancy 23±0.73 and pregnancy of uncertain viability 23±0.92. 48-72 hours after the ultrasound scan (state 3), the mean level of state anxiety was 13±6.3. At this time state anxiety levels ranged from 6.7±0.59 in women with a viable intrauterine pregnancy to 24±0.32 in women with a pregnancy of uncertain viability. State anxiety levels in women with a non-viable intrauterine pregnancy, ectopic pregnancy and pregnancy of unknown location were 14±1.072, 17±1.1 and 22±1.8 respectively (Table 4.6 and Figure 4.6).

Prior to the ultrasound scan (state 1) there was no statistically significant difference in the level of state anxiety between the specific diagnoses, $F(4,155)=0.16$, $p=0.96$, partial $\eta^2=0.004$ (Table 4.6 and Figure 4.6).

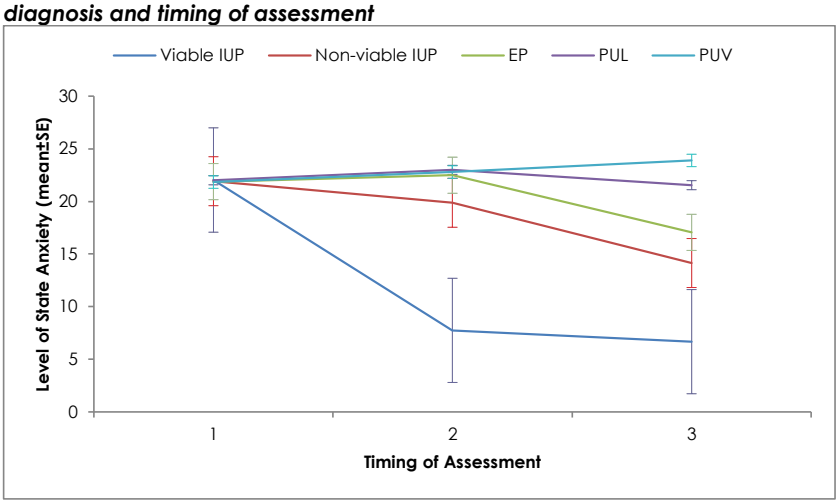
***Table 4.6: State anxiety levels according to specific ultrasonographic diagnosis and timing of assessment***

| | State Anxiety Level (mean±SD) | | | | | |
|---|---|---|---|---|---|---|
| | Certain | | | Uncertain | | p-value (group) |
| | Viable IUP (n=64) | Non-viable IUP (n=48) | EP (n=16) | PUL (n=13) | PUV (n=19) | |
| **State 1** | 22±1.1 | 22±1.1 | 22±1.1 | 22±1.1 | 22±1.1 | ns |
| **State 2** | 7.8±1.1 | 20±1.1 | 23±0.73 | 23±0.58 | 23±0.92 | <0.005 |
| **State 3** | 6.7±0.59 | 14±1.1 | 17±1.1 | 22±1.8 | 24±0.32 | <0.005 |
| **p-value (time)** | <0.005 | <0.005 | <0.005 | <0.05 | <0.005 | |

There was a statistically significant difference however in the level of state anxiety between the specific diagnoses immediately after the ultrasound scan (state 2), $F(4,155)=1712$, $p<0.005$, partial $\eta^2=0.98$ (Table 4.4 and Figure 4.4). At this time, state anxiety levels were significantly higher ($M=12$, $SE=0.20$, $p<0.005$) in women with a non-viable intrauterine pregnancy compared to women with a viable intrauterine pregnancy. State anxiety levels were also significantly higher ($M=2.6$, $SE=0.30$, $p<0.005$) in women with ectopic pregnancy compared to women with a non-viable intrauterine pregnancy. State anxiety levels were

slightly (but not statistically significantly) higher (*M*=0.29, *SE*=0.35, *p*=0.92) in women with a pregnancy of uncertain viability compared to women with an ectopic pregnancy. Similarly, state anxiety levels were slightly (but not statistically significantly) higher (*M*=0.21, *SE*=0.37, *p*=0.98) in women with a pregnancy of unknown location compared to women with a pregnancy of uncertain viability.

***Figure 4.6: State anxiety levels according to specific ultrasonographic diagnosis and timing of assessment***




48-72 hours after the ultrasound scan (state 3) there was a statistically significant difference in the level of state anxiety between the five different specific diagnoses, $F(4,155)=1734$, $p<0.005$, partial $\eta^2 = 0.98$ (Table 4.4 and Figure 4.4). At this time, state anxiety levels were significantly higher (*M*=7.5, *SE*=0.18, *p*<0.005) in women with a non-viable intrauterine pregnancy compared to women with a viable intrauterine pregnancy. Similarly, state anxiety levels were significantly higher (*M*=2.9, *SE*=0.27, *p*<0.005) in women with an ectopic pregnancy compared to women with a non-viable intrauterine pregnancy. Unlike immediately after the ultrasound scan (state 2), 48-72 hours after the ultrasound scan (state 3), state anxiety levels were significantly higher (*M*=4.5, *SE*=0.35, *p*<0.005) in women with a pregnancy of unknown location compared to women with an ectopic pregnancy and furthermore, those with a pregnancy of uncertain viability were significantly more anxious (*M*=2.4, *SE*=0.34, *p*<0.005) than those with a pregnancy of unknown location.

## Simple Main Effect For Time

In women with a viable intrauterine pregnancy, state anxiety levels were highest (22±1.1) prior to the ultrasound scan (state 1) and lowest (6.7±0.59) 48-72 hours after the ultrasound (state 3). State anxiety levels were 7.8±1.1 immediately after the ultrasound scan (state 2). Similarly, in women with a non-viable intrauterine pregnancy, state anxiety levels were highest (21.92±1.13) prior to the ultrasound scan (state 1) and lowest (14±1.1) 48-72 hours after the ultrasound (state 3). State anxiety levels were 20±1.1 immediately after the ultrasound scan (state 2). In women with an ectopic pregnancy, state anxiety levels were highest (23±0.73) immediately after the ultrasound scan (state 2) and lowest (17±1.1) 48-72 hours after the ultrasound scan (state 3). State anxiety levels were 22±1.2 before the ultrasound scan (state 1). Similarly in women with a pregnancy of unknown location, state anxiety levels were highest (23±0.58) immediately after the ultrasound scan (state 2) and lowest (22±1.8) 48-72 hours after the ultrasound scan (state 3). State anxiety levels were 22±1.1 before the ultrasound scan (state 1). In women with a pregnancy of uncertain viability, state anxiety levels were lowest (22±1.1) before the ultrasound scan (state 1) and highest (24±0.32) 48-72 hours after the ultrasound scan (state 3). State anxiety levels were 23±0.92 immediately after the ultrasound scan (state 2) (Table 4.6 and Figure 4.6).

There was a statistically significant effect of time on state anxiety levels for women with a viable intrauterine pregnancy, $F(2, 126)=6084$, $p<0.0005$, partial $\eta^2=0.990$ (Table 4.4 and Figure 4.4). In women with a viable intrauterine pregnancy, state anxiety levels were significantly lower ($M=-14$, $SE = 0.17$, $p<0.001$) immediately after the ultrasound scan (state 2) compared to state anxiety levels prior to the ultrasound scan (state 1). State anxiety levels were also statistically significantly lower ($M=-1.1$, $SE=0.14$, $p<0.001$) 48-72 hours after the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

There was a statistically significant effect of time on state anxiety levels for women with a non-viable intrauterine pregnancy, $F(2,94)=604$, $p<0.005$, partial $\eta^2=0.93$ (Table 4.4 and Figure 4.4). In women with a non-viable intrauterine pregnancy, state anxiety levels were significantly lower ($M=-2.0$, $SE=0.26$, $p<0.005$) immediately after the ultrasound scan (state 2) compared to state

anxiety levels prior to the ultrasound scan (state 1). State anxiety levels were also significantly lower (*M*=-5.7, *SE*=0.13, *p*<0.005) 48-72 hours after the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

There was a statistically significant effect of time on state anxiety levels for women with an ectopic pregnancy, *F*(2,30)=120, *p*<0.005, partial $\eta^2$=0.89 (Table 4.4 and Figure 4.4). In women with an ectopic pregnancy, state anxiety levels were not statistically significantly different (*M*=0.63, *SE*=0.38, *p*<0.35) immediately following the ultrasound scan (state 2) compared to before the ultrasound scan (state 1) but 48-72 hours after the ultrasound scan (state 3), state anxiety levels were statistically significantly lower (*M*=-5.4, *SE*=0.29, *p*<0.005) than immediately after the ultrasound scan (state 2).

There was a statistically significant effect of time on the level of state anxiety for women with a pregnancy of unknown location, *F*(2,24)=5.5, *p*<0.05, partial $\eta^2$=0.32 (Table 4.4 and Figure 4.4). In women with a pregnancy of unknown location, state anxiety levels were slightly, (but not statistically significantly), higher (*M*=1.0, *SE*=0.38, *p*=0.062) immediately following the ultrasound scan (state 2) compared to before the ultrasound scan (state 1). 48-72 hours after the ultrasound scan (state 3), state anxiety levels were statistically significantly lower (*M*=-1.5, *SE*=0.46, *p*<0.05) than they had been immediately after the ultrasound scan (state 2). State anxiety levels were not however statistically significantly different (*M*=-0.46, *SE*=0.50, *p*=1.0) 48-72 hours after the ultrasound scan (state 3) than they had been before it (state 1).

There was a statistically significant effect of time on the level of state anxiety for women with a pregnancy of uncertain viability, *F*(2,36)=27, *p*<0.005, partial $\eta^2$=0.60 (Table 4.4 and Figure 4.4). In women with a pregnancy of uncertain viability, state anxiety levels were statistically significantly higher (*M*=0.95, *SE*=0.35, *p*<0.05) immediately following the ultrasound scan (state 2) compared to before the ultrasound scan (state 1). Similarly, the level of state anxiety was statistically significantly higher (*M*=1.1, *SE*=0.22, *p*<0.005) 48-72 hours after the ultrasound scan (state 3) compared to immediately after the ultrasound scan (state 2).

## 4.6 Discussion

Our results demonstrate that the propensity for anxiety, as measured by Spielberger's trait anxiety subscale, in women presenting to Early Pregnancy Assessment Units with abdominal pain and/or vaginal bleeding, was fairly low, reflected by a mean score on the trait anxiety inventory of 33.28±7.13 (the maximum score being 80). Of interest however is that despite this general tendency towards low levels of anxiety, very high levels of state anxiety (21.96±1.11) were reported by women in all groups prior to the ultrasound (the maximum score being 24). All women scored 20 or more at this time indicating that the experience of abdominal pain and/or vaginal bleeding in early pregnancy is highly anxiogenic. For women with certain diagnoses following the ultrasound, whether associated with positive or negative connotations, anxiety levels significantly decreased but for women with uncertain diagnoses, especially a pregnancy of uncertain viability, anxiety levels increased. Hence it appears to be the certainty of diagnosis that affects anxiety levels rather than the positive or negative connotations associated with it *per se*. Accordingly we can accept all of the aforementioned hypotheses.

With regards to the viable intrauterine pregnancy group, it is unsurprising that anxiety levels significantly decreased immediately following the ultrasound scan and continued to do so with time. It is important to note that the minimum score on the STAI-SSF is six (not zero), hence in our study, most women who presented to the Early Pregnancy Assessment Unit with abdominal pain and/or vaginal bleeding but who were subsequently diagnosed with a viable intrauterine pregnancy report feeling only very slightly anxious 48-72 hours after the ultrasound scan (6.67±0.59).

It is interesting that in the non-viable intrauterine pregnancy group, state anxiety levels also decreased following the ultrasound. However, Spielberger's STAI only measures anxiety and not other constructs. The high Cronbach $\alpha$ coefficient reflects this. The fact that these women have been given a certain diagnosis, albeit a negative one, appears to have abated their initial anxiety, which has perhaps been replaced by other feelings, for example grief, sadness and depression, which are well-documented amongst women who have miscarried [186-191].

Given the threats to health and future fertility, it is unsurprising that state anxiety levels in the ectopic pregnancy group increased following the ultrasound. What is surprising is that this increase in anxiety was not significant. This may be because there were only sixteen women in this group and their anxiety levels were already high, or it may be indicative of the general naivety of the population towards the diagnosis, or a reflection of the communication skills of those imparting the diagnosis. Most ectopic pregnancies are treated within 48-72 hours of diagnosis, which explains the significant decrease in anxiety levels observed at this time. Again, it is worth remembering that the STAI does not measure constructs such as sadness and depression. There is very little evidence regarding the psychological sequelae following an ectopic pregnancy [193], which, for many women is considered in the same light as a miscarriage. It is, after all, a pregnancy that will never result in a child.

In the pregnancy of unknown location group, anxiety levels were shown to increase slightly following the ultrasound scan. Although this increase was not found to be statistically significant, this can perhaps be explained by the fact that anxiety levels before the scan were already high (22.00±1.08) and the maximum score obtainable on the scale is 24. Anxiety levels then decreased significantly over the next 48-72 hours, most likely because serial serum βhCGs were taken, the results of which may have provided clarification on the situation one way or another. However, whilst anxiety levels significantly decreased, they were still high (21.54±1.81).

In women with a pregnancy of uncertain viability, there was a significant increase in anxiety levels immediately after the ultrasound and again 48-72 hours after. In the UK, pregnancy of uncertain viability management protocols advocate a repeat ultrasound 7-10 days after the initial scan [138]. Since no other investigations were performed, no additional information was available to alter anxiety levels. Even if further investigations are not definitive, there is evidence to suggest that women would benefit psychologically from tests that give them an indication of what a subsequent ultrasound might show [168].

It is extremely important that the psychological wellbeing of women undergoing investigation for abdominal pain and/or vaginal bleeding in early pregnancy is not overlooked. Our study has demonstrated that the experience of these symptoms alone is highly anxiogenic and that for some

women, particularly those given uncertain diagnoses, anxiety levels remain elevated for at least 48-72 hours. This is concerning since women who feel anxious in the first trimester of pregnancy have a much higher risk of miscarriage than those who do not [184].

This is the first study to assess anxiety levels of women presenting to Early Pregnancy Assessment Units with abdominal pain and/or vaginal bleeding and to determine how this changes over time with different types of diagnoses. The study was prospective, included a considerable number of women and utilized a widely adopted, validated and reliable measure of anxiety.

Although only 14.3% of eligible women were sampled, the proportions of the different diagnoses were representative of the population hence there should not be any sampling bias. Additionally, due to the relative rarity of diagnosis, the numbers of women in the ectopic pregnancy, pregnancy of unknown location and pregnancy of uncertain viability groups were small. However, since the statistical analyses took this into consideration and the results were extremely significant reflected by p-values of <0.005, this should not affect the validity of our results or conclusions drawn.
Whilst we collected very little specific demographic and clinical data, all participants were of reproductive age and presented with abdominal pain and/or vaginal bleeding in early pregnancy. Although perhaps interesting to collect more data, our remit was to determine *how* anxious women were, not *why* they were this anxious. Furthermore, we wanted to make the questionnaires as brief as possible to encourage participation.

A further weakness of our study is that follow-up was only for 48-72 hours. This is because we wanted to focus on the impact of the diagnosis itself on anxiety levels and felt that with longer follow-up other factors might come into play for example, anxiety about further management and future reproductive performance. Longer follow-up would however enable us to assess how anxiety levels in women diagnosed with a viable, non-viable or ectopic pregnancy after a period of uncertainty compare to those given a certain diagnosis at the outset – does an initial uncertain diagnosis reduce the psychological burden of a subsequent negative diagnosis or does it cause women to remain anxious for the rest of an ongoing pregnancy?

## 4.7 Conclusion

In conclusion, this study has proven that women who present to Early Pregnancy Assessment Units with abdominal pain and/or vaginal bleeding in early pregnancy and who are subsequently given an uncertain diagnosis have significantly higher levels of anxiety than their counterparts who are given certain diagnoses, even if those certain diagnoses are not associated with an ongoing pregnancy. Healthcare providers should be aware of this when communicating uncertain diagnoses. Women with non-viable intrauterine pregnancies, and to a lesser extent those with ectopic pregnancies, have access to different support groups. Women with uncertain diagnoses have no such psychological support and this must be addressed if we are to improve the holistic nature of care provided to women with complications of early pregnancy. Further research should focus on reducing the number of women given uncertain diagnoses in early pregnancy and/or minimising the duration of uncertainty so that we can: reduce anxiety levels for women and their partners; redistribute valuable and limited NHS resources; decrease the number of women presenting to hospital in a state of haemodynamic compromise following delayed diagnosis of early pregnancy complications; and enable women to choose more conservative forms of management for non-viable intrauterine and ectopic pregnancies if they wish.

# 5. Accuracy Of First Trimester Ultrasound In The Diagnosis Of An Intrauterine Pregnancy Prior To Visualization Of The Yolk Sac

*NB; An abridged version of this chapter has been published in the journal Ultrasound in Obstetrics and Gynaecology [214]*

## 5.1 Introduction

Having established in the previous chapter that uncertain diagnoses in early pregnancy generate considerable anxiety for women, we can now move on to attempting to minimise the number of women given uncertain diagnoses in early pregnancy, which is the focus of the remainder of this thesis. As discussed in chapter one, the earliest reliable ultrasonographic sign of an intrauterine pregnancy is visualization of the gestation sac. This is first seen as a uniformly round, hypoechoic ring-like structure with an echogenic rim. It can be identified using transvaginal ultrasound from 28 days' gestation [12]. Initially, the structure does not contain any internal echoes and at this stage can be difficult to differentiate from a pseudosac, that is, an intrauterine fluid collection, that occurs in up to 15% of ectopic pregnancies [215].

The yolk sac is the first structure to appear within the gestation sac, and indicates an intrauterine pregnancy with a positive predictive value of 100% [216]. Identification of the yolk sac therefore excludes the possibility of an ectopic pregnancy in most cases with the very rare exception of a heterotopic pregnancy. The yolk sac is first visualized with transvaginal ultrasound from around 35 days gestation [19] appearing as a spherical hyperechoic ring situated eccentrically within the gestation sac. It increases in size to a maximum diameter of 6mm at ten weeks gestation and then begins to regress, usually disappearing completely by twelve weeks gestation [217].

Many health care professionals opt to wait until the yolk sac is visualized before confirming the presence of a true gestation sac. This may improve the accuracy of ultrasound for the detection of an intrauterine pregnancy, but there is a distinct interval of at least seven days during which a gestation sac may be visible but a yolk sac may not. The potential diagnoses that could be made during this interval are an early intrauterine pregnancy, which could be viable or non-viable, or an ectopic pregnancy, whereby the 'sac' visualized is actually a pseudosac. Prompt differentiation between these two would be preferable, as it would minimise the level of anxiety for women, prevent unnecessary investigations for those with intrauterine pregnancies and permit earlier, potentially less invasive intervention for women with ectopic pregnancies.

Several different ultrasonographic signs have been proposed to help differentiate a true gestation sac from a pseudosac prior to development of the yolk sac including the intradecidual [218], double decidual sac [219] and chorionic rim [220] signs. The intradecidual sign consists of a well-defined endometrial stripe with an echogenic area eccentrically embedded into the thickened decidua on one side of the uterine cavity [218]. This differs from that of a pseudosac, which appears as fluid, surrounded by the echogenic endometrial lining only. The double decidual sac sign appears as an intra-endometrial fluid collection with two surrounding concentric echogenic rings that impress upon the endometrial stripe in a normal early pregnancy [219]. In an ectopic pregnancy, the decidual reaction consists of only a single ring around the fluid collection [19]. The chorionic rim sign consists of a curvilinear echogenic rim separate from the underlying decidua bordering an outwardly convex fluid collection [220]. In clinical practice however, due to varying degrees of accuracy reported in individual studies, none of these signs are relied upon to confirm pregnancy location.

## 5.2 Aims

We therefore undertook a systematic review of the literature and meta-analysis to determine the accuracy of these first trimester ultrasonographic signs in the diagnosis of an intrauterine pregnancy prior to the appearance of a yolk sac, in women with or without symptoms of abdominal pain and/or vaginal bleeding in early pregnancy.

# 5.3 Methods

## Protocol and Registration

Search methods, criteria for inclusion and outcomes were specified in advance and documented in the protocol, which was registered with PROSPERO (http://www.crd.york.ac.uk/PROSPERO) on 4th October 2012. The registration number is CRD42012003046. Prospero is an international database of prospectively registered systematic reviews in health and social care, welfare, public health, education, crime, justice and international development, where there is a health related outcome. Key features from the review protocol are recorded and maintained as a permanent record. PROSPERO aims to provide a comprehensive listing of systematic reviews registered at inception to help avoid duplication and reduce opportunity for reporting bias by enabling comparison of the complete review with what was planned in the protocol. PROSPERO is produced by the Centre for Reviews and Dissemination and funded by the National Institute for Health Research.

Transparent and complete reporting of the systematic review and meta-analysis was ensured by adhering to the recommendations set out in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [221]. PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses. As with all research, the value of a systematic review depends on what was done, what was found and the clarity of reporting. As with other publications, the reporting quality of systematic reviews varies, limiting readers' ability to assess the strengths and weaknesses of those reviews. Several studies have evaluated the quality of systematic review reports and found them lacking [222, 223].

In 1999, to address the suboptimal reporting of meta-analyses, an international group developed a guidance called the QUOROM Statement (Quality Of Reporting Of Meta-analyses), which focused on the reporting of meta-analyses of randomized controlled trials [224]. In 2009, the guideline was updated to address several conceptual and practical advances in the science of systematic reviews, and was renamed PRISMA. The PRISMA Statement consists of a checklist, which contains 27 items, which pertain to the

content of a systematic review and meta-analysis including the title, abstract, methods, results, discussion and funding, and a flow diagram, which depicts the flow of information through the different phases of a systematic review. It maps out the number of records identified, included and excluded and the reasons for exclusions (Appendix 11).

## Information Sources

The following databases were searched electronically for relevant citations: MEDLINE (1951-March week 3 2013), Embase (1980-2013 week 11) and the Cochrane Library (2013). We used a combination of text words, Medical Subject and Emree Headings to generate two subsets of citations, one indexing ultrasound ('ultraso$' OR 'sonograph$') and the other indexing terms related to early pregnancy location or viability ('ectopic pregnancy' OR 'tubal pregnancy' OR 'viab$ pregnancy' OR 'failing pregnancy' OR 'miscarr$' OR 'abort$' OR 'intrauterine pregnancy') or ultrasonographic signs of either an intrauterine pregnancy ('gestation$ sac' OR 'yolk sac' OR 'f$etal pole' OR 'intradecidual sign' OR 'double decidual sac sign' OR 'double decidual sac' OR 'double decidual sign' OR 'chorionic rim sign' OR 'chorionic rim') or an ectopic pregnancy ('empty uterus' OR 'pseudosac' 'free fluid' OR 'cul de sac fluid' OR 'adnexal mass' OR 'tubal ring' OR 'donut sign' OR 'doughnut sign'). These two subsets were then combined with 'AND' to generate a subset of citations relevant to the research question. The search was limited to human subjects and the English language. Duplicates were removed during the process of assessing the full text articles for eligibility. The search was last run on 3rd July 2014. Further relevant papers were searched by examination of the reference lists of all included studies, reviews and other previously identified papers. A comprehensive database of relevant articles was constructed.

## Study Selection

Primary studies that reported original data regarding the ultrasonographic diagnosis of either an intrauterine (viable or non-viable) pregnancy were included. Case reports and case studies where the sample size was less than 10 were excluded due to the high risk of bias. Commentaries, narrative

reviews and letters were also excluded. There was no limitation on publication date or publication status.

Studies were selected in a two-stage process. Firstly, two reviewers independently examined the titles and abstracts of all of the citations produced by the electronic searches. The full manuscripts of citations that met the predefined selection criteria were then obtained. Secondly, examination of the full manuscripts led to a final decision regarding inclusion or exclusion. In cases of duplicates, the most recent version was selected. Any disagreements concerning selection were resolved by consensus or arbitration by a third reviewer.

## Data Collection Process

Two review authors independently extracted the data from included studies using a data extraction form designed and pilot-tested by the authors. One reviewer independently checked the extracted data. If there were data queries the corresponding author of the study was contacted. Disagreements were resolved by consensus. The names of article authors and titles of the included studies were juxtaposed to identify duplicate publications; in case of duplicates both articles were considered as a unique study.

## Data Items

The following data were extracted from included studies using a standardized data extraction form, designed and pilot-tested by the authors: study characteristics (first author, year of publication, population, age group, inclusion and exclusion criteria); study methodology (study design, study period, recruitment method); details of the intervention (ultrasound approach i.e. transabdominal or transvaginal, frequency/resolution of ultrasound machine, operator; ultrasonographic feature under evaluation i.e. intradecidual sign, double decidual sac sign, chorionic rim sign, gestation sac, yolk sac); outcome investigated (intrauterine pregnancy, ectopic pregnancy) and the quality and accuracy of the results. Accuracy data were used to construct 2x2 tables of ultrasonographic findings and pregnancy location.

## Risk of Bias in Individual Studies

One reviewer completed the quality assessment using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 checklist (Appendix 12) [225]. This checklist is designed to assess the quality of primary diagnostic accuracy studies and consists of four key domains covering patient selection, the index test, the reference standard and flow of patients through the study and timing of the index test(s) and reference standard. The tool is completed in four phases: 1) state the review question; 2) develop review specific guidance; 3) review the published flow diagram for the primary study or construct a flow diagram if none is reported; 4) judgement of bias and applicability. Each domain is assessed in terms of the risk of bias and the first three are also assessed in terms of concerns regarding applicability. To help reach a judgement on the risk of bias, signaling questions are included. These flag aspects of study design related to the potential for bias and aim to help reviewers make risk of bias judgements.

## Summary Measures

All data were inserted into the Review Manager 5.1 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark, 2011) for producing summary tables. Accuracy measures of the various different ultrasonographic signs were calculated, including sensitivity, specificity and likelihood ratios. When there were more than three studies reporting on the ultrasonographic sign, meta-analysis was performed. Individual study estimates of sensitivities and specificities were plotted in summary receiver operating characteristic (ROC) space and forest plots for visual examination of heterogeneity. We used the statistical package STATA version 12 (College Station, TX, USA) to meta-analyze the sensitivity and specificity from each included study using the hierarchical summary ROC (HSROC) approach [226, 227]. This approach estimates the position and shape of the summary ROC curve and takes into account both within and between study variations. The summary ROC curve includes the pairs of sensitivity and specificity for individual studies showing the differences in precision between them and the overall sensitivity and specificity for the test when all studies are pooled together. When all the parameters of the HSROC model could not be

estimated due to limited number of studies, it was simplified by assuming a symmetrical shape for the summary ROC curve. When only one study was available we calculated the sensitivities, specificities, 95% confidence intervals, likelihood ratios and pre-test with post-test probabilities for that study. Post-test probabilities were calculated using the summary likelihood ratios and the median prevalence values with their ranges as the pre-test probabilities.

### Risk of Bias Across Studies

The potential impact of publication and reporting bias were minimised by performing a comprehensive search for eligible studies and by looking for duplication of data.

## 5.4 Results

### Study Selection

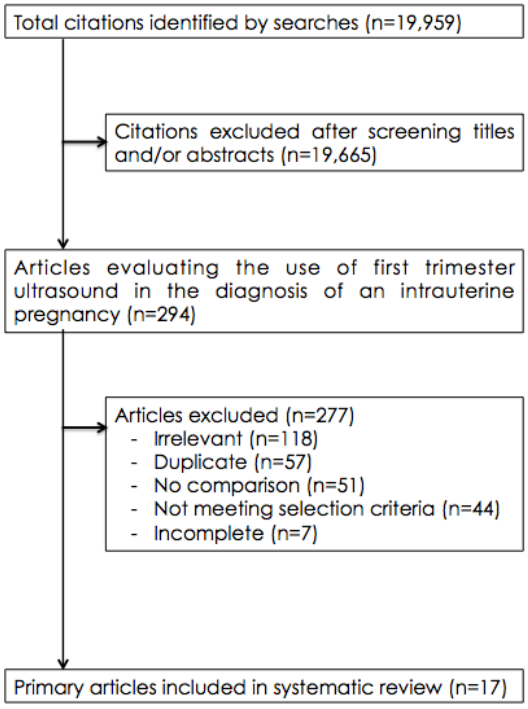
The search identified 19,959 potential papers. Following review of the titles and abstracts, 294 full-text papers were selected for further examination and subsequently 277 of these studies were excluded (Figure 5.1). Seventeen studies [218-220, 228-241], including 2564 women, met the inclusion criteria and were incorporated into the systematic review. The characteristics of the included studies are shown in Table 5.1.

### Diagnostic Accuracy Of The Gestation Sac For Predicting An Intrauterine Pregnancy

Twelve cohort studies [228, 229, 231-240], including 1920 women in early pregnancy, evaluated the diagnostic accuracy of visualization of a gestation sac on ultrasound examination to predict the likelihood of an intrauterine pregnancy. Figure 5.2 shows the sensitivities and specificities of the presence of a gestation sac for predicting an intrauterine pregnancy in the individual studies. The precision estimates for each of the studies and the estimated summary sensitivity and specificity for differentiating between an intrauterine and an extra-uterine pregnancy are shown in Figure 5.3 and Table 5.2. Following meta-analysis of these twelve studies we found that the presence of

a gestation sac predicts an intrauterine pregnancy with a pooled sensitivity of 53% (95% CI, 38–67%), specificity of 98% (95% CI, 94–99%), positive likelihood ratio of 22 (95% CI, 9.8–51) and negative likelihood ratio of 0.48 (95% CI, 0.36–0.66). Of the studies included in this meta-analysis, the median prevalence of an intrauterine pregnancy was 68% (range, 30–88%), however if the gestational sac was present the probability of an intrauterine pregnancy was as high as 98% (range, 96–99%) compared with 50% (range, 48–52%) if the gestational sac was absent.

***Figure 5.1: Flow chart summarizing study selection of papers on first-trimester ultrasound signs in the diagnosis of intrauterine pregnancy prior to visualization of the yolk sac***



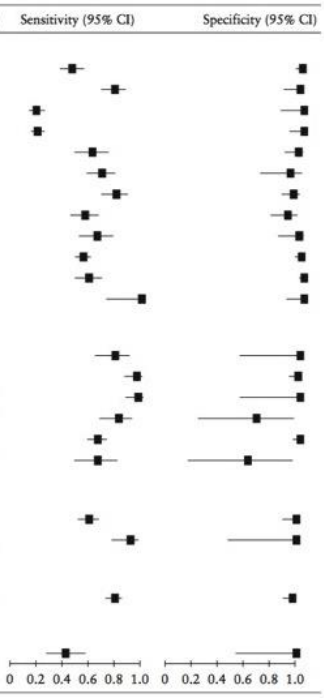

## Diagnostic Accuracy Of The Double Decidual Sac Sign For Predicting An Intrauterine Pregnancy

Six cohort studies [218-220, 238, 239, 241], including 571 women in early pregnancy, evaluated the diagnostic accuracy of the double decidual sac sign for predicting the likelihood of an intrauterine pregnancy. Figure 5.2 shows the sensitivities and specificities of the double decidual sac sign to predict an intrauterine pregnancy in the individual studies. The precision estimates for

each of the studies and the estimated summary sensitivity and specificity for differentiating between an intrauterine and an extra-uterine pregnancy are shown in Figure 5.3 and Table 5.2. Following meta-analysis of these six studies we found that the presence of the double decidual sac sign predicts an intrauterine pregnancy with a pooled sensitivity of 82% (95% CI, 68-90%), specificity of 97% (95% CI, 76-100%), positive likelihood ratio of 30 (95% CI, 2.8-331) and negative likelihood ratio of 0.19 (95% CI, 0.10-0.35). In the studies included in this meta-analysis the median prevalence of an intrauterine pregnancy was 89% (range, 49-91%), but if the double decidual sac sign was present the probability of an intrauterine pregnancy was as high as 100% (range, 97-100%) compared with 61% (range, 15%-64%) probability if the double decidual sac sign was absent.

***Figure 5.2: Forest plots for the performance of each ultrasonographic sign for predicting an intrauterine pregnancy***

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| *Gestational sac* | | | | | | |
| Ankum (1993) | 49 | 1 | 70 | 88 | 0.41 (0.32, 0.51) | 0.99 (0.94, 1.00) |
| Bateman (1990) | 68 | 1 | 24 | 33 | 0.74 (0.64, 0.83) | 0.97 (0.85, 1.00) |
| Dart (1997) | 20 | 0 | 124 | 19 | 0.14 (0.09, 0.21) | 1.00 (0.82, 1.00) |
| Dart (1998) | 29 | 0 | 167 | 32 | 0.15 (0.10, 0.21) | 1.00 (0.89, 1.00) |
| Enk (1990) | 34 | 2 | 26 | 45 | 0.57 (0.43, 0.69) | 0.96 (0.85, 0.99) |
| Kadar (1981) | 50 | 2 | 28 | 17 | 0.64 (0.52, 0.75) | 0.89 (0.67, 0.99) |
| Nyberg (1987) | 57 | 6 | 19 | 68 | 0.75 (0.64, 0.84) | 0.92 (0.83, 0.97) |
| Nyberg (1988) | 45 | 6 | 43 | 42 | 0.51 (0.40, 0.62) | 0.88 (0.75, 0.95) |
| Nyberg (1988) | 35 | 1 | 23 | 25 | 0.60 (0.47, 0.73) | 0.96 (0.80, 1.00) |
| Romero (1985) | 139 | 2 | 140 | 102 | 0.50 (0.44, 0.56) | 0.98 (0.93, 1.00) |
| Tongsong (1993) | 52 | 0 | 44 | 105 | 0.54 (0.44, 0.64) | 1.00 (0.97, 1.00) |
| Weckstein (1985) | 11 | 0 | 0 | 26 | 1.00 (0.72, 1.00) | 1.00 (0.87, 1.00) |
| *Double decidual sac sign* | | | | | | |
| Bradley (1982) | 34 | 0 | 10 | 6 | 0.77 (0.62, 0.89) | 1.00 (0.54, 1.00) |
| Nyberg (1983) | 59 | 1 | 4 | 64 | 0.94 (0.85, 0.98) | 0.98 (0.92, 1.00) |
| Nyberg (1987) | 54 | 0 | 3 | 6 | 0.95 (0.85, 0.99) | 1.00 (0.54, 1.00) |
| Nyberg (1988) | 36 | 2 | 9 | 4 | 0.80 (0.65, 0.90) | 0.67 (0.22, 0.96) |
| Parvey (1996) | 108 | 0 | 61 | 69 | 0.64 (0.56, 0.71) | 1.00 (0.95, 1.00) |
| Yeh (1986) | 23 | 2 | 13 | 3 | 0.64 (0.46, 0.79) | 0.60 (0.15, 0.95) |
| *Intradecidual sign* | | | | | | |
| Chiang (2004) | 92 | 0 | 61 | 34 | 0.60 (0.52, 0.68) | 1.00 (0.90, 1.00) |
| Yeh (1986) | 33 | 0 | 3 | 5 | 0.92 (0.78, 0.98) | 1.00 (0.48, 1.00) |
| *Chorionic rim sign* | | | | | | |
| Parvey (1996) | 135 | 2 | 34 | 67 | 0.80 (0.73, 0.86) | 0.97 (0.90, 1.00) |
| *Yolk sac* | | | | | | |
| Nyberg (1988) | 19 | 0 | 26 | 6 | 0.42 (0.28, 0.58) | 1.00 (0.54, 1.00) |

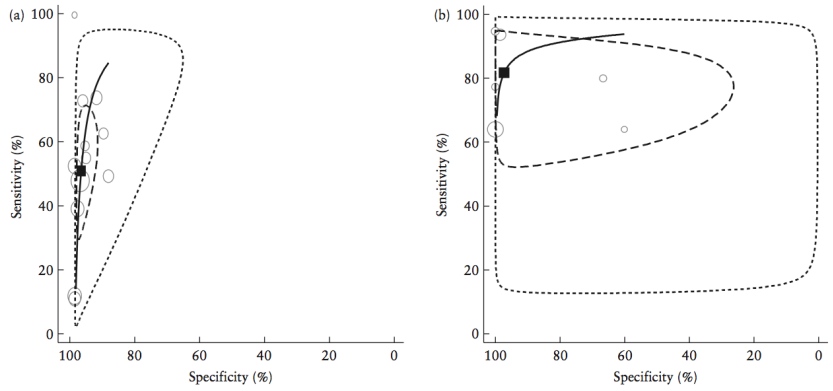



## Diagnostic Accuracy Of The Intradecidual Sign For Predicting An Intrauterine Pregnancy

Two cohort studies [218, 230], including 228 women in early pregnancy, evaluated the diagnostic accuracy of the intradecidual sign for predicting the

likelihood of an intrauterine pregnancy. Figure 5.2 shows the sensitivities and specificities of an intradecidual sign to predict an intrauterine pregnancy in the individual studies. The precision estimates for each of the studies and the estimated summary sensitivity and specificity for differentiating between an intrauterine and an extra-uterine pregnancy are shown in Table 5.2. Following meta-analysis of these two studies we found that the presence of the intradecidual sign predicts an intrauterine pregnancy with a pooled sensitivity of 66% (95% CI, 59-73%), specificity of 100% (95% CI, 91-100%), positive likelihood ratio of 21 (95% CI, 3.1-141) and negative likelihood ratio of 0.22 (95% CI, 0.06-0.88). The median prevalence of an intrauterine pregnancy was 85%, but if the intradecidual sign was present the probability of an intrauterine pregnancy was as high as 99% compared with 56% if the intradecidual sign was absent.

***Figure 5.3: Summary receiver operating characteristics (ROC) plot of the ability of a gestation sac (a) and the double decidual sac sign (b) to predict an intrauterine pregnancy.***




*o, Study estimate; ––––, hierarchal summary ROC curve; -----, 95% prediction region; □, summary point; – – –, 95% confidence region*

## Diagnostic Accuracy Of The Chorionic Rim Sign For Predicting An Intrauterine Pregnancy

One cohort study [220], including 238 women in early pregnancy, evaluated the diagnostic accuracy of the chorionic rim sign for predicting the likelihood of an intrauterine pregnancy. The estimated summary sensitivity and specificity for differentiating an intrauterine from an extra-uterine pregnancy are shown in Figure 5.2 and Table 5.2. This study found that the presence of the chorionic rim sign predicts an intrauterine pregnancy with a sensitivity of 80% (95% CI, 73-

86%), specificity of 97% (95% CI, 90-100%), positive likelihood ratio of 28 (95% CI, 7.0-108) and negative likelihood ratio of 0.21 (95% CI, 0.15-0.28). In the study the prevalence of an intrauterine pregnancy was 71%, but if the chorionic rim sign was present the probability of an intrauterine pregnancy was as high as 99% compared with 66% if the chorionic rim sign was absent.

## Diagnostic Accuracy Of The Yolk Sac For Predicting An Intrauterine Pregnancy

One cohort study [239], including 51 women in early pregnancy, evaluated the diagnostic accuracy of the presence of the yolk sac for predicting the likelihood of an intrauterine pregnancy. The estimated summary sensitivity and specificity for differentiating an intrauterine from an extra-uterine pregnancy are shown in Figure 5.2 and Table 5.2. This study found that the presence of a yolk sac predicts an intrauterine pregnancy with a sensitivity of 42% (95% CI, 28-58%), specificity of 100% (95% CI, 54%-100%), positive likelihood ratio was infinite and negative likelihood ratio of 0.58 (95% CI, 0.45-0.74). In this study the prevalence of an intrauterine pregnancy was 88%, but if a yolk sac was present the probability of an intrauterine pregnancy was 100% compared with 19% probability if a yolk sac was absent.

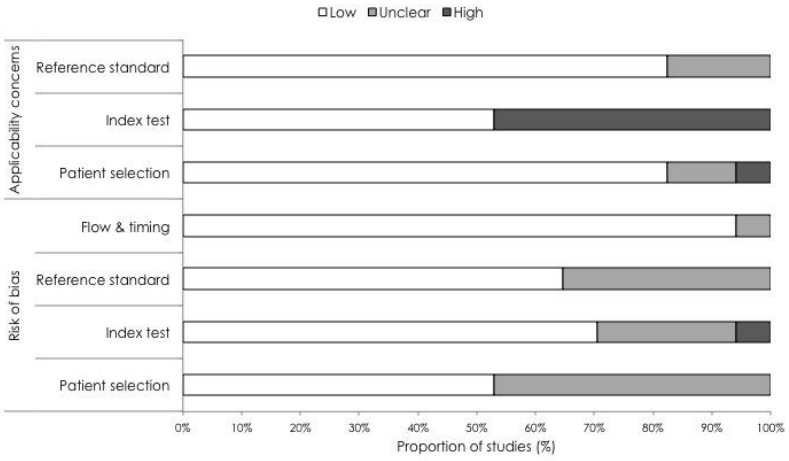
## Risk Of Bias Within Studies

Figure 5.4 summarizes the risks of bias and applicability concerns of studies based on QUADAS-2 (the assessment of each individual study is presented in Table 5.3). Although some high quality studies were included in the systematic review [220, 230, 236], the quality of most of the studies was considered mediocre. Six studies were retrospective in nature [219, 230, 232, 236, 237, 241], five were small (including fewer than 100 participants) [218, 219, 232, 235, 240], and twelve studies were undertaken more than 20 years ago [219, 228, 229, 231-235, 238-241]. Many studies did not fully describe the methods of patient selection, most notably with respect to whether a consecutive or random sample of patients was selected, and hence it is unclear whether the selection of patients could have introduced bias[229, 231, 233, 234, 238-240].

The inclusion criteria for the different studies were also variable. In some studies the only inclusion criteria appeared to be that of a positive pregnancy test [219, 234, 239] whilst most others additionally required symptoms

suggestive of an ectopic or failing pregnancy namely abdominal pain and/or vaginal bleeding [218, 220, 228, 229, 231-233, 235, 236, 238, 240]. Other studies had more specific inclusion criteria. The study by Chiang *et al*., [230], for example, included patients who were pregnant and whose ultrasonographic findings revealed the presence of either an intrauterine fluid collection associated with an early intrauterine pregnancy of less than 5.5 weeks' gestation (defined as a mean sac diameter of less than or equal to 8mm) or an ectopic pregnancy. In contrast, Dart *et al*., [237] included symptomatic pregnant women with indeterminate transvaginal ultrasound scans and either a serum human Chorionic Gonadotropin (hCG) level of greater than 3000mIU/ml or women whose last menstrual period was more than 38 days before examination. The results of these studies with more specific inclusion criteria may be less generalizable.

***Figure 5.4: Risk of bias and applicability concerns based on quality assessment of diagnostic accuracy studies (QUADAS)-2 across included studies***




The degree of blinding in the studies was also unclear. Many studies did not explicitly state whether the ultrasound images were interpreted without knowledge of the final diagnosis. In the prospective studies, it is probable that this was the case owing to the inevitable passage of time that occurred whilst waiting for the clinical follow-up (reference standard) to occur. It is less clear in the retrospective studies [219, 230, 232, 236, 237, 241]. Furthermore three studies did not clearly define the ultrasonographic feature under surveillance

[228, 232, 235] and in those studies that did give a clear definition, there were often considerable differences between the studies. Most studies that investigated the accuracy of a gestation sac defined a gestation sac as being an anechoic intrauterine fluid collection surrounded by an echogenic border [229, 233, 238-240] but two studies included the presence of internal echoes in the definition [231, 234] and two others incorporated a size limitation [236, 237]. Therefore the conduct and/or interpretation of the index test could have introduced bias. Some of the older studies utilized transabdominal ultrasound [219, 232, 233, 238, 239] and the ultrasound approach was not stated in others [235] and hence their results may not be applicable to current practice. Seven studies [218, 219, 232-234, 236, 237] did not clearly define the reference standard and in the majority of studies it was unclear if the results of the reference standard were interpreted without knowledge of the index test. Patient flow was considered to be appropriate in all the studies.

# 5.5 Discussion

## Summary of Evidence

This systematic review and meta-analysis summarizes the diagnostic accuracy of commonly used ultrasonographic signs for indicating the location of a pregnancy, and shows that the presence of any of the ultrasonographic features evaluated, namely a gestation sac, double decidual sac sign, intradecidual sign or chorionic rim sign, increases substantially the probability that a pregnancy is of intrauterine location. Therefore, the presence of these signs indicates an intrauterine pregnancy and can be used to guide clinical practice. The exception to this is the use of the presence of the gestation sac as this test is slightly less specific than the others for predicting an intrauterine pregnancy. The absence of these signs does not exclude the diagnosis of an intrauterine pregnancy, and a negative test result therefore cannot be relied upon to inform clinical practice.

## Strengths and Weaknesses of Study

We conducted a prospective and extensive systematic search of electronic databases using a predefined protocol which was registered with PROSPERO. The high number of included studies in our meta-analyses for a gestation sac

and the double decidual sac sign strengthened the power of these conclusions and enabled us to define the diagnostic accuracy of these signs in confirming an intrauterine pregnancy with relative precision. Our findings for the other ultrasonographic features i.e. the intradecidual sign, chorionic rim sign and yolk sac, were, however, limited by the small number of included studies.

An additional strength is that we performed an assessment of quality of the included studies. However, the quality of the included studies was relatively poor as there was a substantial risk of bias and concerns regarding the applicability to current clinical practice. Furthermore, many of the studies reported a different prevalence of pregnancy outcomes compared with more recent studies, which may affect the generalizability of the findings to clinical practice in a variety of settings.

The main limitation of our study is that our conclusions with regard to evaluating the accuracy of visualization of a yolk sac for determining the location of an intrauterine pregnancy have been drawn from one small study. Other studies investigating the significance of a yolk sac in early pregnancy were identified by the search strategy but these did not meet the pre-specified inclusion criteria. These studies were largely considered to be irrelevant, as they were more concerned with the relative size, shape or position of the yolk sac with regard to predicting pregnancy viability than with the actual presence of the yolk sac confirming identification of a true gestation sac and, ultimately, an intrauterine pregnancy, which was the focus of our review. It is surprising that no other studies have been conducted to investigate the performance of visualization of the yolk sac on ultrasound for determining the true nature of an intrauterine fluid collection. It can be speculated that this may be because, embryologically, the yolk sac is derived from migrating hypoblast cells of the inner cell mass and could therefore only occur within a true gestation sac. In the case of a pseudosac, which is merely a fluid filled space with no gestational tissue, there is no potential to develop a yolk sac. Given this fact and the 100% specificity found in the one included study, further studies to investigate the accuracy of a yolk sac for predicting an intrauterine pregnancy may have been considered unnecessary.

A further limitation of our study is that there is wide variation in sensitivity and specificity between studies reporting on the same ultrasonographic sign. For

example, the sensitivity of a gestation sac for predicting an intrauterine pregnancy ranged from 14% in the study by Dart *et al* [237] to 100% in the study by Weckstein *et al* [235]. This is probably because of the considerable population heterogeneity between the studies. Dart *et al* [237] included only pregnant women with abdominal pain and/or vaginal bleeding with an indeterminate ultrasound scan who had either a serum hCG level greater than 3000iu/l or whose last menstrual period was more than 38 days before examination. It was conducted using transvaginal ultrasound and a gestation sac was defined as an empty anechoic intrauterine fluid collection with a hyperechoic border and mean sac diameter of less than 10mm. In contrast, Weckstein *et al* [235] included a less specific group of patients, a broader definition of what constitutes a gestation sac and a less accurate ultrasonographic approach. It is therefore of no surprise that these studies, with their inherent differences in study design, have yielded considerably different accuracy measures.

A final limitation of this study is that no information regarding pregnancy viability can be inferred from the results. The finding of the double decidual sac sign, for example, suggests an intrauterine pregnancy with a sensitivity of 82% and specificity of 97%, whether that pregnancy is viable or not cannot be concluded from our results. However it was the aim of this systematic review to determine the accuracy of first trimester ultrasonographic signs in predicting intrauterine pregnancy location, and this has been accomplished. In order to achieve this, of studies that considered three separate outcomes including viable intrauterine, non-viable intrauterine and ectopic pregnancies, all intrauterine pregnancies were combined prior to construction of the 2x2 tables [219, 229-234, 236, 237, 239-241]. Of studies that did not differentiate between viable and non-viable intrauterine pregnancies no such combination was required [218, 220, 228, 235, 238].

## 5.6 Conclusion

This review is the first to comprehensively collate evidence of the accuracy of various different ultrasonographic features for predicting an intrauterine pregnancy prior to ultrasonographic visualization of the yolk sac. The findings are limited by the relatively small number and poor quality of the included studies and by the heterogeneity seen between the tests and outcome

assessment. An appropriately powered study following STARD guidelines [242] using transvaginal ultrasound and an appropriate reference standard, is required to establish standards for the accurate prediction of an intrauterine pregnancy. In the interim, it would be prudent to continue the current practice of waiting until a yolk sac is visualized before confirming that a pregnancy is intrauterine.

***Table 5.1: Characteristics of studies included in the systematic review and meta-analysis***

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Ankum, 1993 (n=208)[228] | Prospective | Positive UPT, pain +/- PVB, risk factors for EP | | TV 5.5-7.5MHz | GS | Not stated | IUP: uneventful clinical course<br>EP: all were confirmed surgically |
| Bateman, 1990 (n=126)[229] | Prospective | Suspected EP | | TV 5MHz | GS | Anechoic intrauterine fluid collection with mean sac diameter <10mm and a regular echogenic border | Normal IUP: normal progression into the second trimester<br>EP: documented surgically<br>SA: terminated before 12 weeks gestation |
| Bradley, 1982 (n=50)[219] | Retrospective | Positive UPT | Fetal pole or abnormal GS; lost to follow-up | TA 3.5-5MHz | DDSS | Two concentric rings within the endometrial cavity | Normal IUP, abnormal IUP or EP: determined from medical records |
| Chiang, 2004 (n=187)[230] | Retrospective | Intrauterine fluid collection with mean sac diameter <8mm or suspected EP | No intrauterine fluid collection or fluid collection with a DDSS, YS or fetal pole; heterotopic pregnancy; no TV US; Lost to follow up | TV 5-9MHz | IDS | Echogenic area embedded in thickened decidua that is eccentrically located on one side of the uterine cavity which appears as a well-defined endometrial stripe | Normal IUP: subsequent US shows fetal heart or a hospital record reported delivery of a live infant corresponding to the index pregnancy<br>Abnormal IUP: histology shows chorionic villi<br>EP: surgical records show an EP or if the US findings show classic findings of an EP with an extra-ovarian adnexal mass in patients treated with methotrexate or showed free fluid with debris with follow-up and hCG levels showing inappropriate increases over >7 days |
| Dart, 1997 (n=163)[237] | Retrospective | hCG>3000iu/l and/or LMP>38 days ago with pain ± PVB and positive UPT and indeterminate US | Recent ERPC/delivery; hCG or LMP unknown; final diagnosis unknown | TV 5MHz | GS | An empty anechoic intrauterine fluid collection with a hyperechoic border and a mean sac diameter <10mm | Normal IUP, abnormal IUP and EP: determined from hospital records and results from radiology, laboratory and pathology department databases |
| Dart, 1998 (n=228)[236] | Retrospective | Positive UPT, pain ± PVB, indeterminate TVUS | Recent ERPC/delivery; hCG or LMP unknown; final diagnosis unknown | TV 5MHz | GS | An empty anechoic intrauterine fluid collection with hyperechoic border and mean sac diameter <10mm | Normal IUP, abnormal IUP and EP: determined from hospital records and results from radiology, laboratory and pathology department databases |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Enk, 1990 (n=107)[231] | Prospective | Positive UPT, pain ± PVB, risk factors for EP | Repeat US | TV 7MHz | GS | Hypoechoic area including internal echoes | Normal IUP: ultrasonographic evidence of normal fetal development in utero<br>SA: absence of ultrasonographic evidence of fetal development with histopathological confirmation of trophoblasts in curettage material<br>EP: Histopathological confirmation of extra-uterine trophoblasts obtained at laparotomy/laparoscopy |
| Kadar, 1981 (n=97)[232] | Retrospective | Suspected EP | hCG>25000mIU/ml | TA | GS | Not stated | Normal IUP, abnormal IUP and EP: not stated |
| Nyberg, 1983 (n=128)[241] | Retrospective | Intrauterine fluid collection associated with pregnancy or suspected EP | Fetal pole visible on US; inadequate US; Lost to follow up | TA 3.5-5MHz | DDSS | Two concentric rings surrounding a portion of the gestational sac | Normal IUP: patients carried the gestation to term or had a living intrauterine fetus confirmed subsequently<br>Abnormal IUP: presence of chorionic villi on histopathology<br>EP: all confirmed surgically |
| Nyberg, 1987 (n=63/150 DDSS/GS)[238] | Prospective | Pain ± PVB, TVS and hCG | IUP on US; subsequent non-viable IUP; repeat US | TA 3.5MHz | DDSS<br>GS | Not stated<br>Central sonolucency surrounded by an echogenic ring | IUP: shown to have normal outcome by follow-up sonograms or clinical evaluation<br>EP: all confirmed surgically |
| Nyberg, 1988 (n=51/136 DDSS&YS/GS)[240] | Prospective | Positive UPT | Lost to follow up | TA | DDSS<br><br>GS<br><br><br>YS | Two concentric rings surrounding portion of GS<br>Intrauterine fluid collection surrounded by echogenic ring<br>Sonolucent rounded sac like structure within the GS | Normal IUP: repeat sonograms and/or clinical evaluation demonstrated normal growth of a living fetus<br>Abnormal IUP: failure of normal growth and development on follow-up sonograms or clinical examination usually supported by declining serum hCG levels<br>EP: all confirmed surgically |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Nyberg, 1988 (n=84)[243] | Prospective | Positive UPT, pain ± PVB | Fetal heart on US, repeat US | TA 3.5-5MHz TV5-7.5MHz | GS | Intrauterine fluid collection surrounded by echogenic ring | Normal IUP: normal gestational growth on clinical examination or demonstration of a living fetus >10 weeks gestation<br>Abnormal IUP: lack of normal growth demonstrated clinically or by later ultrasonographic examination usually accompanied by declining serial hCG levels<br>EP: all confirmed surgically |
| Parvey, 1996 (n=238)[220] | Prospective | Positive UPT, pain ± PVB | Fetal heart on US | TA 5MHz<br>TV 5-7 MHz | CRS<br><br><br>DDSS | Single concentric echogenic ring around an early intrauterine GS<br>Two concentric echogenic rings around an early intrauterine GS | Live IUP: repeat sonogram showing YS or fetal pole ± fetal heart or documented term delivery of a live infant<br>Abnormal IUP: pathological examination after curettage<br>EP: surgical and pathological confirmation of products of conception in the fallopian tube |
| Romero, 1985 (n=383)[233] | Prospective | Positive UPT, Suspected EP | | TA 3.5MHz | GS | Hypoechogenic area surrounded by an echogenic rim | Normal IUP: resulting in a viable infant<br>Abnormal IUP: not stated<br>EP: not stated |
| Tongsong, 1993 (n=201)[234] | Prospective | Haemodynamically stable, known hCG, | Repeat US; lost to follow up | TV 5MHz | GS | GS with a fetal pole+/- fetal heart or DDSS | Normal IUP, abnormal IUP or EP: determined following review of surgical and clinical records |
| Weckstein, 1985 (n=37)[235] | Prospective | Positive UPT, pain | Haemodynamically unstable; febrile | Not stated | GS | Not stated | IUP: not stated<br>EP: confirmed surgically |
| Yeh, 1986 (n=41)[218] | Prospective | Suspected EP | | TA 3.5MHz | DDSS<br>IDS | Not stated<br>GS or echogenic area of early implantation, remains within thickened decidua on one side of the uterine cavity which is relatively undisturbed and appears as a straight line | IUP or EP: proven by subsequent scanning, clinical follow-up, delivery, curettage or surgery |

***Table 5.2: Summary estimates of each ultrasonographic sign for predicting an intrauterine pregnancy***

| **Ultrasonographic Sign** | **Studies n [N]** | **Sensitivity (95% CI)(%)** | **Specificity (95% CI)(%)** | **LR+ (95% CI)** | **LR- (95% CI)** | **Pre- and post-test probability (range)(%)** | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | **Pre-test** | **Post-test if test positive** | **Post-test if test negative** |
| GS | 12 (1920) | 53 (38-67) | 98 (94-99) | 22 (9.8-51) | 0.48 (0.36-0.66) | 68 (30-88.) | 98 (96-99) | 50 (48-52) |
| DDSS | 6 (571) | 82 (68-90) | 97 (76-100) | 30 (2.8-331) | 0.19 (0.10-0.35) | 89 (49-91) | 100 (97-100) | 61 (15-64) |
| IDS* | 2 (228) | 66 (59-73) | 100 (91-100) | 21 (3.1-141) | 0.22 (0.06-0.88) | 85 | 99 | 56 |
| CRS* | 1 (238) | 80 (73-86) | 97 (90-100) | 28 (7.0-108) | 0.21 (0.15-0.28) | 71 | 99 | 66 |
| YS* | 1 (51) | 42 (27-58) | 100 (54-100) | ∞ | 0.58 (0.45-0.74) | 88 | 100 | 19 |

**Probability ranges not applicable for ultrasonographic signs with fewer than three studies. n[N], number of studies [number of women]*

*Table 5.3: Quality assessment of included studies in the systematic review using quality assessment of diagnostic accuracy studies (QUADAS)-2*

| Study | Risk of Bias | | | Applicability Concerns | | | |
|---|---|---|---|---|---|---|---|
| | Patient Selection | Index Test | Reference Standard | Flow and Timing | Patient Selection | Index Test | Reference Standard |
| Ankum (1993)[228] | Low | Unclear | Unclear | Low | Low | Low | Low |
| Bateman (1990)[229] | Unclear | Low | Low | Low | Unclear | Low | Low |
| Bradley (1982)[219] | Low | Unclear | Unclear | Low | Low | High | Unclear |
| Chiang (2004)[230] | Low | Low | Low | Low | Low | Low | Low |
| Dart (1997)[237] | Unclear | Low | Low | Low | High | Low | Low |
| Dart (1998) [236] | Low | Low | Low | Low | Low | Low | Low |
| Enk (1990)[231] | Unclear | Low | Low | Low | Low | Low | Low |
| Kadar (1981)[232] | Low | Unclear | Unclear | Unclear | Unclear | High | Unclear |
| Nyberg (1983)[241] | Low | High | Unclear | Low | Low | High | Low |
| Nyberg (1987)[238] | Unclear | Low | Low | Low | Low | High | Low |
| Nyberg (1988)[240] | Unclear | Low | Low | Low | Low | High | Low |
| Nyberg (1988)[243] | Unclear | Low | Low | Low | Low | Low | Low |
| Parvey (1996)[220] | Low | Low | Low | Low | Low | Low | Low |
| Romero (1985)[233] | Unclear | Low | Low | Low | Low | High | Unclear |
| Tongsong (1993)[234] | Unclear | Low | Low | Low | Low | Low | Low |
| Weckstein (1985)[235] | Low | Unclear | Unclear | Low | Low | High | Low |
| Yeh (1986)[218] | Low | Low | Unclear | Low | Low | High | Low |

# 6. Accuracy Of First Trimester Ultrasound For Diagnosis Of Tubal Ectopic Pregnancy In The Absence Of An Obvious Extra-Uterine Embryo

*NB; An abridged version of this chapter has been published in the journal Ultrasound in Obstetrics and Gynaecology [244]*

## 6.1 Introduction

The incidence of ectopic pregnancy has risen over the last few decades. Fortunately both maternal morbidity and mortality associated with the condition has declined during this period, largely owing to greater awareness and earlier diagnosis (Figure 2.3). Despite this, ectopic pregnancy still accounts for 3.4% of maternal mortality in the UK [245]. Early diagnosis of ectopic pregnancy is essential for reducing maternal mortality. Although diagnostic laparoscopy is considered the gold standard, it has a false positive rate of 5% and a false negative rate of 3-4% [77]. The advent of high-resolution transvaginal ultrasound has revolutionized the diagnosis of ectopic pregnancy.

Unfortunately, the ultrasonographic appearances of ectopic pregnancies vary considerably. Diagnosis should be based on the positive visualization of an extra-uterine mass rather than the inability to recognize an intrauterine pregnancy [114]. As discussed in chapter two, a living embryo located outside the uterus is the only pathognomonic sign of an ectopic pregnancy but is only reported in 8-26% of ectopic pregnancies detected on transvaginal ultrasound [77]. In the absence of an obvious ectopic pregnancy, several different ultrasonographic signs have been proposed to help detect ectopic pregnancy, with variable sensitivities and specificities. These include an empty uterus (i.e. one that does not contain a gestation sac, pseudosac or retained products of conception), a pseudosac, free fluid, and an adnexal mass [3].

An empty uterus is a non-specific finding, which not only occurs in an ectopic pregnancy but also in a very early intrauterine pregnancy, following a complete miscarriage and in a non-pregnant woman.

A pseudosac represents a thickened decidual reaction surrounding an intrauterine fluid collection. They occur in up to 15% of ectopic pregnancies [215]. They can be difficult to differentiate from a true gestation sac. The presence of a pseudosac alone cannot be used to diagnose an ectopic pregnancy. Indeed, when an intrauterine smooth-walled anechoic cystic structure is the only ultrasonographic finding in a woman with a positive pregnancy test, the probability of a tubal ectopic pregnancy is, according to one study, only 0.02% [246], hence a small intrauterine anechoic cystic structure is actually more likely to be an early gestation sac than a pseudosac.

An adnexal mass that is separate from the ovary is the most common finding of a tubal ectopic pregnancy and is seen in as many as 89-100% of patients [164, 247]. In a meta-analysis of ten studies [248], a non-cystic or inhomogeneous mass was the most appropriate criterion on which to diagnose an ectopic pregnancy with a specificity of 99%, sensitivity of 84%, positive predictive value of 96% and negative predictive value of 95%. However, in addition to being non-discriminatory (occurring in other conditions such as ruptured corpus lutea, haemorrhagic cysts, pelvic inflammatory disease and endometriosis), adnexal masses are also non-existent in 15-35% of patients with an ectopic pregnancy [249].

Free fluid may be visible on ultrasound but this is also a non-specific finding and can occur in other pathologies, for example, ruptured ovarian cysts, pelvic inflammatory disease, ovarian torsion and miscarriage. Transvaginal ultrasound has detected free fluid in up to 63% of ectopic pregnancies but it also occurs in 25-31% of intrauterine pregnancies [77]. Echogenic fluid has been reported in 28-56% of ectopic pregnancies [215, 250]. It may signify tubal rupture but most commonly is due to blood leaking from the fimbrial end of a fallopian tube. The presence of free fluid however may help confirm a suspicion of an ectopic pregnancy in lieu of, or in combination with, other ultrasonographic findings.

## 6.2 Aims

We performed a systematic review and meta-analysis of the literature to determine the accuracy of these commonly described, non-specific ultrasonographic signs in the diagnosis of a tubal ectopic pregnancy in the absence of a living embryo located outside the uterus, in women with or without symptoms of abdominal pain and/or vaginal bleeding in early pregnancy.

## 6.3 Methods

The protocol was registered with PROSPERO on 4th October 2012 available from http://www.crd.york.ac.uk/PROSPERO (registration number: CRD42012003046). The systematic review and meta-analysis was carried out according to the methodology described in chapter 5 .3.

## 6.4 Results

### Study Selection

The search identified 19,959 potential papers. Following review of the titles and abstracts, 294 full-text papers were selected for further examination and subsequently 263 of these studies were subsequently excluded (Figure 6.1)). Thirty-one studies [73, 229, 234-236, 239, 240, 250-274], including 5858 women, met the inclusion criteria and were incorporated into the systematic review. The characteristics of the included studies are shown in Table 6.1.

### Diagnostic Accuracy Of An Empty Uterus For The Prediction Of Tubal Ectopic Pregnancy

Thirteen cohort studies [234, 236, 239, 251, 255-257, 259, 261, 262, 265, 274], including 2499 women in early pregnancy, evaluated the diagnostic accuracy of an empty uterus on ultrasound examination to predict the likelihood of an ectopic pregnancy. Figure 6.2 shows the sensitivities and specificities of the presence of an empty uterus for predicting an ectopic

pregnancy in the individual studies. The precision estimates for each of the studies and the estimated summary sensitivity and specificity for differentiating between an ectopic and an intrauterine pregnancy are shown in Figure 6.3a and Table 6.2. Following meta-analysis of these thirteen studies we found that the presence of an empty uterus predicts an ectopic pregnancy with a pooled sensitivity of 32% (95% CI, 20-48%), specificity of 93% (95% CI, 85-97%), positive likelihood ratio of 4.8 (95% CI, 1.6-15) and negative likelihood ratio of 0.72 (95% CI, 0.57-0.93). Of the studies included in this meta-analysis, the median prevalence of an ectopic pregnancy was 34% (range, 7.2% to 78%), however if an empty uterus was present the probability of an ectopic pregnancy was as high as 71% (range, 38% to 94%) compared with 27% (range, 5.3% to 72%) if the uterus was not empty.

***Figure 6.1: Flow chart summarizing study selection of papers on first-trimester ultrasound in the diagnosis of tubal ectopic pregnancy in the absence of an obvious extra-uterine embryo***

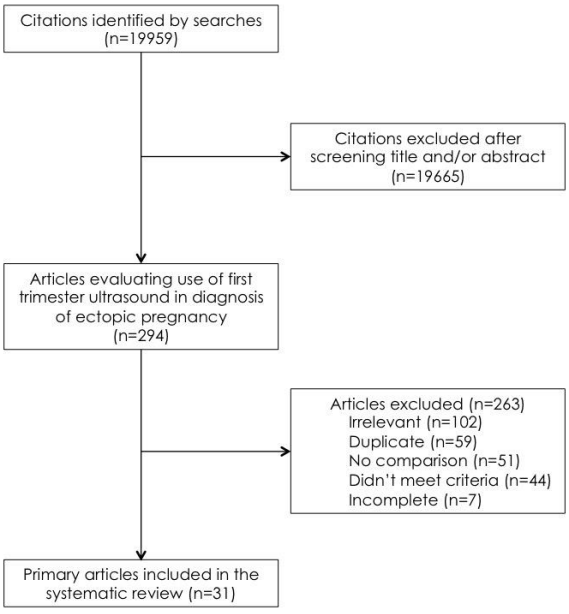
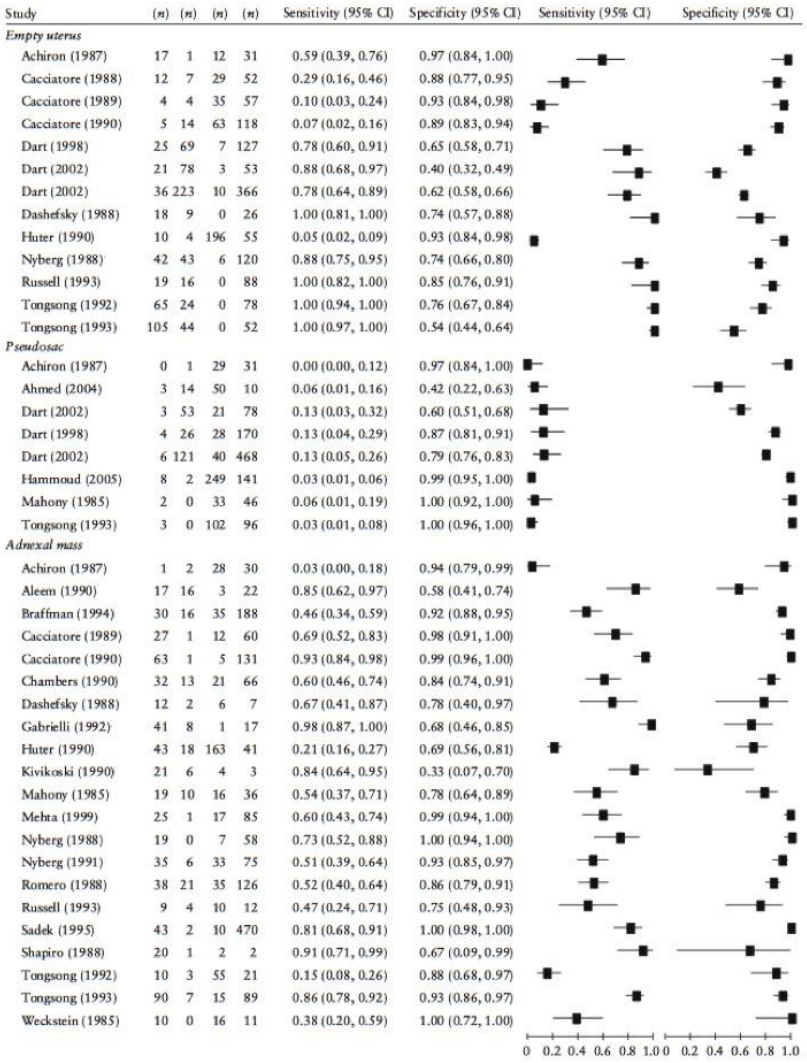


***Figure 6.2: Forest plot for the performance of an empty uterus, a pseudosac and an adnexal mass for predicting a tubal ectopic pregnancy***

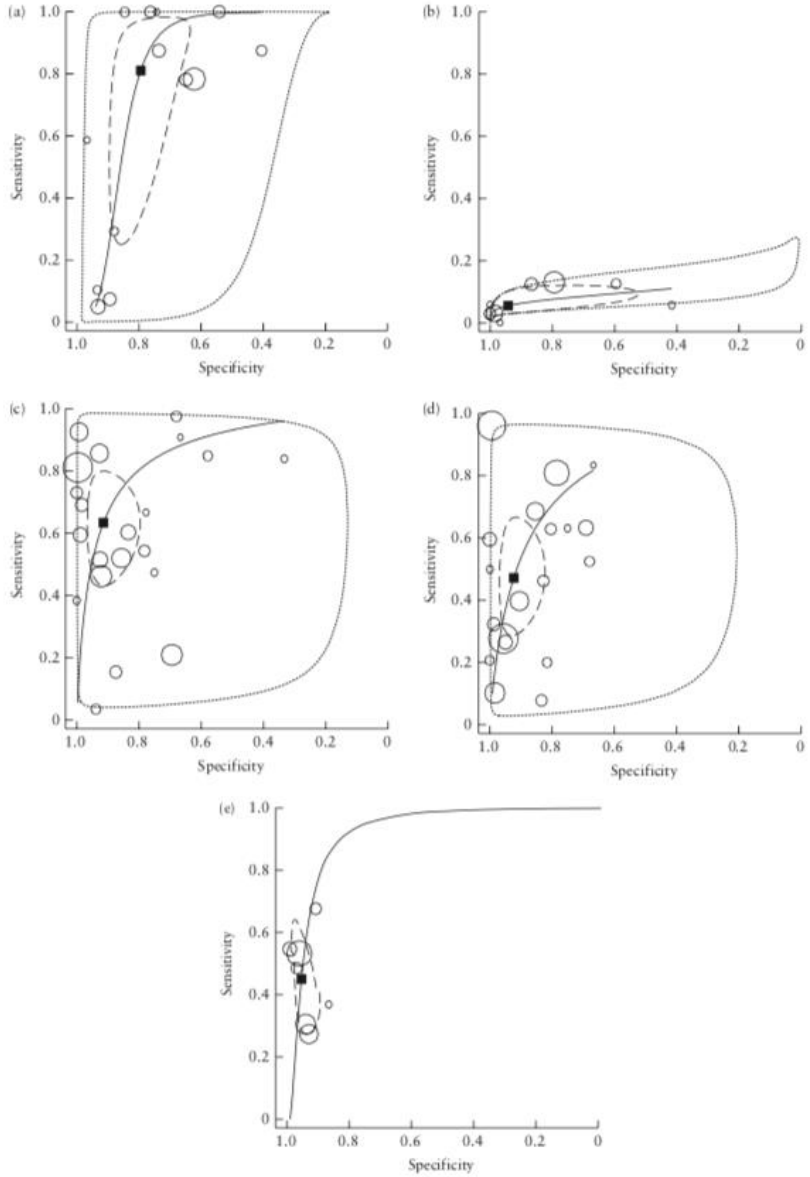



## Diagnostic Accuracy Of A Pseudosac For The Prediction Of Tubal Ectopic Pregnancy

Eight cohort studies [234, 236, 251, 252, 259, 261, 264, 267], including 1838 women in early pregnancy, evaluated the diagnostic accuracy of a pseudosac on ultrasound to predict the likelihood of an ectopic pregnancy. Figure 6.2 shows the sensitivities and specificities of a pseudosac to predict an ectopic pregnancy in the individual studies. The precision estimates for each of the studies and the estimated summary sensitivity and specificity for

differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Figure 6.3 b and Table 6.2.

***Figure 6.3: Summary receiver operating characteristics (ROC) plot of the ability of an empty uterus (a), pseudosac (b), adnexal mass (c), free fluid (d) and the combination of an adnexal mass and free fluid (e) to predict tubal ectopic pregnancy.***




o, Study estimate; ——, hierarchal summary ROC curve; -----, 95% prediction region; □, summary point; – – –, 95% confidence region

Following meta-analysis of these eight studies we found that the presence of a pseudosac predicts an ectopic pregnancy with a pooled sensitivity of 5,5% (95% CI, 3.3-9.0%), specificity of 94% (95% CI, 76-99%), positive likelihood ratio of 0.96 (95% CI, 0.26-3.5) and negative likelihood ratio of 1.0 (95% CI, 0.93-1.1). Of the studies included in this meta-analysis the median prevalence of an ectopic pregnancy was 32% (range, 7.2-69%), however if a pseudosac was present the probability of an ectopic pregnancy fell to 12% (range, 8-16%) compared to staying fairly static (35% (range, 34-35%) probability) if a pseudosac was absent.

## Diagnostic Accuracy Of An Adnexal Mass For The Prediction Of Tubal Pregnancy

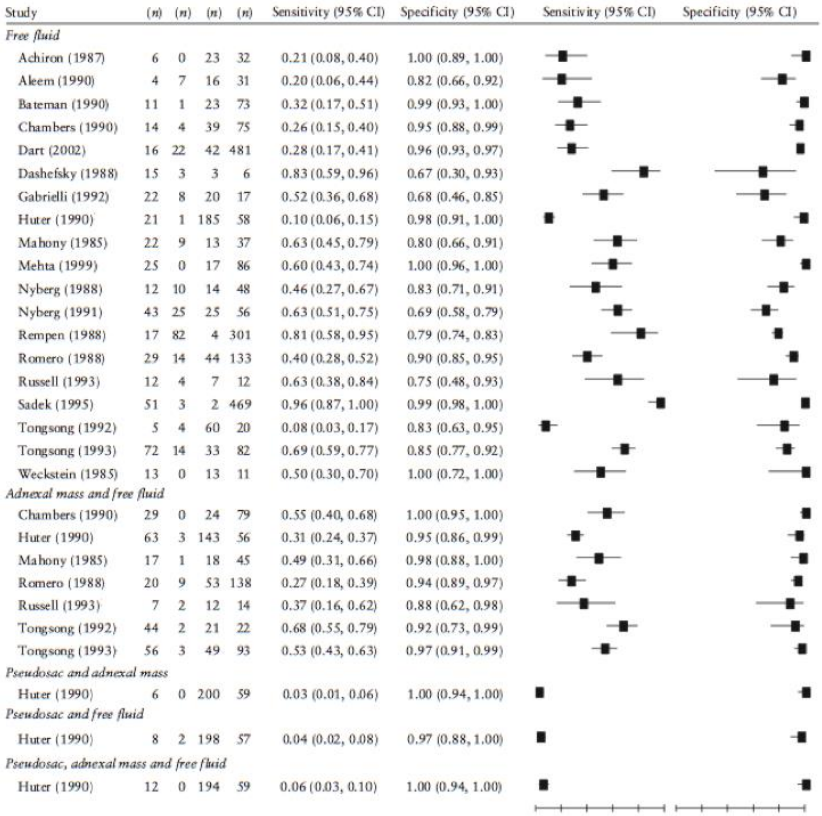
Twenty-one cohort studies [73, 234, 235, 240, 250, 251, 253-256, 258, 262, 263, 265-268, 270-274], including 2787 women in early pregnancy, evaluated the diagnostic accuracy of an adnexal mass for predicting the likelihood of an ectopic pregnancy. Figure 6.2 shows the sensitivities and specificities of an adnexal mass to predict an ectopic pregnancy in the individual studies. The precision estimates for each of the studies and the estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Figure 6.3c and Table 6.2. Following meta-analysis of these 21 studies we found that the presence of an adnexal mass predicts an ectopic pregnancy with a pooled sensitivity of 64% (95% CI, 49-76%), specificity of 91% (95% CI, 84-96%), positive likelihood ratio of 7.4 (95% CI, 3.6-15) and negative likelihood ratio of 0.40 (95% CI, 0.27 to 0.59). Of the studies included in this meta-analysis, the median prevalence of an ectopic pregnancy was 39% (range, 24-88%), however if an adnexal mass was present the probability of an ectopic pregnancy was as high as 83% (range, 80-85%) compared with 21% (range, 19-22%) if there was no evidence of an adnexal mass.

## Diagnostic Accuracy Of Free Fluid For The Prediction Of Tubal Pregnancy

Nineteen cohort studies [229, 234, 235, 240, 250, 251, 253, 258, 260, 262, 263, 265, 267-272, 274] including 3232 women in early pregnancy, evaluated the diagnostic accuracy of free fluid to predict the likelihood of an ectopic pregnancy. Figure 6.4 shows the sensitivities and specificities of free fluid to

predict an ectopic pregnancy in the individual studies. The estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Figure 6.3d and Table 6.2. Following meta-analysis of these nineteen studies we found that the presence of free fluid predicts an ectopic pregnancy with a pooled sensitivity of 47% (95% CI, 33-62%), specificity of 92% (95% CI, 86-96%), positive likelihood ratio of 6.1 (95% CI, 3.1-12.) and negative likelihood ratio of 0.57 (95% CI, 0.43-0.76). Of the studies included in this meta-analysis, the median prevalence of an ectopic pregnancy was 31% (range, 5.2-78%), but if free fluid was present the probability of an ectopic pregnancy was as high as 73% (range, 70-76%) compared with 20% (range, 19-21%) if there was no free fluid.

***Figure 6.4: Forest plot for the performance of free fluid, the combination of an adnexal mass and free fluid, the combination of a pseudosac and an adnexal mass, the combination of a pseudosac and free fluid and the combination of a pseudosac, adnexal mass and free fluid for predicting tubal ectopic pregnancy***



## Diagnostic Accuracy Of The Combination Of An Adnexal Mass And Free Fluid For The Prediction Of Tubal ectopic Pregnancy

Seven cohort studies [234, 258, 265, 267, 270, 271, 274] including 1023 women in early pregnancy, evaluated the diagnostic accuracy of the combination of an adnexal mass and free fluid to predict the likelihood of an ectopic pregnancy. Figure 6.4 shows the sensitivities and specificities of the combination of an adnexal mass and free fluid to predict an ectopic pregnancy in the individual studies. The estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Figure 6.3e and Table 6.2. Following meta-analysis of these seven studies we found that the presence of the combination of an adnexal mass and free fluid predicts an ectopic pregnancy with a pooled sensitivity of 45% (95% CI 34-57%), specificity of 97% (95% CI, 93% to 98%), positive likelihood ratio of 12 (95% CI, 5.9-24) and negative likelihood ratio of 0.57 (95% CI, 0.46-0.71). Of the studies included in this meta-analysis the median prevalence of an ectopic pregnancy was 54% (range, 33-78%), but if the combination of an adnexal mass and free fluid was present the probability of an ectopic pregnancy was as high as 94% (range, 89% to 98%) compared with 40% (range, 38% to 42%) if the combination of an adnexal mass and free fluid was absent.

## Diagnostic Accuracy Of The Combination Of An Adnexal Mass And Pseudosac For The Prediction Of Tubal Pregnancy

One cohort study [265] including 265 women in early pregnancy, evaluated the diagnostic accuracy of the combination of an adnexal mass and pseudosac to predict the likelihood of an ectopic pregnancy. Figure 6.4 shows the sensitivities and specificities of the combination of a pseudosac and adnexal mass to predict an ectopic pregnancy in the individual studies. The estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Table 6.2. This study found that the presence of both an adnexal mass and pseudosac predicts an ectopic pregnancy with a pooled sensitivity of 2.9% (95% CI, 1.1-6.2%), specificity of 100% (95% CI, 93-100%), positive likelihood ratio of infinity and negative likelihood ratio of 0.97 (95% CI, 0.95-0.99). In this study the

median prevalence of an ectopic pregnancy was 78%, but if the combination of an adnexal mass and pseudosac were present the probability of an ectopic pregnancy was as high as 100% compared with 23% if the combination of an adnexal mass and pseudosac were absent.

## Diagnostic Accuracy Of The Combination Of A Pseudosac And Free Fluid For The Prediction Of Tubal Pregnancy

One cohort study [265] including 265 women in early pregnancy, evaluated the diagnostic accuracy of the combination of free fluid and a pseudosac to predict the likelihood of an ectopic pregnancy. Figure 6.4 shows the sensitivities and specificities of the combination of free fluid and a pseudosac to predict an ectopic pregnancy in the individual studies. The estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Table 6.2. This study found that the presence of the combination of free fluid and pseudosac predicts an ectopic pregnancy with a pooled sensitivity of 3.9% (95% CI, 1.7-7.5%), specificity of 97% (95% CI, 88-100%), positive likelihood ratio of 1.2 (95% CI, 0.25-5.3) and negative likelihood ratio of 0.99 (95% CI, 0.94-1.1). In this study the median prevalence of an ectopic pregnancy was 78%, but if the combination of an adnexal mass and pseudosac were present the probability of an ectopic pregnancy was as high as 80% compared with 22% if the combination of an adnexal mass and pseudosac were absent.

## Diagnostic Accuracy Of The Combination Of A Pseudosac, Adnexal Mass And Free Fluid For The Prediction Of Tubal Ectopic Pregnancy

One cohort study [265] including 265 women in early pregnancy, evaluated the diagnostic accuracy of the combination of a pseudosac, adnexal mass and free fluid to predict the likelihood of an ectopic pregnancy. Figure 6.4 shows the sensitivities and specificities of this combination of ultrasonographic features to predict an ectopic pregnancy. The estimated summary sensitivity and specificity for differentiating between an ectopic pregnancy and an intrauterine pregnancy are shown in Table 6.2. This study found that the combination of a pseudosac, adnexal mass and free fluid predicts an ectopic pregnancy with a pooled sensitivity of 5.8% (95% CI, 3.1-10%), specificity of

100% (95% CI, 94-100%), positive likelihood ratio of infinity and negative likelihood ratio of 0.94 (95% CI, 0.91-0.97). In this study the median prevalence of an ectopic pregnancy was 78%, but if the combination of a pseudosac, adnexal mass and free fluid was present the probability of an ectopic pregnancy was as high as 100% compared with 23% if the combination of a pseudosac, adnexal mass and free fluid was absent.

## Risk Of Bias Within Studies

***Figure 6.5: Risk of bias and applicability concerns based on quality assessment of diagnostic accuracy studies (QUADAS)-2 across included studies***

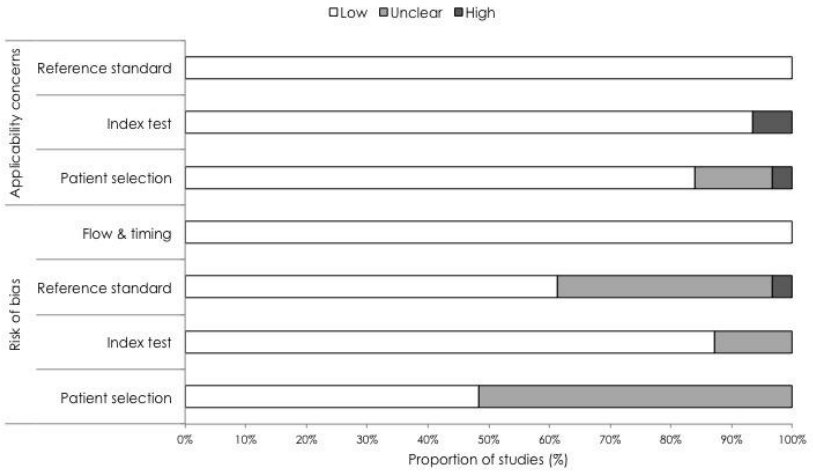



Figure 6.5 summarizes the risk of bias and applicability concerns of studies based on QUADAS-2 (the assessment of each individual study is presented in Table 6.3). The quality of most of the included studies was considered mediocre. Eight studies [236, 240, 252, 261, 262, 264, 265, 271] were retrospective in nature, ten were small (including fewer than 100 participants) [235, 240, 251-253, 262, 263, 266, 267, 273] and 22 were undertaken more than 20 years ago [73, 229, 234, 235, 239, 240, 250, 251, 253, 255-258, 262, 263, 265-267, 269-271, 274]. Many studies [229, 234, 239, 240, 250, 251, 253, 255, 257, 258, 263, 266, 269, 270, 273, 274] did not fully describe the methods of patient selection hence it is unclear whether the selection of patients could have introduced bias. One study only included women who underwent surgery for suspected ectopic pregnancy [265] and three others included women who

were at particularly high risk of ectopic pregnancy as they had risk factors for, as well as symptoms suggestive of, ectopic pregnancy [262, 266, 273].

The degree of blinding was also unclear. The majority of studies did not state explicitly whether the ultrasound images were interpreted without knowledge of the final diagnosis (reference standard result). Four studies [235, 239, 251, 264] did not clearly define the ultrasonographic feature under surveillance and in those studies that did give a clear definition, there was often considerable differences between the studies, for example the study by Braffman *et al* [73] only considered complex adnexal masses, whilst the study by Aleem *et al* [253] included both complex and cystic masses. Similarly, for free fluid, some studies considered merely the presence or absence of free fluid [229, 251, 253, 262, 263, 265, 269, 270, 272, 274] whilst others tried to quantify its volume [240, 267, 268] and two studies included the appearance of the fluid on ultrasound rather than the volume in its definition [234, 258]. Some of the older studies utilized transabdominal ultrasound only [251, 257, 270, 274] and the ultrasound approach was not stated in others [235, 239, 240, 250, 252, 258, 267, 268], hence their results may not be applicable to current practice. Eleven studies did not clearly define the reference standard [73, 235, 251-253, 262, 269, 270, 272-274] and in the majority of studies it was unclear if the results of the reference standard were interpreted without knowledge of the index test. One study clearly stated that the results of the ultrasound were known at the time of surgery and were often an important factor in the decision making process which could have introduced bias [265]. Patient flow was considered to be appropriate in all the studies.

Eight studies had a low risk of bias across all seven domains [236, 256, 259-261, 267, 268, 271] (Table 6.3). Subgroup analysis using only these high quality studies was performed and the estimated summary sensitivities and specificities and positive and negative likelihood ratios of an empty uterus, adnexal mass and free fluid on ultrasound to differentiate between and ectopic pregnancy and intrauterine pregnancy are illustrated in Table 6.4. Utilizing these studies, an adnexal mass and free fluid were both found to be highly specific for an ectopic pregnancy, but not particularly sensitive. An empty uterus on the other hand was less specific but most sensitive.

# 6.5 Discussion

## Summary of Evidence

Our systematic review and meta-analysis summarizes the diagnostic accuracy of commonly used ultrasonographic signs for predicting a tubal ectopic pregnancy and shows that when an obvious extra-uterine pregnancy is not present, the ultrasonographic findings of an empty uterus, a pseudosac, an adnexal mass and/or free fluid have poor sensitivity for identifying a tubal pregnancy. However, the presence of these features on ultrasound has good specificity, especially when found in combination. We can therefore infer ultrasound features are more useful for 'ruling in' a tubal ectopic pregnancy, than for 'ruling out' one.

## Strengths and Weaknesses of Study

We conducted a prospective and extensive systematic search of electronic databases using a predefined protocol that was published. The high number of included studies in our meta-analysis for an empty uterus, the presence of adnexal mass and free fluid strengthened the power of these conclusions and enabled us to define the diagnostic accuracy of these signs in confirming an ectopic pregnancy with relative precision. Our findings for the presence of a pseudosac and various different combinations of ultrasonographic features were limited by the small number of included studies.

An additional strength of our review is that we performed an assessment of quality of the included studies. The quality of most of the included studies was mediocre. The risk of bias and concerns regarding the applicability of the results to current practice were generally low or unclear with only three studies [235, 239, 265] having a high risk of bias or substantial applicability concerns in one or two domains only. We have also conducted a subgroup analysis using the results from eight top quality studies (with a low risk of bias in all seven QUADAS-2 domains).

The main limitation of our study is that the prevalence of an ectopic pregnancy varies considerably between the studies. This is most likely a

reflection of the different inclusion criteria amongst the studies, for example, one study only included women who underwent a diagnostic laparoscopy for suspicion of an ectopic pregnancy [265]. Clearly the prevalence of an ectopic pregnancy in this study will be higher than in those which included women with a positive urinary pregnancy test [240, 269, 270] or in those that included women with symptoms of abdominal pain and/or vaginal bleeding [234, 235, 239, 251, 252, 256, 257, 263, 264, 267, 268, 272, 274]. Studies that included women with risk factors for ectopic pregnancy [262, 266, 273] and those which included symptomatic pregnant women with indeterminate ultrasound scans [73, 236, 250, 253, 255, 259-261, 271] are also likely to have a different prevalence of ectopic pregnancy. In clinical practice it is essential to know how a particular test result predicts the risk of abnormality in the population being evaluated. Sensitivities and specificities do not describe how a particular test result predicts the risk of abnormality. The benefit of likelihood ratios over sensitivity and specificity measures is that they can be used to calculate the probability of abnormality, while adapting for varying *a-priori* probabilities of the chance of abnormality from different contexts. It is essential therefore that the prevalence of ectopic pregnancy in individual early pregnancy assessment/emergency gynaecology units is known before likelihood ratios are applied. Whilst a subgroup analysis based on the level of risk would be interesting and clinically very useful, unfortunately it was not possible in this meta-analysis as only one study included women at high risk of an ectopic pregnancy [265] and only three studies included women at low risk of an ectopic pregnancy [240, 269, 270]. All other studies included women at intermediate risk of an ectopic pregnancy and it was not possible to stratify them further.

A further limitation of our study is that there is wide variation in sensitivity and specificity between studies looking at the same ultrasonographic sign. For example, the sensitivity of an adnexal mass for predicting an ectopic pregnancy ranged from 3.6% in the study by Achiron *et al* [251] to 98% in the study by Gabrielli *et al* [263] and specificity ranged from 33% in the study by Kivikoski *et al* [266] to 100% in the studies by Nyberg *et al* [240], Sadek and Schiotz [272] and Weckstein *et al* [235]. Similarly for free fluid, the sensitivity ranged from 7.7% in the study by Tongsong *et al* [274] to 96% in the study by Sadek and Schiotz [272] and the specificity ranged from 67% in the study by Dashefsky *et al* [262] to 100% in the studies by Achiron *et al [251]*, Mehta *et al [268]* and Weckstein *et al* [235]. This is again likely to be due to the

considerable heterogeneity between the studies involving different populations of women, different ultrasound approaches and different definitions of the signs under evaluation.

Finally, this systematic review demonstrates that a pseudosac is a rare ultrasonographic finding in early pregnancy and is usually absent in women with an ectopic pregnancy. However, when present, it is highly suggestive of an ectopic pregnancy.

Although some experts may disagree [114], many find it difficult to differentiate a gestation sac from a pseudosac prior to the development of a yolk sac or fetal pole. As already discussed in the previous chapter, several different ultrasonographic signs have been proposed to aid in the differentiation, including the intradecidual, double decidual sac and chorionic rim signs, but as the systematic review concluded, whilst the presence of these signs substantially increases the probability that a pregnancy is intrauterine, their absence does not exclude an intrauterine pregnancy and a negative test result cannot be relied upon to guide clinical practice. Furthermore, none of the signs were as accurate in confirming intrauterine pregnancy location as the presence of a yolk sac and hence, in the absence of further research, it is still advisable to wait until a yolk sac is visualized before confirming that a pregnancy is definitely intrauterine. So, whilst a true pseudosac, highly suggestive of an ectopic pregnancy, is a relatively rare ultrasonographic finding in early pregnancy, an empty gestation sac, indicative of an early or failing intrauterine pregnancy, is much more common. It would be preferable to differentiate between these potential diagnoses as soon as possible. Doing so would, not only reduce anxiety for women but also prevent unnecessary investigations for those with an intrauterine pregnancy and minimise morbidity and mortality, permit earlier, potentially less invasive intervention and possibly preserve future fertility for women with an ectopic pregnancy. Therefore this systematic review strengthens the need for a definitive test accuracy study following recommended guidelines [242] to establish standards for the accurate confirmation of an intrauterine pregnancy prior to the development of a yolk sac, or more specifically, the differentiation between an early gestation sac and a pseudosac.

## 6.6 Conclusion

This review is the first to comprehensively collate evidence of the accuracy of various ultrasonographic features to accurately confirm the presence of an ectopic pregnancy in the absence of a live extra-uterine embryo. When an obvious extra-uterine pregnancy is not present, the commonly used ultrasound features have poor sensitivity for identifying a tubal pregnancy, but they have good specificity. We can therefore infer ultrasound features are more useful for 'ruling in' a tubal ectopic pregnancy, than for 'ruling out' one. The findings are limited by the small number and poor quality of the included studies and by the considerable variation in index test and reference standard amongst the different studies.

***Table 6.1 Characteristics of studies included in the systematic review and meta-analysis***

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Achiron, 1987 (n=61)[251] | Prospective | Pain ± PVB | Haemodynamically unstable | TAS | EU<br>PS<br>AM<br>FF | Not stated<br>Not stated<br>Not stated<br>Fluid in the cul-de-sac | EP: histology following laparotomy<br>IUP: not stated |
| Ahmed, 2004 (n=77)[252] | Retrospective | Pain ± PVB | Haemodynamically unstable | Not stated | PS | Any reported sac within uterine cavity without a double decidual sac or yolk sac | EP: not stated<br>IUP: not stated |
| Aleem, 1990 (n=58)[253] | Prospective | Positive UPT, pain ± PVB, indeterminate US | | TVS 5MHz | AM<br>FF | Complex or cystic AM<br>Fluid in the cul-de-sac | EP: not stated<br>IUP: not stated |
| Bateman, 1990 (n=108)[229] | Prospective | Volunteers recruited from infertility unit or referred due to suspected EP | | TVS 5MHz | FF | Fluid in the cul-de-sac | EP: documented surgically<br>IUP: normal progression in to the second trimester or (normal) or histology obtained following curettage (abnormal) |
| Braffman, 1994 (n=269)[73] | Prospective | Positive UPT, pain ± PVB, indeterminate US | Haemodynamically unstable | TAS 3.5-5MHz and TV 5-7.5MHz | AM | Complex AM | EP: not stated<br>IUP: not stated |
| Cacciatore, 1988 (n=100)[257] | Prospective | Positive UPT, pain | | TAS 3-3.5MHz | EU | Neither sac nor adnexal mass | EP: histology following surgery<br>IUP: US revealing a live intrauterine fetus (normal) or histology obtained following curettage (abnormal) |
| Cacciatore, 1989 (n=100)[255] | Prospective | Positive UPT, indeterminate US | | TVS 5MHz | EU<br><br><br>AM | Neither sac nor adnexal mass<br><br><br>Not stated | EP: histology following surgery<br>IUP: US revealing a live intrauterine fetus (normal) or histology obtained following curettage (abnormal) |
| Cacciatore, 1990 (n=200)[30] | Prospective | Positive UPT, pain ± PVB | Haemodynamically unstable | TVS 5-6MHz | EU<br>AM | Neither sac nor adnexal mass<br>A complex AM or a gestation sac-like adnexal ring ± yolk sac or fetal pole | EP: histology following surgery<br>IUP: US revealing a live intrauterine fetus (normal) or histology obtained following curettage (abnormal) |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Chambers, 1990 (n=132)[258] | Prospective | Suspected EP | | Not stated | AM<br>FF | Solid, complex or cystic<br>Transonic fluid or material with low level echoes or of mixed echoity in the pelvic cavity | EP: surgically diagnosed<br>IUP: viable fetus on subsequent US (normal) or no fetus on US and the pregnancy ending in miscarriage (abnormal) |
| Dart, 1998 (n=228)[236] | Retrospective | Positive UPT, pain ± PVB, indeterminate US | Recent delivery/ERPC, lost to follow-up | TVS 5MHz | EU<br><br>FF | Empty endometrial cavity without a thickened endometrium<br>Moderate to large volume of anechoic fluid or any echogenic fluid | EP: determined from hospital records and results from radiology, laboratory and pathology department databases<br>IUP: as above |
| Dart, 2002 (n=155)[261] | Prospective | Positive UPT, pain ± PVB, indeterminate US | Lost to follow-up | TVS 5-10MHz | EU<br><br>FF | Empty endometrial cavity without a thickened endometrium<br>Moderate to large volume of anechoic fluid or any echogenic fluid | EP: visualization of an EP at laparoscopy, in patients managed with methotrexate, identification of an EP on subsequent US or rising hCG levels with no chorionic villi at pathology after ERPC<br>IUP: identification of a fetal heart on subsequent US or delivery of a viable fetus (normal) or falling hCG values with identification of chorionic villi at ERPC, absence of chorionic villi after ERPC with hCG values that fall to zero without further intervention, the patient was observed without intervention and had hCG values that fell to zero or falling hCG values with no EP by laparoscopy (abnormal) |
| Dart, 2002 (n=635)[259] | Prospective | Positive UPT, pain ± PVB, indeterminate US | Recent delivery/ERPC, repeat US, lost to follow-up | TVS 5-10MHz | EU<br><br>FF | Empty endometrial cavity without a thickened endometrium<br>Moderate to large volume of anechoic fluid or any echogenic fluid | EP: visualization of an EP at laparoscopy, in patients managed with methotrexate, identification of an EP on subsequent US or rising hCG levels with no chorionic villi at pathology after ERPC<br>Normal IUP: identification of an IUP with fetal heart on subsequent US or delivery |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Dart, 2002 (n=561)[260] | Retrospective | Positive UPT, pain ± PVB, indeterminate US | Lost to follow-up | TVS 5MHz | FF | Moderate to large volume of anechoic fluid or any echogenic fluid | EP: confirmed by laparoscopy or in patients managed with methotrexate, on the basis of identifying persistently rising hCG levels after a ERPC procedure<br>No EP: an IUP was identified on subsequent US, hCG levels fell to zero without further intervention, a ERPC was performed with identification of chorionic villi at pathology or no EP was visualized during surgery |
| Dashefsky, 1988 (n=53 (EU); 27 (AM/FF))[262] | Retrospective and prospective | Positive UPT, pain ± PVB or risk factors for EP | | TVS 5MHz | EU<br><br>AM<br>FF | No ultrasonographic evidence of an IUP<br>Non-cystic AM<br>No IUP and free adnexal fluid | EP: confirmed by review of the medical and surgical records<br>IUP: as above |
| Gabrielli, 1992 (n=67)[263] | Prospective | Positive UPT, pain ± PVB | | TVS 6.5MHz | AM<br><br>FF | Cystic, complex, tubal ring or ectopic embryo<br>Fluid in the cul-de-sac | EP: confirmed by histological analysis of material obtained at surgery<br>IUP: not stated (normal) or histological demonstration of CV obtained following ERPC (abnormal) |
| Hammoud, 2005 (n=400)[264] | Retrospective | Positive UPT, pain ± PVB | Haemodynamically unstable, definite IUP on US, lost to follow-up | TAS and TVS | PS | Not stated | EP: histopathology following surgery. If no surgery, diagnosis based on clinical evaluation, hormone studies and established ultrasonographic criteria<br>IUP: not stated |
| Huter, 1990 (n=265)[265] | Retrospective | Positive UPT, underwent laparoscopy for suspected EP | | TVS 5-7MHz | EU<br>AM<br>FF | Not stated<br>AM and EU<br>EU and fluid in the cul-de-sac | EP: confirmed surgically<br>IUP: not stated |
| Kivikoski, 1990 (n=34)[266] | Prospective | Positive UPT, pain ± PVB, risk factors for EP | | TVS 5-6.5MHz | AM | Solid or complex , not cystic | EP: confirmed surgically<br>Non-EP: lack of adnexal findings on US and at laparoscopy and/or pathological demonstration of chorionic villi at endometrial curettage |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Mahony, 1985 (n=81)[267] | Prospective | Positive UPT, pain ± PVB | Repeat US, lost to follow-up | Not stated | PS<br><br>AM<br>FF | A single sac like structure with a dense echogenic rind simulating a gestation sac<br>Cystic or non-cystic or adnexal ring<br>Small (only a crescent of fluid seen), moderate (extending into the adjacent adnexa) or large (extending into the lateral paracolic gutters) volume of pelvic intraperitoneal fluid | EP: confirmed by review of the medical/surgical notes<br>IUP: as above |
| Mehta, 1999 (n=128)[268] | Prospective | Positive UPT, pain ± PVB | Definite intrauterine pregnancy on US, lost to follow-up | Not stated | AM<br>FF | Not stated<br>Moderate or large volume of FF | EP: confirmed surgically, by negative findings on ERPC with abnormally rising hCG levels, by ultrasonographic demonstration of an AM separate from the ovary without an IUP or by a combination of these methods<br>IUP: subsequent US demonstrating at least a gestation sac with yolk sac (normal) or documented decreasing serial hCG levels, abnormal US showing an irregular intrauterine sac and/or clinical passage of fetal tissue |
| Nyberg, 1988 (n=211)[243] | Prospective | Positive UPT | Repeat US, lost to follow-up | Not stated | EU | Not stated | EP: all surgically proven<br>IUP: normal gestational growth on clinical examination or demonstration of a living fetus at beyond 10/40 (normal) or lack of normal growth demonstrated clinically or by later ultrasonographic examination usually accompanied by declining serial hCG levels (abnormal) |
| Nyberg, 1988 (n=84)[240] | Retrospective | Positive UPT, pain ± PVB | Fetal heart on US, repeat US | Not stated | AM<br><br>FF | Solid or complex AM not including ovarian cysts<br>Small to moderate volume of fluid in the cul-de-sac | EP: all surgically proven<br>IUP: normal gestational growth on clinical examination or demonstration of a living fetus at beyond 10/40 (normal) or lack of normal growth demonstrated clinically or by later ultrasonographic examination usually accompanied by declining serial hCG levels (abnormal) |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Nyberg, 1991 (n=149)[250] | Prospective | Positive UPT, indeterminate US | | Not stated | AM<br><br><br><br><br>FF | A living extra-uterine embryo or a sac-like structure +/- a mass OR a solid complex mass lacking an embryo or a sac. Obvious cysts or corpus luteum were not considered to be masses and were not included.<br>Absent, small, moderate or large volume of echogenic, anechoic or indeterminate fluid | EP: confirmed by review of surgical and clinical records<br>IUP: as above |
| Rempen, 1988 (n=404)[269] | Prospective | Positive UPT or clinical findings consistent with a pregnancy of 4-13 weeks' gestation | | TVS 5MHz | FF | Fluid in the cul-de-sac | EP: not stated<br>IUP: not stated |
| Romero, 1988 (n=220)[270] | Prospective | hCG<6000iu/l | | TAS 3.5MHz | AM<br>FF | Cystic or non-cystic<br>Fluid in the cul-de-sac | EP: not stated<br>IUP: resulting in a viable infant |
| Russell, 1993 (n=123 (EU);35 (AM/FF))[271] | Retrospective | Positive UPT, pain ± PVB, indeterminate US | Lost to follow-up | TAS 3.5MHz and TVS 5MHz | EU<br><br>AM<br>FF | No IUP (IUP=gestation sac+double decidual sac/yolk sac/fetal pole)<br>Cystic, solid, mixed or ring<br>Pelvic or abdominal fluid; trace, small, moderate or large volume; anechoic or echogenic | EP: confirmed by review of the medical/surgical records<br>IUP: as above |
| Sadek, 1995 (n=525)[272] | Prospective | Positive UPT, pain ± PVB | Haemodynamically unstable | TVS 5MHz | AM<br>FF | Tubal mass<br>Free pelvic fluid | EP: confirmed by surgery and histology<br>IUP: not stated |
| Shapiro, 1988 (n=25)[273] | Prospective | High risk of EP, pain ± PVB, abnormally rising hCG | | TVS 5MHz | AM | Extra-uterine gestation sac ± fetal pole ±fetal heart OR a thick echogenic band surrounding a small hypoechoic core giving the appearance of a donut OR a diffuse echogenic mass within the tube | EP: surgically proven<br>IUP: not stated |

| Study | Type of Study | Inclusion Criteria | Exclusion Criteria | US Details | Sign | Definition | Outcomes Measured |
|---|---|---|---|---|---|---|---|
| Tongsong, 1992 (n=167 (EU); 89 (AM/FF)) [274] | Prospective | Positive UPT, suspected EP | Haemodynamically unstable, lost to follow-up | TAS 3.5MHz | EU<br><br>AM<br>FF | No gestation sac, no PS, no intrauterine fluid or blood collection<br>Complex mass<br>Fluid in the cul-de-sac | EP: not stated<br>IUP: resulting in a viable infant |
| Tongsong, 1993 (n=201) [234] | Prospective | Positive UPT, suspected EP | Haemodynamically unstable, repeat US, lost to follow-up | TVS 5MHz | EU<br><br>PS<br>AM<br><br>FF | No gestation sac, no PS, no intrauterine fluid or blood collection<br>Not stated<br>Complex mass without a sac or embryo<br>Anechoic or echogenic fluid | EP: determined following review of surgical and clinical records<br>IUP: as above |
| Weckstein, 1985 (n=37) [235] | Prospective | Positive UPT, pain | Haemodynamically unstable, febrile | Not stated | AM | Not given | EP: confirmed surgically<br>IUP: not stated |

*Table 6.2 Summary estimates for each ultrasonographic sign for predicting tubal ectopic pregnancy*

| Ultrasonographic Sign | Studies n [N] | Sensitivity (95% CI)(%) | Specificity (95% CI)(%) | LR+ (95% CI) | LR- (95% CI) | Pre- and post-test probability (range)(%) | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Pre-test | Post-test if test positive | Post-test if test negative |
| Empty uterus | 13 (2499) | 81 (42-96) | 80 (69-87) | 4.0 (2.7-5.8) | 0.24 (0.06-0.94) | 30 (7.2-78) | 62 (60-65) | 9 (8-11) |
| Pseudosac | 8 (1838) | 5.5 (3.3-9.0) | 94 (76-99) | 0.96 (0.26-3.5) | 1.0 (0.93-1.1) | 32 (7.2-69) | 12 (8-16) | 35 (34-35) |
| Adnexal mass | 21 (2787) | 64 (49-76) | 91 (84-96) | 7.4 (3.6-15) | 0.40 (0.27-0.59) | 40 (24-88) | 83 (80-85) | 21 (19-22) |
| Free fluid | 19 (3232) | 47 (33-62) | 92 (86-96) | 6.1 (3.08-12) | 0.57 (0.43-0.76) | 31 (5.2-78) | 73 (70-76) | 20 (19-21) |
| Adnexal mass and free fluid | 7 (1023) | 45 (34-57) | 97 (94-98) | 12 (5.9-24) | 0.57 (0.46-0.71) | 54 (33-78) | 93 (90-96) | 40 (38-42) |
| Adnexal mass and pseudosac | 1 (265) | 2.9 (1.1-6.2) | 100 (93.4-100) | ∞ | 0.97 (0.95-0.99) | 78 | 100 | 23 |
| Free fluid and pseudosac | 1 (265) | 3.9 (1.7-7.5) | 97 (88-100) | 1.2 (0.25-5.3) | 0.99 (0.94-1.1) | 78 | 80 | 22 |
| Adnexal mass, free fluid and pseudosac | 1 (265) | 5.8 (3.1-10) | 100 (94-100) | ∞ | 0.94 (0.91-0.97) | 78 | 100 | 23 |

*n[N], number of studies [number of women]*

*Table 6.3: Quality of included studies in the systematic review using quality assessment of diagnostic accuracy studies (QUADAS)-2*

| | Risk of Bias | | | Applicability Concerns | | | |
|---|---|---|---|---|---|---|---|
| | **Patient Selection** | **Index Test** | **Reference Standard** | **Flow and Timing** | **Patient Selection** | **Index Test** | **Reference Standard** |
| Achiron, 1987 [251] | Unclear | Unclear | Unclear | Low | Unclear | Low | Low |
| Ahmed, 2004 [252] | Low | Low | Unclear | Low | Low | Low | Low |
| Aleem, 1990 [253] | Unclear | Low | Unclear | Low | Low | Low | Low |
| Bateman, 1990 [229] | Unclear | Low | Low | Low | Unclear | Low | Low |
| Braffman, 1994 [73] | Low | Low | Unclear | Low | Low | Low | Low |
| Cacciatore, 1988 [257] | Unclear | Low | Low | Low | Low | Low | Low |
| Cacciatore, 1989 [255] | Unclear | Low | Low | Low | Low | Low | Low |
| Cacciatore, 1990 [256] | Low | Low | Low | Low | Low | Low | Low |
| Chambers, 1990 [258] | Unclear | Low | Low | Low | Unclear | Low | Low |
| Dart, 1998 [236] | Low | Low | Low | Low | Low | Low | Low |
| Dart, 2002 [261] | Low | Low | Low | Low | Low | Low | Low |
| Dart, 2002 [259] | Low | Low | Low | Low | Low | Low | Low |
| Dart, 2002 [260] | Low | Low | Low | Low | Low | Low | Low |
| Dashefsky, 1988 [262] | Low | Low | Unclear | Low | Low | Low | Low |
| Gabrielli, 1992 [263] | Unclear | Low | Low | Low | Low | Low | Low |
| Hammoud, 2005 [264] | Low | Unclear | Low | Low | Low | Low | Low |

| Study | Risk of Bias | | | Applicability Concerns | | | |
|---|---|---|---|---|---|---|---|
| | Patient Selection | Index Test | Reference Standard | Flow and Timing | Patient Selection | Index Test | Reference Standard |
| Huter, 1990 [265] | Low | Unclear | High | Low | High | Low | Low |
| Kivikoski, 1990 [266] | Unclear | Low | Low | Low | Low | Low | Low |
| Mahony, 1985 [267] | Low | Low | Low | Low | Low | Low | Low |
| Mehta, 1999 [268] | Low | Low | Low | Low | Low | Low | Low |
| Nyberg, 1988 [239] | Unclear | Low | Low | Low | Low | High | Low |
| Nyberg, 1988 [240] | Unclear | Low | Low | Low | Low | Low | Low |
| Nyberg, 1991 [250] | Unclear | Low | Low | Low | Low | Low | Low |
| Rempen, 1988 [269] | Unclear | Low | Unclear | Low | Unclear | Low | Low |
| Romero, 1988 [270] | Unclear | Low | Unclear | Low | Low | Low | Low |
| Russell, 1993 [271] | Low | Low | Low | Low | Low | Low | Low |
| Sadek, 1995 [272] | Low | Low | Unclear | Low | Low | Low | Low |
| Shapiro, 1988 [273] | Unclear | Low | Unclear | Low | Low | Low | Low |
| Tongsong, 1992 [274] | Unclear | Low | Unclear | Low | Low | Low | Low |
| Tongsong, 1993[234] | Unclear | Low | Low | Low | Low | Low | Low |
| Weckstein, 1985 [235] | Low | Unclear | Unclear | Low | Low | High | Low |

***Table 6.4 Summary estimates for each ultrasonographic sign for predicting tubal ectopic pregnancy using only high quality studies***

| **Ultrasonographic Sign** | **Studies n [N]** | **Sensitivity (95% CI)(%)** | **Specificity (95% CI)(%)** | **LR+ (95% CI)** | **LR- (95% CI)** |
|---|---|---|---|---|---|
| Empty uterus | 5 (1341) | 76 (31-96) | 71 (53-84) | 2.6 (1.5-4.6) | 0.34 (0.093-1.4) |
| Adnexal mass | 4 (444) | 67 (43-85) | 95 (74-99) | 13 (1.7-102) | 0.34 (0.16-0.74) |
| Free fluid | 4 (805) | 52 (36-68) | 94 (74-99) | 8.1 (1.9-35) | 0.51 (0.36-0.72) |

*n[N], number of studies [number of women]*

# 7. Use Of The Double Decidual Sac Sign To Confirm An Intrauterine Pregnancy Prior To Ultrasonographic Visualization Of Embryonic Contents

*NB; An abridged version of this chapter has been published in the journal Ultrasound in Obstetrics and Gynaecology [275]*

## 7.1 Introduction

As discussed in chapter one, a gestation sac is the first ultrasonographic sign of an intrauterine pregnancy. It appears as a uniformly round, hypoechoic structure with an echogenic rim. Initially it does not contain any internal echoes and can therefore be difficult to differentiate from a 'pseudosac', that is, an endometrial fluid collection that occurs in up to 15% of ectopic pregnancies [18]. As demonstrated in chapter six, a pseudosac predicts an ectopic pregnancy with 94% specificity but only 5.5% sensitivity. It is clinically important not to confuse these two structures and hence several different ultrasonographic signs have been proposed to help differentiate between them prior to visualization of any embryonic contents.

The double decidual sac sign is one such sign. It was first described in the 1980s, as two concentric echogenic rings of tissue surrounding an intra-endometrial fluid collection that impress upon the endometrial stripe in an early intrauterine pregnancy [219]. The inner ring represents the decidua capsularis whilst the outer ring represents the decidua basalis. Conversely, in an ectopic pregnancy, the decidual reaction presents as only a single echogenic ring around the endometrial fluid collection [19].

The systematic review and meta-analysis described in chapter five reported that the double decidual sac sign predicted an intrauterine pregnancy with a

sensitivity of 82% and specificity of 97% [214]. Unfortunately, the quality of the studies included was generally poor: five of the six studies were undertaken in the 1980s [218, 219, 238, 241, 243]; two were retrospective [219, 241]; and only one utilized transvaginal ultrasound [220]. Hence the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy may have been under, or indeed over, estimated.

Ultrasound technology has advanced considerably over the last thirty years. Transvaginal ultrasound utilizes higher frequencies with better axial resolution that its transabdominal counterpart and this has revolutionized the ultrasonographic study of very early pregnancy. Not only does transvaginal ultrasound reliably identify normal and abnormal pregnancies at an earlier gestation than transabdominal ultrasound [14] but also threshold values and discriminatory sizes used to distinguish normal and abnormal pregnancies are smaller using higher frequency transvaginal ultrasound compared to lower frequency transvaginal ultrasound [16].

## 7.2 Aims

The aim of this study therefore was to determine the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy prior to visualization of embryonic contents using modern, high-resolution transvaginal ultrasound.

## 7.3 Hypotheses

- Intrauterine fluid collections that demonstrate the double decidual sac sign represent a true gestation sac
- Intrauterine fluid collections that do not demonstrate the double decidual sac sign represent a pseudosac

## 7.4 Methods

### Ethical Approval

Ethical approval was obtained from Nottingham 1 Research Ethics Committee (13-EM-0081) (Appendix 1) and informed written consent was obtained from

all participants (Appendix 4). The study was registered with www.clinicaltrials.gov (NCT02700789) and conducted following STARD guidelines (Standards for Reporting of Diagnostic Accuracy Studies) [276] (Appendix 5).

## Diagnostic Accuracy Studies

A diagnostic accuracy study evaluates the ability of one or more medical tests to correctly classify study participants as having a target condition. This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

Diagnostic accuracy studies are, like other clinical studies, at risk of bias due to shortcomings in design and conduct, and the results of a diagnostic accuracy study may not apply to other patient groups and settings. Readers of study reports need to be informed about study design and conduct in sufficient detail to judge the trustworthiness and applicability of study findings. The standards for reporting of diagnostic accuracy studies (STARD) statement was developed to improve the completely and transparency of reports of diagnostic accuracy studies. STARD contains a list of thirty essential items that can be used as a checklist by authors, reviewers and other readers, to ensure that a report of diagnostic accuracy study contains the necessary information (Appendix 5). The thirty items were identified by an international expert group of methodologists, researchers and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of the conclusions and recommendations. The items include information regarding the population, index test, reference standard and outcome of interest.

### Population

Diagnostic tests perform differently in different populations [277, 278] and it would generally be inappropriate, for example, to evaluate the performance of a test in a secondary care population when it is mainly used in primary care.

Both the frequency and severity of the target condition would be expected to be different in secondary care. It is therefore important to clearly define the population of interest. The ideal study sample for a diagnostic accuracy study is a consecutive or randomly selected series of patients in whom the target condition is suspected, or for screening studies, the target population. Since participant sampling methods are often poorly reported in test accuracy studies [279], using the sampling method as an inclusion/exclusion criteria is likely to result in a substantial reduction in available data. It is therefore more useful to consider the sampling method and/or its reporting as an aspect of study quality and to base the inclusion criteria relating to the population upon participant characteristics.

## Index Test

The test whose accuracy is being evaluated is called the index test. This can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient. Furthermore a study can include a number of different technologies addressing multiple target conditions or compare the performance of an alternative (replacement), less invasive or less costly diagnostic technology with that of the reference standard for the detection of a specified target condition. Evaluating the performance of a test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the reference standard. The traditional concept of diagnostic accuracy often implies the dichotomization of data into test results which are classified as positive (target condition present) or negative (target condition absent). Diagnostic accuracy studies therefore need to consider diagnostic thresholds (points at which results are classified as positive or negative) for each included index test.

## Reference Standard

The reference standard is usually the best available method for establishing the presence or absence of the target condition. It need not necessarily be the test used routinely in practice, and may include information which is not known for some time after the tests have been done. A diagnostic accuracy study is based upon a one-sided comparison between the results of the index

test and those of the reference standard. Any discrepancy is assumed to arise from error in the index test. Selection of the reference standard is therefore critical to the validity of a test accuracy study and the definition of the diagnostic threshold forms part of that reference standard. It is important to note that the assumption of 100% accuracy for the reference standard rarely holds true in practice. This represents a fundamental flaw in the diagnostic accuracy study design, since the index test can never perform better than the reference standard

Where several tests are available to diagnose a target condition, there is often no consensus about which test constitutes the reference standard. In such cases, a composite reference standard, which combines the results of several available tests to produce a better indicator of true disease status may be used [280]. A number of statistical methods have been proposed to estimate the performance of tests in the absence of a single accepted reference standard [281, 282].

There may be instances when it is deemed unethical to use an invasive procedure as a reference standard in a study [283]. In such cases, clinical follow-up and final diagnosis may sometimes be used as a reference standard. The length of follow-up should ideally be defined in advance. Studies using follow-up and clinical outcome in this way may be viewed as prognostic studies in that they are measuring the accuracy with which the test is able to predict a future event, rather than the accuracy with which it is able to determine current status. When using follow-up as the reference standard it may not always be obviously apparent what the clinical outcome/final diagnosis is. In such situations, the use of an adjudication committee may be employed. This is an independent group of experts that reviews all the available data in order to give a consensus opinion on what the most likely clinical outcome/final diagnosis is.

## Outcome of Interest

The outcome of interest may be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used

to estimate the sensitivity of the index test (the proportion of participants *with* the target condition who have a positive index test), and its specificity (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative predictive values of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical precision of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a test positivity cut-off. When multiple such cut-offs can be defined, authors can report a receiver-operating characteristic (ROC) curve, which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The area under the ROC curve informs in a single numerical value about the overall diagnostic accuracy of the index test.

## Ultrasound Scanning Techniques

Sound is a mechanical vibration distinguished by pitch (or frequency) and loudness. The velocity (v) of sound waves is constant (at 1540m/s) and is determined by the wavelength ($\lambda$) multiplied by the frequency (f). Higher frequencies therefore mean that the wavelengths are shorter because the velocity is constant. The frequency is defined as the number of vibrations (or cycles) per second and the unit of frequency is termed Hertz (Hz) (cycles per second). Ultrasound has a frequency above 20,000Hz (20kHz).

The transducers on ultrasound machines have different frequencies and are generally in the range of 2-10MHz. Higher frequency probes have narrower beam widths and give better resolution, which means they are more able to distinguish two targets close together. However, they have decreased penetration. Higher frequency probes should therefore be used to visualize near structures and lower frequency probes for deeper structures. For obstetric scanning, the abdominal probes generally vary from 3-5MHz and the transvaginal probes from 5-7.5MHz, the higher frequency giving better resolution for structures closer to the probe.

Materials exist which produce an electric current when pressure is applied to their surface. This is known as the piezoelectric effect and the inverse effect occurs when a current is applied to the material causing it to expand and contract. An ultrasound transducer for pulse-echo imaging houses one or more slabs of piezoelectric material. An ultrasound image is produced by the reflection of ultrasound. The ultrasound probe acts as both a transmitter and a receiver: ultrasound waves are transmitted by the electrical stimulation of crystals within the probe. These waves are then reflected from the surfaces of the tissue and organs being studied and the returning waves detected by receiving crystals within the probe. This process is repeated in many directions. The ultrasound machine then translates the reflected sound waves into a greyscale visual representation of the organs being studied. Selection of the appropriate probe frequency and style together with manipulation of machine settings including depth, gain, dynamic range, focal zone power level, frequency and harmonic imaging can be used to optimize image quality.

Present-day equipment employs real time imaging, as opposed to static, and this provides an immediate image and reveals movement of the structures being examined. Most of the probes are now electronically rather than mechanically driven and utilize an array system, which comprises scanned transducer elements mounted in line. Sets of elements are pulsed in sequence to produce a rectangular field of view. For obstetric use, curved array transducers are utilized which give a slightly wider field of view and are easier to manipulate on the lower abdomen in early pregnancy. Transvaginal probes work on the same principal.

## 3-Dimensional Ultrasound

Three-dimensional ultrasound allows a tissue volume to be examined and saved for subsequent offline analysis. The volume can be displayed in a series of sections simultaneously acquired in three orthogonal planes providing access to sections previously unachievable by the traditional two-dimensional method - the coronal plane that lies perpendicular to the transducer face.

The most commonly used technique utilizes phased array ultrasound transducers, which fan the ultrasound beam and acquire two-dimensional images at predefined intervals, described as the ‘swept-volume’ technique.

The initial image is the central scan of the volume of interest and the transducer scans from one margin to the other of the total volume to be acquired. The sweep time varies upon the depth, range, angle and image quality. Following acquisition, the three-dimensional ultrasound data can be saved, exported and manipulated or 'post processed'. This technique enables a much more accurate estimation of volume than two-dimensional ultrasound, which uses formulae based on geometric assumptions to derive measurements. Volume calculation methods include planar methods and the more accurate rotational method using Virtual Organ Computer-aided AnaLysis (VOCAL), which involves repeated manual delineation of the object of interest in the multi-planar display. The main disadvantage of three-dimensional ultrasound is its dependence on the image clarity of the two-dimensional images from which it is derived.

## Safety

Within the human species, no harmful bio-effects of ultrasound have been noticed since its introduction into the medical world in the 1940s. The prudent use of ultrasound is always advised, which means maintaining power levels and exposure time as low as reasonably achievable (the ALARA principle).

Bio-effects include the thermal effect of ultrasound and cavitation. The thermal effect refers to the rise in tissue temperature caused by ultrasound. Temperature elevations of 1.5°C above physiological levels can be made safely without limiting the examination. A temperature above 41°C with fetal use is considered harmful. Cavitation refers to the way that gas bubbles react within tissues under the influence of ultrasound. The thermal index (TI) is an indicator of the temperature elevation possible at a particular equipment setting and is defined as the ratio of the acoustic power emitted by the transducer to the acoustic power required to produce a 1°C rise at a particular equipment setting. With a TI of more than 0.7, the exposure to the embryo should be restricted to less than 60minutes. The mechanical index (MI) is an indicator of the likelihood of cavitational events and is defined as the maximal rarefaction pressure or the maximal negative pressure divided by the square root of the frequency. In general the MI should be below 1.9.

## Early Pregnancy Measurements

The following measurements/observations were recorded during each transvaginal ultrasound scan (Appendix 6).

### Endometrium

The uterus was scanned in the sagittal plane from one lateral aspect to the other. The endometrial stripe was measured at its maximum antero-posterior thickness along the longitudinal axis of the uterine body. This measurement included both the anterior and posterior endometrial layers. It was obtained by placing the calipers at the anterior and posterior uterine walls at the margins of the basal layers of the endometrium delineated by the echogenic interface between endometrium and inner myometrium. If the endometrium was focally thickened, the calipers were placed at the site of the maximal antero-posterior diameter. In the presence of intrauterine fluid, the thicknesses of the anterior and posterior endometrial layers was measured separately and summed [284].

### Gestation Sac

Measurements of the gestation sac (Figure 1.3a) were taken in three orthogonal planes (two in a sagittal plane and one in transverse) using the inner borders of the sac from which the mean diameter was calculated [216].

### Double Decidual Sac Sign

The double decidual sac sign (Figure 7.1) was defined as two concentric echogenic rings surrounding an intra-endometrial fluid collection [285].

### Yolk Sac

The yolk sac (Figure 1.3b), if present, was also measured, in three orthogonal planes. Unlike the gestation sac however, it was measured from the outer borders of the sac [216].

***Figure 7.1: An ultrasonogram depicting the double decidual sac sign***

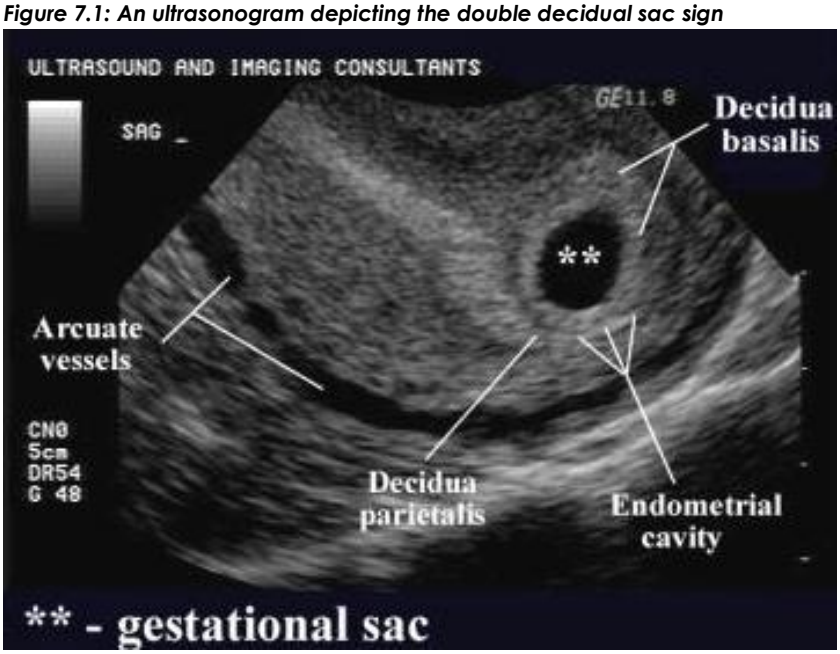



*Image taken from: www.fetalultrasound.com*

## Embryo or Fetus

If present, the embryo was also measured (Figure 1.3c). Initially, flexures have not developed and the cephalic and caudal ends of the embryo cannot be differentiated and hence it was the straight-line length (and not the crown-rump length) that was recorded [216].

## Adnexal Masses

The presence or absence of any adnexal masses as well as their location, size, appearance and mobility was also recorded.

## Free Fluid

The presence of any free fluid in the pouch of Douglas was determined by angling the transvaginal transducer towards the posterior part of the pelvis. If present, an estimate of the volume was made. The volume was considered small if it tracked less than one third up the posterior wall of the uterus on the long axis view. Fluid volume was classed as moderate if it tracked one third to two thirds of the way up the posterior wall of the uterus but was not free

flowing in the pelvis or abdomen. Fluid volume was considered large if it tracked greater than two thirds of the way up the posterior wall of the uterus or if it was free flowing in the pelvis or abdomen [286]. The appearance of any fluid (anechoic, echogenic or indeterminate) was also be noted.

## Study Design

### Population

Participants for the study were recruited prospectively from Nurture Fertility, Nottingham, United Kingdom between 1st January and 31st October 2015. Women were aged between 18 and 45 years of age and had undergone IVF/ICSI treatment using a standard long agonist or antagonist protocol depending on ovarian reserve tests [287]. The study was well advertised within the IVF unit using posters and patient information leaflets. Whenever possible, a clinician involved in recruitment was also present to discuss the study with women following their embryo transfer procedure. All women were invited to participate in the study. Women were excluded from the study if they had a negative urinary pregnancy test (performed 18 days after oocyte retrieval in a fresh cycle or 13-16 days after embryo transfer in a frozen embryo replacement cycle depending on the stage of embryo development at the time of transfer) or if, at the time of the index test, there was either no ultrasonographic evidence of an intrauterine fluid collection, or a yolk sac and/or fetal pole was clearly visible within the intrauterine fluid collection. Women were also excluded if no outcome data were available or if, following the reference standard, the final diagnosis was not known (for example resolving or persistent pregnancies of unknown location).

### Index Test

If the urinary pregnancy test was positive, an early ultrasound scan was scheduled for either 19 or 20 days after oocyte retrieval corresponding to a gestational age of 33 or 34 days. This range was specifically chosen to optimize the chances of a gestation sac being present but a yolk sac or fetal pole being absent [12, 20]. This early ultrasound scan was considered to be the index test. A single investigator with experience in early pregnancy ultrasound performed all of these early scans following standard operating procedures using a Voluson E8 machine with a high frequency (5-9MHz and 9-

12MHz) transvaginal probe. During the early scan the presence or absence of the following structures were recorded: an intrauterine fluid collection (defined as a uniformly round, hypoechoic structure with an echogenic rim); the double decidual sac sign (Figure 7.1); yolk sac (defined as a spherical, hyperechoic ring situated eccentrically within the gestation sac) (Figure 1.3b); and fetal pole (defined as a small linear echogenic structure adjacent to the yolk sac, on the side closest to the gestational sac)(Figure 1.3c). If more than one intrauterine fluid collection was visualized then each was considered as a separate entity. The findings from the index test were interpreted immediately and recorded separate to the main clinical notes. Images were stored for later assessment of inter- and intra-observer reliability (see chapter eight). Referral pathways to local Early Pregnancy Assessment Units were in place for any woman in whom this index test was strongly suggestive of an ectopic pregnancy, for example if there was an empty endometrial cavity and either an inhomogenous adnexal mass or an empty extra-uterine sac or a yolk sac or fetal pole with or without cardiac activity in an extra-uterine sac.

## Reference Standard

All women were scheduled to have a routine viability ultrasound scan at between six and seven weeks gestation (between eight and sixteen days after the index test) as per the fertility unit's standard practice. This viability scan was performed by an appropriately trained doctor or nurse following standard operating procedures using the same ultrasound equipment as the index test. This viability scan plus any subsequent clinical follow-up required i.e. if the diagnosis was not certain following the viability scan alone, constituted the reference standard. Clinical follow-up consisted of a repeat transvaginal ultrasound 7-10 days after the initial viability scan in cases of pregnancies of uncertain viability (defined as the presence of an intrauterine gestation sac of less than 25mm mean diameter with no obvious yolk sac or fetal pole or an intrauterine gestation sac containing a fetal pole of less than 7mm with no obvious fetal heart pulsations) and in cases of pregnancies of uncertain location (defined as no evidence of an intra- or extra-uterine pregnancy or retained products of conception on transvaginal ultrasound scan in the presence of a positive urinary pregnancy test), referral to a local Early Pregnancy Assessment Unit for monitoring of serial serum β-hCG levels and subsequent ultrasonography and possibly surgery where indicated according to departmental protocols until a definitive diagnosis could be made.

Interpretation of the reference standard was performed by an experienced gynaecologist without knowledge of the findings from the index test. Any uncertainty regarding the final diagnosis was dealt with by seeking the opinion of two other senior gynaecologists and gaining a consensus opinion.

### Outcome of Interest

Following the reference standard the possible diagnoses were either an intrauterine or ectopic pregnancy. Intrauterine pregnancies were further classified as either viable or non-viable. A viable intrauterine pregnancy was defined by ultrasonographic identification of an intrauterine gestation sac with a fetal pole of any length with demonstrable fetal heart pulsations. A non-viable intrauterine pregnancy was defined as either an empty intrauterine gestation sac with mean sac diameter greater than 25mm or an intrauterine gestation sac containing a fetal pole with crown rump length greater than 7mm with no demonstrable fetal heart pulsations or in the absence of a viable embryo, no significant increase in the growth of the gestation sac or length of the fetal pole on two ultrasound scans performed more than 7 days apart. Where women underwent surgical or medical management of miscarriage, histological confirmation of the products of conception was also obtained when possible. Ectopic pregnancies were confirmed either by direct visualization of an ectopic pregnancy during surgery with histological confirmation, or, in those managed medically with methotrexate or conservatively, unequivocal identification of an ectopic pregnancy on ultrasound scan. Ultrasonographic appearances indicative of an ectopic pregnancy included: an empty endometrial cavity with either an inhomogenous adnexal mass or an empty extra-uterine sac or a yolk sac or fetal pole with or without cardiac activity in an extra-uterine sac. Following the reference standard, any pregnancy which did not fall into one of these categories was subsequently excluded from the study. These included resolving or persisting pregnancies of unknown location.

## Statistical Analysis

The sensitivity and specificity of the double decidual sac sign for predicting an intrauterine pregnancy were estimated following cross tabulation of the index test results against those of the reference standard. The overall diagnostic accuracy, as well as positive and negative likelihood ratios and predictive

values, were also calculated (Table 7.1 and 7.2). The pre-test probability was estimated to be the prevalence of the disease in our sample. The post-test probabilities were calculated using a likelihood ratio nomogram and confirmed via an online diagnostic test calculator that determines the post-test probability of disease given the pre-test probability and the test characteristics (http://araw.mede.uic.edu/cgi-bin/testcalc.pl). Results are expressed as percentages and 95% confidence intervals are also given.

***Table 7.1: The 2x2 contingency table***

| | | **Reference Standard** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Index Test** | **Positive** | True Positive (TP) | False Positive (FP) |
| | **Negative** | False Negative (FN) | True Negative (TN) |

***Table 7.2: Commonly used measures of test performance and how they are calculated using a 2x2 contingency table***

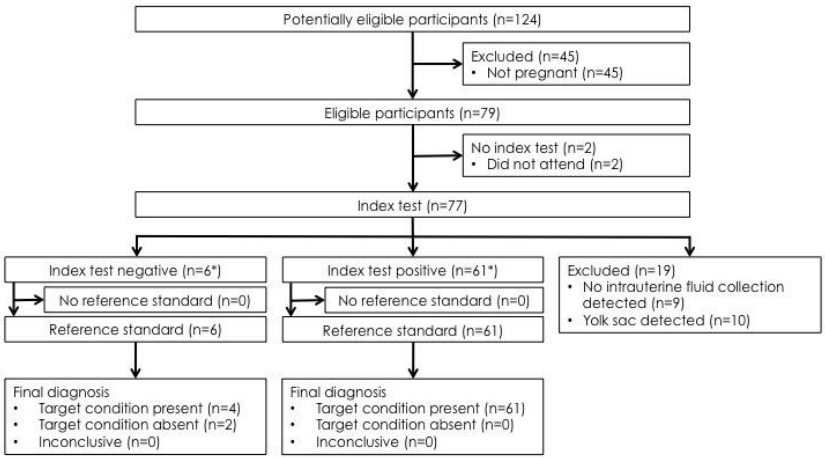
| **Measure** | **Definition** | **Calculation** |
|---|---|---|
| Sensitivity | The proportion of people with the target condition who have a positive test result | $\frac{TP}{TP+FN}$ |
| Specificity | The proportion of people without the target condition who have a negative test result | $\frac{TN}{TN+FP}$ |
| Overall Accuracy | The proportion of people correctly classified by the test | $\frac{TP+TN}{TP+FN+FP+TN}$ |
| Positive Predictive Value | The probability of disease amongst those with a positive result | $\frac{TP}{TP+FP}$ |
| Negative Predictive Value | The probability of non-disease among persons with a negative test result | $\frac{TN}{TN+FN}$ |
| Positive Likelihood Ratio | The probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive | $\frac{Sensitivity}{1\text{-}Specificity}$ |
| Negative Likelihood Ratio | The probability of a person who has the disease testing negative divided by the probability of a person who does not have the disease testing negative | $\frac{1\text{-}Sensitivity}{Specificity}$ |

## Sample Size Calculation

Based on the results of the systematic review and meta-analysis in chapter five, the estimated sensitivity and specificity of the double decidual sac sign to predict an intrauterine pregnancy is 82% and 97% respectively. If we wish to estimate the sensitivity and specificity to within 10% with 95% certainty, then our sample size is at least 69, to include 57 true positive cases and twelve true negative cases [288].

# 7.5 Results

***Figure 7.2: Flow of participants through the study***




*9 women had 2 intrauterine fluid collections detected during the index test

Between 1st January and 31st October 2015, 620 IVF/ICSI cycles were undertaken within the unit. Of these, 124 (20%) women agreed to participate in the study. In addition to these, a further six women were approached at the time of embryo transfer and declined to participate in the study due to various reasons, namely work commitments (n=3), reluctance to have a transvaginal ultrasound (n=2) and distance to travel to the clinic (n=1). 45 (36%) of the 124 women were subsequently excluded as they had a negative urinary

pregnancy test. Of the 79 women who had a positive pregnancy test, two (2.5%) did not attend for the index test and nine (11%) of those that did attend did not have an intrauterine fluid collection present on transvaginal ultrasound and were therefore excluded. 77 intrauterine fluid collections were observed in the remaining 68 women (nine of the women had two intrauterine fluid collections detected). Ten (7.7%) of the 77 intrauterine fluid collections had a definite yolk sac visible and were therefore excluded further from the study leaving 67 intrauterine fluid collections in the study (Figure 7.2). The baseline characteristics of the study participants are illustrated in Table 7.3. These were not significantly different from the baseline characteristics of the general population attending the IVF unit during the same time period.

***Table 7.3: Baseline characteristics of study participants (values refer to mean ± standard deviation unless otherwise indicated)***

| | |
|---|---|
| Age (years) | 34±6.3 |
| Gravidity | 1.0±1.1 |
| Parity | 0.3±0.5 |
| BMI (kg/m$^2$) | 24±4.5 |
| Ethnicity (%) | |
| White British | 84 |
| Asian | 10 |
| Other White | 3.9 |
| African/Black British/Mixed | 1.3 |

Of the 67 intrauterine fluid collections detected during the index test, 61 displayed the double decidual sac sign. Of these, all 61 were subsequently demonstrated to have an intrauterine pregnancy on follow-up. Of the six intrauterine fluid collections that did not display the double decidual sac sign, four were subsequently proven to have an intrauterine pregnancy and two were found to have an ectopic pregnancy (Table 7.4).

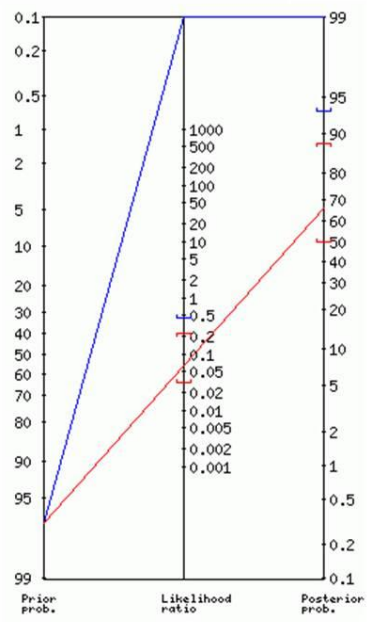
***Table 7.4: A 2x2 contingency table to show the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy using high-resolution transvaginal ultrasound***

| | | **Reference Standard** | | ***Totals*** |
|---|---|---|---|---|
| | | **IUP present** | **IUP absent** | |
| **Index Test** | **DDSS present** | 61 | 0 | 61 |
| | **DDSS absent** | 4 | 2 | 6 |
| ***Totals*** | | 65 | 2 | 67 |

The double decidual sac sign was therefore found to have a sensitivity of 94% (95% CI, 85-98%), specificity of 100% (95% CI, 16-100%) and overall diagnostic accuracy of 94% (95% CI, 88-100%) for predicting an intrauterine pregnancy. The positive and negative predictive values are 100% (95% CI, 94-100%) and 33% (95% CI, 4.3-78%) respectively whilst the positive likelihood ratio was infinite and the negative likelihood ratio was 0.06 (95% CI, 0.02-0.16) (Figure 7.3). In our study, the prevalence of an intrauterine pregnancy was 97% but if the double decidual sac sign was present the probability of an intrauterine pregnancy was 100% compared with 66% if the double decidual sac sign was absent (Table 7.5).

Of the 61 intrauterine pregnancies that demonstrated the double decidual sac sign during the index scan, 58 (95%) were viable. Of the four intrauterine pregnancies that did not display the double decidual sac sign during the index scan, only two (50%) were viable.

***Figure 7.3: Likelihood ratio nomogram for determining post-test probabilities***




Two ectopic pregnancies were identified at the time of the reference standard. Both women were asymptomatic but had ultrasonographic

evidence of an empty endometrial cavity and an inhomogenous adnexal mass. They were referred to the local Early Pregnancy Assessment Unit where the diagnosis was confirmed and one chose to have medical management with Methotrexate, which was successful and the other underwent laparoscopic salpingectomy, which was uncomplicated, and histological examination of the specimen confirmed the diagnosis.

Of the nine women that attended for the index scan but did not have an intrauterine fluid collection present, seven (78%) were subsequently proven to have a viable intrauterine pregnancy, one (11%) had a non-viable intrauterine pregnancy and one (11%) was diagnosed with an ectopic pregnancy at approximately 7 weeks' gestation.

All ten intrauterine fluid collections that contained yolk sacs at the time of the index test were subsequently proven to be viable intrauterine pregnancies.

No adverse events from performing the index test or reference standard were reported.

## 7.6 Discussion

Our results demonstrate that the probability of an intrauterine pregnancy when the double decidual sac sign was present was 100%. The probability of having an ectopic pregnancy when the double decidual sac sign was absent was 33%. Accordingly, only our first hypothesis is accepted. Using modern high-resolution transvaginal ultrasound, the presence or absence of the double decidual sac sign correctly located 94% of pregnancies. The double decidual sac sign was present in 94% of intrauterine pregnancies and absent in both the ectopic pregnancies. In this study, the prevalence of an intrauterine pregnancy was 97%, but if the double decidual sac sign was present the probability of an intrauterine pregnancy was 100% compared with 66% (range, 49-88%) if the double decidual sac sign was absent. These findings suggest that the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy reported in the systematic review and meta-analysis described in chapter four was underestimated.

One of the studies included in the systematic review suggested that whilst the overall sensitivity and specificity of the double decidual sac sign for distinguishing between an intrauterine pregnancy and an ectopic pregnancy was 64% and 100% respectively, in cases where embryonic contents were not yet visible the specificity dropped to 53%, rendering it a useless test in clinical practice [220]. Inherent with a diagnostic test to confirm an intrauterine pregnancy, is the absolute requirement for false positive and false negative rates of zero. This is because of the clinical implications of labeling a pregnancy as intrauterine with subsequently little or no follow-up, when actually it is ectopic and any delay in diagnosis could be detrimental or labeling a pregnancy as an ectopic and initiating treatment when actually it is intrauterine.

In our study however, women were excluded if embryonic contents were visible, hence our results suggest that the double decidual sac sign is actually very valuable in clinical practice: women who present with abdominal pain and/or vaginal bleeding in early pregnancy who have ultrasonographic evidence of an empty sac which demonstrates the double decidual sac sign can, according to our results, be managed following pregnancy of uncertain viability protocols rather than, what has become standard practice in many units, pregnancy of unknown location protocols and consequently not use valuable resources inappropriately nor generate undue anxiety for women.

It is important to note that despite our specificity of 100%, a thorough assessment of the adnexa is still required even when an intrauterine fluid collection does demonstrate the double decidual sac sign because of the possibility, albeit remote, of a heterotopic pregnancy.

Furthermore, even though the sensitivity of the double decidual sac sign to predict an intrauterine pregnancy in this study was extremely high, it was not 100% and therefore intrauterine fluid collections that do not demonstrate the double decidual sac sign cannot be labeled as pseudosacs associated with ectopic pregnancies. Such ultrasonographic findings require follow-up until pregnancy location can be conclusively determined.

The main limitation of our study is that it only included two true negative cases. Whilst neither of these demonstrated the double decidual sac sign resulting in a specificity of 100%, as there were only two, the confidence intervals are

wide. If more women were recruited, more pseudosacs might have been identified but accepting that (1) the incidence of an ectopic pregnancy is relatively low, even following IVF/ICSI treatment [289], and (2) pseudosacs occur in less than 15% of ectopic pregnancies [18], approximately 1700 women would need to participate in the study to identify the required number of pseudosacs from our sample size calculation. Taking into consideration the throughput and success rates of our IVF unit and the recruitment rate observed during our study, this would take over a decade to achieve. A multi-center approach would generate greater numbers more quickly but this would incorporate variations in practice that could affect the validity of our results in other ways.

Another limitation of our study is that all scans were performed and interpreted by a single investigator. Although it would have been ideal if more than one operator performed each ultrasound, this was not incorporated into the study protocol for two reasons. Firstly, many of the scans were performed out-of-hours for the convenience of the participants and logistically it would have been difficult to coordinate these appointments if multiple investigators were simultaneously required. Secondly, following a focus group discussion with potential participants during the planning phase of the study, it was discovered that many women felt apprehensive about having a transvaginal ultrasound when there was no clinical indication. Whilst other women said that they would be willing to undergo one additional transvaginal ultrasound as part of a research study, most women were reluctant to have more than one. However, the inter-observer variability associated with ultrasonographic detection of the double decidual sac sign is an important consideration if it is to be used in clinical practice and this is the focus a subsequent study described in the following chapter.

Finally, all participants in the study were pregnant as a consequence of IVF/ICSI treatment. This was so that we could be certain of the exact gestational age of the pregnancy and schedule the index test to be performed at an appropriate time. From the literature, a gestation sac first appears at approximately 28 days gestation [12] but since the women in our study had been instructed to perform their urinary pregnancy test 18 days after egg collection (corresponding to a gestational age of 32) as part of the fertility units standard practice, we could not perform the index test any earlier than this. We also wanted to perform the index test at a time when there was

a low probability of identifying embryonic contents, which is thought to be around day 35 [20]. It is possible that the accuracy of the double decidual sac sign to predict an intrauterine pregnancy may be related to gestational age and perhaps intrauterine fluid collections identified ultrasonographically prior to day 33 or after day 34 may be more or less likely to demonstrate the double decidual sac sign than those identified in our study. Scheduling the index test for this specific time would have been difficult if using spontaneous conceptions. However, there is no theoretical reason to suggest that the diagnostic accuracy of the double decidual sac sign to predict an intrauterine pregnancy would be any different in spontaneous pregnancies compared to assisted conceptions.

## 7.7 Conclusion

Using modern high-resolution transvaginal ultrasound, the presence of the double decidual sac sign can be used to accurately confirm an intrauterine pregnancy prior to visualization of embryonic contents. The absence of it however does not preclude an intrauterine pregnancy.

*Table 7.5: Summary estimates of the double decidual sac sign to predict an intrauterine pregnancy*

| | Sensitivity (95% CI)(%) | Specificity (95% CI)(%) | PPV (95% CI)(%) | NPV (95% CI)(%) | LR+ (95% CI) | LR- (95% CI) | Pre- and post-test probability | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Pre-test | Post-test if test positive | Post-test if test negative |
| **DDSS** | 94 (85-98) | 100 (16-100) | 100 (94-100) | 33 (4.3-78) | ∞ | 0.06 (0.02-0.16) | 97 | 100 (93-100) | 66 (49-88) |

# 8. Inter- and Intra-Observer Reliability Associated with Ultrasonographic Visualization of the Double Decidual Sac Sign

## 8.1 Introduction

The ability to accurately differentiate a true gestation sac from a pseudosac prior to ultrasonographic visualization of embryonic contents would be very useful clinically. As described in the preceding chapters, several different ultrasonographic signs have been proposed to help discriminate between the two structures [218-220]. The double decidual sac sign is one such sign. It was first described in the early 1980s as two concentric echogenic rings of tissue surrounding an intra-endometrial fluid collection that impress upon the endometrial stripe in an early intrauterine pregnancy [219]. Conversely, in an ectopic pregnancy, the decidual reaction presents as only a single echogenic ring around the endometrial fluid collection [19].

Following the conclusions made in chapter five, we undertook a study to determine the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy prior to visualization of embryonic contents using modern, high-resolution transvaginal ultrasound. This study, reported in chapter seven, demonstrated that the double decidual sac sign had a sensitivity of 94%, specificity of 100% and overall diagnostic accuracy of 94% for predicting an intrauterine pregnancy [275] and concluded that, using high resolution transvaginal ultrasound, the presence of the double decidual sac sign could be used to accurately confirm an intrauterine pregnancy prior to visualization of embryonic contents. These results could potentially revolutionize the management of diagnostic uncertainties in early pregnancy by rationalising the follow-up of women with ultrasonographic evidence of an 'empty sac'.

For various reasons, described in the previous chapter, all of the scans undertaken in the diagnostic accuracy study were performed and interpreted by a single investigator. For the double decidual sac sign to be rendered clinically useful, multiple different observers assessing the same ultrasonogram independently should agree about whether the sign is present or not. As the presence or absence of the double decidual sac sign informs further management decisions, and the consequences of a false positive diagnosis are to overlook an ectopic pregnancy (which could have catastrophic clinical consequences), a high degree of reliability is essential. The more objective and unambiguous the criteria are for a sign, the more likely it is that it will have high inter- and intra-observer agreement and vice versa.

## 8.2 Aims

The aim of this study therefore was to determine the inter- and intra-observer reliability associated with ultrasonographic identification of the double decidual sac sign as well as the diagnostic interpretation and management of 'empty sacs'.

## 8.3 Methods

### Study Design

All ultrasonographic images were collected prospectively as part of the study attempting to determine the diagnostic accuracy of the double decidual sac sign for predicting an intrauterine pregnancy prior to ultrasonographic visualization of embryonic contents as described in chapter seven.

For the assessment of inter-observer reliability, two-dimensional images from a selection of 25 of the study participants were distributed to eighteen individuals with experience in early pregnancy ultrasound to assess. In an attempt to reflect normal clinical practice as far as possible we surveyed observers from different hospitals with different levels of experience. The observers were colleagues of the author who had expressed an interest in participating in the study. Both clinicians and ultrasonographers were included in the study because in the United Kingdom a large proportion of

early pregnancy ultrasound scans are performed by ultrasonographers. The experience of each observer was assessed both subjectively, by asking what level of experience they had i.e. beginner, intermediate or advanced, and objectively, by asking how many years' experience they had.

Each observer was provided with an electronic link to a secure website (Dropbox Inc, San Francisco, USA) where they could view the ultrasound images (Appendix 7). The author selected several images from 25 of the original study participants which they believed depicted intrauterine fluid collections that both did and did not exhibit the double decidual sac sign. Determining the exact sample size and method of selection was difficult. We wanted to ensure a range of cases, including borderline ones, were incorporated into the study, and this would not have been possible by using consecutive cases. Neither were they chosen at random as we needed to ensure that image quality was for the most part reasonable and that we included intrauterine fluid collections that both did, and did not, demonstrate the double decidual sac sign. For practical purposes it was thought that twenty-five cases was an acceptable number for observers to assess. For each of the 25 cases selected, between three and four two-dimensional images were provided for assessment. The images given for review included standard transverse and sagittal planes. The final diagnoses were not known at the time of selection of cases/images for inclusion in the study. All datasets were anonymized to comply with research ethics approval. Observers were informed that all images were taken from asymptomatic women who were between four and five weeks pregnant following IVF treatment. No other clinical history was provided.

Each observer independently performed off-line analysis of the ultrasonographic images and was asked to complete a standardized assessment form. For each of the 25 cases, the observers were asked three questions: (1) is the double decidual sac sign present; (2) what does the structure represent; and (3) what initial follow-up would you recommend? Available options from a drop down menu included: 'yes' and 'no/uncertain' for question one; 'gestation sac' and 'pseudosac/uncertain' for question two; and 're-scan in 7-10 days' and 'serial hCGs' for question three. Although a repeat ultrasound is advocated in addition to serial hCGs in the management of pregnancies of unknown location, the initial management consists of serial hCGs and then, depending on both the trend and the absolute levels, a

repeat ultrasound may be undertaken at some point in the future. Question 3 was primarily concerned with the initial management plan, hence the two options available.

Following the initial assessment of inter-observer reliability, six of the observers who were locally available were provided with training. This training included both theoretical and practical components, static image review and live scanning. Observers had a group tutorial of approximately two hours duration and an individual supervised scanning session in a dedicated early pregnancy setting. To determine the effect of training on the inter-observer reliability, these six observers were asked to assess the same set of images two weeks after completion of training. A further four weeks later, in an attempt to assess the intra-observer reliability, the six local observers were asked to assess the same set of images one third and final time. The order of the cases was randomly manipulated prior to the second and third assessments in an attempt to eliminate recall bias.

## Statistical Analysis

***Table 8.1: Recommended descriptions of numerical kappa (κ) values [290]***

| **κ Statistic** | **Level of Agreement** |
|---|---|
| <0 | Less than chance |
| 0.01-0.20 | Slight |
| 0.21-0.40 | Fair |
| 0.41-0.60 | Moderate |
| 0.61-0.80 | Substantial |
| 0.81-0.99 | Almost perfect |

For the determination of inter- and intra-observer reliability, statistical analysis was performed using Mini Tab 17 (Coventry, UK). To describe inter-observer reliability, the Fleiss-Kappa statistic was used. Fleiss-Kappa is a generic term for several similar measures of agreement used with categorical data, which reflect the classification of objects into different groups or categories. Typically it is used in assessing the degree to which two or more raters, examining the same data, agree when it comes to assigning the data into categories. In this study, the Fleiss-Kappa statistic was used to assess the extent by which the observers vary in their interpretation of the ultrasonographic images provided to them. Complete agreement corresponds to K=1, and lack of agreement

corresponds to K=0. A negative value of kappa would mean negative agreement, usually caused by the rater's tendency toward avoiding a grade assigned to an object by others. Recommended descriptions of numerical K values are illustrated in Table 8.1[290].

## 8.4 Results

Fifteen (83%) of the eighteen observers asked completed the survey. Within the UK, observers were located in London, Nottingham, Cambridge, Winchester, Milton Keynes, Plymouth, Salisbury and High Wycombe. Observers from Denmark and Australia also completed the survey. Twelve (80%) observers were clinicians with a special interest in ultrasonography, two (13%) were ultrasonographers and one (6.7%) was a nurse ultrasonographer. The number of years' experience with early pregnancy ultrasound ranged from between one and fifteen. Three (20%) observers subjectively rated themselves as having a 'beginner' level of experience with early pregnancy ultrasound, five (33%) as 'intermediate' and seven (47%) as 'advanced'. Fourteen (93%) observers had heard of the double decidual sac sign prior to completing the survey but only eight (53%) utilized it in clinical practice.

Overall, there was significant agreement amongst the fifteen observers for question one (is the double decidual sac sign present?) but the level of agreement was only 'fair' ($K=0.25$, $p<0.01$). Although the level of agreement amongst the two ultrasonographer was only 'slight' ($K=0.17$, $p>0.05$) and amongst the twelve clinicians it was 'fair' ($K=0.28$, $p<0.01$), the difference between the two groups was not statistically significant, as reflected by the overlapping confidence intervals. The eight observers who utilized the double decidual sac sign in their clinical practice had 'fair' levels of agreement ($K=0.34$, $p<0.01$), which was significantly higher than the seven observers who did not ($K=0.13$, $p<0.01$). Despite the fact that there was a 'moderate' level of agreement amongst the five observers who subjectively rated their ultrasound skills as 'intermediate' ($K=0.41$, $p<0.01$) and only 'fair' agreement amongst the seven observers who rated their skills as 'advanced' ($K=0.23$, $p<0.01$), the difference between the two groups was not statistically significant. Both groups however had significantly higher levels of agreement than the three observers who rated themselves as a 'beginner'. The levels of agreement were significantly higher in observers with five or more years' experience

(K=0.24, p<0.01) compared to those with less than five years' experience (K=0.16, p<0.01). The highest levels of agreement were witnessed in the three observers who rated themselves as 'advanced' and had five or more years' experience (K=0.41, p<0.01) (Table 8.2).

Overall, there was significant agreement amongst the fifteen observers for question two (what does the structure represent?) but again the level of agreement was only 'fair' (K=0.33, p<0.01). Levels of agreement were 'fair' regardless of job description, use in clinical practice, subjective level of experience and number of years' experience. The highest levels of agreement were witnessed in the eight observers who utilized the double decidual sac sign in their clinical practice (K=0.36, p<0.01) (Table 8.3).

Interpretation of the ultrasonographic findings was in-keeping with the observers impression of whether or not the double decidual sac sign was present in 82% of cases but in 15 of the 187 (8.0%) instances when the double decidual sac sign was considered to be present, the structure was thought to represent a pseudosac and in 52 of the 188 (28%) instances when the double decidual sac sign was considered to be absent, it was thought that the structure represented a gestation sac.

Overall, there was significant agreement amongst the fifteen observers for question three (what follow-up would you recommend?) but again the level of agreement was only 'fair' (K=0.21, p<0.01). There was no significant difference between the levels of agreement amongst the twelve clinicians (K=0.24, p<0.01) and the two ultrasonographers (K=0.40, p<0.05), which were all 'fair'. The eight observers who utilized the double decidual sac sign in their clinical practice also had 'fair' levels of agreement (K=0.32, p<0.01), which was significantly higher than the seven observers who did not (K=0.085, p<0.05). Although there was a trend towards increasing levels of agreement with increasing skill level, the difference between those with beginner, intermediate and advanced ultrasound skills was not statistically significant. There was 'slight' agreement amongst observers with less than five years' experience (K=0.17, p<0.01) and 'fair' agreement amongst those with five or more years' experience (K=0.26, p<0.01) but again there was no statistically significant difference between the two groups (Table 8.4).

The recommended follow-up was considered appropriate in 91% of cases based on the observers interpretation of what they thought the structure represented (irrespective of whether they thought the double decidual sac sign was present or not) but in 28 of the 224 (13%) presumed gestation sacs, serial hCGs were recommended and in six of the 151 (4.0%) presumed pseudosacs, a re-scan in 7-10 days was advised.

Amongst the six local observers, there were three clinicians with a special interest in ultrasonography, two ultrasonographers and one nurse ultrasonographer. The number of years' experience with early pregnancy ultrasound ranged from between one and eight. Three observers subjectively rated themselves as having a 'beginner' level of experience with early pregnancy ultrasound, two as 'intermediate' and one as 'advanced'. All six had heard of the double decidual sac sign prior to being asked to participate in the study but only two utilized it in clinical practice. Prior to training, there was significant ($p<0.01$) agreement amongst the six observers for questions 1, 2 and 3 but the level of agreement was only 'fair' for questions 1 and 2 and 'slight' for question 3 reflected by kappa scores of 0.23, 0.37 and 0.12 respectively (Table 8.5). After training, there was 'substantial' agreement for questions 1 and 3 and 'moderate' for question 2 reflected by kappa scores of 0.70, 0.63 and 0.53. This improvement was statistically significant for questions 1 and 3 (Table 8.5).

After training, the level of agreement within the six local observers when reassessing the same set of still images ranged from 'substantial' (observer 4, question 3, K=0.65) to 'almost perfect' (observer 1, question 1, K=0.92) (Table 8.6). Consistency was greatest for question two as five of the six observers had 'almost perfect' levels of agreement. Least consistency occurred for question 3 but at worst there was still a 'substantial' level of agreement (observer 4, K=0.65) and at best an 'almost perfect' level of agreement (observer 1, K=0.84).

## 8.5 Discussion

This aim of this study was to determine the inter- and intra-observer reliability associated with ultrasonographic identification of the double decidual sac sign as well as the diagnostic interpretation and management of 'empty

sacs'. Although similar studies exist [246], this is the first study to utilize multiple observers from different institutions with varying levels of experience. Previous studies did not consider intra-observer reliability or the effect of training. Our study is also the first to utilize images that were obtained prospectively, using high-resolution transvaginal ultrasound, with the specific intention of demonstrating the double decidual sac sign.

Although our results suggest that agreement between observers when determining the presence or absence of the double decidual sac sign were generally only 'fair', the fact that it was significantly higher amongst observers who utilized the sign in clinical practice and were more experienced implies that the double decidual sac sign could be useful after training. Indeed, our secondary analysis, incorporating six of the original observers, proves that training significantly improves the levels of agreement.

As described in chapter five, due to previous studies reporting variable degrees of accuracy, with sensitivities between 64-95% [220, 238] and specificities between 60-100% [219, 220, 238], the clinical utility of the double decidual sac sign has been questioned [220], perhaps negating the need for it to be included in ultrasound training. Current clinical practice advocates waiting until a yolk sac and/or fetal pole is visualized ultrasonographically before stating that a pregnancy is definitely intrauterine. This is because the yolk sac is a reliable indicator of an intrauterine pregnancy with a positive predictive value of 100% [243]. However, since we have demonstrated in chapter seven that the double decidual sac sign also has a specificity and positive predictive value of 100% [275], an intrauterine pregnancy can be accurately diagnosed prior to ultrasonographic visualization of embryonic contents, and hence the double decidual sac sign should now be incorporated into standard ultrasound training in early pregnancy because of the potential clinical benefit that could be gained from it.

Whilst variation between observers can introduce bias and make it difficult to interpret results, variation within an observer is random and, as such, does not cause bias in itself but affects precision, which is also an important consideration if the double decidual sac sign is to be used clinically. The high levels of intra-observer reliability found in our study adds further value to its use in clinical practice.

The main limitation of our study is that static ultrasound images were used. Although multiple images were provided, still images don't adequately reflect the full assessment made when performing the scan. Although it would have been preferable to have had each observer perform each scan independently, this was not feasible for reasons described in the previous chapter. A cineloop recording in addition to the still images would have enhanced the observers assessment in lieu of performing the scan themselves which may have improved the levels of agreement reported.

Additionally, the observers were not blinded to the clinical information and this may have altered their responses, particularly to questions two and three. However, the clinical history was the same for each case and suitably vague.

Finally, the Fleiss Kappa statistic was used to determine reliability and although the appropriate test to measure agreement between more than two observers when responses fall into categories, there are several problems inherent with both its use and the use of the recommended descriptions of numerical K values leading to an under, or indeed over, estimation of reliability. Firstly, it depends on the true proportions of subjects in each category, secondly it assumes the sample is representative of the underlying population and thirdly, it ignores the influence experimental conditions have on the magnitude of estimated agreement coefficients [291-293]. Another problem with the Fleiss Kappa statistic is the interpretation of what a statistically significant test means in clinical practice. In a practical situation in which ratings are compared across clinicians, agreement will usually be better than that expected by chance and specifying a zero value for kappa in the null hypothesis i.e. no agreement, is therefore not very meaningful [294, 295]. Thus, the value in the null hypothesis should usually be set at a level below which would be considered clinically unacceptable. This will of course, depend on the clinical context. In the context of the double decidual sac sign, the minimal acceptable level of agreement on this categorization should be set fairly high, for example, greater than 0.60, so that decisions on patient management are made with a high degree of consistency.

## 8.6 Conclusion

In conclusion, using modern high resolution transvaginal ultrasound, the double decidual sac sign is not only accurate in its ability to confirm an intrauterine pregnancy prior to visualization of embryonic contents, but also, as demonstrated by this study, has the potential (after training) to be both reliable and precise, making it a very useful ultrasonographic sign in clinical practice.

*Table 8.2: Inter-observer reliability for question 1 'Is the double decidual sac sign present?'*

| | | n | κ | Level of Agreement | SE | 95%CI | p-value |
|---|---|---|---|---|---|---|---|
| **ALL** | | 15 | 0.25 | Fair | 0.20 | -0.13-0.63 | <0.01 |
| **JOB DESCRIPTION** | **Clinician** | 12 | 0.28 | Fair | 0.025 | 0.23-0.33 | <0.01 |
| | **Ultrasonographer** | 2 | 0.17 | Slight | 0.20 | -0.23-0.56 | ns |
| **KNOWLEDGE** | | 14 | 0.26 | Fair | 0.021 | 0.22-0.30 | <0.01 |
| **PRACTICE** | **Yes** | 8 | 0.34 | Fair | 0.038 | 0.27-0.4182 | <0.01 |
| | **No** | 7 | 0.13 | Slight | 0.044 | 0.043-0.2135 | <0.01 |
| **LEVEL OF EXPERIENCE** | **Beginner** | 3 | -0.036 | Less than Chance | 0.12 | -0.26-0.1907 | ns |
| | **Intermediate** | 5 | 0.41 | Moderate | 0.063 | 0.28-0.5301 | <0.01 |
| | **Advanced** | 7 | 0.23 | Fair | 0.044 | 0.15-0.3164 | <0.01 |
| **YEARS EXPERIENCE** | **< 5** | 5 | 0.16 | Slight | 0.63 | -1.1-1.4 | <0.01 |
| | **≥ 5** | 10 | 0.24 | Fair | 0.030 | 0.18-0.30 | <0.01 |
| **ADVANCED AND ≥5 YRS EXPERIENCE** | | 3 | 0.41 | Moderate | 0.12 | 0.18-0.63 | <0.01 |

*Table 8.3: Inter-observer reliability for question 2 'What does the structure represent?'*

| | | n | κ | Level of Agreement | SE | 95%CI | p-value |
|---|---|---|---|---|---|---|---|
| **ALL** | | 15 | 0.33 | Fair | 0.20 | -0.049-0.72 | <0.01 |
| **JOB DESCRIPTION** | **Clinician** | 12 | 0.31 | Fair | 0.025 | 0.26-0.36 | <0.01 |
| | **Ultrasonographer** | 2 | 0.35 | Fair | 0.20 | -0.041-0.74 | <0.05 |
| **KNOWLEDGE** | | 14 | 0.34 | Fair | 0.021 | 0.30-0.38 | <0.01 |
| **PRACTICE** | **Yes** | 8 | 0.36 | Fair | 0.038 | 0.29-0.44 | <0.01 |
| | **No** | 7 | 0.33 | Fair | 0.044 | 0.24-0.41 | <0.01 |
| **LEVEL OF EXPERIENCE** | **Beginner** | 3 | 0.30 | Fair | 0.12 | 0.077-0.53 | <0.01 |
| | **Intermediate** | 5 | 0.33 | Fair | 0.063 | 0.21-0.46 | <0.01 |
| | **Advanced** | 7 | 0.34 | Fair | 0.044 | 0.26-0.43 | <0.01 |
| **YEARS EXPERIENCE** | **< 5** | 5 | 0.26 | Fair | 0.63 | -0.98-1.5 | <0.01 |
| | **≥ 5** | 10 | 0.34 | Fair | 0.030 | 0.28-0.40 | <0.01 |
| **ADVANCED AND ≥5 YRS EXPERIENCE** | | 3 | 0.22 | Fair | 0.12 | -0.0041-0.45 | <0.05 |

*Table 8.4: Inter-observer reliability for question 3 'What follow-up would you recommend?'*

| | | n | κ | Level of Agreement | SE | 95%CI | p-value |
|---|---|---|---|---|---|---|---|
| **ALL** | | 15 | 0.21 | Fair | 0.20 | -0.17-0.60 | *<0.01* |
| **JOB DESCRIPTION** | **Clinician** | 12 | 0.24 | Fair | 0.025 | 0.19-0.28 | *<0.01* |
| | **Ultrasonographer** | 2 | 0.40 | Fair | 0.20 | 0.013-0.80 | *<0.05* |
| **KNOWLEDGE** | | 14 | 0.24 | Fair | 0.021 | 0.20-0.28 | *<0.01* |
| **PRACTICE** | **Yes** | 8 | 0.32 | Fair | 0.038 | 0.24-0.39 | *<0.01* |
| | **No** | 7 | 0.085 | Slight | 0.044 | -0.0005-0.17 | *<0.05* |
| **LEVEL OF EXPERIENCE** | **Beginner** | 3 | -0.023 | Less than chance | 0.12 | -0.25-0.20 | *ns* |
| | **Intermediate** | 5 | 0.16 | Slight | 0.063 | 0.035-0.28 | *<0.01* |
| | **Advanced** | 7 | 0.31 | Fair | 0.044 | 0.22-0.40 | *<0.01* |
| **YEARS EXPERIENCE** | **< 5** | 5 | 0.17 | Slight | 0.63 | -1.1-1.41 | *<0.01* |
| | **≥ 5** | 10 | 0.26 | Fair | 0.030 | 0.20-0.32 | *<0.01* |
| **ADVANCED AND ≥5 YRS EXPERIENCE** | | 3 | 0.22 | Fair | 0.12 | -0.0041-0.45 | *<0.05* |

*Table 8.5: Inter-observer reliability for questions 1, 2 and 3 for six observers before (assessment 1) and after (assessment 2) training*

| | Q1: Is the DDSS present? | | | | Q2: What does the structure represent? | | | | Q3: What follow-up would you recommend? | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | κ | Level of Agreement | 95%CI | p-value | κ | Level of Agreement | 95%CI | p-value | κ | Level of Agreement | 95%CI | p-value |
| **1** | 0.23 | Fair | 0.13-0.33 | <0.01 | 0.37 | Fair | 0.27-0.47 | <0.01 | 0.12 | Slight | 0.023-0.23 | <0.01 |
| **2** | 0.70 | Substantial | 0.60-0.80 | <0.01 | 0.53 | Moderate | 0.43-0.63 | <0.01 | 0.63 | Substantial | 0.53-0.73 | <0.01 |

*Table 8.6: Intra-observer reliability for questions 1, 2 and 3 for six observers*

| Question | κ statistic (95% CI) | | | | | |
|---|---|---|---|---|---|---|
| | Observer 1 | Observer 2 | Observer 3 | Observer 4 | Observer 5 | Observer 6 |
| **1. Is the DDSS present?** | 0.92 (0.53-1.3) | 0.84 (0.45-1.2) | 0.75 (0.36-1.1) | 0.73 (0.34-1.1) | 0.84 (0.45-1.2) | 0.92 (0.53-1.3) |
| **2. What does the structure represent?** | 0.82 (0.42-1.2) | 0.86 (0.47-1.3) | 0.83 (0.43-1.2) | 0.76 (0.37-1.1) | 0.90 (0.50-1.3) | 0.92 (0.53-1.3) |
| **3. What follow-up would you recommend?** | 0.84 (0.45-1.2) | 0.76 (0.37-1.2) | 0.68 (0.28-1.1) | 0.65 (0.26-1.0) | 0.76 (0.37-1.2) | 0.84 (0.45-1.2) |

# 9. Predicting Outcome in Pregnancies of Uncertain Viability Using Novel Serum Biomarkers

*NB; An abridged version of this chapter has been accepted for publication in the Journal of Obstetrics and Gynaecology*

## 9.1 Introduction

As discussed in chapter three, in cases of known intrauterine pregnancy, viability will be uncertain in approximately 10% of women attending an Early Pregnancy Assessment Unit [138]. The diagnosis of a pregnancy of uncertain viability is made following ultrasonographic visualization of either a small gestation sac (mean diameter ≤25mm) with no obvious yolk sac or fetal pole, or a small fetal pole (crown rump length ≤7mm) with no obvious fetal heart activity. These findings may represent a normal early pregnancy of between four and six weeks' gestation or they may be failed or failing pregnancies with arrested development which are destined to miscarry.

The diagnosis of a pregnancy of uncertain viability generates significant anxiety for women, as discussed in chapter four, and a considerable workload for Early Pregnancy Assessment Units as they all need to be followed up until a definitive diagnosis of either a viable or non-viable intrauterine pregnancy can be made, which takes at least seven days to achieve [138].

Unfortunately, as mentioned in chapter three, the prediction of outcome in pregnancies of uncertain viability is challenging [296, 297]. Many studies have attempted to establish methods of differentiating viable from non-viable pregnancies in cases of uncertainty, some using hCG levels [34, 38, 106, 150, 298], others progesterone concentrations [39-42], and others still various different ultrasonographic features [23, 25, 299-304]. No studies however, not even those that incorporate multiple demographic, clinical, ultrasonographic and serological parameters [305, 306] have been shown to predict pregnancy viability with absolute certainty. It is necessary for a test to have a sensitivity of

100% to avoid the unintentional termination of a viable pregnancy resulting from a false negative diagnosis.

This predicament has led to the search for alternative prognostic markers and substances involved in either vascular development or immune responses have been prime candidates of late. Recent studies have shown that, in women with certain diagnoses (i.e. viable, non-viable or ectopic pregnancies), serum angiopoietin-1 (Ang-1), angiopoietin-2 (Ang-2) and soluble fms-like tyrosine kinase-1 (Flt-1) concentrations are significantly higher [307-310] and levels of interleukin-15 (IL-15) and TNF-related apoptosis inducing ligand (TRAIL) significantly lower [311, 312] in women with viable intrauterine pregnancies compared to those with non-viable intrauterine or ectopic pregnancies.

## 9.2 Aims

The aim of our study was to determine whether serum concentrations of Ang-1, Ang-2, Flt-1, IL-15 or TRAIL can be used to predict pregnancy viability in cases of uncertainty.

## 9.3 Hypotheses

- Women with a pregnancy of uncertain viability and a high serum concentration of Ang-1, Ang-2 and/or Flt-1 are more likely to subsequently be diagnosed with a viable intrauterine pregnancy than women with a low serum concentration of Ang-1, Ang-2 and/or Flt-1.
- Women with a pregnancy of uncertain viability and a low serum concentration of IL-15 and/or TRAIL are more likely to subsequently be diagnosed with a viable intrauterine pregnancy than women with a high serum concentration of IL-15 and/or TRAIL.

## 9.4 Methods

### Ethical Approval

Ethical approval was obtained from Nottingham 1 Research Ethics Committee (13-EM-0081) (Appendix 1) and informed written consent was obtained from all participants (Appendix 8).

## Study Design

Participants were prospectively recruited from the Early Pregnancy Assessment Unit at the Queen's Medical Centre in Nottingham between 17th June 2014 and the 1st September 2015.

Women given a diagnosis of a pregnancy of uncertain viability following a transvaginal ultrasound were eligible to take part in the study. A pregnancy of uncertain viability was defined as the presence of an intrauterine gestation sac of less than 25mm mean diameter with no obvious yolk sac or fetal pole or an intrauterine gestation sac containing a fetal pole of less than 7mm with no obvious fetal heart pulsations [313, 314]. Women were excluded from the study if they were haemodynamically unstable, had attended the unit previously in the same pregnancy and had already had an ultrasound scan, if it was subsequently proven to be a multiple pregnancy or if the final diagnosis was not known.

Following the diagnosis of a pregnancy of uncertain viability, women were managed according to departmental protocols that involved a repeat ultrasound seven to ten days after the initial ultrasound. After follow-up, the final diagnosis of either a viable or non-viable intrauterine pregnancy was recorded. A viable intrauterine pregnancy was defined by subsequent ultrasonographic identification of an intrauterine gestation sac with a fetal pole of any length with demonstrable fetal heart pulsations. A non-viable intrauterine pregnancy was defined as either an empty intrauterine gestation sac with mean sac diameter greater than 25mm or an intrauterine gestation sac containing a fetal pole with crown rump length greater than 7mm with no demonstrable fetal heart pulsations or, in the absence of a viable embryo, no significant increase in the growth of the gestation sac or length of the fetal pole on two ultrasound scans performed more than 7 days apart. If the gestation sac and/or fetal pole had increased in size since the initial ultrasound, but obvious fetal heart pulsations were still not visible, a further ultrasound was scheduled ten to fourteen days later. Histological confirmation of all products of conception was also sought when available.

Demographic details, current and past obstetric history and ultrasonographic findings were recorded at the time of entry into the study (Appendix 9) and 10mls of blood in a serum separating tube was taken from the antecubital

fossa by a trained practitioner using aseptic technique. Blood was processed by centrifuge (4000rpm for 20 minutes) within two hours of collection and the supernatants were stored at -80◦C until analyzed. Serum samples were assayed for Ang-1, Ang-2, TRAIL, Flt-1, IL-15 and hCG by trained technicians utilizing the methods described below. The final diagnosis, determined by clinical follow-up as described above, was not known at the time of interpreting the biomarker assays, nor were the assay results known by those interpreting the final diagnosis.

As this was a phase I exploratory prognostic factor study, where possible, we have followed the recommendations for the REporting of tumour MARKer prognostic studies (REMARK) [315]. These guidelines, similar to the successful CONSORT initiative for randomized controlled trials [316] and the STARD statement for diagnostic studies [276], aim to encourage transparent and complete reporting so that the relevant information will be available to others to help them judge the usefulness of the data and understand the context in which the conclusions apply by ensuring relevant information about the study design, preplanned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods is provided. Although initially developed for tumour marker prognostic studies, the guidelines have been labeled as applying to clinical prognostic studies in general (Appendix 10).

## Candidate Biomarkers

### Angiopoietin-1 and 2

The angiopoietin family of angiogenic proteins are involved in vasculature development and angiogenesis [317]. They are also important in the process of placental maturation and growth from early pregnancy onwards. A major developmental step during placentation is the remodeling of maternal vasculature to gain a uteroplacental circulation and development of a competent fetoplacental vasculature within the trophoblast. Angiopoietins have been shown to be critically involved in vasculature development and angiogenesis, especially in the female reproductive tract [317-319]. Angiopoietin 1 (Ang-1) and 2 (Ang-2) are specific ligands of the Tie2 receptor and both bind to it with similar affinity to promote vessel assembly and maturation by mediating survival signals for endothelial cells and regulating the recruitment of mural cells. Ang-1 acts in a paracrine agonistic manner

inducing Tie2 phosphorylation and subsequent vessel stabilization. In contrast, Ang-2 is produced by endothelial cells and acts as an autocrine antagonist of Ang-1-mediated Tie2 activation. Ang-2 thereby primes the vascular endothelium to exogenous cytokines and induces vascular destabilization at higher concentrations [320].

Ang-1 and Ang-2 are both produced by the developing syncytiotrophoblast and cytotrophoblast and Ang-2 is also produced by placental macrophages and endothelial cells. Ang-1 and Ang-2 help in the formation of new blood vessels from pre-existing ones during vascular development of the placenta. As syncytiotrophoblasts are in direct contact with maternal blood in the intervillous space (Figures 1.1b-e), Ang-1 and Ang-2 concentrations in the feto-placental unit may be reflected in maternal serum. Decidual tissues of healthy pregnancies and miscarriages have shown an altered Ang-2/Ang-1 ratio based on different oxygen levels in both groups (since hypoxia upregulates Ang-2 transcription and destabilizes Ang-1 in healthy pregnancies). Any change in the hypoxic conditions of an early pregnancy therefore, such as that which occurs during pregnancy failure, could alter angiopoietins expression and ratio in the feto-placental unit, which may then be reflected in the maternal serum concentration.

In a recent study, maternal serum Ang-1 and Ang-2 concentrations were measured at between 6 and 8 weeks gestation in 33 women with a viable intrauterine pregnancy and 60 women with a failed pregnancy (either a missed miscarriage (n=30) or a ruptured ectopic (n=30)) [307]. Both Ang-1 and Ang-2 concentrations were lower in the ectopic pregnancy (median 689 and 302pg/ml respectively) and missed miscarriage (median 810 and 402pg/ml respectively) groups compared to the viable intrauterine pregnancy group (median 963 and 1477pg/ml). All the differences were found to be statistically significant. In addition, Ang-1 was also able to differentiate between an ectopic pregnancy and a missed miscarriage (p=0.011).

## Serum TNF-Related Apoptosis-Inducing Ligand

Successful implantation, trophoblast invasion and placental insufficiency are key processes in early pregnancy and placental insufficiency increases the risk of miscarriage and other pregnancy complications. Due to the difficulties in

obtaining first trimester placental tissue material, only a limited number of investigations on targeted gene expression in placentas from miscarriage cases have been conducted.

Increased expression of TNF-related apoptosis-inducing ligand (TRAIL) has been detected in placentas of women suffering from recurrent miscarriage [311]. TRAIL is the protein product of this gene found in the serum. It has been indicated as a potential predictive biomarker in maternal serum for early pregnancy complications.

Maternal TRAIL concentrations were found to be significantly elevated both at the recurrent miscarriage event (33.6 ± 4.3pg/ml) and in pregnancies which were uncomplicated at the time of sampling but later resulted in an unexpected first trimester miscarriage (28.5 ± 4.4pg/ml) compared to normal first trimester pregnancies (16.1 ± 1.6pg/ml). The differences were found to be statistically significant (p=0.00027 and 0.039 respectively) [311].

## Interleukin-15

During implantation and early pregnancy, the immunological processes that take place within the uterus are modulated by pro- and anti-inflammatory cytokines and their altered expression in the maternal serum may play a role in early pregnancy failure [2]. Successful pregnancy is considered a T helper (Th) 1 – Th2 cooperation phenomenon, with a predominantly Th2-type lymphocyte response and specific cytokine production [321]. Th2 responses favour a cytokine environment that promotes the induction of autoantibodies and several studies have attempted to link pregnancy failures with the presence of specific cytokines and autoantibodies, for example, interleukin-15 (IL-15) [322].

In women with recurrent miscarriage, there is an up-regulation of IL-15 expression in trophoblasts [323], suggesting that it may be a marker of a failing pregnancy. IL-15 is a cytokine that is expressed by human placental tissue culture and its serum levels correlate with the duration of the pregnancy [321, 324].

A recent study measured IL-15 levels in 33 women with a viable intrauterine pregnancy and 60 women with a failed pregnancy (either a missed miscarriage (n=30) or a ruptured ectopic (n=30)) [312] and found that IL-15

concentrations were higher in the women with a failed pregnancy (median 19.86pg/ml) compared to women with a viable intrauterine pregnancy (median 15.06pg/ml) and this difference was found to be highly significant ($p<0.001$). This is mainly due to the fact that the concentration of IL-15 in women with an ectopic is significantly higher than the concentration of IL-15 in women with a viable intrauterine pregnancy (median 24.9pg/ml and 15.06pg/ml respectively; $p<0.001$). IL-15 is able to discriminate a viable intrauterine pregnancy from an ectopic pregnancy with high diagnostic accuracy (AUC 0.818), and, at the threshold of 16.1pg/ml, has a sensitivity, specificity and clinically important negative predictive value of 92%, 68% and 0.999 respectively. Furthermore, IL-15 was also able to distinguish an ectopic pregnancy from a missed miscarriage ($p=0.015$).

## Soluble FMS-like Tyrosine Kinase-1

Angiogenic factors are involved in the formation of new blood vessels required for placental development and function. They are therefore critical for fetal growth and development. Soluble FMS-like Tyrosine Kinase (Flt-1), the soluble form of vascular endothelial growth factor, is an anti-angiogenic protein that inhibits the formation of new blood vessels resulting in potential pregnancy complications. It is expressed at very high levels in the trophoblast [325] and its production is highly increased in hypoxic conditions [326].

In a systematic review looking at the Flt-1 concentrations in the first trimester and subsequent development of pregnancy complications [327], twelve relevant studies were identified. The review found no clear evidence of an association between Flt-1 levels in the first trimester and adverse pregnancy outcomes. However, as the findings were affected by methodological, biological and testing variations between studies, no meta-analysis was performed. Furthermore, the search strategy included the terms 'adverse pregnancy outcomes', 'pregnancy complications' 'still birth', 'preterm birth', 'prematurity', 'small for gestational age', 'SGA', 'IUGR' and 'pre-eclampsia'. They did not specifically include early pregnancy complications such as miscarriage and ectopic pregnancy.

Two recent studies [308, 309] have however looked at the relationship between serum Flt-1 concentrations and the development of early pregnancy complications. In the first of these [309], Flt-1 levels were measured in 21

women with threatened miscarriage who subsequently went on to have a live birth, 19 women with threatened miscarriage who did later miscarry, 32 asymptomatic pregnant women in the first trimester and 14 non-pregnant women. Flt-1 levels were significantly ($p<0.001$) higher in pregnant women compared to the non-pregnant women. Similarly Flt-1 levels were significantly higher in women with threatened miscarriage who subsequently had a live birth compared to women with threatened miscarriage who did in fact miscarry ($p\leq0.001$).

In the second study [308], Flt-1 concentrations were measured at between 6 and 8 weeks gestation in 50 women with a viable intrauterine pregnancy, 40 women with a missed miscarriage and 38 women with a ruptured ectopic. The concentration of Flt-1 was lower in women with ectopic pregnancies (178.16±76.03pg/ml) and missed miscarriages (399.42±337.54pg/ml) compared to women with viable intrauterine pregnancies (1390.32±655.37pg/ml). Similarly, Flt-1 concentrations were higher in viable intrauterine pregnancies (1390pg/ml) compared to all pregnancy failures (375.76pg/ml) and the difference was found to be statistically significant ($p<0.001$). Flt-1 also had the ability to discriminate an ectopic pregnancy from a missed miscarriage ($p=0.033$). Flt-1 was also able to accurately discriminate between a normal pregnancy and a missed miscarriage (AUC of 0.771) and a normal pregnancy and an ectopic (AUC 0.758). For the clinically important discrimination between a missed miscarriage and an ectopic pregnancy, Flt-1 had a sensitivity and specificity of 62.5% and 92.1% respectively at the threshold value of 228.08pg/ml. The optimal threshold value for differentiating between a normal intrauterine pregnancy and an abnormal pregnancy (missed miscarriage or ectopic pregnancy) were 741.5pg/ml for Flt-1. At these thresholds, the sensitivity, specificity, positive and negative predictive values were 88%, 96.2%, 0.815 and 0.977 respectively.

The results from these studies suggest that Ang-1, Ang-2, TRAIL, IL-15 and Flt-1 may be potential prognostic factors in women with pregnancies of uncertain viability (Table 9.1).

## Enzyme-Linked Immunosorbent Assays

Enzyme-linked immunosorbent assay (ELISA) is a biochemical technique used mainly in immunology to detect the presence of an antibody or antigen in a

sample. It is a popular format of wet-lab type analytic biochemistry assay that involves detection of an analyte in a liquid sample by a method that continues to use liquid reagents during the analysis that stays liquid and remains inside a reaction chamber or well needed to keep the reactants contained. ELISAs have been used as diagnostic tools in medicine and plant pathology as well as a quality-control check in various industries, including the food industry.

***Table 9.1: Summary of relative Ang-1 [307], Ang-2 [307], TRAIL [311], IL-15 [312] and Flt-1 [308, 309] concentrations in women with viable and non-viable intrauterine pregnancies***

| Biomarker | Viable Pregnancy | Failing Pregnancy | Ability to discriminate between non-viable and ectopic |
|---|---|---|---|
| Ang-1 | High | Low | Yes |
| Ang-2 | High | Low | No |
| TRAIL | Low | High | Not assessed |
| IL-15 | Low | High | Yes |
| Flt-1 [308, 309] | High | Low | Yes |

Performing an ELISA involves at least one antibody with specificity for a particular antigen. The sample, with an unknown concentration of antigen, is immobilized on a solid support, usually a polystyrene microtiter plate, either non-specifically, via adsorption to the surface, or specifically, via capture by another antibody specific to the same antigen. After the antigen is immobilized, the detection antibody is added, forming a complex with the antigen. The detection antibody can be covalently linked to an enzyme, or can itself be detected by secondary antibody that is linked to an enzyme through bioconjugation. Between each step, the plate is typically incubated and then washed with a mild detergent solution to remove any proteins or antibodies that are non-specifically bound. After the final wash step, the plate is developed by adding an enzymatic substrate to produce a visible signal, which is proportional to the quantity of antigen in the sample. The qualitative reading is usually based on the intensity of transmitted light detected by spectrophotometry, which involves quantitation of transmission of some specific wavelength of light through the liquid.

There are various different types of ELISAs including indirect, direct, sandwich and competitive. Sandwich ELISAs are used in our study. These are a less common variant of ELISA but are highly efficient in sample antigen detection.

The sandwich ELISA quantifies antigens between two layers of antibody i.e. capture and detection antibody. The antigen to be measured must therefore contain at least two antigenic epitopes capable of binding to antibody since at least two antibodies act in the sandwich. Either monoclonal or polyclonal antibodies can be used as the capture and detection antibodies in Sandwich ELISA systems. Monoclonal antibodies recognize a single epitope that allows fine detection and quantification of small differences in antigen. A polyclonal is often used as the capture antibody to pull down as much of the antigen as possible. The advantage of Sandwich ELISA is that the sample does not have to be purified before analysis, and the assay can be very sensitive.

The steps in a sandwich ELISA are as follows: (1) prepare a surface to which a known quantity of capture antibody is bound (Figure 9.1(1)); (2) block any non-specific binding sites on the surface; (3) apply the antigen-containing sample to the plate (Figure 9.1(2)); (4) wash the plate, so that unbound antigen is removed; (5) add a specific antibody which binds to the antigen (Figure 9.1(3)); (6) apply enzyme-linked secondary antibodies as detection antibodies that also bind specifically to the antibody's Fc region (Figure 9.1(4)); (7) wash the plate, so that the unbound antibody-enzyme conjugates are removed; (8) apply a chemical that is converted by the enzyme into a color or fluorescent or electrochemical signal (Figure 9.1(5)); and (9) measure the absorbency or fluorescence or electrochemical signal of the plate wells to determine the presence and quantity of antigen.

***Figure 9.1: Schematic procedure of a sandwich ELISA***

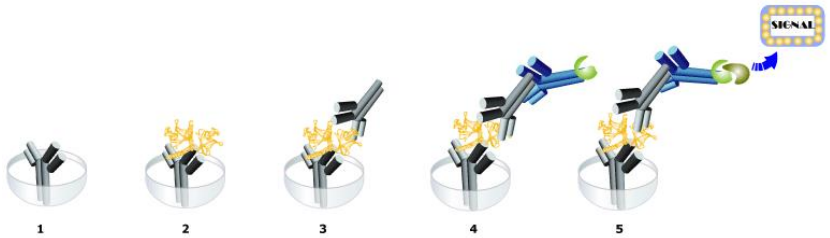



*Image taken from: https://www.genwaybio.com/services/sandwich-elisa*

Figure 9.1 includes the use of a secondary antibody conjugated to an enzyme, though, in the technical sense, this is not necessary if the primary antibody is conjugated to an enzyme. However, use of a secondary-antibody conjugate avoids the expensive process of creating enzyme-linked antibodies

for every antigen one might want to detect. By using an enzyme-linked antibody that binds the Fc region of other antibodies, this same enzyme-linked antibody can be used in a variety of situations. Without the first layer of capture antibody, any proteins in the sample (including serum proteins) may competitively adsorb to the plate surface, lowering the quantity of antigen immobilized. Use of the purified specific antibody to attach the antigen to the plastic eliminates a need to purify the antigen from complicated mixtures before the measurement, simplifying the assay, and increasing the specificity and the sensitivity of the assay.

## Biomarker Assay Technique

Serum concentrations of hCG were determined in the hospital biochemistry laboratory by appropriately trained technicians using a chemiluminometric two-site sandwich immunoassay (ADIVA Centaur® XP System).

Quantitative enzyme-linked immunosorbent assays were performed to measure the serum levels of Ang-1, Ang-2, IL-15, TRAIL and Flt-1 in accordance with the manufacturer's instructions (Quantikine, R&D Systems, Minneapolis, MN). Briefly, monoclonal antibodies specific for Ang-1, Ang-2, TRAIL, Flt-1 or IL-15 were pre-coated onto separate 96-well polystyrene microplates. Eight standards were prepared for each biomarker from a pre-made, readily available stock solution as instructed by the manufacturer (Table 9.2). For Ang-1, Ang-2, IL-15 and TRAIL, 50µl of the individual standards, samples and controls were pipetted into the wells containing 100µl of assay diluent specific for the biomarker. Serum samples were diluted 50-fold for Ang-1 and 5-fold for Ang-2. For Flt-1, 100µl of the individual standards, samples and controls was used. For Ang-1, Ang-2, TRAIL and Flt-1, the plates were then incubated for two hours at room temperature on a horizontal orbital microplate shaker set at 500±50rpm. For IL-15, the plates were incubated for three hours at room temperature. After four washes to remove any unbound substances, 200µl of an enzyme-linked monoclonal antibody specific for the corresponding biomarker was added to each well. For Ang-1, Ang-2, TRAIL and Flt-1, the plates were again incubated for two hours at room temperature on a horizontal orbital microplate shaker set at 500±50 rpm. For IL-15, the plates were incubated for one hour at room temperature. Following four washes to remove any unbound antibody-enzyme reagent, 200µl of substrate solution (hydrogen peroxide/tetramethyl benzidine) was added to the wells, which were then

incubated for 30 minutes at room temperature in the dark. The colour development reaction was stopped using 50µl of 2 N sulphuric acid and the absorbance values (optical densities) were determined in a microplate reader (Biorad Benchmark) at 450nm with the correction wavelength set at 540nm. The concentrations of the serum levels of each specific biomarker were calculated with the data analysis software (Microplate Manager; Bio-Rad). All tests were done in duplicate and the average of the duplicate readings was used for the analysis.

***Table 9.2: Concentration of standards used for the different ELISAs***

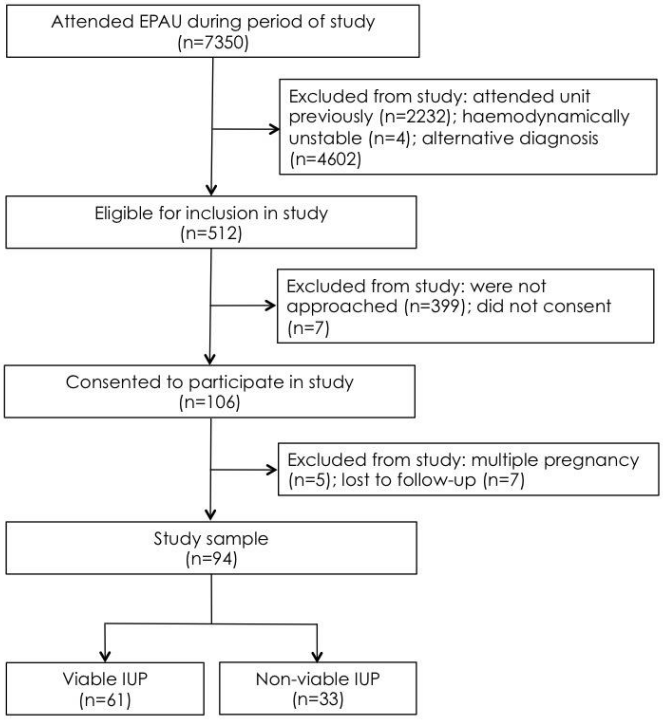
| Standard | Concentration (pg/ml) | | | | |
|---|---|---|---|---|---|
| | **Ang-1** | **Ang-2** | **Flt-1** | **TRAIL** | **IL-15** |
| **1** | 4000 | 3000 | 2000 | 1000 | 250 |
| **2** | 2000 | 1500 | 1000 | 500 | 125 |
| **3** | 1000 | 750 | 500 | 250 | 62.5 |
| **4** | 500 | 375 | 250 | 125 | 31.3 |
| **5** | 250 | 187.5 | 125 | 62.5 | 15.6 |
| **6** | 125 | 93.7 | 62.5 | 31.3 | 7.8 |
| **7** | 62.5 | 46.9 | 31.3 | 15.6 | 3.9 |
| **8** | 0 | 0 | 0 | 0 | 0 |

## Statistical Analysis

Statistical analysis was performed using SPSS 21 (Statistical Package for Social Sciences; IBM, Chicago, IL, USA). The continuous exposure variables (Ang-1, Ang-2, Flt-1, TRAIL and IL-15 concentrations) were first categorized and the nature of the association between each and the binary outcome measure (viable or non-viable intrauterine pregnancy) explored by computing odds ratios and risk ratios (with corresponding 95% confidence intervals) for each exposure category relative to the lowest category. Chi-squared test and chi-squared test for trend were undertaken to determine the statistical significance of the associations. Multivariate logistic regression was then performed to explore whether hCG acted as a confounder in each of the associations, and adjusted risk ratios presented if controlling for hCG materially altered the magnitude of the risk ratios (difference of greater than ±10%).

# 9.5 Results

***Figure 9.2: Flow of participants through the study***




Between 17th June 2014 and the 1st September 2015, 7350 women were assessed in the Early Pregnancy Assessment Unit at the Queen's Medical Centre in Nottingham. Of these, 512 (6.9%) were given the diagnosis of a pregnancy of uncertain viability following a transvaginal ultrasound. 106 (21%) women with a pregnancy of uncertain viability agreed to participate in the study. Five women were excluded as they were subsequently proven to have a multiple pregnancy and a further seven women were excluded because the final diagnosis was not known. Of the remaining 94 women, 61 (65%) had a viable intrauterine pregnancy and the remaining 33 (35%) had a non-viable intrauterine pregnancy confirmed on subsequent ultrasound (Figure 9.2).

The baseline characteristics of the 94 women that formed our study sample are given in Table 9.3. Most participants were non-smoking, nulliparous women, in their early thirties with a slightly raised body mass index, presenting at between 6 and 7 weeks gestation with abdominal pain and vaginal bleeding. An empty sac measuring approximately 7mm was the most common ultrasound finding.

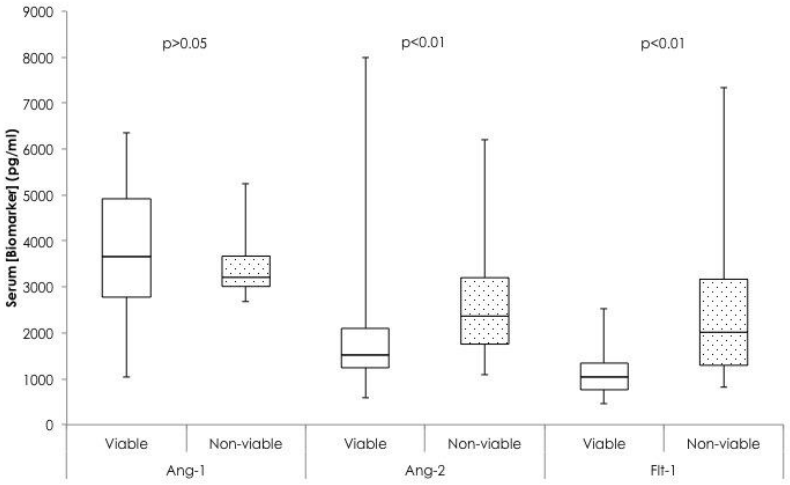
***Table 9.3: Baseline characteristics of study participants***

| | n (%) | | |
|---|---|---|---|
| | **Overall (n=94)** | **Viable (n=61)** | **Non-viable (n=33)** |
| **Age (mean±SD)** | 32±11 | 33±12 | 32±6.8 |
| **Gravidity (mean±SD)** | 2.5±1.6 | 2.4±1.7 | 2.9±1.5 |
| **Parity (mean±SD)** | 0.8±1.0 | 0.6±1.0 | 1.0±1.0 |
| **Uncertain LMP** | 9 (10) | 5 (8.2) | 4 (12) |
| **EGA (mean±SD)** | 45±13 | 38±8.5 | 57±12 |
| **Presenting complaint** | | | |
| Pain | 21 (22) | 13 (21.3) | 8 (24) |
| Bleeding | 26 (28) | 19 (31.2) | 7 (21) |
| Pain and bleeding | 47 (50) | 29 (47.5) | 18 (55) |
| **Previous miscarriage** | 28 (30) | 15 (24.6) | 13 (39) |
| **Previous ectopic** | 7 (7.4) | 7 (11.5) | 0 (0.0) |
| **Body mass index (mean±SD)** | 26±5.6 | 25±4.9 | 27.4±7.1 |
| **Smoking status** | | | |
| Never | 64 (68) | 45 (74) | 19 (58) |
| Ex | 17 (18) | 8 (13) | 9 (27) |
| Current | 13 (14) | 8 (13) | 5 (15) |
| **US findings** | | | |
| MSD (mean±SD) | 7.3±5.0 | 5.6±3.2 | 11±6.2 |
| Empty GS | 64 (68) | 44 (72) | 20 (61) |
| GS with YS only | 16 (17) | 12 (20) | 4 (12) |
| GS with YS and FP | 14 (15) | 5 (8.2) | 9 (27) |
| CRL (mean±SD) | 3.3±1.2 | 2.9±1.3 | 3.6±1.1 |

Eighty-one of the 94 (86%) samples had TRAIL concentrations less than 15.6pg/ml i.e. the lowest standard (Table 9.2) and therefore could not be accurately interpreted. There were too few samples remaining to perform any meaningful statistical analyses. Similarly, all 94 samples had IL-15

concentrations less than 3.9pg/ml and therefore could not be accurately interpreted.

***Figure 9.3 Serum biomarker concentrations according to subsequent pregnancy viability (median/IQR)***




[Ang-1] and [Flt-1] concentrations have been scaled up/down to enable easier graphical representation. For actual levels, multiply by 10 for [Ang-1] and divide by 10 for [Flt-1]

***Table 9.4: Serum biomarker concentration according to subsequent pregnancy viability (median/IQR)***

| Biomarker | Serum concentration (pg/ml) | | p-value* |
|---|---|---|---|
| | Viable (n=61) | Non-viable (n=33) | |
| **Ang-1** | 36660 (27684-49070) | 32100 (30099-36702) | >0.05 |
| **Ang-2** | 1510 (1250-2093) | 2365 (1754-3200) | <0.01 |
| **Flt-1** | 103 (77-134) | 202 (129-317) | <0.01 |

**Mann-Witney U Test*

The median concentrations of Ang-1, Ang-2 and Flt-1 according to subsequent pregnancy viability are illustrated in Figure 9.3 and Table 9.4. The median concentrations of Ang-2 and Flt-1 were significantly lower in pregnancies of uncertain viability that were subsequently proven to be viable than those that were subsequently proven to be non-viable. There was no

significant difference in the median concentration of Ang-1 in pregnancies of uncertain viability that were subsequently proven to be viable compared to those that were subsequently proven to be non-viable.

There was a statistically significant (p=0.026), non-linear (p-value for trend=0.20) association between Ang-1 concentration and pregnancy viability such that women with Ang-1 concentrations between 30337pg/ml and 38391pg/ml were 30% less likely to have a viable intrauterine pregnancy than women with Ang-1 concentrations less than or equal to 30337pg/ml and women with Ang-1 concentrations greater than or equal to 38391pg/ml were 31% more likely to have a viable intrauterine pregnancy than women with Ang-1 concentrations less than or equal to 30337pg/ml (Table 9.5 and Figure 9.4).

***Table 9.5: Likelihood of subsequent pregnancy viability according to serum Ang-1, Ang-2 and Flt-1 concentrations***

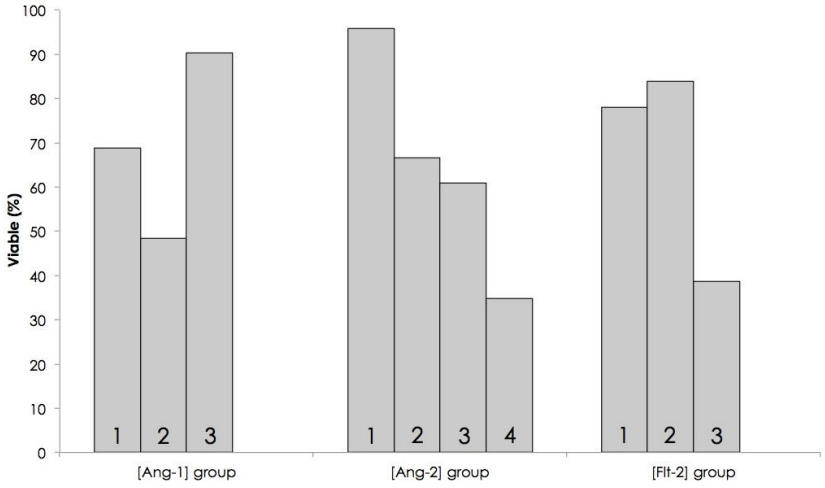
| Biomarker | Group | Cut-off (pg/ml) | n | Viable n (%) | RR (95% CI) |
|---|---|---|---|---|---|
| Ang-1 | 1 | ≤30337 | 32 | 22 (69) | 1.0 |
| | 2 | 30338-38391 | 31 | 15 (48) | 0.70 (0.46-1.1) |
| | 3 | ≥38391 | 31 | 28 (90) | 1.3 (1.0-1.7) |
| Ang-2 | 1 | ≤1382 | 24 | 23 (96) | 1.00 |
| | 2 | 1383-1772 | 24 | 16 (67) | 0.70 (0.52-0.93) |
| | 3 | 1773-2666 | 23 | 14 (61) | 0.64 (0.45-0.89) |
| | 4 | ≥2666 | 23 | 8 (35) | 0.36 (0.21-0.64) |
| Flt-1 | 1 | ≤87 | 32 | 25 (78) | 1.00 |
| | 2 | 88-142 | 31 | 26 (84) | 1.1 (0.84-1.4) |
| | 3 | ≥142 | 31 | 12 (39) | 0.50 (0.31-0.80) |

There was a statistically significant (p<0.01), linear (p-value for trend <0.01) association between Ang-2 concentration and pregnancy viability such that women with Ang-2 concentrations greater than or equal to 2666pg/ml were 64% less likely to have a viable intrauterine pregnancy than women with Ang-2 concentrations less than or equal to 1382pg/ml (Table 9.5 and Figure 9.4).

Similarly, there was a statistically significant (p<0.01), linear (p-value for trend=0.01) association between Flt-1 concentration and pregnancy viability such that women with Flt-1 concentrations greater than or equal to 142pg/ml were 50% less likely to have a viable intrauterine pregnancy than women with Flt-1 concentrations less than or equal to 87pg/ml (Table 9.5 and Figure 9.4).

Controlling for serum hCG concentration had no material effect on any of the associations.

***Figure 9.4: Percentage chance of subsequent pregnancy viability according to serum Ang-1, Ang-2 and Flt-1 group***




# 9.6 Discussion

The aim of our study was to determine whether serum concentrations of Ang-1, Ang-2, Flt-1, IL-15 or TRAIL could potentially be used to predict pregnancy viability in cases of uncertainty. Although previous studies have demonstrated statistically significant differences in the concentrations of these biomarkers between women diagnosed with viable and non-viable intrauterine pregnancies [307-312], this is the first study to investigate the prognostic potential of these serum biomarkers to predict pregnancy outcome in cases of uncertainty.

The findings of our prospective study incorporating 94 women presenting with abdominal pain and/or vaginal bleeding in early pregnancy and diagnosed, following transvaginal ultrasound, with pregnancies of uncertain viability suggest that Ang-2 and Flt-1, and to a lesser extent Ang-1, may be useful in the prediction of pregnancy viability in cases of uncertainty. Our results demonstrate that women with low serum concentrations of Ang-2 and Flt-1

are significantly more likely to have viable intrauterine pregnancies than women with high serum concentrations.

These findings are not altogether consistent with previous studies [307-310, 312] which reported significantly higher concentrations of Flt-1 and Ang-2 in viable pregnancies. This phenomenon is not uncommon in prognostic marker studies: often initially reported studies of a marker show great promise but subsequent studies on the same or related markers yield inconsistent conclusions or stand in direct contradiction to the promising results [315]. In these earlier studies, in women with viable intrauterine pregnancies at the time of sampling, serum Ang-1, Ang-2 and Flt-1 concentrations were significantly higher than women with non-viable intrauterine pregnancies [307-310]. This led us to hypothesize that women with pregnancies of uncertain viability and a high serum concentration of Ang-1, Ang-2 and/or Flt-1 would be more likely to be subsequently diagnosed with viable intrauterine pregnancies than women with pregnancies of uncertain viability and low serum concentrations of Ang-1, Ang-2 and/or Flt-1. In fact, our results have suggested that the opposite may be true.

The reasons for these discrepancies are not fully understood but are likely to be due to fundamental differences in study design as well as the complex processes involved in implantation and placentation in normal, and abnormal, early pregnancy development.

The aim of the studies by Daponte et al [307, 308], was to assess whether a single serum measurement of Ang-1, Ang-2 and Flt-1 at 6-8 weeks gestation, could contribute to the differential diagnosis between failed pregnancies, whether ectopic or non-viable intrauterine, and viable intrauterine pregnancies whilst the aim of our study was to determine whether serum concentrations of Ang-1, Ang-2 and Flt-1 could be used to predict pregnancy viability in cases of uncertainty. This subtle but important difference relies upon a different study design (case-control versus cohort) and statistical analysis, which may, at least partially, be responsible for the differences in the results observed.

The methodology described in the Daponte studies [307, 308] is not altogether clear. Although it states that Ang-1, Ang-2 and Flt-1 concentrations were measured in a group of controls with healthy intrauterine pregnancies of

between six and eight weeks of gestation, it does not state whether they underwent an ultrasound scan to confirm either the gestation or the viability of the pregnancy. In the study regarding the angiopoietins [307], the cases appear to consist of women who presented with abdominal pain and/or vaginal bleeding at between six and eight weeks gestation, in whom the precise diagnosis could not be made following the initial transvaginal ultrasound and hence were admitted to hospital for further investigation. Serum samples were apparently collected 'at the initial visit before treatment'. It is not clear from the methodology described whether the serum samples were taken at the time of uncertainty i.e. at the initial visit when they were admitted to hospital, or after the miscarriage or ectopic pregnancy had been diagnosed but before treatment was initiated. In the Flt-1 study [308], serum samples were collected from women presenting with mild abdominal pain or vaginal bleeding between six and eight weeks' gestation who were subsequently diagnosed with a missed miscarriage. How the gestation was calculated (i.e. from last menstrual period or by measurement of gestation sac diameter) or how the diagnosis of a missed miscarriage was made (i.e. clinically or ultrasonographically) is not fully described. These ambiguities may explain some of the differences in biomarker concentrations observed.

Our study sample consisted of women who had been given the ultrasonographic diagnosis of a pregnancy of uncertain viability following presentation with abdominal pain and/or vaginal bleeding in early pregnancy. Serum samples were taken immediately after the ultrasound, at the time of diagnostic uncertainty, and women were followed up until a definitive diagnosis of either a viable or non-viable intrauterine pregnancy was made. Although in our study we did not control for gestation in the sense that we recruited any woman with a pregnancy of uncertain viability, irrespective of the date of her last menstrual period, we did measure serum hCG concentrations in all women and control for hCG concentration in our statistical analysis. Whilst, as already mentioned in chapter one, hCG is not a precise indicator of gestation [27, 28], neither is dating a pregnancy by last menstrual period alone [142-144].

The situation is complicated by the fact that the physiological processes involved in implantation and placentation are complex. As described in chapter one, the process of placentation is initiated once the blastocyst makes contact with the epithelium of the uterus. An initial trophoblastic shell is

penetrated by columns of proliferating extra-villous cytotrophoblast that form the anchoring villi and provide specialized invasive cells that transform the decidual and proximal portions of the decidual spiral arteries [328]. During the initial phase of implantation and uterine wall invasion, the main role of extra-villous trophoblast is to form plugs that occlude capillaries in the endometrial gland stroma; this prevents maternal haemorrhage from disrupting the conceptus and maternal blood from entering the lacunar spaces of the trophoblastic shell. Embryogenesis thus takes place in a hypoxic environment for the first ten weeks of pregnancy because oxygen tension within the placenta is much lower than in the surrounding endometrial glands [329-332]. The plugging mechanism protects the growing embryo and the primitive placental villi against oxidative damage. After approximately ten weeks, the trophoblastic plugs are breached by maternal blood that then enters the inter-villous space. This sudden perfusion is thought to render the placenta less hypoxic than in earlier pregnancies. Utero-placental blood flow increases exponentially from less than 50 mL/min in the non-pregnant state to approximately 350 mL/min by full-term. The demands of this large rise in uteroplacental blood flow (up to 20% of the total maternal cardiac output), require large adaptations in maternal physiology [333]. Approximately two-thirds of miscarriages are attributed to defective placentation associated with an absence of physiological change in maternal spiral arteries and premature onset of maternal circulation through the placenta [334].

Concentrations of hCG, Ang-1, Ang-2 and Flt-1 are dependent on both pregnancy gestation and pregnancy viability. The concentration of hCG in the maternal blood increases with advancing gestation reaching a peak at between ten and twelve weeks of gestation. The concentration then declines to a stable plateau for the remainder of the pregnancy (Figure 1.4)[26]. In a non-viable pregnancy, the rate of hCG decrease is described by a quadratic profile, with a faster decline in hCG value with higher presentation levels [32]. Unfortunately such nomograms for serum concentrations of Ang-1, Ang-2 and Flt-1 do not exist to the same extent as they do for serum hCG concentrations. We do know however, that the local balance between Ang-1 and its antagonist Ang-2 determines whether blood vessels grow, are maintained or regress. We also know that the relative levels of Ang-1 and Ang-2 are regulated by local oxygen tension by different mechanisms and this may be important during normal human placentation [335]. Furthermore, high levels of Flt-1 occur in normal first trimester pregnancies as a consequence of

excessive placental production under hypoxic conditions [326] but following the onset of maternal circulation through the placenta, the placenta becomes less hypoxic and Flt-1 concentrations may alter to reflect this change. Hence it is difficult to ascertain whether certain concentrations of hCG, Ang-1, Ang-2 and Flt-1 are observed because the gestation is early but progressing or more advanced and failing.

The median concentrations of all biomarkers were substantially different between our study and the studies by Daponte et al (Table 9.8). In the Daponte studies [307, 308], although all women were recruited at between six and eight weeks gestation, the median concentration of hCG in those with a viable pregnancy was vastly different from those with a non-viable pregnancy (59668mIU/ml and 56130mIU/ml versus 3000mIU/ml and 4710mIU/ml). Whilst a serum hCG concentration of 59668mIU/ml correlates to a gestation of approximately seven weeks and one or two days, a serum hCG concentration of 3000mIU/ml is more suggestive of a pregnancy of less than five weeks' gestation. Granted the latter were failing pregnancies, which may account for the lower hCG concentrations observed, but a difference of approximately 15-fold is quite pronounced particularly since hCG concentrations remain elevated in the maternal serum for some time following a miscarriage, especially when the trophoblastic tissue is still in situ [32] as it was in these women. Furthermore, in our study, the median hCG concentrations of women with a pregnancy of uncertain viability that was subsequently proven to be viable was 2313.3iu/l (correlating to a gestational age of approximately four weeks and five days) and the median hCG concentrations of women with a pregnancy of uncertain viability that was subsequently proven to be non-viable was 6754.0iu/l (correlating to a gestational age of five weeks and three days). Using hCG as a crude marker of gestation, it therefore appears that we have sampled our viable intrauterine pregnancies at an earlier gestation and our non-viable intrauterine pregnancies at a later gestation than the Daponte studies. This may account for the differences in the absolute levels of Ang-1, Ang-2 and Flt-1 concentrations observed as well as the differences in the trends between the viable and non-viable pregnancies (Table 9.6).

As part of the analyses in the Daponte studies [307, 308], the median concentration of Ang-1, Ang-2 and Flt-1 in women with a viable intrauterine pregnancy was compared to the median concentration of Ang-1, Ang-2 and Flt-1 in women with a non-viable intrauterine pregnancy. Using the Kruskal-

Wallis test, concentrations of Ang-1, Ang-2 and Flt-1 were found to be significantly higher in women with viable intrauterine pregnancies compared to those with non-viable intrauterine pregnancies. In our study, a similar analysis demonstrated (1) no significant difference in the median concentrations of Ang-1 between women with pregnancies of uncertain viability that were subsequently proven to be viable compared to those with pregnancies of uncertain viability that were subsequently proven to be non-viable and (2) significantly lower concentrations of Ang-2 and Flt-1 in women with pregnancies of uncertain viability that were subsequently proven to be viable compared to those with pregnancies of uncertain viability that were subsequently proven to be non-viable (Figure 9.3 and Table 9.4). However, the median concentrations of Ang-1, Ang-2 and Flt-1 in our study were very different to those reported in the Daponte studies (Table 9.6) despite exactly the same assays being performed in both. These discrepancies may be related to differences in the timing of serum sampling, either in terms of gestational age or in relation to the miscarriage event.

## Strengths and Limitations

This is the first study of its kind to investigate the prognostic potential of Ang-1, Ang-2 and Flt-1 to predict pregnancy outcome in cases of uncertainty. We included 94 participants and whilst only 33 of these were subsequently found to have non-viable intrauterine pregnancies, this is reflective of the clinical course of pregnancies of uncertain viability [306]. Our study was prospective in nature, had clearly defined hypotheses based upon the best available evidence at conception of the study, had explicit inclusion and exclusion criteria and the selection of patients was not related to outcome. Furthermore, our statistical analysis controlled for hCG concentration as a marker of gestation. Whilst hCG is not a precise indicator of gestation [27, 28], neither is dating a pregnancy by last menstrual period alone [142-144].

The multi-variable character of prognostic research makes it notoriously difficult to estimate the required sample size [315]. One weakness of our study is that we were unable to do an *a priori* sample size calculation due to absence of the relevant data in the scientific literature. In the lieu of this, we undertook a *post-hoc* power calculation which revealed that on the basis of the means, the between-groups comparison effect sizes observed, and the size of each of our two groups, a statistical power of >0.90 was observed for

Ang-2 and Flt-1. Unfortunately a statistical power of only 0.26 was observed for Ang-1. This limited statistical power generated by our modest sample size only limits the statistical comparisons conducted for Ang-1 and fortunately our main conclusions relate to Ang-2 and Flt-1.

Another weakness of our study is that viability was defined as visible fetal heart pulsations on a subsequent ultrasound scan rather than viability at twelve weeks gestation. Approximately 25% of pregnancies of uncertain viability that are initially found to be viable subsequently fail [306]. However, it was felt that because the mechanisms surrounding miscarriage are multifold and complex, a marker of initial pregnancy viability would still be useful clinically.

Only symptomatic women with pregnancies of uncertain viability were included in our study. It is not clear whether our results could be extrapolated to all women in early pregnancy in an attempt reassure women that their pregnancy is progressing normally or identify failing pregnancies at a very early gestation.

As with all prognostic factor studies, discoveries made in small studies such as this are prone to overestimating or underestimating the actual association. Evidence from multiple studies, in particular large studies, is necessary to appreciate the discriminating ability of these emerging prognostic factors. Rapid clinical adoption in the absence of such evidence may lead to wasted resources. It is therefore imperative that we replicate and confirm our findings before implementing them in clinical practice.

Furthermore, if validated, it is necessary to determine appropriate threshold levels, above or below which we can confidently diagnose a viable or non-viable intrauterine pregnancy. The sensitivity needs to be 100% because the clinical consequences of a false negative diagnosis could be to induce, either medically or surgically, termination of a wanted, viable intrauterine pregnancy. It may be that on their own our biomarkers are not capable of generating this high degree of accuracy, but in combination with other demographic, clinical, serological and radiological parameters they may be used to generate a prognostic research model capable of predicting individual risk of a future outcome. This would, if proven to be cost effective, reduce the strain on limited resources and alleviate anxiety for women.

## 9.7 Conclusion

Serum concentrations of Ang-2 and Flt-1, and to a lesser extent Ang-1, may be able to predict outcome in women with pregnancies of uncertain viability. Although further work is required to confirm these findings before implementation into clinical practice, they appear to be promising prognostic factors.

***Table 9.6: Differences in serum concentrations of hCG, Ang-1, Ang-2 and Flt-1 between our study and the studies by Daponte et al. [307, 308]***

| | **hCG** | | | **Ang-1** | | **Ang-2** | | **Flt-1** | |
|---|---|---|---|---|---|---|---|---|---|
| | **Richardson** | **Daponte [307]** | **Daponte [308]** | **Richardson** | **Daponte [307]** | **Richardson** | **Daponte [307]** | **Richardson** | **Daponte [308]** |
| Viable | 2313 | 59688 | 56130 | 36660 | 964 | 1510 | 1477 | 103 | 1390 |
| Non-viable | 6754 | 3000 | 4710 | 32100 | 811 | 2365 | 402 | 202 | 399 |

# 10. Clinical Impact & Future Research Recommendations

## 10.1 Clinical Impact

The aim of this research was to try to minimise the number of women given uncertain diagnoses in early pregnancy, and if that was not possible, to at least minimise the duration of uncertainty for them. The premise for this work was that the current clinical management of women with diagnostic uncertainties in early pregnancy was often haphazard and protracted and utilized valuable, limited resources. Furthermore, in the time taken to make a definitive diagnosis, a stable woman with an unknown ectopic pregnancy or miscarriage might become unstable, and require immediate resuscitation, life-saving blood transfusion and/or emergency surgery.

In the first of our studies, described in chapter four, we describe another reason why it is important to minimise the number of women given uncertain diagnoses in early pregnancy and that is because of the considerable levels of anxiety that diagnoses such as pregnancies of uncertain viability and unknown location generate. We demonstrate that women who present to Early Pregnancy Assessment Units with abdominal pain and/or vaginal bleeding in early pregnancy and who are subsequently given an uncertain diagnosis have significantly higher levels of anxiety than their counterparts who are given certain diagnoses, even if those certain diagnoses are not associated with ongoing pregnancies. Although perhaps intuitive to some, this has never before been described in the scientific literature.

A number of studies have noted anxiety, distress [336], depression [337] and hostility [338] in response to non-definitive ultrasound findings. Some of these emotions have been reported to be stronger following an uncertain diagnosis than a negative, but less ambiguous one [338]. The levels of anxiety reported by the women in our study prior to the ultrasound scan suggest that they are aware that the experience of symptoms such as abdominal bleeding and/or vaginal bleeding in early pregnancy may be indicative of a problem. Whilst they may anticipate being given bad news, it is likely that they do not expect to hear uncertain news, or that there may be threats to their own health, or

that they require a plethora of further investigations, or that they may not get a definitive answer for several weeks. It is of no surprise therefore that uncertain diagnoses in early pregnancy generate considerable anxiety.

It is crucial that having demonstrated that uncertain diagnoses in early pregnancy generate considerable anxiety, we address this potential for anxiety in clinical practice. We can do this is several ways. Firstly, in an attempt to minimise the number of women given an uncertain diagnosis in early pregnancy, it is imperative that asymptomatic women are not offered an ultrasound scan when the likelihood of an uncertain diagnosis is high. Based on a study of 1442 women [172], the commonest transvaginal ultrasound finding prior to 35 days gestation was a pregnancy of unknown location and from 42 days gestation it was a viable intrauterine pregnancy although the chance of confirming viability increased rapidly per day of gestation until 49 days and thereafter plateaued. A miscarriage could not be diagnosed on initial transvaginal ultrasound prior to 35 days gestation. Between 35 and 41 days gestation the commonest transvaginal ultrasound finding was a pregnancy of uncertain viability. It is sensible therefore that in asymptomatic women, with no previous ectopic pregnancy, transvaginal ultrasound should be delayed until 49 days gestation which will decrease the number of inconclusive ultrasound scans performed without an associated increase in morbidity from missed ectopic pregnancy. Symptomatic women however should not have an ultrasound scan deferred, regardless of gestation, due to the risk of serious morbidity associated with a missed diagnosis of an ectopic pregnancy. Whilst performing an ultrasound scan too early in pregnancy may lead to an uncertain diagnosis, deferring an ultrasound scan based on arbitrary limits for gestation may be associated with physical and psychological morbidity or mortality due to a delay in the diagnosis of a miscarriage [170] or ectopic pregnancy [171] and this is unacceptable.

Secondly, others have previously suggested that patient education about potential findings on routine ultrasound examinations should be made explicit to patients before they undergo the procedure [336, 339, 340]. This approach would also be useful for non-routine ultrasound examinations in early pregnancy such as those which occur when women present with symptoms of abdominal pain and/or vaginal bleeding. If women are given information regarding all the possible diagnoses, including the uncertain ones, prior to

having the ultrasound scan, then the initial shock of an uncertain diagnosis may be slightly abated. They are less likely to feel that they are being 'fobbed off' if they have been previously informed that an uncertain diagnosis in early pregnancy is not only a recognized phenomenon but also a fairly common one at that. This information need not be extensive, it merely serves to set the scene and focus the mind prior to the scan. It could be as simple as an information leaflet given to all patients when they report to the front desk for their scheduled appointment.

Finally, once an uncertain diagnosis has been given, women need to be supported. Those with non-viable intrauterine pregnancies, and to a lesser extent those with ectopic pregnancies, have access to different support groups and forums, either locally, nationally or via the internet. Women with uncertain diagnoses have no such psychological support and this must be addressed if we are to improve the holistic nature of care provided to women with complications of early pregnancy. Initially this should be face-to-face, but given that it is likely that women will be acutely distressed immediately following the ultrasound scan, this should be accompanied by detailed written information. Most women will seek additional information beyond that which is given by the provider. The value of these outside sources is generally to reinforce information that is provided at the time of the diagnosis, at a time when the woman is more receptive, in an environment that is less emotionally charged. In today's world the use of the internet as a primary tool for seeking this information is a reality, hence providers should consider developing a brief list of sites that contain accurate information and distributing it at the time of diagnosis e.g. www.miscarriageassociation.org.uk, www.ectopic.org.uk and www.nhs.uk. Women with a pregnancy of unknown location will likely be returning to the Early Pregnancy Assessment Unit 48 hours later for a repeat hCG blood test and it is important that time is taken during this visit to see how they are, both physically and psychologically. Women with a pregnancy of uncertain viability will not be returning to the Unit for at least seven days and it is important that they are not left to fester at home. The results from our study show that their anxiety levels continue to increase 48-72 hours after the ultrasound scan, so this would be the opportune time for a courtesy telephone call to see how they are. Women need to be reassured that they are not on their own in this limbo period.

Having demonstrated that diagnostic uncertainties in early pregnancy generate considerable anxiety for women, it is now imperative that further research focuses on reducing the number of women to whom uncertain diagnoses are given and this is what the studies described in chapters five to eight have tried to accomplish. Equally, if an uncertain diagnosis is absolutely unavoidable, future research efforts should try to at least minimise the duration of uncertainty, so that women have to wait the shortest amount of time possible for an accurate and definitive diagnosis. This was the aim of the study reported in chapter nine.

In an attempt to minimise the number of women given the diagnosis of a pregnancy of unknown location, we chose to focus on the 'pseudosac'. One of the biggest challenges in early pregnancy ultrasonography is trying to differentiate between an empty early gestation sac and a pseudosac. Pseudosacs occur in up to 15% of ectopic pregnancies and although rare, if present they are highly suggestive of an ectopic pregnancy, hence it is an important distinction to make, ideally as soon as possible. Although experts may claim that it is an easy distinction to make, in clinical practice, many of those undertaking the early pregnancy ultrasound scans do not claim to be experts. Traditional teaching has always been that one should wait until a yolk sac or fetal pole has been visualized before confirming that an intrauterine fluid collection is indeed an intrauterine gestation sac. Although safe, the problem with this approach is that using modern transvaginal ultrasound, an intrauterine fluid collection may be visible at 28 days gestation whilst a yolk sac may not appear for another seven days at least and if an ultrasound is undertaken during this time uncertainty will ensue.

We undertook a systematic review of the literature in an attempt to see if there were any ultrasonographic signs that could be used to reliably differentiate a true gestation sac from a pseudosac prior to visualization of a yolk sac. This systematic review, described in chapters five and six, identified three potential signs: the double decidual sac sign; the intradecidual sign; and the chorionic rim sign. Meta-analysis revealed that the double decidual sac sign was the most promising candidate with a sensitivity of approximately 82% and a specificity of 97%. Unfortunately however only six small studies were included in the meta-analysis, the majority of which were considered to be of mediocre quality at best, mainly because they were performed 20 to 30 years ago using

ultrasound technology vastly inferior to what we have available today resulting in high applicability concerns with regards to the index test.

Our systematic review and meta-analysis also demonstrated that a pseudosac has a sensitivity of 5.5% and a specificity of 94% for predicting an ectopic pregnancy in the absence of an obvious extra-uterine embryo. So, whilst a true pseudosac, highly suggestive of an ectopic pregnancy is a relatively rare finding in early pregnancy and usually absent in women with an ectopic pregnancy, an empty gestational sac, indicative of an early or failing intrauterine pregnancy, is much more common.

This prompted us to undertake our own study, reported in chapter seven, to determine the diagnostic accuracy of the double decidual sac sign to predict an intrauterine pregnancy prior to ultrasonographic visualization of embryonic contents using modern, high-resolution transvaginal ultrasound. This study found that the overall diagnostic accuracy of the double decidual sac sign to predict an intrauterine pregnancy prior to visualization of embryonic contents was 94% suggesting that the results from the meta-analysis in chapter five were underestimated.

One of the main limitations of this study is that, for a variety of reasons, all of the ultrasound scans were performed and interpreted by a single investigator. However, if we want to start using the double decidual sac sign in clinical practice, to inform management decisions, it is important that we not only demonstrate that it is accurate, but also reliable. In our penultimate study, described in chapter eight, we assessed the inter- and intra-observer reliability associated with ultrasonographic visualization of the double decidual sac sign. Although the inter-observer reliability was only fair, it did improve significantly after a period of training and as such, may still be rendered clinically useful after incorporation into ultrasound training and widespread dissemination into the relevant communities. This is in progress following presentation of our findings at national and international conferences and publication in relevant peer-reviewed journals.

Our study looking at the inter- and intra-observer reliability associated with ultrasonographic visualization of the double decidual sac sign also highlights that the management of women with ultrasonographic evidence of an 'empty sac' is inconsistent. Irrespective of whether or not the double decidual

sac sign was thought to be present or not, when the intrauterine fluid collection was considered to be a gestation sac, the vast majority of observers advised a rescan in at least seven days, in line with UK guidelines for pregnancies of uncertain viability [341]. However, observers recommended serial hCG blood tests in 13% of cases. Since serial hCGs cannot reliably discriminate between viable and non-viable pregnancies [34, 38, 106, 150, 298], the main indication for them is to help rule in, or indeed out, an ectopic pregnancy in which case these 'empty sacs' are being managed according to pregnancy of unknown location guidelines [341] despite the fact that the ultrasonographic finding of an 'empty sac' does not fulfill the criteria for a pregnancy of unknown location according to the current definition [108]. Similarly, when the intrauterine fluid collection was considered to be a pseudosac, the vast majority of observers advised serial serum hCG measurements be taken, consistent with pregnancy of unknown location guidelines. However, observers recommended a rescan in 4% of cases.

Therefore the management of an 'empty sac' appears to fall into a grey area – are they pregnancies of uncertain viability even though they lack a yolk sac or fetal pole and as such are not, by traditional teaching, definitely intrauterine pregnancies – or are they pregnancies of unknown location, despite the fact that there is some evidence of a potential, albeit not definite, intrauterine pregnancy? Should a rescan to confirm viability or serial hCGs to confirm location be advised? This ambiguity needs clarification. Some women seem to be being over-investigated which takes time, costs money and generates undue anxiety but, of greater concern, we are under-investigating other women, which jeopardizes patient safety.

Based on the results of the reliability study reported in chapter eight, and of the diagnostic accuracy study reported in chapter seven, we propose that the current definitions for pregnancies of unknown location and uncertain viability be refined to include the ultrasonographic finding of a small 'empty sac' and we suggest that the definitions take into consideration the double decidual sac sign. For example, if there is ultrasonographic evidence of an 'empty sac' which demonstrates the double decidual sac sign, according to the results of our diagnostic accuracy study, it is definitely an intrauterine pregnancy, and it should therefore be regarded as a pregnancy of uncertain viability and follow-up should consist of a repeat ultrasound scan at an interval of at least seven days. If there is ultrasonographic evidence of an 'empty sac'

which does not demonstrate the double decidual sac sign, then notwithstanding the fact that pseudosacs are extremely rare [244], it should prudently be regarded as a potential ectopic pregnancy and managed according to pregnancy of unknown location protocols with serial serum hCGs measurements. Adopting this strategy will enhance the clinical care of women with the ultrasonographic finding of an empty sac in early pregnancy by improving consistency, appropriately rationalising follow-up and minimising error.

Arguably, this doesn't actually reduce uncertainty; merely convert what might previously have been regarded as a pregnancy of unknown location to one of uncertain viability. However, in clinical practice it is the pregnancies of unknown location that are more concerning given the immediate threats to health and potential threats to future fertility. It is also the pregnancies of unknown location that have a more haphazard follow-up, with different institutions, and even different individuals within the same institution, using different algorithms [34, 37, 38]. This follow-up also utilizes more resources and is therefore also more costly.

The aim of our final study, reported in chapter nine, was to minimise the duration of uncertainty for women with pregnancies of uncertain viability using novel serum biomarkers. Although subject to confirmatory studies, our preliminary findings suggest that serum concentrations of Ang-2 and Flt-1, and to a lesser extent Ang-1, may be able to predict outcome in women with pregnancies of uncertain viability. The findings of our prospective study incorporating 94 women presenting with abdominal pain and/or vaginal bleeding in early pregnancy and diagnosed, following transvaginal ultrasound, with pregnancies of uncertain viability, suggest that women with low serum concentrations of Ang-2 and Flt-1 are significantly more likely to have viable intrauterine pregnancies than women with high serum concentrations.

If these findings are validated and appropriate threshold levels for our biomarkers determined, it may be possible to minimise the duration of uncertainty for women with pregnancies of uncertain viability to hours (i.e. the turn-around time for the blood work) rather than weeks. It may be that on their own our biomarkers are not capable of generating the high degree of accuracy required, but in combination with other demographic, clinical,

serological and radiological parameters they may be used to generate a prognostic research model capable of predicting individual risk of a future outcome which would be clinically welcome, and if found to be cost-effective, extremely useful.

Hence, if, using ultrasound and the double decidual sac sign, we can successfully convert the vast majority of uncertainties regarding location to those regarding viability, we can then potentially use our biomarkers to minimise the duration of uncertainty to hours rather than weeks and we will therefore, using a combination of approaches, have achieved the overall aim of this thesis in minimising diagnostic uncertainties in early pregnancy.

## 10.2 Future Research Recommendations

This thesis has made steps towards improving the care of women with diagnostic uncertainties in early pregnancy. However, all research, and this work is no exception, invariably produces further questions that need to be answered. The cycle of identifying and prioritizing research agendas, carrying out research to answer a question and subsequently identifying further areas for exploration is paramount to moving practice forward. For women who experience diagnostic uncertainties in early pregnancy, this work is a stepping-stone, and it is with great hope that this project will stimulate others to take an interest in this area and help to make a difference to the clinical care of women through continuing the research cycle.

Themes identified from each of our individual studies should be further explored on a larger scale. Our understanding of diagnostic uncertainties in early pregnancy, the distress they cause and how to minimise them, could be improved by future research to address the following questions:

- Do women given a certain negative diagnosis following an initial uncertain diagnosis fare better psychologically than their counterparts who are given an immediate diagnosis of either a non-viable or ectopic pregnancy. Does the period of uncertainty prepare them better for a negative diagnosis or does a diagnosis of a non-viable or ectopic pregnancy still come as a shock?

- What happens to those women given a certain positive diagnosis after an initial uncertain diagnosis – do anxiety levels quickly abate or do they remain elevated for the remainder of the pregnancy?
- Do simple measures to reduce anxiety, such as those described in the previous section, make a significant difference to women given an uncertain diagnosis following the ultrasound? Are any additional resources required to do this justified?
- In a larger sample, containing more ectopic pregnancies, how accurate is the double decidual sac sign for predicting an intrauterine pregnancy prior to visualization of embryonic contents?
- If we adopt a strategy for the management of empty sacs based on the presence or absence of the double decidual sac sign, what are the chances of delaying the diagnosis of an ectopic pregnancy?
- Are ultrasonographers ready to incorporate the double decidual sac sign into their clinical practice? Does training improve the inter-observer reliability associated with the double decidual sac sign?
- Is it possible to replicate our findings with regards to the prognostic capability of Ang-2 and Flt-1? If so, what threshold levels should we use and how accurate are they?
- Are these new approaches to minimise diagnostic uncertainties in early pregnancy cost effective?

## 10.3 Concluding Remarks

Minimising diagnostic uncertainties in early pregnancy has many advantages, the importance of which cannot be over-emphasized. Not only will it decrease anxiety for women and their partners but it also permits Early Pregnancy Assessment Units to redistribute their workload so that they can see more new patients, more quickly, as they will be seeing fewer women for follow-up appointments. Earlier diagnosis of non-viable and ectopic pregnancies will minimise the number of women presenting to hospital in a state of haemodynamic compromise and enable women to choose more conservative forms of management for their early pregnancy complications if they wish.

# References


1. Larsen, W.J., *Human embryology*. 2nd ed. 1997, New York: Churchill Livingstone. xvi, 512 p.
2. Dey, S.K., et al., *Molecular cues to implantation.* Endocr Rev, 2004. **25**(3): p. 341-73.
3. Quenby, S. and J.J. Brosens, *Human implantation: a tale of mutual maternal and fetal attraction.* Biol Reprod, 2013. **88**(3): p. 81.
4. Salker, M.S., et al., *Disordered IL-33/ST2 activation in decidualizing stromal cells prolongs uterine receptivity in women with recurrent pregnancy loss.* PLoS One, 2012. **7**(12): p. e52252.
5. Gellersen, B., et al., *Human endometrial stromal cell-trophoblast interactions: mutual stimulation of chemotactic migration and promigratory roles of cell surface molecules CD82 and CEACAM1.* Biol Reprod, 2013. **88**(3): p. 80.
6. Teklenburg, G., et al., *Natural selection of human embryos: decidualizing endometrial stromal cells serve as sensors of embryo quality upon implantation.* PLoS One, 2010. **5**(4): p. e10258.
7. Grewal, S., et al., *Implantation of the human embryo requires Rac1-dependent endometrial stromal cell migration.* Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16189-94.
8. Schwenke, M., et al., *Control of human endometrial stromal cell motility by PDGF-BB, HB-EGF and trophoblast-secreted factors.* PLoS One, 2013. **8**(1): p. e54336.
9. Grewal, S., et al., *Human endometrial stromal cell rho GTPases have opposing roles in regulating focal adhesion turnover and embryo invasion in vitro.* Biol Reprod, 2010. **83**(1): p. 75-82.
10. Gellersen, B., et al., *Invasiveness of human endometrial stromal cells is promoted by decidualization and by trophoblast-derived signals.* Hum Reprod, 2010. **25**(4): p. 862-73.
11. Brosens, J.J. and B. Gellersen, *Something new about early pregnancy: decidual biosensoring and natural embryo selection.* Ultrasound Obstet Gynecol, 2010. **36**(1): p. 1-5.
12. Timor-Tritsch, I.E., D. Farine, and M.G. Rosen, *A close look at early embryonic development with the high-frequency transvaginal transducer.* Am J Obstet Gynecol, 1988. **159**(3): p. 676-81.
13. Donald, I. and U. Abdulla, *Ultrasonics in obstetrics and gynaecology.* Br J Radiol, 1967. **40**(476): p. 604-11.
14. Kaur, A. and A. Kaur, *Transvaginal ultrasonography in first trimester of pregnancy and its comparison with transabdominal ultrasonography.* J Pharm Bioallied Sci, 2011. **3**(3): p. 329-38.
15. Rowling, S.E., et al., *Sonography during early pregnancy: dependence of threshold and discriminatory values on transvaginal transducer frequency.* AJR Am J Roentgenol, 1999. **172**(4): p. 983-8.
16. Benacerraf, B.R., T.D. Shipp, and B. Bromley, *Does the 10-MHz transvaginal transducer improve the diagnostic certainty that an intrauterine fluid collection is a true gestational sac?* J Clin Ultrasound, 1999. **27**(7): p. 374-7.
17. Shapiro, B.S., et al., *A model-based prediction for transvaginal ultrasonographic identification of early intrauterine pregnancy.* Am J Obstet Gynecol, 1992. **166**(5): p. 1495-500.
18. Marks, W.M., et al., *The decidual cast of ectopic pregnancy: a confusing ultrasonographic appearance.* Radiology, 1979. **133**(2): p. 451-4.
19. Lazarus, E., *What's new in first trimester ultrasound.* Radiol Clin North Am, 2003. **41**(4): p. 663-79.

20. Jauniaux, E., et al., *Development of the secondary human yolk sac: correlation of sonographic and anatomical features.* Hum Reprod, 1991. **6**(8): p. 1160-6.
21. Jurkovic, D., K. Gruboeck, and S. Campbell, *Ultrasound features of normal early pregnancy development.* Curr Opin Obstet Gynecol, 1995. **7**(6): p. 493-504.
22. Tezuka, N., et al., *Embryonic heart rates: development in early first trimester and clinical evaluation.* Gynecol Obstet Invest, 1991. **32**(4): p. 210-2.
23. Levi, C.S., et al., *Endovaginal US: demonstration of cardiac activity in embryos of less than 5.0 mm in crown-rump length.* Radiology, 1990. **176**(1): p. 71-4.
24. Goldstein, S.R., *Significance of cardiac activity on endovaginal ultrasound in very early embryos.* Obstet Gynecol, 1992. **80**(4): p. 670-2.
25. Brown, D.L., et al., *Diagnosis of early embryonic demise by endovaginal sonography.* J Ultrasound Med, 1990. **9**(11): p. 631-6.
26. Braunstein, G.D., et al., *Serum human chorionic gonadotropin levels throughout normal pregnancy.* Am J Obstet Gynecol, 1976. **126**(6): p. 678-81.
27. Daya, S., et al., *Transvaginal ultrasound scanning in early pregnancy and correlation with human chorionic gonadotropin levels.* J Clin Ultrasound, 1991. **19**(3): p. 139-42.
28. Nyberg, D.A., et al., *Abnormal pregnancy: early diagnosis by US and serum chorionic gonadotropin levels.* Radiology, 1986. **158**(2): p. 393-6.
29. Barnhart, K., et al., *Prompt diagnosis of ectopic pregnancy in an emergency department setting.* Obstet Gynecol, 1994. **84**(6): p. 1010-5.
30. Cacciatore, B., U.H. Stenman, and P. Ylostalo, *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): p. 904-8.
31. Condous, G., et al., *Role of biochemical and ultrasonographic indices in the management of pregnancies of unknown location.* Ultrasound Obstet Gynaecol, 2002. **20 Suppl 1**: p. 36-37.
32. Barnhart, K., et al., *Decline of serum human chorionic gonadotropin and spontaneous complete abortion: defining the normal curve.* Obstet Gynecol, 2004. **104**(5 Pt 1): p. 975-81.
33. Sagili, H. and K. Mohamed, *Pregnancy of unknown location: an evidence based approach to management.* The Obstetrician and Gynaecologist, 2008. **10**: p. 224-230.
34. Kadar, N., B.V. Caldwell, and R. Romero, *A method of screening for ectopic pregnancy and its indications.* Obstet Gynecol, 1981. **58**(2): p. 162-6.
35. Kadar, N., M. Freedman, and M. Zacher, *Further observations on the doubling time of human chorionic gonadotropin in early asymptomatic pregnancies.* Fertil Steril, 1990. **54**(5): p. 783-7.
36. Fritz, M.A. and S.M. Guo, *Doubling time of human chorionic gonadotropin (hCG) in early normal pregnancy: relationship to hCG concentration and gestational age.* Fertil Steril, 1987. **47**(4): p. 584-9.
37. Barnhart, K.T., et al., *Symptomatic patients with an early viable intrauterine pregnancy: HCG curves redefined.* Obstet Gynecol, 2004. **104**(1): p. 50-5.
38. Seeber, B.E., et al., *Application of redefined human chorionic gonadotropin curves for the diagnosis of women at risk for ectopic pregnancy.* Fertil Steril, 2006. **86**(2): p. 454-9.
39. Hahlin, M., J. Thorburn, and I. Bryman, *The expectant management of early pregnancies of uncertain site.* Hum Reprod, 1995. **10**(5): p. 1223-7.
40. Banerjee, S., et al., *Expectant management of early pregnancies of unknown location: a prospective evaluation of methods to predict spontaneous resolution of pregnancy.* BJOG, 2001. **108**(2): p. 158-63.

41. Condous, G., E. Okaro, and T. Bourne, *The conservative management of early pregnancy complications: a review of the literature.* Ultrasound Obstet Gynecol, 2003. **22**(4): p. 420-30.
42. Mol, B.W., et al., *The accuracy of single serum progesterone measurement in the diagnosis of ectopic pregnancy: a meta-analysis.* Hum Reprod, 1998. **13**(11): p. 3220-7.
43. Kirk, E., G. Condous, and T. Bourne, *Pregnancies of unknown location.* Best Pract Res Clin Obstet Gynaecol, 2009. **23**(4): p. 493-9.
44. Cahill, D.J., R. Swingler, and P. Wardle, *Bleeding and Pain in Early Pregnancy*, in *High Risk Pregnancy Management Options*, D. James, Editor. 2011, Elsevier. p. 57-73.
45. Macklon, N.S., J.P.M. Geraedts, and B.C.J.M. Fauser, *Conception to ongoing pregnancy: the 'black box' of early pregnancy loss.* Human Reproduction Update, 2002. **8**(4): p. 333-343.
46. Wilcox, A.J., et al., *Incidence of early loss of pregnancy.* New England Journal of Medicine, 1988. **319**: p. 189-194.
47. Zinamen, M.J., et al., *Estimates of human fertility and pregnancy loss.* Fertil Steril, 1996. **65**: p. 503-509.
48. Impey, L., *Disorders of Early Pregnancy*, in *Obstetrics and Gynaecology*, L. Impey, Editor. 2002, Blackwell Science. p. 94-102.
49. Kolte, A.M., et al., *Terminology for pregnancy loss prior to viability: a consensus statement from the ESHRE early pregnancy special interest group.* Hum Reprod, 2015. **30**(3): p. 495-8.
50. Ptettyman, R.J., C.J. Cordle, and G.D. Cook, *A three-month follow-up of psychological morbidity after early miscarriage.* British Journal of Medical Psychology, 1993. **66**(4): p. 363-372.
51. *Saving Mothers' Lives: Reviewing maternal deaths to make motherhood safer: 2006–2008.* BJOG: An International Journal of Obstetrics & Gynaecology, 2011. **118**: p. 1-203.
52. Trinder, J., et al., *Management of miscarriage: expectant, medical, or surgical? Results of randomised controlled trial (miscarriage treatment (MIST) trial).* BMJ, 2006. **332**(7552): p. 1235-40.
53. Faller, E., et al., *[Full term abdominal pregnancy].* J Gynecol Obstet Biol Reprod (Paris), 2006. **35**(7): p. 732-5.
54. Nama, V., et al., *Secondary abdominal appendicular ectopic pregnancy.* J Minim Invasive Gynecol, 2007. **14**(4): p. 516-7.
55. Nielsen, K.G. and A.T. Pedersen, *An ectopic pregnancy under the liver.* BMJ Case Rep, 2010. **2010**.
56. Gang, G., Y. Yudong, and G. Zhang, *Successful laparoscopic management of early splenic pregnancy: case report and review of literature.* J Minim Invasive Gynecol, 2010. **17**(6): p. 794-7.
57. Chopra, S., et al., *Primary omental pregnancy: case report and review of literature.* Arch Gynecol Obstet, 2009. **279**(4): p. 441-2.
58. Vermesh, M., et al., *Management of unruptured ectopic gestation by linear salpingostomy: a prospective, randomized clinical trial of laparoscopy versus laparotomy.* Obstet Gynecol, 1989. **73**(3 Pt 1): p. 400-4.
59. Lundorff, P., et al., *Laparoscopic surgery in ectopic pregnancy. A randomized trial versus laparotomy.* Acta Obstet Gynecol Scand, 1991. **70**(4-5): p. 343-8.
60. Gray, D.T., et al., *A cost-effectiveness study of a randomised trial of laparoscopy versus laparotomy for ectopic pregnancy.* Lancet, 1995. **345**(8958): p. 1139-43.
61. Mol, F., et al., *Salpingotomy versus salpingectomy in women with tubal pregnancy (ESEP study): an open-label, multicentre, randomised controlled trial.* Lancet, 2014. **383**(9927): p. 1483-9.

62. Donald, I., *Ultrasonics in diagnosis (Sonar).* Proc R Soc Med, 1969. **62**(2): p. 442-446.
63. Bottomley, C., A. Daemen, and F. Mukri, *Assessing first trimester growth: the influence of ethnic background and maternal age.* Hum Reprod, 2009. **24**(2): p. 284-290.
64. (RCOG), R.C.o.O.a.G., *Green-top guideline no. 25 The management of early pregnancy loss*. 2006, London: RCOG Press.
65. ACOG, *Practice Bulletin No. 101: Ultrasonography in prenancy.* Obstet Gynecol, 2009. **113**(2): p. 451-61.
66. THKCoOa, G., *Guidelines for first trimester ultrasound examination: part 1.* HKCOG Guidelines, 2004. **10**(1): p. 1-6.
67. Morin, L.a.V.d.H., M.C., *SOGC clinical practice guidelines: ultrasound evaluation of first trimester complications.* Int J Gynaecol Obstet, 2006. **93**(1): p. 77-81.
68. Jeve, Y., Rana, R., Bhide A., and Thangaratinam, S., *Accuracy of first-trimester ultrasound in the diagnosis of early embryonic demise: a systematic review.* Ultrasound Obstet Gynecol, 2011. **38**(5): p. 489-96.
69. Abdallah, Y., Daemen, A., Kirk, E., Pexsters, A., Naji, O., Stalder, C., Gould, D., Ahmed, S., Guha, S., Syed, S., Bottomley, C., Timmerman, D. and Bourne, T., *Limitations of current definitions of miscarriage using mean gestational sac diameter and crown–rump length measurements: a multicenter observational study.* Ultrasound Obstet Gynecol, 2011. **38**(5): p. 497-502.
70. Pexsters A, L.J., Van Schoubroeck D, Bottomley C, Van Calster B, Van Huffel S, Abdallah Y, D'Hooghe T, Lees C, Timmerman D, Bourne T., *Clinical implications of intra- and interobserver reproducibility of transvaginal sonographic measurement of gestational sac and crown-rump length at 6-9 weeks' gestation.* Ultrasound Obstet Gynecol, 2011. **38**(5): p. 510-15.
71. Kirk, E., et al., *The diagnostic effectiveness of an initial transvaginal scan in detecting ectopic pregnancy.* Hum Reprod, 2007. **22**(11): p. 2824-8.
72. Condous, G., et al., *The accuracy of transvaginal ultrasonography for the diagnosis of ectopic pregnancy prior to surgery.* Hum Reprod, 2005. **20**(5): p. 1404-9.
73. Braffman, B.H., et al., *Emergency department screening for ectopic pregnancy: a prospective US study.* Radiology, 1994. **190**: p. 797-802.
74. Atri, M., et al., *Effect of transvaginal sonography on the use of invasive procedures for evaluating patients with a clinical diagnosis of ectopic pregnancy.* J Clin Ultrasound, 2003. **31**(1): p. 1-8.
75. Shalev, E., et al., *Transvaginal sonography as the ultimate diagnostic tool for the management of ectopic pregnancy: experience with 840 cases.* Fertil Steril, 1998. **69**(1): p. 62-5.
76. Mavrelos, D., et al., *Efficacy and safety of a clinical protocol for expectant management of selected women diagnosed with a tubal ectopic pregnancy.* Ultrasound Obstet Gynecol, 2013. **42**(1): p. 102-7.
77. Atri, M., et al., *Role of endovaginal sonography in the diagnosis and management of ectopic pregnancy.* Radiographics, 1996. **16**(4): p. 755-74; discussion 775.
78. Kirk, E., et al., *Why are some ectopic pregnancies characterized as pregnancies of unknown location at the initial transvaginal ultrasound examination?* Acta Obstet Gynecol Scand, 2008. **87**(11): p. 1150-4.
79. Nahum, G.G., *Rudimentary uterine horn pregnancy. The 20th-century worldwide experience of 588 cases.* J Reprod Med, 2002. **47**(2): p. 151-63.
80. Jermy, K., et al., *The conservative management of interstitial pregnancy.* BJOG, 2004. **111**(11): p. 1283-8.
81. Rottem, S., et al., *Criteria for transvaginal sonographic diagnosis of ectopic pregnancy.* J Clin Ultrasound, 1990. **18**(4): p. 274-9.

82. Ushakov, F.B., et al., *Cervical pregnancy: past and future.* Obstet Gynecol Surv, 1997. **52**(1): p. 45-59.
83. Jurkovic, D., E. Hacket, and S. Campbell, *Diagnosis and treatment of early cervical pregnancy: a review and a report of two cases treated conservatively.* Ultrasound Obstet Gynecol, 1996. **8**(6): p. 373-80.
84. Timor-Tritsch, I.E., et al., *Successful management of viable cervical pregnancy by local injection of methotrexate guided by transvaginal ultrasonography.* Am J Obstet Gynecol, 1994. **170**(3): p. 737-9.
85. Comstock, C., K. Huston, and W. Lee, *The ultrasonographic appearance of ovarian ectopic pregnancies.* Obstet Gynecol, 2005. **105**(1): p. 42-5.
86. Shiau, C.S., C.L. Hsieh, and M.Y. Chang, *Primary ovarian pregnancy.* Int J Gynaecol Obstet, 2007. **96**(2): p. 127.
87. Rotas, M.A., S. Haberman, and M. Levgur, *Cesarean scar ectopic pregnancies: etiology, diagnosis, and management.* Obstet Gynecol, 2006. **107**(6): p. 1373-81.
88. Godin, P.A., S. Bassil, and J. Donnez, *An ectopic pregnancy developing in a previous caesarian section scar.* Fertil Steril, 1997. **67**(2): p. 398-400.
89. Jurkovic, D., et al., *First-trimester diagnosis and management of pregnancies implanted into the lower uterine segment Cesarean section scar.* Ultrasound Obstet Gynecol, 2003. **21**(3): p. 220-7.
90. Timor-Tritsch, I.E., et al., *The diagnosis, treatment, and follow-up of cesarean scar pregnancy.* Am J Obstet Gynecol, 2012. **207**(1): p. 44 e1-13.
91. Seow, K.M., J.L. Hwang, and Y.L. Tsai, *Ultrasound diagnosis of a pregnancy in a Cesarean section scar.* Ultrasound Obstet Gynecol, 2001. **18**(5): p. 547-9.
92. Timor-Tritsch, I.E. and A. Monteagudo, *Unforeseen consequences of the increasing rate of cesarean deliveries: early placenta accreta and cesarean scar pregnancy. A review.* Am J Obstet Gynecol, 2012. **207**(1): p. 14-29.
93. Vial, Y., P. Petignat, and P. Hohlfeld, *Pregnancy in a cesarean scar.* Ultrasound Obstet Gynecol, 2000. **16**(6): p. 592-3.
94. Osborn, D.A., T.R. Williams, and B.M. Craig, *Cesarean scar pregnancy: sonographic and magnetic resonance imaging findings, complications, and treatment.* J Ultrasound Med, 2012. **31**(9): p. 1449-56.
95. Gerli, S., et al., *Early ultrasonographic diagnosis and laparoscopic treatment of abdominal pregnancy.* Eur J Obstet Gynecol Reprod Biol, 2004. **113**(1): p. 103-5.
96. Worley, K.C., M.D. Hnat, and F.G. Cunningham, *Advanced extrauterine pregnancy: diagnostic and therapeutic challenges.* Am J Obstet Gynecol, 2008. **198**(3): p. 297 e1-7.
97. Felmus, L.B. and P. Pedowitz, *Interstitial pregnancy; a survey of 45 cases.* Am J Obstet Gynecol, 1953. **66**(6): p. 1271-9.
98. Eddy, C.A. and C.J. Pauerstein, *Anatomy and physiology of the fallopian tube.* Clin Obstet Gynecol, 1980. **23**(4): p. 1177-93.
99. Tulandi, T. and A. Saleh, *Surgical management of ectopic pregnancy.* Clin Obstet Gynecol, 1999. **42**(1): p. 31-8; quiz 55-6.
100. Moore, K.L., A.M.R. Agur, and A.F. Dalley, *Essential clinical anatomy*. 4th ed. 2011, Baltimore, MD: Lippincott Williams & Wilkins. xxviii, 703 p.
101. Ackerman, T.E., et al., *Interstitial line: sonographic finding in interstitial (cornual) ectopic pregnancy.* Radiology, 1993. **189**(1): p. 83-7.
102. Araujo Junior, E., et al., *Three-dimensional transvaginal sonographic diagnosis of early and asymptomatic interstitial pregnancy.* Arch Gynecol Obstet, 2007. **275**(3): p. 207-10.
103. Rastogi, R., et al., *Interstitial ectopic pregnancy: A rare and difficult clinicosonographic diagnosis.* J Hum Reprod Sci, 2008. **1**(2): p. 81-2.

104. Filhastre, M., et al., *Interstitial pregnancy: role of MRI.* Eur Radiol, 2005. **15**(1): p. 93-5.
105. Tamai, K., T. Koyama, and K. Togashi, *MR features of ectopic pregnancy.* Eur Radiol, 2007. **17**(12): p. 3236-46.
106. Seeber, B.E. and K.T. Barnhart, *Suspected ectopic pregnancy.* Obstet Gynecol, 2006. **107**(2 Pt 1): p. 399-413.
107. McCord, M.L., et al., *Single serum progesterone as a screen for ectopic pregnancy: exchanging specificity and sensitivity to obtain optimal test performance.* Fertil Steril, 1996. **66**(4): p. 513-6.
108. Barnhart, K., et al., *Pregnancy of unknown location: a consensus statement of nomenclature, definitions, and outcome.* Fertil Steril, 2011. **95**(3): p. 857-66.
109. Kirk, E. and T. Bourne, *Pregnancy of Unknown Location.* Obstetrics, Gynaecology and Reproductive Medicine, 2008. **19**(3): p. 80-83.
110. Condous, G., et al., *Pregnancies of unknown location: consensus statement.* Ultrasound Obstet Gynecol, 2006. **28**(2): p. 121-2.
111. Banerjee, S., et al., *The expectant management of women with early pregnancy of unknown location.* Ultrasound Obstet Gynecol, 1999. **14**(4): p. 231-6.
112. Banerjee S, A.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**: p. 158-163.
113. Kirk, E., et al., *Rationalizing the follow-up of pregnancies of unknown location.* Hum Reprod, 2007. **22**(6): p. 1744-50.
114. Kirk, E. and T. Bourne, *Diagnosis of ectopic pregnancy with ultrasound.* Best Pract Res Clin Obstet Gynaecol, 2009. **23**(4): p. 501-8.
115. Condous, G., et al., *Failing pregnancies of unknown location: a prospective evaluation of the human chorionic gonadotrophin ratio.* BJOG, 2006. **113**(5): p. 521-7.
116. Morse, C.B., et al., *Performance of human chorionic gonadotropin curves in women at risk for ectopic pregnancy: exceptions to the rules.* Fertil Steril, 2012. **97**(1): p. 101-6 e2.
117. Day, A., et al., *Use of serum progesterone measurements to reduce need for follow-up in women with pregnancies of unknown location.* Ultrasound Obstet Gynecol, 2009. **33**(6): p. 704-10.
118. Kobayashi, F., et al., *Maternal serum CA125 levels in early intrauterine and tubal pregnancies.* Arch Gynecol Obstet, 1993. **252**(4): p. 185-9.
119. Schmidt, T., et al., *Prognostic value of repeated serum CA 125 measurements in first trimester pregnancy.* Eur J Obstet Gynecol Reprod Biol, 2001. **97**(2): p. 168-73.
120. Condous, G., et al., *Do levels of serum cancer antigen 125 and creatine kinase predict the outcome in pregnancies of unknown location?* Hum Reprod, 2005. **20**(12): p. 3348-54.
121. Chandra, L. and A. Jain, *Maternal serum creatine kinase as a biochemical marker of tubal pregnancy.* Int J Gynaecol Obstet, 1995. **49**(1): p. 21-3.
122. Duncan, W.C., et al., *Measurement of creatine kinase activity and diagnosis of ectopic pregnancy.* Br J Obstet Gynaecol, 1995. **102**(3): p. 233-7.
123. Korhonen, J., et al., *Failure of creatine kinase to predict ectopic pregnancy.* Fertil Steril, 1996. **65**(5): p. 922-4.
124. Qasim, S.M., et al., *Evaluation of serum creatine kinase levels in ectopic pregnancy.* Fertil Steril, 1996. **65**(2): p. 443-5.
125. Vandermolen, D.T. and J.F. Borzelleca, *Serum creatine kinase does not predict ectopic pregnancy.* Fertil Steril, 1996. **65**(5): p. 916-21.

126. Zorn, J.R., et al., *Evaluation of maternal plasma creatine kinase activity as a marker of abnormal early pregnancy.* Hum Reprod, 1997. **12**(11): p. 2534-7.
127. Kirk, E., et al., *The use of serum inhibin A and activin A levels in predicting the outcome of 'pregnancies of unknown location'.* Hum Reprod, 2009. **24**(10): p. 2451-6.
128. Chetty, M., et al., *The use of novel biochemical markers in predicting spontaneously resolving 'pregnancies of unknown location'.* Hum Reprod, 2011. **26**(6): p. 1318-23.
129. Horne, A.W., et al., *Evaluation of ADAM-12 as a diagnostic biomarker of ectopic pregnancy in women with a pregnancy of unknown location.* PLoS One, 2012. **7**(8): p. e41442.
130. Condous, G., et al., *The use of a new logistic regression model for predicting the outcome of pregnancies of unknown location.* Hum Reprod, 2004. **19**(8): p. 1900-10.
131. Gevaert, O., 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): p. 1824-31.
132. Condous, G., et al., *Prospective cross-validation of three methods of predicting failing pregnancies of unknown location.* Hum Reprod, 2007. **22**(4): p. 1156-60.
133. Condous, G., et al., *Prediction of ectopic pregnancy in women with a pregnancy of unknown location.* Ultrasound Obstet Gynecol, 2007. **29**(6): p. 680-7.
134. Barnhart, K.T., 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): p. 2434-40.
135. Condous, G., et al., *A prospective evaluation of a single-visit strategy to manage pregnancies of unknown location.* Hum Reprod, 2005. **20**(5): p. 1398-403.
136. Cordina, M., et al., *Introduction of a single visit protocol in the management of selected patients with pregnancy of unknown location: a prospective study.* BJOG, 2011. **118**(6): p. 693-7.
137. Bottomley, C. and T. Bourne, *Diagnosing miscarriage.* Best Practice & Research in Clinical Obstetrics & Gynaecology, 2009. **23**(4): p. 463-77.
138. Hinshaw, K., *Greentop Guideline No. 25: The management of early pregnancy loss*. RCOG Greentop Guideline. October 2006, London: RCOG Press.
139. Hately, W., J. Case, and S. Campbell, *Establishing the death of an embryo by ultrasound: a report of a public inquiry with recommendations.* Ultrasound Obstet Gynaecol, 1995. **5**: p. 353-357.
140. Gynaecologists, R.C.o.O.a., *Management of early pregnancy loss. Guideline no. 25.* RCOG Press, 2006.
141. Nyberg, D.A. and R.A. Filly, *Predicting pregnancy failure in 'empty' gestational sacs.* Ultrasound Obstet Gynaecol, 2003. **21**: p. 9-12.
142. Savitz, D.A., et al., *Comparison of pregnancy dating by last menstrual period, ultrasound scanning and their combination.* Am J Obstet Gynecol, 2002. **187**: p. 1660-1666.
143. Warren, W.B., et al., *Dating the early pregnancy by sequential appearance of embryonic structures.* Am J Obstet Gynecol, 1989. **161**: p. 747-753.
144. Dietz, P.M., et al., *A comparison of LMP-based and ultrasound based estimates of gestational age using linked California livebirth and prenatal screening records* Paediatr Perinat Epidemiol, 2007. **21**(Suppl 2): p. 62-71.

145. Mukri, F., et al., *Evidence of early first-trimester growth restriction in pregnancies that subsequently end in miscarriage.* BJOG, 2008. **115**(10): p. 1273-8.
146. Bora, S.A., et al., *Twin growth discrepancy in early pregnancy.* Ultrasound Obstet Gynaecol, 2009. **34**: p. 38-42.
147. Falco, P., et al., *Sonography of pregnancies with first trimester bleeding and a viable embryo: a study of prognostic indicators by logistic regression analysis.* Ultrasound Obstet Gynaecol, 1996. **7**: p. 165-169.
148. Reljic, M., *The significance of crown-rump length measurement for predicting adverse pregnancy outcome of threatened abortion.* Ultrasound Obstet Gynaecol, 2001. **17**: p. 510-512.
149. Blaas, H.G., S.H. Eik-Nes, and J.B. Bremnes, *The growth of the human embryo. A longitudinal biometric assessment from 7 to 12 weeks of gestation.* Ultrasound Obstet Gynaecol, 1998. **12**: p. 346-354.
150. Bateman, B.G., et al., *Vaginal sonography findings and hCG dynamics of early intrauterine and tubal pregnancies.* Obstet Gynecol, 1990. **75**(3 Pt 1): p. 421-7.
151. Elson, J., et al., *Prediction of early pregnancy viability in the absence of an ultrasonically detectable embryo.* Ultrasound Obstet Gynaecol, 2003. **21**: p. 57-61.
152. McCord, M.L., et al., *Single serum progesterone as a screen for ectopic pregnancy: exchanging specificity and sensitivity to obtaine optimal test performance.* Fertil Steril, 1996. **66**: p. 513-516.
153. Tongsong, T., C. Wanapirak, and J. Srisomboon, *Transvaginal ultrasound in threatened abortions with empty gestational sacs.* International Journal of Gynaecology & Obstetrics, 1994. **46**(3): p. 297-301.
154. Bernard, K.G. and P.L. Cooperberg, *Sonographic differentiation between blighted ovum and early viable pregnancy.* Am J Roentgenol, 1985. **144**: p. 597-602.
155. Küçük, T., et al., *Yolk sac size and shape as predictors of poor pregnancy outcome.* J Perinat Med, 1999. **27**: p. 316-320.
156. Rôlo, L., et al., *Yolk sac volume assessed by three-dimensional ultrasonography using the VOCAL method.* Acta Obstet Gynecol Scand, 2008. **87**: p. 499-502.
157. Babinszki, A., et al., *Three-dimensional measurement of gestational and yolk sac volumes as predictors of pregnancy outcome in the first trimester.* Am J Perinatol 2001. **18**: p. 203-211.
158. Figueras, F., et al., *Three-dimensional yolk and gestational sac volume: a prospective study of prognostic value.* J Reprod Med, 2003. **48**(4): p. 252-256.
159. Goldstein, B.H., *Endovaginal sonography in very early pregnancy: new observations.* Radiology, 1990. **176**: p. 7-8.
160. Pennell, R.G., L. Needleman, and T. Pajak, *Prospective comparison of vaginal and abdominal sonography in normal early pregnancy.* J Ultrasound Med, 1991. **10**: p. 63-67.
161. Schouwink, M.H., et al., *Ultrasonographic criteria for non-viability of first trimester intra-uterine pregnancy.* Early Pregnancy, 2000. **4**: p. 203-213.
162. Filly, M.R., et al., *The yolk stalk sign: evidence of death in small embryos without heartbeats.* J Ultrasound Med, (of Publication: Feb 2010): p. 29 (2) (pp 237-241), 2010.
163. Koornstra, G. and E. N, *Echography in the first pregnancy trimester has prognostic value.* Ned Tijdschr Geneeskd, 1991. **135**: p. 2231-2235.
164. Nyberg, D.A., et al., *Distinguishing normal from abnormal gestational sac growth in early pregnancy.* J Ultrasound Med, 1987. **6**(1): p. 23-7.
165. C. Bottomley, C., et al., *Functional linear discriminant analysis: a new longitudinal approach to the assessment of embryonic growth.* Hum Reprod 2009. **24**: p. 278-283.

166. Abdallah, Y., et al., *Gestational sac and embryonic growth are not useful as criteria to define miscarriage: a multicenter observational study.* Ultrasound Obstet Gynaecol, 2011. **38**: p. 503-509.
167. Bottomley, C., et al., *A model and scoring system to predict outcome of intrauterine pregnancies of uncertain viability.* Ultrasound Obstet Gynecol, (of Publication: May 2011): p. 37 (5) (pp 588-595), 2011.
168. Davison, A.Z., et al., *The psychological effects and patient acceptability of a test to predict viability in early pregnancy: a prospective randomised study.* Eur J Obstet Gynecol Reprod Biol, 2014. **178**: p. 95-9.
169. Campbell, S., et al., *Routine ultrasound screening for the prediction of gestational age.* Obstet Gynecol, 1985. **65**(5): p. 613-20.
170. Nikcevic, A.V., S.A. Tunkel, and K.H. Nicolaides, *Psychological outcomes following missed abortions and provision of follow-up care.* Ultrasound Obstet Gynecol, 1998. **11**(2): p. 123-8.
171. Mol, B.W. and F. Van der Veen, *A study of ruptured tubal ectopic pregnancy.* Obstet Gynecol, 1997. **90**(5): p. 866-7.
172. Bottomley, C., et al., *The optimal timing of an ultrasound scan to assess the location and viability of an early pregnancy.* Human Reproduction, (of Publication: August 2009): p. 24 (8) (pp 1811-1817), 2009.
173. Richardson, A., et al., *Anxiety associated with diagnostic uncertainty in early pregnancy.* Ultrasound Obstet Gynecol, 2017. **50**(2): p. 247-254.
174. Budner, S., *Intolerance of ambiguity as a personality variable.* J Pers, 1962. **30**: p. 29-50.
175. McIntosh, J., *Patients' awareness and desire for information about diagnosed but undisclosed malignant disease.* Lancet, 1976. **2**(7980): p. 300-3.
176. Mishel, M.H., *Perceived uncertainty and stress in illness.* Res Nurs Health, 1984. **7**(3): p. 163-71.
177. Suls, J. and B. Mullen, *Life events, perceived control and illness: the role of uncertainty.* J Human Stress, 1981. **7**(2): p. 30-4.
178. Rozanski, A., J.A. Blumenthal, and J. Kaplan, *Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy.* Circulation, 1999. **99**(16): p. 2192-217.
179. Kemeny, M.E. and M. Schedlowski, *Understanding the interaction between psychosocial stress and immune-related diseases: a stepwise progression.* Brain Behav Immun, 2007. **21**(8): p. 1009-18.
180. Buljevac, D., et al., *Prospective study on the relationship between infections and multiple sclerosis exacerbations.* Brain, 2002. **125**(Pt 5): p. 952-60.
181. Cohen, S., *Keynote Presentation at the Eight International Congress of Behavioral Medicine: the Pittsburgh common cold studies: psychosocial predictors of susceptibility to respiratory infectious illness.* Int J Behav Med, 2005. **12**(3): p. 123-31.
182. Chida, Y. and K. Vedhara, *Adverse psychosocial factors predict poorer prognosis in HIV disease: a meta-analytic review of prospective investigations.* Brain Behav Immun, 2009. **23**(4): p. 434-45.
183. Chida, Y., et al., *Do stress-related psychosocial factors contribute to cancer incidence and survival?* Nat Clin Pract Oncol, 2008. **5**(8): p. 466-75.
184. Maconochie, N., et al., *Risk factors for first trimester miscarriage--results from a UK-population-based case-control study.* BJOG, 2007. **114**(2): p. 170-86.
185. Lazarus, R.S., et al., *Stress and adaptational outcomes. The problem of confounded measures.* Am Psychol, 1985. **40**(7): p. 770-85.
186. Thapar, A.K. and A. Thapar, *Psychological sequelae of miscarriage: a controlled study using the general health questionnaire and the hospital anxiety and depression scale.* Br J Gen Pract, 1992. **42**(356): p. 94-6.
187. Conway, K. and G. Russell, *Couples' grief and experience of support in the aftermath of miscarriage.* Br J Med Psychol, 2000. **73 Pt 4**: p. 531-45.

188. Friedman, T., *Distress caused by miscarriage.* Br J Psychiatry, 1989. **155**: p. 567.
189. Lok, I.H. and R. Neugebauer, *Psychological morbidity following miscarriage.* Best Pract Res Clin Obstet Gynaecol, 2007. **21**(2): p. 229-47.
190. Geller, P.A., D. Kerns, and C.M. Klier, *Anxiety following miscarriage and the subsequent pregnancy: a review of the literature and future directions.* J Psychosom Res, 2004. **56**(1): p. 35-45.
191. Moscrop, A., et al., *Primary care follow-up and measured mental health outcomes among women referred for ultrasound assessment of pain and/or bleeding in early pregnancy: a quantitative questionnaire study.* BMJ Open, 2013. **3**(4).
192. Farren, J., et al., *Post-traumatic stress, anxiety and depression following miscarriage or ectopic pregnancy: a prospective cohort study.* BMJ Open, 2016. **6**(11): p. e011864.
193. Farhi, J., Z. Ben-Rafael, and D. Dicker, *Suicide after ectopic pregnancy.* N Engl J Med, 1994. **330**(10): p. 714.
194. Spielberger, C.D., *State-trait anxiety inventory : a comprehensive bibliography*. 1984, Palo Alto, CA: Consulting Psychologists Press. 102 p.
195. Beck, A.T., et al., *An inventory for measuring clinical anxiety: psychometric properties.* J Consult Clin Psychol, 1988. **56**(6): p. 893-7.
196. Knight, R.G., H.J. Waal-Manning, and G.F. Spears, *Some norms and reliability data for the State--Trait Anxiety Inventory and the Zung Self-Rating Depression scale.* Br J Clin Psychol, 1983. **22 (Pt 4)**: p. 245-9.
197. Addolorato, G., et al., *State and trait anxiety in women affected by allergic and vasomotor rhinitis.* J Psychosom Res, 1999. **46**(3): p. 283-9.
198. Kvaal, K., et al., *The Spielberger State-Trait Anxiety Inventory (STAI): the state scale in detecting mental disorders in geriatric patients.* Int J Geriatr Psychiatry, 2005. **20**(7): p. 629-34.
199. Fydrich, T., D. Dowdall, and D.L. Chambless, *Reliability and validity of the Beck Anxiety Inventory.* J Anxiety Disord, 1993. **6**: p. 55-61.
200. Creamer, M., J. Foran, and R. Bell, *The Beck Anxiety Inventory in a non-clinical sample.* Behav Res Ther, 1995. **33**(4): p. 477-85.
201. Osman, A., et al., *The Beck Anxiety Inventory: psychometric properties in a community population.* J Psychopath Behav Assess, 1993. **15**: p. 287-297.
202. VanDyke, M.M., et al., *Anxiety in rheumatoid arthritis.* Arthritis Rheum, 2004. **51**(3): p. 408-12.
203. Kabacoff, R.I., et al., *Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients.* J Anxiety Disord, 1997. **11**(1): p. 33-47.
204. Kennedy, B.L., et al., *Assessment of state and trait anxiety in subjects with anxiety and depressive disorders.* Psychiatr Q, 2001. **72**(3): p. 263-76.
205. Morin, C., et al., *The Beck Anxiety Inventory: psychometric properties with older adults.* J Clin Geropsychol, 1999. **5**: p. 19-29.
206. Brown, G.K., et al., *A comparison of focused and standard cognitive therapy for panic disorder.* J Anxiety Disord, 1997. **11**(3): p. 329-45.
207. Lee, Y.W., et al., *Impact of Psoriasis on Quality of Life: Relationship between Clinical Response to Therapy and Change in Health-related Quality of Life.* Ann Dermatol, 2010. **22**(4): p. 389-96.
208. Marteau, T.M. and H. Bekker, *The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI).* Br J Clin Psychol, 1992. **31 ( Pt 3)**: p. 301-6.
209. Tavakol, M. and R. Dennick, *Making sense of Cronbach's alpha.* International Journal of Medical Education, 2011. **2**: p. 53-55.
210. Bland, J.M. and D.G. Altman, *Cronbach's alpha.* BMJ, 1997. **314**(7080): p. 572.

211. DeVellis, R.F., *Scale development : theory and applications*. 2nd ed. Applied social research methods series. 2003, Thousand Oaks, Calif.: Sage Publications, Inc. viii, 171 p.
212. George, D. and P. Mallory, *SPSS for Windows step by step: A simple guide and reference 11.0 Update*. 4th ed. 2003, Boston: Allyn & Bacon.
213. Streiner, D.L., *Starting at the beginning: an introduction to coefficient alpha and internal consistency.* J Pers Assess, 2003. **80**(1): p. 99-103.
214. Richardson, A., et al., *Accuracy of first-trimester ultrasound in diagnosis of intrauterine pregnancy prior to visualization of the yolk sac: a systematic review and meta-analysis.* Ultrasound Obstet Gynecol, 2015. **46**(2): p. 142-9.
215. Fleischer, A.C., et al., *Ectopic pregnancy: features at transvaginal sonography.* Radiology, 1990. **174**(2): p. 375-8.
216. Bottomley, C. and T. Bourne, *Dating and growth in the first trimester.* Best Pract Res Clin Obstet Gynaecol, 2009. **23**(4): p. 439-52.
217. Stampone, C., et al., *Transvaginal sonography of the yolk sac in normal and abnormal pregnancy.* J Clin Ultrasound, 1996. **24**(1): p. 3-9.
218. Yeh, H.C., et al., *Intradecidual sign: a US criterion of early intrauterine pregnancy.* Radiology, 1986. **161**(2): p. 463-7.
219. Bradley, W.G., C.E. Fiske, and R.A. Filly, *The double sac sign of early intrauterine pregnancy: use in exclusion of ectopic pregnancy.* Radiology, 1982. **143**(1): p. 223-6.
220. Parvey, H.R., et al., *The chorionic rim and low-impedance intrauterine arterial flow in the diagnosis of early intrauterine pregnancy: evaluation of efficacy.* AJR Am J Roentgenol, 1996. **167**(6): p. 1479-85.
221. Stroup, D.F., et al., *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.* JAMA, 2000. **283**(15): p. 2008-12.
222. Mulrow, C.D., *The medical review article: state of the science.* Ann Intern Med, 1987. **106**(3): p. 485-8.
223. Sacks, H.S., et al., *Meta-analysis: an update.* Mt Sinai J Med, 1996. **63**(3-4): p. 216-24.
224. Moher, D., et al., *Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses.* Lancet, 1999. **354**(9193): p. 1896-900.
225. Whiting, P.F., et al., *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.* Ann Intern Med, 2011. **155**(8): p. 529-36.
226. Rutter, C.M. and C.A. Gatsonis, *Regression methods for meta-analysis of diagnostic test data.* Acad Radiol, 1995. **2 Suppl 1**: p. S48-56; discussion S65-7, S70-1 pas.
227. Harbord, R.M., et al., *A unification of models for meta-analysis of diagnostic accuracy studies.* Biostatistics, 2007. **8**(2): p. 239-51.
228. Ankum, W.M., et al., *Transvaginal sonography and human chorionic gonadotrophin measurements in suspected ectopic pregnancy: a detailed analysis of a diagnostic approach.* Hum Reprod, 1993. **8**(8): p. 1307-11.
229. Bateman, B.G., et al., *Vaginal sonography findings and hCG dynamics of early intrauterine and tubal pregnancies.* Obstet Gynecol, 1990. **75**(3): p. 421-427.
230. Chiang, G., et al., *The intradecidual sign: is it reliable for diagnosis of early intrauterine pregnancy?* AJR Am J Roentgenol, 2004. **183**(3): p. 725-31.
231. Enk, L., et al., *The value of endovaginal sonography and urinary human chorionic gonadotropin tests for differentiation between intrauterine and ectopic pregnancy.* J Clin Ultrasound, 1990. **18**(2): p. 73-8.
232. Kadar, N., G. DeVore, and R. Romero, *Discriminatory hCG zone: its use in the sonographic evaluation for ectopic pregnancy.* Obstet Gynecol, 1981. **58**(2): p. 156-61.

233. Romero, R., et al., *Diagnosis of ectopic pregnancy: value of the discriminatory human chorionic gonadotropin zone.* Obstet Gynecol, 1985. **66**(3): p. 357-60.
234. Tongsong, T. and S. Pongsatha, *Transvaginal sonographic features in diagnosis of ectopic pregnancy.* Int J Gynecol Obstet, 1993. **43**(3): p. 277-283.
235. Weckstein, L.N., et al., *Accurate diagnosis of early ectopic pregnancy.* Obstet Gynecol, 1985. **65**(3): p. 393-397.
236. Dart, R. and K. Howard, *Subclassification of indeterminate pelvic ultrasonograms: stratifying the risk of ectopic pregnancy.* Acad Emerg Med, 1998. **5**(4): p. 313-319.
237. Dart, R., et al., *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): p. 967-71.
238. Nyberg, D.A., et al., *Ectopic pregnancy. Diagnosis by sonography correlated with quantitative HCG levels.* J Ultrasound Med, 1987. **6**(3): p. 145-50.
239. Nyberg, D.A., et al., *Value of the yolk sac in evaluating early pregnancies.* J Ultrasound Med, 1988. **7**(3): p. 129-135.
240. Nyberg, D.A., et al., *Early pregnancy complications: endovaginal sonographic findings correlated with human chorionic gonadotropin levels.* Radiology, 1988. **167**(3): p. 619-622.
241. Nyberg, D.A., et al., *Ultrasonographic differentiation of the gestational sac of early intrauterine pregnancy from the pseudogestational sac of ectopic pregnancy.* Radiology, 1983. **146**(3): p. 755-9.
242. Bossuyt, P.M., et al., *The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration.* Ann Intern Med, 2003. **138**(1): p. W1-12.
243. Nyberg, D.A., et al., *Value of the yolk sac in evaluating early pregnancies.* J Ultrasound Med, 1988. **7**(3): p. 129-35.
244. Richardson, A., et al., *Accuracy of first-trimester ultrasound in diagnosis of tubal ectopic pregnancy in the absence of an obvious extrauterine embryo: systematic review and meta-analysis.* Ultrasound Obstet Gynecol, 2016. **47**(1): p. 28-37.
245. Nama, V. and I. Manyonda, *Tubal ectopic pregnancy: diagnosis and management.* Arch Gynecol Obstet, 2009. **279**(4): p. 443-53.
246. Doubilet, P.M. and C.B. Benson, *Double sac sign and intradecidual sign in early pregnancy: interobserver reliability and frequency of occurrence.* J Ultrasound Med, 2013. **32**(7): p. 1207-14.
247. Atri, M., J. de Stempel, and P.M. Bret, *Accuracy of transvaginal ultrasonography for detection of hematosalpinx in ectopic pregnancy.* J Clin Ultrasound, 1992. **20**(4): p. 255-61.
248. Brown, D.L. and P.M. Doubilet, *Transvaginal sonography for diagnosing ectopic pregnancy: positivity criteria and performance characteristics.* J Ultrasound Med, 1994. **13**(4): p. 259-66.
249. Lin, E.P., S. Bhatt, and V.S. Dogra, *Diagnostic clues to ectopic pregnancy.* Radiographics, 2008. **28**(6): p. 1661-71.
250. Nyberg, D.A., et al., *Extrauterine findings of ectopic pregnancy at transvaginal US: importance of echogenic fluid.* Radiology, 1991. **178**: p. 823-826.
251. Achiron, R., E. Schejter, and H. Zakut, *Combined pelvic sonography and serum beta hCG versus laparoscopy for the diagnosis of stable patient suspected of ectopic pregnancy* . Clinical and experimental obstetrics and gynecology, 1987 . **14**: p. 15-22 .
252. Ahmed, A., B. Tom, and P. Calabrese, *Ectopic pregnancy diagnosis and the pseudo-sac.* Fertility and Sterility, 2004. **81**(5): p. 1225-1228.

253. Aleem, F.A., M. DeFazio, and J. Gintautas, *Endovaginal sonography for the early diagnosis of intrauterine and ectopic pregnancies.* Human Reproduction, 1990. **5**(6): p. 755-758.
254. Blaivas, M. and M. Lyon, *Reliability of adnexal mass mobility in distinguishing possible ectopic pregnancy from corpus luteum cysts.* J Ultrasound Med, 2005. **24**: p. 599-603.
255. Cacciatore, B., U.H. Stenman, and P. Ylostalo, *Comparison of abdominal and vaginal sonography in suspected ectopic pregnancy.* Obstetrics and Gynecology, 1989. **73**(5): p. 770-774.
256. Cacciatore, B., U.H. Stenman, and P. Ylostalo, *Diagnosis of ectopic pregnancy by vaginal ultrasonography in combination with a discriminatory serum hCG level of 1000IU/l (IRP).* BJOG, 1990. **97**: p. 904-908.
257. Cacciatore, B., et al., *Suspected ectopic pregnancy: ultrasound findings and hCG levels assessed by an immunofluorometric assay.* BJOG, 1988. **95**: p. 497-502.
258. Chambers, S.E., B.B. Muir, and N.G. Haddad, *Ultrasound evaluation of ectopic pregnancy including correlation with human chorionic gonadotrophin levels.* The British Journal of Radiology, 1990. **63**: p. 246-250.
259. Dart, R., G. Burke, and L. Dart, *Subclassification of indeterminate pelvic ultrasonography: prospective evaluation of the risk of ectopic pregnancy.* Annals of Emergency Medicine, 2002. **39**(4): p. 382-388.
260. Dart, R., S. McLean, and L. Dart, *Isolated fluid in the cul-de-sac: how well does it predict ectopic pregnancy?* American Journal of Emergency Medicine, 2002. **20**(1): p. 1-4.
261. Dart, R., P. Ramanujam, and L. Dart, *Progesterone as a predictor of ectopic pregnancy when the ultrasound is indeterminate.* American Journal of Emergency Medicine, 2002. **20**(7): p. 575-579.
262. Dashefsky, S.M., et al., *Suspected ectopic pregnancy: endovaginal and transvesical US.* Radiology, 1988. **169**: p. 181-184.
263. Gabrielli, S., et al., *Accuracy of transvaginal ultrasound and serum hCG in the diagnosis of ectopic pregnancy.* Ultrasound Obstet Gynecol, 1992. **2**: p. 110-115.
264. Hammoud, A.O., et al., *The role of sonographic endometrial patterns and endometrial thickness in the differential diagnosis of ectopic pregnancy.* American Journal of Obstetrics and Gynecology, 2005. **192**: p. 1370-5.
265. Huter, O., et al., *Diagnosis of extrauterine pregnancy with transvaginal ultrasound.* Gynecol Obstet Invest, 1990. **30**: p. 204-206.
266. Kivikoski, A.I., C.M. Martin, and J.S. Smeltzer, *Transabdominal and transvaginal ultrasonography in the diagnosis of ectopic pregnancy: a comparative study.* American Journal of Obstetrics and Gynecology, 1990. **163**(1): p. 123-128.
267. Mahony, B.S., et al., *Sonographic evaluation of ectopic pregnancy.* J Ultrasound Med, 1985. **4**: p. 221-228.
268. Mehta, T.S., D. Levine, and C.R. McArdle, *Lack of sensitivity of endometrial thickness in predicting the presence of an ectopic pregnancy.* J Ultrasound Med, 1999. **18**: p. 117-122.
269. Rempen, A., *Vaginal sonography in ectopic pregnancy.* J Ultrasound Med, 1988. **7**: p. 381-387.
270. Romero, R., et al., *The value of adnexal sonographic findings in the diagnosis of ectopic pregnancy.* American Journal of Obstetrics and Gynecology, 1988. **158**(1): p. 52-55.
271. Russell, S.A., R.A. Filly, and N. Damato, *Sonographic diagnosis of ectopic pregnancy with endovaginal probes: what really has changed?* J Ultrasound Med, 1993. **3**: p. 145-151.

272. Sadek, A.L. and H.A. Schiotz, *Transvaginal sonography in the management of ectopic pregnancy.* Acta Obstet Gynecol Scand, 1995. **74**: p. 293-296.
273. Shapiro, B.S., et al., *Transvaginal ultrasonography for the diagnosi sof ectopic pregnancy.* Fertility and Sterility, 1988. **50**(3): p. 425-429.
274. Tongsong, T., et al., *Songraphic evaluation of clinical suspicion for ectopic pregnancy.* Asia-Oceania J. Obstet Gynaecol, 1992. **18**(2): p. 115-120.
275. Richardson, A., et al., *Use of the double decidual sac sign to confirm intrauterine pregnancy location prior to ultrasonographic visualisation of embryonic contents: a diagnostic accuracy study* Ultrasound Obstet Gynecol, 2016. **(in print)**.
276. Bossuyt, P.M., et al., *STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies.* BMJ, 2015. **351**: p. h5527.
277. Whiting, P., et al., *Sources of variation and bias in studies of diagnostic accuracy: a systematic review.* Ann Intern Med, 2004. **140**(3): p. 189-202.
278. Ransohoff, D.F. and A.R. Feinstein, *Problems of spectrum and bias in evaluating the efficacy of diagnostic tests.* N Engl J Med, 1978. **299**(17): p. 926-30.
279. Smidt, N., et al., *Quality of reporting of diagnostic accuracy studies.* Radiology, 2005. **235**(2): p. 347-53.
280. Alonzo, T.A. and M.S. Pepe, *Using a combination of reference tests to assess the accuracy of a new diagnostic test.* Stat Med, 1999. **18**(22): p. 2987-3003.
281. Hui, S.L. and X.H. Zhou, *Evaluation of diagnostic tests without gold standards.* Stat Methods Med Res, 1998. **7**(4): p. 354-70.
282. Rutjes, A.W., et al., *Evaluation of diagnostic tests when there is no gold standard. A review of methods.* Health Technol Assess, 2007. **11**(50): p. iii, ix-51.
283. Knottnerus, J.A., C. van Weel, and J.W. Muris, *Evaluation of diagnostic procedures.* BMJ, 2002. **324**(7335): p. 477-80.
284. Bredella, M.A., et al., *Measurement of endometrial thickness at US in multicenter drug trials: value of central quality assurance reading.* Radiology, 2000. **217**(2): p. 516-20.
285. Bradley WG, F.C., Filly RA., *The double sac sign of early intrauterine pregnancy: use in exclusion of ectopic pregnancy.* Radiology, 1982. **143**(1): p. 223-6.
286. Dart, R., S.A. McLean, and L. Dart, *Isolated fluid in the cul-de-sac: how well does it predict ectopic pregnancy?* Am J Emerg Med, 2002. **20**(1): p. 1-4.
287. Jayaprakasan, K., et al., *A randomised controlled trial of 300 versus 225 IU recombinant FSH for ovarian stimulation in predicted normal responders by antral follicle count.* BJOG, 2010. **117**(7): p. 853-62.
288. Hajian-Tilaki, K., *Sample size estimation in diagnostic test studies of biomedical informatics.* J Biomed Inform, 2014. **48**: p. 193-204.
289. Malak, M., et al., *Risk factors for ectopic pregnancy after in vitro fertilization treatment.* J Obstet Gynaecol Can, 2011. **33**(6): p. 617-9.
290. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
291. Fleiss, J.L., *Measuring nominal scale agreement among many raters.* Psychol Bull, 1968. **70**: p. 213-20.
292. Cohen, J., *Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit.* Psychol Bull, 1968. **70**(4): p. 213-20.
293. Gwet, K.L., *Benchmarking Inter-Rater Reliability Coefficients*, in *Handbook of Inter-Rater Reliability*. 2014, Advanced Analytics, LLC.
294. Posner, K.L., et al., *Measuring interrater reliability among multiple raters: an example of methods for nominal data.* Stat Med, 1990. **9**(9): p. 1103-15.
295. Peterson, I.S., *Using the kappa coefficient as a measure of reliability or reproducibility.* Chest, 1998. **114**: p. 946-947.

296. Nyberg, D.A. and R.A. Filly, *Predicting pregnancy failure in 'empty' gestational sacs.* Ultrasound Obstet Gynecol, 2003. **21**(1): p. 9-12.
297. Hately, W., J. Case, and S. Campbell, *Establishing the death of an embryo by ultrasound: report of a public inquiry with recommendations.* Ultrasound Obstet Gynecol, 1995. **5**(5): p. 353-7.
298. Elson, J., et al., *Prediction of early pregnancy viability in the absence of an ultrasonically detectable embryo.* Ultrasound Obstet Gynecol, 2003. **21**(1): p. 57-61.
299. Tongsong, T., et al., *Transvaginal ultrasound in threatened abortions with empty gestational sacs.* Int J Gynaecol Obstet, 1994. **46**(3): p. 297-301.
300. Bernard, K.G. and P.L. Cooperberg, *Sonographic differentiation between blighted ovum and early viable pregnancy.* AJR Am J Roentgenol, 1985. **144**(3): p. 597-602.
301. Babinszki, A., et al., *Three-dimensional measurement of gestational and yolk sac volumes as predictors of pregnancy outcome in the first trimester.* Am J Perinatol, 2001. **18**(4): p. 203-11.
302. Goldstein, R.B., *Endovaginal sonography in very early pregnancy: new observations.* Radiology, 1990. **176**(1): p. 7-8.
303. Pennell, R.G., et al., *Prospective comparison of vaginal and abdominal sonography in normal early pregnancy.* J Ultrasound Med, 1991. **10**(2): p. 63-7.
304. Schouwink, M.H., et al., *Ultrasonographic criteria for non-viability of first trimester intra-uterine pregnancy.* Early Pregnancy, 2000. **4**(3): p. 203-13.
305. Falco, P., et al., *Sonography of pregnancies with first-trimester bleeding and a viable embryo: a study of prognostic indicators by logistic regression analysis.* Ultrasound Obstet Gynecol, 1996. **7**(3): p. 165-9.
306. Bottomley, C., et al., *A model and scoring system to predict outcome of intrauterine pregnancies of uncertain viability.* Ultrasound Obstet Gynecol, 2011. **37**(5): p. 588-95.
307. Daponte, A., et al., *Angiopoietin-1 and angiopoietin-2 as serum biomarkers for ectopic pregnancy and missed abortion: a case-control study.* Clin Chim Acta, 2013. **415**: p. 145-51.
308. Daponte, A., et al., *Soluble FMS-like tyrosine kinase-1 (sFlt-1) and serum placental growth factor (PlGF) as biomarkers for ectopic pregnancy and missed abortion.* J Clin Endocrinol Metab, 2011. **96**(9): p. E1444-51.
309. Muttukrishna, S., et al., *Soluble Flt-1 and PlGF: new markers of early pregnancy loss?* PLoS One, 2011. **6**(3): p. e18041.
310. Horne, A.W., et al., *Placental growth factor: a promising diagnostic biomarker for tubal ectopic pregnancy.* J Clin Endocrinol Metab, 2011. **96**(1): p. E104-8.
311. Rull, K., et al., *Increased placental expression and maternal serum levels of apoptosis-inducing TRAIL in recurrent miscarriage.* Placenta, 2013. **34**(2): p. 141-8.
312. Daponte, A., et al., *Interleukin-15 (IL-15) and Anti-C1q Antibodies as Serum Biomarkers for Ectopic Pregnancy and Missed Abortion.* Clin Dev Immunol, 2013. **2013**: p. 637513.
313. Abdallah, Y., et al., *Limitations of current definitions of miscarriage using mean gestational sac diameter and crown-rump length measurements: a multicenter observational study.* Ultrasound Obstet Gynecol, 2011. **38**(5): p. 497-502.
314. Jeve, Y., et al., *Accuracy of first-trimester ultrasound in the diagnosis of early embryonic demise: a systematic review.* Ultrasound Obstet Gynecol, 2011. **38**(5): p. 489-96.
315. McShane, L.M., et al., *REporting recommendations for tumour MARKer prognostic studies (REMARK).* Br J Cancer, 2005. **93**(4): p. 387-91.

316. Moher, D., et al., *The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials.* JAMA, 2001. **285**(15): p. 1987-91.
317. Folkman, J. and M. Klagsbrun, *Angiogenic factors.* Science, 1987. **235**(4787): p. 442-7.
318. Davis, S., et al., *Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning.* Cell, 1996. **87**(7): p. 1161-9.
319. Maisonpierre, P.C., et al., *Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis.* Science, 1997. **277**(5322): p. 55-60.
320. Thomas, M. and H.G. Augustin, *The role of the Angiopoietins in vascular morphogenesis.* Angiogenesis, 2009. **12**(2): p. 125-37.
321. Dimitriadis, E., et al., *Cytokines, chemokines and growth factors in endometrium related to implantation.* Human Reproduction Update, 2005. **11**(6): p. 613-30.
322. Girardi, G., et al., *Complement activation in animal and human pregnancies as a model for immunological recognition.* Mol Immunol, 2011. **48**(14): p. 1621-30.
323. Toth, B., et al., *Placental interleukin-15 expression in recurrent miscarriage.* Am J Reprod Immunol, 2010. **64**(6): p. 402-10.
324. Kitaya, K., et al., *IL-15 expression at human endometrium and decidua.* Biol Reprod, 2000. **63**(3): p. 683-7.
325. Clark, D.E., et al., *Comparison of expression patterns for placenta growth factor, vascular endothelial growth factor (VEGF), VEGF-B and VEGF-C in the human placenta throughout gestation.* J Endocrinol, 1998. **159**(3): p. 459-67.
326. Nagamatsu, T., et al., *Cytotrophoblasts up-regulate soluble fms-like tyrosine kinase-1 expression under reduced oxygen: an implication for the placental vascular development and the pathophysiology of preeclampsia.* Endocrinology, 2004. **145**(11): p. 4838-45.
327. Jacobs, M., et al., *Levels of soluble fms-like tyrosine kinase one in first trimester and outcomes of pregnancy: a systematic review.* Reprod Biol Endocrinol, 2011. **9**: p. 77.
328. Irving, J.A., et al., *Characteristics of trophoblast cells migrating from first trimester chorionic villus explants and propagated in culture.* Placenta, 1995. **16**(5): p. 413-33.
329. Burton, G.J., E. Jauniaux, and A.L. Watson, *Maternal arterial connections to the placental intervillous space during the first trimester of human pregnancy: the Boyd collection revisited.* Am J Obstet Gynecol, 1999. **181**(3): p. 718-24.
330. Jauniaux, E., et al., *Comparison of ultrasonographic and Doppler mapping of the intervillous circulation in normal and abnormal early pregnancies.* Fertil Steril, 2003. **79**(1): p. 100-6.
331. Jauniaux, E., et al., *Doppler ultrasonographic features of the developing placental circulation: Correlation with anatomic findings.* Am J Obstet Gynecol, 1992. **166**(2): p. 585-7.
332. Rodesch, F., et al., *Oxygen measurements in endometrial and trophoblastic tissues during early pregnancy.* Obstet Gynecol, 1992. **80**(2): p. 283-5.
333. Palmer, S.K., et al., *Quantitative estimation of human uterine artery blood flow and pelvic blood flow redistribution in pregnancy.* Obstet Gynecol, 1992. **80**(6): p. 1000-6.
334. Jauniaux, E., L. Poston, and G.J. Burton, *Placental-related diseases of pregnancy: Involvement of oxidative stress and implications in human evolution.* Hum Reprod Update, 2006. **12**(6): p. 747-55.

335. Zhang, E.G., et al., *The regulation and localization of angiopoietin-1, -2, and their receptor Tie2 in normal and pathologic human placentae.* Mol Med, 2001. **7**(9): p. 624-35.
336. Mitchell, L.M., *Women's experiences of unexpected ultrasound findings.* J Midwifery Womens Health, 2004. **49**(3): p. 228-34.
337. Leithner, K., et al., *Affective state of women following a prenatal diagnosis: predictors of a negative psychological outcome.* Ultrasound Obstet Gynecol, 2004. **23**(3): p. 240-6.
338. Sparling, J.W., J.W. Seeds, and D.C. Farran, *The relationship of obstetric ultrasound to parent and infant behavior.* Obstet Gynecol, 1988. **72**(6): p. 902-7.
339. Baillie, C., G. Mason, and J. Hewison, *Scanning for pleasure.* Br J Obstet Gynaecol, 1997. **104**(11): p. 1223-4.
340. Baillie, C. and J. Hewison, *Antenatal screening. Obtaining selective consent to scanning, rather than screening, is possible.* BMJ, 1999. **318**(7186): p. 805.
341. NICE, *Ectopic Pregnancy and Miscarriage: Diagnosis and Initial Management*. 2012: London.

# Appendix 1

NHS
Health Research Authority
**NRES Committee East Midlands - Nottingham 1**
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Telephone: 0115 883 9390

16 April 2013

Mr Nick Raine-Fenning
NURTURE
B Floor, East Block, Queens Medical Centre
Nottingham
NG7 2UH

Dear Mr Raine-Fenning,

| | |
|---|---|
| **Study title:** | **Pregnancies of Uncertain Location or Viability Research (PULoVR)** |
| **REC reference:** | **13/EM/0081** |
| **Protocol number:** | **13008** |
| **IRAS project ID:** | **121905** |

Thank you for your letter of 11 April 2013, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the Co-ordinator Liza Selway, NRESCommittee.EastMidlands-Nottingham1@nhs.net.

**Confirmation of ethical opinion**

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

**Ethical review of research sites**

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS sites

**Conditions of the favourable opinion**

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

*Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.*

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

*Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.*

*For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.*

*Sponsors are not required to notify the Committee of approvals from host organisations*

**It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).**

**Approved documents**

The final list of documents reviewed and approved by the Committee is as follows:

| *Document* | *Version* | *Date* |
|---|---|---|
| Advertisement | 1.0 | 25 January 2013 |
| Covering Letter | | 30 January 2013 |
| Evidence of insurance or indemnity | | 06 August 2012 |
| Investigator CV | Nicholas Raine-Fenning | 04 December 2012 |
| Investigator CV | Bruce Campbell | 28 January 2013 |
| Investigator CV | Alison Richardson | 25 January 2013 |
| Letter from Sponsor | | 25 January 2013 |
| Other: Pregnancies of Uncertain Location or Viability Outcome Letter | 1.1 | 26 March 2013 |
| Other: Referee Report - Peer Review | | |
| Participant Consent Form: Part One (NURTURE (IVF/ICSI and Natural Conceptions)) | 1.0 | 25 January 2013 |
| Participant Consent Form: Part two (EPU) | 1.1 | 01 April 2013 |
| Participant Consent Form: Part one (Nurture(IVF/ICSI/Natural Conception | 1 | 25 January 2013 |
| Participant Information Sheet: NURTURE (Natural Conception) | 1.1 | 01 April 2013 |
| Participant Information Sheet: NURTURE (IVF/ICSI) | 1.1 | 01 April 2013 |
| Participant Information Sheet: Early Pregnancy Unit | 1.1 | 01 April 2013 |
| Protocol | 1.0 | 25 January 2013 |
| REC application | 121905/406634/1/291 | 28 January 2013 |
| Response to Request for Further Information | | 11 April 2013 |

**Statement of compliance**

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

**After ethical review**

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

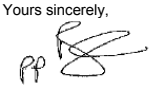
Further information is available at National Research Ethics Service website > After Review

| **13/EM/0081** | **Please quote this number on all correspondence** |
|---|---|

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

Yours sincerely,



**Mr Robert Johnson**
**Chair**

Email:NRESCommittee.EastMidlands-Nottingham1@nhs.net

*Enclosures:* "After ethical review – guidance for researchers" *[SL-AR2]*

*Copy to:* *Mr Paul Cartledge*
*Miss Shabina Sadiq, Nottingham University Hospitals NHS Trust*

NHS
Health Research Authority

**NRES Committee East Midlands - Nottingham 1**
Royal Standard Place
Nottingham
NG1 6FS

Tel: 0115 8839697

17 July 2014

Mr Nick Raine-Fenning
Associate Professor and Reader in Reproductive Medicine and Surgery
University of Nottingham
NURTURE
B Floor, East Block, Queens Medical Centre
Nottingham
NG7 2UH

Dear Mr Raine-Fenning,

| | |
|---|---|
| **Study title:** | **Pregnancies of Uncertain Location or Viability Research (PULoVR)** |
| **REC reference:** | **13/EM/0081** |
| **Protocol number:** | **13008** |
| **Amendment number:** | **1.1** |
| **Amendment date:** | **15 July 2014** |
| **IRAS project ID:** | **121905** |

Thank you for submitting the above amendment, which was received on 15 July 2014. I can confirm that this is a valid notice of a substantial amendment and will be reviewed by the Sub-Committee of the REC at its next meeting.

**Documents received**

The documents to be reviewed are as follows:

| *Document* | *Version* | *Date* |
|---|---|---|
| Notice of Substantial Amendment (non-CTIMP) | | 15 July 2014 |
| Participant consent form [(Nurture (IVF/ICSI)] | 1.1 | 10 July 2014 |
| Participant information sheet (PIS) [Nurture (IVF/ICSI)] | 1.2 | 10 July 2014 |
| Research protocol or project proposal | 1.1 | 10 July 2014 |

**Notification of the Committee's decision**

The Committee will issue an ethical opinion on the amendment within a maximum of 35 days from the date of receipt.

**R&D approval**

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval for the research.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

| 13/EM/0081: | Please quote this number on all correspondence |
|---|---|

Yours sincerely,

**Ms Rachel Nelson**
**REC Assistant**

E-mail: NRESCommittee.EastMidlands-Nottingham1@nhs.net

*Copy to:* *Miss Shabina Sadiq*

*Mr Paul Cartledge*

**NRES Committee East Midlands - Nottingham 1**
Royal Standard Place
Nottingham
NG1 6FS
Tel: 0115 8839428

06 February 2015

Mr Nick Raine-Fenning
Associate Professor and Reader in Reproductive Medicine and Surgery
University of Nottingham
NURTURE
B Floor, East Block, Queens Medical Centre
Nottingham
NG7 2UH

Dear Mr Raine-Fenning

| | |
|---|---|
| **Study title:** | **Pregnancies of Uncertain Location or Viability Research (PULoVR)** |
| **REC reference:** | **13/EM/0081** |
| **Protocol number:** | **13008** |
| **Amendment number:** | **Substantial Amendment 2** |
| **Amendment date:** | **14 January 2015** |
| **IRAS project ID:** | **121905** |

The above amendment was reviewed on 27 January 2015 by the Sub-Committee in correspondence.

**Ethical opinion**

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

**Approved documents**

The documents reviewed and approved at the meeting were:

| *Document* | *Version* | *Date* |
|---|---|---|
| Non-validated questionnaire [Pregnancies of Uncertain Location or Viability Research Anxiety Questionnaire] | 1.1 | 4 February 2015 |
| Notice of Substantial Amendment (non-CTIMP) | | 14 January 2015 |
| Participant consent form [Tracked] | 1.2 | 15 January 2015 |
| Participant information sheet (PIS) [Tracked] | 1.3 | 4 February 2015 |
| Research protocol or project proposal [Tracked] | 1.2 | 13 January 2015 |

**Membership of the Committee**

The members of the Committee who took part in the review are listed on the attached sheet.

**R&D approval**

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

**Statement of compliance**

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

| **13/EM/0081:** | **Please quote this number on all correspondence** |
|---|---|

Yours sincerely

**Dr Carl Edwards**
**Chair**

E-mail: NRESCommittee.EastMidlands-Nottingham1@nhs.net

*Enclosures:* *List of names and professions of members who took part in the review*

*Copy to:* *Miss Shabina Sadiq, Nottingham University Hospitals NHS Trust*
*Mr Paul Cartledge*

**NRES Committee East Midlands - Nottingham 1**

**Attendance at Sub-Committee of the REC meeting on 27 January 2015**

**Committee Members:**

| *Name* | *Profession* | *Present* | *Notes* |
|---|---|---|---|
| Dr Carl Edwards | Development Lead, Research Delivery and Support Unit | Yes | |
| Dr Ursula Holdsworth | Retired Staff Grade Community Paediatrician | Yes | |

**Also in attendance:**

| *Name* | *Position (or reason for attending)* |
|---|---|
| Miss Nicola Kohut | REC Assistant |

# Appendix 2

Nottingham University Hospitals NHS
NHS Trust

| | |
|---|---|
| Title of Guideline (must include the word "Guideline" (not protocol, policy, procedure etc) | **Guidelines on management of pregnancy of unknown location (PUL)** |
| Contact Name and Job Title (author) | Dr. Shilpa Deb<br>Consultant Obstetrician and Gynaecologist |
| Directorate & Speciality | Family Health<br>Obstetrics and Gynaecology |
| Date of submission | March 2016 |
| Date on which guideline must be reviewed (this should be one to three years) | March 2021 |
| Explicit definition of patient group to which it applies (e.g. inclusion and exclusion criteria, diagnosis) | Patients with an early pregnancy of ≤12 weeks |
| Abstract | This guideline is aimed at management of women with an early pregnancy when the location of pregnancy is not known. |
| Key Words | Ultrasound, early pregnancy, intra-uterine, ectopic, hCG |
| Statement of the evidence base of the guideline – has the guideline been peer reviewed by colleagues?<br><br>Evidence base: (1-5) | Literature review, evidence ranging from 1 to 5.<br><br>Peer-reviewed by the risk management group |
| Consultation Process | Risk Management Group<br>Consultant Gynaecologists<br>Ward Sisters<br>Gynaecology Nurse Specialists<br>Practice Development Matron |
| | |

| | |
|---|---|
| 1a | meta analysis of randomised controlled trials |
| 1b | at least one randomised controlled trial |
| 2a | at least one well-designed controlled study without randomisation |
| 2b | at least one other type of well-designed quasi-experimental study |
| 3 | well –designed non-experimental descriptive studies (ie comparative / correlation and case studies) |
| 4 | expert committee reports or opinions and / or clinical experiences of respected authorities |
| 5 | recommended best practise based on the clinical experience of the guideline developer |

| | |
|---|---|
| Target audience | All the medical, nursing and admin staff involved with emergency gynaecology |

**This guideline has been registered with the trust. However, clinical guidelines are guidelines only. The interpretation and application of clinical guidelines will remain the responsibility of the individual clinician. If in doubt contact a senior colleague or expert. Caution is advised when using guidelines after the review date**.

## Introduction

In the absence of either an intra- or extra-uterine pregnancy or retained products of conception in a woman with a positive pregnancy test, the pregnancy should be described as, “pregnancy of unknown location”. Even with expert use of TVS (Transvaginal Scanning) using agreed criteria, it may not be possible to confirm if a pregnancy is intrauterine or extra-uterine in 8–31% of cases at the first visit. These women should be classified as having a pregnancy of unknown location. In specialised scanning units, the overall incidence of pregnancy of unknown location is as low as 8–10%. In cases of known intrauterine pregnancy, viability will be uncertain in approximately 10% of women at their first EPAU (Early Pregnancy Assessment Unit) visit. The number of cases falling into these two groups can be kept to a minimum by using a thorough and critical approach to TVS in conjunction with strict diagnostic criteria. The sonographer should record whether an ‘apparently empty’ sac is eccentrically placed in the fundus, whether it exhibits a ‘double-ring’ pattern, and so on. These findings will help to delineate whether this is likely to be an intra- or extra-uterine pregnancy.

At levels above 1500 iu/l, an ectopic pregnancy will usually be visualised with TVS. However, the importance of levels that plateau below 1000 iu/l must be recognised. In these cases, an ectopic pregnancy and miscarriage are both possible outcomes. The potential for rarer

diagnoses, such as gestational trophoblastic disease or cranial germ cell tumour, must be considered although, in these cases, serum hCG (Human Chorionic Gonadotropin) levels are likely to be greater than 1000 IU/l. With a history and TVS findings suggestive of complete miscarriage, a 5.9% incidence of ectopic pregnancy has been reported and therefore the importance of performing serial hCG.

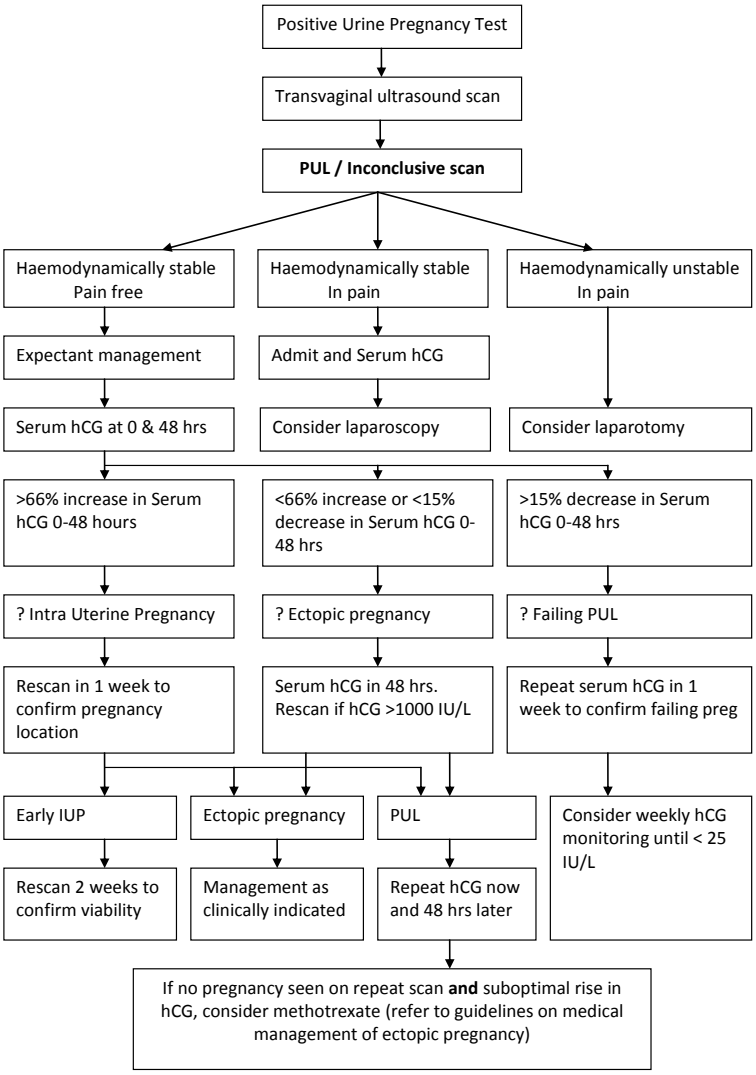
Anti D immunoglobulin should be administered when a surgical recourse (ERPC - Evacuation of Retained Products of Conception / laparoscopy) is taken to managing pregnancy of unknown location.

**Guidelines for management of Inconclusive Scan Result after the Initial Visit to EPU using Serum hCG and TVS**

| hCG IU/L | Ultrasound | Pattern of change Of hCG level after 48 hours | Management |
|---|---|---|---|
| **≤1000** | No intrauterine sac<br>No Adnexal mass<br>No fluid in POD<br>No symptoms | hCG rise > 66% or doubled | If hCG>1000 repeat ultrasound **or** If hCG < 1000 repeat hCG |
| **>1000** | **No intrauterine sac** | | |
| | 1. No adnexal mass<br>No fluid POD<br>No symptoms | | Repeat hCG and repeat ultrasound 2 days later |
| | | A. Falling hCG | Serial hCG levels until hCG <20 |
| | | B. Rising Or plateauing hCG x 3<br>Diagnosis: Ectopic or PUL | Laparoscopy (if symptomatic) Or Methotrexate (if asymptomatic)<br>(refer to guidelines on ectopic pregnancy) |

| | | | |
|---|---|---|---|
| | 2. Suspicious adnexal mass <3.5cm<br>No fluid POD<br>Asymptomatic | | Repeat hCG and repeat ultrasound 2 days later |
| | | A. Falling hCG | Serial hCG levels until hCG <20 |
| | | B. Rising/ plateauing hCG x 3 | Laparoscopy +/-D&C or Methotrexate |
| | 3. Ad.Mass ≥3.5cm<br>Or<br>Fluid POD<br>Or<br>Symptomatic | | Laparoscopy |
| **>2400** | No intrauterine sac<br>Adnexal findings +/-<br>Asymptomatic | Fluctuating x3<br>Diagnosis: Ectopic<br>Or PUL | Laparoscopy or Methotrexate |

# Guidelines on management of pregnancy of unknown location (PUL)



## References

Banerjee S, Aslam N, Woelfer B, Lawrence A. Elson J, Jurkovic D. Expectant management of pregnancies of unknown location:a prospective evaluation of methods to predict spontaneous resolution of pregnancy. *BJOG* 2001;108:158–63.

Condous G, Okaro E, Bourne T. The conservative management of early pregnancy complications: a review of the literature. *Ultrasound Obstet Gynecol* 2003;22:420–30.

Condous G, Okaro E, Khalid A, Bourne T. Do we need to follow up complete miscarriages with serum human chorionic gonadotrophin levels? *BJOG* 2005;112:827–9.

Hahlin M, Thorburn J, Bryman I. The expectant management of early pregnancy of uncertain site.*HumReprod* 1995;10:1223–7.

Jauniaux E, Johns J, Burton GJ. The role of ultrasound imaging in diagnosing and investigating early pregnancy failure. *Ultrasound Obstet Gynecol* 2005;25:613–24.

NICE Guideline CG154, *Ectopic pregnancy and miscarriage – Diagnosis and initial management in early pregnancy of ectopic pregnancy and miscarriage*. December 2012.

Royal College of Obstetricians and Gynaecologists. *Use of Anti-D Immunoglobulin for Rh Prophylaxis*. Guideline No. 22. London: RCOG; 2002.

Royal college of Obstetrician and Gynaecologist. *The management of early pregnancy loss.* Guideline No. 25, London: RCOG; 2004.

Royal college of Obstetrician and Gynaecologist. *The management of tubal pregnancy.* Guideline No. 21, London: RCOG; 2004, reviewed 2010.

# Appendix 3

## RESEARCH INTO STRESS AND ANXIETY IN EARLY PREGNANCY

We are trying to determine how women visiting the Early Pregnancy Unit feel at various stages of their visit. We appreciate that pain and bleeding in early pregnancy can cause a lot of stress and anxiety for women and we would like to try to quantify this further. By completing this short questionnaire you will be giving us invaluable information that we can use to improve the services we provide for you, and other women like you, in the future.

There are three short questionnaires to fill in: one before your ultrasound scan, one just after and one optional one 24-48 hours later. If you are willing to complete this optional one, please leave your contact details when requested on page 3.

We appreciate that some people might not feel able to complete the questionnaires after the ultrasound scan if they have been given bad news but please do try if possible because your feelings at this time are particularly important to us.

Please be reassured that your answers will be kept strictly confidential. The only reason we ask for your initials and date of birth is so that we can match your questionnaire answers with your ultrasound findings when we analyse the data.

Please follow the instructions on the questionnaire and feel free to ask a member of the Early Pregnancy Team if you have any questions.

Many thanks

Dr Alison Richardson

**Initials** _____

**Date of birth** _____ /_____ /_____

**Date today** _____ /_____ /_____

**PART I (to be completed in the waiting area before the ultrasound scan)**

A number of statements which people have used to describe themselves are given below. Read each statement and then write the number in the blank to the right of the statement that indicates ***how you generally feel***. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you *generally* feel. Please make sure you answer all the questions.

**1 = Almost never 2 = Sometimes 3 = Often 4 = Almost always**

1. I feel pleasant ____
2. I feel nervous and restless ____
3. I feel satisfied with myself ____
4. I wish I could be as happy as others seem to be ____
5. I feel like a failure ____
6. I feel rested ____
7. I am 'cool, calm and collected' ____
8. I feel that difficulties are piling up so that I cannot overcome them ____
9. I worry too much over something that really doesn't matter ____
10. I am happy ____
11. I have disturbing thoughts ____
12. I lack self-confidence ____
13. I feel secure ____
14. I make decisions easily ____
15. I feel inadequate ____
16. I am content ____
17. Some unimportant thought runs through my mind and bothers me ____
18. I take disappointments so keenly that I can't put them out of my mind ____
19. I am a steady person ____
20. I get in a state of tension or turmoil as I think over my recent concerns and interests ____

**PART II (to be completed in the waiting area before the ultrasound scan)**

Now read each statement and circle the number that best indicates ***how you feel right now***.

| | **Not at all** | **Somewhat** | **Moderately** | **Very much** |
|---|---|---|---|---|
| 21. I feel calm | 1 | 2 | 3 | 4 |
| 22. I am tense | 1 | 2 | 3 | 4 |
| 23. I feel upset | 1 | 2 | 3 | 4 |
| 24. I am relaxed | 1 | 2 | 3 | 4 |
| 25. I feel content | 1 | 2 | 3 | 4 |
| 26. I am worried | 1 | 2 | 3 | 4 |

***PLEASE KEEP HOLD OF THIS FOR THE TIME BEING AND COMPLETE THE NEXT PAGE AFTER YOU HAVE HAD YOUR ULTRASOUND SCAN***

**PART III (to be completed soon after the ultrasound scan)**

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the number that corresponds to the statement that best indicates ***how you feel right now***. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best. Please make sure you answer all the questions.

| | Not at all | Somewhat | Moderately | Very much |
|---|---|---|---|---|
| 27. I feel calm | 1 | 2 | 3 | 4 |
| 28. I am tense | 1 | 2 | 3 | 4 |
| 29. I feel upset | 1 | 2 | 3 | 4 |
| 30. I am relaxed | 1 | 2 | 3 | 4 |
| 31. I feel content | 1 | 2 | 3 | 4 |
| 32. I am worried | 1 | 2 | 3 | 4 |

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

Would you be willing for us to contact you in 24-48 hours to answer the same 6 questions one final time? This would greatly help our research. If you are happy to consent to this, please indicate your preferred contact method below and clearly write your contact details:

☐ Telephone _____ ☐ Morning ☐ Afternoon ☐ Evening ☐ Any time

☐ E-mail _____

Signed _____Date _____/_____/_____

If you choose telephone, please be aware that the call may come from a withheld number. If you choose e-mail, please look out for an email from alison.richardson@nottingham.ac.uk with the subject 'Early Pregnancy Research Questionnaire'.

***Many thanks for taking the time to complete this questionnaire.***
***It is very much appreciated.***

***PLEASE RETURN THIS COMPLETED QUESTIONNAIRE TO A MEMBER OF THE EARLY PREGNANCY TEAM BEFORE YOU LEAVE THE DEPARTMENT***

For EPAU use only: VIUP ☐ NVIUP ☐ EP ☐ PUL ☐ PUV ☐

**PART IV (to be completed by a member of the research team 24-48 hours after the ultrasound scan (optional))**

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the number that corresponds to the statement that best indicates ***how you feel right now***. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best. Please make sure you answer all the questions.

| | **Not at all** | **Somewhat** | **Moderately** | **Very much** |
|---|---|---|---|---|
| 33. I feel calm | 1 | 2 | 3 | 4 |
| 34. I am tense | 1 | 2 | 3 | 4 |
| 35. I feel upset | 1 | 2 | 3 | 4 |
| 36. I am relaxed | 1 | 2 | 3 | 4 |
| 37. I feel content | 1 | 2 | 3 | 4 |
| 38. I am worried | 1 | 2 | 3 | 4 |

For EPAU use only: VIUP ☐ NVIUP ☐ EP ☐ PUL ☐ PUV ☐

# Appendix 4

**CONSENT FORM**
**(Final version 1.1: 10th July 2014)**

**Title of Study: Pregnancies of Uncertain Location or Viability Research (PULoVR)**

**REC ref: 13/EM/0081**

**Name of Researcher**: Dr Alison Richardson

**Name of Participant:** **Please initial box**

1. I confirm that I have read and understand the information sheet version number ……dated..... for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐

3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4. I understand and agree that a transvaginal ultrasound scan will be performed for analysis of very early pregnancy development. ☐

5. I agree to be contacted regarding the outcome of the pregnancy if the information required cannot be obtained from my hospital records. ☐

6. I agree to take part in the above study. ☐

| Name of Participant | Date | Signature |
|---|---|---|
| Name of Person taking consent | Date | Signature |

# Appendix 5

| Section & Topic | No | Item |
|---|---|---|
| **TITLE OR ABSTRACT** | | |
| | **1** | Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC) |
| **ABSTRACT** | | |
| | **2** | Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts) |
| **INTRODUCTION** | | |
| | **3** | Scientific and clinical background, including the intended use and clinical role of the index test |
| | **4** | Study objectives and hypotheses |
| **METHODS** | | |
| *Study design* | **5** | Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study) |
| *Participants* | **6** | Eligibility criteria |
| | **7** | On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry) |
| | **8** | Where and when potentially eligible participants were identified (setting, location and dates) |
| | **9** | Whether participants formed a consecutive, random or convenience series |
| *Test methods* | **10a** | Index test, in sufficient detail to allow replication |
| | **10b** | Reference standard, in sufficient detail to allow replication |
| | **11** | Rationale for choosing the reference standard (if alternatives exist) |
| | **12a** | Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory |
| | **12b** | Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory |
| | **13a** | Whether clinical information and reference standard results were available to the performers/readers of the index test |
| | **13b** | Whether clinical information and index test results were available to the assessors of the reference standard |
| *Analysis* | **14** | Methods for estimating or comparing measures of diagnostic accuracy |
| | **15** | How indeterminate index test or reference standard results were handled |
| | **16** | How missing data on the index test and reference standard were handled |
| | **17** | Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory |
| | **18** | Intended sample size and how it was determined |
| **RESULTS** | | |
| *Participants* | **19** | Flow of participants, using a diagram |
| | **20** | Baseline demographic and clinical characteristics of participants |
| | **21a** | Distribution of severity of disease in those with the target condition |
| | **21b** | Distribution of alternative diagnoses in those without the target condition |
| | **22** | Time interval and any clinical interventions between index test and reference standard |
| *Test results* | **23** | Cross tabulation of the index test results (or their distribution) by the results of the reference standard |
| | **24** | Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals) |
| | **25** | Any adverse events from performing the index test or the reference standard |
| **DISCUSSION** | | |
| | **26** | Study limitations, including sources of potential bias, statistical uncertainty, and generalisability |
| | **27** | Implications for practice, including the intended use and clinical role of the index test |
| **OTHER INFORMATION** | | |
| | **28** | Registration number and name of registry |
| | **29** | Where the full study protocol can be accessed |
| | **30** | Sources of funding and other support; role of funders |

# STARD 2015

## AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

## EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

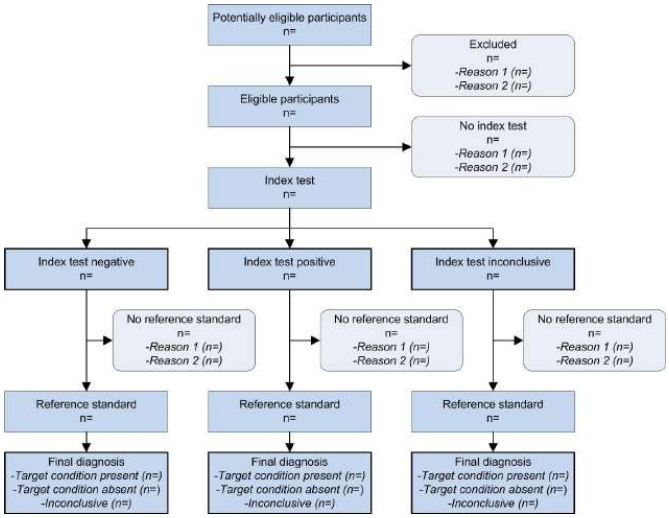
Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

## DEVELOPMENT

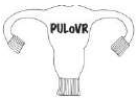
This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

**Prototypical STARD diagram to report flow of participants through the study**



# Appendix 6




# Study Number:_____

**Name** _____

**Unit Number** _____ **DOB** _____ / _____ / _____

**Email** _____

**Gravidity** _____ **Parity** _____

**Ethnicity** _____

**Height** _____ cm OR _____ ft _____ inches

**Weight** _____ kg OR _____ st _____ lbs

**AMH** _____ **AFC** _____

**Date of TVOR** _____ / _____ / _____ **No. of eggs collected** _____

**Date of ET** _____ / _____ / _____

**Stage of ET** d2 / d3 / d5

**No. of embryos** sET / dET

**Status of embryo(s)** Fresh / Frozen

**Grade of embryo(s)** _____

**Date of official UPT** _____ / _____ / _____ **Result** Positive / Negative

**Date of US** _____ / _____ / _____

**Calculated GA** _____ / 40

**Uterus** AV / RV **ET** _____ mm

**GS** present / absent **Number** _____

**GS dimensions** _____ x _____ x _____ mm **Vol** _____ $cm^3$

**DDSS** present / absent

**YS** present / absent **Size** _____ x _____ x _____ mm **Vol** _____ $cm^3$

**FP** present / absent **Length** _____ mm **FH** present / absent

**RO** _____ **LO** _____

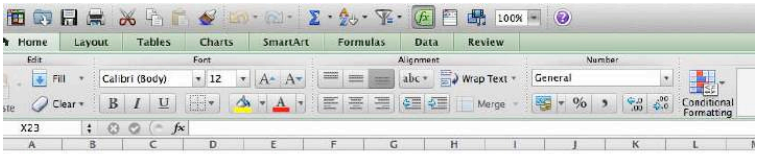
**AM** present / absent **Side** left / right **Size** _____ x _____ x _____ mm

**FF** present / absent **Volume** min / mod / large

**Initial Outcome** Viable IUP / Non-viable IUP / EP / PUL

**Final Outcome** Miscarriage / TOP / Livebirth / Stillbirth

# Appendix 7




**Off-line Analysis of ultraSonographic Images Study (OASIS): *The Double Decidual Sac Sign***

Thank you for agreeing to take part in this study looking at the inter-observer variability associated with the ultrasonographic identification of the double decidual sac sign (DDSS) in very early pregnancy.

The Dropbox link to the right will take you to a secure site where you will find 25 different files labelled 1 to 25. Each file contains several transvaginal ultrasonographic images of uteri from asymptomatic women who are between 4 and 5 weeks pregnant following IVF treatment. All the uteri contain an intrauterine fluid collection which may or may not represent a gestation sac. We would like you to tell us (1) whether or not you think the DDSS is present, (2) what you think the structure is and (3) what follow-up you would recommend.

Please note, not all the images are of good quality – we wanted to replicate daily practice. Some are particularly hard to interpret!

Please complete the 'Proforma' on the next sheet with your results. We suggest you save it to your computer and save your entries as you work. Each cell has a dropdown menu to make this as quick and as easy as possible. The form also asks you to provide some baseline data including your name, position and level of experience with early pregnancy ultrasound. Please feel free to add anything else you consider to be important.

Further information on the DDSS and the rational behind this research can be found on the 'Information' sheet of this spreadsheet.

Once finished, please email the Excel file to alison.richardson@nottingham.ac.uk, ideally before 15th September 2015.

Many thanks

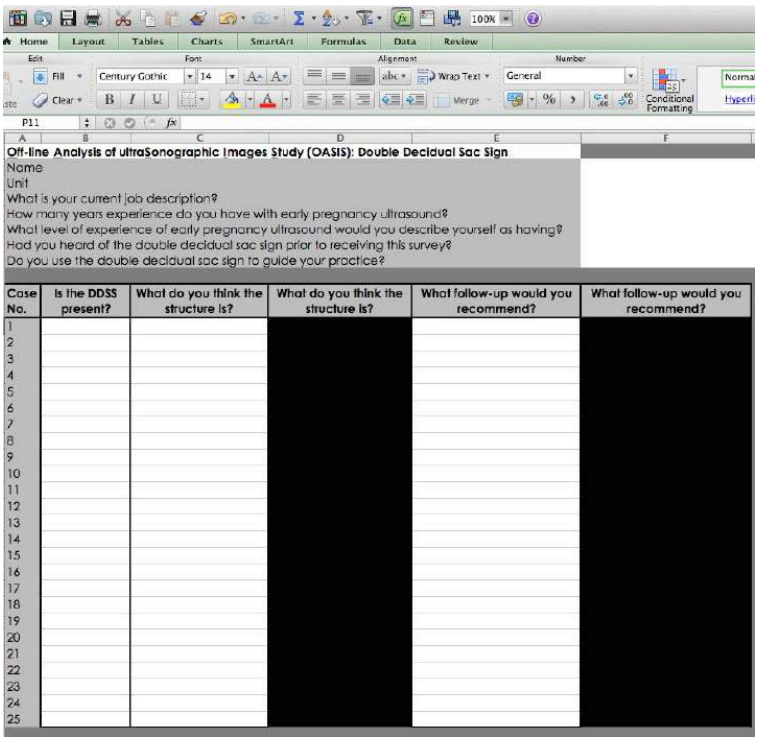
Alison Richardson & Nick Raine-Fenning

**(Please note some people have had difficulty getting to the Dropbox folder if using a browsers other than Internet Explorer - if this is the case just copy and paste the web address into Internet Explorer and it should work. Any problems, please let us know.)**

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Nurture Fertility

Dropbox



**Off-line Analysis of ultraSonographic Images Study (OASIS): Double Decidual Sac Sign**

Name
Unit
What is your current job description?
How many years experience do you have with early pregnancy ultrasound?
What level of experience of early pregnancy ultrasound would you describe yourself as having?
Had you heard of the double decidual sac sign prior to receiving this survey?
Do you use the double decidual sac sign to guide your practice?

| Case No. | Is the DDSS present? | What do you think the structure is? | What do you think the structure is? | What follow-up would you recommend? | What follow-up would you recommend? |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |
| 19 | | | | | |
| 20 | | | | | |
| 21 | | | | | |
| 22 | | | | | |
| 23 | | | | | |
| 24 | | | | | |
| 25 | | | | | |

# Appendix 8

**CONSENT FORM**
**(Final version 1.1: 10th July 2014)**

**Title of Study: Pregnancies of Uncertain Location or Viability Research (PULoVR)**

**REC ref: 13/EM/0081**

**Name of Researcher**: Dr Alison Richardson

**Name of Participant:** **Please initial box**

1. I confirm that I have read and understand the information sheet version number ……dated..... for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐

3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4. I understand and agree that serum samples will be taken for analysis of very early pregnancy development. ☐

5. I agree to be contacted regarding the outcome of the pregnancy if the information required cannot be obtained from my hospital records. ☐

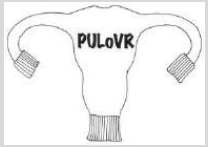
6. I agree to take part in the above study. ☐

| Name of Participant | Date | Signature |
|---|---|---|
| Name of Person taking consent | Date | Signature |

# Appendix 9

**Study Number:** PULoVR
Date: ____ /____ /_____
**Patient Details:**
Name:_____
NHS/Hospital No.:_____
DOB: _____/_____/_____ Age:_____yrs
G_____P_____ LMP _____/_____/_____
Weight:_____kg Height :_____cm
Smoking status: Never/Current/Ex
**Presenting complaint*:**
☐ Pain
☐ PVB
☐ Previous ectopic
☐ Previous miscarriage
☐ Other:_____
**Risk factors for ectopic*:**
☐ Previous ectopic
☐ Previous PID/pelvic infection
☐ Tubal surgery/sterilisation
☐ IVF pregnancy
☐ Coil in situ
**US Findings:**_____
_____
_____PUL ☐ PUV ☐

***Lab Use Only***




**Assays:**
☑ hCG _____
☑ Progesterone_____
☑ sTRAIL _____
☑ sFlt-1 _____
☑ ANG-1 _____
☑ ANG-2 _____
☑ IL-15 _____
☑ PIGF_____

***Research Team Use Only***

**Outcome:**
☐ Viable IUP
☐ Non-viable IUP*
☐ EP*
☐ Resolving PUL
☐ Persisting PUL

***Histology available:** Yes ☐ No ☐

*tick all that apply

# Appendix 10

**Table 1** REporting recommendations for tumour MARKer prognostic studies (REMARK)

**Introduction**

1. State the marker examined, the study objectives, and any prespecified hypotheses.

**Materials and Methods**

**Patients**

2. Describe the characteristics (e.g. disease stage or comorbidities) of the study patients, including their source and inclusion and exclusion criteria.
3. Describe treatments received and how chosen (e.g. randomised or rule-based).

**Specimen characteristics**

4. Describe type of biological material used (including control samples), and methods of preservation and storage.

**Assay methods**

5. Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study end point.

**Study design**

6. State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g. by stage of disease or age) was employed. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7. Precisely define all clinical end points examined.
8. List all candidate variables initially examined or considered for inclusion in models.
9. Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.

**Statistical analysis methods**

10. Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.
11. Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

**Results**

**Data**

12. Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.
13. Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumour marker, including numbers of missing values.

**Analysis and presentation**

14. Show the relation of the marker to standard prognostic variables.
15. Present univariate analyses showing the relation between the marker and outcome, with the estimated effect (e.g. hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analysed. For the effect of a tumour marker on a time-to-event outcome, a Kaplan–Meier plot is recommended.
16. For key multivariable analyses, report estimated effects (e.g. hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.
17. Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their significance.
18. If done, report results of further investigations, such as checking assumptions, sensitivity analyses, internal validation.

**Discussion**

19. Interpret the results in the context of the prespecified hypotheses and other relevant studies; include a discussion of limitations of the study.
20. Discuss implications for future research and clinical value.

# Appendix 11

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | |

## PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | |

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

# Appendix 12

# QUADAS-2: Background Document

## QUADAS-2

QUADAS-2 is designed to assess the quality of primary diagnostic accuracy studies; it is not designed to replace the data extraction process of the review and should be applied in addition to extracting primary data (e.g. study design, results etc) for use in the review. It consists of four key domains covering patient selection, index test, reference standard, and flow of patients through the study and timing of the index test(s) and reference standard ("flow and timing") (Table 1). The tool is completed in four phases: 1) state the review question; 2) develop review specific guidance; 3) review the published flow diagram for the primary study or construct a flow diagram if none is reported; 4) judgement of bias and applicability. Each domain is assessed in terms of the *risk of bias* and the first three are also assessed in terms of *concerns regarding applicability*. To help reach a judgement on the risk of bias, *signalling questions* are included. These flag aspects of study design related to the potential for bias and aim to help reviewers make risk of bias judgements.

**Phase 1: Review Question**

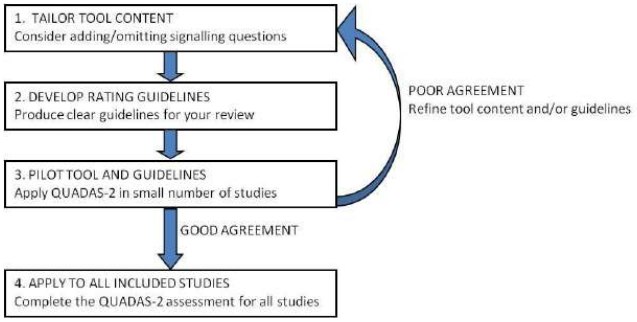
Review authors are first asked to report their systematic review question in terms of patients, index test(s), and reference standard and target condition. As the accuracy of a test may depend on where in the diagnostic pathway it will be used, review authors are asked to describe patients in terms of setting, intended use of the index test, patient presentation and prior testing.(1;2)

**Phase 2: Review Specific Tailoring (Figure 1)**

It is essential to tailor QUADAS-2 to each review by adding or omitting signalling questions and developing review-specific guidance on how to assess each signalling question and use this information to judge the risk of bias. The first step is to consider whether any signalling question does not apply to the review or whether any specific issues for the review are not adequately covered by the core signalling questions. For example, for a review of an objective index test it may be appropriate to omit the signalling question relating to blinding of the test interpreter to results of the reference standard. Review authors should avoid

complicating the tool by adding too many signalling questions. Once tool content has been agreed, review-specific rating guidance should be developed. The tool should be piloted independently by at least two people. If agreement is good, the tool can be used to rate all included studies. If agreement is poor, further refinement may be needed.

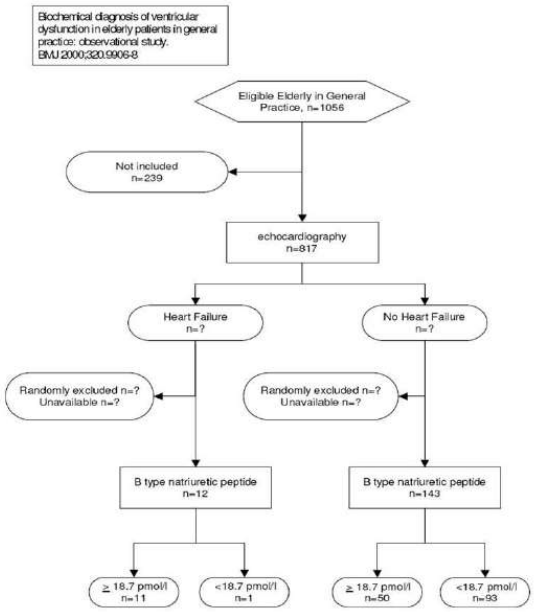
**Figure 1: Process for tailoring QUADAS-2 to your systematic review**




**Phase 3: Flow Diagram**

The next stage is to review the published flow diagram for the primary study or to draw one if none is reported or the published diagram is not adequate. The flow diagram will facilitate judgments of risk of bias, and should provide information about the method of recruitment of patients (e.g. based on a consecutive series of patients with specific symptoms suspected of having the target condition, or of cases and controls), the order of test execution, and the number of patients undergoing the index test and the reference standard. A hand drawn diagram is sufficient as this step does not need to be reported as part of the QUADAS-2 assessment. Figure 2 shows an example based on a primary study of B type natriuretic peptide for the diagnosis of heart failure.

**Figure 2: Flowchart based on diagnostic cohort study of BNP for diagnosing heart failure**




**Phase 4: Judgments on bias and applicability**

*Risk of bias*

The first part of each domain concerns bias and comprises three sections: 1) information used to support the risk of bias judgment, 2) signalling questions, and 3) judgment of risk of bias. By recording the information used to reach the judgment ("support for judgment"), we aim to make the rating transparent and facilitate discussion between review authors completing assessments independently.(3) The additional signalling questions are included

to assist judgments. They are answered as “yes”, “no”, or “unclear”, and are phrased such that “yes” indicates low risk of bias.

Risk of bias is judged as “low”, “high”, or “unclear”. If all signalling questions for a domain are answered “yes” then risk of bias can be judged “low”. If any signalling question is answered “no” this flags the potential for bias. Review authors then need to use the guidelines developed in phase 2 to judge risk of bias. The “unclear” category should be used only when insufficient data are reported to permit a judgment.

*Applicability*

Applicability sections are structured in a similar way to the bias sections, but do not include signalling questions. Review authors are asked to record the information on which the judgment of applicability is made and then to rate their concern that the study does not match the review question. Concerns regarding applicability are rated as “low”, “high” or “unclear”. Applicability judgments should refer to the first phase, where the review question was recorded. Again, the “unclear” category should only be used when insufficient data are reported.

The following sections provide brief explanations of the signalling questions and risk of bias/concerns regarding applicability questions for each domain.

**DOMAIN 1: PATIENT SELECTION**

***Risk of bias: Could the selection of patients have introduced bias?***

*Signalling question 1: Was a consecutive or random sample of patients enrolled?*

*Signalling question 2: Was a case-control design avoided?*

*Signalling question 3: Did the study avoid inappropriate exclusions?*

A study should ideally enrol all consecutive, or a random sample of, eligible patients with suspected disease – otherwise there is potential for bias. Studies that make inappropriate exclusions, e.g. excluding “difficult to diagnose” patients, may result in overoptimistic estimates of diagnostic accuracy. In a review of anti-CCP antibodies for the diagnosis of rheumatoid arthritis, we found that some studies enrolled consecutive patients who had confirmed diagnoses. These studies showed greater sensitivity of the anti-CCP test than

studies that included patients with suspected disease but in whom the diagnosis had not been confirmed – "difficult to diagnose" patients.(4) Similarly, studies enrolling patients with known disease and a control group without the condition may exaggerate diagnostic accuracy.(5;6) Exclusion of patients with "red flags" for the target condition, who may be easier to diagnose, may lead to underestimation of diagnostic accuracy.

***Applicability: Are there concerns that the included patients and setting do not match the review question?***

There may be concerns regarding applicability if patients included in the study differ, compared to those targeted by the review question, in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols. For example, larger tumours are more easily seen with imaging tests than smaller ones, and larger myocardial infarctions lead to higher levels of cardiac enzymes than small infarctions making them easier to detect and so increasing estimates of sensitivity.(7)

**DOMAIN 2: INDEX TEST**

***Risk of Bias: Could the conduct or interpretation of the index test have introduced bias?***

*Signalling question 1: Were the index test results interpreted without knowledge of the results of the reference standard?*

This item is similar to "blinding" in intervention studies. Interpretation of index test results may be influenced by knowledge of the reference standard.(6) The potential for bias is related to the subjectivity of index test interpretation and the order of testing. If the index test is always conducted and interpreted prior to the reference standard,this item can be rated "yes".

*Signalling question 2: If a threshold was used, was it pre-specified?*

Selecting the test threshold to optimise sensitivity and/or specificity may lead to overoptimistic estimates of test performance, which is likely to be poorer in an independent sample of patients in whom the same threshold is used.(8)

***Applicability: Are there concerns that the index test, its conduct, or interpretation differ from the review question?***

Variations in test technology, execution, or interpretation may affect estimates of its diagnostic accuracy. If index tests methods vary from those specified in the review question there may be concerns regarding applicability. For example, a higher ultrasound transducer frequency has been shown to improve sensitivity for the evaluation of patients with abdominal trauma.(9)

**DOMAIN 3: REFERENCE STANDARD**

***Risk of Bias: Could the reference standard, its conduct, or its interpretation have introduced bias?***

*Signalling question 1: Is the reference standard likely to correctly classify the target condition?*

Estimates of test accuracy are based on the assumption that the reference standard is 100% sensitive and specific disagreements between the reference standard and index test are assumed to result from incorrect classification by the index test.(10;11)

*Signalling question 2: Were the reference standard results interpreted without knowledge of the results of the index test?*

This item is similar to the signalling question related to interpretation of the index test. Potential for bias is related to the potential influence of prior knowledge on the interpretation of the reference standard.(6)

***Applicability: Are there concerns that the target condition as defined by the reference standard does not match the question?***

The reference standard may be free of bias but the target condition that it defines may differ from the target condition specified in the review question. For example, when defining urinary tract infection the reference standard is generally based on specimen culture but the threshold above which a result is considered positive may vary.(12)

**DOMAIN 4: FLOW AND TIMING**

***Risk of Bias: Could the patient flow have introduced bias?***

*Signalling question 1: Was there an appropriate interval between index test and reference standard?*

Ideally results of the index test and reference standard are collected on the same patients at the same time. If there is a delay or if treatment is started between index test and reference standard, misclassification may occur due to recovery or deterioration of the condition. The length of interval leading to a high risk of bias will vary between conditions. A delay of a few days may not be a problem for chronic conditions, while for acute infectious diseases a short delay may be important. Conversely, when the reference standard involves follow-up a minimum follow-up period may be required to assess the presence or absence of the target condition. For example, for the evaluation of magnetic resonance imaging for the early diagnosis of multiple sclerosis, a minimum follow-up period of around 10 years is required to be confident that all patients who will go on to fulfil diagnostic criteria for multiple sclerosis will have done so.(13)

*Signalling question 2: Did all patients receive the same reference standard?*

Verification bias occurs when not all of the study group receive confirmation of the diagnosis by the same reference standard. If the results of the index test influence the decision on whether to perform the reference standard or which reference standard is used, estimated diagnostic accuracy may be biased.(5;14) For example, a study evaluating the accuracy of the D-dimer test for the diagnosis of pulmonary embolism carried out ventilation perfusion scans (reference standard 1) in those testing positive and used clinical follow-up to determine whether or not those testing negative had a pulmonary embolism (reference standard 2). This may result in misclassifying some of the false negatives as true negatives as some patients who had a pulmonary embolism but were index test negative may be missed by clinical follow-up and so be classified as not having a pulmonary embolism. This misclassification will overestimate sensitivity and specificity.

*Signalling question 3: Were all patients included in the analysis?*

All patients who were recruited into the study should be included in the analysis.(15) There is a potential for bias if the number of patients enrolled differs from the number of patients included in the 2x2 table of results, for example because patients lost to follow-up differ systematically from those who remain.

## Incorporating QUADAS-2 assessments in diagnostic accuracy reviews

We emphasise that QUADAS-2 should not be used to generate a summary "quality score", because of the well-known problems associated with such scores.(16;17) If a study is judged as "low" on all domains relating to bias or applicability then it is appropriate to have an overall judgment of "low risk of bias" or "low concern regarding applicability" for that study. If a study is judged "high" or "unclear" on one or more domains then it may be judged "at risk of bias" or as having "concerns regarding applicability".

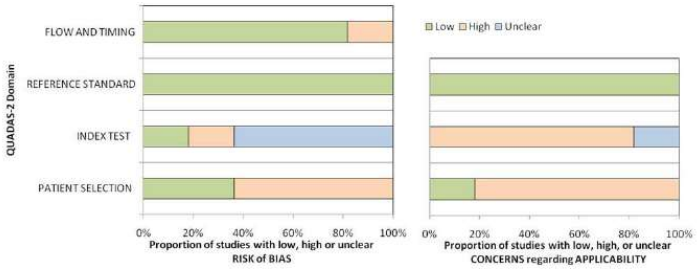
At minimum, reviews should present a summary of the results of the QUADAS-2 assessment for all included studies. This could include summarising the number of studies that found low, high or unclear risk of bias/concerns regarding applicability for each domain. If studies are found to consistently rate well or poorly on particular signalling questions then reviewers may choose to highlight these. Tabular (Table) and graphical (Figure 3) displays are helpful to summarise QUADAS-2 assessments.

**Table: Suggested tabular presentation for QUADAS-2 results**

| Study | RISK OF BIAS | | | | APPLICABILITY CONCERNS | | |
|---|---|---|---|---|---|---|---|
| | PATIENT SELECTION | INDEX TEST | REFERENCE STANDARD | FLOW AND TIMING | PATIENT SELECTION | INDEX TEST | REFERENCE STANDARD |
| Study 1 | ☺ | ☺ | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 2 | ☺ | ☺ | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 3 | ☹ | ☹ | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 4 | ☹ | ☹ | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 5 | ☹ | ? | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 6 | ☹ | ? | ☺ | ☺ | ☹ | ? | ☺ |
| Study 7 | ☹ | ? | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 8 | ☹ | ? | ☺ | ☺ | ☹ | ? | ☺ |
| Study 9 | ☹ | ? | ☺ | ☺ | ☹ | ☺ | ☺ |
| Study 10 | ☹ | ? | ☺ | ☹ | ☹ | ☺ | ☺ |
| Study 11 | ☺ | ? | ☺ | ☹ | ☺ | ☺ | ☺ |

☺ Low Risk ☹ High Risk ? Unclear Risk

**Figure 3: Suggested Graphical Display for QUADAS-2 results**




Review authors may choose to restrict the primary analysis so that only studies at low risk of bias and/or low concern regarding applicability for all or specified domains are included. It may be appropriate to restrict inclusion to the review based on similar criteria, but it is often preferable to review all relevant evidence and then investigate possible reasons for heterogeneity.(13;18) Subgroup and or sensitivity analysis can be conducted by investigating how estimates of accuracy of the index test vary between studies rated as high, low, or unclear on all or selected domains. Domains or signalling questions can be included as items in meta-regression analyses, to investigate their association with estimated accuracy.

## Website

The QUADAS website (www.quadas.org) contains QUADAS-2, information on training, a bank of additional signalling questions, more detailed guidance for each domain, examples of completed QUADAS-2 assessments, and downloadable resources including a Microsoft Access™ database for data extraction, an Excel™ spreadsheet to produce graphical displays of results, and templates for Word™ tables to summarise results.

## References


(1) Bossuyt PM, Leeflang MMG. Chapter 6: Developing Criteria for Including Studies. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration; 2009.

(2) Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008; 149(12):889-897.

(3) Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD et al. The Cochrane Collaboration's tool for assessing risk of bias in randomized trials. BMJ. In press 2011.

(4) Whiting PF, Smidt N, Sterne JA, Harbord R, Burton A, Burke M et al. Systematic review: accuracy of anti-citrullinated Peptide antibodies for diagnosing rheumatoid arthritis. Ann Intern Med 2010; 152(7):456-464.

(5) Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999; 282(11):1061-1066.

(6) Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. Ann Intern Med 2004; 140(3):189-202.

(7) Reitsma J, Rutjes A, WP, Vlassov V, Leeflang M, Deeks J. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration; 2009.

(8) Leeflang MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. Clinical Chemistry 2008; 54(4):729-737.

(9) Stengel D, Bauwens K, Rademacher G, Mutze S, Ekkernkamp A. Association between compliance with methodological standards of diagnostic research and reported test accuracy: meta-analysis of focused assessment of US for trauma. Radiology 2005; 236(1):102-111.

(10) Biesheuvel C, Irwig L, Bossuyt P. Observed differences in diagnostic test accuracy between patient subgroups: is it real or due to reference standard misclassification? Clin Chem 2007; 53(10):1725-1729.

(11) van Rijkom HM, Verdonschot EH. Factors involved in validity measurements of diagnostic tests for approximal caries--a meta-analysis. Caries Research 1995; 29(5)):364-70.

(12) Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J et al. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model. Health Technol Assess 2006; 10(36):iii-xiii, 1.

(13) Whiting P, Harbord R, Main C, Deeks JJ, Filippini G, Egger M et al. Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review. BMJ 2006; 332(7546):875-884.

(14) Rutjes A, Reitsma J, Di NM, Smidt N, Zwinderman A, Van RJ et al. Bias in estimates of diagnostic accuracy due to shortcomings in design and conduct: empirical evidence [abstract]. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26 31; Barcelona, Spain 2003;45.

(15) Macaskill P, Gatsonis C, Deeks JJ, Harbord R, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration; 20010.

(16) Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999; 282(11):1054-1060.

(17) Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. BMC Med Res Methodol 2005; 5:19.

(18) Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Medical Research Methodology 2006; 6:9.