ASSESSMENT OF BIOMARKERS IN PUL (PREGNANCY OF UNKNOWN LOCATION) AND ECTOPIC PREGNANCY

THE ABPEP STUDY

Shabnam Bobdiwala

MBBS (hons) BSc MRCOG

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

August 2019

Supervisors: Professor Tom Bourne and Professor Phillip Bennett

Department of Surgery and Cancer
Imperial College London

PhD Student

Shabnam Bobdiwala MBBS (hons) BSc MRCOG
Clinical Research Fellow
Division of Surgery and Cancer
Institute of Reproductive and Developmental Biology
Imperial College London
Hammersmith Hospital Campus
Du Cane Road, London, W12 0NN
Email: s.bobdiwala14@imperial.ac.uk Tel: +44 (0)7958707224

Supervisors

Professor Tom Bourne PhD FRCOG
Consultant Gynaecologist and Professor of Practice
Early Pregnancy Assessment Unit
Imperial College London, Hammersmith Campus
Queen Charlotte's and Chelsea Hospital
Du Cane Road, London W12 0HS
Email: t.bourne@imperial.ac.uk Tel: +44 (0)7768076797

Professor Phillip Bennett BSc PhD MD FRCOG
Professor in Obstetrics and Gynaecology
Director of the Institute of Reproductive and Developmental Biology
Division of Surgery and Cancer
Institute of Reproductive and Developmental Biology
Imperial College London
Hammersmith Hospital Campus
Du Cane Road, London, W12 0NN
Email: p.bennett@imperial.ac.uk Tel: +44 (0)207 594 2176

Declaration

Statement of originality

All work presented in this thesis was performed by myself unless otherwise stated.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Shabnam Bobdiwala

August 2019

Dedicated to the women I hope this work will help

Acknowledgments

My opening thank you unequivocally goes to Professor Tom Bourne for taking that first chance on me and providing the huge number of opportunities that have followed. I hope I spend the rest of my career making you proud.

Thank you to Professor Phil Bennett for always turning a 'no' into a 'yes' and providing his invaluable support when I have needed it the most. I also want to acknowledge Dr David MacIntyre for providing new opportunities researching the microbiome in early pregnancy. Your help and advice, even when inundated, were greatly appreciated. Thank you to doctors Yooni Lee and Pamela Pruski for generously providing their time and invaluable expertise in the lab. I am grateful to the team in the Women's Health Research Centre, especially Ramona Mannino and Maria Pipi for running and storing my numerous samples. Sophie Tapp, my research assistant extraordinaire – I look forward to watching you become the remarkable doctor I know you will be.

I have been fortunate enough to collaborate with the distinguished team at KU Leuven. Thank you to Professor Dirk Timmerman for welcoming me so warmly to his research group and to Professor Ben Van Calster for caring about mathematical models which sound like motorways as much as I do. Professor Jan Verbakels' statistical expertise in performing the meta-analysis was also greatly appreciated.

None of this work would have been possible without my famille de travail in the Early Pregnancy Unit at Queen Charlotte's and Chelsea Hospital: Miss Catriona Stalder, Miss Shyamaly Sur, Sister Maureen Mazibuko and a special mention to Sister Maeve Tuomey – your kindness and wicked humour have lifted me on the most difficult of days. You all spawned an army of research children and this experience would not have been half as enjoyable without the camaraderie of my fellow researchers: Maya Al-Memar, Jessica Farren, Nicola Mitchell-Jones, Srdjan Saso and now Chris Kyriacou. Charlotte's would not have been the same without you all. I am also grateful for the tremendous amount of effort and enthusiasm put in by the researchers at each of the participating sites for the clinical studies. Thank you to Francis Ayim, Baljinder Chohan, Osama Abughazza, Emma Kirk, Sharmistha Guha, Veluppillai Vathanan, Bramara Guruwadayarhalli, Cecilia Bottomley and Debbie Gould for so willingly giving up the little time you have amidst your hectic clinical schedules. I am also grateful to the team at NIHR CLAHRC NWL, especially Ganesh Sathyamoorthy and Wendy Carnegie, for their support and allowing me free rein.

I am indebted to the women who selflessly gave up their time to take part in this project – thank you for believing in the work that we do and allowing me to be a part of the most intimate and often difficult moments of your lives.

To Ahmedhusain and Mamuna – these words insufficiently express the gratitude I feel for the sacrifices you made so I could have the opportunities you were never given. Thank you, alongside Sophia and Farah, for always believing in me. My final thank you goes to Seb – merci d'être toujours là pour moi. J'ai hâte de pouvoir continuer à grandir à tes côtés.

Thesis abstract

A pregnancy of unknown location (PUL) describes a clinical situation where a woman presents with a positive urine pregnancy test but the location of the pregnancy cannot be determined when a transvaginal ultrasound scan is performed. The concern is that the patient may be harbouring an ectopic pregnancy (a pregnancy located outside of the uterus, most commonly in the fallopian tube), which remains the leading cause of early-pregnancy related maternal mortality in the UK. There remains a great deal of heterogeneity in the way women with a PUL are managed and this is not always evidence-based. There is therefore an unmet clinical need to improve the care of these patients. This manuscript outlines original research on how biomarkers, both novel and those in current clinical use, can be best used to optimise the clinical care women with a PUL receive.

Participants classified as a PUL at their initial ultrasound scan were recruited and underwent serial ultrasound scans and serum blood sampling in the first trimester. Women underwent routine follow-up and treatment as was clinically indicated. Those participants recruited to study 5 (see below) underwent additional bio-fluid sampling (urine and vaginal swabs, as well as serum blood testing). Pregnancy outcomes were collected using hospital paper records and IT systems.

7,139 women were recruited to the studies in total. Five separate analyses were conducted to meet the aims of this thesis:

Study 1: I performed a systematic review and meta-analysis to review the current evidence base for managing women with a pregnancy of unknown location. A total of 8,847 titles were screened and 43 studies analysed, with 23,802 women with a PUL recruited by these studies in total. This analysis demonstrated a logistic regression model called the M4 model was the current best diagnostic protocol for managing women with a PUL and predicting a final outcome of ectopic pregnancy.

Study 2: I undertook a clinical implementation study involving 1,022 patients initially classified as a PUL across three centres in the UK to assess the test performance and safety profile of using the M4 model in 'real-time' clinical practice. It demonstrated good test performance (classifying 70% PUL as low-risk after just two hospital visits, with a negative predictive value of 97% and a sensitivity for correctly predicting ectopic pregnancy at 82%).

Study 3: As there remained scope for improvement in test performance, a newer logistic regression model (the M6 model) was developed on 2,753 patients from two UK centres. We carried out a diagnostic accuracy study which demonstrated the M6 model had superior test performance to the M4 model, classifying 62% PUL as low-risk after just two hospital visits, with a negative predictive value of 99% and a sensitivity for correctly predicting ectopic pregnancy at 92%.

Study 4: I undertook a clinical implementation study involving 3,272 patients initially classified as a PUL across eight centres in the UK to assess the test performance and safety profile of using the M6 model (as part of a two-step management protocol) in 'real-time' clinical practice. It demonstrated good test performance (classifying 16% PUL as low-risk after just one hospital visit and a further 54% after just two visits with a negative predictive value of 98% and 99% respectively and a sensitivity for correctly predicting ectopic pregnancy at 86%). There were no clinical safety issues.

Study 5: I analysed the vaginal microbial composition in 92 women initially classified as a PUL. This demonstrated reduced *Lactobacillus* species abundance in pregnancies with a final outcome of an ectopic pregnancy compared to controls (i.e. those with a viable intrauterine pregnancy) that was independent of vaginal bleeding and preceded diagnosis of the final pregnancy outcome.

Summary conclusions

The studies carried out as part of this thesis demonstrate that a two-step triage protocol consisting of:

Step 1 – a progesterone cut-off of 2nmol/ l and

Step 2 – the M6 model

is the most effective way to manage women with a PUL and accurately predict a final outcome of an ectopic pregnancy. These findings can be used to rationalize the management of these women so they can be given more consistent and evidence-based care. Novel biomarker studies looking at the vaginal microbiota have shown promising preliminary results but further work is required in this area to establish the mechanistic pathways underlying the pathophysiology of ectopic pregnancy.

Table of Contents

ASSESSMENT OF BIOMARKERS IN PUL AND ECTOPIC PREGNANCY 1

DECLARATION 3

ACKNOWLEDGEMENTS 5

THESIS ABSTRACT 6

TABLE OF CONTENTS 9

PUBLICATIONS AND PRESENTATIONS 13

Publications 13

Authored book chapters 14

Research grants/ funding awarded 14

Prizes awarded 15

Oral presentations 16

Poster presentations 20

LIST OF TABLES 23

LIST OF FIGURES 24

ABBREVIATIONS 27

INTRODUCTION 30

The early pregnancy period 31

The establishment of specialised Early Pregnancy Units 31

Background to the studies in this thesis 32

Pregnancy of Unknown Location: terms & definitions 33

Ectopic Pregnancy 34

Pregnancy of Unknown Location: high-risk versus low-risk 38

PUL: biomarkers used in current clinical practice 39

Evaluating test performance in diagnostic accuracy studies 42

Novel serum biomarkers in the management of PUL 44

The microbiome 52

Summary of biomarkers 55

Justification for project 56

Project aims 57

Project hypotheses 57

METHODS **58**

Studies 59

Study design and ethical approval 59

Inclusion criteria 60

Exclusion criteria 60

Withdrawal 60

Data collection 61

Ultrasonography protocol (studies 2-5) 62

Sample collection and analysis 69

Study outcome measures and definitions 69

Statistical analysis 69

Contribution of various authors to each study 69

STUDY 1

DIAGNOSTIC PROTOCOLS FOR THE MANAGEMENT OF PREGNANCIES OF UNKNOWN LOCATION: A SYSTEMATIC REVIEW AND META-ANALYSIS **72**

Chapter abstract 73

Introduction 74

Methods 75

Results 78

Discussion 90

Final conclusions and future work 92

STUDY 2

A CLINICAL INTERVENTION STUDY USING THE M4 LOGISTIC REGRESSION MODEL TO MANAGE WOMEN WITH A PUL **93**

Chapter abstract 94

Introduction 95

Methods 96

Results 101

Discussion 111

Final conclusions and future work 114

STUDY 3

UTILIZING SERUM PROGESTERONE LEVELS AS WELL AS HCG – DEVELOPING A 2-STEP PROTOCOL USING THE M6 MODEL 115

Chapter abstract 116

Introduction 117

Methods 118

Results 121

Discussion 128

Final conclusions and future work 131

STUDY 4

A CLINICAL INTERVENTION STUDY USING A TWO-STEP TRIAGE PROTOCOL INCLUDING THE M6 LOGISTIC REGRESSION MODEL TO MANAGE WOMEN WITH A PUL 132

Chapter abstract 133

Introduction 135

Methods 136

Results 140

Discussion 150

Final conclusions and future work 153

STUDY 5

THE VAGINAL MICROBIOME IN PREGNANCIES OF UNKNOWN LOCATION 154

Chapter abstract 155

Introduction 157

Methods 159

Results 164

Discussion 171

Final conclusions and future work 174

DISCUSSION AND FUTURE PROSPECTS..... 175

Summary of the clinical problem 176

Clinical relevance and future work 178

Recommendations for further research from this thesis 179

Final Conclusions 181

REFERENCES 182

APPENDICES 200

Appendix 1: Early pregnancy clerking proforma 201

Appendix 2: Baseline characteristics (according to STARD criteria) of the 43 studies included in the systematic review and meta-analysis 205

Appendix 3: Inclusion and exclusion criteria of each of the 43 studies included in the systematic review 211

Appendix 4: The test performance of various management protocols in women with a pregnancy of unknown location (PUL) 213

Appendix 5: Development of the $M6_P$ and $M6_{NP}$ models 225

Appendix 6: Data on West Middlesex University Hospital (WMUH) 228

Appendix 7: Vaginal microbiome in PUL – Standard Operating Procedure (SOP) 230

Publications and presentations relevant to this thesis

Peer-reviewed publications

Bobdiwala S, Christodolou E, Farren J, Mitchell-Jones N, Kyriacou C, Al-Memar M, Ayim F, Chohan B, Kirk E, Abughazza O, Guruwadahyarhalli B, Guha S, Vathanan V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Triaging women with a pregnancy of unknown location using a two-step triage protocol including the M6 model: a clinical implementation study. Ultrasound Obstet Gynecol. 2019 (accepted for publication)

Bobdiwala S, Saso S, Verbakel JY, Al-Memar M, Van Calster B, Timmerman D, Bourne T. Diagnostic protocols for the management of pregnancy of unknown location: a systematic review and meta-analysis. BJOG. 2019;126(2):190-198

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

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. The clinical performance of the M4 decision support model to triage women with a pregnancy of unknown location as at low or high risk of complications. Hum Reprod. 2016; 31(7):1425-35

Van Calster B*, **Bobdiwala S** ***,** Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Condous G, Sur S, Stalder C, Timmerman D, Bourne T. Managing pregnancy of unknown location based on initial serum progesterone and serial serum hCG: development and validation of a two-step triage protocol. Ultrasound Obstet Gynecol. 2016;48(5):642-9
**joint first authors*

Al-Memar M, **Bobdiwala S**, Madhra M, Saso S, De Cock B, Van Calster B, Brown JK, Mukri F, Bottomley C, Papageorghiou A, Timmerman D, Horne AW, Bourne T. The potential value of activin B and fibronectin for the triage of pregnancies of unknown location and prediction of first trimester viability. Australasian Journal of Ultrasound in Medicine 2018, 21: 138-146.

Van Calster B, Van Hoorde K, Vergouwe Y, **Bobdiwala S**, Condous G, Kirk E, Bourne T, Steyerberg EW. Validation and updating of risk models based on multinomial logistic regression. Diagn Progn Res. 2017;1:2

Articles submitted to peer-reviewed journals

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. The potential use of urinary hCG measurements in the management of pregnancies of unknown location. Human Fertility. 2019 (submitted)

Authored book chapters

Early Pregnancy Ultrasound: A Practical Guide (2017)
Bobdiwala S, Bourne T. Chapter: Pregnancy of Unknown Location.

Ectopic Pregnancy: A Clinical Casebook (2015)
Bobdiwala S, Bourne T. Chapter: The Medical Management of Ectopic Pregnancy.

Research grants/ funding awarded

IMPERIAL HEALTH CHARITY grant

2019: £62,839 awarded to contribute towards the salary of the research fellow I co-supervise in continuing work related to biomarker studies in women with a PUL (pregnancy of unknown location)

GENESIS RESEARCH TRUST grant

2018: £12,760 awarded to conduct microbiome analysis of samples obtained from women with a PUL/ ectopic pregnancy

IMPERIAL CHARITY (SPECIAL PURPOSE FUND) grant

2016: £6,160 awarded to conduct metabonomic and microbiome analysis of samples obtained from women with a PUL/ ectopic pregnancy with the Clinical Phenome Centre (CPC) at Imperial College London

NIHR CLAHRC RDIP (Research, Development and Improvement Project) grant

2015: £99,717.67 awarded to conduct a multi-centre trial using a logistic regression model to manage PUL

NIHR CLAHRC eHealth grant

2015: £7,470 awarded to develop an app for smartphones to manage PUL

Prizes

WINNER of TOP ABSTRACT AWARD: International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Singapore)

2018: competitively shortlisted for this award (top abstract out of 950 submitted)

A multi-centre trial on the performance of a two-step triage protocol based on initial serum progesterone & serial hCG used to manage pregnancies of unknown location (PUL)

BEST ORAL COMMUNICATION: ISUOG World Congress (Singapore)

2018: Evaluating cut-off values for progesterone, single hCG and hCG ratio to define pregnancy viability and location in women with a pregnancy of unknown location (PUL)

BEST ORAL COMMUNICATION: Wessex Deanery Annual General Meeting

2018: Association of vaginal microbiota composition and outcomes in women found to have a pregnancy of unknown location (PUL) on an initial early pregnancy ultrasound scan

BEST ORAL COMMUNICATION: Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting

2017: The vaginal microbiome and its' association with final pregnancy outcome in women with a pregnancy of unknown location (PUL)

BEST ORAL COMMUNICATION: International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Vienna)

2017: Association of vaginal microbiota composition and outcomes in women found to have a pregnancy of unknown location (PUL) on an initial early pregnancy ultrasound scan

Shortlisted for YOUNG INVESTIGATOR OF THE YEAR AWARD: ISUOG World Congress

2017: competitively shortlisted for this award (top 9 out of 1251 submitted abstracts)

BEST ORAL COMMUNICATION: RCOG National Trainees Conference

2016: Managing pregnancy of unknown location (PUL) based on initial serum progesterone and serial serum hCG: a multi-centre trial on the performance and complications associated with using a two-step triage protocol

BEST ORAL COMMUNICATION: CLAHRC NWL Autumn Collaborative Learning Event

2016: for a 'Quality Improvement in Action' project - the RAMP (Risk Assessment Model in PUL) project

BEST ORAL POSTER COMMUNICATION: International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome)

2016: The potential use of urinary βhCG for managing pregnancies of unknown location (PUL): correlating urinary & serum βhCG levels using two immunoassays

BEST ORAL POSTER COMMUNICATION: Association for Early Pregnancy Units (AEPU) Annual Scientific Meeting

2015: Progesterone to predict ultimate viability in pregnancies of unknown location

BEST ORAL POSTER COMMUNICATION: ROYAL SOCIETY OF MEDICINE (RSM), Herbert Reiss and Dame Josephine Barnes prize meeting

2015: Managing pregnancy of unknown location: a two-step approach using initial serum progesterone and a novel risk prediction model based on serial serum hCG levels

Oral presentations

International oral presentations:

Bobdiwala S, Saso S, Verbakel J, Al-Memar M, Van Calster B, Timmerman D, Bourne T. Managing pregnancy of unknown location: a meta-analysis and evidence from a multi-centre trial on a two-step triage protocol based on initial serum progesterone & serial hCG. *RCOG World Congress (London, UK), June 2019*

Bobdiwala S. How to rationalize the management of pregnancies of unknown location (PUL): serum progesterone, the hCG ratio and novel prediction models. *Expert Pregnancy Imaging (London) – invited speaker, May 2019*

Bobdiwala S. Ultrasound in early pregnancy: pregnancy of unknown location (PUL) & ectopic pregnancy. *International IOTA (International Ovarian Tumor Analysis) Congress (Belgium) – invited speaker, April 2019*

Bobdiwala S. The basics of a gynaecology ultrasound (TA and TV walk-through). *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Singapore) – invited speaker, October 2018*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Ahmed S, Gould D, Sur S, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre trial on the performance of a two-step triage protocol based on initial serum progesterone & serial hCG used to manage pregnancies of unknown location (PUL). *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Singapore) – winner of the top abstract award, October 2018*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Evaluating cut-off values for progesterone, single hCG and hCG ratio to define pregnancy viability and location in women with a pregnancy of unknown location (PUL). *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Singapore), October 2018*

Bobdiwala S. Do we scan too many women too early? A practical validated approach to managing women with pregnancy of unknown location. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Vienna) – invited speaker, September 2017*

Bobdiwala S, Pruski P, Al-Memar M, Lee YS, Smith A, Marchesi JR, Bennett PR, Bourne T, MacIntyre DA. Association of vaginal microbiota composition and outcomes in women found to have a pregnancy of unknown location (PUL) on an initial early pregnancy ultrasound scan. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Vienna) – shortlisted for Young Investigator of the Year award, September 2017*

Bobdiwala S, Saso S, Verbakel J, Al-Memar M, Van Calster B, Timmerman D, Bourne T. Diagnostic protocols for the management of pregnancy of unknown location (PUL): a systematic review and meta-analysis. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Vienna), September 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Abughazza O, Al-Memar M, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Endometrial thickness and its value in triaging women with a pregnancy of unknown location. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Vienna), September 2017*

Bobdiwala S. Assessing the normal and abnormal pregnancy between 4 and 10 weeks. *RCOG World Congress (South Africa) – invited speaker, March 2017*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. The potential use of urinary βhCG for managing pregnancies of unknown location (PUL): correlating urinary & serum βhCG levels using two immunoassays. *RCOG World Congress (South Africa), March 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Abughazza O, Al-Memar M, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Assessing the value of endometrial thickness in triaging women with a pregnancy of unknown location. *RCOG World Congress (South Africa), March 2017*

Bobdiwala S. The role of serum progesterone alone and prediction models using hCG and progesterone to rationalise the follow up of women with a pregnancy of unknown location (PUL). *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome) – invited speaker, September 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. The role of progesterone for the prediction of viability in pregnancies of unknown location (PUL): a multicentre study*. International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome), September 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre study to predict outcome in pregnancies of unknown location (PUL) - the role of the bleeding score. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome), September 2016*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. The potential use of urinary βhCG for managing pregnancies of unknown location (PUL): correlating urinary & serum βhCG levels using two immunoassays. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome), September 2016*

Bobdiwala S. Pre-congress workshop: Interstitial ectopic pregnancy. *RCOG World Congress – invited speaker, June 2016*

Bobdiwala S. Pre-congress workshop: The practical management of Pregnancy of Unknown Location (PUL). *RCOG World Congress – invited speaker, June 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B,

Bourne T. Managing pregnancy of unknown location based on initial serum progesterone and serial serum hCG: a multi-centre trial on the performance and complications associated with using a two-step triage protocol. *RCOG World Congress, June 2016*

Bobdiwala S. Case examples – using the "early pregnancy APP" to predict outcome for PUV and PUL. *Early Pregnancy & Acute Gynaecology conference – invited speaker, November 2015*

Bobdiwala S. Case examples and how to manage PUL. *Early Pregnancy & Acute Gynaecology conference – invited speaker, November 2015*

Bobdiwala S. How to manage women with a pregnancy of unknown location: the role of progesterone, hCG and prediction models. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Montreal) – invited speaker, October 2015*

Bobdiwala S. What if you cannot see a pregnancy on ultrasound in early pregnancy? Managing pregnancy of unknown location (PUL) and features of tubal ectopic *Obstetric Ultrasound Foundation Course (London) – invited speaker, April 2015*

Bobdiwala S. The practical application of prediction models to manage pregnancies of unknown location. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Barcelona) – invited speaker, September 2015*

Bobdiwala S. How to classify pregnancies of unknown location as low or high risk. *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), Early Pregnancy and Infertility Education Conference – invited speaker, June 2014*

National oral presentations:

Bobdiwala S, Pruski P, Al-Memar M, Lee YS, Smith A, Marchesi JR, Bennett PR, Bourne T, MacIntyre DA. The vaginal microbiome and its' association with final pregnancy outcome in women with a pregnancy of unknown location (PUL). *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Abughazza O, Al-Memar M, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre study assessing the value of endometrial thickness in predicting outcome in pregnancies of unknown location (PUL). *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2017*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. Comparing urine & serum βHCG levels using two immunoassays: the potential utility of urine βHCG measurements in the management of pregnancies of unknown location (PUL). *Herbert Reiss and Dame Josephine Barnes prize meeting, Royal Society of Medicine, March 2017*

Bobdiwala S, Pruski P, Al-Memar M, Lee YS, Smith A, Marchesi JR, Bennett PR, Bourne T, MacIntyre DA. Ectopic pregnancy is associated with increased vaginal bacterial diversity and reduced *Lactobacillus* species dominance. *RCOG Annual Academic Meeting, March 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Managing pregnancy of unknown location (PUL) based on initial serum progesterone and serial serum hCG: a multi-centre trial on the performance and complications associated with using a two-step triage protocol**.** *RCOG National Trainees Conference, December 2016*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. Correlating urinary & serum βhCG levels using two immunoassays to risk-stratify pregnancies of unknown location (PUL). *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Progesterone to predict ultimate viability in pregnancies of unknown location. *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2015*

Bobdiwala S, Van Calster B, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Taheri M, Coundous G, Sur S, Stalder C, Timmerman D, Bourne T. Managing pregnancy of unknown location: a two-step approach using initial serum progesterone and a novel risk prediction model based on serial serum hCG levels. *London Specialty School of Obstetrics and Gynaecology (LSSOG) conference, March 2015*

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. A multi-centre trial on the performance and complications associated with using the M4 risk prediction model for triaging women with a pregnancy of unknown location (PUL). *Herbert Reiss and Dame Josephine Barnes prize meeting, Royal Society of Medicine, March 2015*

Bobdiwala S, Guha S, Ayim F, Ludlow J, Sayasneh A, Condous G, Kirk E, Stalder C, Timmerman D, Bourne T, Van Calster B. Managing Pregnancy of Unknown Location: testing the performance of progesterone and serial hCG protocols. *London Specialty School of Obstetrics and Gynaecology (LSSOG) conference, March 2014*

Poster presentations

International poster presentations:

Bobdiwala S, Pruski P, Al-Memar M, Lee YS, Smith A, Marchesi JR, Bennett PR, Bourne T, MacIntyre DA. Ectopic pregnancy is associated with increased vaginal bacterial diversity and reduced *Lactobacillus* species dominance. *Society for Reproductive Investigation 64th Annual Scientific Meeting, March 2017*

Bobdiwala S, Pruski P, Al-Memar M, Lee YS, Smith A, Marchesi JR, Bennett PR, Bourne T, MacIntyre DA. Ectopic pregnancy is associated with increased vaginal bacterial diversity and reduced *Lactobacillus* species dominance. *RCOG World Congress, March 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre trial on the performance and complications of a two-step triage protocol based on initial serum progesterone and serial serum hCG to manage pregnancy of unknown location (PUL). *International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), World Congress (Rome), September 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Taheri M, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre study to predict ultimate viability in pregnancies of unknown location (PUL) – the role of progesterone. *RCOG World Congress, June 2016*

National poster presentations:

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Abughazza O, Al-Memar M, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Assessing the value of endometrial thickness measurement in triaging women with a pregnancy of unknown location (PUL). *RCOG Annual Academic Meeting, March 2017*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. The potential use of urinary β human chorionic gonadotrophin (βhCG) for managing pregnancies of unknown location (PUL): correlating urinary & serum βhCG levels using two immunoassays. *RCOG Annual Academic Meeting, March 2017*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre trial on the performance and complications of a two-step triage protocol based on initial serum progesterone and serial serum hCG to manage pregnancy of unknown location (PUL). *RCOG Annual Academic Meeting, March 2017*

Bobdiwala S, Harvey R, Abdallah Y, Al-Memar M, Fisher RA, Stalder C, Timmerman D, Bourne T. The potential use of urinary βhCG for managing pregnancies of unknown location (PUL): correlating urinary & serum βhCG levels using two immunoassays. *RCOG National Trainees Conference, December 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. A multi-centre trial on managing pregnancy of unknown location (PUL) based on initial serum progesterone and serial serum hCG. *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Predicting ultimate viability in pregnancies of unknown location – the role of progesterone. *RCOG Annual Academic Meeting, March 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Predicting ultimate viability in pregnancies of unknown location – the role of the bleeding score. *RCOG Annual Academic Meeting, March 2016*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Progesterone to predict ultimate viability in pregnancies of unknown location. *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2015*

Bobdiwala S, Farren J, Mitchell-Jones N, Ayim F, Chohan B, Guruwadahyarhalli B, Kirk E, Abughazza O, Al-Memar M, Guha S, Veluppillai V, Bottomley C, Gould D, Stalder C, Timmerman D, Van Calster B, Bourne T. Bleeding score to predict ultimate viability in pregnancies of unknown location. *Association of Early Pregnancy Units (AEPU) Annual Scientific Meeting, November 2015*

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. The clinical performance of the M4 decision support model to triage women with a pregnancy of unknown location as at low or high risk of complications. *RCOG National Trainees Conference, December 2015*

Bobdiwala S, Van Calster B, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Coundous G, Sur S, Stalder C, Timmerman D, Bourne T. Using a two-stage strategy: serum progesterone at presentation and a novel risk model based on progesterone and serial serum hCG to manage pregnancy of unknown location (PUL). *RCOG National Trainees Conference, December 2015*

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. A multi-centre trial on the performance and complications associated with using the M4 risk prediction model for triaging women with a pregnancy of unknown location (PUL). *London Specialty School of Obstetrics and Gynaecology (LSSOG) conference, March 2015*

Bobdiwala S, Van Calster B, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Coundous G, Sur S, Stalder C, Timmerman D, Bourne T. Managing pregnancy of unknown location: a two-step approach using initial serum progesterone and a novel risk prediction model based on serial serum hCG levels. *Herbert Reiss and Dame Josephine Barnes prize meeting, Royal Society of Medicine, March 2015*

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. A multi-centre trial on the performance and complications associated with using the M4 risk prediction model for triaging women with a pregnancy of unknown location (PUL). *RCOG Annual Academic Meeting, January 2015*

Bobdiwala S, Van Calster B, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Coundous G, Sur S, Stalder C, Timmerman D, Bourne T. Managing pregnancy of unknown location: a two-step approach using initial serum progesterone and a novel risk prediction model based on serial serum hCG levels. *RCOG Annual Academic Meeting, January 2015*

Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. The safety of using the M4 risk prediction model to triage women with a pregnancy of unknown location. *Selected abstract at the Victor Bonney Society Winter Academic Meeting, September 2014*

Bobdiwala S, Van Calster B, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Coundous G, Sur S, Stalder C, Timmerman D, Bourne T. Risk prediction models and pregnancy of unknown location: a novel two-step approach. *Selected abstract at the Victor Bonney Society Winter Academic Meeting, September 2014*

Bobdiwala S, Guha S, Ayim F, Ludlow J, Sayasneh A, Condous G, Kirk E, Stalder C, Timmerman D, Bourne T, Van Calster B. Triaging pregnancies of unknown location: the performance of protocols based on single serum progesterone or repeated serum HCG levels. *Annual Academic Meeting, RCOG, December 2013*

List of tables

Table 1: Pooled sensitivity and specificity data for the management protocols assessed in the systematic review and meta-analysis

Table 2: Data on demographics, symptoms and serum biochemistry results for patients in the study with pregnancy of unknown location (PUL)

Table 3: The 17 high-risk patients misclassified as 'low risk' by the model and the interventions they had

Table 4: The performance of [i] the M4 model, [ii] triage based on hCG ratio cut-offs (low risk if hCG ratio <0.87 or >1.66), and [iii] triage based on a single progesterone cut-off (low risk if progesterone ≤10nmol/l)

Table 5: Descriptive statistics presented as median (interquartile range) in the two participating hospitals

Table 6: Development data performance of different initial progesterone (measured in nmol/l) cut-offs to define low-risk at presentation

Table 7: Performance of PUL classification approaches. Confidence intervals are given between parentheses

Table 8: Patient recruitment by each participating centre

Table 9: Pooled descriptive statistics from all participating early pregnancy units, n=2625

Table 10: Overview of patient triage, overall and by final PUL outcome

Table 11: Patient-specific information of women with an ectopic pregnancy that were misclassified by the two-step algorithm (n=15)

Table 12: Protocol deviations and other events associated with two-step algorithm use (n=2625). Step 1 = progesterone cut-off of ≤2nmol/l. Step 2 = M6 risk prediction model

Table 13: Breakdown of adverse events (AE) in the study cohort

Table 14: Metadata for PUL final pregnancy outcome groups

List of figures

Figure 1: Potential sites for an ectopic pregnancy

Figure 2: A tubal ectopic pregnancy seen as an inhomogenous mass ('blob sign')

Figure 3: A tubal ectopic pregnancy seen as an inhomogenous mass with a gestation sac ('bagel sign')

Figure 4: A tubal ectopic pregnancy seen as an inhomogenous mass with a gestation sac, yolk sac and fetal pole that may contain positive fetal cardiac activity

Figure 5: A large haemoperitoneum (blood in the pelvis) seen on transvaginal ultrasonography

Figure 6: Schematic for triaging women with a pregnancy of unknown location (PUL) as either low-risk or high-risk of clinical complications

Figure 7: Comparison of ROC curves for the prediction of ectopic pregnancy as a final pregnancy outcome in women with a PUL by the M1, M2 and M3 models

Figure 8: representation of a ROC (Receiver Operating Curve) and AUC (Area under the Receiver Operating Curve)

Figure 9: Five key areas historically explored in novel serum biomarker discovery for ectopic pregnancies.

Figure 10: Phases of biomarker research

Figure 11: Thesis outline and rationale

Figure 12: The pictorial blood loss assessment chart (PBAC) used to quantify the vaginal bleeding experienced

Figure 13: The Visual Analogue Scale (VAS) used to quantify the patients' pain score

Figure 14: Examples of endometrial thickness measurement

Figure 15: Example of a fluid filled structure seen within the uterus

Figure 16: Examples of measurement of the ovaries

Figure 17: Examples of free fluid seen in the pelvis (specifically, the Pouch of Douglas)

Figure 18: example of an intra-uterine gestation sac (IUGS) and measurement of the mean sac diameter taken in three orthogonal planes

Figure 19: Example a yolk sac and measurements taken in three orthogonal planes

Figure 20: Example of a fetal pole and CRL (crown rump length) measurement

Figure 21: Example an adnexal mass, seen on the right side of the image and medial to the right ovary

Figure 22: PRISMA flowchart of the literature search and article selection for a systematic review and meta-analysis on the test performance of protocols used in the management of women with a pregnancy of unknown location (PUL)

Figure 23: QUADAS-2 data for the included studies

Figure 24: SROC (Summary Receiver Operating Curve) for the test performance of all management protocols predicting a final outcome of ectopic pregnancy (EP). The number on each curve refers to the summary threshold for each protocol used in the original papers

Figure 25: SROC (Summary Receiver Operating Curve) for the test performance of all management protocol predicting a final outcome of failed pregnancy of unknown location (FPUL). The number on each curve refers to the summary threshold for each protocol used in the original papers

Figure 26: SROC (Summary Receiver Operating Curve) for the test performance of all management protocol predicting a final outcome of intrauterine pregnancy (IUP). The number on each curve refers to the summary threshold for each protocol used in the original papers

Figure 27: The three possible predictions and management plans as recommended by the M4 model after inputting the initial (0h) and 48h serum hCG results

Figure 28: Flowchart of final outcome data and correct versus incorrect risk stratification according to the M4 model

Figure 29: Flowchart of patient recruitment and protocol deviations

Figure 30: Breakdown of adverse and serious adverse events in the study cohort

Figure 31: Flow diagram of two-step approach for managing pregnancies of unknown location (PUL)

Figure 32: Flow diagram of patient flow through the study

Figure 33: Scatterplot of initial serum progesterone level against initial serum human chorionic gonadotropin (hCG) level for (A) development dataset and (B) validation dataset for managing pregnancies of unknown location (PUL)

Figure 34: ROC curves for discriminating EP from FPUL/IUP for the M6P, M6NP and M4 models

Figure 35: Flow diagram of two-step approach for managing pregnancies of unknown location (PUL)

Figure 36: Flowchart of patient recruitment, final outcome data and correct versus incorrect risk stratification by the two-step protocol

Figure 37: Demonstration of the user-friendly app for smartphones developed to enable use of the M6 model by clinicians

Figure 38: Analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) outcomes (n=70). (A) heatmap of genera data; (B) association between the vaginal bleeding score and percentage Lactobacillus dominance; (C) percentage of Lactobacillus dominance in viable and adverse pregnancy outcome groups; (D) Shannon index analysis of vaginal bacterial diversity between the two outcome groups

Figure 39: Analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) outcomes (n=70). (A) heatmap of species data; (B) proportion of community state types in viable and adverse pregnancy outcome groups; (C) extended error bar for microbial biomarkers identified to be significantly different between the two outcome groups. Note, none of these withstood correction

Figure 40: LefSe (latent discriminatory analysis effect size) analysis identifying microbial biomarkers signficaintly different between PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) early pregnancy outcomes (n=70)

Figure 41: Analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus an ectopic pregnancy (n=19). (A) proportion of community state types in viable and ectopic pregnancy outcome groups; (B) Shannon index analysis of vaginal bacterial diversity between the two outcome groups

Figure 42: LefSe (latent discriminatory analysis effect size) analysis identifying microbial biomarkers signficaintly different between PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus an ectopic pregnancy (n=19)

Abbreviations

°C	Degrees Celsius
2D	Two-dimensional
48h	48 hours
ABPEP	Assessment of Biomarkers in PUL and Ectopic Pregnancy
A&E	Accident and emergency
AE	Adverse event
AFP	α-fetoprotein
AUC	Area under the receiver-operating-characteristic curves
BMI	Body mass index
BV	Bacterial vaginosis
CA125	Cancer antigen 25
CI	Confidence interval
CK	Creatine kinase
CPA	Clinical Pathology Accreditation
CRL	Crown rump length
CST	Community state type
CWH	Chelsea and Westminster Hospital
D&C	Dilatation and curettage
D2	Day 2
DNA	Deoxyribonucleic acid
DNTP	Deoxyribonucleotide triphosphate
E2	Oestradiol
ED	Emergency department
EDTA	Ethylenediaminetetraacetic acid
EP	Ectopic pregnancy
EPAU	Early pregnancy assessment unit
EPU	Early pregnancy unit
ET	Endometrial thickness
EtOH	Ethanol
FFN	Fetal fibronectin
FN	False negative
FP	False positive
FPUL	Failed pregnancy of unknown location
FSH	Follicle-stimulating hormone
GCP	Good clinical practice
GP	General practitioner
HCA	Hierarchical clustering analysis
hCG	Human chorionic gonadotrophin
HCl	Hydrochloric acid
HH	Hillingdon Hospital
HPL	Human placental lactogen
ICSI	Intracytoplasmic sperm injection
IL-2	Interleukin 2
IL-6	Interleukin 6
IL-8	Interleukin 8
IQR	Interquartile range

IUGS	Intra-uterine gestation sac
IUI	Intrauterine insemination
IUP	Intrauterine pregnancy
IVF	In vitro fertilisation
kg	Kilograms
LDA	Linear discriminative analysis
LefSe	Latent discriminatory analysis effect size
LFU	Lost to follow-up
LH	Luteinizing hormone
LIF	Leukaemia inhibitory factor
LMP	Last menstrual period
log	Logarithm
LR	Logistic regression
MBRRACE	Mothers and babies: reducing risk through audits and confidential enquiries
MI	Multiple imputation
MTX	Methotrexate
MUC16	Cancer antigen 25
MVA	Manual vacuum aspiration
MYH11	Smooth muscle heavy-chain myosin
N	Number
N/A	Not applicable
NC	Not calculable
NC	Negative control
NGS	Next generation sequencing
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NIPT	Non-invasive prenatal testing
NMH	North Middlesex Hospital
NPV	Negative predictive value
NS	Not specified
NVIUP	Non-viable intrauterine pregnancy
OTU	Operational taxonomic unit
PAEP	Placenta-protein 14
PAEP	Glycodelin
PAPP-A	Pregnancy-associated plasma protein A
PBAC	Pictorial blood loss assessment chart
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
POC	Products of conception
PPROM	Preterm pre-labour rupture of membranes
PPUL	Persistent PUL
PPV	Positive predictive value
PRISMA	Preferred reporting systems for systematic reviews and meta-analyses
Prog	Progesterone
PROSPERO	Prospective register of systematic reviews
pts	Points
PUL	Pregnancy of unknown location
PVB	Per vaginal bleeding

QCCH	Queen Charlottes' and Chelsea Hospital
QUADAS	Quality assessment tool for diagnostic accuracy studies
RCOG	Royal College of Obstetricians & Gynaecologists
RDP	Ribosomal Database Project
REC	Research ethics committee
ROC	Receiver operating curve
RPOC	Retained products of conception
rRNA	ribosomal Ribonucleic acid
RSH	Royal Surrey Hospital
SAE	Serious adverse event
SD	Standard deviation
SGH	St Georges' Hospital
SMH	St. Marys' Hospital
SMM	Surgical management of miscarriage
SOP	Standard operating procedure
spp.	Species
SROC	Summary receiver operating curve
STAMP	Statistical Analysis of Metagenomic Profiles
STARD	Standards for reporting diagnostic accuracy
TBE	Tris/Borate/EDTA
TN	True negative
TNF	Tumour necrosis factor
TP	True positive
TR	Trophoblast regression
TrisHCL	Tris hydrochloride
TSH	Thyroid-stimulating hormone
TVS	Transvaginal ultrasound scan
UK	United Kingdom
UPT	Urine pregnancy test
US	Ultrasound
USA	United States of America
UV	Ultraviolet
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
VIUP	Viable intrauterine pregnancy
WMUH	West Middlesex University Hospital
WXP	Wexham Park Hospital

INTRODUCTION

The early pregnancy period

Research into pregnancy has largely focussed on the second and third trimesters. The early pregnancy period commences at the point of conception up until the 14th complete week of pregnancy. Current routine antenatal care in the United Kingdom consists of regular visits, with only one in the first trimester (the 'booking' appointment, where the pregnancy is booked with the womans' hospital of choice) whereas routine appointments in the third trimester may be as frequent as every week[1]. This model of care seems to focus on late detection rather than early prediction, which is in contrast to many other areas of medicine where prevention of disease is the primary objective[2]. It is now becoming apparent that many of the complications which may occur in the late first trimester, the latter part of pregnancy and longer lasting effects on the neonate may actually be initiated much earlier than initially thought[3-5]. This provides a window of opportunity to predict adverse events much earlier and therefore undertake preventative measures that may stop these adverse events from occurring.

The establishment of specialised Early Pregnancy Units

The main adverse pregnancy outcomes that occur in the early pregnancy period itself are miscarriage and an ectopic pregnancy. Ectopic pregnancy continues to be a leading cause of maternal mortality in the UK. The 2019 MBRRACE-UK report, a national triennial report covering lessons to be learnt from causes of maternal mortality, showed that in 2015-17 6 out of 99 women with a pregnancy <24 weeks gestation died from complications related to early pregnancy and 5 of them had an ectopic pregnancy[6]. This equated to 0.18 (95% CI 0.05-4.49) deaths per every 100,000 maternities. In all cases, a delay in prompt diagnosis contributed significantly to the cause of death and the authors commented in all cases "*care could have been improved and… improved care may have made a difference to outcome*".

It is therefore a key priority in national guidance[7], which states "*be aware that women with a pregnancy of unknown location could have an ectopic pregnancy until the location is determined*" and that "*improvement in the diagnosis and management of early pregnancy (problems)… is therefore of vital importance, in order to… avoid the unnecessary deaths of women with ectopic pregnancies*". This has led

to the establishment of Early Pregnancy Units (EPUs), which provide specialised care to women with problems in the early pregnancy period. They have been shown to result in women receiving higher quality care as well as being cost-effective[8, 9]. A systematic review showed patients value the more patient-centred approach this service provides with staff appropriately trained in breaking bad news and providing up to date information on complications that may occur in early pregnancy[10].

Background to the studies in this thesis

A pregnancy of unknown location (PUL) is defined as the clinical scenario where a woman presents with a positive urine pregnancy test but the location of the pregnancy cannot be determined when a transvaginal ultrasound scan (TVS) is performed. It is important to note that the term 'PUL' is not a final diagnosis but an intermediary classification until the final outcome of the pregnancy can be established.

The number of early pregnancies initially classified as PUL has been described as anywhere between 8 and 42%[11]. A lower PUL rate may correlate with higher quality ultrasonography and the ability of examiners to better identify pregnancy location at an initial ultrasound scan[11]. Most women with an early pregnancy problem in the UK are seen in the context of a specialised EPU. Broadly speaking, a PUL rate of <15% is accepted as a sign that the quality of scanning in an EPU is satisfactory[12]. Other factors, however (for example the gestational age women present at), independent of scanning quality can also affect the PUL rate.

Despite dedicated EPUs, there remains a great deal of heterogeneity in the way women with a PUL are managed and this is not always evidence-based. There is therefore an unmet clinical need to improve the care of these patients. This manuscript outlines original research on how biomarkers, both novel and those in current clinical use, can be best used to optimise the treatment women with a PUL are given.

Pregnancy of Unknown Location: terms & definitions

There is significant variation in the way final pregnancy outcomes in PUL are defined. Several consensus papers have been published to clarify this issue[12-14], but there unfortunately remains variability in their conclusions.

Broadly speaking, the final pregnancy outcome in a woman initially classified as a PUL falls into one of four categories:

(i) Intrauterine pregnancy (IUP – this can either be viable (VIUP) or non-viable (NVIUP))

(ii) Failed PUL (FPUL)

(iii) Ectopic pregnancy (EP)

(iv) Persistent PUL (PPUL)

An IUP is defined as the presence of an intra-uterine gestation sac (IUGS) within the endometrial cavity on TVS. A VIUP is when the IUGS contains an embryo with visible cardiac activity. A NVIUP is defined as a pregnancy that is confirmed as intra-uterine but subsequently miscarries (meeting the national criteria for diagnosis of a miscarriage[15, 16]).

A FPUL is defined as a pregnancy where the location is never conclusively determined on TVS but biochemically behaves like a pregnancy that is failing/ miscarrying. Serum hCG (human chorionic gonadotrophin) levels, i.e. the 'pregnancy hormone', are routinely used in this assessment. These patients rarely require intervention and resolution of the pregnancy is usually confirmed with a negative urine pregnancy test two weeks after their initial consultation.

An EP is defined as an extra-uterine mass seen on TVS or confirmed via histopathological examination after a salpingectomy.

The definition of a PPUL remains contentious but is commonly described as when >3 serial serum hCG levels increase or decrease by less than 15% and the location of the pregnancy remains unclear on TVS. These pregnancies are likely to be EP that are too small to be visualised on TVS.

Ectopic Pregnancy

An ectopic pregnancy (EP) is defined as a pregnancy outside of the uterine cavity, of which approximately 95% will be within a fallopian tube (either the isthmic, ampullary or fimbrial portions of the tube, most commonly the ampullary region). The remaining 5% are located in either the interstitium, cervix, a previous caesarean section scar, ovary or abdominal cavity (figure 1).

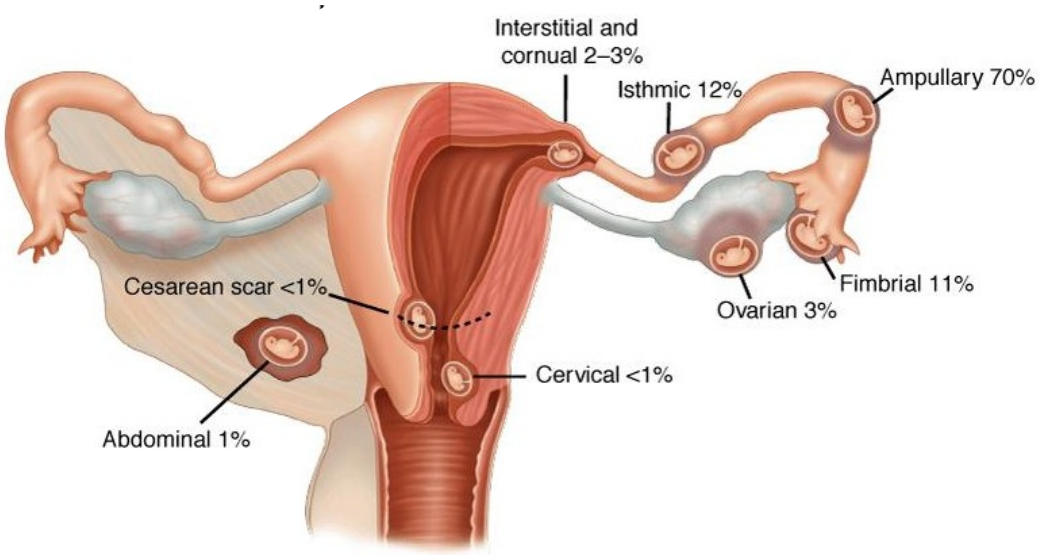


Figure 1: potential sites for an ectopic pregnancy. Adapted from Cunningham et al, Williams Obstetrics, 23rd Edition, McGraw-Hill, USA

The incidence varies depending on the population but in the UK, 11/1000 spontaneously conceived pregnancies are ectopic[17] and this can increase to 4% in assisted conception patients[11]. The estimated maternal mortality is 0.2/1000 ectopic pregnancies[17] and it remains the leading cause of early-pregnancy related maternal death in the UK[18]. In the developing world, it has been estimated that "*10% of women admitted to hospital with a diagnosis of EP ultimately die from the condition*" [19].

Consequently, determining which PUL will go onto have a final outcome of EP versus those that don't is an essential component of early pregnancy care.

TVS remains the non-interventional gold standard for diagnosing EP[20]. The majority (60%) of tubal EP will have the appearance of an inhomogenous mass (figure 2) in the adnexa, commonly referred to as the 'blob sign'. 20% tubal EP will have a 'bagel sign', where an inhomogenous mass containing a gestation sac is seen in the adnexa (figure 3). The final 20% are easiest to visualize on TVS and consist of an inhomogenous mass with a gestation sac, yolk sac and fetal pole that may contain positive fetal cardiac activity (figure 4)[20, 21]. An EP may also be strongly suspected if there is evidence of a haemoperitoneum in the pelvis, indicating that the EP has ruptured (figure 5). This requires immediate surgical intervention.

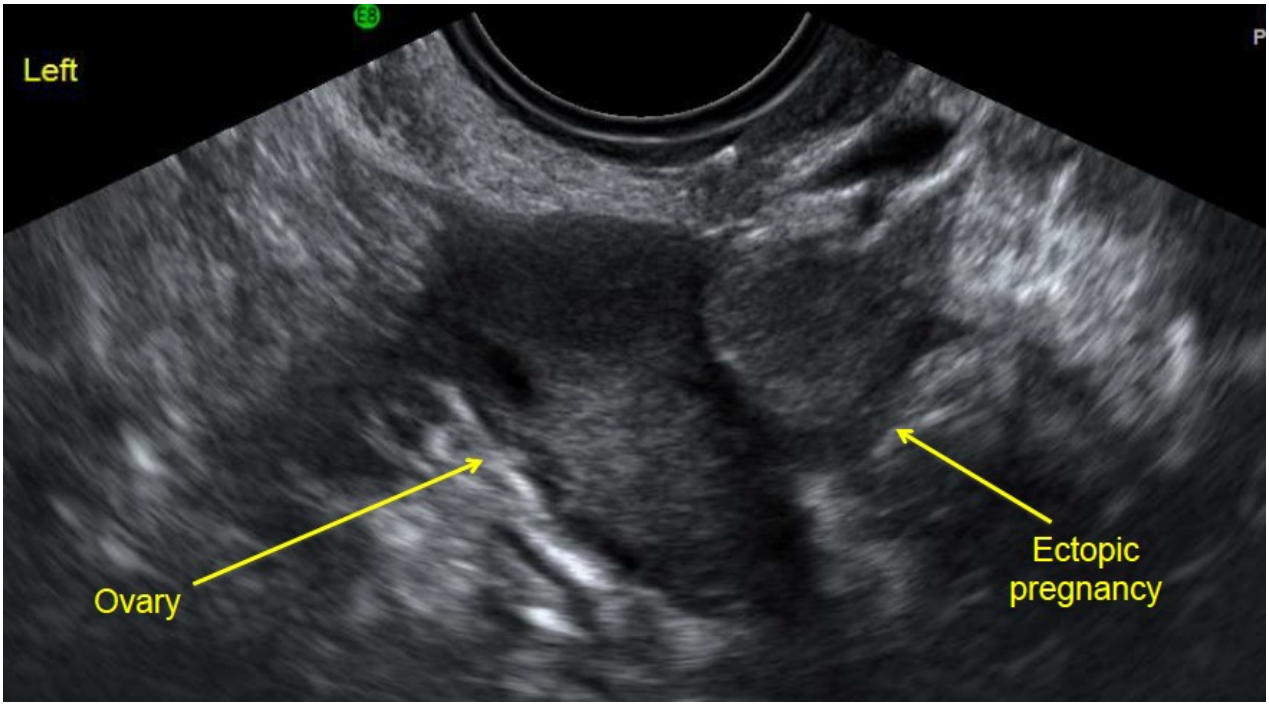


Figure 2: a tubal ectopic pregnancy seen as an inhomogenous mass ('blob sign')

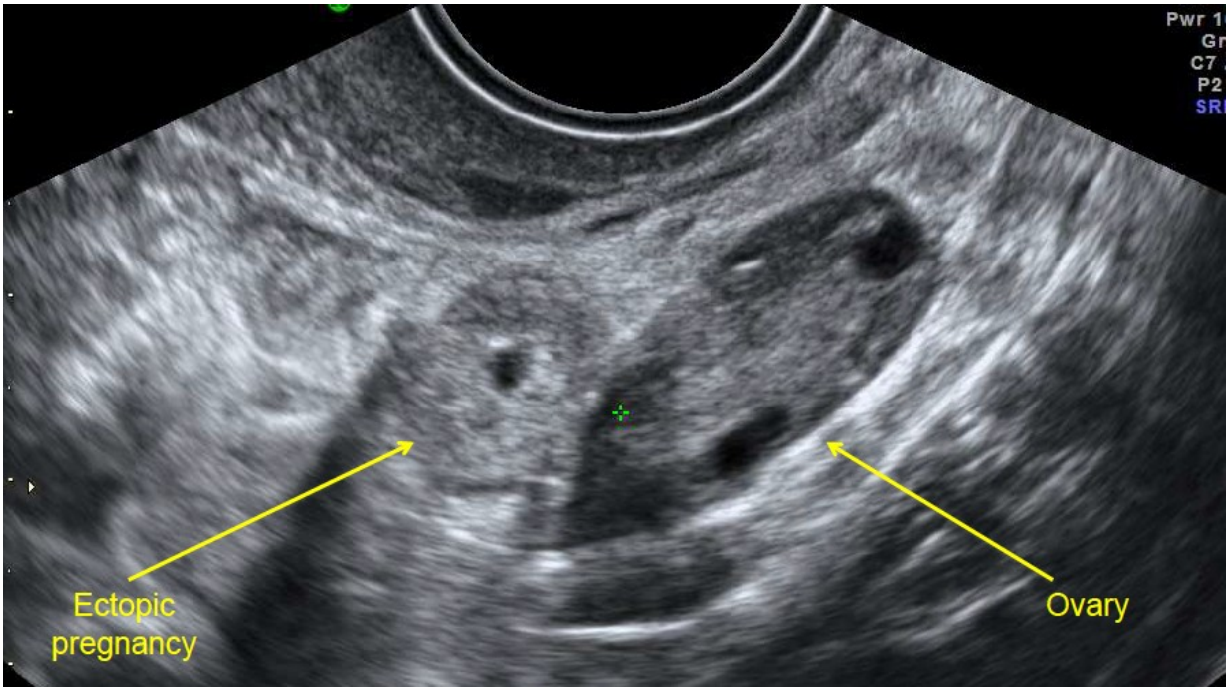


Figure 3: a tubal ectopic pregnancy seen as an inhomogenous mass with a gestation sac ('bagel sign')

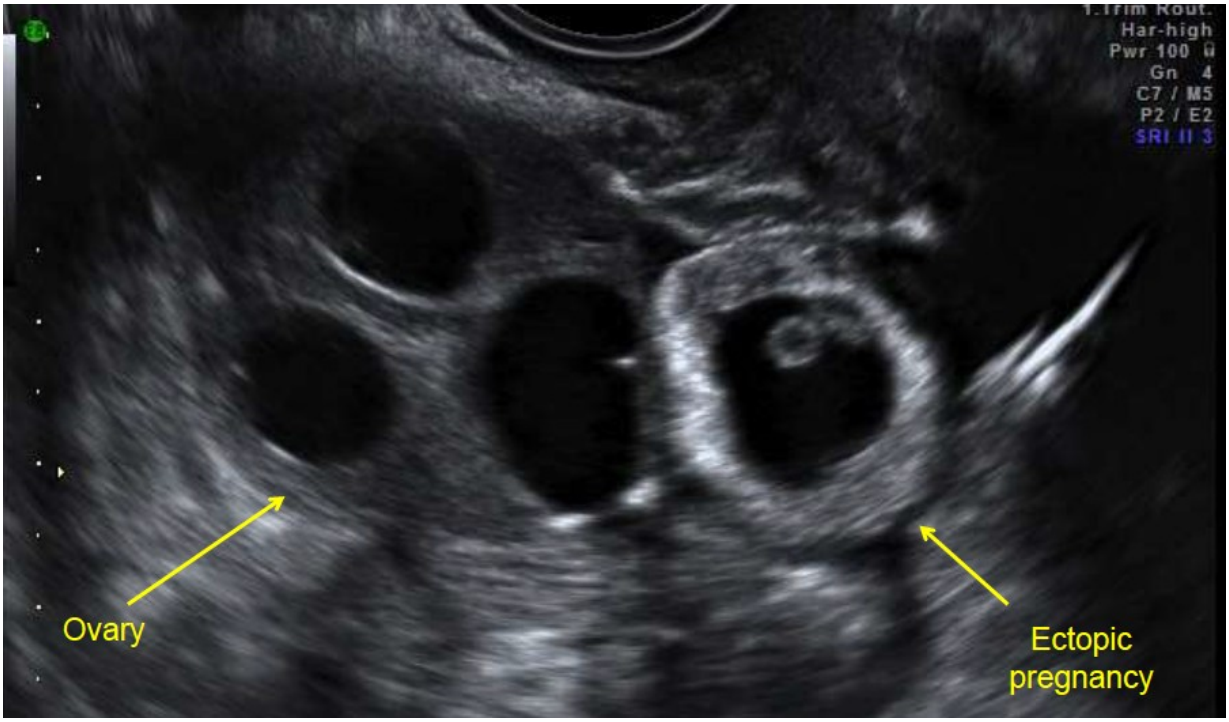


Figure 4: a tubal ectopic pregnancy seen as an inhomogenous mass with a gestation sac, yolk sac and fetal pole that may contain positive fetal cardiac activity

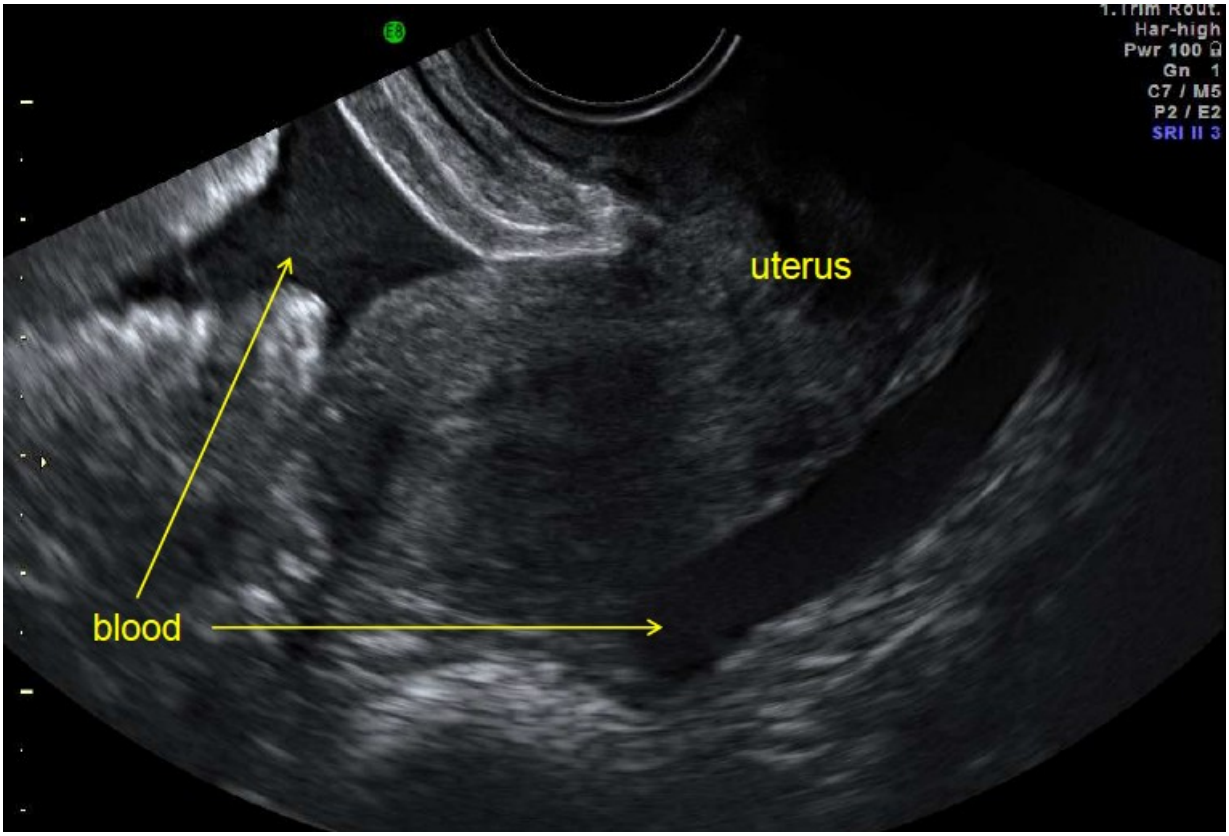


Figure 5: a large haemoperitoneum (blood in the pelvis) seen on transvaginal ultrasonography

Applicability in a low-income resource setting

It is important to state that the presence of dedicated early pregnancy units requires significant financial resources as it involves up to date ultrasonography equipment, facilities to perform serum biochemistry tests and trained specialists who can provide expert care. Early pregnancy care of this type, therefore, is largely confined to developed countries.

Pregnancy of Unknown Location: high-risk versus low-risk

Women with a PUL often undergo multiple blood tests, ultrasound scans and hospital visits in the fear that an ectopic pregnancy (EP) is being missed. In reality, the majority of PUL have a final outcome of an IUP (17-41%) or FPUL (47-70%) rather than EP (8-16%)[22]. Consequently, there is a clinical need to rationalize management so few women with a final outcome of EP are misdiagnosed and unnecessary follow-up for the remaining women with a PUL is avoided. There has, therefore, been an effort towards managing patients based on the probability of complications related to EP, i.e. as either 'low-risk of complications' (a predicted final outcome of IUP or FPUL) or 'high-risk of complications' (predicted final outcome of EP or PPUL), rather than focusing on definitively determining the location of the pregnancy (see figure 6). This risk-stratification approach allows finite healthcare resources to be directed towards PUL deemed at high-risk of complications and the unnecessary follow-up of women at low-risk of complications to be reduced.

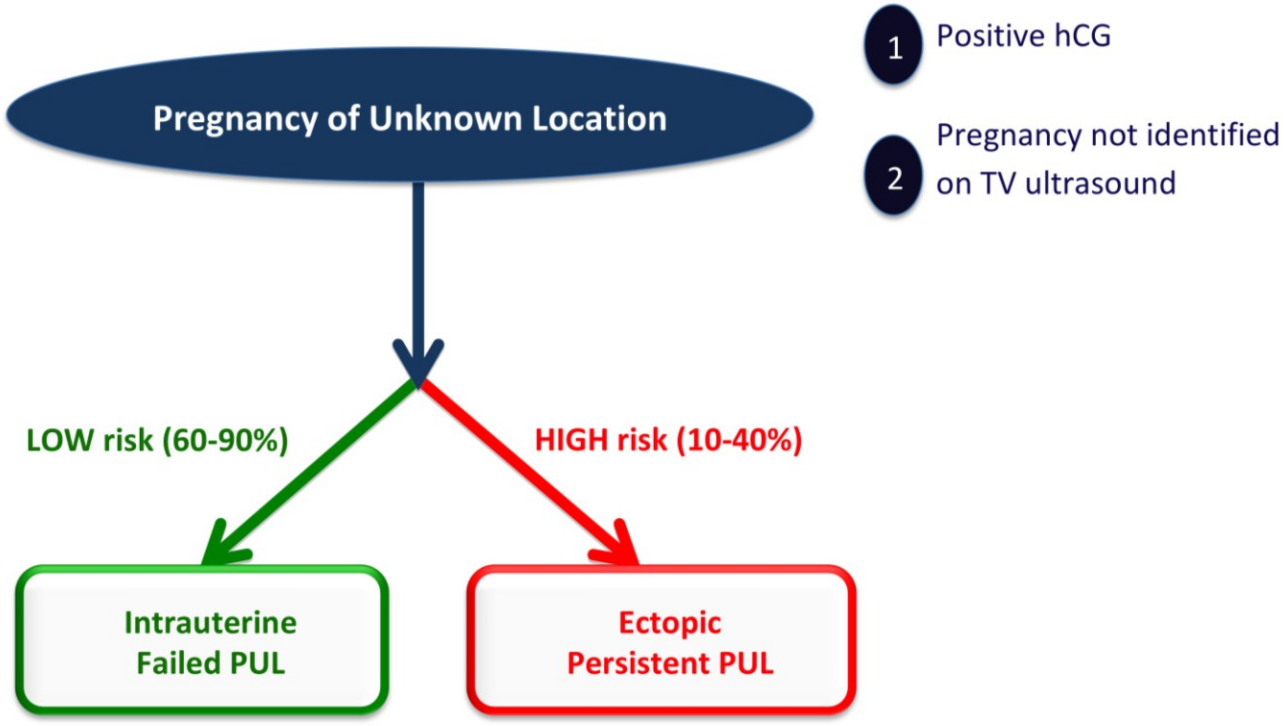


Figure 6: Schematic for triaging women with a pregnancy of unknown location (PUL) as either low-risk or high-risk of clinical complications. hCG = human chorionic gonadotrophin; TV = transvaginal

Pregnancy of Unknown Location: biomarkers used in current clinical practice

Most women classified as a PUL will present with vaginal bleeding and/ or lower abdominal pain. Symptoms, however, are not predictive of which PUL are at high-risk of complications[23, 24]. Several management protocols have been proposed in trying to achieve reliable risk stratification. These focus on the serum biomarkers progesterone and human chorionic gonadotrophin (hCG).

Serum progesterone

Progesterone is a steroid hormone essential for female reproduction, adequate functioning of the central nervous system and as a precursor to the production of other endogenous steroids such as corticosteroids and sex hormones. With regards to female reproduction, progesterone is essential for the maintenance of pregnancy, hence its name, derived from the Latin *pro gestationem*. It accounts for the secretory phase of the menstrual cycle and prepares the endometrium for potential implantation of a fertilised ovum. If this does not occur, decreasing levels of progesterone lead to a withdrawal bleed, i.e. menstruation[25]. If pregnancy does occur, the production of progesterone is taken over by the corpus luteum of the ovary in early pregnancy and by the placenta in later pregnancy. It is also produced in much smaller quantities by the adrenal glands. As well as preparing the uterus for implantation, progesterone is involved in modulating the maternal immune response[26-31] and allowing the acceptance of a pregnancy. It suppresses uterine contractility[32-34], improving luteal phase support in early pregnancy and the utero-placental circulation at later gestations[35, 36]. As hCG is responsible for sustaining the production of progesterone via the corpus luteum, the two biomarkers are intrinsically linked in maintaining an early pregnancy.

In the context of PUL, progesterone appears to be a good predictor of pregnancy viability rather than location[37]. However, serum progesterone levels of ≤10nmol/l have been found to be associated with a final outcome of FPUL[38, 39], which could be used to select women for appropriately reduced follow-up. Selecting a threshold of 10nmol/l, however, is associated with more EP being misclassified into the low-risk group when compared to lower cut-offs[40] and the optimal progesterone cut-off value remains contentious.

Serum hCG

Serum hCG is the most common biomarker used to clinically manage PUL[41]. hCG is a glycoprotein produced by syncytiotrophoblast cells, the precursors to the fetal placenta. It is heterodimeric, and consists of an alpha and a beta subunit[42]. The alpha subunit is identical to that found in other hormones such as follicle-stimulating hormone (FSH), luteinizing hormone (LH) and thyroid-stimulating hormone (TSH). The beta subunit is unique to hCG and is therefore clinically used to quantify hCG levels[43, 44]. For the purposes of this thesis, all reference to serum hCG levels implies the beta-subunit was measured. hCG is responsible for maintaining the corpus luteum in pregnancy and therefore the production of progesterone. A meta-analysis by Van Mello et al[45] found that single hCG measurements have little diagnostic value and hCG trends help determine the final outcome in a PUL. This finding is supported by UK national guidance on the subject[15]. Several methods in interpreting hCG trends are in current clinical use:

hCG ratio

This is described as the 48 hour hCG level divided by the initial (0 hour) hCG. A ratio of <0.87 (consistent with a decrease in hCG levels of ≥13% over 48 hours) is likely to indicate a final outcome of FPUL[46]. A ratio of >1.66 indicates a final outcome of IUP[47] and a ratio between those two values is likely to result in a final outcome of EP/ PPUL[21, 41].

Mathematical risk prediction models

Logistic regression models are also used in the management of PUL. Of these, the most widely published are a series of 'M' models.

The M1 model

Model 1 (M1) assessed the test performance of the hCG ratio[48]. Of a total of 388 consecutively recruited women with a PUL, 189 cases were included in a training set and 199 cases in a test set. The M1 model had a bias corrected AUC (95% CI) of 0.823 (0.752 – 0.929) and 0.914 (0.842 – 0.989) in the training

and test sets respectively in correctly distinguishing between the three different PUL final pregnancy outcomes (IUP; EP; FPUL). It had a sensitivity of 83.3%, specificity of 88.0%, PPV (positive predictive value) of 44.1% and NPV (negative predictive value) of 98.8% in correctly predicting a final outcome of EP.

The M2 model

The M2 model[48] was based on the average serum progesterone result (the mean of the serum progesterone result and 0 and 48 hours) and had a bias corrected AUC (95% CI) of 0.867 (0.805 – 0.939) and 0.728 (0.578 – 0.849) in correctly distinguishing between the three different PUL final pregnancy outcomes.

The M3 model

The M3 model[48] utilised the hCG ratio, average serum progesterone result and the patients' age as variables. It had a bias corrected AUC (95% CI) of 0.902 (0.843 – 0.951) and 0.834 (0.756 – 0.946) in correctly distinguishing between the three different PUL final pregnancy outcomes. See figure 7.

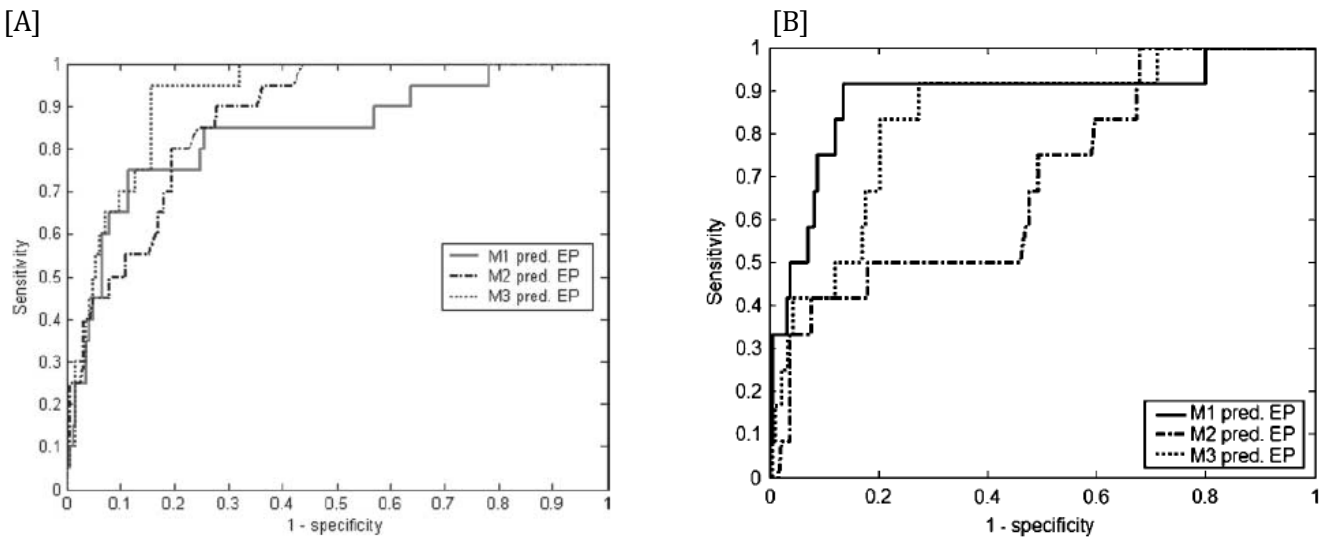


Figure 7: Comparison of the ROC curves for the prediction of ectopic pregnancy as a final pregnancy outcome in women with a PUL by the M1, M2 and M3 models: [A} training data set (n=185), [B] test data set (n=195). Reproduced from Condous et al[48].

The M4 model

The M4 model[49] is based on the initial serum hCG and hCG ratio as predictive variables. In a diagnostic accuracy study on 1962 PUL, it classified 69.6% PUL as at low-risk of complications with a NPV of 97.5%[22]. It had a sensitivity of 88.0%, i.e. correctly classified 88.0% of PUL with a final outcome of EP/PPUL as at high-risk of complications. This model was superior in test performance to a single progesterone cut-off of <10nmol/l or the hCG ratio[50] and outperformed the M1, M2 and M3 models[21].

The M5 model

To assess whether clinical variables such as the amount of pain or bleeding experienced by the patient had any utility, the M5 model was built. It was based on the log of the average serum hCG, amount of vaginal bleeding experienced by the patient, the hCG ratio and the quadratic effect of the hCG ratio as variables. The authors concluded that clinical information did not significantly improve the performance of logistic regression models in predicting PUL outcomes and that biochemical data was more useful[23].

Evaluating test performance in diagnostic accuracy studies

The commonly used test performance measures in diagnostic accuracy studies include:

Percentage of low-risk classifications

These are PUL cases selected by the index test (model) for reduced follow-up

Negative predictive value (NPV)

This is defined as the probability that subjects with a negative screening test truly don't have the disease. Specifically for this thesis, it is the percentage of patients with a final pregnancy outcome of FPUL or IUP amongst those classified as low-risk by the model being used

Sensitivity

This is defined as the proportion of actual positives that are correctly identified as such. Specifically for this thesis, it is the percentage of EP correctly classified as high-risk by the model being used

False positive rate

This is also known as as a type I (alpha) error and is the probability of falsely rejecting the null hypothesis for a particular test. Specifically for this thesis, it is the percentage of patients with a final pregnancy outcome of FPUL or IUP incorrectly classified as high-risk

Area under the receiver operating curve (AUC)

This is also a measure of test performance measurement. A ROC (Receiver Operating Curve) is a probability curve. Specifically for this thesis, the AUC shows how much a model is capable of distinguishing between different final pregnancy outcomes. The higher the AUC, the better the model is at risk prediction (see figure 8).

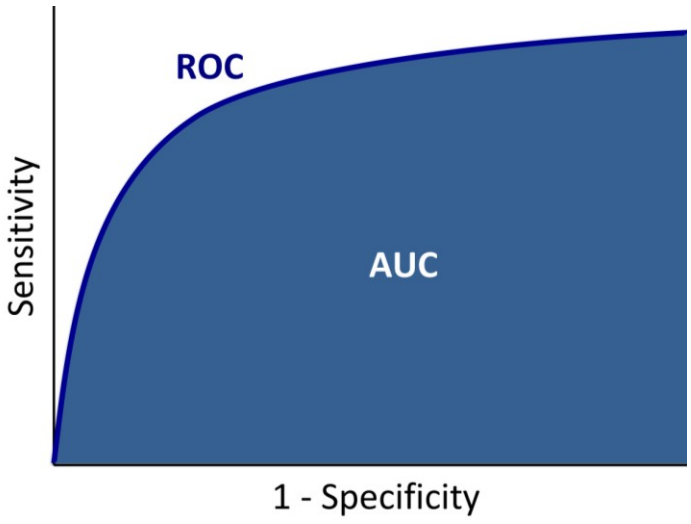


Figure 8: representation of a ROC (Receiver Operating Curve) and AUC (Area under the Receiver Operating Curve)

Novel serum biomarkers in the management of PUL

Biomarker discovery in the context of PUL and EP has previously used anatomical landmarks and centred around five key areas (*adapted from* Cartwright et al, see figure 9[51]), focussing on biomarker abnormalities in the:

- fallopian tube
- uterus
- embryo/ trophoblast
- corpus luteum
- inflammation/ peritoneal irritation

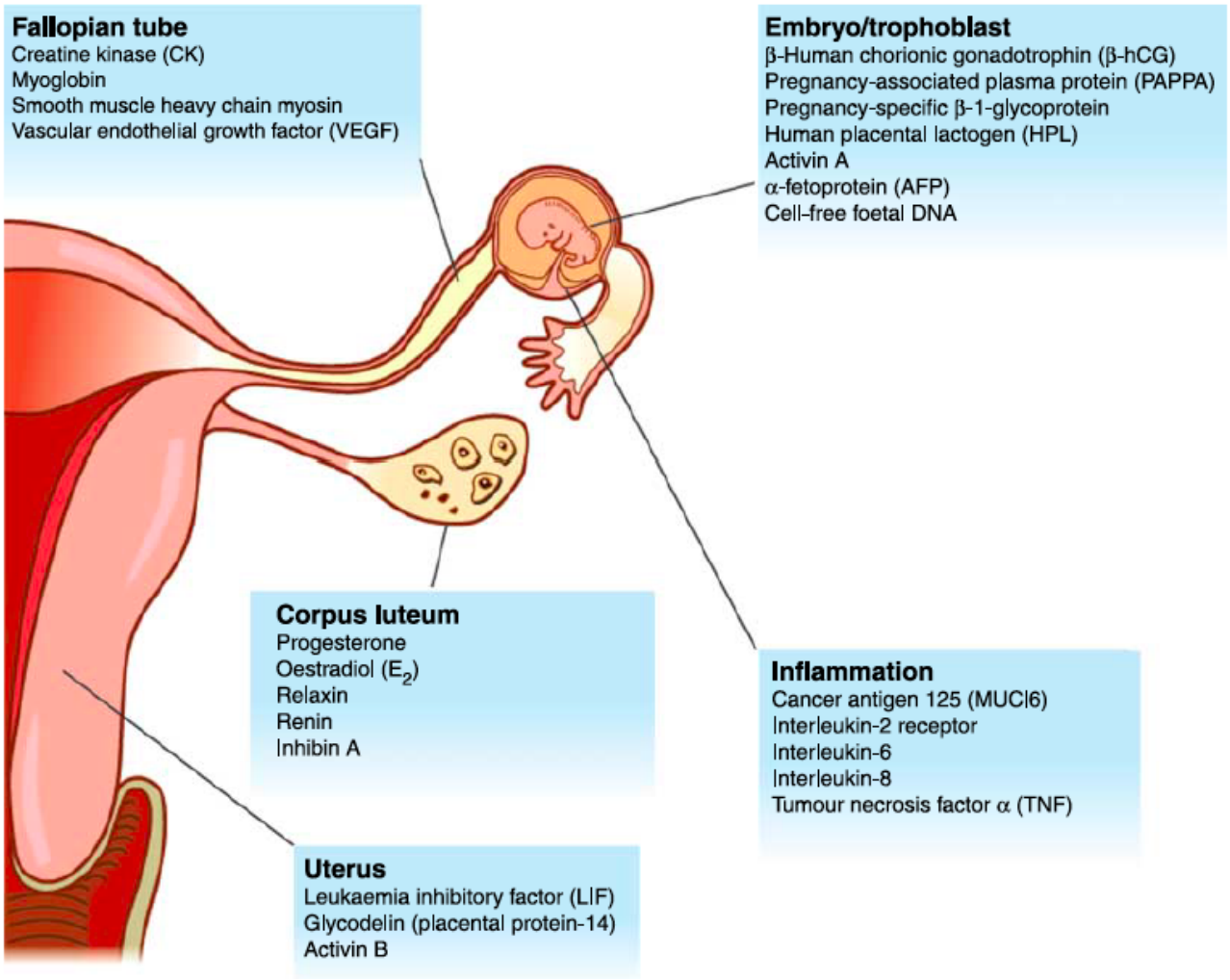


Figure 9: Five key areas historically explored in novel serum biomarker discovery for ectopic pregnancies. Reproduced from Cartwright et al[51]

Fallopian tube

The fallopian tube contains circular smooth muscle which is responsible for its' peristaltic action. It has a cross-sectional diameter of less than 1cm. This is stretched and disrupted by the presence of a tubal ectopic pregnancy so it has been hypothesized that this damage potentially leads to the release of detectable biomarkers in serum, which can then be used to diagnose the presence of a tubal EP. The markers focussed on are largely indicators of muscle injury and include creatine kinase (CK), myoglobin, smooth muscle heavy-chain myosin (MYH11) and vascular endothelial growth factor (VEGF).

CK is an enzyme released from damaged muscle cells and was historically used as a marker of myocardial infarction[52] but has since been superceded by the discovery of Troponin I. It is now used in clinical practice to diagnose rhabdomyolysis or myositis. An initial study determined a CK cut-off level of >45IU/L could discriminate between EP, VIUP and NVIUP with little overlap (n=51, 17 in each group)[53]. Further work regarding its diagnostic utility in the diagnosis of EP, however, have not been able to replicate these findings[54-61]. CK is also increased in pelvic inflammatory disease and appendicitis[62], but to a lesser extent than with an EP[63-66], particularly when the ectopic location within the fallopian tube is at the isthmus rather than the ampullary or fimbrial regions[67]. This is unlikely to be helpful as the majority (70%, see figure 1) of tubal EP are within the ampullary region of the tube. CK is elevated to the greatest extent with a ruptured EP[68]. It, therefore, may be best used as a marker of tubal rupture but this is unlikely to have clinical utility as the diagnosis of a ruptured EP is usually clear from the clinical presentation.

Both myoglobin and MYH11 are protein components of muscle fibres and elevated levels are found with striated muscle damage[69]. Myoglobin is present to a much lesser extent with smooth muscle damage[70, 71], the main muscular component of the fallopian tube. Studies, therefore, have not shown it to adequately discriminate between healthy IUP and tubal EP in the first trimester[72]. There may be some utility for MYH11, which does appear to be increased in women with an EP and showed a NPV of 91% but a PPV of only 22%[65] so has not been looked into further as a discriminatory test for EP.

VEGF is a potent angiogenic factor, stimulating intense angiogenesis at the site of normal implantation into the endometrium[73] and at the site of the corpus luteum[74]. It is activated by tissue hypoxia[75] and it was therefore hypothesized that a tubal EP implantation takes place in the presence of greater tissue hypoxia than within the uterus and elevated levels of VEGF may be found as a result of this. Several pilot studies[76-78] did demonstrate levels with an EP were higher compared to those with a VIUP, even after ovarian stimulation prior to an IVF cycle[79] when a cut-off of 700 pg/ml was used[80]. A cut-off level of >200 pg/ml was found to be discriminatory with one study (n=15) demonstrating a sensitivity of 88% and a specificity of 100%[81] but this finding was not replicated in a larger study (n=85, sensitivity 56%; specificity 51%)[82]. Although pilot work on VEGF as a marker of EP was initially promising, it is best at discriminating between EP and VIUP, which does not take into the account the other potential final pregnancy outcomes in women with a PUL, the majority of which will actually be a failing pregnancy. Furthermore, a larger subsequent study has disproven the initial preliminary results.

Uterus

The normal implantation of a fertilised egg into the endometrium releases a variety of potentially measurable biomarkers and it has been hypothesized that the levels of these may be altered when the process of implantation is aberrant, as with miscarriage or in the incorrect location, as with EP. The main biomarkers investigated are leukaemia inhibitory factor (LIF), glycodelin (placenta-protein 14, i.e. PAEP) and activin B.

LIF is a glycopeptide cytokine and member of the interleukin-6 family, a pro-inflammatory cytokine[83]. A small pilot study (n=11) found serum LIF levels to be decreased in women with an EP with a sensitivity of 73% and specificity of 72%[84]. A larger study[77] failed to find any difference in LIF levels between different early pregnancy outcomes and it is therefore unlikely to be useful in the prediction of outcome in women with a PUL.

PAEP is secreted by the endometrium and therefore higher levels are seen throughout the menstrual cycle, with a peak in the secretory phase[85, 86]. Levels steadily rise in early pregnancy, plateauing between 8-10 weeks gestation and decline thereafter[85]. Lower levels of PAEP have been observed in women with an EP[87, 88], irrespective of whether the tube was intact or the EP had ruptured[89]. Once again, larger studies[77] failed to consistently find any real difference in PAEP levels between different PUL final pregnancy outcomes.

Activin B is a member of the transforming growth factor B1 superfamily and is produced by the pregnant decidua. It is linked to FSH and increases its' levels. It has been found to be lower in women with an EP in a study undertaking shotgun genomic profiling[90]. No other studies have replicated these results, however, so the preliminary positive findings should be viewed with caution.

Embryo/ trophoblast

The majority of work trying to determine the pathophysiology of ectopic pregnancies has focussed on potential markers of aberrant implantation. The only one used clinically is βhCG. Other investigated biomarkers include pregnancy-associated plasma protein A (PAPP-A), pregnancy-specific β-1-glycoprotein, human placental lactogen (HPL), activin A and α-fetoprotein (AFP).

PAPP-A is a protease produced by placental syncytiotrophoblast cells[91]. Lower levels were found in women with an EP[92-94] but the detectable levels are generally very low prior to 7 weeks' gestation[77, 78]. As the majority of EP will be diagnosed at or before that gestation, it is unlikely to have substantial utility as a single predictive marker of EP.

Pregnancy-specific β-1-glycoprotein is a glycoprotein also produced by placental trophoblast cells[95] and lower levels were found in women with an EP compared to other early pregnancy outcomes[96, 97]. Serial measurements were also assessed and patients with an EP had a suboptimal rise in levels whereas those with a viable pregnancy had a steeper increase[78, 98] but the studies performed were small in size and further validation data is required.

HPL is a placental product detectable in the first trimester[99]. It is primarily involved in lipid and glucose homeostasis and increases the production of insulin[100]. Studies on HPL levels in EP have had conflicting results[77, 78, 101] and it has therefore not been proven to have any real utility in the diagnosis of EP.

Activin A is a glycoprotein produced by the pregnant decidua and ovary[102, 103] and was found to have excellent test performance in one study[104] with a sensitivity of 100% and specificity of 99.6% (n=486) when a cut-off of <0.37ng/ml was applied but this was not replicated in further studies[105] and so its' use cannot be advised in current clinical practice without further, larger validation studies.

Unlike most other markers related to implantation that appear to be lower in women with an EP, AFP levels have been found to be elevated[106]. It is produced by the fetal liver and yolk sac[107] in early pregnancy and elevated levels have also been found in women with a threatened miscarriage, indicating it may be a generic marker of aberrant implantation but not specifically of an EP.

Corpus luteum

The primary function of the corpus luteum is to produce progesterone and provide luteal phase support in early pregnancy before the placenta takes over progesterone production. It has therefore been hypothesized that reduced levels of progesterone are related to adverse pregnancy outcomes[108-110]. Studies following on from these, however, reported mixed results[111-113] and that progesterone cut-off levels are unlikely to be useful as a single predictor of EP[78, 98, 101, 106, 114-119]. Meta-analyses have shown that it is a good marker of pregnancy viability but not of pregnancy location and therefore not a good diagnostic marker for EP[37, 120]. Other biomarkers such as oestradiol (E2), relaxin, renin and inhibin A have also been proposed as markers of abnormal corpus luteal function.

Oestradiol is a sex steroid produced by the ovary and corpus luteum in response to rising hCG levels. Levels have been found to be reduced in EP[98] and one study[112] assessing 205 patients found a cut-off of <650pg/ml discriminated between EP and viable pregnancies with a sensitivity of 100% and specificity

of 99%. These findings, however, were not replicated in later studies[101, 106, 113] and levels overlapped considerably with other final pregnancy outcomes. Single values of E2 are therefore unlikely to be clinically useful.

Relaxin is a peptide hormone produced by the corpus luteum specifically in pregnancy[121]. Unlike many other tested biomarkers, which only begin to increase after 7 weeks' gestation, levels are increased soon after conception[122]. Lower levels have been shown to correlate with EP[123] but there was significant overlap with other pregnancy outcomes[113], rendering it a poor diagnostic marker for EP.

Renin is an enzyme involved in the regulation of blood volume and is largely produced by the kidneys. In pregnancy, the ovaries have also been shown to produce it[124], with lower levels seen in EP[125]. Once again, studies have demonstrated too poor a sensitivity and specificity (75% and 76% in one study[125]) for this to be used in current clinical practice.

Inhibin A levels are regulated by hCG as it is a peptide product of the corpus luteum[126]. It is also produced, to a lesser extent, by trophoblast cells[127]. Levels has been shown to be lower in EP[128] and decrease more significantly after successful treatment of EP compared to hCG, likely secondary to its' shorter half life[129]. It therefore, may be a better marker of the efficacy of expectant or medical management of an EP rather than the diagnosis.

Inflammation/ peritoneal irritation

Pelvic pain is a common presenting complaint in women with an EP and this is likely to be secondary to peritoneal irritation/ inflammation. It has therefore been hypothesized that biomarkers released secondary to this may prove to be of diagnostic importance and become elevated in women with an EP. Pro-inflammatory cytokines such as interleukin 6 (IL-6), IL-8, IL-2 receptors and tumour necrosis factor (TNF) have all been investigated. Cancer antigen 25 (CA125, also known as MUC16) has also been examined as a surrogate marker for peritoneal irritation.

CA125 is a cell-surface antigen found in the female genital tract and pregnant decidua. It is used in current clinical practice in the diagnosis and management of ovarian cancer[130] but can also be raised in benign gynaecological conditions such as endometriosis. It has been shown to be elevated in early pregnancy[131, 132] but was found to be a better marker of miscarriage (NVIUP) rather than EP when compared to VIUP[133]. Contrary to what was expected, levels were in fact reduced rather than increased[113, 134] in EP or no different[101] from other early pregnancy outcomes. It is therefore unlikely to be sufficient as a single diagnostic biomarker in women with a PUL.

IL-2 receptors are expressed on the surface of T-lymphocytes. A study of 72 women found no difference in levels between those with a VIUP (n=33), NVIUP (n=22) or EP (n=17)[135]. IL-6 is secreted by T-lymphocytes, IL-8 is a chemokine produced by macrophages and TNF is involved in the acute phase inflammatory immune response, integral to the process of apoptotic cell death. In the same study[135], all three cytokines were found to be significantly higher in women with an EP compared to those with a VIUP or NVIUP. No other studies have been performed to validate these findings.

hCG and progesterone continue to be the only serum biomarkers used clinically in the management of PUL. Several novel biomarkers that may be beneficial either in combination with, or as a replacement for, hCG and progesterone have been explored but none have had enough efficacy to be implemented into clinical practice.

Limitations of current novel biomarker studies

Four main phases of novel biomarker discovery have been described[136] – see figure 10.

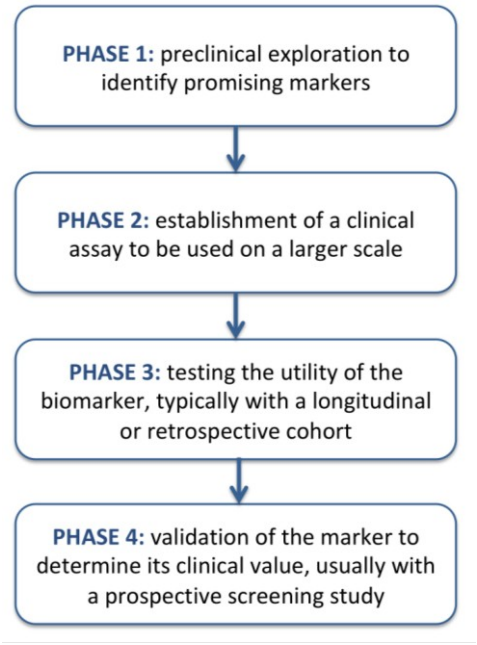


Figure 10: phases of biomarker research. Adapted from Rausch et al[136].

The initial phases have been pursued for multiple serum biomarkers, most of which were detailed in the sections above. The main limitation of these studies is a lack of validation work to support preliminary findings and progression to stage 4. None of these novel markers have been able to consistently differentiate ectopic pregnancies from other early pregnancy outcomes. It is therefore unlikely that there is a single novel biomarker that can accurately determine which PUL are at high-risk of complications. If we look at the NHS combined screening programme[1], a national screening tool offered to all women in the first trimester of pregnancy to determine their risk of common trisomies (13, 18 and 21), it consists of the measurement of serum biomarkers hCG and PAPP-A, combined with the ultrasound-based measurement of nuchal translucency. Future work in this area, therefore, centres on developing multiplexed biomarker tests[136-138] which potentially include ultrasound or clinical variables that can then be integrated into logistic regression models.

Another avenue worth exploring is samples that don't rely on biomarkers being present in the systemic circulation (and therefore measurable in serum samples) but closer to the site of the pathology itself. This includes vaginal swab samples taken in early pregnancy and work on the vaginal microbiome is increasingly being explored in the area of obstetrics and gynaecology.

The microbiome

The term *microbiome* refers to the collection of genomes from all the microorganisms in the environment, whereas the *microbiota* describes the community of microorganisms present within a specific environment[139]. The last decade has seen a huge increase in the study of the microbiome in medicine, with environments such as the gastrointestinal and reproductive tracts being explored. This led to initiatives such as the Human Microbiome Project[140, 141], which commenced in 2008 and aimed to comprehensively characterise microbial communities in the healthy state at five main sites in the body: oral, lung, skin, gut and vagina. The project revolutionised our understanding of the microbial community structures at these sites[142, 143].

Traditional methods for detecting microorganisms centre on culture-based techniques. Newer, culture-independent techniques such as next generation sequencing (NGS) now allow us to detect far more microorganisms than before, revealing bacterial communities that resist traditional cultivation methods. This has potentially enormous implications for our understanding of disease and subsequent therapeutic targets. The study of the human microbiome has focussed on cross-sectional analyses comparing microbial communities detected in healthy controls against those with disease and discovered huge variations in the microbiota of subjects without any apparent clinical disease. A number of variables appear to impact the vaginal microbiome and these include ethnicity[144], the menstrual cycle and sexual intercourse[145-149]. Longitudinal studies have allowed researchers to monitor the effect of treatments or the progression of disease on the composition of commensal microbial communities.

With NGS techniques, bacterial DNA is extracted from samples and amplified by polymerase chain reaction (PCR) using either universal or specific primers. PCR amplicons are then put through DNA sequence analysis. The 16S ribosomal RNA (rRNA) gene is exclusive to bacteria and has highly conserved flank regions ideal for the targeted binding of PCR primers. It contains nine 'hyper-variable' regions (V1 – V9) that differ amongst bacterial species, enabling accurate differentiation between

microbial profiles[150]. The DNA sequences obtained can then be compared with international databases of 16S rRNA sequences to determine the precise species of bacteria identified[151].

The normal vaginal microbiota and pregnancy

Bacteria that colonise the vagina are collectively referred to as the vaginal microbiome. There is now good evidence to suggest the composition of this has an impact on reproductive health. Women of reproductive age have been found to have a vaginal microbiota that is dominated by *Lactobacillus* species[145, 146, 152]. These make up the normal vaginal flora and have a protective effect against more pathogenic, anaerobic bacteria. They exert these effects by producing lactic acid and maintaining a low vaginal pH where pathogenic organisms cannot thrive[153], as well as producing directly bacteriocidal products[154]. The effects of *Lactobacillus* species are mediated by oestrogen, which increases glycogen production by the vaginal epithelium, the main substrate for Lactobacillus growth[155].

When assessed at a species level, the vaginal microbiota is most commonly divided into five main community state types (CSTs)[146]:

CST I – *Lactobacillus* crispatus

CST II – *Lactobacillus* gasseri

CST III – *Lactobacillus* iners

CST IV – 'high diversity'

CST V – *Lactobacillus* jensenii

Four of the CSTs are dominated by *Lactobacillus* species but CST IV describes a vaginal microbiota that is Lactobacillus deplete and comprised of a number of anaerobic and enteric bacteria similar to those found with bacterial vaginosis (BV). Maintaining a stable Lactobacillus-rich vaginal microbiome, especially L. *crispatus*, across pregnancy has been shown to be an important feature of favourable pregnancy outcomes[145, 156]. CST III is the only *Lactobacillus* species that lowers stability and may precede a transition towards a dysbiotic (CST IV) microbiome[157, 158]. Vaginal dysbiosis and a dominance of BV- related pathobionts such as Gardnerella *vaginalis* is associated with adverse pregnancy outcomes[159-161].

The vaginal microbiome and adverse pregnancy outcomes

The Human Microbiome Project was the first time NGS techniques were applied to the study of the vaginal microbiome in pregnancy. Aagaard et al described a consistently stable, Lactobacillus-rich, vaginal microbiome in pregnancy compared to the non-pregnant state[162]. Any disruption of this and a move towards a more dysbiotic microbiome was shown to be associated with adverse pregnancy outcomes. Within the field of preterm birth, it has been hypothesized that dysbiosis is associated with an increased expression of inflammatory mediators into cervicovaginal fluid[163, 164]. This pro-inflammatory environment then compromises the cervico-epithelial barrier[165], activating downstream inflammatory pathways, which results in poor obstetric outcomes such as preterm labour and preterm pre-labour rupture of membranes (PPROM)[156, 166-171].

In early pregnancy, infection has been proposed as a potential cause of both miscarriage[172] and ectopic pregnancy[173]. Previous work has shown an association between miscarriage and bacterial vaginosis[159-161]. The presence or previous infection with Chlamydia *trachomatis* is a known risk factor for developing an EP[174, 175]. These studies have either relied on culture-dependent techniques or sampling has been undertaken late in the diagnosis of an EP. Using NGS in women with a PUL not only allows the detection of pathogenic bacteria that were not amenable to recognition with traditional culture-based techniques, but may enable detection of women at risk of developing an ectopic pregnancy before it can be clinically diagnosed via transvaginal ultrasound. Early detection raises the possibility of discovering potentially modifiable risk factors and therapeutic targets. This has already been explored in the field of preterm birth and late miscarriage, where the presence of bacterial vaginosis is treated with antibiotics[176].

For the purposes of this thesis, the vaginal bacterial microbiome will be primarily discussed in a cohort of women with a PUL. Those with a final pregnancy outcome of a viable intrauterine pregnancy will be used as controls and compared to all other (adverse) pregnancy outcomes, including ectopic pregnancy.

Summary of biomarkers

hCG (and in part progesterone) remain the gold standard serum biomarkers used to appropriately risk-stratify women with a PUL and improve our ability to correctly diagnose an early ectopic pregnancy. The way they are interpreted, however, is incredibly variable across the United Kingdom and clinicians do not always follow the current best evidence. This is because there are a multitude of studies in the literature, all advocating slightly different methods of interpretation. Furthermore, current national guidance on the subject is vague, contributing to the confusion.

When used appropriately, however, these biomarkers can have great utility in the management of women with a PUL. One could therefore argue there is little benefit in exploring the value of additional novel biomarkers, particularly as the vast majority of novel serum biomarker discovery has been unfruitful. Nonetheless, we still lack any real understanding of the mechanisms behind the development of an ectopic pregnancy and assessing the vaginal microbiome in women with a PUL is an avenue that may provide new, useful information on not only the differences between those that have a favourable versus adverse early pregnancy outcome, but the mechanistic pathways that underlie their aetiology. This could not only have diagnostic value but possibly provide therapeutic targets, allowing improved monitoring of the disease process and patient response to treatment.

Justification for project

Existing biomarkers

Although specialised EPUs provide high quality care to women in the early pregnancy period, there remains a great deal of confusion over the management of women with a PUL. This leads to inconsistent care and concerns regarding health inequality. The work presented in this thesis aims to address this. Firstly, I will comprehensively review the current evidence-base via a systematic review and meta-analysis. I will then provide evidence from large, prospective clinical implementation studies on the most effective diagnostic support tools currently available, their test performance and most importantly their clinical safety profiles.

Novel biomarkers

Although much work has been carried out in the field of novel serum biomarker discovery to aid the management of women with a PUL, there has been no examination of the vaginal microbiome in this cohort of patients. The work presented in this thesis aims to address this, focussing on differences in the vaginal microbiota in women who have a healthy pregnancy versus those with an adverse outcome such as ectopic pregnancy.

Project aims

Existing biomarkers

1. Evaluate existing biomarkers used in the management of pregnancy of unknown location (PUL) to develop management protocols that have sufficient diagnostic accuracy to safely rationalize management in clinical practice

Novel biomarkers

2. Assess the vaginal microbiota in women with a PUL and compare results between those that have an adverse pregnancy outcome and those that go onto have a viable intrauterine pregnancy

Project hypotheses

1. Logistic regression models have the best test performance in triaging women with a PUL and predicting a final outcome of ectopic pregnancy

2. The clinical implementation of diagnostic support tools (such as logistic regression models) is feasible and safe

3. There will be a greater prevalence of vaginal dysbiosis in women with a PUL who have an adverse pregnancy outcome compared to those that go onto have a viable intrauterine pregnancy

METHODS

Studies

Figure 11 outlines the five studies that address the aims and objectives of this thesis. Each study uses data from the ABPEP study cohort.

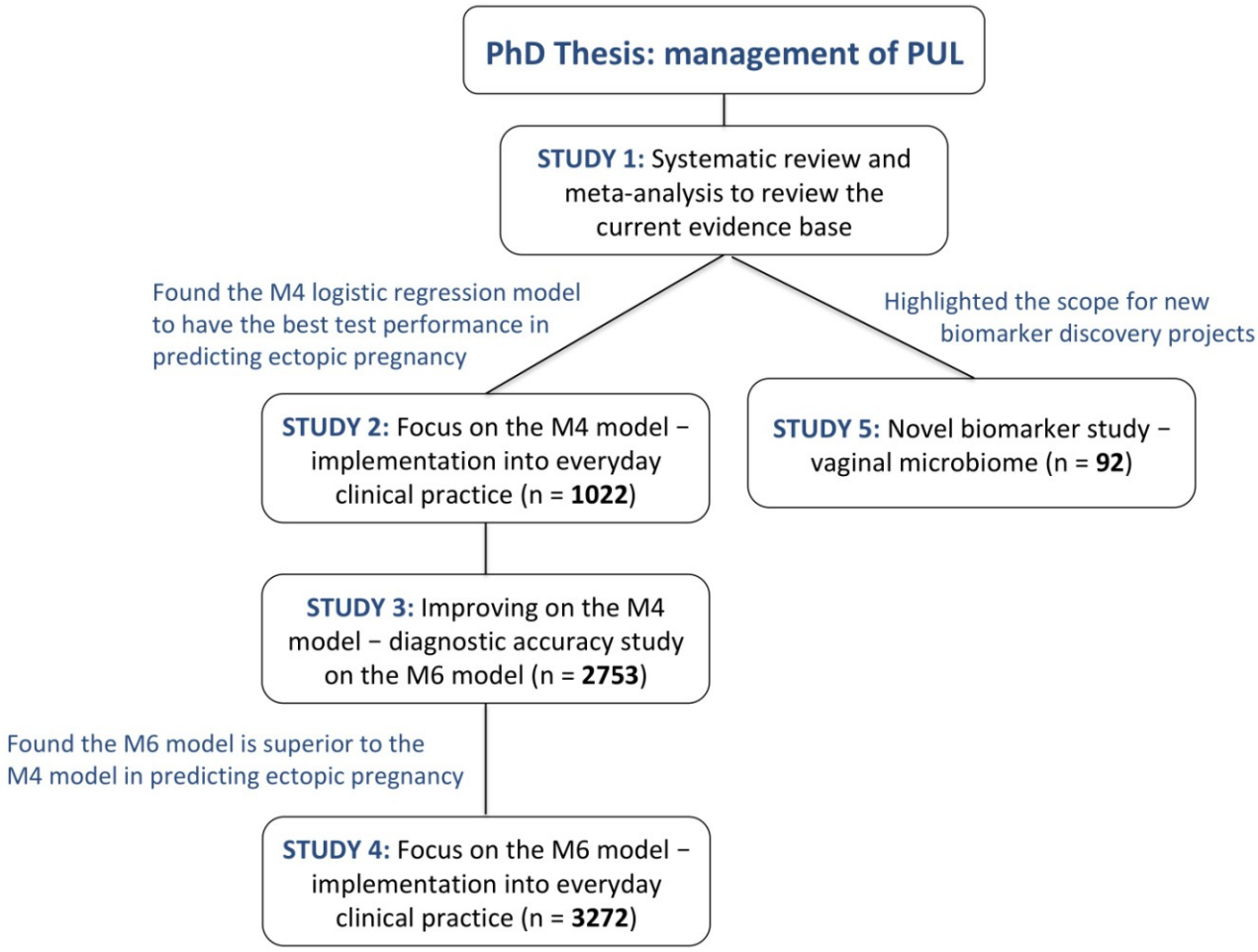


Figure 11: Thesis outline and rationale

This section outlines the overall methodology of the studies within this thesis including patient recruitment, data collection, procedures for bio-fluid sample collection and the reported outcome measures. The specific methodologies relating to individual analyses undertaken are described within each separate study.

Study design and ethical approval

Study 1: this was a systematic review and meta-analysis and therefore did not require ethical approval

Studies 2-4: these were studies carried out from prospectively collected data on women attending an EPU at a participating centre and were classified as a PUL. After discussion with the North of Scotland Research Ethics Committee, it was felt that written consent was not required from participants for these studies. They were therefore registered as audits at each participating centre.

Study 5: Ethical approval was provided by North of Scotland Research Ethics Committee (REC reference no. 14/NS/1078) and all participants provided informed written consent.

Inclusion criteria

Women classified as a PUL after attending for an ultrasound scan in the first trimester of pregnancy and:

- Under 14 weeks gestation
- Aged between 18-50 years old

Exclusion criteria

- Any patient under 18 years or over 50 years of age
- Any woman that was seen in a participating EPU and not classified as a PUL at her initial ultrasound scan
- Any patient unsuitable for outpatient management with serial serum hCG levels (which included haemodynamically unstable women)
- Anyone with a final diagnosis of a molar pregnancy (as the serum hCG levels in these patients behave in a way that is not suitable for interpretation by any management protocol used for PUL)
- Women whom in the opinion of the researcher, by virtue of language or learning impairment, would be unable to give fully informed consent to the studies

Withdrawal

Participants could withdraw from study 5 at any stage and all data related to their participation could be destroyed at their request.

Data collection

As part of their routine clinical care in early pregnancy, patients complete a clerking questionnaire. This includes detailed data on patient demographics, previous past medical, surgical, obstetric and gynaecology history (see appendix 1). We also collected details on their presenting symptoms:

Vaginal bleeding

The gold standard for quantifying the extent of vaginal bleeding experienced is by weighing tampons/sanitary towels. This, in reality however, is impractical. The pictorial blood loss assessment chart (PBAC) acts as the next best alternative (figure 12), scoring the amount of vaginal bleeding experienced from 0 to 4. It has been validated as a reported symptom score with comparable test performance to more objective measurement techniques[177, 178]. Furthermore, it is commonly used in the field of early pregnancy[64, 179]. We therefore used this to quantify the amount of vaginal bleeding experienced by patients when they first attended an EPU.

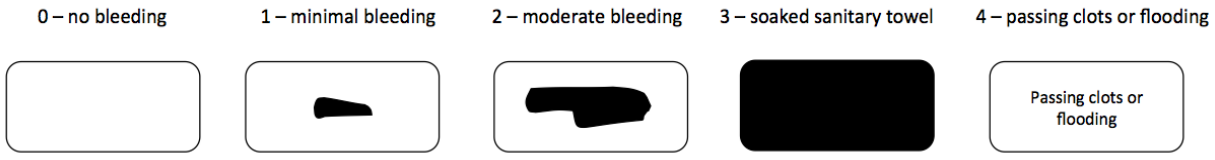


Figure 12: The pictorial blood loss assessment chart (PBAC) used to quantify the vaginal bleeding experienced

Pelvic pain

Quantifying pain is difficult as pain perception is highly subjective. There is evidence to suggest, however, that pain is best quantified on a continuous rather than categorical scale[180]. A Visual Analogue Scale (VAS)[180] is an appropriate representation of that and was used in this study so women could describe the pelvic pain they had experienced when seen at their initial visit (see figure 13). A VAS is usually a horizontal line that is 10cm in length and bookmarked by opposite descriptors at each end (i.e. 'no pain' and 'worst pain'). The distance from the left side of the line to where the patient had marked it was then measured with a ruler so an objective measurement could be derived. Participants

were asked to mark their pain score on the VAS on the day of their attendance and for the worst pain they had experienced (appendix 1). A higher score indicated greater pain intensity. The VAS was validated in a cohort of post-operative patients who had been asked to grade their post-operative pain as none, mild, moderate or severe. This study recommended the following correlating measurements on the VAS: no pain (0-4mm), mild pain (5–44 mm), moderate pain (45–74 mm), and severe pain (75–100 mm)[180].



Figure 13: The Visual Analogue Scale (VAS) used to quantify the patients' pain score

Ultrasonography protocol (studies 2-5)

Women attend an EPU up to 14 weeks' gestation and are either asymptomatic and undergo ultrasonography for reassurance or present with symptoms such as pain, bleeding and/or hyperemesis. All women in this study had a TVS as part of routine clinical practice within the EPU at their participating centre and were classified as a PUL using definitions from an international consensus paper[13]. Routine measurements were taken at the initial TVS, including the measurement of:

- endometrial thickness (figure 14)
- any fluid filled structure seen within the uterus (figure 15)
- the ovaries (figure 16)
- any adnexal masses
- any free fluid within the pelvis (figure 17)

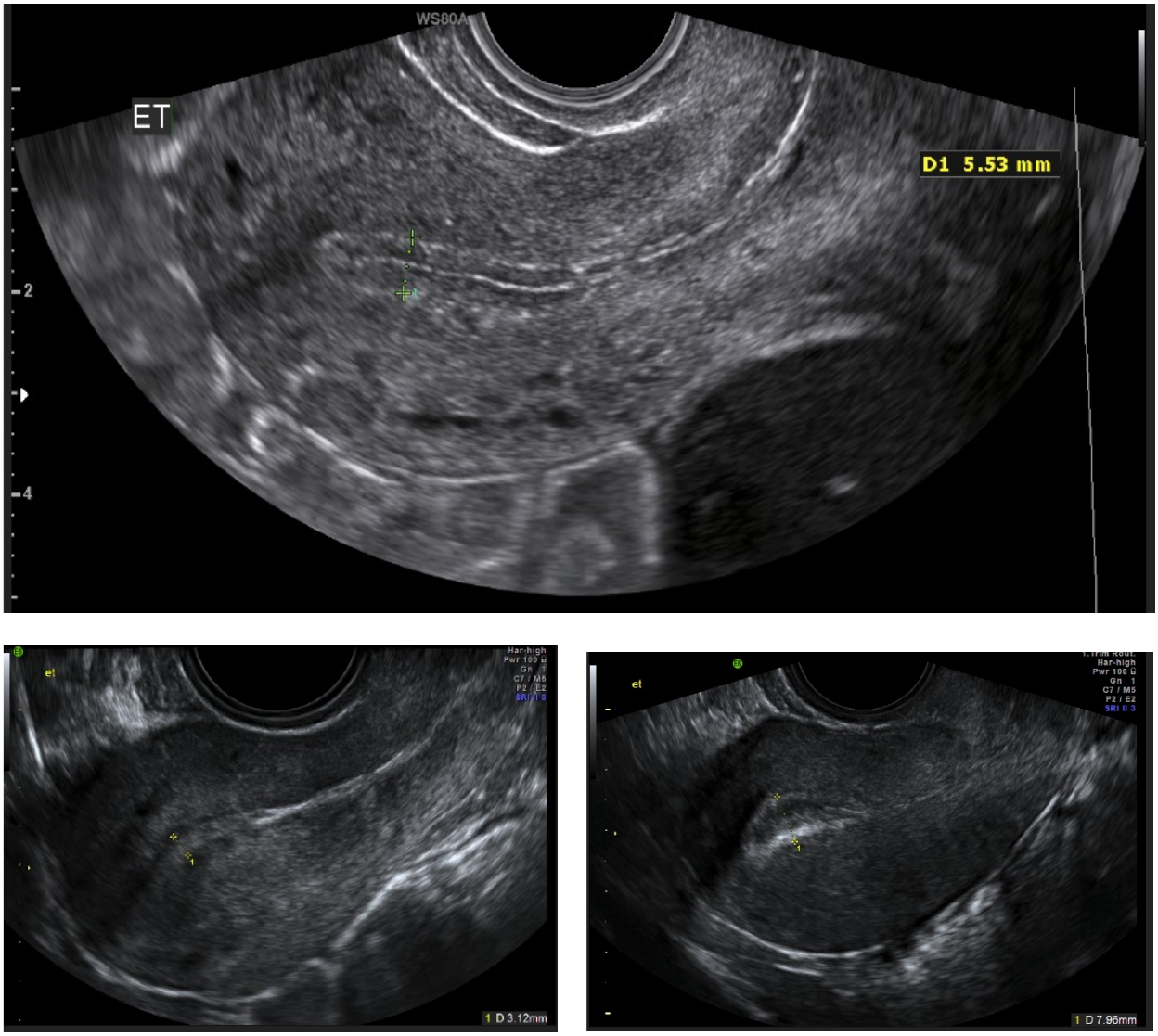


Figure 14: examples of endometrial thickness measurement. The measurement was taken perpendicular to the long line of the uterus.

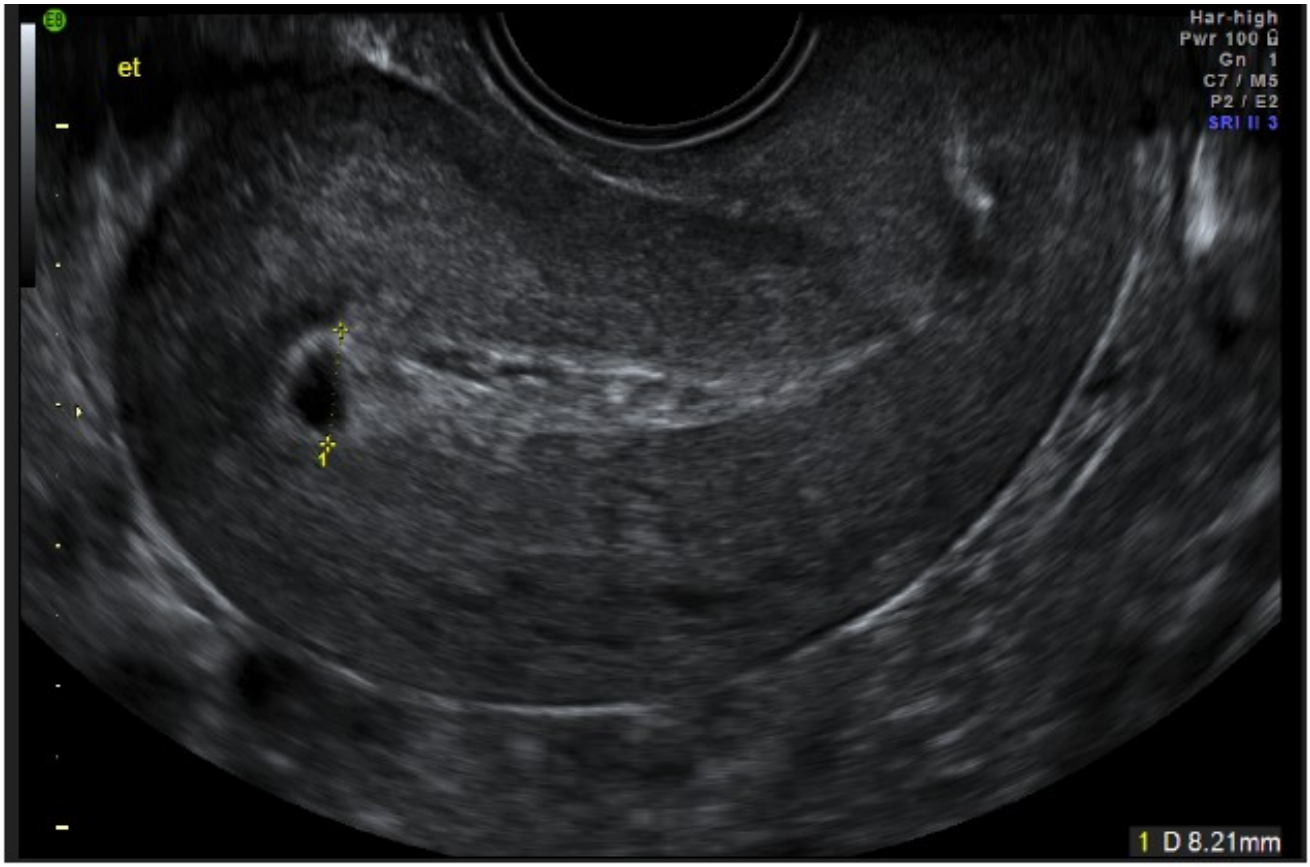


Figure 15: example of a fluid filled structure seen within the uterus. et = endometrial thickness

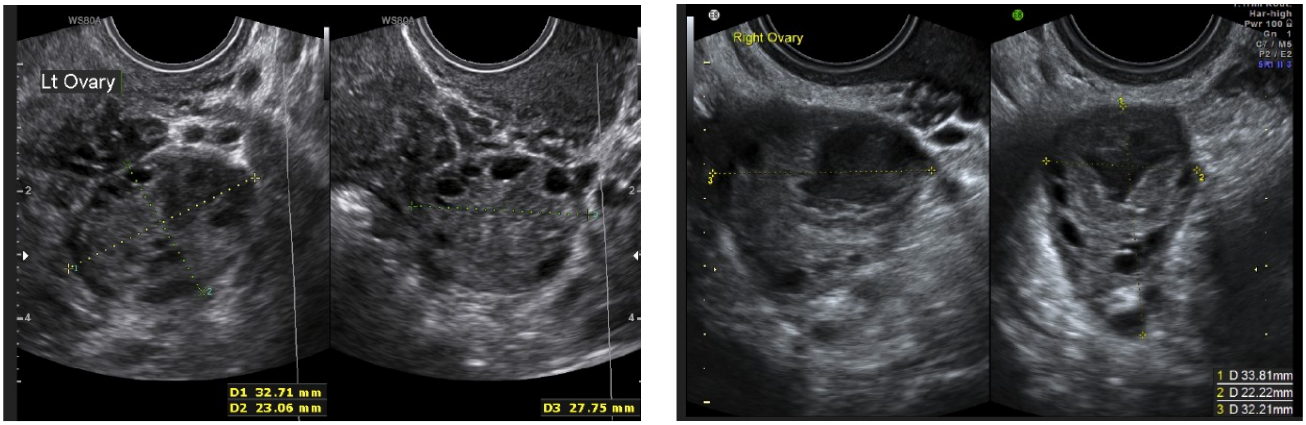


Figure 16: examples of measurement of the ovaries. Note, the ovary in the right image contains a corpus luteum

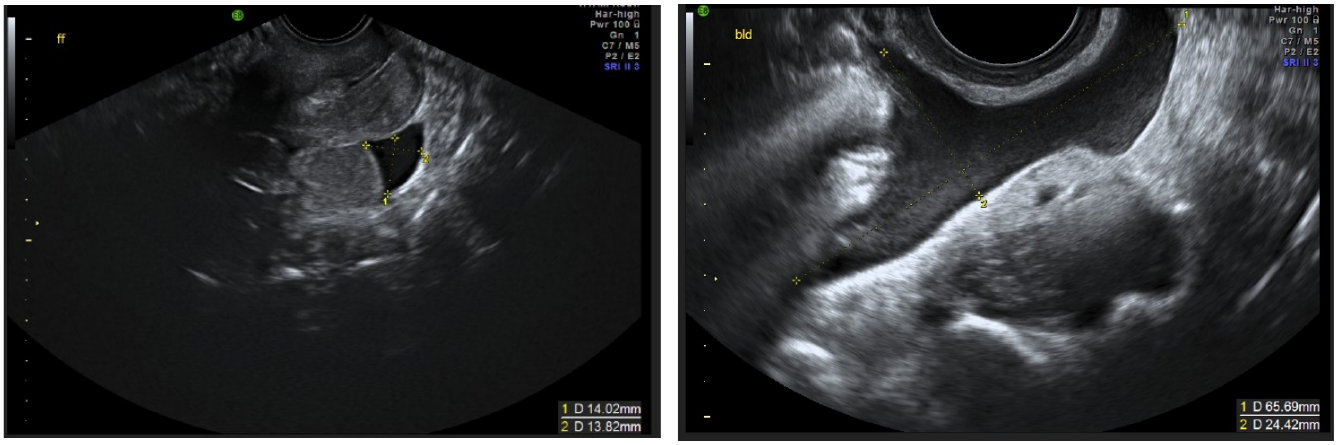


Figure 17: examples of free fluid seen in the pelvis (specifically, the Pouch of Douglas). Note, the free fluid in the left image is physiological whereas that seen in the right image is excessive and the increased echogenicity is likely to represent blood.

All measurements were recorded in millimetres (mm) using a transvaginal probe and routine 2D B-mode settings. Follow-up TVS were performed as clinically indicated and until the final pregnancy outcome could be determined. This included:

(v) Intrauterine pregnancy (IUP – either viable (VIUP) or non-viable (NVIUP))

(vi) Failed PUL (FPUL)

(vii) Ectopic pregnancy (EP)

(viii) Persistent PUL (PPUL)

An IUP was confirmed when an intra-uterine gestation sac (figure 18) with or without a yolk sac (figure 19) was seen within the endometrial cavity.

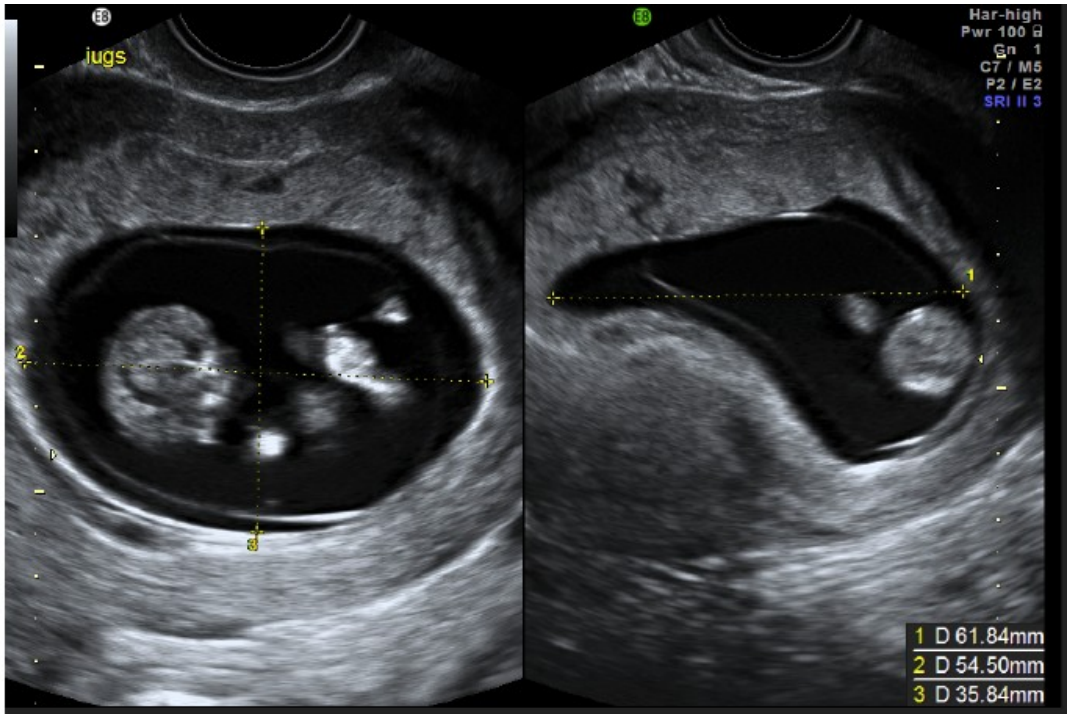


Figure 18: example of an intra-uterine gestation sac (IUGS) and measurement of the mean sac diameter taken in three orthogonal planes

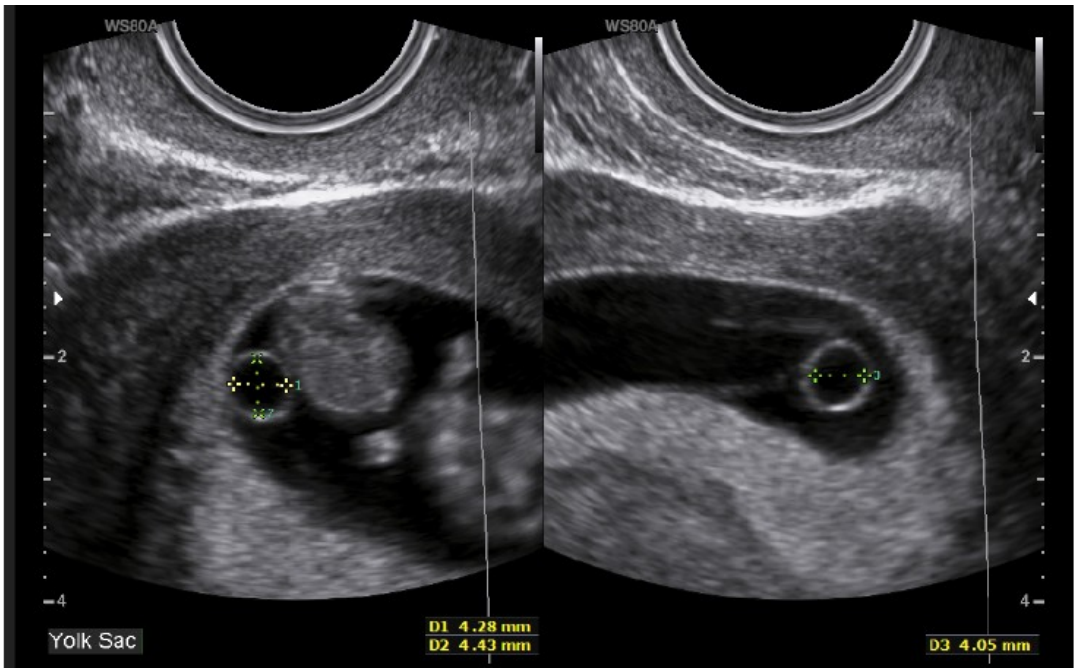


Figure 19: example a yolk sac and measurements taken in three orthogonal planes

A VIUP was defined as the presence of a fetal pole (figure 20) with visible fetal cardiac activity. A NVIUP was defined as the presence of a fetal pole measuring ≥7mm with no evidence of fetal cardiac activity, an empty gestation sac with a mean sac diameter measuring ≥25mm or if any other national criteria set for the diagnosis of miscarriage were met[7, 16].

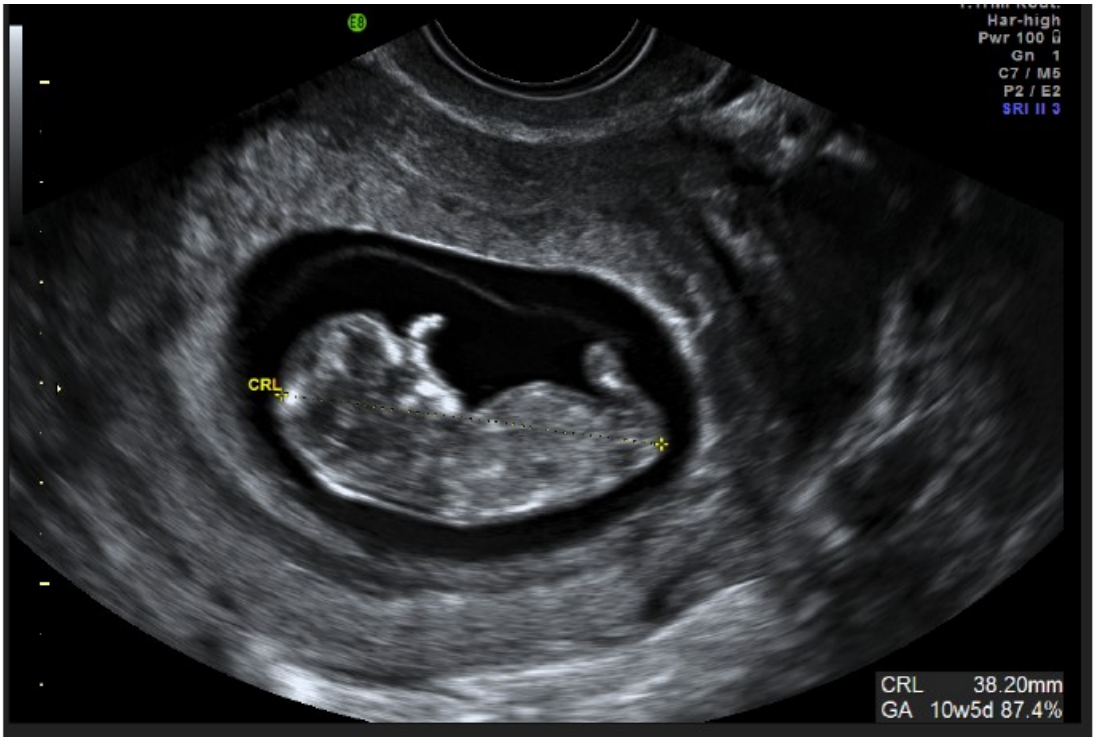


Figure 20: example of a fetal pole and CRL (crown rump length) measurement

A FPUL was diagnosed when the pregnancy location was never clearly determined on TVS but serial serum hCG measurements declined by ≥50% when repeated at 48 hours and a urine pregnancy test was negative two weeks from the initial consultation.

An EP was diagnosed by either the positive visualization of an adnexal mass on TVS (figure 21) or confirmation of chorionic villi on histopathological examination of the removed specimen after a salpingectomy.

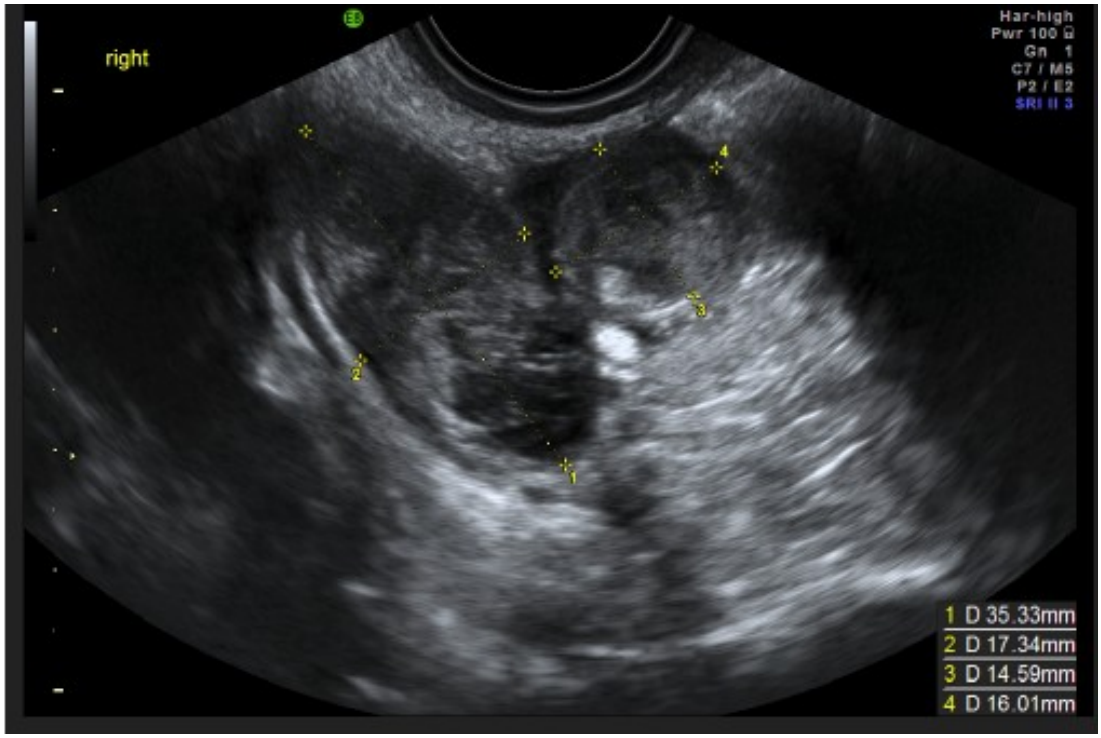


Figure 21: example an adnexal mass, seen on the right side of the image and medial to the right ovary

A PPUL was diagnosed when the pregnancy location was never clearly determined on TVS but >3 serial serum hCG levels, each taken 48 hours apart, increased or decreased by <15% (i.e. 'suboptimal' rise or fall in serum hCG levels).

Basic ultrasound principles

All ultrasound scans were performed by trained sonographers, gynaecologists or nurse specialists. The equipment used was compliant with national standards and had at a minimum the following capabilities:

- real-time, grey-scale, two-dimensional (2D) ultrasound
- transabdominal and transvaginal ultrasound transducers
- adjustable acoustic power output controls with output display standards
- freeze frame and zoom capabilities
- electronic calipers
- capacity to print/store images
- regular servicing and maintenance

Sample collection and analysis

Once cases were clinically classified as a PUL, women had blood tests taken as part of their routine care to measure serum progesterone and hCG levels at the initial visit and then hCG again 48 hours later. All samples were processed and tested in line with standard medical laboratory practice including daily QC analysis of samples to assess assay accuracy, precision and further checks as specified by Clinical Pathology Accreditation (CPA) quality standards. Additional samples taken exclusively for study 5 are described within the individual study.

Study outcome measures and definitions

These are described within each individual study.

Statistical analysis

Each study within this thesis has been conducted using a different type of statistical analysis and will be described in detail within each study.

Contribution of various authors to each study

Study 1

For the meta-analysis and systematic review, I Shabnam Bobdiwala (SB) was involved in the design of the study, screened all titles, screened all abstracts, reviewed full articles, performed data extraction for the purposes of meta-analyses, interpreted the data, created an initial draft of the article submitted for peer-reviewed publication and revised all subsequent versions to create the final published manuscript. I was therefore first author on the published paper resulting from this work. Mr Srdjan Saso (SS), clinical lecturer at Imperial College London, was involved in the design of the study, reviewed full articles, performed data extraction for the purposes of meta-analyses and was involved in interpreting the data. Professor Jan Verbakel (JV), a clinical researcher based at the Nuffield Department of Primary Care Health Sciences, University of Oxford and the Department of Development and Regeneration, KU Leuven, Belgium, performed all analyses and was involved in interpreting the data. Professor Tom

Bourne (TB), Professor of Practice at Imperial College London and Visiting Professor at KU Leuven, Belgium, was a third reviewer in cases where consensus could not be reached by SB and SS, interpreted the data and was the lead supervisor for this study.

Study 2

I (SB) acquired patient data, was involved in data analysis, interpreted the results, created an initial draft of the article for peer-reviewed publication and revised all subsequent versions to create the final published manuscript. I was therefore first author on the published paper resulting from this work. Ben Van Calster (BVC), professor of statistics at the Department of Development and Regeneration, KU Leuven, Belgium and Department of Biomedical Data Sciences, Leiden University Medical Centre, the Netherlands, participated in the conception and design of this study, performed the statistical analysis and interpreted the results. Miss Sharmitha Guha (SG), consultant obstetrician and gynaecologist at Chelsea and Westminster Hospital, London, UK, was involved in the conception and design of this study, acquiring patient data and interpreting the results. She was therefore the joint first author on this study. Miss Nicola Mitchell-Jones (NMJ), specialist trainee at Chelsea and Westminster Hospital, London, UK, Mr Francis Ayim (FA), consultant obstetrician and gynaecologist at Hillingdon Hospital, London, UK and Miss Hannah Mitchell (HM), specialist trainee at Hillingdon Hospital, London, UK all acquired patient data. TB participated in the conception and design of this study, interpreted the data and was the lead supervisor for this study.

Study 3

I (SB) was involved in data analysis, interpreting the results, creating an initial draft of the article for peer-reviewed publication and revising all subsequent versions to create the final published manuscript. I was therefore joint first author on the published paper resulting from this work. BVC participated in the conception and design of this study, created the M6 model, performed the statistical analysis and interpreted the results. He was therefore the other joint first author on the published paper resulting from this work. SG, Miss Jessica Farren (JF), specialist trainee at St Marys' Hospital, London, UK, Miss Emma Kirk (EK), consultant obstetrician and gynaecologist at Royal Free NHS

Foundation Trust, UK and George Condous (GC), professor of gynaecology at the University of Sydney, Australia, all acquired patient data. TB participated in the conception and design of this study, interpreted the data and was the lead supervisor for this study.

Study 4

I (SB) was involved in the conception and design of this study, coordinated all participating centres in this multi-centre piece of work, acquired patient data, interpreted the results, created an initial draft of the article for publication and revised all subsequent versions to create the final published manuscript. I was therefore first author on the published paper resulting from this work. SG, NMJ, FA, JF, EK, Mr Velluppillai Vathanan (VV), consultant obstetrician and gynaecologist at Wexham Park Hospital, UK, Miss Baljinder Chohan (BC), specialist trainee at Wexham Park Hospital, UK and Mr Osama Abughazza (OA), consultant obstetrician and gynaecologist at Royal Surrey Hospital, UK all acquired patient data. BVC performed the statistical analysis and interpreted the results. TB participated in the conception and design of this study, interpreted the data and was the lead supervisor for this study.

Study 5

I (SB) was involved in the conception and design of this study, patient recruitment, sample collection, performing all experiments, data collection, data analyses and interpretation, generating figures and tables, creating the initial draft of the manuscript written for peer-reviewed publication and revising all subsequent versions. I am therefore first author on the paper due for submission for publication and resulting from this work. Dr David MacIntyre (DAM), reader in reproductive systems medicine at Imperial College London, was involved in the conception and design of this study, data analysis and interpretation, generating figures and tables and reviewing the draft of the manuscript written for peer-reviewed publication. Professor Phillip Bennett (PB), professor of obstetrics and gynaecology at Imperial College London, participated in the conception and design of this study, interpreted the data and was a secondary supervisor for this study. TB participated in the conception and design of this study, interpreted the data and was the lead supervisor for this study.

STUDY 1

DIAGNOSTIC PROTOCOLS FOR THE MANAGEMENT OF PREGNANCIES OF UNKNOWN LOCATION:

A SYSTEMATIC REVIEW AND META-ANALYSIS

Chapter Abstract

This chapter refers to original work published in Bobdiwala et al[181].

Introduction

The clinical care women with a pregnancy of unknown location receive is varied and there is no clear international consensus on which diagnostic protocol should be used in their management. I aimed to perform a systematic review and meta-analysis to summarize the current evidence base.

Methods

I searched MEDLINE, COCHRANE and DARE databases from 01/01/1984 to 31/01/2017. The primary outcome was the accurate prediction of a final outcome of ectopic pregnancy (high-risk) as opposed to either a failed PUL or intrauterine pregnancy (low-risk). We included forty-three studies and used QUADAS-2 criteria to evaluate the risk of bias. A novel linear mixed effects model was used to perform a meta-analysis and summary receiver operating characteristic (SROC) curves were constructed for the thresholds of interest.

Results

Most studies showed a high level of differential verification bias. We were able to perform meta-analyses on the accuracy of risk prediction on (i) single hCG cut-off levels, (ii) hCG ratio (hCG at 48 hours / initial hCG), (iii) single progesterone cut-off levels and (iv) the 'M4 model' (a logistic regression risk prediction model based on the initial hCG and hCG ratio). For predicting a final outcome of ectopic pregnancy in the PUL population, the AUCs (95% CI) for the four management protocols were: (i) 0.42 (0.00-0.99), (ii) 0.69 (0.57-0.78), (iii) 0.69 (0.54-0.81) and (iv) 0.87 (0.83-0.91), respectively.

Conclusions

The M4 model was the best available method for predicting a final outcome of ectopic pregnancy in the PUL population.

Introduction

Numerous protocols for the management of women classified as a PUL exist and the majority are based on serum hCG levels[21, 22, 40, 41, 47, 49, 50, 182-199]. Published studies include the evaluation of single hCG levels and the hCG ratio (hCG at 48 hours / initial hCG). Logistic regression models such as the 'M4 model'[21, 199] also rely on the interpretation of hCG. Other serum biomarkers such as progesterone[38, 39, 119, 200], novel serum biomarkers[104, 201, 202], clinical markers such as the amount of vaginal bleeding seen and ultrasound-based markers such as the endometrial thickness (ET)[203-205] have also been assessed. There is, however, no clear international consensus on which management protocol is best and therefore the management of women with a PUL is often suboptimal and inconsistent.

The primary aim of this study was to assess the test performance of management protocols used in the clinical care of women with a PUL by performing a comprehensive systematic review and meta-analysis of the current evidence base.

Methods

Study protocol and registration

We created a study protocol that specified in advance the literature search terms and criteria, inclusion and exclusion criteria, study outcomes to be measured, tools for quality assessment and the plan for data synthesis. This was registered with PROSPERO: (http://www.crd.york.ac.uk/PROSPERO, registration number CRD42017058776). A PRISMA checklist was used to perform the systematic review.

Inclusion criteria

- All studies on women who were classified as a PUL on ultrasound at their first visit
- Studies written in the English language
- Included study designs: randomised control trials, case-control studies, cohort studies and observational studies
- Both prospective and retrospective studies

Exclusion criteria

- Studies assessing pregnancies that were not a PUL on the initial ultrasound scan (i.e. an intrauterine or ectopic pregnancy)
- Studies not written in the English language
- Studies not in humans
- Excluded study designs: reviews, case reports and letters

Literature search criteria

I used predefined search strings and MeSH terms to undertake a literature search on MEDLINE, COCHRANE and DARE databases from 01/01/1984 to 31/01/2017. The free text search strings used were: “ectopic pregnancy OR pregnancy of unknown location”. The MeSH terms used were: Pregnancy, Tubal[Majr] OR "Pregnancy, Ectopic"[Majr] OR ectopic pregnancy OR PUL OR pregnancy of unknown

location. A start date of 01/01/1984 was used as this was when transvaginal ultrasound began to be routinely used in the assessment of early pregnancy problems. I also assessed the references of eligible papers and other systematic reviews on early pregnancy problems to identify papers that were absent from the initial database searches.

Study selection

I was one of two reviewers evaluating potentially eligible papers in a three-step process:

[1] All study titles were screened

[2] If there was any doubt about eligibility, further screening included review of the study abstract. This was to ensure no papers were overlooked

[3] The full text of selected articles were reviewed

Data extraction

A second review author and I extracted data from the selected studies using a standardized data extraction spreadsheet I designed. This spreadsheet collected details on study characteristics (primary author, publication year, study groups), study methodology and type (prospective or retrospective), study design, recruitment period, recruitment location (an accident and emergency department or dedicated EPU), details on the definition used for a PUL, the index test (i.e. management protocol used) and the recorded final pregnancy outcome of the women with a PUL included in each study. Binary diagnostic accuracy data was then extracted and we constructed 2x2 tables from all included studies for all thresholds reported in those studies. If applicable, 2x2 tables were also constructed for each participating centre in multi-centre studies.

Assessment of the risk of bias

I and the other review author assessed the risk of bias in the included studies by using QUADAS-2 criteria[206]. A third review author was used when the first two reviewers could not reach a consensus.

Outcomes of interest

The primary outcome assessed was correct diagnosis of the final pregnancy outcome. This was either: [1] ectopic pregnancy (EP), [2] failed PUL (FPUL) or [3] an intrauterine pregnancy (IUP), either viable (VIUP) or non-viable (NVIUP). Persistent PUL (PPUL) was not assessed as a final outcome in any of the selected studies and was therefore not included in our analysis.

Statistical Analysis

All analyses were performed using R version 3.2.4 and RevMan version 5.3. Paired forest plots of sensitivity and specificity (with 95% confidence intervals) were created for the three main outcomes of interest: EP, IUP and FPUL for each management protocol.

Several studies reported more than one threshold (e.g. a single hCG cut-off of 1000, 1500 and 2000IU/L). We therefore applied a novel parametric approach which enabled several thresholds to be analysed per study[207]. Assuming a logistic distribution, it estimates the distribution parameters in both diseased and non-diseased groups, applying a linear mixed effects model to the transformed data. This model allows the assessment of heterogeneity between studies (random effect). We used the inverse variance as prior weights of the studies scaled to mean one. Once the model parameters were estimated, we determined the underlying distribution functions. The pooled sensitivity and specificity values at each relevant threshold could then be extracted. Summary receiver operating characteristic (SROC) curves to compare the different management protocols were derived with corresponding 95% confidence regions. We then calculated the Area under the Receiver-Operating-Characteristic curves (AUCs) using numerical integration over the areas in the SROC curves. Likelihood ratios were not calculated as this is not possible with the type of parametric approach used in this study.

Results

Study selection

MEDLINE, COCHRANE and DARE electronic database searches retrieved 8847 papers after the removal of duplicates. The flow of study selection is summarized in figure 22. Ninety-nine papers were selected for full text review. Forty-three papers met the criteria for qualitative synthesis (systematic review). A total of 27,493 PUL were recruited and 23,802 PUL analysed in these studies. Thirty-seven papers met the criteria for quantitative synthesis (meta-analysis). Of the management protocols assessed, fifteen studies evaluated the predictive value of single hCG cut-off levels, 8 studies the hCG ratio (hCG 48 hours / hCG at 0 hours), 12 studies progesterone cut-off levels, 2 studies used a combination of hCG and progesterone levels, 16 studies assessed logistic regression (LR) models which used a range of predictive variables, 5 studies assessed the value of endometrial thickness and 4 studies evaluated novel biomarkers such as CA125, activin A and inhibin A. Of these, 30 were prospective and the remaining 13 retrospective. Of the 37 studies included in the meta-analysis, 15 studies evaluated single hCG measurements, 8 studies the hCG ratio, 13 studies progesterone cut-offs and 8 studies the M4 model. Meta-analysis could not be performed on additional LR models as there were not enough studies on any other model.

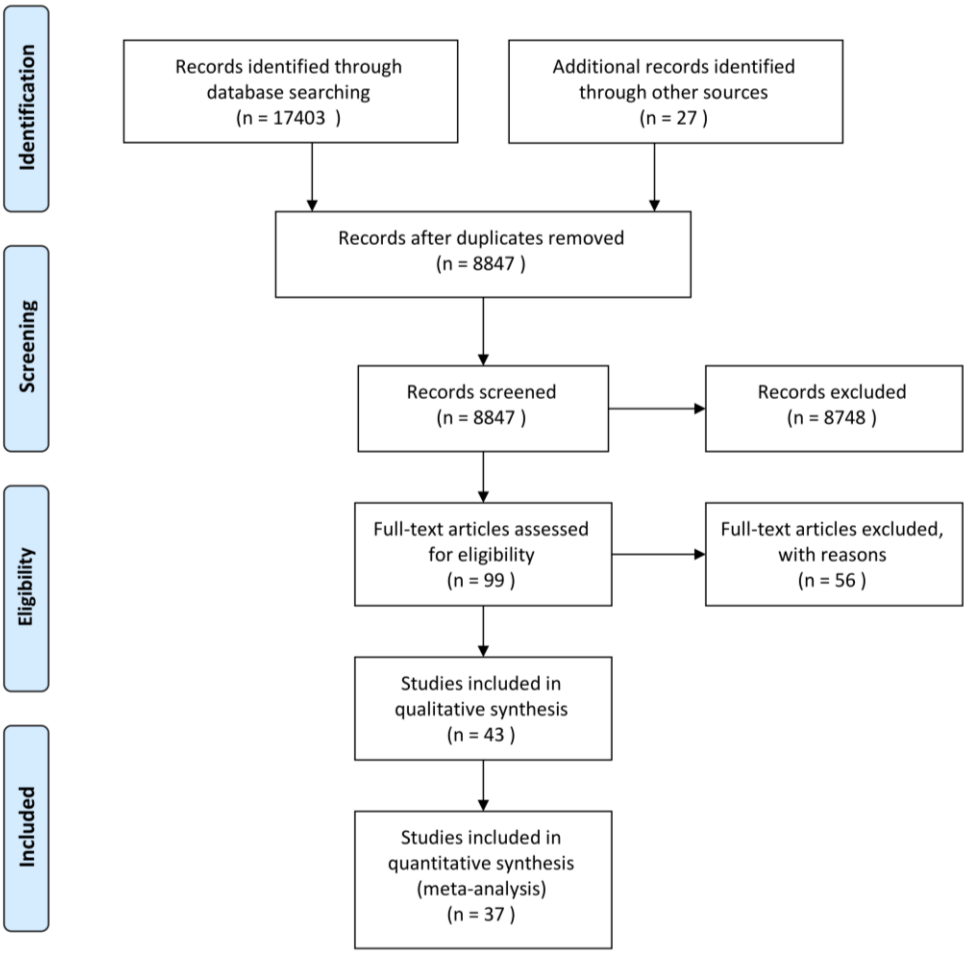


Figure 22: PRISMA flowchart of the literature search and article selection for a systematic review and meta-analysis on the test performance of protocols used in the management of women with a pregnancy of unknown location (PUL). Reproduced from supplementary material published by Bobdiwala et al[181].

Characteristics of included studies

The characteristics of the 43 studies selected for final analysis is summarised in appendix 2. As the term PUL is not a final diagnosis, there was a great deal of heterogeneity amongst studies in both the definition of PUL and the potential final pregnancy outcomes. 'PUL' was most commonly defined as a TVS that failed to find evidence of an IUP or EP in a woman presenting with a positive urine pregnancy test. Some studies extended this to include no evidence of RPOC (retained products of conception) on TVS.

Not all three (IUP, EP, FPUL) potential final outcomes were assessed in each study. A final outcome of PPUL was not included in any of the studies so we removed this from our data analyses.

An 'IUP' was usually defined as the visualization of an intrauterine gestation sac on ultrasound +/- confirmation of chorionic villi on histopathological examination of tissue removed during uterine curettage. Some studies differentiated between VIUP and NVIUP, but most classified them under the umbrella term of IUP.

An 'EP' was most commonly defined as an extra-uterine/ adnexal mass seen on TVS or trophoblast tissue confirmed on histopathological examination of tissue removed after a salpingectomy.

Prior to 1999, 'failed PUL' was not commonly used but met similar definitions to other terms used in earlier papers such as 'trophoblast regression' (TR). We therefore, for the purposes of this analysis, analysed TR as FPUL. 'FPUL' was usually defined as persistent negative ultrasound findings and the spontaneous resolution of a pregnancy with a corresponding decrease in serum hCG levels (usually to <5 IU/L). Some papers also included an initial progesterone level of <20mmol/l in the definition of FPUL.

Most studies defined specific inclusion and exclusion criteria and there was also heterogeneity amongst these. This is summarised in appendix 3.

Risk of bias within studies

All studies included in a systematic review or meta-analysis carry a potential risk of bias. This was formally assessed in this study using QUADAS 2 criteria[206] (figure 23). There are four main categories where bias in diagnostic accuracy studies can occur:

[i] *Patient Selection:* i.e. if the samples are consecutive or random and if the research group have avoided inappropriate exclusions. Broadly speaking, retrospective studies are more likely to be at high-risk of patient selection bias compared to prospective studies.

[ii] *Index Test:* this refers to the management protocol being assessed. If the reference standard (see below) was known prior to applying the index test, then the study is at high-risk of bias in this category. Accordingly, retrospective studies are more likely to be at high-risk of index test bias compared to prospective studies.

[iii] *Reference Standard:* this refers to the gold standard test used to confirm the specific final pregnancy outcome. The reference standard item in this meta-analysis was always considered high-risk as the reference standard varied depending on the final pregnancy outcome (i.e. differential verification bias). As an example, the reference standard for confirming an IUP is a TVS demonstrating an intrauterine gestation sac with an embryo and a visible heart beat, whereas that for an EP could be confirmation of chorionic villi on histological examination of a fallopian tube post-salpingectomy. The only occasion where the reference standard risk was low was when all patients had the same reference test/ intervention, e.g. uterine curettage.

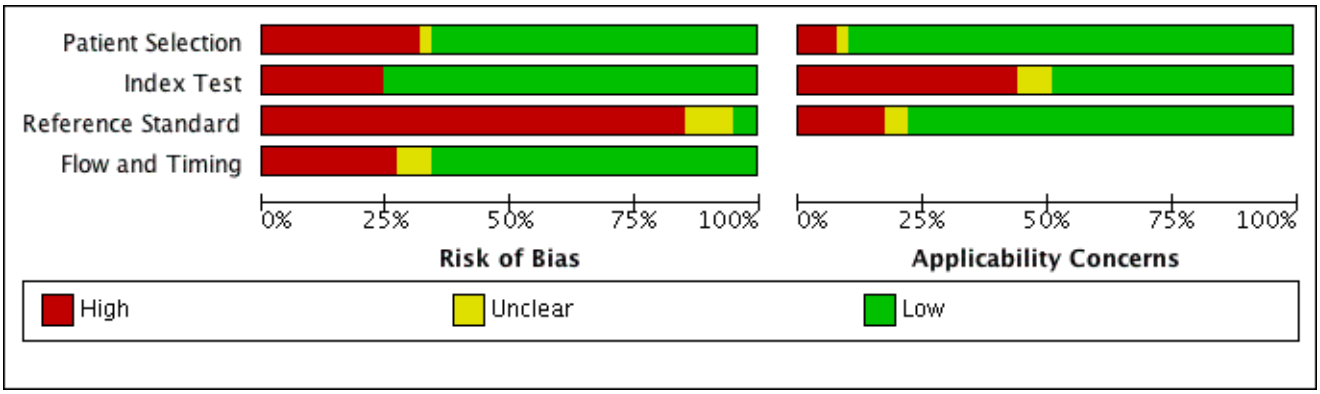
[iv] *Flow and Timing:* this includes an assessment of whether there was an appropriate interval between the index test and reference standard, whether all patients had the same reference standard applied and whether all patients were included in the final analyses.

There is also a concern about whether the findings from the studies included in the meta-analysis have applicability to other patient populations. This is also formally assessed by QUADAS-2 criteria via three main domains:

[i] *Patient Selection:* i.e. if the patients included in the study do not match those described in the review question, then there are 'high' concerns about applicability

[ii] *Index Test:* i.e. if the index test is insufficient to answer the review question, then there are high concerns about applicability

[iii] *Reference Standard:* i.e. if the target condition, as defined by the reference standard, does not match the review question, then there are high concerns about applicability



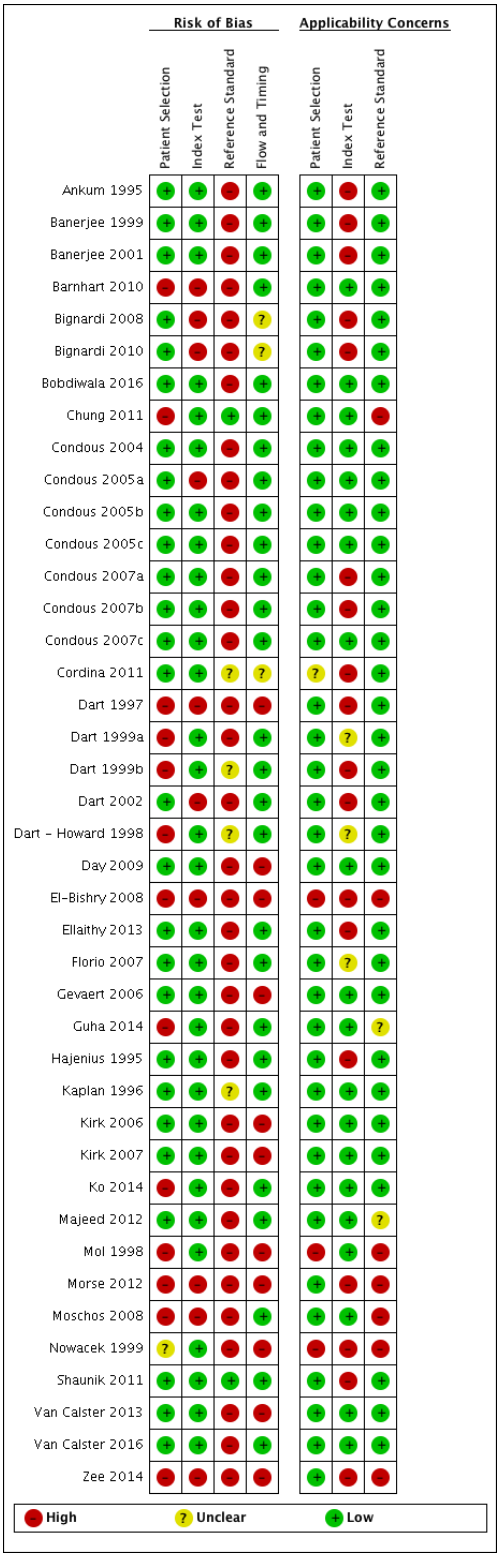


Figure 23: QUADAS 2 data for the included studies. Reproduced from supplementary material published by Bobdiwala et al[181].

The diagnostic performance of management protocols for PUL

The diagnostic performance of each management protocol for each of the three potential final outcomes (EP, IUP, FPUL) is summarised in appendix 4. The paired forest plots of sensitivity and specificity of each study for the three outcome categories are summarised in table 1.

PUL Outcome	Management Protocol	Threshold	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Summary AUC (for all thresholds for this protocol)
EP	**Single hCG**	1000 IU/L	**0.42** (0.28 - 0.57)	**0.52** (0.26 - 0.76)	**0.42** (0.00-0.99)
	hCG ratio	0.87	**0.55** (0.33 - 0.76)	**0.67** (0.63 - 0.70)	**0.69** (0.57-0.78)
		1.66	**0.52** (0.29 - 0.74)	**0.69** (0.65 - 0.73)	
	Progesterone	10 nmol/L	**0.66** (0.50 - 0.80)	**0.58** (0.45 - 0.70)	**0.69** (0.54-0.81)
	M4 model	5% risk of EP	**0.82** (0.79 - 0.85)	**0.80** (0.77 - 0.82)	**0.87** (0.83-0.91)
IUP	**Single hCG**	1000 IU/L	**0.91** (0.68 - 0.98)	**0.33** (0.19 - 0.50)	**0.71** (0.01-1.00)
	hCG ratio	1.66	**0.86** (0.57 - 0.97)	**0.97** (0.95 - 0.98)	**0.97** (0.87-0.99)
	Progesterone	10 nmol/L	**0.91** (0.82 - 0.96)	**0.59** (0.48 - 0.69)	**0.89** (0.76-0.96)
	M4 model	Risk EP <5% and risk IUP > risk FPUL	**0.59** (0.38 - 0.78)	**0.71** (0.48 - 0.87)	**0.86** (0.39-0.99)
FPUL	**Single hCG**	1000 IU/L	**0.51** (0.13 - 0.88)	**0.59** (0.40 - 0.76)	**0.59** (0.00-1.00)
	hCG ratio	0.87	**0.86** (0.84 - 0.88)	**0.97** (0.96 - 0.98)	**0.98** (0.86-1.00)
	Progesterone	10 nmol/L	**0.82** (0.77 - 0.85)	**0.79** (0.65 - 0.89)	**0.87** (0.76-0.94)
	M4 model	Risk EP <5% and risk FPUL > risk IUP	**0.71** (0.46 - 0.87)	**0.57** (0.30 - 0.81)	**0.84** (0.15-1.00)

Table 1: Pooled sensitivity and specificity data for the management protocols assessed in the systematic review and meta-analysis. The summary AUC column provides an overall AUC for all thresholds analysed within that management protocol, not just that which is specified in the 'threshold' column. PUL = pregnancy of unknown location, EP = ectopic pregnancy; IUP = intrauterine pregnancy, FPUL = failed pregnancy of unknown location, hCG = human chorionic gonadotrophin, CI = confidence interval, AUC = area under the curve. Reproduced from supplementary material published by Bobdiwala et al[181].

Meta-analysis: primary outcome

Predicting a final outcome of EP:

A final outcome of EP puts women classified as a PUL at the greatest risk of clinical harm. This is, therefore, the most important outcome to accurately predict or exclude. Summary analysis of all management protocols used to predict EP as a final outcome showed AUC of: 0.42 (95% CI 0.00-0.99) for single hCG cut-offs, 0.69 (95% CI 0.57-0.78) for the hCG ratio and 0.69 (95% CI 0.54-0.81) for progesterone cut-offs. The M4 model was associated with an AUC of 0.87 (95% CI 0.83-0.91). See figure 24.

Meta-analysis: secondary outcomes

Predicting a final outcome of FPUL:

Summary analysis of all management protocols used to predict FPUL as a final outcome showed AUC of: 0.59 (0.00-1.00) for single hCG cut-offs, 0.98 (0.86-1.00) for the hCG ratio, 0.87 (0.76-0.94) for progesterone cut-offs and 0.84 (0.15-1.00) for the M4 model. See figure 25.

Predicting a final outcome of IUP:

Summary analysis of all management protocols used to predict IUP as a final outcome showed AUC of: hCG 0.71 (0.01-1.00) for single hCG cut-offs, 0.97 (0.87-0.99) for the hCG ratio, 0.89 (0.76-0.96) for progesterone cut-offs and 0.86 (0.39-0.99) for the M4 model. See figure 26.

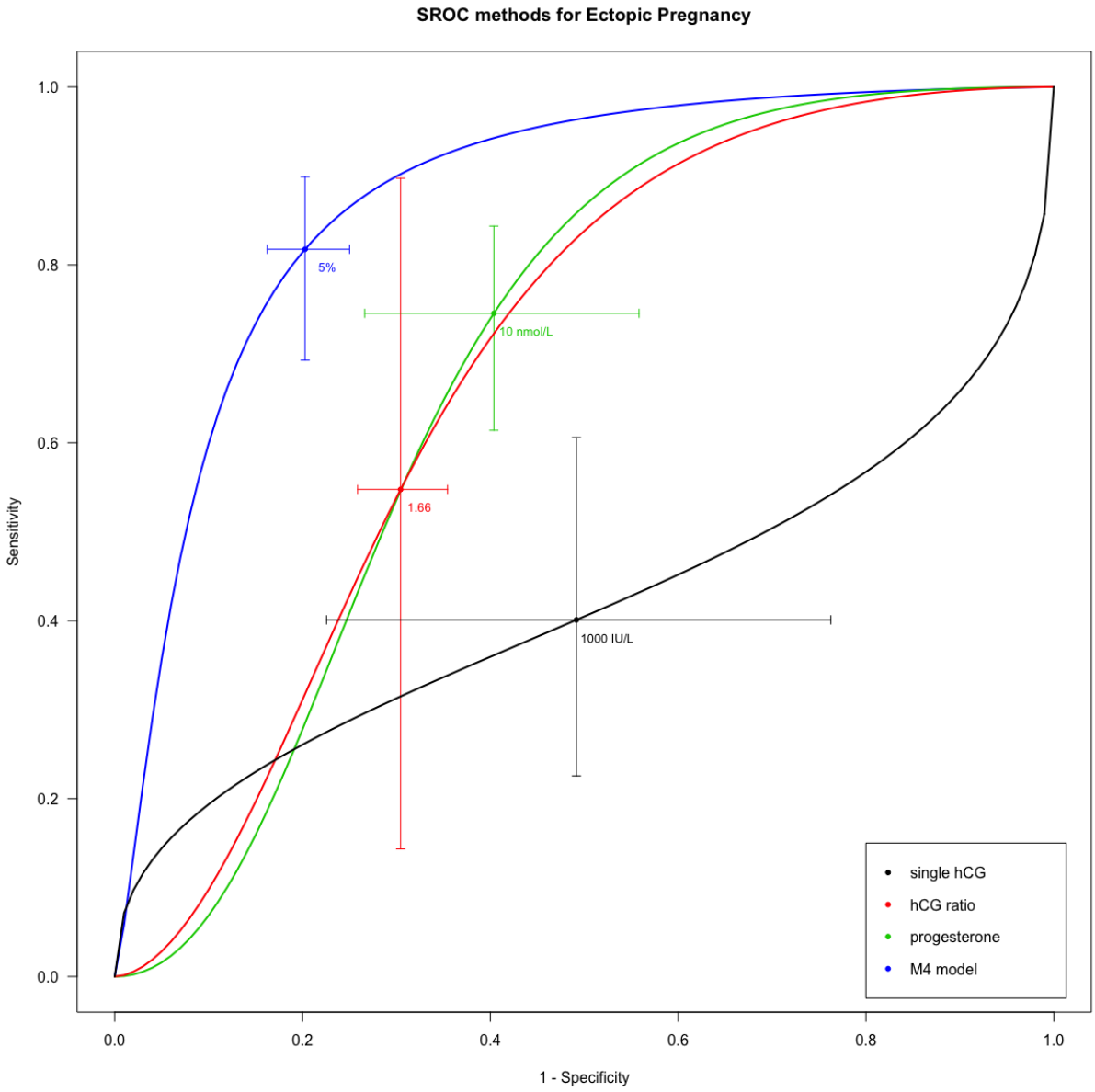


Figure 24: SROC (Summary Receiver Operating Curve) for the test performance of all management protocols predicting a final outcome of ectopic pregnancy (EP). The number on each curve refers to the summary threshold for each protocol used in the original papers.
hCG = human chorionic gonadotrophin.

Management Protocol	**AUC (95% CI) for EP prediction**
Single hCG	0.42 (0.00-0.99)
hCG ratio	0.69 (0.57-0.78)
Progesterone	0.69 (0.54-0.81)
M4 model	0.87 (0.83-0.91)

Reproduced from Bobdiwala et al[181].

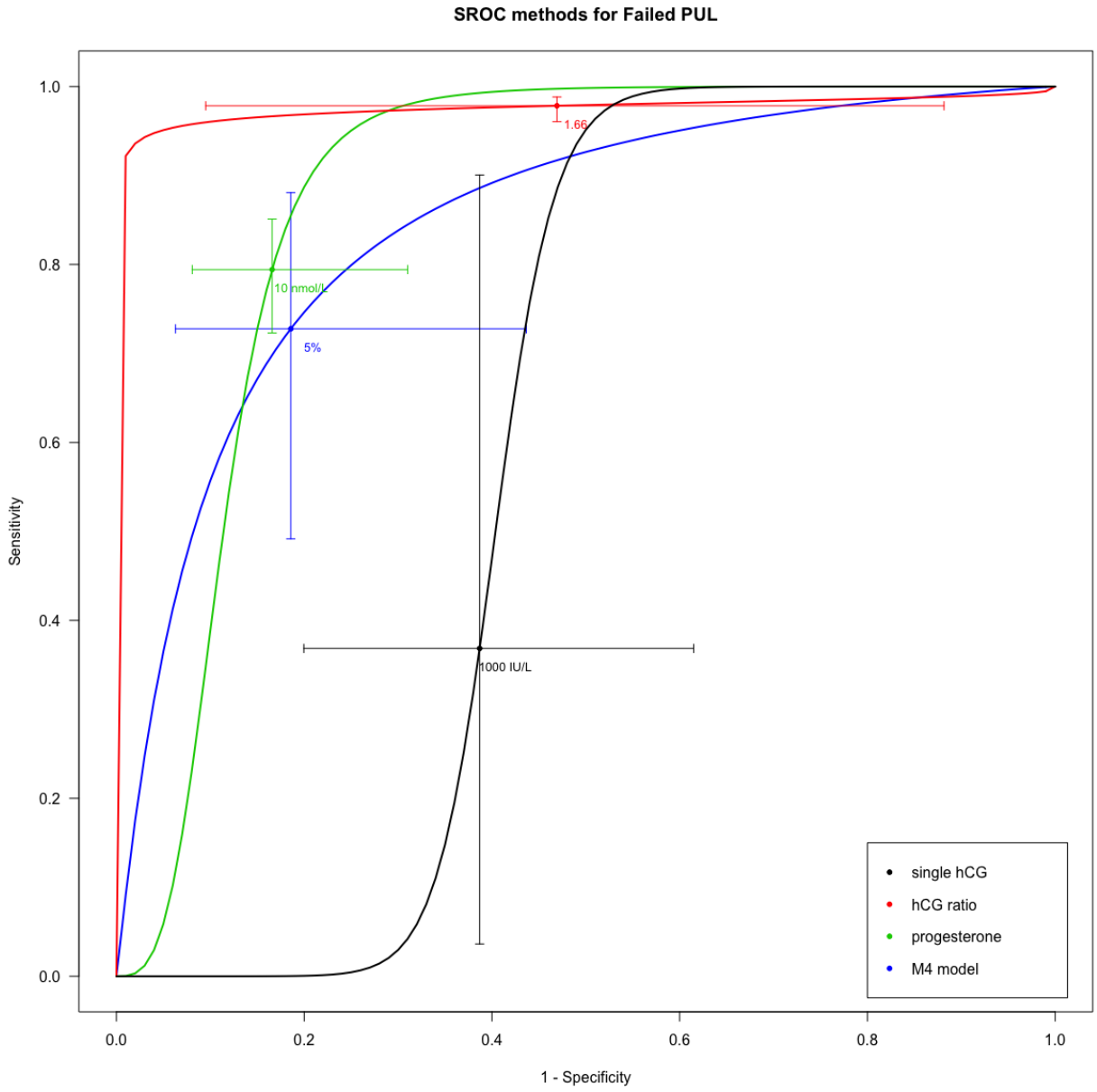


Figure 25: SROC (Summary Receiver Operating Curve) for the test performance of all management protocol predicting a final outcome of failed pregnancy of unknown location (FPUL). The number on each curve refers to the summary threshold for each protocol used in the original papers.
hCG = human chorionic gonadotrophin.

Reproduced from supplementary material published by Bobdiwala et al[181].

Management Protocol	**AUC (95% CI) for FPUL prediction**
Single hCG	0.59 (0.00-1.00)
hCG ratio	0.98 (0.86-1.00)
Progesterone	0.87 (0.76-0.94)
M4 model	0.84 (0.15-1.00)

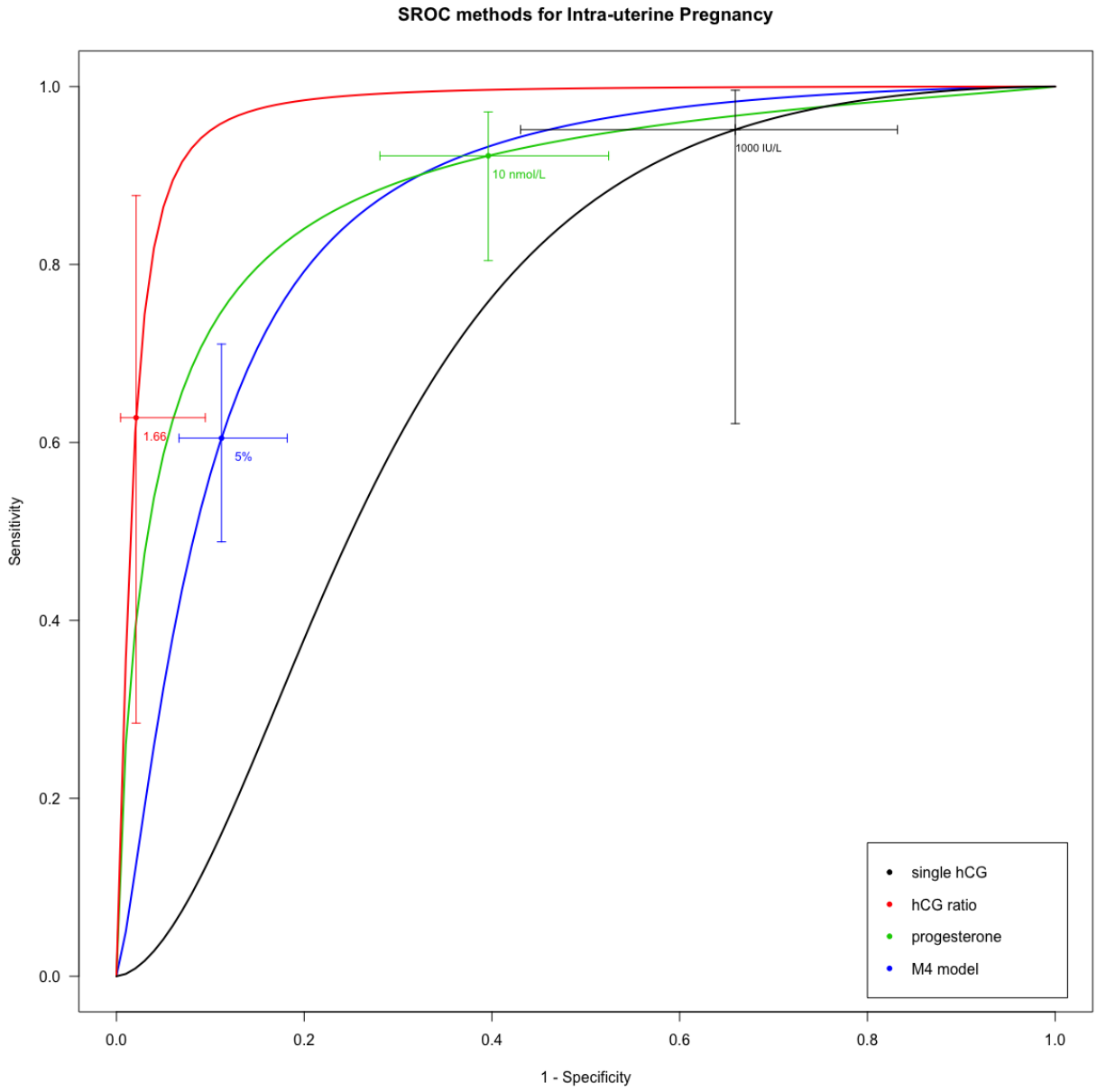


Figure 26: SROC (Summary Receiver Operating Curve) for the test performance of all management protocol predicting a final outcome of intrauterine pregnancy (IUP). The number on each curve refers to the summary threshold for each protocol used in the original papers.
hCG = human chorionic gonadotrophin.

Reproduced from supplementary material published by Bobdiwala et al[181].

Management Protocol	**AUC (95% CI) for IUP prediction**
Single hCG	0.71 (0.01-1.00)
hCG ratio	0.97 (0.87-0.99)
Progesterone	0.89 (0.76-0.96)
M4 model	0.86 (0.39-0.99)

Discussion

Main findings

This systematic review provides healthcare professionals with a comprehensive summary of the current evidence on PUL - a clinical scenario that lacks a clear evidence-base and international consensus on how it should be managed. A delay in the diagnosis of an EP can lead to maternal morbidity and mortality[18]. It is therefore clinically important to rationalise the management of PUL whilst retaining a high level of test performance for predicting a final outcome of EP. We have shown that the M4 model performed better than other management protocols in predicting a final outcome of EP (AUC 0.87). It is also a good discriminator of failing pregnancies (AUC 0.84) and on-going intrauterine pregnancies (AUC 0.86). Consequently, the M4 model can assist clinicians in reducing unnecessary follow-up for low risk PUL, enabling resources to be freed up to monitor PUL predicted to be EP more closely. The AUCs for predicting IUP or FPUL using the hCG ratio and serum progesterone cut-offs are better than the M4 model but are both limited by having an AUC for the correct prediction of EP of only 0.69.

Interpretation

It is important to state that management protocols for PUL are meant to be predictive rather than diagnostic. The management protocol used should guide clinicians on how intensive further follow-up should be and not used to justify unnecessary, potentially harmful, interventions. Most women with a PUL are clinically well and premature intervention, such as treatment with systemic methotrexate, poses the danger of inadvertent termination of what could have otherwise been a healthy IUP. Even when a final diagnosis of EP is confirmed, not all cases will require medical or surgical intervention, as evidence exists to support the safe, expectant management of select EP[208].

Strengths and limitations

There was a large amount of heterogeneity in the studies we meta-analysed. This included the definition used for PUL and how the data was presented. The identified studies also had a high risk of verification bias due to the differences in how final pregnancy outcomes were assessed. There was heterogeneity in the management protocols themselves, e.g. studies evaluated several different hCG cut-offs rather than the same one. To account for this, we applied a novel approach taking into account the multiple thresholds per study. This meant we did not need to limit analyses to an arbitrarily chosen threshold per study and could analyse all available data. Finally, it is also important to recognize that the test performance of any management protocol will be affected by the population it is used in, the available resources and the PUL rate (the number of patients classified as a PUL / all early pregnancy patients scanned in a unit, it is generally accepted that a 'good' PUL rate is <15%[12]).

Summary

This systematic review and meta-analysis provides a summary of the current best evidence for the management of PUL. Each available protocol will have its' strengths and limitations. The hCG ratio and progesterone cut-offs are good at predicting pregnancy viability and failure, but poor at correctly predicting ectopic pregnancy. The M4 model is less capable at predicting viability but better at predicting ectopic pregnancy. Our study showed that the M4 model has the best overall test performance. Focusing further on refining and validating the use of risk prediction models in the setting of a clinical trial is likely to be the best way forward in rationalising the management of PUL.

Final conclusions and future work

This meta-analysis provided a summary of the current available evidence regarding management protocols used to look after women with a PUL. The M4 model had the best test performance overall, particularly in the prediction of ectopic pregnancy as the final pregnancy outcome.

I therefore set out to validate the use of the M4 model by carrying out a prospective multi-centre interventional trial focusing on assessing its' safety as well as its' test performance when used in everyday clinical practice.

STUDY 2

A CLINICAL INTERVENTION STUDY USING THE M4 LOGISTIC REGRESSION MODEL TO MANAGE WOMEN WITH A PUL

Chapter Abstract

This chapter refers to original work published in Bobdiwala et al[199].

Introduction

Van Calster et al[22] previously published a diagnostic accuracy study on the M4 model, a logistic regression model based on the initial hCG and hCG ratio (hCG at 48 hours/ hCG at 0 hours) results of women with a pregnancy of unknown location (PUL). The M4 model classified PUL as being at either low- or high-risk of an ectopic pregnancy with good test performance**.** We aimed to assess the test performance and any adverse outcomes associated with using the M4 model in everyday clinical practice.

Methods

This was a prospective multi-centre cohort study of 1022 women presenting with a PUL to the early pregnancy units of three London hospitals between August 2012 and December 2013. The hCG results at 0 and 48 hours of participants were entered into the M4 model algorithm, which assigned a percentage risk of having an ectopic pregnancy (≥5% used as the cut-off for 'high-risk'). Adverse clinical events associated with its' use were recorded.

Results

835/1022 (82%) PUL managed according to the M4 model had complete final outcome data. The M4 model predicted 70% (585/835) PUL as low-risk, of which 568 (97%) were confirmed as FPUL or IUP. It correctly classified 77/94 (82%) EP as high-risk. There were 17/835 (2.0%) adverse events and no serious adverse events associated with using the M4 model in everyday clinical practice.

Conclusions

The M4 model maintains good test performance when used in clinical practice and is an effective way of triaging women with a PUL as being at high or low-risk of an ectopic pregnancy. The number of adverse events was small (2.0%) and indicates that the expectant management of PUL using the M4 model is safe.

Introduction: the M4 model

The management of women with a PUL and differentiating between low-risk and high-risk cases relies on different methods of interpreting biomarker results – mainly serum progesterone and hCG. Risk prediction models, which include logistic regression models and Bayesian networks, utilise some of the most up to date evidence in managing these women. It is important to make the distinction between a risk prediction model and a diagnostic model. The latter aims to determine what the diagnosis is (i.e. in this scenario the final pregnancy outcome in a women classified as a PUL). The former aims to risk-stratify a patient as either low- or high-risk of a specific complication (i.e. EP in this scenario) and gives evidence-based recommendations on the next steps in managing a patient (i.e. how frequent their follow-up should be). These risk prediction models utilize biomarker results, but may also include clinical variables such as the extent of vaginal bleeding and ultrasound-based variables such as endometrial thickness[21, 23, 49, 209, 210]. One of the most extensively evaluated risk prediction models is M4[21, 22], a logistic regression model based on the patients' initial serum hCG result and the hCG ratio as its' predictive variables.

The M4 model has been shown to reduce unnecessary follow-up in 70% of PUL with a negative predictive value (NPV) of 97.5%[22], which alleviates the burden multiple hospital attendances have on patients as well as having significant cost-saving implications for the NHS. Furthermore, it has been shown to outperform other PUL management strategies such as a progesterone cut-off ≤10 nmol/l and the hCG ratio alone[50].

The M4 model has undergone both temporal and external validation but a prospective assessment of any complications coupled with its use in 'real world' clinical practice has not been performed. The primary aim of this study was to evaluate any complications or adverse outcomes associated with using the M4 model in everyday clinical care. A secondary aim was to further prospectively validate M4 model use on a new patient cohort.

Methods

Design and settings

This was a prospective multi-centre cohort study. Consecutive women that presented to the Early Pregnancy Units (EPUs) of three UK hospitals based in London and were classified as a PUL at their first ultrasound scan were eligible for inclusion. Two of the sites were university teaching hospitals: Queen Charlottes' and Chelsea Hospital (QCCH) and Chelsea and Westminster Hospital (CWH). The third site was a district general hospital: Hillingdon Hospital (HH). The study took place between August 2012 and December 2013 at QCCH (16 months); February to September 2013 at HH (7 months) and March to September 2013 at CWH (6 months). As the M4 model had undergone previous validation and peer-reviewed publication[22, 50] the research and development departments at the participating hospitals were advised that formal ethical approval for the study was not required and it was registered as an audit.

Inclusion and exclusion criteria

As outlined in the methods chapter, see page 60.

Sample collection and analysis

As outlined in the methods chapter, see page 69.

Triage using the M4 risk-prediction model

The algorithm for the M4 model was embedded into a simple protected Microsoft Excel™ file that all staff could access, including junior doctors on-call out of normal working hours. Initial and 48 hour serum hCG results were inputted into the excel document and the M4 model algorithm would estimate the risk of EP, 'high-risk' pre-determined as a risk of EP of ≥5% [22].

We considered it sensible to refer women with a PUL for more extensive follow-up if their risk of having an EP was at least 5%. This value of ≥5% was partly subjective as there is no

internationally accepted cut-off in this scenario. A value of ≥5% has been found to be a safe and reasonable cut-off level in previous, similar studies[22]. One could argue that the interpreting clinician, based on their preference, could change the threshold. This has certainly been proposed in other clinical scenarios such as the prediction of ovarian masses as either benign or malignant based on logistic regression models[211, 212]. There is always a balance required between additional avoidable workload and the clinical concerns around misclassifying an EP[213-215]. Lower cut-offs would result in fewer missed EP but would also result in increased unnecessary follow-up for patients that do not require it (i.e. increased false positive rate). When a cut-off of ≥3% rather than ≥5% was assessed for the M4 model[22], more EP were correctly classified as high-risk (90.2% versus 88.0%) but this correlated with fewer PUL being classified as low-risk (66.5% versus 69.6%) and a higher false positive rate (23.6% versus 20.4%).

In this study, we on balance chose to use a cut-off of ≥5%. If the risk was <5%, the PUL was classified as 'low-risk, probable FPUL' or low-risk, probable IUP'. Women with a 'low-risk probable FPUL', were advised to perform a urine pregnancy test in two weeks with confirmation of a negative result via telephone follow-up by trained EPU staff. If the urine pregnancy test was positive, patients were brought back for further blood tests and ultrasound scans. PUL defined as 'low-risk probable IUP' were followed up with a repeat TVS after one week to confirm the presence of an IUP. PUL triaged as 'high-risk probable EP' were instructed to return in 48 hours for a repeat ultrasound scan by a senior clinician (figure 27).

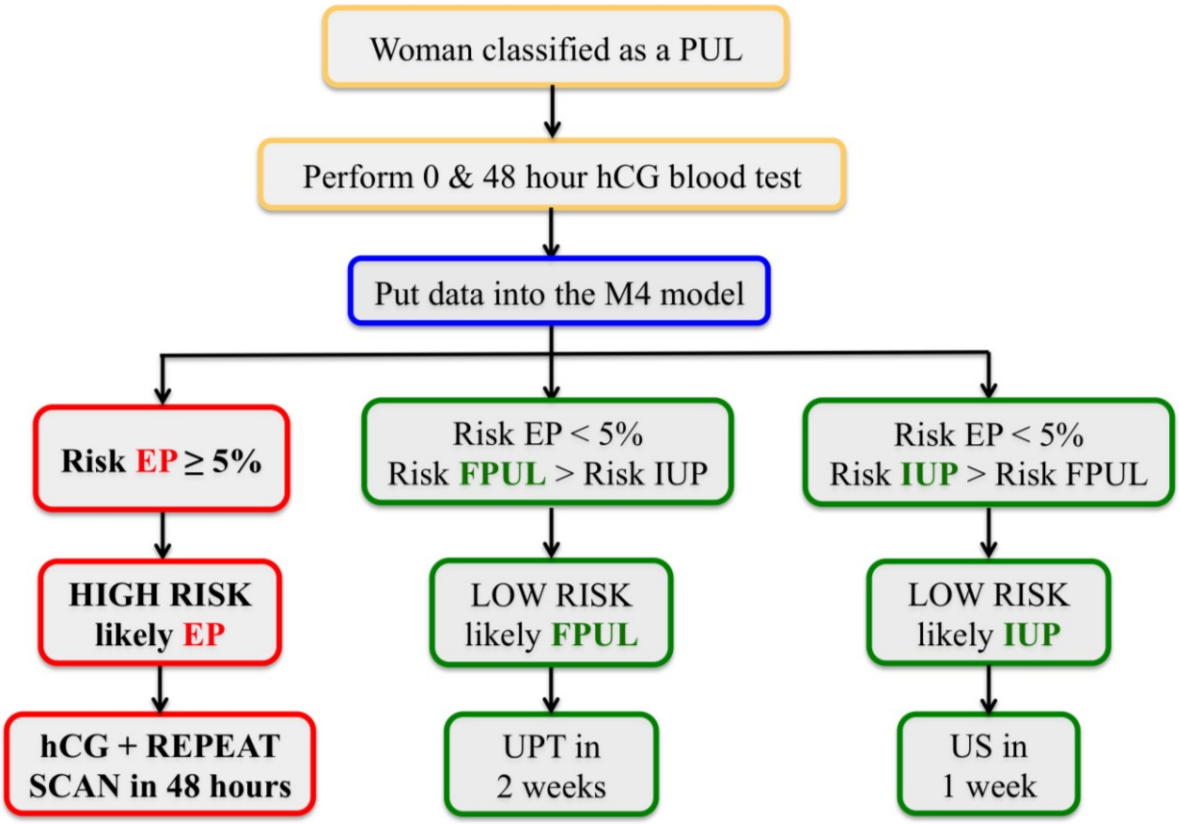


Figure 27: The three possible predictions and management plans as recommended by the M4 model after inputting the initial (0h) and 48h serum hCG results. PUL, pregnancy of unknown location; EP, ectopic pregnancy; FPUL, failed pregnancy of unknown location; IUP, intrauterine pregnancy; US, ultrasound; UPT, urine pregnancy test. Adapted from Bobdiwala et al[199].

All patients were warned that any symptoms of concern, such as increased levels of pain or vaginal bleeding, warranted medical review as an EP had not been ruled out. All clinicians using the M4 model were aware that it was a tool to guide management but should not supersede their judgment if they felt that the clinical situation required a different management plan. The M4 model only provided guidance on follow-up and the decision for intervention always lay with the clinical team in charge.

Reference standard: pregnancy outcome definitions

All final outcomes of the pregnancy were defined at the time point the pregnancy would be expected to be 12 weeks in gestation. The definitions used for final outcomes are as described in the methods chapter (pages 65-8).

Deviations from protocol

A protocol deviation was defined as any deviation from the management recommended by the M4 model. These were categorized under three subheadings:

[1] *Patient reason:* the patient attended for an unscheduled visit to the EPU or A&E (accident and emergency) departments for increased symptoms (most commonly pain or vaginal bleeding)

[2] *Clinician reason:* the clinician in charge of that patients' care felt that, based on their judgment of the overall clinical scenario, the management recommended should not be followed

[3] *Timing reason:* all participating EPUs were only open Monday-Friday, 09:00 – 17:00, so if a follow-up scan recommended by the M4 model fell on a weekend or bank holiday, this would usually be moved to the following working day. Blood tests were still taken out of hours by on-call staff. This category also included appointments that were rescheduled by patients due to other commitments.

Adverse events

The national Good Clinical Practice (GCP) guidelines[216] cover the legislation, guidance and good practice relating to the conduct of clinical trials in the UK. An 'adverse event' (AE) is defined as "*any untoward medical occurrence in a patient or clinical study subject*". A 'serious adverse event' (SAE) is defined as "*any untoward and unexpected medical occurrence or effect that results in death (or) is life-threatening*". We extended the classification of a SAE to include any untoward occurrence that may have caused the death of a potentially viable intrauterine pregnancy.

Statistical Analysis

Statistical analysis was performed with SAS v9.4 (SAS Institute, Cary, USA).

Results

Patient flow

Of 1022 women classified as a PUL, 514 were from QCCH, 306 from CWH and 202 from HH. Nine participants met the exclusion criteria. Sixty-nine (7%) patients were lost to follow-up (final outcome unknown): 39 (8%) from QCCH, 26 (9%) from CWH and 4 (2%) from HH. The 48-hour serum hCG result was missing for 109 (11%) of the remaining 944 patients, largely because the initial hCG had been low (<50 IU/L). In those cases, there was a 'clinician-reason' for deviating from protocol as women were asked to perform a follow-up urinary pregnancy test in two weeks rather than a serum hCG at 48 hours. In all cases the urinary pregnancy test result was negative. Additional reasons for missing 48-hour hCG results included: failure of the patient to attend for the 48-hour blood test (n=28), visualization of either an intra- or extra-uterine pregnancy on a repeat ultrasound scan (n=10), a definitive intervention was performed prior to the second blood test (n=2) and loss of the blood sample (n=1).

This left 835 women with a PUL for final analysis. Data on demographics, presenting symptoms and serum biochemistry results are shown in table 2. The final pregnancy outcome was a FPUL in 443 (53%) patients, IUP in 298 (36%) and EP in 94 (11%) patients.

Characteristic	All (n=835)	FINAL PREGNANCY OUTCOME FPUL (n=443)	IUP (n=298)	EP (n=94)
Age (years)	32 (28-36)	33 (28-38)	31 (26-35)	33 (29-36)
Initial serum hCG (IU/L)	571 (183-1914)	418 (115-1937)	825 (363-2081)	486 (189-1187)
48h serum hCG (IU/L)	462 (123-1707)	155 (48-510)	1680 (701-3439)	504 (251-1516)
hCG ratio	0.74 (0.33-1.84)	0.35 (0.24-0.46)	2.04 (1.55-2.42)	1.18 (0.93-1.49)
Initial progesterone (nmol/L)	12 (4-40)	4 (2-9)	49 (31-64)	19 (7-32)
Mean pain score	2.1	2.2	2.0	2.1
Vaginal bleeding				
No bleeding (0)	211 (25%)	22 (5%)	169 (57%)	20 (21%)
Minimal bleeding (1)	230 (28%)	114 (26%)	75 (25%)	41 (44%)
Moderate bleeding (2)	187 (22%)	130 (29%)	33 (11%)	24 (26%)
Soaked sanitary towel (3)	92 (11%)	80 (18%)	7 (2%)	5 (5%)
Clots or flooding (4)	115 (14%)	97 (22%)	14 (5%)	4 (4%)
Mean score	1.6	2.3	0.7	1.3
History of EP				
Any	45 (5%)	10 (2%)	25 (8%)	10 (11%)
Laparoscopy	32 (4%)	7 (2%)	19 (6%)	6 (6%)
Methotrexate	7 (1%)	2 (<1%)	3 (1%)	2 (2%)
Expectant	6 (1%)	1 (<1%)	3 (1%)	2 (2%)
Indication for scan				
Bleeding and pain	432 (52%)	299 (67%)	79 (27%)	54 (57%)
Bleeding only (no pain)	184 (22%)	119 (27%)	47 (16%)	18 (19%)
Pain only (no bleeding)	159 (19%)	13 (3%)	131 (44%)	15 (16%)
Previous ectopic	14 (2%)	4 (1%)	9 (3%)	1 (1%)
Unsure dates	10 (1%)	2 (<1%)	8 (3%)	0 (0%)
Previous miscarriage	2 (<1%)	0 (0%)	2 (1%)	0 (0%)
Maternal reassurance	1 (<1%)	0 (0%)	0 (0%)	1 (1%)
Other	33 (4%)	6 (1%)	22 (7%)	5 (5%)

Table 2: Data on demographics, symptoms and serum biochemistry results for patients in the study with pregnancy of unknown location (PUL). Results are shown as median (interquartile range) for age and biomarkers, mean score for pain and vaginal bleeding, and n (%) for vaginal bleeding, history of ectopic pregnancy (EP), and indication for scan. IUP = intrauterine pregnancy. There were 248 missing values for initial progesterone (30%), 177 for pain score (21%), and 2 for history of EP (0.2%). Reproduced from Bobdiwala et al[199].

M4 model test performance

PUL predicted low-risk

570/835 (68%) PUL were classified as low-risk by the M4 model. Of these, 553 (97%) were indeed low-risk, with final outcomes of FPUL (n = 393) and IUP (n = 160). 17/94 (18%) EP were misclassified as low-risk.

PUL predicted high-risk

94 PUL had a final outcome of EP. Of these, 77 (82%) were correctly classified as high-risk by the M4 model. 188/741 (25%) FPUL or IUP were also misclassified as high-risk. See figure 28.

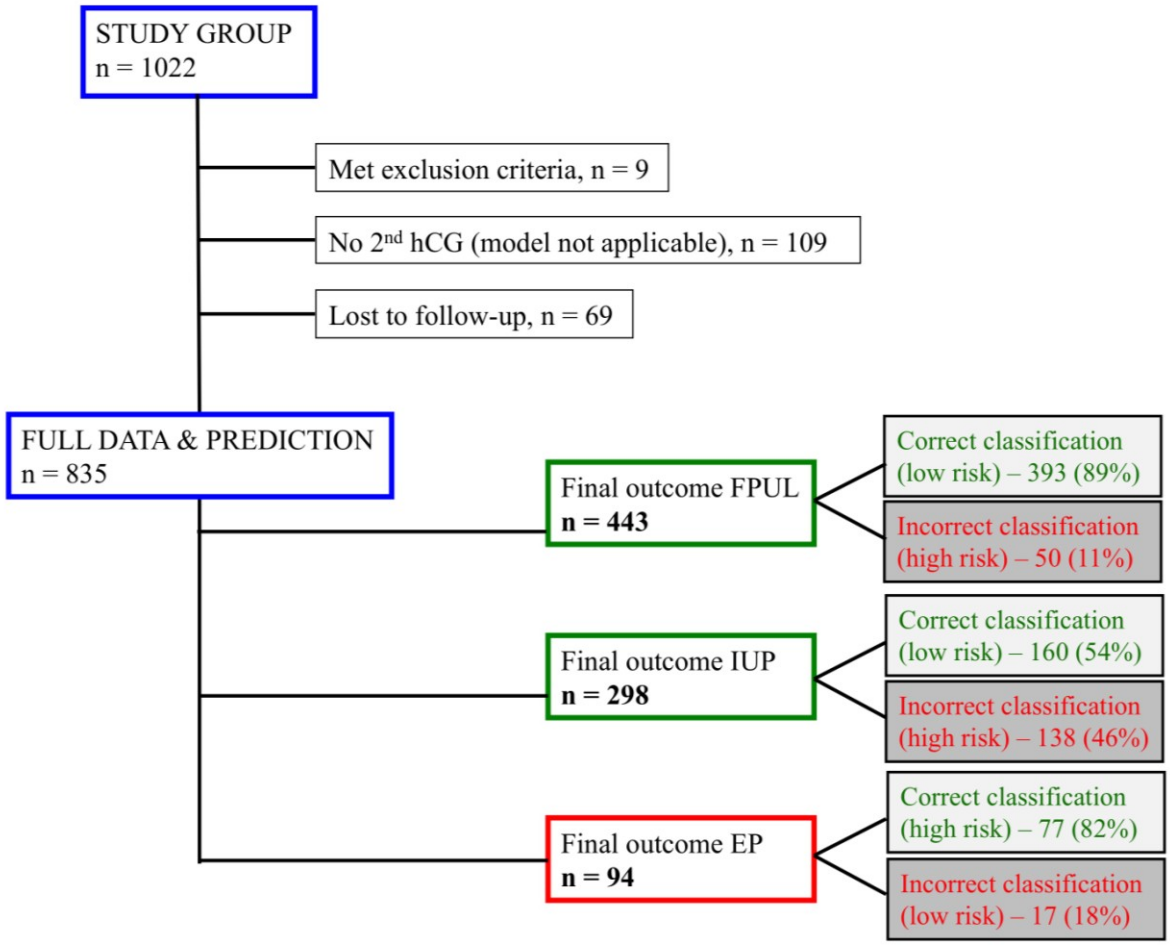


Figure 28: Flowchart of final outcome data and correct versus incorrect risk stratification according to the M4 model. Reproduced from Bobdiwala et al[199].

EP/ PPUL misclassified by the M4 model

77/94 (82%) EP/ PPUL were correctly predicted to be high-risk by the M4 model. Of these, 19 (25%) were managed expectantly, 28 (36%) underwent medical management (with intramuscular methotrexate) and 23 (30%) had surgical intervention with a laparoscopic salpingectomy. Of the 17/94 (18%) EP/ PPUL misclassified as low-risk by the M4 model, 5 (29%) had expectant management, 7 (41%) had medical management and 5 (29%) had surgical management (see table 3).

Patient	hCG 0 hours	hCG 48 hours	hCG ratio	Prog 0 hours	M4 model: prediction of outcome	Actual final outcome	Intervention	Indication for intervention
1	626	408	0.652		FPUL	Ectopic pregnancy	**Expectant management**	
2	330	700	2.121	19	IUP			
3	485	281	0.579	44	FPUL			
4	1550	436	0.281	54	FPUL			
5	1805	926	0.513	16	FPUL			
6	84	256	3.048	28	IUP	Ectopic pregnancy	**Methotrexate**	Ectopic seen on one week follow-up ultrasound
7	105	83	0.790	4	FPUL			Later rise in hCG and ectopic seen on ultrasound
8	129	340	2.636	7	IUP			Later rise in hCG and ectopic seen on ultrasound
9	76	184	2.421	76	IUP			Later rise in hCG and ectopic seen on ultrasound
10	525	208	0.396	18	FPUL			Later rise in hCG and ectopic seen on ultrasound
11	187	132	0.706		FPUL	Persisting PUL		Negative laparoscopy & SMM, methotrexate given for persistent hCG
12	182	126	0.692	20	FPUL			Two methotrexate doses required
13	1164	2496	2.144	71	IUP	Ectopic pregnancy	**Laparoscopy**	Surgery as methotrexate not suitable (renal disease)
14	438	941	2.148	26	IUP			Patient chose to have surgery
15	235	624	2.655		IUP			Re-admitted with pain
16	162	389	2.401		IUP			Re-admitted with pain
17	64	40	0.625	37	FPUL			Patient chose to have surgery

Table 3: The 17 high-risk patients misclassified as 'low risk' by the model and the interventions they had. Prog = progesterone, PPUL = persistent pregnancy of unknown location, FPUL = failed pregnancy of unknown location, SMM = surgical management of miscarriage. Adapted from Bobdiwala et al[199].

Test performance compared to other PUL management protocols

We compared the test performance of other PUL management protocols in current clinical use such as the hCG ratio and a progesterone cut-off of ≤10nmol/l against that of the M4 model on this dataset. Data from CWH was omitted from the sub-analysis for progesterone cut-offs as progesterone levels were not consistently measured in that hospital.

The hCG ratio classified more PUL as low-risk when compared to the M4 model (81% versus 70%) but had a much lower sensitivity for correctly classifying EP/ PPUL as high-risk (68% versus 82%). A progesterone cut-off of ≤10nmol/l classified far fewer PUL as low-risk (47%) and also had a lower sensitivity for picking up EP (68%). See table 4.

	PUL classified as low risk	NPV	Sensitivity	FPR
Using all data (n=835)				
M4 decision	70% (67 to 73)	97% (95 to 98)	82% (73 to 88)	23% (20 to 27)
hCG ratio cut-offs	81% (78 to 83)	96% (94 to 97)	68% (58 to 77)	13% (11 to 15)
Difference	-11% (-13 to -9)	nc	14% (7 to 21)	10% (8 to 13)
Using QCCH and HH patients with progesterone levels (n=560)				
M4 decision	68% (64 to 72)	97% (94 to 98)	82% (72 to 89)	25% (21 to 29)
Progesterone cut-off	47% (43 to 51)	91% (87 to 94)	68% (57 to 78)	51% (46 to 55)
Difference	21% (16 to 26)	nc	14% (-1 to 28)	-26% (-31 to -21)

Table 4: The performance of [i] the M4 model, [ii] triage based on hCG ratio cut-offs (low risk if hCG ratio <0.87 or >1.66), and [iii] triage based on a single progesterone cut-off (low risk if progesterone ≤10nmol/l). 95% confidence intervals are shown in parentheses. NPV = negative predictive value; FPR = false positive rate; QCCH = Queen Charlottes' & Chelsea Hospital; HH = Hillingdon Hospital; nc = not computed. Adapted from Bobdiwala et al[199]

Protocol deviations

In 710/835 (85%) cases, the management recommended by the M4 model was followed. Of the remaining 125 cases, there were 19 protocol deviations secondary to a patient reason, 38 secondary to a clinician reason and 68 secondary to an issue with the timing of a follow-up blood test or ultrasound scan (see figure 29).

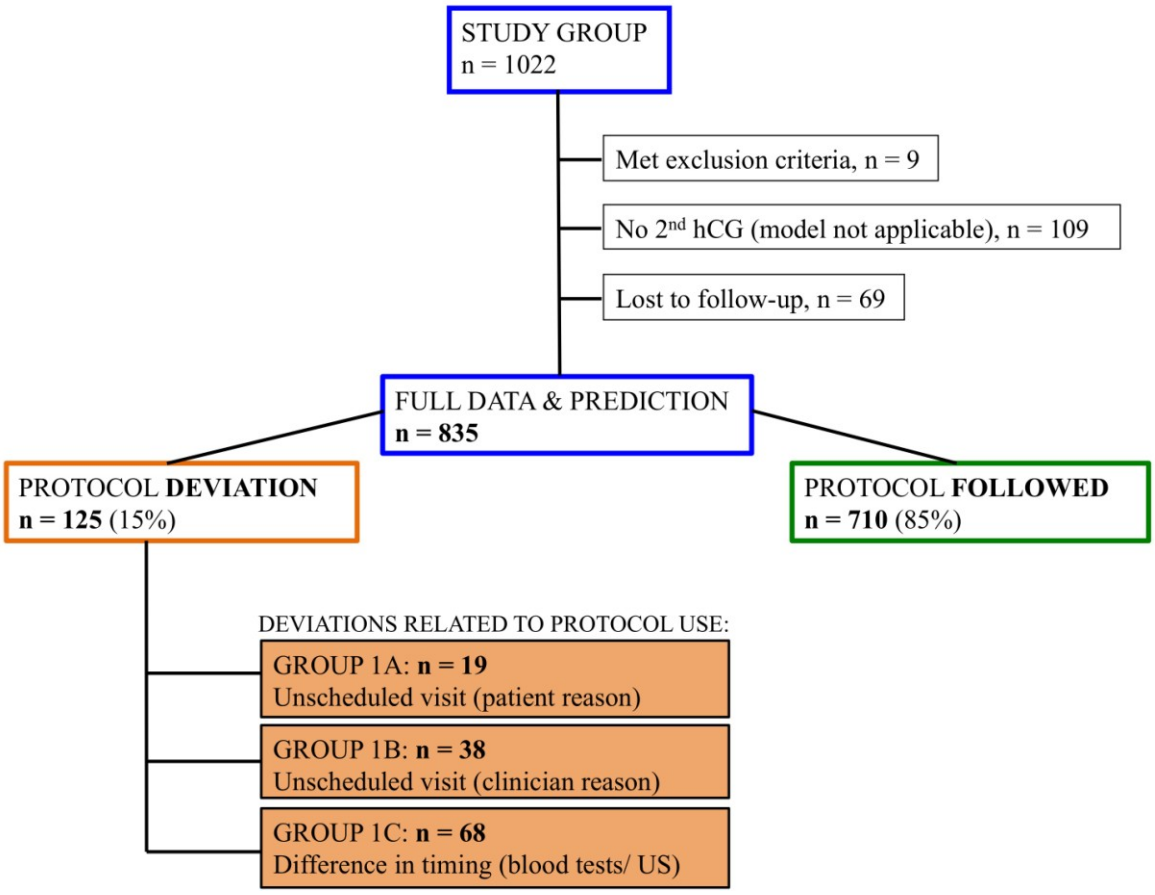


Figure 29: Flowchart of patient recruitment and protocol deviations. Adapted from Bobdiwala et al[199]

Deviations secondary to a patient reason

These 19 cases represent women who had an unscheduled hospital visit for increased symptoms of pelvic pain or vaginal bleeding. The final outcomes were 5 FPUL, 5 NVIUP, 1 VIUP and 8 EP/PPUL. The M4 model incorrectly classified 6/8 of these EP/PPUL as low-risk. Of these, 3 underwent a laparoscopy, 2 had successful medical management with methotrexate and 1 had

persistent hCG levels despite medical management and had a laparoscopic salpingectomy as second line treatment. None of the cases that underwent surgical intervention were found to have a ruptured EP at the time of surgery.

Deviations secondary to a clinician reason

These 38 cases represent women who had their follow-up hospital appointment (including the timing of a blood test or scan) either earlier or later than recommended by the M4 model as the clinician in charge of that patients' care felt, in their judgment of the whole case, that this was necessary.

Deviations secondary to a timing reason

These 68 cases represent women who had their follow-up hospital appointment (including the timing of a blood test or scan) either earlier or later than recommended by the M4 model as their follow-up date fell on a weekend/ bank holiday when the EPUs were not open. This also included women who phoned to change their follow-up to a date/ time that was more convenient for them.

Analysis of adverse events (AE) & serious adverse events (SAE)

Twenty-six women had an AE and 1 woman a SAE. These were subcategorized into AE/ SAE related to M4 model use, an incorrect clinical decision or an unscheduled admission of a patient to hospital (see figure 30).

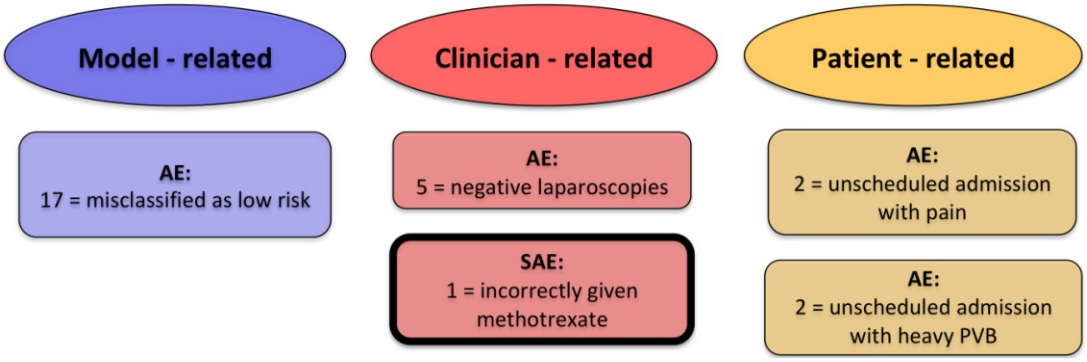


Figure 30: Breakdown of adverse and serious adverse events in the study cohort. AE, adverse event; SAE, serious adverse event; PPUL, persistent pregnancy of unknown location; PVB, per vaginal bleeding. Adapted from Bobdiwala et al[199]

AE: M4 model-related

These 17 cases represent women with a final outcome of EP/PPUL incorrectly classified as low risk by the M4 model (see table 3).

AE: clinician-related

Five patients were felt, by the clinician in charge, to require surgical intervention for a presumed adnexal mass or haemoperitoneum in the pelvis on a follow-up ultrasound scan but all had a negative laparoscopy. Interestingly, none of these patients had re-presented with pelvic pain. Four cases had been classified as low-risk by the M4 model and eventually had a final outcome of VIUP (n=1), NVIUP (n=1) and FPUL (n=2). One case had been classified as high-risk by the M4 model and ultimately had a final outcome of a NVIUP that was confirmed on histology after the patient underwent surgical management of miscarriage.

AE: patient-related

These 4 cases represent women who had emergency hospital admissions. Two patients had attended with heavy vaginal bleeding but were successfully managed expectantly and did not require emergency surgery (final outcomes were FPUL and a NVIUP). The remaining two women attended with severe pain, appropriately underwent diagnostic laparoscopies but both of these were negative.

SAE

One patient had a SAE. Her serum hCG levels were 24 rather than 48 hours apart and the M4 model was inappropriately used. A void 'high-risk' prediction was given and the patient erroneously received methotrexate. The pregnancy was ultimately found to be an IUP and was terminated after extensive counselling due to the recognized teratogenic effects of methotrexate on a developing fetus.

Discussion

Test performance

This study verified that the M4 risk prediction model can be incorporated into everyday clinical practice and used to triage women with a PUL as being at low or high-risk of complications with good test performance comparable to what was predicted in a previous diagnostic accuracy study[22]. Sixty-eight percent of PUL cases were classified as low-risk after just two hospital visits with a NPV of 97%. This is of real clinical relevance as women with a PUL often undergo prolonged unnecessary follow-up and streamlining this has the potential for reducing the burden of repeated hospital visits on patients and the avoidable cost burden on the NHS.

Clinical safety

This study also evaluated the safety of using the M4 model in clinical care, demonstrating that it was not associated with any serious clinical adverse events (AE) that caused harm to patients. Although 17 EP/PPUL were misclassified as low-risk by the M4 model, there were no cases with a ruptured EP, none required a blood transfusion, prolonged hospital stay or any other event that increased their morbidity. Two women in the study population (1:511), having undergone M4 model triage, were readmitted and required surgery because of pain. This would be classified as an uncommon event (between 1:100 - 1:1000) and may be useful information when counselling women on the risks associated with expectant management when using the M4 model. In both these cases, no EP was found at the time of laparoscopy.

The case resulting in a SAE in this study was given methotrexate out of hours based on ultrasound scan findings of a presumed adnexal mass and two hCG levels incorrectly taken 24, rather than 48, hours apart. The M4 model was inappropriately used and the case was classified as at 'high-risk'. This patient had a final outcome of an IUP. After counselling that included the principles of duty of candour, this pregnancy was terminated due to the known teratogenic

effects of methotrexate on the developing fetus. This error highlights the importance of ruling out an IUP before treating any pregnant patient with methotrexate.

Protocol deviations

There were protocol deviations in 15% of the final study population. Thirty percent (38/125) of protocol deviations were clinician-related, i.e. the result of doctors overruling the follow-up recommended by the M4 model. This implies that the M4 model was being used as it was intended – an adjunct rather than a replacement for clinical decision-making. Interestingly however, when clinicians did not follow the M4 model, there was a tendency to over-intervene.

It is also reassuring that 15% (19/125) of protocol deviations were secondary to unscheduled attendances by patients as it implies they were appropriately counselled about seeking medical assessment should their symptoms worsen. No management algorithm for PUL will ever be perfect as it is well established that the hCG trend in 10% of EP will mimic that of a VIUP and 20% will mimic a FPUL[11, 217, 218]. These results demonstrate that as long as women are correctly instructed to re-present to hospital should they have increased pain/ vaginal bleeding, an appropriately timed diagnosis and intervention can still occur without increased clinical harm.

The majority (54%, 68/125) of deviations from the recommended management protocol were secondary to timing issues with blood tests or scans. This can be a practical problem in service provision as in the UK most EPUs are only open during standard working hours or because women simply cannot attend at the required time.

Strengths and limitations

A strength of the study is that it was performed in multiple centres that included both teaching and district general hospitals. These have different patient populations, staffing levels and resources. The two university teaching hospitals had dedicated early pregnancy clinical research fellows and more than one consultant. The district general hospital ran with a nurse specialist, one consultant and rotating middle grade doctors. These results, therefore, should be generalizable to the UK population. The main limitations of the study are that 69 (7%) of PUL were lost to follow-up. Forty-eight hour hCG levels were also missing for 109 (11%) patients, although 68 of these had initial hCG levels <50 with negative urinary pregnancy tests 2 weeks later, so we can clearly define these cases as having a final outcome of FPUL.

Summary

The clinical care of women with a PUL has moved towards effective risk prediction and triage. This study has shown good test performance of the M4 model when used in 'real-world' clinical practice in over 1000 patients with no concerns regarding its' safety. Clinicians appeared willing to use it with overall good compliance with M4 model use.

In summary, the M4 model is a user-friendly predictive tool that rationalizes follow-up and can guide healthcare professionals when making decisions about the management of PUL.

Final conclusions and future work

This study established that a multi-centre study incorporating a risk prediction model for PUL into everyday clinical practice was feasible. More importantly, there was good test performance with few adverse events related to using the M4 model and none that caused any significant clinical harm.

The study also highlighted how there was scope to improve test performance. The M4 model is based on the patients' initial hCG and hCG ratio results as predictive variables. It does not incorporate progesterone levels into its' algorithm, a biomarker that is in clinical use in other EPUs in the UK[38, 39]. Accordingly, the next step was to build and evaluate a risk prediction model for women with a PUL that incorporated both progesterone and hCG as predictive variables.

STUDY 3

UTILIZING SERUM PROGESTERONE LEVELS AS WELL AS HCG – DEVELOPING A 2-STEP PROTOCOL USING THE M6 MODEL

Chapter Abstract

This chapter refers to original work published in Van Calster/ Bobdiwala (joint first authors) et al[40].

Introduction

Mathematical models, such as the M4 model, have been developed and tested as methods for improving the management of women classified as having a pregnancy of unknown location (PUL). They have demonstrated good test performance but there still remains scope for improvement. We developed a two-step triage protocol based on the presenting serum progesterone (step 1) and hCG ratio two days later (step 2) to select PUL at high-risk of ectopic pregnancy (EP).

Methods

This was a secondary analysis of prospective and consecutively collected data on 2753 cases of PUL at two UK-based university teaching hospitals. The data was chronologically split, with 1449 PUL cases used in the development dataset and 1304 cases in the validation dataset. Step 1 triage selects low-risk PUL at the initial presentation using a serum progesterone cut-off value. Step 2 triage utilises a novel logistic regression risk model (the 'M6 model') based on the hCG ratio and initial serum progesterone, defining low-risk as an EP risk <5%.

Results

In the validation dataset, step 1 triage with an initial serum progesterone cut-off of ≤2nmol/l, classified 16.1% PUL cases as low-risk. Step 2 triage with the M6 model classified a further 46.0% PUL as low-risk. Overall, the two-step protocol classified 62.1% of PUL as low-risk, with an NPV of 98.6% and a sensitivity of 92.0%.

Conclusions

This two-step PUL triage protocol can successfully classify PUL into being at high- or low-risk of complications with good test performance. It outperforms other validated risk prediction models, including the M4 model, used in the management of women with a PUL.

Introduction: a two-step strategy and the M6 model

It is recognized that the current clinical management of women with a PUL is often chaotic and lacks a clear evidence-base. Measurements of serum hCG[45, 46, 219, 220] and/or progesterone[38, 39] are primarily used to determine management but how they are interpreted varies greatly. The M4 model is a logistic regression model based on the initial hCG and hCG ratio as variables[21]. It has been prospectively validated and demonstrated good test performance[22, 199]. As outlined in study 2, a multi-centre prospective cohort study of 1022 patients[199] showed it correctly classified 97% (568/585) PUL as low-risk and 82% (77/94) PUL as high-risk. This test performance was superior to other published management protocols for PUL based on hCG or progesterone such as the hCG ratio alone or single serum progesterone cut-offs[50].

There is potential value, however, in attempting to combine the strengths of these different strategies. We therefore devised a two-step protocol based on:

Step 1: a progesterone cut-off level, and

Step 2: a logistic regression model (the M6 model) that utilized both the hCG ratio and progesterone as variables

Methods

Design, setting and participants

This was an observational cohort study to assess the diagnostic accuracy of a two-step protocol to manage PUL. It involved a secondary analysis of data that was prospectively and consecutively collected from the Early Pregnancy Units (EPU) of two UK-based university teaching hospitals: St Georges' hospital, London (SGH) between July 2003 and February 2007 and Queen Charlottes' & Chelsea hospital, London (QCCH) between April 2009 and December 2013. Data was collected in accordance with a standardized protocol[21, 41], as outlined in the methods chapter, page 55.

Inclusion and exclusion criteria

As outlined in the methods chapter, page 60.

The data from the two EPUs were combined and chronologically split into a development and validation dataset.

Sample collection and analysis

As outlined in the methods chapter, page 69.

Reference standard: pregnancy outcome definitions

All final outcomes of the pregnancy were defined at the time point the pregnancy would be expected to be 12 weeks in gestation. The definitions used for final outcomes are as described in methods chapter, pages 65-8.

Development of a two-step triage protocol for PUL

The aim of any PUL triage protocol is to risk-stratify patients into either a low-risk or high-risk category whilst minimizing false positive or false negative results. The two-step triage protocol

combines the benefits of previously individually developed PUL management strategies based on serum progesterone and hCG levels. The algorithm is summarized in figure 31.

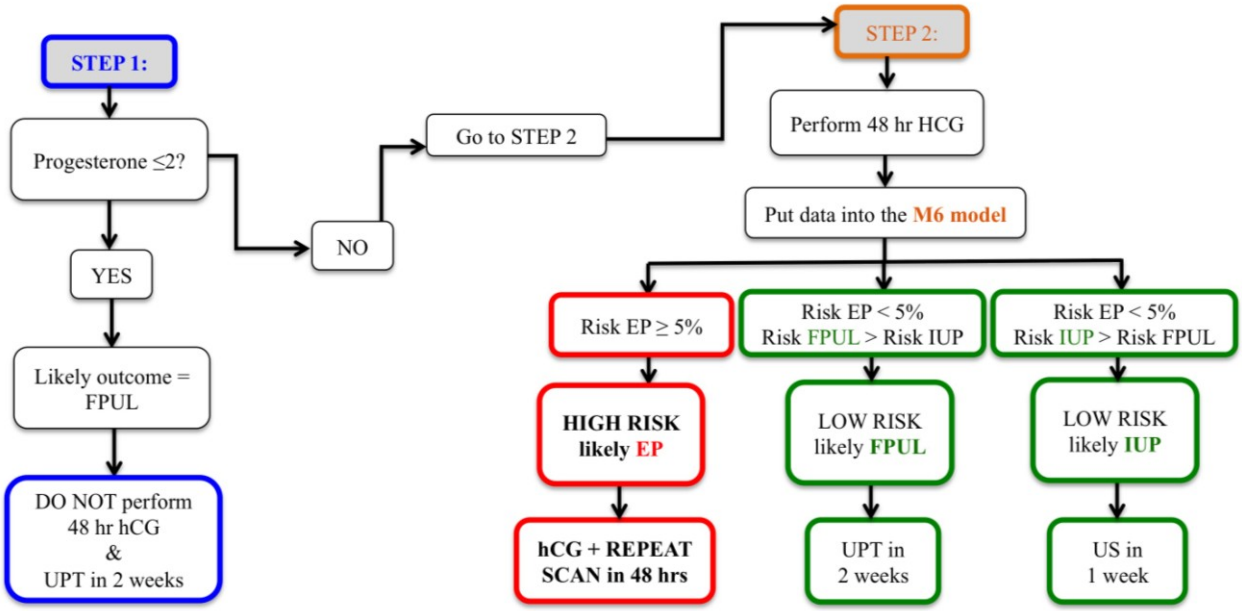


Figure 31: Flow diagram of two-step approach for managing pregnancies of unknown location (PUL). EP, ectopic pregnancy; FPUL, failed PUL; hCG, human chorionic gonadotropin; IUP, intrauterine pregnancy; UPT, urine pregnancy test; US, ultrasound examination. Adapted from Van Calster/ Bobdiwala et al[40].

Data was available on 2753 cases of PUL from the two centres. The triage protocol was developed on the first set of patients from each centre (n=1449, of which 785 FPUL, 501 IUP, and 163 EP), and validated on the second half (n=1304, of which 665 FPUL, 501 IUP, and 138 EP). Missing values for initial serum progesterone and hCG levels were handled using multiple imputation[221]. Second hCG levels taken more than three days after the first hCG level were considered missing.

Step 1 development:

This involved determining an appropriate serum progesterone level cut-off that enables the safe discharge of women with a PUL likely to have a final outcome of FPUL. Various different cut-offs were assessed, further details are provided in the results section.

Patients not classified as low-risk at step 1 then proceed to step 2 (figure 31).

Step 2 development:

Step 2 is risk prediction based on a multinomial logistic regression model called the M6 model.

Two sub-models were developed: $M6_P$ (P = progesterone) and $M6_{NP}$ (NP = no progesterone).

- $M6_P$ uses the initial serum hCG and progesterone levels and the hCG ratio as key predictors
- $M6_{NP}$ uses the initial serum hCG levels and the hCG ratio as key predictors. Progesterone is omitted from this sub-model so that it can be used in the small proportion of patients that are known to be on exogenous progesterone supplementation.

Any PUL with an estimated risk of EP <5% is labelled as low-risk. The M6 model can also be used in isolation, i.e. without step 1.

Two-step protocol validation

The main measures of test performance evaluated were:

- Negative predictive value (NPV): the percentage of non-EP amongst PUL classified as low-risk
- Sensitivity: the percentage of EP correctly classified as high-risk
- Overall percentage of PUL classified as low-risk (as this cohort would receive reduced follow-up)
- False positive rate (FPR): percentage of non-EP classified as high-risk

Performance was also compared to other published triage tools for PUL: the M4 model (also using an estimated risk of EP <5% to define low-risk[22]), and a single visit protocol using a serum progesterone level cut-off of ≤10nmol/l to define low-risk[39].

Results: a diagnostic accuracy study

Patient flow

Of 3152 cases of PUL recruited to the study, 2058 were from SGH and 1094 from QCCH. 271 cases were excluded: 58 from SGH that were lost to follow-up, 212 cases with an initial hCG ≤25IU/L, and 1 patient that was taking progesterone supplementation. This left 2881 patients (1858 SGH, 1023 QCCH). There were a further 164 cases with missing values for serum hCG at 48 hours (5.7%) and 188 for initial progesterone (6.5%). For 128 PUL, the final outcome of the pregnancy was not known (4.4%). These were also excluded, leaving 2753 patients (1786 from SGH, 967 from QCCH). Descriptive data on these cases is shown in table 5.

	FINAL PREGNANCY OUTCOME		
Centre **Variable**	**Failed PUL**	**IUP**	**Ectopic**
SGH	N=898	N=719	N=169
Age (years)	31 (26-36)	30 (24-34)	31 (28-35)
Initial hCG (IU/L)	257 (87-742)	539 (280-982)	419 (198-1039)
48 hour hCG (IU/L)	112 (37-326)	1212 (610-2052)	518 (221-1238)
hCG ratio	0.40 (0.28-0.61)	2.18 (1.90-2.48)	1.18 (0.94-1.45)
Initial progesterone (nmol/l)	5 (3-10)	68 (52-91)	23 (12-43)
QCCH	N=552	N=283	N=132
Age (years)	33 (27-37)	31 (26-35)	33 (29-37)
Initial hCG (IU/L)	425 (129-1726)	712 (322-1747)	480 (188-1139)
48 hour hCG (IU/L)	180 (58-604)	1329 (635-2785)	550 (234-1409)
hCG ratio	0.37 (0.26-0.52)	2.03 (1.36-2.44)	1.20 (0.98-1.54)
Initial progesterone (nmol/l)	5 (2-10)	51 (31-70)	21 (11-32)

Table 5: Descriptive statistics presented as median (interquartile range) in the two participating hospitals. SGH = St Georges' Hospital; QCCH = Queen Charlottes' & Chelsea Hospital; PUL = pregnancy of unknown location; IUP = intrauterine pregnancy; hCG = human chorionic gonadotrophin. Reproduced from Van Calster/ Bobdiwala et al[40].

Of the remaining 2753 PUL, there were 1450 (52.7%) FPUL, 1002 (36.4%) IUP, and 301 (10.9%) EP. The EP rate was 9.5% at SGH and 13.7% at QCCH. The development dataset included data from the first 921 (52%) patients from SGH and the first 528 (55%) patients from QCCH (1449 in total). The remaining patients formed the validation data (1304 in total, 865 from SGH and 439 from QCCH), see figure 32.

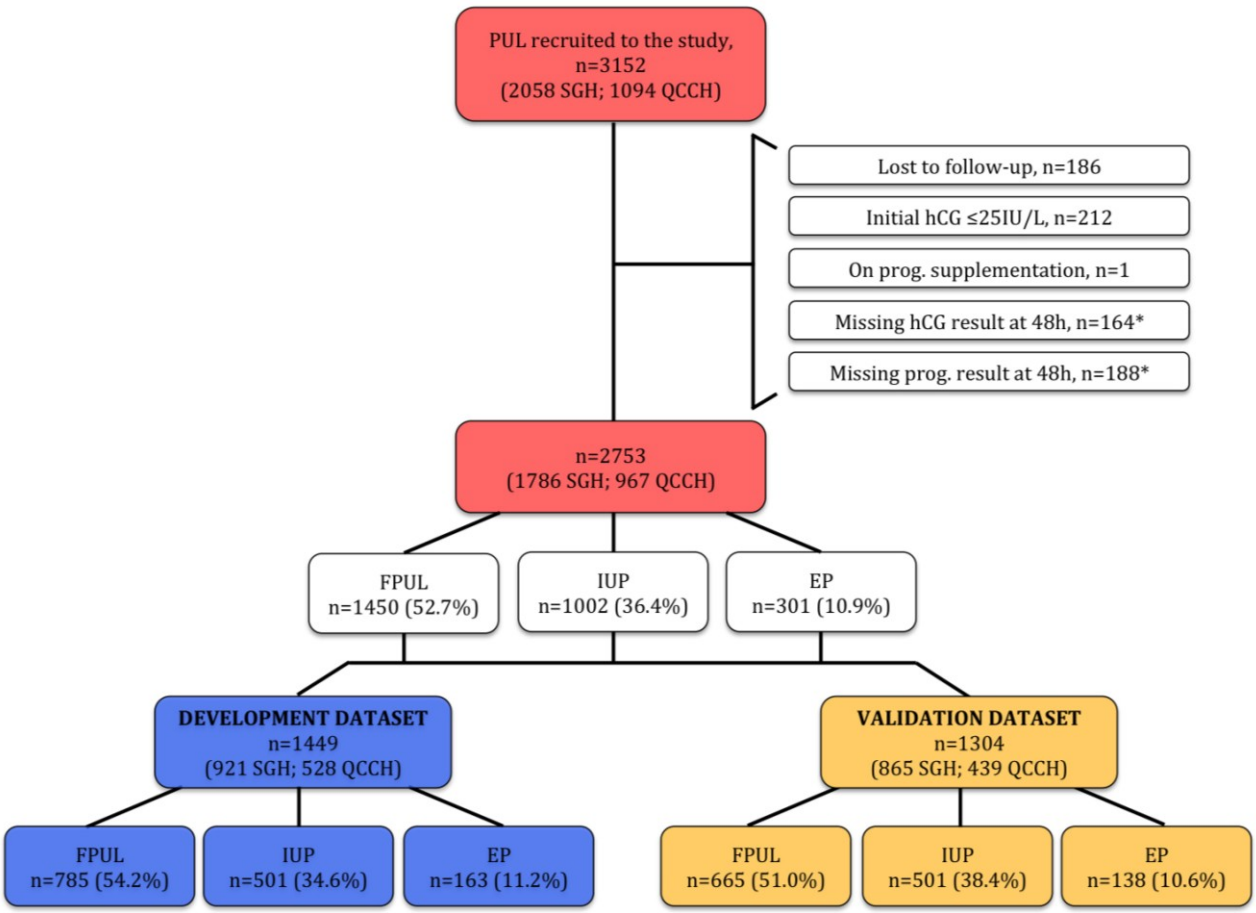


Figure 32: Flow diagram of patient flow through the study. SGH = St Georges' Hospital; QCCH = Queen Charlottes' & Chelsea Hospital; FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; EP = ectopic pregnancy; hCG = human chorionic gonadotrophin; prog. = progesterone.
* these values were multiply imputed back into the dataset

Development and validation of step 1 and step 2

Development of step 1 (progesterone cut-offs)

We evaluated the performance of initial progesterone cut-offs ranging from ≤2 to ≤10nmol/l on the development dataset (table 6; figure 33). Based on misclassifying as few ectopic pregnancies as possible at this first step, we selected a serum progesterone cut-off of ≤2nmol/l.

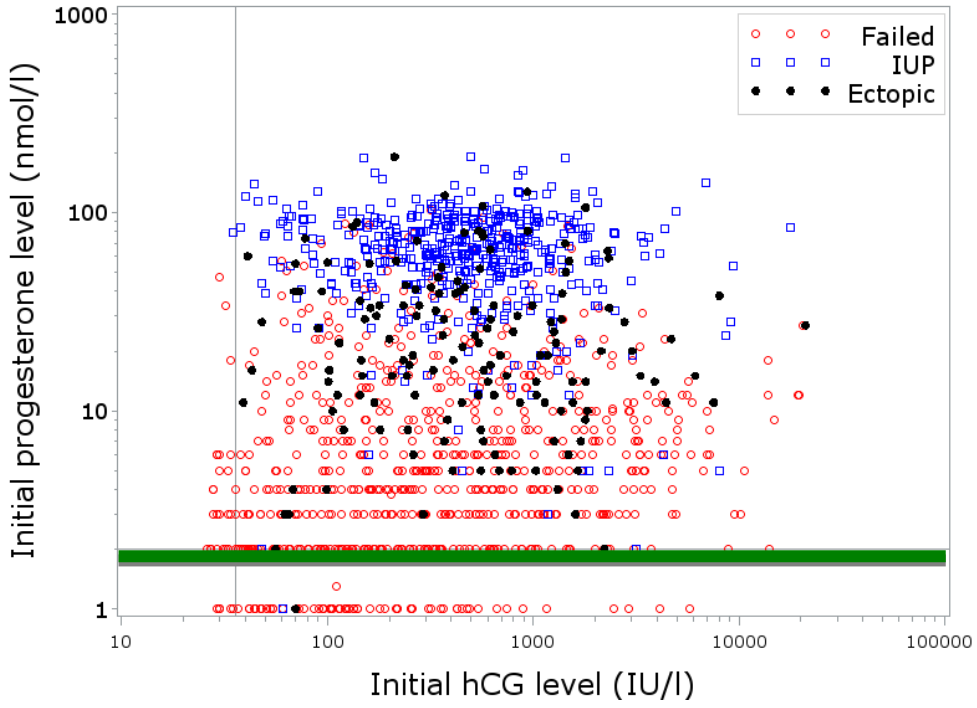
Validation of step 1

In the validation dataset, this cut-off classified 210/1304 (16.1%) patients as low-risk at the first visit, with a NPV of 98.1% (206/210) (table 7).

Progesterone cut-off to define low-risk (nmol/l)	PUL classified as low-risk	Negative predictive value	Sensitivity for EP	False positive rate
≤2	200/1449, **13.8%**	197/200, **98.5%**	160/163, **98.2%**	1089/1286, **84.7%**
≤3	293/1449, **20.2%**	286/293, **97.6%**	156/163, **95.7%**	1000/1286, **77.8%**
≤4	377/1449, **26.0%**	367/377, **97.3%**	153/163, **93.9%**	919/1286, **71.5%**
≤5	445/1449, **30.7%**	429/445, **96.4%**	147/163, **90.2%**	857/1286, **66.6%**
≤6	505/1449, **34.8%**	486/505, **96.2%**	144/163, **88.3%**	800/1286, **62.2%**
≤7	549/1449, **37.9%**	526/549, **95.8%**	140/163, **85.8%**	760/1286, **59.1%**
≤8	583/1449, **40.2%**	555/583, **95.3%**	136/163, **83.3%**	731/1286, **56.8%**
≤9	609/1449, **42.0%**	579/609, **95.2%**	134/163, **81.9%**	707/1286, **55.0%**
≤10	636/1449, **43.9%**	603/636, **94.9%**	130/163, **79.9%**	683/1286, **53.1%**

Table 6: Development data performance of different initial progesterone (measured in nmol/l) cut-offs to define low-risk at presentation. EP = ectopic pregnancy; PUL = pregnancy of unknown location; Prog = progesterone. Reproduced from Van Calster/ Bobdiwala et al[40].

(A)



(B)

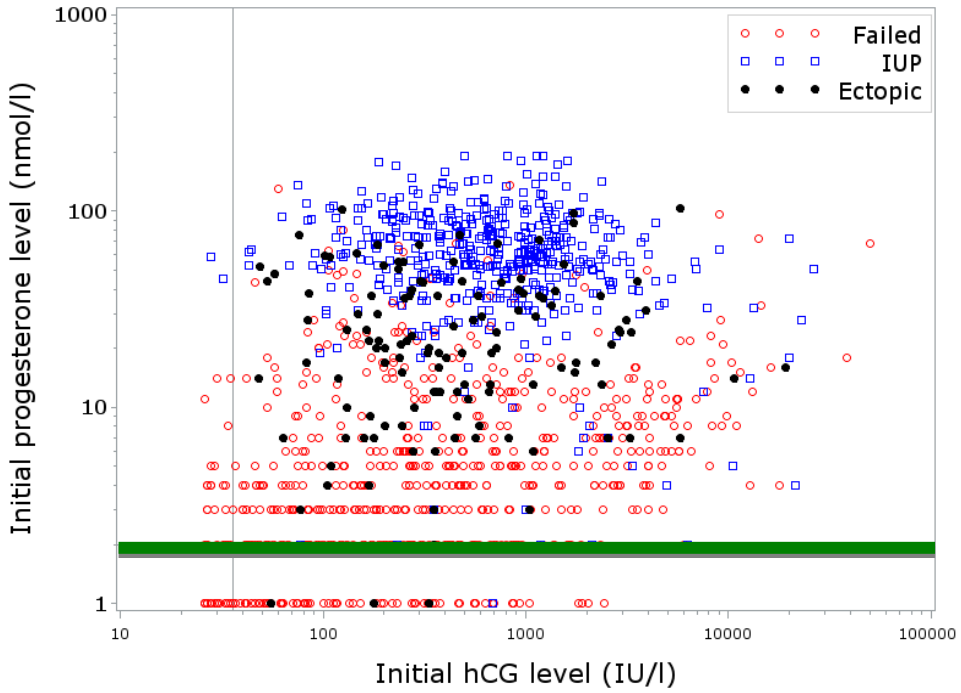


Figure 33: Scatterplot of initial serum progesterone level against initial serum human chorionic gonadotropin (hCG) level for (A) development dataset and (B) validation dataset for managing pregnancies of unknown location (PUL). A cut-off of ≤2 nmol/l is used for the initial progesterone level to classify PUL as low risk at step 1 of the two-step approach, indicated by a horizontal green line. Individual data points are plotted according to final pregnancy outcome as a failed PUL, intrauterine pregnancy or ectopic pregnancy. Reproduced from Van Calster/ Bobdiwala et al[40].

Development of step 2 ($M6_P$ (or $M6_{NP}$) models)

For the $M6_P$ ($_P$ = 'progesterone' included as a variable) sub-model, the following variables were used: log(initial hCG), log(initial progesterone), log(hCG ratio), log(hCG ratio) squared, and the interaction between log(hCG ratio) and log(initial progesterone). The $M6_{NP}$ ($_{NP}$ = 'no progesterone', i.e. progesterone was not included as a variable) sub-model does not include log(initial progesterone) and the interaction between log(initial progesterone) and log(hCG ratio) as variables. See appendix 5 for further details.

Validation of step 2

The $M6_P$ model classified 60.5% PUL (789/1304) as low-risk with a negative predictive value (NPV) of 99.1% (782/789), sensitivity for EP of 94.9% EP (131/138) (i.e. 7 EP misclassified as low-risk) and a false positive rate (FPR) of 32.9% (385/1166) (table 7). The AUC for EP vs. other PUL outcomes was 0.903 (95% CI 0.880 - 0.922) (figure 34).

The $M6_{NP}$ model classified 54.5% PUL as low-risk with a NPV of 98.6%, sensitivity of 92.5% (128/138), and a FPR of 39.9%. The AUC for EP vs. other PUL outcomes was 0.870 (95% CI 0.837 - 0.897).

The two-step protocol: overall test performance

When combining steps 1 and 2, the two-step protocol classified 62.1% PUL as low-risk (810/1304) with a NPV of 98.6% (799/810), sensitivity of 92.0% (127/138) (i.e. 11 EP were misclassified as low-risk) and a FPR of 31.5% (367/1166) (table 7).

If the $M6_{NP}$ model was used as the second step (for women who are on progesterone supplementation), 57.7% PUL were classified as low-risk with a NPV of 98.1%, sensitivity of 89.6% and a FPR of 36.6%.

Data **Classification approach**	**PUL classified as low-risk**	**Negative predictive value**	**Sensitivity for EP**	**False positive rate**
Development dataset				
Step 1 only: Progesterone cut-off ≤2 nmol/l	200/1449, **13.8%**	197/200, **98.5%**	160/163, **98.2%**	1089/1286, **84.7%**
Step 2 only: $M6_P$ model	880/1449, **60.8%**	871/880, **99.0%**	154/163, **94.4%**	415/1286, **32.3%**
Step 2 only: $M6_{NP}$ model	762/1449, **52.6%**	751/762, **98.6%**	152/163, **93.4%**	535/1286, **41.6%**
Combined two-step protocol: Step 1 + $M6_P$ model	896/1449, **61.9%**	884/896, **98.7%**	151/163, **92.6%**	402/1286, **31.2%**
Combined two-step protocol: Step 1 + $M6_{NP}$ model	797/1449, **55.0%**	783/797, **98.3%**	149/163, **91.7%**	503/1286, **39.1%**
Validation dataset				
Step 1 only: Progesterone cut-off ≤2 nmol/l	210/1304, **16.1%** (14.2-18.2)	206/210, **98.1%** (95.0-99.3)	134/138, **97.1%** (92.4-98.9)	960/1166, **82.3%** (80.0-84.5)
Step 2 only: $M6_P$ model	789/1304, **60.5%** (57.1-63.8)	782/789, **99.1%** (98.1-99.6)	131/138, **94.9%** (89.4-97.6)	384/1166, **32.9%** (29.5-36.5)
Step 2 only: $M6_{NP}$ model	716/1304, **54.5%** (49.8-59.2)	706/716, **98.6%** (97.3-99.2)	128/138, **92.5%** (86.4-96.1)	460/1166, **39.9%** (35.0-45.1)
Combined two-step protocol: Step 1 + $M6_P$ model	810/1304, **62.1%** (58.8-65.3)	799/810, **98.6%** (97.5-99.3)	127/138, **92.0%** (85.9-95.6)	367/1166, **31.5%** (28.1-35.0)
Combined two-step protocol: Step 1 + $M6_{NP}$ model	754/1304, **57.7%** (53.2-62.1)	740/754, **98.1%** (96.8-98.9)	124/138, **89.6%** (83.0-93.9)	426/1166, **36.6%** (32.0-41.5)
Comparison to other protocols				
M4 model	921/1304, **70.6%** (68.0-73.1)	895/921, **97.2%** (95.9-98.1)	112/138, **81.4%** (73.9-87.2)	271/1166, **23.2%** (20.8-25.8)
Single visit progesterone cut-off ≤10nmol/l	572/1304, **43.8%** (41.1-46.6)	542/572, **94.7%** (92.5-96.3)	108/138, **78.1%** (70.3-84.3)	625/1166, **53.6%** (50.7-56.5)

Table 7: Performance of PUL classification approaches. Confidence intervals are given between parentheses. EP = ectopic pregnancy; PUL = pregnancy of unknown location; Prog = progesterone. Adapted from Van Calster/ Bobdiwala et al[40].

Comparison to other PUL triage protocols

The M4 model

The M4 model classified 70.6% PUL as low-risk with a NPV of 97.2%, sensitivity for EP of 81.4% and a FPR of 23.2% (table 7). The AUC for EP vs. other PUL outcomes was 0.847 (95% CI 0.808 - 0.880) (figure 34), i.e. 0.056 lower than the $M6_P$ model (95% CI 0.029 - 0.083) and 0.023 lower than the AUC for the $M6_{NP}$ model (95% CI 0.009 - 0.036).

Single visit progesterone protocol (cut-off ≤10 nmol/l)

The existing single visit progesterone protocol (≤10nmol/l indicating a low-risk PUL) classified 43.8% PUL as low-risk with a NPV of 94.7%, sensitivity for EP of 78.1% and a FPR of 53.6% (table 7).

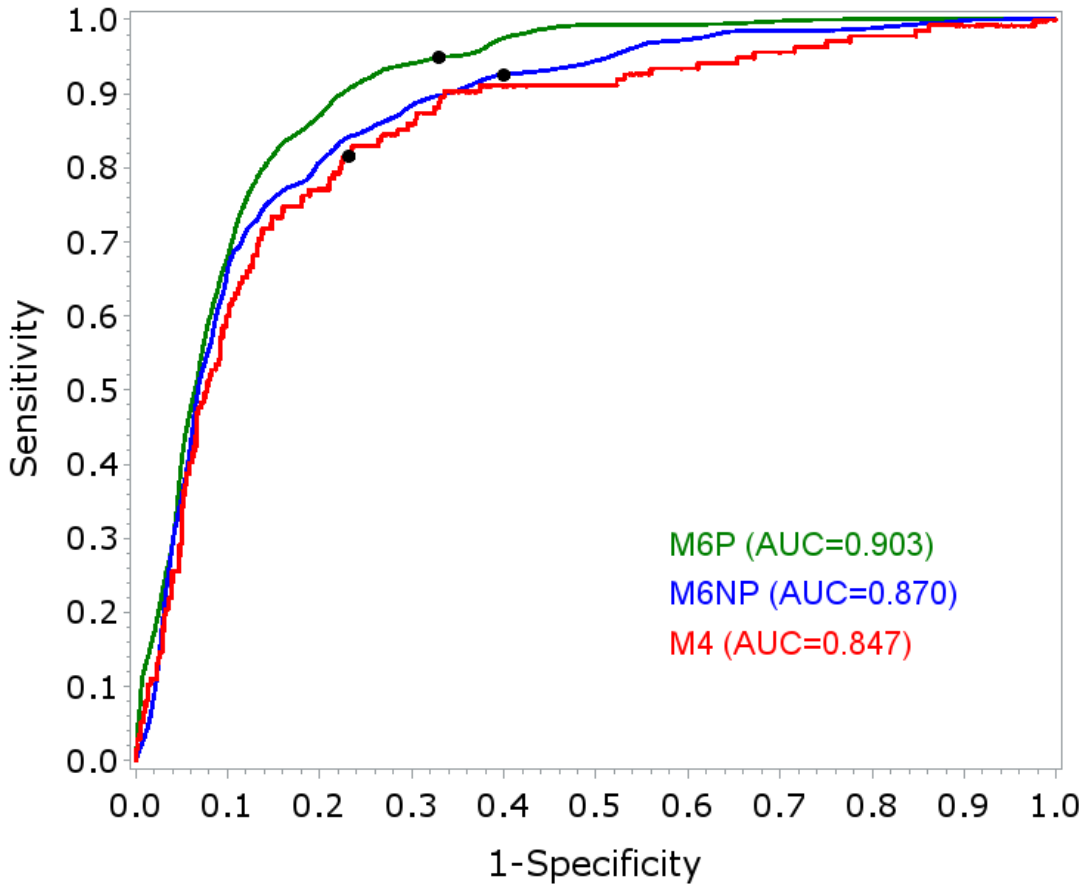


Figure 34: ROC curves for discriminating EP from FPUL/IUP for the M6P, M6NP and M4 models. The performance at the risk threshold of 5% is indicated by black dots. Reproduced from supplementary material published by Van Calster/ Bobdiwala et al[40] .

Discussion

The management of women with a PUL currently relies on either the use of single progesterone levels or hCG. Levels of hCG are most commonly interpreted in the context of the hCG ratio or risk prediction models. Both biomarkers have their benefits and drawbacks and have not been evaluated in conjunction. We aimed to combine both of these biomarkers into a two-step protocol and compare the test performance of this protocol against existing protocols. In this study, I demonstrated that the two-step protocol classifies the majority of women with a PUL as low-risk or high-risk with a high level of confidence and outperforms other published management strategies for PUL, including the M4 model.

Step 1 relies on a progesterone cut-off of ≤2 nmol/l and step 2 on a logistic regression model called the M6 model based on the hCG ratio +/- the initial progesterone level.

Step 1

I selected a much lower initial progesterone cut-off of ≤2 nmol/l (compared to the previously published cut-off of ≤10 nmol/l[38, 39]) as I wanted to minimize the number of missed EP at this first step (2.8% versus 20.1% if a cut-off of ≤10 nmol is used).

Despite such a low cut-off, a progesterone of ≤2 nmol/l still enables the discharge of around 1 in 6 PUL after that first consultation with just a follow-up urinary pregnancy test at home in two weeks to confirm resolution of the pregnancy. Being able to discharge such a significant proportion of patients with such a high level of confidence after just one hospital attendance could confer a financial benefit to the NHS and reduce unnecessary follow-up with its' associated anxiety for women. The follow-up urinary pregnancy test provides a safety net for women, confirmation of the correct diagnosis if the test is negative and prompts repeat follow-up if the test is positive and an EP may have been missed. Women should of course also be counselled

that they should contact the hospital/ EPU should they have any increase in symptoms or concerns.

There is an argument to suggest that at low progesterone levels (below 10 nmol/l), even if an EP is missed, there is a high likelihood that it will resolve spontaneously and therefore misclassification does not lead to clinical harm or a delay in intervention. There is, however, currently insufficient data to confirm that this misclassification is safe. Due to the lack of safety data, I used 2 nmol/l as the cut-off as I believe this reaches a workable balance between the size of the group selected for reduced follow-up (1 in 6 in this study) and the number of EP that are misclassified into this group.

Step 2

The second step uses either the $M6_P$ or $M6_{NP}$ models, which can also be used in isolation (i.e. without step 1). I have demonstrated that they have superior test performance and provide more accurate estimates of risk than the M4 model[22], the most widely published risk prediction model for managing women with a PUL and single-step protocols such as a serum progesterone cut-off value of ≤10 nmol/l[38, 39] (table 7). In fact, using the $M6_P$ model in isolation had higher sensitivity for classifying EP as high-risk than the two-step protocol (94.9% versus 92.0%). Including step 1, however, has the benefit of enabling the discharge of 16% women with a PUL after just one hospital attendance.

Strengths of the study

A key strength of the study is the large sample size (n=2753) and the inclusion of both development and temporal validation datasets. The study also collates data from two centres. This increases the potential applicability of the model(s) to other populations. Furthermore, the two-step protocol is based solely on measurements of progesterone and hCG (objective

numerical variables) rather than more subjective information (such as the amount of pain or bleeding experienced). This should also increase generalizability to other PUL populations.

Limitations of the study

Serum progesterone level results were missing in 188 (6.5%) and serum hCG levels in 164 (5.7%) cases. Multiple imputation was used to deal with these missing values, which is in fact statistically recommended over the use of complete case analysis[221]. Another limitation is that progesterone levels cannot be accurately interpreted in women taking progesterone supplementation, as they tend to have artificially raised serum levels. We could not perform a sub-analysis to investigate M6 model performance in women taking progesterone supplementation as progesterone is not given as routinely in the UK as it is in other European countries. In fact, there was only one patient known to be taking progesterone supplementation in our patient cohort. This case was excluded from the final analysis (figure 32). A low serum progesterone level at the first visit (step 1) can still be used to indicate a failing pregnancy but progesterone cannot be incorporated into a risk prediction model for these cases. In view of this, we developed the $M6_{NP}$ model, which does not use progesterone as a variable. Test performance of the model is better when progesterone is included as a variable (table 7) although using the $M6_{NP}$ model is only likely to apply to a small proportion of patients.

Final conclusions and further work

The two-step protocol for managing PUL is an effective strategy for selecting low-risk cases where follow-up can be reduced to a minimum and direct resources to those PUL at highest risk of having an EP. It also outperforms all other management strategies currently used in clinical practice, including the previously trialled M4 model. This protocol, in an evidence-based way, simplifies the interpretation of serum hormone levels used to manage PUL. It could improve the care given to these women, particularly by less experienced/ junior staff by allowing more consistent decision-making and reducing unnecessary visits and interventions.

The next stage was to then implement this two-step protocol into everyday clinical practice and assess its' 'real-world' test performance and clinical safety. I therefore set out to assess the two-step protocol within the setting of a clinical intervention trial.

STUDY 4

A CLINICAL INTERVENTION STUDY USING A TWO-STEP TRIAGE PROTOCOL INCLUDING THE M6 LOGISTIC REGRESSION MODEL TO MANAGE WOMEN WITH A PUL

Chapter Abstract

Introduction

The M6 model was published as part of a two-step protocol where step 1 used an initial progesterone cut-off of ≤2nmol/l to identify failing pregnancies and step 2 used the M6 model. The aims of this study were to validate the triage performance of this protocol in clinical practice.

Methods

This was a prospective multi-centre interventional study of 3272 women with a PUL was carried out between January 2015 and January 2017 in four district general hospitals and four university teaching hospitals within the United Kingdom. If the initial progesterone was ≤2nmol/l, patients were discharged with a follow-up urine pregnancy test in two weeks. If the progesterone was >2nmol/l or had not been taken, a 48 hour hCG level was taken and results entered into the M6 model. Patients were managed according to their predicted outcome: those classified with pregnancies likely to resolve (FPUL) were advised to perform a urine pregnancy test in two weeks and those with a likely IUP were asked to attend for a scan a week later. Women with a PUL classified as high-risk (i.e. those with a risk of EP ≥ 5%) were reviewed clinically within 48 hours.

Results

Of 3272 women with a PUL, 317 met the exclusion criteria, were lost to follow-up or were recruited in a unit that used a progesterone cut-off of ≤10nmol/l and so were analysed separately. Progesterone results were available for 2392 (91%). 407 (15.5%) patients were classified as low-risk at step 1 (progesterone ≤2nmol/l), of which 7 (1.7%) were ultimately diagnosed with an EP. 1038 women with a PUL were classified as low-risk at step 2, of which 8 (0.8%) were EP. 275/320 (85.9%) EP were correctly classified as high-risk. Overall, 1445/2625

PUL (55.0%) were classified as low-risk, of which 15 (1.0%) were EP. None of these 15 EP ruptured or came to significant clinical harm.

Conclusions

This study has shown that this two-step protocol effectively triaged 55% of women with a PUL as being at low-risk of an EP, minimizing the follow-up required for these patients after just two visits. There were few misclassified EP and none of these patients suffered a serious adverse clinical event. This two-step protocol is an effective and clinically safe way of rationalizing the management of women with a PUL.

Introduction: the two-step protocol, which incorporates the M6 model

As demonstrated in study 3, the M6 model has been shown to be superior to other management protocols in risk-stratifying women with a PUL as being at low or high-risk of harbouring an ectopic pregnancy[40]. M6 was published as part of a two-step protocol where:

- *Step 1* used an initial progesterone cut-off of ≤2nmol/l to identify failing pregnancies
- *Step 2* used the M6 model

The aims of this study were to validate the triage performance of this protocol in clinical practice by assessing:

(1) the number of protocol-related adverse events

(2) if patients were correctly triaged

Methods

Design and settings

This was a UK-based prospective multi-centre interventional study carried out on consecutive patients classified as having a PUL at their initial visit to a dedicated EPU in one of eight participating hospitals. Four were university teaching hospitals: Queen Charlottes' and Chelsea Hospital (QCCH), St. Marys' Hospital (SMH), Chelsea and Westminster Hospital (CWH) and West Middlesex University Hospital (WMUH). The remaining four EPUs were in district general hospitals: Hillingdon Hospital (HH), North Middlesex Hospital (NMH), Wexham Park Hospital (WXP) and Royal Surrey Hospital (RSH).

This study was registered as an audit, after consultation with a Research Ethics Committee (REC) and the research and development departments within each participating centre established that formal ethical approval and written consent was not required from patients. This was because this was validating a protocol that had already been internally validated and undergone publication[40] after peer-review.

Inclusion and exclusion criteria

As outlined in the methods chapter, page 60.

Sample collection and analysis

As outlined in the methods chapter, page 69.

Triage using the two-step protocol

The two-step protocol was embedded into a protected Microsoft excel document that enabled data collection and risk prediction at each participating centre. Once classified as a PUL, patients had initial serum hCG and progesterone levels taken. These results were entered onto the excel data collection sheet. The two-step approach used to manage PUL is shown below (figure 35).

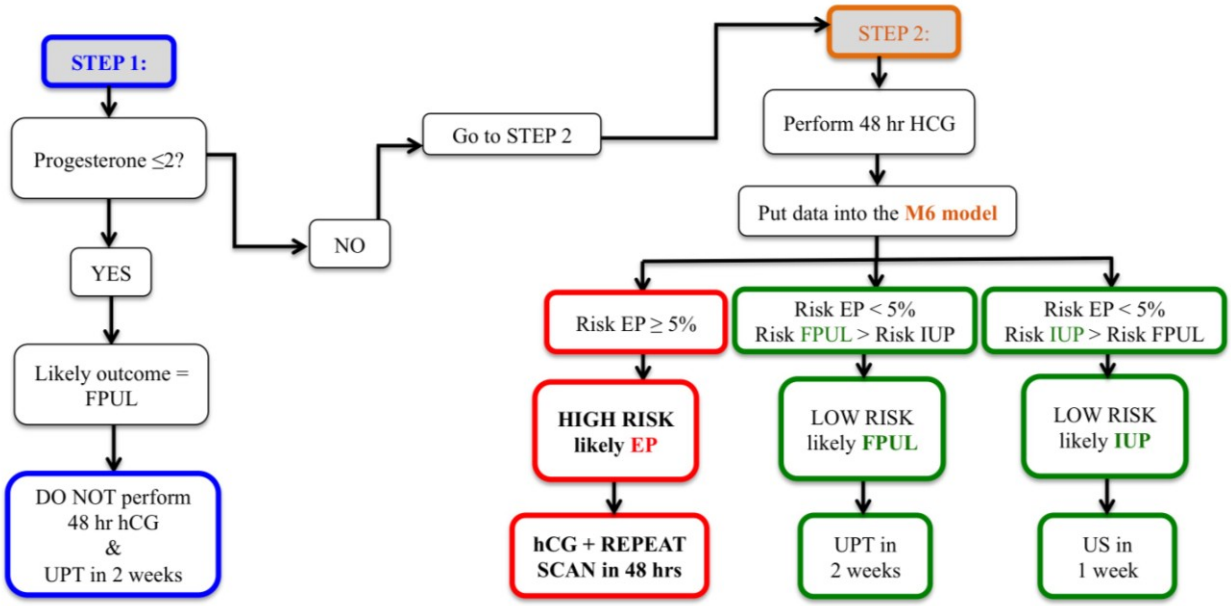


Figure 35: Flow diagram of two-step approach for managing pregnancies of unknown location (PUL). EP, ectopic pregnancy; FPUL, failed PUL; hCG, human chorionic gonadotropin; IUP, intrauterine pregnancy; UPT, urine pregnancy test; US, ultrasound examination. Adapted from Van Calster/ Bobdiwala et al[40].

Step 1: If the initial progesterone result was ≤2nmol/l, the protocol predicted a final outcome of FPUL and advised the patient could be discharged and asked to perform a urine pregnancy test two weeks later to ensure the result was negative. There was a telephone follow-up service at each hospital and patients were contacted to provide reassurance and counselling, as well as confirm the urine pregnancy test result. If the result was positive, women were asked to re-attend the EPU for clinical review. If measuring serum progesterone levels was not part of a units' routine practice, they were advised they could move straight to step 2.

Step 2: If the progesterone result was >2nmol/l or had not been taken, women were advised to re-attend 48 hours later for a repeat serum hCG blood test. Units were advised to not use the second hCG result in step 2 if it had been taken 8 hours before or after the 48 hour time point as interpretation may then be unreliable. The initial hCG and progesterone and 48 hours hCG results were then entered onto the excel datasheet which contained an embedded algorithm for

the M6 model. This provided an instant risk prediction and recommended management plan to the clinician using it.

The M6 model classified a PUL case as 'high-risk' (likely diagnosis of EP/ PPUL) if the predicted risk of EP was ≥5%[22]. If the risk was <5%, the predicted final outcome was either likely 'low-risk, probable FPUL' or 'low-risk, probable IUP', dependent on which of the two outcomes was most likely. Those predicted as 'low-risk probable FPUL' were advised to perform a urine pregnancy test two weeks later to confirm a negative result. If it was positive, the patient was brought back for review with a senior clinician and a repeat serum hCG and/or TVS. Patients predicted 'low-risk probable IUP' had a repeat TVS to confirm the presence of an IUP after one week. Patients predicted to be 'high-risk, probable EP' were advised to attend the EPU within 48 hours for senior clinical review and a repeat TVS (figure 35). If the pregnancy could still not be located, a repeat serum hCG was taken and dependent on the clinical situation, an individualized plan was made. The rationale for the clinical pathway summarised in figure 35 was derived from previous publications by our research group[21, 41] and the data analysed in studies 2 and 3, which had not highlighted any problems or safety issues with these intervals of follow-up.

At each visit, women were advised to seek a medical opinion promptly if they had any concerns or worsening of their symptoms. Clinicians were informed that the two-step protocol was a guide, offering advice on how to manage women with a PUL and should not be followed if the clinical situation warranted a different clinical approach. Clinicians were also advised that the M6 model should be used for guidance on the intensity of follow-up and not used to determine whether medical/ surgical intervention was necessary – this responsibility always rested with the managing clinician.

Reference standard: pregnancy outcome definitions

All final outcomes of the pregnancy were defined at the time point the pregnancy would be expected to be 12 weeks in gestation. The definitions used for final outcomes are as described in the *methods* chapter, pages 65-8.

Deviations from protocol

As outlined in study 2, page 99.

Adverse events

As outlined in study 2, page 100.

Statistical Analysis

Descriptive statistics were used to analyze data, counts and percentages for categorical variables and medians and ranges for continuous variables. 95% confidence intervals (CI) were calculated using Wilson's score method. We did not impute missing data so that we could more accurately describe what actually happened when the protocol was used in 'real-time' clinical practice. The statistical package R (version 3.5.2) was used.

Results

Patient flow

A total of 3272 women from the eight participating centres were classified as a PUL during the study period (N per centre ranged between 202 and 551), see table 8.

	Recruited	Excluded		Included			
Centre	**N**	**Exclusion criteria met, N**	**Lost to follow-up, N**	**All, N (%)[a]**	**FPUL, N (%)[b]**	**IUP, N (%)[b]**	**EP/PPUL, N (%)[b]**
St. Mary's Hospital	551	1	66	484 (88)	235 (49)	177 (37)	72 (15)
Chelsea & Westminster Hospital	541	1	30	510 (94)	292 (57)	164 (32)	54 (11)
Hillingdon Hospital	534	0	64	470 (88)	288 (61)	138 (29)	44 (9)
Queen Charlotte's & Chelsea Hospital	510	0	46	464 (91)	238 (51)	168 (36)	58 (13)
Wexham Park	351	0	57	294 (84)	142 (48)	108 (37)	44 (15)
West Middlesex University Hospital	340	0	10	330 (97)	250 (76)	64 (19)	16 (5)
North Middlesex Hospital	243	4	14	225 (93)	142 (63)	56 (25)	27 (12)
Royal Surrey Hospital	202	0	24	178 (88)	96 (54)	61 (34)	21 (12)
Total	**3272**	**6**	**311**	**2955 (90)**	**1433 (55)**	**872 (33)**	**320 (12)**

[a] Percentage of recruited patients. [b] Percentage of included patients.

Table 8: Patient recruitment by each participating centre

Six participants met the exclusion criteria and were omitted from the final analysis, while 311 (10%) were lost to follow-up (figure 36). One centre (WMUH) used a progesterone cut-off of ≤10nmol/l and was analysed separately (see appendix 6). This left 2625 patients for the main analysis. Descriptive data for these patients is presented in table 9.

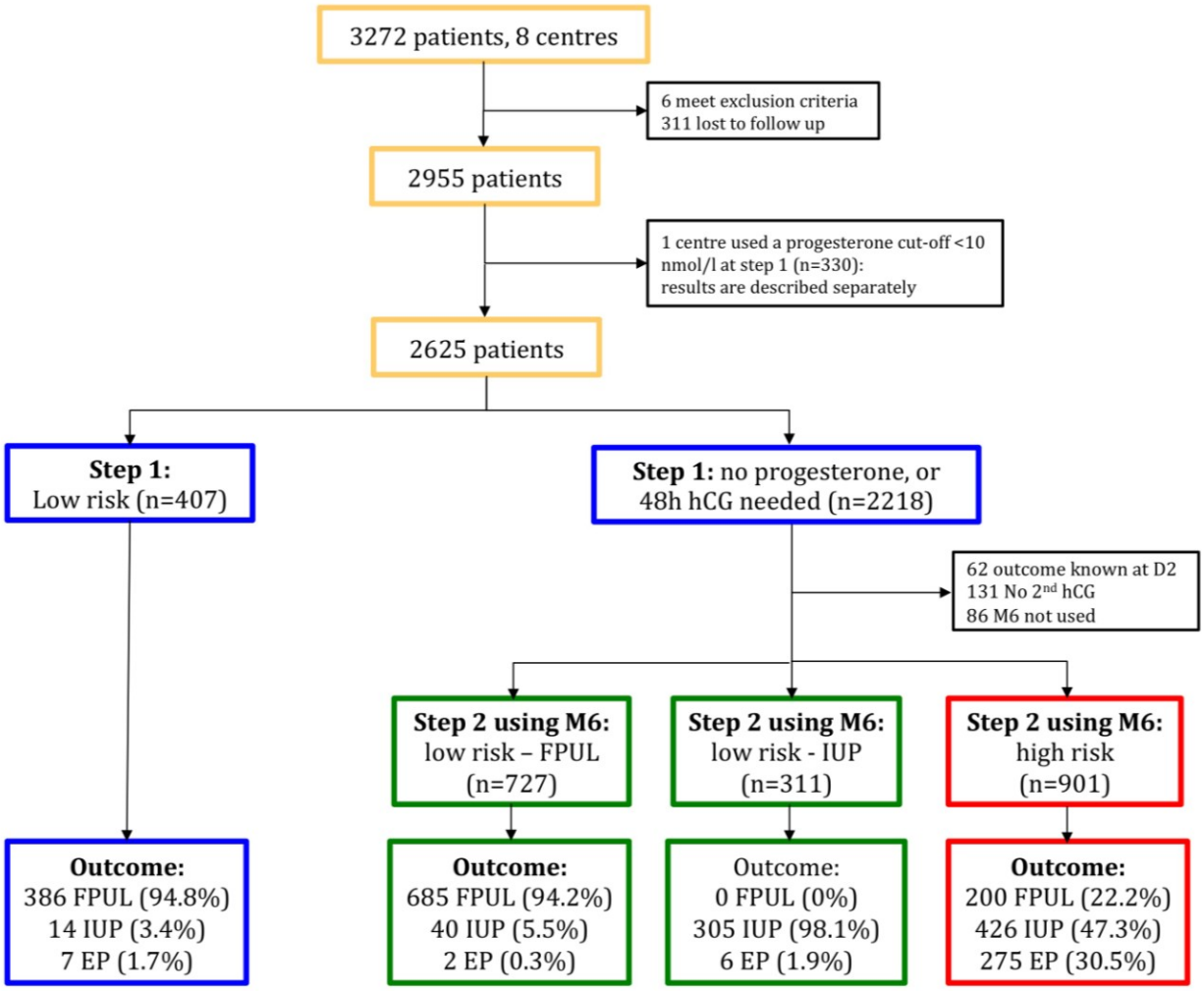


Figure 36: Flowchart of patient recruitment, final outcome data and correct versus incorrect risk stratification by the two-step protocol. FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; EP = ectopic pregnancy; hCG = human chorionic gonadotrophin; prog = progesterone

Patient characteristic	Median (IQR, range), or N (%)	Missing values, N (%)
Age	32 (27-36, 14-50)	2 (0.1%)
Initial progesterone (nmol/L)	10 (4-38, 0.3-219)	291 (11.1%)*
Initial serum hCG (IU/L)	407 (92-1257, 0.1-105006)	0 (0%)
2nd serum hCG (IU/L)$	496 (121-1345, 0.5-109658)	500 (19.0%)#
2nd serum hCG if at D2 (IU/L)	515 (126-1361, 0.8-109658)	511 (19.5%)
hCG ratio$	0.88 (0.34-1.91, 0.002-25.0)	500 (19.0%)#
hCG ratio if 2nd hCG at D2	0.96 (0.36-1.95, 0.01-6.31)	511 (19.5%)
Pain score (0-10)	3 (0-4, 0-10)	113 (4.3%)
Vaginal bleeding		9 (0.3%)
No (0)	688 (26%)	
Minimal (1)	728 (28%)	
Moderate (2)	616 (24%)	
Soaked (3)	295 (11%)	
Clots (4)	289 (11%)	
History of EP		7 (0.3%)
None	2446 (93%)	
Any	172 (7%)	
Laparoscopy	124 (5%)	
MTX	26 (1%)	
Expectant	22 (1%)	
Indication for scan		3 (0.1%)
Bleeding and pain	1212 (46%)	
Bleeding only	705 (27%)	
Pain only	485 (18%)	
Maternal reassurance	72 (3%)	
Unsure dates	65 (2%)	
Previous ectopic	42 (2%)	
Previous miscarriage	41 (2%)	
Type of PUL◆		5 (0.2%)
True PUL	1323 (50%)	
Probable miscarriage	831 (32%)	
Probable intra-uterine	338 (13%)	
Probable ectopic	128 (5%)	

$ This describes any 2nd hCG level recorded in the database, irrespective of interval with the initial measurement and irrespective of whether this 2nd level was used to calculate the M6 result

* This includes known progesterone levels from cases who were on progesterone supplements

This includes cases where step 1 triage was 'low-risk' and therefore a 2nd hCG was not indicated

◆ As presumed by the sonographer at the initial ultrasound scan

Table 9: Pooled descriptive statistics from all participating early pregnancy units, n=2625. PUL = pregnancy of unknown location; hCG = human chorionic gonadotrophin; MTX = methotrexate; IQR = interquartile range

Two-step protocol: test performance

Step 1 PUL prediction

233 participants did not have an initial progesterone level taken. One centre (HH, n=534), had a lab assay for serum progesterone with a lower limit of 5 nmol/l so all patients from this centre automatically required triage by step 2. Step 1 of the protocol was therefore applied to 2392 PUL. 407/2392 (15.5% of 2625, 95% CI 14.2-16.9) women with a PUL were classified as low-risk at step 1. 386 (94.8%) had a final outcome of FPUL, 14 (3.4%) an IUP and 7 (1.7%, 95% CI 0.8-3.5) an EP.

Step 2 PUL prediction

2218 women with a PUL underwent a 48 hour hCG blood test for step 2 (the M6 model): 233 without initial progesterone and 1985 where step 1 recommended M6 model use. In 62 cases, the PUL outcome was known by the time M6 should be implemented whereas for 217 cases step 2 was not applied, usually secondary to a clinical deviation. Of the remaining 1038 women with a PUL, 727 were classified as 'low-risk, likely FPUL', of which 685 (94.2%) were FPUL, 40 (5.5%) IUP and 2 (0.3%, 95% CI 0.1-1.0) EP. 311/1038 women with a PUL were classified as 'low-risk, likely IUP', of which 305 (98.1%) IUP and 6 (1.9%, 95% CI 0.9-4.1) EP. 901/1038 women with a PUL were classified as 'high-risk', of which 200 (22.2%) FPUL, 426 (47.3%) IUP and 275 (30.5%, 95% CI 27.6-33.6) EP. The triage results by final pregnancy outcome are summarized in table 10.

Final classification		Outcome		
	N	FPUL N (%)	IUP N (%)	EP N (%)
Low-risk				
at Step 1 (probable FPUL)	407 (16%)	386 (27%)	14 (2%)	7 (2%)
at Step 2 (all)	1038 (40%)	685 (48%)	345 (40%)	8 (3%)
Probable FPUL	727	685	40	2
Probable IUP	311	0	305	6
High risk (probable EP)	901 (34%)	200 (14%)	426 (49%)	275 (86%)
Outcome known at D2	62 (2%)	1 (<1%)	47 (5%)	14 (4%)
Not classified	217 (8%)	161 (11%)	40 (5%)	16 (5%)
Total	**2625**	**1433**	**872**	**320**

Table 10: Overview of patient triage, overall and by final PUL outcome. D2 = day two. FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; EP = ectopic pregnancy

EP/ PPUL and the two-step protocol

320/2625 cases of PUL had a final outcome of EP. 275 (85.9%, 95% CI 81.7-89.3) were correctly classified as high-risk at step 2. 7 (2.2%, 95% CI 1.1-4.4) EP were classified as low-risk at step 1 and 8 (2.5%, 95% CI 1.3-4.9) as low-risk at step 2. 14 EP were visualized on ultrasonography at 48 hours (i.e. when M6 should have been used) and for 16 EP the M6 model was not used. Therefore, where the M6 model was used, 275/283 (97.2%, 95% CI 94.5-98.6) EP were classified as high-risk. Of the 275 correctly classified EP, 99 (36.0%) underwent successful expectant management, 71 (25.8%) were successfully managed medically with methotrexate and 100 (36.4%) were had surgical treatment via laparoscopy. Five (1.8%) cases underwent other interventions such as uterine dilatation and curettage.

In 9/16 EP where step 2 was not used, results were absent due to a technical error. When applying the triage protocol retrospectively, all 9 cases would have been classified as high-risk. In 4/16, second hCG results were not used by the investigator because they were taken 5-8 days after the initial hCG. In 3/16, a second hCG result was not available.

Step 1 misclassifications:

Seven EP were misclassified at step 1 (all as an assumed FPUL), of which 4 patients underwent successful expectant management and 3 had treatment with methotrexate. Three patients were identified as they had a positive urine pregnancy test two weeks later and the other four cases were brought back due to the individual clinical situation (1 patient had previously had hCG readings in the IVF unit which had been suboptimal, for 2 cases there was a suspicion of an adnexal mass on the initial TVS and 1 case had ongoing PVB and pain). None of these cases resulted in a ruptured EP or came to any significant clinical harm. See table 11 for individual case details.

Step 2 misclassifications:

Eight EP were misclassified at step 2. Two patients were incorrectly classified as FPUL and both underwent successful expectant management. Both these EP were identified as they were bought back by the clinician for an unscheduled TVS as there had been previously suspicion of an adnexal mass. Six patients were misclassified as an IUP. 1 patient was managed expectantly and had her EP identified as the managing clinician organized an unscheduled TVS and the EP was seen on a repeat scan on day 4. Five patients required surgical intervention (laparoscopy) and the EP was diagnosed as the patient either re-presented with symptoms or a repeat ultrasound scan was carried out against protocol as the managing clinician felt the initial hCG was too high. One of these cases was an interstitial EP that underwent surgical resection but was misclassified as low-risk as the hCG levels were 60 hours apart. This EP was diagnosed due to an initial clinical suspicion of EP on the first TVS. None of these cases resulted in a ruptured EP or came to significant clinical harm. See table 11 for individual case details.

Patient	0h prog	0h hCG	2nd hCG$	hCG ratio	When misclassified	Misclassified as	Intervention
1	2	237	70	0.30	Step 1	FPUL	Expectant
2	1*	348	419	1.20	Step 1	FPUL	Expectant
3	2	1166	1058	0.91	Step 1	FPUL	Expectant
4	1	85	NA	NA	Step 1	FPUL	Expectant
5	1	62	89	1.44	Step 1	FPUL	Methotrexate
6	1	37	46	1.24	Step 1	FPUL	Methotrexate
7	1	167	NA	NA	Step 1	FPUL	Methotrexate
8	16.7	690	251	0.36	Step 2	FPUL	Expectant
9	4.8	82.7	37.3	0.45	Step 2	FPUL	Expectant
10	NA	552.1	863.5	1.56	Step 2	IUP	Expectant
11	NA	187	491	2.63	Step 2	IUP	Laparoscopy
12	73	852	2497	2.93	Step 2	IUP	Laparoscopy
13	86.7	1480	2721	1.84	Step 2	IUP	Laparoscopy
14	304*	145	368	2.54	Step 2	IUP	Laparoscopy
15	75	142	309	2.18	Step 2	IUP	Resection

$ 2nd hCG, irrespective of interval

* Patient on progesterone supplementation

For this patient, a wrong hCG ratio was used in M6 (technical error with the excel spreadsheet); if the correct ratio is used, M6 classifies the patient as high risk

Table 11: Patient-specific information of women with an ectopic pregnancy that were misclassified by the two-step algorithm (n=15). *NB cases from WMUH are not included in this analysis as they used a different step 1 cut-off of ≤10nmol/l rather than ≤2nmol/l, which was used by all other participating units.* hCG = human chorionic gonadotrophin (IU/L); h= hour; FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; NA = not applicable; prog = progesterone (nmol/l)

Protocol deviations

This section relates to cases where the management recommended by the two-step protocol was not adhered to (see table 12). In 1918/2625 (73%) cases, clinicians followed the recommended management and there was no protocol deviation.

Deviations secondary to a clinician reason

328 (13%) had a protocol deviation for a 'clinician reason' (276 were non-technical and the remaining 52 related to a technical error). A technical deviation relates to incorrect data entry or a software malfunction causing an incorrect risk prediction by the M6 model. A non-technical clinician deviation relates to the clinician making a judgment that they would not adhere to the management recommended by the M6 model. This included reasons such as an initial hCG which was felt to warrant earlier repeat TVS and suspicious features on the initial TVS such as free fluid or a potential adnexal mass. This sometimes led to unnecessary interventions (e.g. a negative laparoscopy).

Deviations secondary to a patient reason

There were 92 (4%) deviations secondary to a 'patient reason'.

Deviations secondary to a timing reason

There were 287 (11%) deviations secondary to a 'timing reason'. The vast majority of 'timing deviations' were because the second hCG level was not taken +/-8 hours of 48 hours.

Protocol deviation	n (%)
No protocol deviation	1918 (73.1%)
Clinician reason - non-technical - technical	328 (12.5%) - 276 (10.5%) - 52 (2.0%)
Patient reason	92 (3.5%)
Timing reason	287 (10.9%)

Table 12: Protocol deviations and other events associated with two-step algorithm use (n=2625). Step 1 = progesterone cut-off of ≤2nmol/l. Step 2 = M6 risk prediction model. hCG = human chorionic gonadotrophin

Analysis of adverse events (AE) & serious adverse events (SAE)

These results are summarised in table 13.

AE: Two-step protocol-related

7/320 women with a final pregnancy outcome of EP/ PPUL were misclassified as low-risk at step 1 and 8/320 at step 2. These were discussed in the section above (also see table 11). Sixteen women with an EP/ PPUL were not triaged by the M6 model because the second hCG level was not taken at 48 hours or there was a technical error with the excel data collection sheet in providing a risk prediction.

AE: clinician-related

Ten patients underwent a negative laparoscopy. Nine had a final outcome of FPUL/ IUP and one was ultimately a PPUL.

AE: patient-related

Fourteen patients had an unscheduled attendance/ admission with abdominal pain and 7 with vaginal bleeding. The majority underwent an earlier than scheduled repeat TVS. Thirteen patients underwent a laparoscopy which confirmed a diagnosis of ectopic pregnancy after their unscheduled attendance but none were a ruptured EP. A further three patients had a negative laparoscopy. Two patients underwent SMM but did not require a blood transfusion or have a prolonged hospital admission.

SAE

There were no SAE associated with implementing this two-step protocol into clinical practice.

Adverse event	N
Two-step protocol related	
EP classified as at low risk in step 1	7/320
EP classified as at low risk in step 2	8/320
EP not classified (no 2nd bloods, technical error)	16/320
Clinician related	
Negative laparoscopies*	10
Incorrectly given MTX#	0
Patient related	
Unscheduled attendance with pain	14
Unscheduled attendance with bleeding	7

* 9 were in patients who had a final outcome of FPUL/ IUP, 1 had a final outcome of PPUL
# MTX given to women with an intrauterine pregnancy

Table 13: Breakdown of adverse events (AE) in the study cohort. EP = ectopic pregnancy; MTX = methotrexate; PVB = per vaginal bleeding

Ruptured ectopic pregnancies

There were 7 ruptured EP in this study. Two cases re-presented before a 48-hour hCG could be taken so there was no M6 model risk prediction. In the other 5 cases, all were correctly classified as high-risk by the M6 model but had unsuccessful attempts at expectant/ medical management. All cases were managed promptly with surgery and none required a laparotomy, blood transfusion or prolonged hospital stay.

Discussion

Test performance and safety

We have shown that this two-step protocol using a serum progesterone cut-off level of ≤2nmol/l (step 1) and the M6 decision support model (step 2) demonstrated good test performance and was associated with few adverse clinical outcomes. Of 320 women with a final pregnancy outcome of EP/PPUL, only 15 were misclassified as low-risk and none came to any significant clinical harm.

Strengths of the study

This was a large (n=2625) multi-centre study, involving a number of district general and university hospitals. Furthermore, the study was carried out prospectively by a range of healthcare professionals including nurses, junior doctors and more senior clinicians. These elements are more likely to make our findings generalizable to other countries. We also had two versions of the M6 model so it could be used with or without serum progesterone results to accommodate centres where measuring progesterone levels is not routinely performed or for women who are on progesterone supplementation.

Limitations of the study

Human factor errors related to incorrect data entry are not uncommon in clinical studies and in 2% (n=63) of our patients there was a technical error with using the excel data collection sheet to provide a risk prediction (table 12). An excel sheet was used as we needed to collect detailed demographic and clinical data on all participants. We have, however, made the M6 model available for free use via a website (www.earlypregnancycare.co.uk), which may reduce operator-led errors. One could also argue that model use will deter clinicians from using their clinical acumen and become more protocol driven (especially junior staff), although the work in this study looking at protocol deviations showed this was not the case. There may be a lack of generalizability

to other non-EPU led patient populations. Further research to assess this is therefore warranted.

Protocol deviations

Most cases of misclassified EP were picked up due to a clinician or patient-related protocol deviation. Clinicians were instructed to not follow the management recommended if there was a clinical reason to do so and it was reassuring to see that there were sensible protocol deviations when necessary. Furthermore, they were instructed to counsel patients (with written information where possible) about the risks of a potential EP and to re-present if they had any worrying symptoms. The patient-related protocol deviations demonstrate that participating patients understood and followed this advice.

Summary

The management of women with a PUL is often inconsistent and lacking in a clear evidence-base. We have shown in this large prospective, multi-centre study that the two-step protocol incorporating the M6 model is able to assist clinicians in their decision making so they can provide consistent, evidence-based care. It reduces unnecessary follow-up for patients that do not need it and the allocation of finite healthcare resources to patients at greatest clinical risk of an EP. It has good test performance with few adverse events when used by clinicians in their daily clinical practice. It is accessible as a free online application at www.earlypregnancycare.co.uk and as an application for smartphones (search 'early pregnancy Leuven', see figure 37), allowing clinicians to incorporate its use into everyday clinical practice.

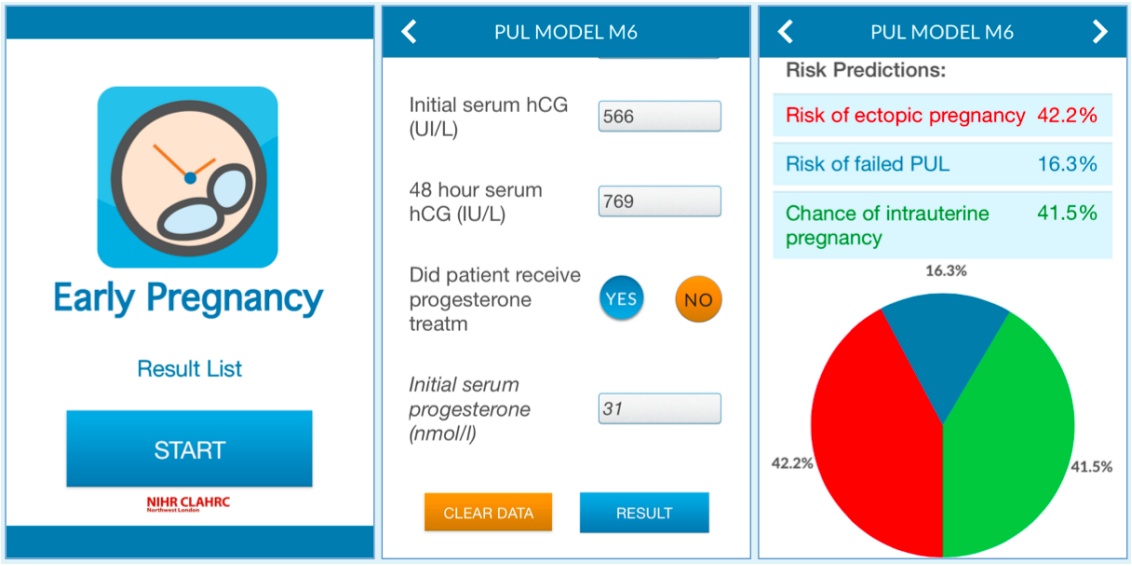


Figure 37: demonstration of the user-friendly app for smartphones developed to enable use of the M6 model by clinicians

It is also important to consider using the app to further validate M6 model use and that the ideal situation would be a dynamic and adaptive model that improves the more it is used via the use of AI (artificial intelligence) technology. There are however, logistical and financial limitations with regulating data input and model use via the app. This may cause unintended negative consequences such as less robust data collection, specifically the collection of safety data. There is also greater potential for misuse (i.e. M6 is not used as a risk prediction model but as a diagnostic model). One way around this may be the use of an online dataminer and future work could focus on this.

Final conclusions and future work

This study demonstrated that the two-step protocol for managing women with a PUL has better test performance at predicting a final outcome of an ectopic pregnancy than any other currently available management protocol for PUL. It is available freely for clinicians to use (www.earlypregnancycare.co.uk). The next steps are to widen its' use nationally throughout the UK via official approval by NHS organisations such as NICE (National Institute for Health and Clinical Excellence) and the RCOG (Royal College of Obstetricians & Gynaecologists). These organisations develop clinical guidelines that dictate clinical care and will be able to facilitate the uptake of this protocol so it becomes a routine, consistent way of caring for these women.

STUDY 5

THE VAGINAL MICROBIOME IN PREGNANCIES OF UNKNOWN LOCATION

Chapter Abstract

Introduction

We aimed to characterise vaginal microbiota composition in early pregnancy and investigate its' relationship with ectopic pregnancy within the context of PUL.

Methods

Miseq sequencing of 16S rRNA gene amplicons was used to prospectively profile and compare vaginal bacteria composition from as early as four weeks gestation in pregnancies that resulted in an ectopic pregnancy, miscarriage (NVIUP) or viable intrauterine pregnancy matched for age, gestation and body mass index. The relative abundance of vaginal bacteria was categorized by culture-independent next generation bacterial DNA sequencing techniques and the assessment of bacterial diversity.

Results

There were a total of 92 pregnancies, of which 22 patients had a final outcome of a VIUP, 15 a NVIUP, 27 a FPUL, 19 an EP and 9 a PPUL. Adverse final pregnancy outcomes were associated with reduced *Lactobacillus* species abundance (95.5% vs. 67.1%, P=0.033) and higher alpha diversity (Shannon index, P=0.0057) when compared to viable pregnancies. This association was independent of vaginal bleeding and observed prior to the diagnosis of the final pregnancy outcome (i.e. before the pregnancy was visible on ultrasound). Ectopic pregnancy had an even stronger association with *Lactobacillus* spp. depletion when compared to viable pregnancies (P=0.0007) with increased alpha diversity (P<0.0001).

Conclusions

Early pregnancy vaginal microbial stability was seen in women with a PUL who go onto have a viable intrauterine pregnancy. A vaginal microbiota deplete in *Lactobacillus* spp. precedes a diagnosis of ectopic pregnancy in a PUL population. These findings suggest that vaginal microbiota composition is a risk factor for ectopic pregnancy and that an unfavourable microbial composition is established very early in pregnancy. Further studies are required to clarify the role of infection in ectopic pregnancy and determine the mechanistic pathways by which vaginal microbial composition increases risk. It represents a potential therapeutic target for agents that could modify a womans' vaginal bacterial composition and improve early pregnancy outcomes.

Introduction

As outlined in previous chapters, the accurate and timely diagnosis of a potential EP in women with a PUL is clinically important as EP remains the leading early pregnancy-related cause of maternal mortality[18]. The mechanisms leading to the development of an EP are poorly understood, although pathogenic bacterial species such as Chlamydia *trachomatis*, a sexually transmitted infection, have been implicated in its' aetiology[222]. *Lactobacillus* species are commonly found in the healthy reproductive tract. They are dependent on high levels of oestrogen allowing high levels of glycogen production that *Lactobacillus* species can then use as a substrate. Lactobacillus have been shown to protect against pathogenic infection in the reproductive tract via several mechanisms: [i] producing lactic acid and lowering the vaginal pH[223], creating an acidic environment where pathogenic organisms cannot thrive; [ii] producing bacterocidins[224] (e.g. L. *gasseri* produces gasserin) which directly target pathogenic bacteria and by reducing levels of pro-inflammatory cytokines such as IL-6, IL-8 and TNFα and directly upregulating the production of anti-inflammatory cytokines such as IL-10[225]. An in vitro study has provided experimental evidence for the specific inhibition of C. *trachomatis* by *Lactobacillus* species via these protective mechanisms[226].

The vaginal microbiota is grouped into five main community state types (CST). CST I, II, III and V are dominated by *Lactobacillus* species and CST IV is characterised by high bacterial diversity and a dominance of anaerobic organisms[146]. Having a *Lactobacillus* deplete vaginal microbiota (CST IV) has been implicated in pathologies such as preterm labour[148, 227] and the progression of cervical intraepithelial neoplasia[228].

Bacterial diversity

Biodiversity refers to the variety of living organisms within a particular habitat. The assessment of bacterial diversity is categorised into three main domains:

[i] *Alpha diversity* accounts for diversity of bacteria within the same sample and is a measure of richness (i.e. the count or number of species present in a sample).

[ii] *Beta diversity* is a comparison of samples to each other and is a measure of evenness (i.e. the distribution or relative abundance of different species between samples).

[iii] *Gamma diversity* is a measure of overall diversity within the entire sample set.

I used the Shannon Diversity Index as measures of richness and evenness, which is a commonly used weighted measure of the different types of bacteria in an ecological niche, i.e. when there is only one type of bacterium present, the SDI is zero and there is no diversity.

Study aims

I aimed to compare the vaginal bacterial community composition of PUL that went onto have different final pregnancy outcomes. I hypothesized that PUL with a final outcome of VIUP would have a vaginal microbiota dominated by *Lactobacillus* species and all other outcomes (which are essentially adverse) would be *Lactobacillus* deplete. I tested this by comparing early vaginal microbial composition in 92 cases of PUL.

Methods

Study population

All women presenting to the EPU at Queen Charlottes' and Chelsea Hospital between 25-03-15 and 27-05-16 and classified as a PUL at their initial TVS were eligible for participation in the study. Ethical approval was obtained from the North of Scotland Research Ethics Service (reference no. 14/NS/1078) and all participants provided informed written consent. A detailed medical and gynaecological history was collected (see appendix 1). Ethnicity was self-reported as White, Asian or Black and other ethnicities were grouped as 'mixed/ other'. The final outcome of the pregnancy was determined when the pregnancy would have been 11-14 weeks gestation.

Inclusion and exclusion criteria

As outlined in the methods chapter, see page 60.

Materials

- Elution-swab system (Copan)
- Lysozyme (Sigma L6876 -chicken egg white)
- Mutanolysin (Sigma M9901/10KU)
- Lysostaphin (Sigma L9043)
- QIAMP DNA mini kit (Qiagen Cat No 51304)
- Glass beads (Sartorius Stedim Cat No BBI-8541400)
- Micro Dismembrator (Sartorius)
- Amplitaq Gold DNA polymerase (Applied Biosystems) OR High fidelity DNA polymerase. NEB One Taq Polymerase M0480
- 10mM dNTP (Sigma cat no D7295)
- DNase/RNase free water (VWR Cat no 436912C)
- Agarose (electrophoresis grade from Invitrogen cat no 15510-019)

- SYBR Safe DNA gel Stain (Invitrogen Cat No S33102) or from Appleton Woods
- Molecular weight marker (Bioline Hyperladder 100bp Cat no BIO-33056)

Sample collection and processing

A swab was taken at a single time point for each patient from the mid-vagina using a BBLTM CultureSwab™ containing liquid Amies (Becton Dickinson, Oxford, UK) and a Transwab MW170 with rayon bud type (Medical Wire & Equipment). These were immediately placed on ice and stored at −80°C. This was usually at the second hospital visit (as the first visit involved a TVS that requires lubricant gel which would contaminate swab samples). If this was not possible (e.g. the patient attended for their second visit over a weekend), a swab was taken at a subsequent hospital visit.

DNA extraction

The full SOP (Standard Operating Procedure) I wrote is attached as appendix 7. To summarise, whole-Genomic bacterial DNA was extracted from the BBL CultureSwab swabs after defrosting to isolate bacterial DNA from the collected cervicovaginal fluid. In addition to the vaginal swabs sampling study participants (n=92), 4 negative control swabs were analysed. Swabs were then vortexed after the addition of transport buffer (Amies Liquid Medium) to re-suspend cells.

Initial cell lysis step: this was carried out using an enzymatic method. 50 µL of lyzosyme (10 mg/mL), 6 µL of mutanolysin (25,000 U/mL; Sigma- Aldrich), 3 µL of lysostaphin (4,000 U/mL in sodium acetate; Sigma- Aldrich), 41 µL of TE50 buffer (10 mM Tris.HCL and 50 mM EDTA, pH 8.0) were added to the cells and incubated for 1 hour at 37°C[146]. Acid washed glass beads were then added to each sample, to induce mechanical cell lysis using a Mikro-Dismembrator (Sartorius UK Ltd, Surrey, United Kingdom) for 1 min at 1000 x g. The resulting lysate was then processed and purified using the QIAamp

DNA Mini kit (Qiagen, Venlo, Netherlands). The resulting DNA within each sample was eluted in 100 µl AE buffer (10 mM Tris-Cl, 0.5 mM EDTA; pH 9.0). Confirmation of the extracted bacterial DNA was carried out by PCR amplification using universal primers[145].

Illumina MiSeq sequencing of 16S rRNA gene amplicons

The V1-V2 hypervariable regions of 16S rRNA genes were amplified by PCR using a forward and reverse fusion primer as previously described[145]. The forward primer was an Illumina i5 adapter (5′-AATGATACGGCGACCACCGAGATCTACAC-3′), an 8–base pair (bp) bar code, a primer pad (forward, 5′-TATGGTAATT-3′), and the 28F primer (5′-GAGTTTGATCNTGGCTCAG-3′)[229]. The reverse fusion primer was created with an Illumina i7 adapter (5′-CAAGCAGAAGACGGCATACGAGAT-3′), an 8-bp bar code, a primer pad (reverse, 5′-AGTCAGTCAG-3′), and the 388R primer (5′-TGCTGCCTCCCGTAGGAGT-3′). Sequencing was conducted at the Research and Testing Laboratory (Lubbock, TX, USA) using an Illumina MiSeq platform (Illumina Inc.).

16S rRNA gene sequence analysis

Sequence data were analysed in Mothur using the MiSeq SOP Pipeline[230]. Sequence reads were quality checked and normalised to the lowest number of reads (n=3945), with a cut-off value of 97% used to avoid sequencing bias. Sequence alignment was performed using the Silva bacterial database (www.arb-silva.de/) and undertaken using the RDP (Ribosomal Database Project) database reference sequence files and the Wang method[231]. OTUs (Operational Taxonomic Units) were defined (from phylum to genus) using the RDP MultiClassifier script, with USEARCH used for species-level taxonomies[232]. Result data was analysed using the Vegan package within the R statistical package for assessment of microbial composition and diversity (R Development Core Team 2008).

Statistical analysis

Subjects were analysed in five different groups according to their final clinical outcome: VIUP, NVIUP, FPUL, EP and PPUL. The definitions used for final outcomes are as described in the methods chapter (pages 65-8).

Analysis of statistical differences between the vaginal microbiome of patient groups was performed using the Statistical Analysis of Metagenomic Profiles (STAMP)[233] and ClustVis[234] packages. Data were subjected to multivariate analysis using hierarchical clustering analysis (HCA) by nearest neighbour linkage with a clustering density threshold of 0.75. To assess for bias in the selected clinical characteristics with respect to the five outcome categories, we performed Fisher's exact test for age and ethnicity.

Analyses were performed using STAMP and false discovery rate adjustment (Benjamin & Hochberg) was applied to correct p-values. P-values <0.05 were considered significant. Samples were a proportion of the total number of reads per sample into the following groups: normal (>93% *Lactobacillus* spp.), intermediate (67-81% *Lactobacillus* spp.), or high diversity (<17% *Lactobacillus* spp.). Bacterial species data were classified into CSTs as described by Ravel et al[146]: CST I (Lactobacillus *crispatus*), CST II (Lactobacillus *gasseri*), CST III (Lactobacillus *iners*), CST IV (mixed bacterial species), and CST V (Lactobacillus *jensenii*).

In the negative controls, contaminants were considered to be those OTUs that were detected in at least half of the 4 negative control swabs and at a proportional ratio of greater than 1 to 1. This assumes that all sequences found in negative controls arise from contamination in the kit itself.

The linear discriminative analysis (LDA) Effect Size (LEfSe) method[235] allows the discovery of microbial biomarkers that are different between two outcome groups. I therefore used it to identify discriminatory features at different taxonomic levels between different PUL outcome groups (VIUP vs all other (adverse) outcomes and VIUP vs EP). This identifies differentially abundant features that highlight both statistical significance and biological relevance. A value of 0.01 was used for factorial Kruskal-Wallis test between classes and a threshold of 3.0 was used for logarithmic LDA score for discriminative features. The Mann-Whitney test was used to analyse the diversity of bacterial species between participants using the Shannon index.

Results

Baseline characteristics of study subjects

Women who attended the EPU at Queen Charlottes' and Chelsea Hospital and were classified as a PUL at their initial TVS were prospectively recruited. Of 92 PUL, 22 patients had a final outcome of a viable intrauterine pregnancy (VIUP), 15 a non-viable intrauterine pregnancy (NVIUP), 27 a failed PUL (FPUL), 19 an ectopic pregnancy (EP) and 9 a persistent PUL (PPUL). PUL with a final outcome of VIUP were used as a control group.

Two sub-analyses were performed: (i) VIUP versus all other (adverse) PUL outcomes and (ii) VIUP versus EP. Demographic data, including age, gestation and ethnicity were comparable (table 14).

	Viable	Adverse	p-value	Ectopic	p-value[+]
n (%)	22	70		19	NA
Age (median ± SD)	31.0 ±5.7	35.0±5.8	0.093	35.0±5.3	0.29
Ethnicity, n (%)			0.557		0.131
Arab	2 (9.1)	6 (8.6)	ns	0 (0.0)	ns
Asian	1 (4.5)	4 (5.7)	ns	3 (15.8)	ns
Black	2 (9.1)	7 (10.0)	ns	2 (10.5)	ns
White	17 (77.3)	45 (64.3)	ns	11 (57.9)	ns
Mixed/Other	0 (0.0)	8 (11.4)	ns	3 (15.8)	ns
Gestation at presentation (days), (median ± SD)	37	45.6		43.3	
Bleeding score					
0	17	16		4	
1	4	24		9	
2	0	13		3	
3	1	11		3	
4	0	6		0	

Table 14: metadata for PUL final pregnancy outcome groups. Note the adverse group contains the ectopic group. SD = standard deviation

The impact of final pregnancy outcome on the vaginal microbiome

A total of 92 swabs samples underwent bacterial DNA sequencing. The data set consisted of 354,902 high-quality gene sequences, with a mean sequence read count of 2063 per sample (range, 20 to 206,338). A total of 172 bacterial species were identified. Samples were classified into two groups using hierarchical clustering of relative abundance data from the top 100 bacterial species (accounting for >95% of all sequence reads). These clusters were characterised by the relative abundance of *Lactobacillus* spp. and classified according to their vaginal bacterial communities as normal (>67% *Lactobacillus* spp.) or deplete (<17% *Lactobacillus* spp.), see figure 38A.

Reduced Lactobacillus spp. abundance was associated with adverse final pregnancy outcomes after adjustment for vaginal bleeding

For initial analyses, the bacterial composition of all adverse final pregnancy outcome samples (n=70) was compared to the control (VIUP) group (n=22). The amount of vaginal bleeding experienced can significantly alter the microbiota. All participants completed a questionnaire (appendix 1) that included quantifying their vaginal bleeding via a validated PBAC (Pictoral Blood Loss Assessment Chart). Patients with a bleeding score of 4 (n=6) had microbiota disproportionately deplete in Lactobacillus (figure 38B, *P*=0.035). These 6 patients were therefore excluded from further analysis. Viable pregnancies demonstrated a greater likelihood of having a vaginal microbiota composition dominated by *Lactobacillus* spp. (95.5% vs. 67.1%, *P*=0.033, figure 38C). Consistent with this, adverse outcome samples demonstrated significantly higher alpha diversity (Shannon index *P*=0.0057, figure 38D).

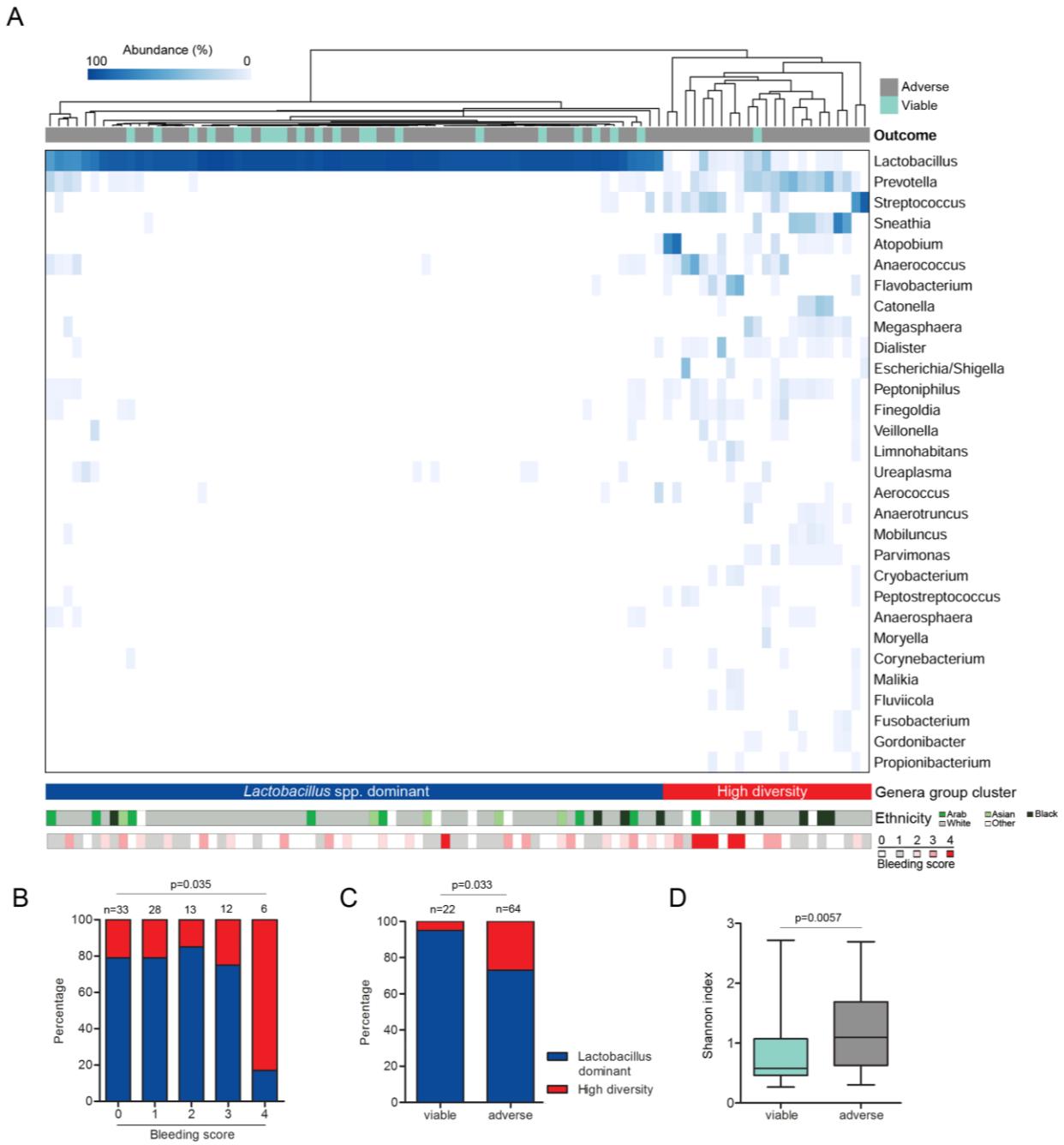


Figure 38: analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) outcomes (n=70)

(A) heatmap of genera data; (B) association between the vaginal bleeding score and percentage Lactobacillus dominance; (C) percentage of Lactobacillus dominance in viable and adverse pregnancy outcome groups; (D) Shannon index analysis of vaginal bacterial diversity between the two outcome groups

CST II is associated with a viable outcome and CST III and IV with adverse final pregnancy outcomes

CST II (Lactobacillus *gasseri*) was dominant in the viable group, whereas the adverse group had a microbiota dominant in CST III (Lactobacillus *iners*) and CST IV (figure 39A and 39B, *P*=0.0003). There was a trend in the adverse group towards dominance of anaerobic organisms such as Prevotella *timonensis*, Prevotella *bivia*, *Peptoniphilus*, Peptostreptococcus *anaerobius*, Mobiluncus *mulieris* and *Dialister* (figure 39C), but none of these withstood correction. LefSe[235] analysis confirmed the dominance of CST II in the viable group and CST III in the adverse group. It also revealed further microbial biomarkers in the adverse group, namely *Anaerococcus*, *Lachnospiraceae* and *Clostridia* (figure 40).

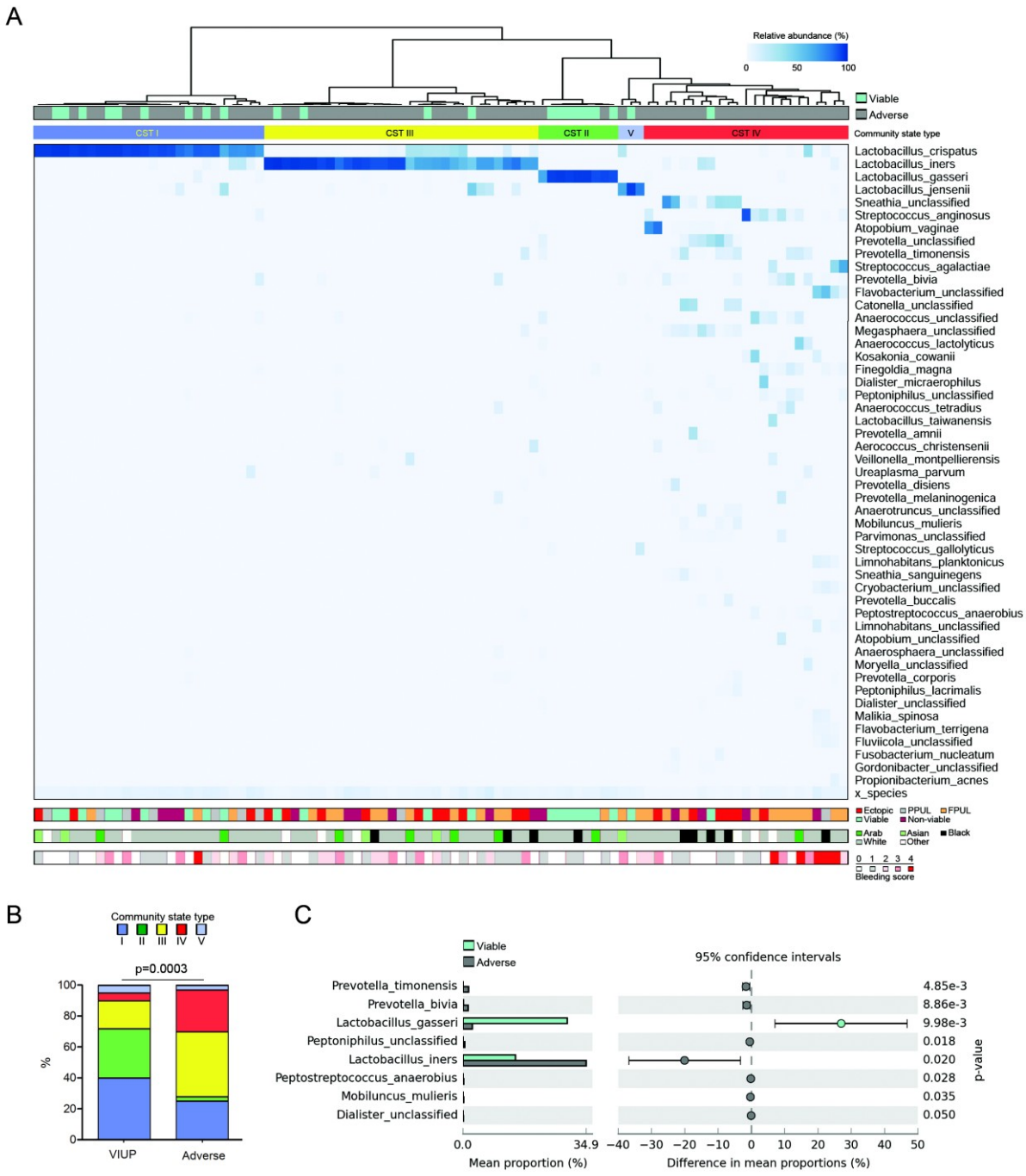


Figure 39: analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) outcomes (n=70)

(A) heatmap of species data; (B) proportion of community state types in viable and adverse pregnancy outcome groups; (C) extended error bar for microbial biomarkers identified to be significantly different between the two outcome groups. Note, none of these withstood correction

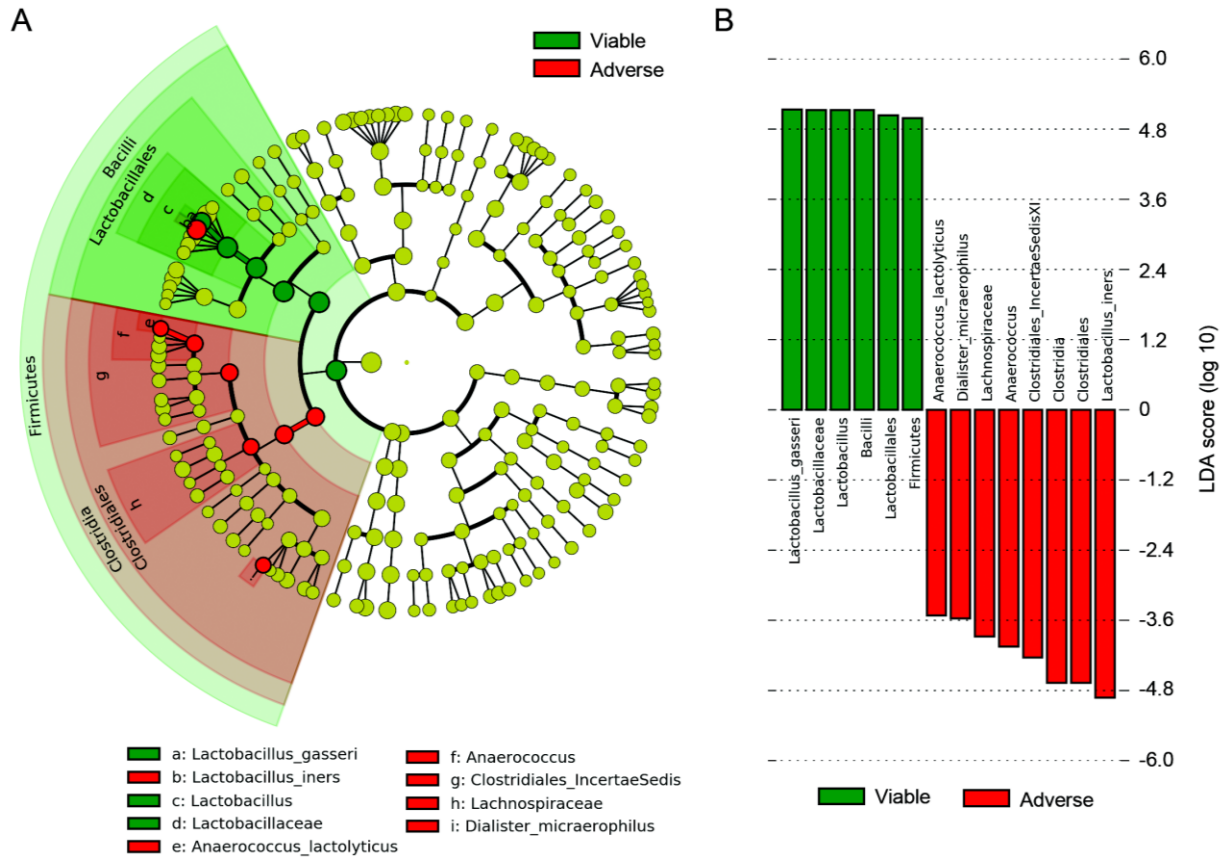


Figure 40: LefSe (latent discriminatory analysis effect size) analysis identifying microbial biomarkers significantly different between PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus all other (adverse) early pregnancy outcomes (n=70)

Reduced *Lactobacillus* spp. abundance was associated with ectopic pregnancy and precedes its' diagnosis

For secondary analyses, bacterial composition of all ectopic pregnancy samples (n=19) was compared to the control (VIUP) group (n=22). There was a dominance of CST III and IV in the ectopic group (figure 41A, *P*=0.0007). Shannon index analysis revealed a significant difference in bacterial diversity, with much higher diversity seen in the ectopic group (figure 41B, *P*<0.0001). LefSe analysis confirmed the dominance of CST II in the viable group. Microbial biomarkers dominant in the adverse group included L *iners*, *Sneathia*, *Fusobacteria*, *Catonella*, *Dialister*, *Prevotella* and *Lachnospiraceae* (figure 42).

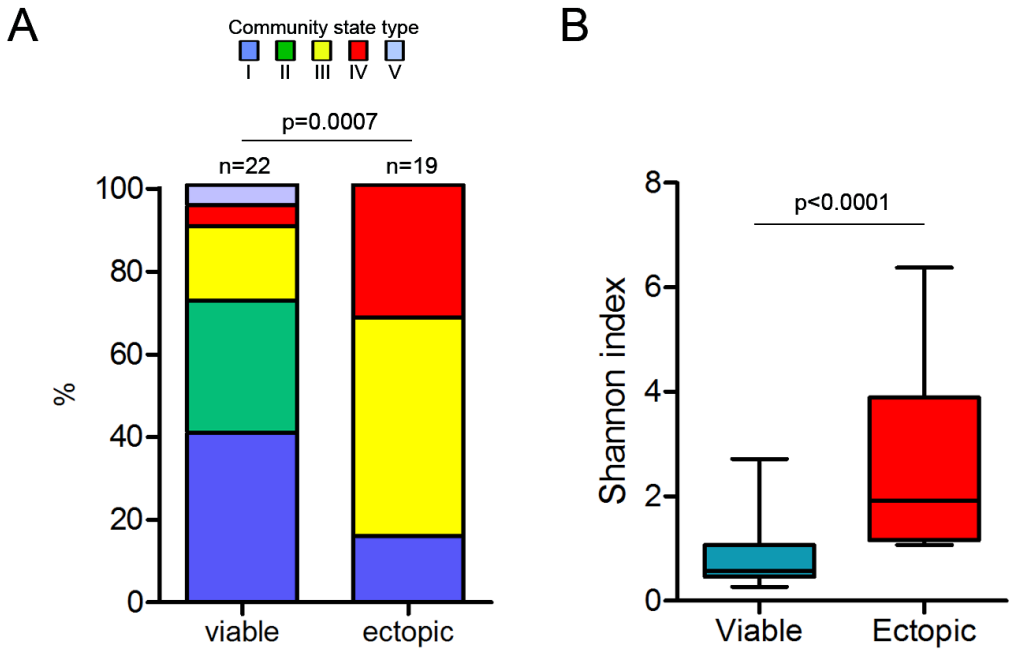


Figure 41: analysis of PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus an ectopic pregnancy (n=19). (A) proportion of community state types in viable and ectopic pregnancy outcome groups. (B) Shannon index analysis of vaginal bacterial diversity between the two outcome groups

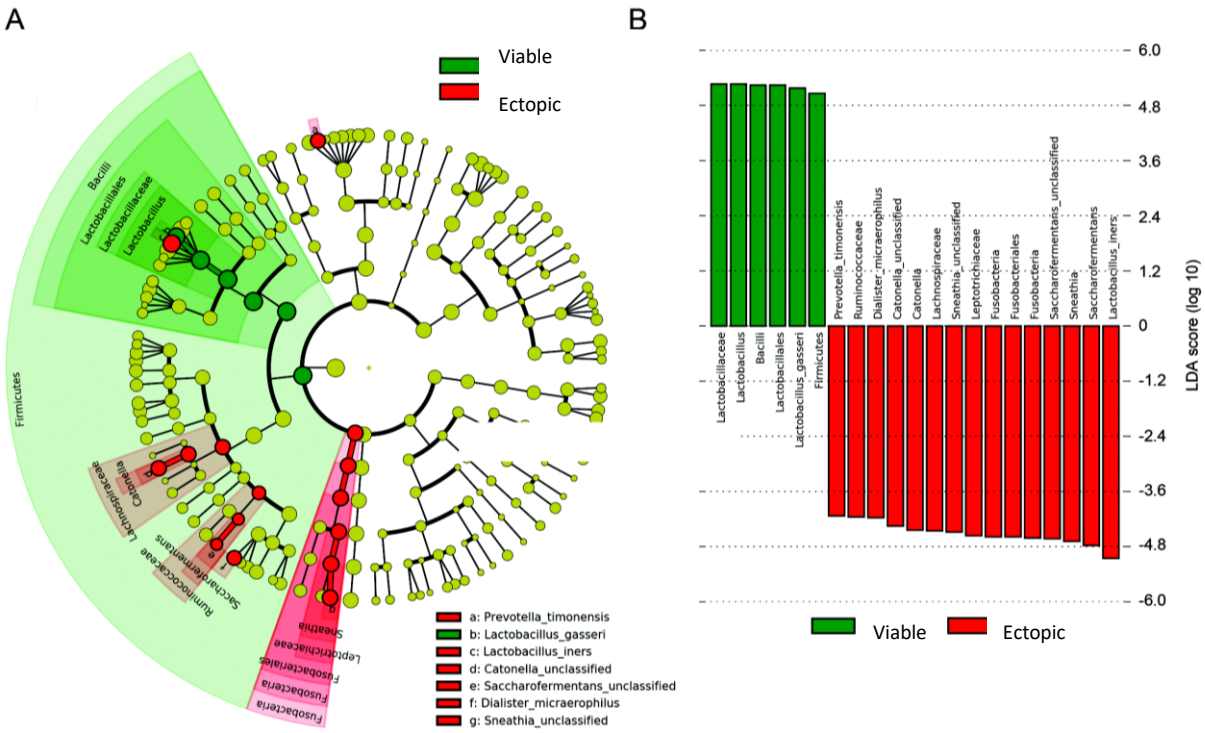


Figure 42: LefSe (latent discriminatory analysis effect size) analysis identifying microbial biomarkers significantly different between PUL with a final outcome of a viable intrauterine pregnancy (n=22) versus an ectopic pregnancy (n=19)

Discussion

Main findings

Lactobacillus spp. deplete, high diversity vaginal microbial community composition precedes a diagnosis of ectopic pregnancy in a PUL population.

Early pregnancy vaginal microbial stability was present in women with a PUL who go onto have a viable intrauterine pregnancy.

Strengths and limitations

Strengths of this study include the analysis of a well-characterised cohort with information on vaginal bleeding. We also used culture-independent bacterial DNA sequencing approaches which permitted the identification and evaluation of the relative abundance of a broad range of commensal and pathogenic vaginal bacteria, including relevant *Lactobacillus* spp.

Limitations include the relatively low number of samples and the absence of pre-conception vaginal samples, which would have provided more information on whether Lactobacillus depletion precedes or is caused by an adverse early pregnancy outcome. There may also be issues with platform capability and if the MiSeq platform cannot be used in the future, it would cause concerns about data reproducibility. Furthermore, these techniques are currently at a high cost with low throughput, requiring specialist training and equipment. This is understandable considering how new these techniques are, but any advancement towards use in everyday clinical practice would require a streamlining of cost and intensity of work.

Chlamydia trachomatis detection

All sequencing techniques have limitations, irrespective of the hypervariable region studied. 16S rRNA sequencing is currently unable to reliably identify Chlamydia *trachomatis*. We intend to, however, perform targeted PCR on these samples to specifically look for this bacterium. The reality is that we are unlikely to detect a significant amount of Chlamydia *trachomatis* in this cohort as there is good epidemiological evidence to suggest that only 2 in every 1000 women between the ages of 16 and 44 confirmed to have had a genital tract infection positive for Chlamydia *trachomatis* go onto have an EP during their reproductive lifetime[236]. The findings of dysbiosis in this study are far more stark compared to that figure.

Interpretation

Predicting which women with a PUL are likely to have a final outcome of an ectopic pregnancy is clinically important. A host of novel serum and plasma-based biomarkers have been investigated[137] and none, outside of hCG and progesterone, have reliable test performance enabling their use in clinical practice. In this study, we investigated the role of vaginal microbial biomarkers, comparing PUL with a final outcome of VIUP against all other (adverse) clinical outcomes and then EP specifically. These results are the first of their kind to demonstrate that women with a PUL that go onto have a positive pregnancy outcome (i.e. VIUP) have a vaginal microbiota that is largely Lactobacillus dominant, particularly L *gasseri* (CST II). In contrast, PUL with an adverse final outcome have a microbiota characterised by Lactobacillus depletion.

It may be that a vaginal microbiota deplete in Lactobacillus disturbs local inflammatory processes required for adequate implantation, predisposing a woman to developing an ectopic pregnancy. Bacteria such as Sneathia, which was dominant in the EP group, are known to cause inflammation within the female genital tract[237, 238]. Alternatively, it may

be that the aberrant implantation related to having an ectopic pregnancy versus a healthy viable intrauterine pregnancy is reflected in the composition of the vaginal microbiota.

Whatever the mechanism, microbiome structure appears to be linked to the final pregnancy outcome in PUL. The findings of this study are associative and further studies to investigate the mechanisms behind these microbial differences are warranted. Vaginal bleeding is a common symptom in early pregnancy[239] but can also cause vaginal microbial instability and *Lactobacillus* spp. depletion resulting in colonisation with potentially pathogenic anaerobic organisms[145]. The relationship between *Lactobacillus* spp. depletion and ectopic pregnancy identified in this study, however, appears to be independent of bleeding as it persisted following the exclusion of samples with reported heavy vaginal bleeding.

Performing this study within the context of PUL and obtaining a vaginal swab sample prior to determining the final clinical outcome gives rise to the potential for developing a new biomarker that can predict pregnancy outcome before any other test can do so. This would have real clinical benefit in determining which PUL patients require intensive follow-up versus those that can be safely discharged. Further work may also lead to the development of therapeutic agents that can modify a womans' vaginal bacterial composition through the use of targeted antibiotic, prebiotic or probiotic treatments and improve early pregnancy outcomes. Not all women with a final outcome of an EP had a microbiota that was Lactobacillus deplete. The role of host adaptive immune responses, therefore, is likely to play a key role and warrants further study.

Summary

A vaginal microbiota deplete in *Lactobacillus* spp. precedes a diagnosis of ectopic pregnancy in a PUL population. Early pregnancy vaginal microbial stability was seen in women with a PUL who go onto have a viable intrauterine pregnancy.

Final conclusions and future work

This study provides an insight into the first trimester vaginal microbiota of women with a pregnancy of unknown location and its potential influence on the development of an ectopic pregnancy. Lactobacillus depletion is known to augment local inflammatory responses and interference with key inflammatory pathways that allow adequate implantation. Further work should therefore focus on analysing pro- and anti-inflammatory cytokines that have been linked to microbiome structure to try and elucidate the potential mechanistic pathways responsible for altered vaginal microbial composition in women with a PUL who experience an adverse final pregnancy outcome. This may shed light on the pathophysiology behind ectopic pregnancies and how high-risk vaginal microbiota communities are established. Further studies are needed to better understand this relationship in order to clarify the role of infection in ectopic pregnancy and how treatment strategies may help us prevent them.

DISCUSSION AND FUTURE PROSPECTS

Summary of the clinical problem

The clinical care provided to women with a pregnancy of unknown location (PUL) is often inconsistent and lacking in a clear evidence base. Somewhere between 10-40% women initially classified as a PUL will have a final pregnancy outcome of an ectopic pregnancy (EP). Despite research into this area and the advent of dedicated Early Pregnancy Units, this remains the leading early pregnancy-related cause of maternal death in the UK. It is therefore clinically important to determine the best management protocol for these women so those at highest risk of harbouring an EP can be monitored more intensively and those at lowest risk can safely reduce the amount of any unnecessary follow-up.

The ABPEP (Assessment of Biomarkers in PUL and Ectopic Pregnancy) study aimed to assess whether existing serum biomarkers (hCG and progesterone) used in the management of PUL can be used to develop management protocols that have sufficient diagnostic accuracy to safely rationalize management in clinical practice. In addition, it also aimed to assess whether novel biomarkers, such as those seen in the vaginal microbiome, might identify those at risk of an ectopic pregnancy or provide new information regarding the aetiology of EP.

To achieve these aims, I carried out five separate studies:

Study 1: Management protocols used in pregnancy of unknown location: a systematic review and meta-analysis

Study 2: Using the M4 logistic regression model to manage women with a pregnancy of unknown location: a clinical implementation study

Study 3: Development and validation of the M6 model – a logistic regression model utilizing serum hCG and progesterone as part of a two-step protocol

Study 4: Using the M6 model as part of a two-step protocol to manage women with a pregnancy of unknown location: a clinical implementation study

Study 5: The vaginal microbiome in pregnancies of unknown location and its' role in ectopic pregnancy

I initially performed a systematic review and meta-analysis to review the current evidence-base (study 1). Based on those findings, I coordinated three clinical studies (studies 2-4) recruiting 7,047 patients in total. I also obtained ethical approval to collect additional samples from recruited women to perform novel biomarker studies. This arm of the thesis recruited 794 participants in total. From these, I was able to carry out a pilot study assessing the vaginal microbiota of 92 women with a PUL (study 5). For the duration of the ABPEP study, I performed 1,797 ultrasound scans and collected over 15,000 bio-fluid samples.

This thesis has successfully met the aims above and reported novel findings within each study.

Clinical relevance and future work

Strengths

This thesis is based on data from a prospectively recruited and well-characterised patient cohort. Demographic and clinical data was collected prospectively and in detail, using validated symptom scores.

Each participating centre in the clinical studies followed a pre-determined study protocol with careful attention paid to the early detection of any issues relating to clinical risk. I coordinated regular update meetings with the clinical leads at each participating centre to address queries, discuss cases and provide interim updates on the data collected. Each clinical study is based on a large number of patients and all studies are multi-centre, increasing the potential generalizability of our findings to other patient populations.

For the novel biomarker studies, detailed SOPs (standard operating procedures) were followed for the collection, processing and storage of bio-fluid samples.

Finally, the work presented in this thesis confirms that the initial aims of my PhD have been met.

Limitations

All clinical studies were undertaken in hospitals with dedicated EPUs and senior clinicians with an expertise and interest in the field of early pregnancy. This is standard practice within the UK and some countries in Europe but not in others such as the USA where women are often scanned by professionals with less specialised training within Accident & Emergency departments. This potentially results in different patient

populations being seen and therefore our findings of excellent test performance with the M6 model as a diagnostic tool for predicting ectopic pregnancy may not be reproduced in these settings.

The study on the vaginal microbiota in PUL was a small pilot study and further validation studies are required to determine whether the findings are reproducible. Furthermore, our findings were associative and further work is required to elucidate whether changes in the vaginal microbiome have a causative role in the pathophysiology of EP. If there is an infectious aetiology to the development of EP, studies looking into the host immune response and specific inflammatory pathways are required.

Recommendations for further research from this thesis

Given the findings described in this thesis, the following suggestions for further work within this field are recommended:

- The M4 and M6 models are purely based on the interpretation of serum biomarker (i.e. hCG and progesterone) results. The development of newer models incorporating clinical variables such as the extent of vaginal bleeding or ultrasound-based variables such as endometrial thickness may provide additional diagnostic benefit.

- Only one novel biomarker study was carried out as part of this thesis. As the ABPEP study has collected a large bio-bank of urine, serum, plasma and vaginal swab samples from recruited participants, there is huge scope for further biomarker discovery and validation work.

- The pathophysiology of ectopic pregnancy remains poorly understood and therefore its' incidence remains unchanged. There has always been a hypothesis

that it may have an infectious aetiology (particularly Chlamydia *trachomatis*). Further work looking into the vaginal microbiome, the bacteria detected by novel culture-independent techniques and associated inflammatory pathways may provide useful information into the mechanisms causing ectopic pregnancy and therefore potential therapeutic targets. Furthermore, the pilot study on the vaginal microbiome could be used as the basis for applying for larger research grants for work in this field.

- To further examine the aetiological role of the vaginal microbiome in ectopic pregnancy, microbial sampling in other parts of the female reproductive tract (e.g. endometrium or fallopian tube tissue in women undergoing a salpingectomy as treatment for their EP) may provide additional information. Pre-conception sampling would provide useful information on the baseline vaginal microbial composition of women and how this may change dependent on their final pregnancy outcome.

- Further studies combining the information on existing biomarkers (i.e. hCG and progesterone) and more novel biomarker data may provide additional diagnostic benefit.

- MicroRNAs (miRNAs) are small non-coding RNA molecules that have been found to be associated with numerous malignancies. They also appear to to play a modulatory role in pregnancy, particularly in the onset of labour. Novel work looking at the utility of microRNAs in early pregnancy may yield new diagnostic biomarkers.

Final Conclusions

Research findings we can implement to change our clinical practice now

The two-step protocol for managing PUL and interpreting serum hCG and/ or progesterone levels has better test performance in correctly predicting a final outcome of EP when compared to all other existing management protocols for PUL. The work in this thesis has externally validated its' use within a large multi-centre clinical implementation study. It demonstrated an excellent clinical safety profile with no reported serious clinical adverse events. It is freely available for clinicians to use in their routine clinical practice (www.earlypregnancycare.co.uk). The next steps are to now widen its' use and carry out further validation studies confirming the findings of this thesis so the two-step protocol can be incorporated into national guidance for the management of these women. This would be an affirmative way to ensure the care women with a PUL receive is consistent and evidence-based, irrespective of which hospital they are seen at.

Research findings revealing avenues for future research

Managing women with a PUL centres around accurate risk prediction. Clinically this is important as it enables the appropriate allocation of finite healthcare resources, reducing the risk of clinical harm and the honest counselling of women about the likely outcome of their pregnancy. Events in early pregnancy are likely to affect or be affected by clinical, biochemical and microbial factors. The study assessing the vaginal microbiota in women with a PUL yielded new information on how novel biomarkers may be used to predict the risk of an ectopic pregnancy and provide insight into the underlying mechanisms of the disease process. Predicting adverse outcomes at an early stage is not only essential for risk prediction but may elucidate modifiable risk factors amenable to treatment or preventative strategies best employed pre-conception or in these early stages to positively alter the outcome of a pregnancy. Future work within this field may provide more answers to these questions.

REFERENCES

1. Antenatal Care: Routine Care for the Healthy Pregnant Woman. National Institute for Health and Clinical Excellence: Guidance. London2008.
2. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ (Clinical research ed). 2009;338:b606.
3. Cha J, Sun X, Dey SK. Mechanisms of implantation: strategies for successful pregnancy. Nat Med. 2012;18(12):1754-67.
4. Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. BMC Med. 2013;11:154.
5. Wilcox AJ, Baird DD, Weinberg CR. Time of implantation of the conceptus and loss of pregnancy. The New England journal of medicine. 1999;340(23):1796-9.
6. Knight M BK, Tuffnell D, Shakespeare J, Kotnis R, Kenyon S, Kurinczuk JJ (Eds.). on behalf of MBRRACE-UK. Saving Lives, Improving Mothers' Care - Lessons learned to inform maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2015-17. Oxford: National Perinatal Epidemiology Unit, University of Oxford 2019. 2019.
7. Ectopic Pregnancy and Miscarriage: Diagnosis and Initial Management in Early Pregnancy of Ectopic Pregnancy and Miscarriage. National Institute for Health and Clinical Excellence: Guidance. London2012.
8. Bigrigg MA, Read MD. Management of women referred to early pregnancy assessment unit: care and cost effectiveness. BMJ (Clinical research ed). 1991;302(6776):577-9.
9. van den Berg MM, Goddijn M, Ankum WM, van Woerden EE, van der Veen F, van Wely M, Hajenius PJ. Early pregnancy care over time: should we promote an early pregnancy assessment unit? Reproductive biomedicine online. 2015;31(2):192-8.
10. van den Berg MMJ, Dancet EAF, Erlikh T, van der Veen F, Goddijn M, Hajenius PJ. Patient-centered early pregnancy care: a systematic review of quantitative and qualitative studies on the perspectives of women and their partners. Hum Reprod Update. 2018;24(1):106-18.
11. Kirk E, Bottomley C, Bourne T. Diagnosing ectopic pregnancy and current concepts in the management of pregnancy of unknown location. Hum Reprod Update. 2014;20(2):250-61.
12. Condous G, Timmerman D, Goldstein S, Valentin L, Jurkovic D, Bourne T. Pregnancies of unknown location: consensus statement. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2006;28(2):121-2.
13. Barnhart K, van Mello NM, Bourne T, Kirk E, Van Calster B, Bottomley C, Chung K, Condous G, Goldstein S, Hajenius PJ, Mol BW, Molinaro T, O'Flynn O'Brien KL, Husicka R, Sammel M, Timmerman D. Pregnancy of unknown location: a consensus statement of nomenclature, definitions, and outcome. Fertil Steril. 2011;95(3):857-66.
14. Kolte AM, Bernardi LA, Christiansen OB, Quenby S, Farquharson RG, Goddijn M, Stephenson MD, Eshre Special Interest Group EP. Terminology for pregnancy loss prior to viability: a consensus statement from the ESHRE early pregnancy special interest group. Human reproduction (Oxford, England). 2015;30(3):495-8.

15. National Collaborating Centre for Ws, Children's H. National Institute for Health and Clinical Excellence: Guidance. Ectopic Pregnancy and Miscarriage: Diagnosis and Initial Management in Early Pregnancy of Ectopic Pregnancy and Miscarriage. London: Rcog
National Collaborating Centre for Women's and Children's Health.; 2012.
16. Abdallah Y, Daemen A, Kirk E, Pexsters A, Naji O, Stalder C, Gould D, Ahmed S, Guha S, Syed S, Bottomley C, Timmerman D, Bourne T. Limitations of current definitions of miscarriage using mean gestational sac diameter and crown-rump length measurements: a multicenter observational study. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2011;38(5):497-502.
17. Diagnosis and Management of Ectopic Pregnancy: Green-top Guideline No. 21. Bjog. 2016;123(13):e15-e55.
18. Cantwell R, Clutton-Brock T, Cooper G, Dawson A, Drife J, Garrod D, Harper A, Hulbert D, Lucas S, McClure J, Millward-Sadler H, Neilson J, Nelson-Piercy C, Norman J, O'Herlihy C, Oates M, Shakespeare J, de Swiet M, Williamson C, Beale V, Knight M, Lennox C, Miller A, Parmar D, Rogers J, Springett A. Saving Mothers' Lives: Reviewing maternal deaths to make motherhood safer: 2006-2008. The Eighth Report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. BJOG. 2011;118 Suppl 1:1-203.
19. Sivalingam VN, Duncan WC, Kirk E, Shephard LA, Horne AW. Diagnosis and management of ectopic pregnancy. The journal of family planning and reproductive health care. 2011;37(4):231-40.
20. Kirk E, Papageorghiou AT, Condous G, Tan L, Bora S, Bourne T. The diagnostic effectiveness of an initial transvaginal scan in detecting ectopic pregnancy. Human reproduction (Oxford, England). 2007;22(11):2824-8.
21. Condous G, Van Calster B, Kirk E, Haider Z, Timmerman D, Van Huffel S, Bourne T. Prediction of ectopic pregnancy in women with a pregnancy of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2007;29(6):680-7.
22. Van Calster B, Abdallah Y, Guha S, Kirk E, Van Hoorde K, Condous G, Preisler J, Hoo W, Stalder C, Bottomley C, Timmerman D, Bourne T. Rationalizing the management of pregnancies of unknown location: temporal and external validation of a risk prediction model on 1962 pregnancies. Human reproduction (Oxford, England). 2013;28(3):609-16.
23. Condous G, Van Calster B, Kirk E, Haider Z, Timmerman D, Van Huffel S, Bourne T. Clinical information does not improve the performance of mathematical models in predicting the outcome of pregnancies of unknown location. Fertil Steril. 2007;88(3):572-80.
24. Ayim F, Tapp S, Guha S, Ameye L, Al-Memar M, Sayasneh A, Bottomley C, Gould D, Stalder C, Timmerman D, Bourne T. Can risk factors, clinical history and symptoms be used to predict risk of ectopic pregnancy in women attending an early pregnancy assessment unit? Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2016;48(5):656-62.
25. Taraborrelli S. Physiology, production and action of progesterone. Acta obstetricia et gynecologica Scandinavica. 2015;94 Suppl 161:8-16.

26. Di Renzo GC, Giardina I, Clerici G, Mattei A, Alajmi AH, Gerli S. The role of progesterone in maternal and fetal medicine. Gynecol Endocrinol. 2012;28(11):925-32.
27. Druckmann R, Druckmann MA. Progesterone and the immunology of pregnancy. J Steroid Biochem Mol Biol. 2005;97(5):389-96.
28. Szekeres-Bartho J, Barakonyi A, Par G, Polgar B, Palkovics T, Szereday L. Progesterone as an immunomodulatory molecule. Int Immunopharmacol. 2001;1(6):1037-48.
29. Schwartz N, Xue X, Elovitz MA, Dowling O, Metz CN. Progesterone suppresses the fetal inflammatory response ex vivo. Am J Obstet Gynecol. 2009;201(2):211 e1-9.
30. Raghupathy R, Al-Mutawa E, Al-Azemi M, Makhseed M, Azizieh F, Szekeres-Bartho J. Progesterone-induced blocking factor (PIBF) modulates cytokine production by lymphocytes from women with recurrent miscarriage or preterm delivery. Journal of reproductive immunology. 2009;80(1-2):91-9.
31. Szekeres-Bartho J, Par G, Dombay G, Smart YC, Volgyi Z. The antiabortive effect of progesterone-induced blocking factor in mice is manifested by modulating NK activity. Cell Immunol. 1997;177(2):194-9.
32. Fanchin R, Ayoubi JM, Olivennes F, Righini C, de Ziegler D, Frydman R. Hormonal influence on the uterine contractility during ovarian stimulation. Human reproduction (Oxford, England). 2000;15 Suppl 1:90-100.
33. Perusquia M, Jasso-Kamel J. Influence of 5alpha- and 5beta-reduced progestins on the contractility of isolated human myometrium at term. Life Sci. 2001;68(26):2933-44.
34. Chanrachakul B, Broughton Pipkin F, Warren AY, Arulkumaran S, Khan RN. Progesterone enhances the tocolytic effect of ritodrine in isolated pregnant human myometrium. Am J Obstet Gynecol. 2005;192(2):458-63.
35. Liu J, Matsuo H, Laoag-Fernandez JB, Xu Q, Maruo T. The effects of progesterone on apoptosis in the human trophoblast-derived HTR-8/SV neo cells. Molecular human reproduction. 2007;13(12):869-74.
36. Czajkowski K, Sienko J, Mogilinski M, Bros M, Szczecina R, Czajkowska A. Uteroplacental circulation in early pregnancy complicated by threatened abortion supplemented with vaginal micronized progesterone or oral dydrogesterone. Fertil Steril. 2007;87(3):613-8.
37. Verhaegen J, Gallos ID, van Mello NM, Abdel-Aziz M, Takwoingi Y, Harb H, Deeks JJ, Mol BW, Coomarasamy A. Accuracy of single progesterone test to predict early pregnancy outcome in women with pain or bleeding: meta-analysis of cohort studies. BMJ (Clinical research ed). 2012;345:e6077.
38. Day A, Sawyer E, Mavrelos D, Tailor A, Helmy S, Jurkovic D. Use of serum progesterone measurements to reduce need for follow-up in women with pregnancies of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2009;33(6):704-10.
39. Cordina M, Schramm-Gajraj K, Ross JA, Lautman K, Jurkovic D. Introduction of a single visit protocol in the management of selected patients with pregnancy of unknown location: a prospective study. Bjog. 2011;118(6):693-7.
40. Van Calster B, Bobdiwala S, Guha S, Van Hoorde K, Al-Memar M, Harvey R, Farren J, Kirk E, Condous G, Sur S, Stalder C, Timmerman D, Bourne T. Managing

pregnancy of unknown location based on initial serum progesterone and serial serum hCG levels: development and validation of a two-step triage protocol. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2016;48(5):642-9.
41. Kirk E, Condous G, Van Calster B, Van Huffel S, Timmerman D, Bourne T. Rationalizing the follow-up of pregnancies of unknown location. Human reproduction (Oxford, England). 2007;22(6):1744-50.
42. Dirnhofer S, Madersbacher S, Bidart JM, Ten Kortenaar PB, Spottl G, Mann K, Wick G, Berger P. The molecular basis for epitopes on the free beta-subunit of human chorionic gonadotrophin (hCG), its carboxyl-terminal peptide and the hCG beta-core fragment. J Endocrinol. 1994;141(1):153-62.
43. Cole LA. Immunoassay of human chorionic gonadotropin, its free subunits, and metabolites. Clinical chemistry. 1997;43(12):2233-43.
44. Spencer K, Coombes EJ, Mallard AS, Ward AM. Free beta human choriogonadotropin in Down's syndrome screening: a multicentre study of its role compared with other biochemical markers. Ann Clin Biochem. 1992;29 (Pt 5):506-18.
45. van Mello NM, Mol F, Opmeer BC, Ankum WM, Barnhart K, Coomarasamy A, Mol BW, van der Veen F, Hajenius PJ. Diagnostic value of serum hCG on the outcome of pregnancy of unknown location: a systematic review and meta-analysis. Hum Reprod Update. 2012;18(6):603-17.
46. Condous G, Kirk E, Van Calster B, Van Huffel S, Timmerman D, Bourne T. Failing pregnancies of unknown location: a prospective evaluation of the human chorionic gonadotrophin ratio. Bjog. 2006;113(5):521-7.
47. Condous G, Kirk E, Lu C, Van Huffel S, Gevaert O, De Moor B, De Smet F, Timmerman D, Bourne T. Diagnostic accuracy of varying discriminatory zones for the prediction of ectopic pregnancy in women with a pregnancy of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2005;26(7):770-5.
48. Condous G, Okaro E, Khalid A, Timmerman D, Lu C, Zhou Y, Van Huffel S, Bourne T. The use of a new logistic regression model for predicting the outcome of pregnancies of unknown location. Human reproduction (Oxford, England). 2004;19(8):1900-10.
49. Condous G, Van Calster B, Kirk E, Timmerman D, Van Huffel S, Bourne T. Prospective cross-validation of three methods of predicting failing pregnancies of unknown location. Human reproduction (Oxford, England). 2007;22(4):1156-60.
50. Guha S, Ayim F, Ludlow J, Sayasneh A, Condous G, Kirk E, Stalder C, Timmerman D, Bourne T, Van Calster B. Triaging pregnancies of unknown location: the performance of protocols based on single serum progesterone or repeated serum hCG levels. Human reproduction (Oxford, England). 2014;29(5):938-45.
51. Cartwright J, Duncan WC, Critchley HO, Horne AW. Serum biomarkers of tubal ectopic pregnancy: current candidates and future possibilities. Reproduction (Cambridge, England). 2009;138(1):9-22.
52. Costa TN, Cassaro Strunz CM, Nicolau JC, Gutierrez PS. Comparison of MB fraction of creatine kinase mass and troponin I serum levels with necropsy findings in acute myocardial infarction. Am J Cardiol. 2008;101(3):311-4.

53. Lavie O, Beller U, Neuman M, Ben-Chetrit A, Gottcshalk-Sabag S, Diamant YZ. Maternal serum creatine kinase: a possible predictor of tubal pregnancy. Am J Obstet Gynecol. 1993;169(5):1149-50.
54. Duncan WC, Sweeting VM, Cawood P, Illingworth PJ. Measurement of creatine kinase activity and diagnosis of ectopic pregnancy. British journal of obstetrics and gynaecology. 1995;102(3):233-7.
55. Darai E, Vlastos G, Benifla JL, Sitbon D, Hassid J, Dehoux M, Madelenat P, Durand Gaucher G, Nunez E. Is maternal serum creatine kinase actually a marker for early diagnosis of ectopic pregnancy? European journal of obstetrics, gynecology, and reproductive biology. 1996;68(1-2):25-7.
56. Garcia-Velasco JA, Requena A, Santisteban J. Maternal serum creatine kinase: a possible predictor of tubal pregnancy? Am J Obstet Gynecol. 1996;175(1):238-40.
57. Korhonen J, Alfthan H, Stenman UH, Ylostalo P. Failure of creatine kinase to predict ectopic pregnancy. Fertil Steril. 1996;65(5):922-4.
58. Lincoln SR, Dockery JR, Long CA, Rock WA, Jr., Cowan BD. Maternal serum creatine kinase does not predict tubal pregnancy. Journal of assisted reproduction and genetics. 1996;13(9):702-4.
59. Qasim SM, Trias A, Sachdev R, Kemmann E. Evaluation of serum creatine kinase levels in ectopic pregnancy. Fertil Steril. 1996;65(2):443-5.
60. Vandermolen DT, Borzelleca JF. Serum creatine kinase does not predict ectopic pregnancy. Fertil Steril. 1996;65(5):916-21.
61. Plewa MC, Ledrick D, Buderer NF, King RW. Serum creatine kinase is an unreliable predictor of ectopic pregnancy. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine. 1998;5(4):300-3.
62. Chandra L, Jain A. Maternal serum creatine kinase as a biochemical marker of tubal pregnancy. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 1995;49(1):21-3.
63. Birkhahn RH, Gaeta TJ, Leo PJ, Bove JJ. The utility of maternal creatine kinase in the evaluation of ectopic pregnancy. The American journal of emergency medicine. 2000;18(6):695-7.
64. Guha S, Van Belle V, Bottomley C, Preisler J, Vathanan V, Sayasneh A, Stalder C, Timmerman D, Bourne T. External validation of models and simple scoring systems to predict miscarriage in intrauterine pregnancies of uncertain viability. Human reproduction (Oxford, England). 2013;28(11):2905-11.
65. Birkhahn RH, Gaeta TJ, Paraschiv D, Bove JJ, Suzuki T, Katoh H, Nagai R. Serum levels of myoglobin, creatine phosphokinase, and smooth muscle heavy-chain myosin in patients with ectopic pregnancy. Ann Emerg Med. 2001;38(6):628-32.
66. Kurzel RB, Mazdisnian F, Paige S, Liu P. Serum creatine kinase is not a reliable indicator of ectopic pregnancy. Int J Fertil Womens Med. 2001;46(6):300-3.
67. Soundravally R, Soundara Raghavan S, Selvaraj N. Serum creatine kinase as a predictor of tubal ectopic pregnancy. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2007;98(3):253-4.

68. Develioglu OH, Askalli C, Uncu G, Samli B, Daragenli O. Evaluation of serum creatine kinase in ectopic pregnancy with reference to tubal status and histopathology. BJOG. 2002;109(2):121-8.
69. Wodzig KW, Kragten JA, Hermens WT, Glatz JF, van Dieijen-Visser MP. Estimation of myocardial infarct size from plasma myoglobin or fatty acid-binding protein. Influence of renal function. Eur J Clin Chem Clin Biochem. 1997;35(3):191-8.
70. Enoki Y, Morimoto T. Gizzard myoglobin contents and feeding habits in avian species. Comp Biochem Physiol A Mol Integr Physiol. 2000;125(1):33-43.
71. Klocke FJ, Copley DP, Krawczyk JA, Reichlin M. Rapid renal clearance of immunoreactive canine plasma myoglobin. Circulation. 1982;65(7):1522-8.
72. Katoh H, Suzuki T, Yokomori K, Suzuki S, Ohtaki E, Watanabe M, Yazaki Y, Nagai R. A novel immunoassay of smooth muscle myosin heavy chain in serum. J Immunol Methods. 1995;185(1):57-63.
73. Fraser HM, Wilson H, Silvestri A, Morris KD, Wiegand SJ. The role of vascular endothelial growth factor and estradiol in the regulation of endometrial angiogenesis and cell proliferation in the marmoset. Endocrinology. 2008;149(9):4413-20.
74. Fraser HM, Duncan WC. Vascular morphogenesis in the primate ovary. Angiogenesis. 2005;8(2):101-16.
75. Ladoux A, Frelin C. Hypoxia is a strong inducer of vascular endothelial growth factor mRNA expression in the heart. Biochem Biophys Res Commun. 1993;195(2):1005-10.
76. Daniel Y, Geva E, Lerner-Geva L, Eshed-Englender T, Gamzu R, Lessing JB, Bar-Am A, Amit A. Levels of vascular endothelial growth factor are elevated in patients with ectopic pregnancy: is this a novel marker? Fertil Steril. 1999;72(6):1013-7.
77. Daponte A, Pournaras S, Zintzaras E, Kallitsaris A, Lialios G, Maniatis AN, Messinis IE. The value of a single combined measurement of VEGF, glycodelin, progesterone, PAPP-A, HPL and LIF for differentiating between ectopic and abnormal intrauterine pregnancy. Human reproduction (Oxford, England). 2005;20(11):3163-6.
78. Mueller MD, Raio L, Spoerri S, Ghezzi F, Dreher E, Bersinger NA. Novel placental and nonplacental serum markers in ectopic versus normal intrauterine pregnancy. Fertil Steril. 2004;81(4):1106-11.
79. Licht P, Neuwinger J, Fischer O, Siebzehnrubl E, Wildt L. Peripheral levels of vascular endothelial growth factor (VEGF) are higher in gonadotropin stimulated as compared to natural ovarian cycles. Exp Clin Endocrinol Diabetes. 2001;109(6):345-9.
80. Fasouliotis SJ, Spandorfer SD, Witkin SS, Liu HC, Roberts JE, Rosenwaks Z. Maternal serum vascular endothelial growth factor levels in early ectopic and intrauterine pregnancies after in vitro fertilization treatment. Fertil Steril. 2004;82(2):309-13.
81. Felemban A, Sammour A, Tulandi T. Serum vascular endothelial growth factor as a possible marker for early ectopic pregnancy. Human reproduction (Oxford, England). 2002;17(2):490-2.
82. Kucera-Sliutz E, Schiebel I, Konig F, Leodolter S, Sliutz G, Koelbl H. Vascular endothelial growth factor (VEGF) and discrimination between

abnormal intrauterine and ectopic pregnancy. Human reproduction (Oxford, England). 2002;17(12):3231-4.
83. Senturk LM, Arici A. Leukemia inhibitory factor in human reproduction. American journal of reproductive immunology (New York, NY : 1989). 1998;39(2):144-51.
84. Wegner NT, Mershon JL. Evaluation of leukemia inhibitory factor as a marker of ectopic pregnancy. Am J Obstet Gynecol. 2001;184(6):1074-6.
85. Ruge S, Sorensen S, Vejtorp M, Vejerslev LO. The secretory endometrial protein, placental protein 14, in women with ectopic gestation. Fertil Steril. 1992;57(1):102-6.
86. Julkunen M, Koistinen R, Sjoberg J, Rutanen EM, Wahlstrom T, Seppala M. Secretory endometrium synthesizes placental protein 14. Endocrinology. 1986;118(5):1782-6.
87. Stabile I, Olajide F, Chard T, Grudzinskas JG. Circulating levels of placental protein 14 in ectopic pregnancy. British journal of obstetrics and gynaecology. 1994;101(9):762-4.
88. Pedersen JF, Ruge S. Serum level of secretory endometrial protein PP-14 in intact ectopic pregnancy. British journal of obstetrics and gynaecology. 1991;98(4):414.
89. Foth D, Romer T. Glycodelin serum levels in women with ectopic pregnancy. European journal of obstetrics, gynecology, and reproductive biology. 2003;108(2):199-202.
90. Horne AW, van den Driesche S, King AE, Burgess S, Myers M, Ludlow H, Lourenco P, Ghazal P, Williams AR, Critchley HO, Duncan WC. Endometrial inhibin/activin beta-B subunit expression is related to decidualization and is reduced in tubal ectopic pregnancy. The Journal of clinical endocrinology and metabolism. 2008;93(6):2375-82.
91. Lin TM, Halbert SP, Kiefer D. Quantitative analysis of pregnancy-associated plasma proteins in human placenta. J Clin Invest. 1976;57(2):466-72.
92. Bischof P, Reyes H, Herrmann WL, Sizonenko PC. Circulating levels of pregnancy-associated plasma protein-A (PAPP-A) and human chorionic gonadotrophin (hCG) in intrauterine and extrauterine pregnancies. British journal of obstetrics and gynaecology. 1983;90(4):323-5.
93. Sinosich MJ, Ferrier A, Teisner B, Porter R, Westergaard JG, Saunders DM, Grudzinskas JG. Circulating and tissue concentrations of pregnancy-associated plasma protein-A (PAPP-A) in tubal ectopic gestation. Clin Reprod Fertil. 1985;3(4):311-7.
94. Sjoberg J. Pregnancy-associated plasma protein A in pregnancy-related gynaecologic emergency. Human reproduction (Oxford, England). 1987;2(7):615-6.
95. Horne CH, Towler CM, Pugh-Humphreys RG, Thomson AW, Bohn H. Pregnancy specific beta1-glycoprotein--a product of the syncytiotrophoblast. Experientia. 1976;32(9):1197.
96. Ho PC, Jones WR. Pregnancy-specific beta 1-glycoprotein as a prognostic indicator in complications of early pregnancy. Am J Obstet Gynecol. 1980;138(3):253-6.
97. Tornehave D, Chemnitz J, Westergaard JG, Teisner B, Poulsen HK, Bolton AE, Grudzinskas JG. Placental proteins in peripheral blood and tissues of ectopic pregnancies. Gynecologic and obstetric investigation. 1987;23(2):97-102.

98. Mantzavinos T, Phocas I, Chrelias H, Sarandakou A, Zourlas PA. Serum levels of steroid and placental protein hormones in ectopic pregnancy. European journal of obstetrics, gynecology, and reproductive biology. 1991;39(2):117-22.
99. Handwerger S, Freemark M. The roles of placental growth hormone and placental lactogen in the regulation of human fetal growth and development. J Pediatr Endocrinol Metab. 2000;13(4):343-56.
100. Walker WH, Fitzpatrick SL, Barrera-Saldana HA, Resendez-Perez D, Saunders GF. The human placental lactogen genes: structure, function, evolution and transcriptional regulation. Endocr Rev. 1991;12(4):316-28.
101. Kuscu E, Vicdan K, Turhan NO, Oguz S, Zorlu G, Gokmen O. The hormonal profile in ectopic pregnancies. Materia medica Polona Polish journal of medicine and pharmacy. 1993;25(3-4):149-52.
102. Petraglia F, Sawchenko P, Lim AT, Rivier J, Vale W. Localization, secretion, and action of inhibin in human placenta. Science. 1987;237(4811):187-9.
103. Wen X, Tozer AJ, Li D, Docherty SM, Al-Shawaf T, Iles RK. Human granulosa-lutein cell in vitro production of progesterone, inhibin A, inhibin B, and activin A are dependent on follicular size and not the presence of the oocyte. Fertil Steril. 2008;89(5 Suppl):1406-13.
104. Florio P, Severi FM, Bocchi C, Luisi S, Mazzini M, Danero S, Torricelli M, Petraglia F. Single serum activin a testing to predict ectopic pregnancy. The Journal of clinical endocrinology and metabolism. 2007;92(5):1748-53.
105. Kirk E, Papageorghiou AT, Van Calster B, Condous G, Cowans N, Van Huffel S, Timmerman D, Spencer K, Bourne T. The use of serum inhibin A and activin A levels in predicting the outcome of 'pregnancies of unknown location'. Human reproduction (Oxford, England). 2009;24(10):2451-6.
106. Grosskinsky CM, Hage ML, Tyrey L, Christakos AC, Hughes CL. hCG, progesterone, alpha-fetoprotein, and estradiol in the identification of ectopic pregnancy. Obstet Gynecol. 1993;81(5 (Pt 1)):705-9.
107. Jones EA, Clement-Jones M, James OF, Wilson DI. Differences between human and mouse alpha-fetoprotein expression during early development. J Anat. 2001;198(Pt 5):555-9.
108. Matthews CP, Coulson PB, Wild RA. Serum progesterone levels as an aid in the diagnosis of ectopic pregnancy. Obstet Gynecol. 1986;68(3):390-4.
109. Yeko TR, Gorrill MJ, Hughes LH, Rodi IA, Buster JE, Sauer MV. Timely diagnosis of early ectopic pregnancy using a single blood progesterone measurement. Fertil Steril. 1987;48(6):1048-50.
110. Buck RH, Joubert SM, Norman RJ. Serum progesterone in the diagnosis of ectopic pregnancy: a valuable diagnostic test? Fertil Steril. 1988;50(5):752-5.
111. Riss PA, Radivojevic K, Bieglmayer C. Serum progesterone and human chorionic gonadotropin in very early pregnancy: implications for clinical management. European journal of obstetrics, gynecology, and reproductive biology. 1989;32(2):71-7.
112. Guillaume J, Benjamin F, Sicuranza BJ, Deutsch S, Seltzer VL, Tores W. Serum estradiol as an aid in the diagnosis of ectopic pregnancy. Obstet Gynecol. 1990;76(6):1126-9.
113. Witt BR, Wolf GC, Wainwright CJ, Johnston PD, Thorneycroft IH. Relaxin, CA-125, progesterone, estradiol, Schwangerschaft protein, and human chorionic gonadotropin as predictors of outcome in threatened and nonthreatened pregnancies. Fertil Steril. 1990;53(6):1029-36.

114. Daily CA, Laurent SL, Nunley WC, Jr. The prognostic value of serum progesterone and quantitative beta-human chorionic gonadotropin in early human pregnancy. Am J Obstet Gynecol. 1994;171(2):380-3; discussion 3-4.
115. Ledger WL, Sweeting VM, Chatterjee S. Rapid diagnosis of early ectopic pregnancy in an emergency gynaecology service--are measurements of progesterone, intact and free beta human chorionic gonadotrophin helpful? Human reproduction (Oxford, England). 1994;9(1):157-60.
116. McCord ML, Muram D, Buster JE, Arheart KL, Stovall TG, Carson SA. Single serum progesterone as a screen for ectopic pregnancy: exchanging specificity and sensitivity to obtain optimal test performance. Fertil Steril. 1996;66(4):513-6.
117. O'Leary P, Nichols C, Feddema P, Lam T, Aitken M. Serum progesterone and human chorionic gonadotrophin measurements in the evaluation of ectopic pregnancy. The Australian & New Zealand journal of obstetrics & gynaecology. 1996;36(3):319-23.
118. Buckley RG, King KJ, Disney JD, Riffenburgh RH, Gorman JD, Klausen JH. Serum progesterone testing to predict ectopic pregnancy in symptomatic first-trimester patients. Ann Emerg Med. 2000;36(2):95-100.
119. Condous G, Okaro E, Khalid A, Lu C, Van Huffel S, Timmerman D, Bourne T. A prospective evaluation of a single-visit strategy to manage pregnancies of unknown location. Human reproduction (Oxford, England). 2005;20(5):1398-403.
120. Mol BW, Lijmer JG, Ankum WM, van der Veen F, Bossuyt PM. The accuracy of single serum progesterone measurement in the diagnosis of ectopic pregnancy: a meta-analysis. Human reproduction (Oxford, England). 1998;13(11):3220-7.
121. Weiss G, O'Byrne EM, Steinetz BG. Relaxin: a product of the human corpus luteum of pregnancy. Science. 1976;194(4268):948-9.
122. Szlachter BN, Quagliarello J, Jewelewicz R, Osathanondh R, Spellacy WN, Weiss G. Relaxin in normal and pathogenic pregnancies. Obstet Gynecol. 1982;59(2):167-70.
123. Garcia A, Skurnick JH, Goldsmith LT, Emmi A, Weiss G. Human chorionic gonadotropin and relaxin concentrations in early ectopic and normal pregnancies. Obstet Gynecol. 1990;75(5):779-83.
124. Sealey JE, Cholst I, Glorioso N, Troffa C, Weintraub ID, James G, Laragh JH. Sequential changes in plasma luteinizing hormone and plasma prorenin during the menstrual cycle. The Journal of clinical endocrinology and metabolism. 1987;65(1):1-5.
125. Meunier K, Mignot TM, Maria B, Guichard A, Zorn JR, Cedard L. Predictive value of the active renin assay for the diagnosis of ectopic pregnancy. Fertil Steril. 1991;55(2):432-5.
126. Illingworth PJ, Groome NP, Duncan WC, Grant V, Tovanabutra S, Baird DT, McNeilly AS. Measurement of circulating inhibin forms during the establishment of pregnancy. The Journal of clinical endocrinology and metabolism. 1996;81(4):1471-5.
127. Treetampinich C, O'Connor AE, MacLachlan V, Groome NP, de Kretser DM. Maternal serum inhibin A concentrations in early pregnancy after IVF and embryo transfer reflect the corpus luteum contribution and pregnancy outcome. Human reproduction (Oxford, England). 2000;15(9):2028-32.

128. Seifer DB, Lambert-Messerlian GM, Canick JA, Frishman GN, Schneyer AL. Serum inhibin levels are lower in ectopic than intrauterine spontaneously conceived pregnancies. Fertil Steril. 1996;65(3):667-9.
129. D'Antona D, Mamers PM, Lowe PJ, Balazs N, Groome NP, Wallace EM. Evaluation of serum inhibin A as a surveillance marker after conservative management of tubal pregnancy. Human reproduction (Oxford, England). 1998;13(8):2305-7.
130. Canney PA, Moore M, Wilkinson PM, James RD. Ovarian cancer antigen CA125: a prospective clinical assessment of its role as a tumour marker. Br J Cancer. 1984;50(6):765-9.
131. Halila H, Stenman UH, Seppala M. Ovarian cancer antigen CA 125 levels in pelvic inflammatory disease and pregnancy. Cancer. 1986;57(7):1327-9.
132. Sadovsky Y, Pineda J, Collins JL. Serum CA-125 levels in women with ectopic and intrauterine pregnancies. The Journal of reproductive medicine. 1991;36(12):875-8.
133. Katsikis I, Rousso D, Farmakiotis D, Kourtis A, Diamanti-Kandarakis E, Panidis D. Receiver operator characteristics and diagnostic value of progesterone and CA-125 in the prediction of ectopic and abortive intrauterine gestations. European journal of obstetrics, gynecology, and reproductive biology. 2006;125(2):226-32.
134. Brumsted JR, Nakajima ST, Badger G, Riddick DH, Gibson M. Serum concentration of CA-125 during the first trimester of normal and abnormal pregnancies. The Journal of reproductive medicine. 1990;35(5):499-502.
135. Soriano D, Hugol D, Quang NT, Darai E. Serum concentrations of interleukin-2R (IL-2R), IL-6, IL-8, and tumor necrosis factor alpha in patients with ectopic pregnancy. Fertil Steril. 2003;79(4):975-80.
136. Rausch ME, Barnhart KT. Serum biomarkers for detecting ectopic pregnancy. Clinical obstetrics and gynecology. 2012;55(2):418-23.
137. Senapati S, Barnhart KT. Biomarkers for ectopic pregnancy and pregnancy of unknown location. Fertil Steril. 2013;99(4):1107-16.
138. Barnhart K, Speicher DW. Molecular diagnosis of ectopic pregnancy. Expert Rev Mol Diagn. 2011;11(8):759-62.
139. Young VB. The role of the microbiome in human health and disease: an introduction for clinicians. BMJ (Clinical research ed). 2017;356:j831.
140. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14.
141. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449(7164):804-10.
142. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.
143. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220-30.
144. Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, Foster JA, Forney LJ. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J. 2007;1(2):121-33.

145. MacIntyre DA, Chandiramani M, Lee YS, Kindinger L, Smith A, Angelopoulos N, Lehne B, Arulkumaran S, Brown R, Teoh TG, Holmes E, Nicoholson JK, Marchesi JR, Bennett PR. The vaginal microbiome during pregnancy and the postpartum period in a European population. Sci Rep. 2015;5:8988.
146. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A. 2011;108 Suppl 1:4680-7.
147. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, Ravel J. Temporal dynamics of the human vaginal microbiota. Sci Transl Med. 2012;4(132):132ra52.
148. Hyman RW, Fukushima M, Jiang H, Fung E, Rand L, Johnson B, Vo KC, Caughey AB, Hilton JF, Davis RW, Giudice LC. Diversity of the vaginal microbiome correlates with preterm birth. Reprod Sci. 2014;21(1):32-40.
149. Chaban B, Links MG, Jayaprakash TP, Wagner EC, Bourque DK, Lohn Z, Albert AY, van Schalkwyk J, Reid G, Hemmingsen SM, Hill JE, Money DM. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. Microbiome. 2014;2:23.
150. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods. 2007;69(2):330-9.
151. Ma J, Prince A, Aagaard KM. Use of whole genome shotgun metagenomics: a practical guide for the microbiome-minded physician scientist. Seminars in reproductive medicine. 2014;32(1):5-13.
152. Walther-Antonio MR, Jeraldo P, Berg Miller ME, Yeoman CJ, Nelson KE, Wilson BA, White BA, Chia N, Creedon DJ. Pregnancy's stronghold on the vaginal microbiome. PLoS One. 2014;9(6):e98514.
153. Linhares IM, Summers PR, Larsen B, Giraldo PC, Witkin SS. Contemporary perspectives on vaginal pH and lactobacilli. Am J Obstet Gynecol. 2011;204(2):120 e1-5.
154. Reid G, Younes JA, Van der Mei HC, Gloor GB, Knight R, Busscher HJ. Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat Rev Microbiol. 2011;9(1):27-38.
155. Ayre WB. The glycogen-estrogen relationship in the vaginal tract. The Journal of clinical endocrinology and metabolism. 1951;11(1):103-10.
156. Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Bieda J, Chaemsaithong P, Miranda J, Chaiworapongsa T, Ravel J. The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. Microbiome. 2014;2:18.
157. Africa CW, Nel J, Stemmet M. Anaerobes and bacterial vaginosis in pregnancy: virulence factors contributing to vaginal colonisation. Int J Environ Res Public Health. 2014;11(7):6979-7000.
158. Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, Vaneechoutte M. Longitudinal analysis of the vaginal microflora in pregnancy suggests that L. crispatus promotes the stability of the normal vaginal microflora and that L. gasseri and/or L. iners are more conducive to the occurrence of abnormal vaginal microflora. BMC Microbiol. 2009;9:116.

159. Ralph SG, Rutherford AJ, Wilson JD. Influence of bacterial vaginosis on conception and miscarriage in the first trimester: cohort study. BMJ (Clinical research ed). 1999;319(7204):220-3.
160. van Oostrum N, De Sutter P, Meys J, Verstraelen H. Risks associated with bacterial vaginosis in infertility patients: a systematic review and meta-analysis. Human reproduction (Oxford, England). 2013;28(7):1809-15.
161. Nelson DB, Hanlon AL, Wu G, Liu C, Fredricks DN. First Trimester Levels of BV-Associated Bacteria and Risk of Miscarriage Among Women Early in Pregnancy. Matern Child Health J. 2015;19(12):2682-7.
162. Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, Raza S, Rosenbaum S, Van den Veyver I, Milosavljevic A, Gevers D, Huttenhower C, Petrosino J, Versalovic J. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One. 2012;7(6):e36466.
163. Sakai M, Ishiyama A, Tabata M, Sasaki Y, Yoneda S, Shiozaki A, Saito S. Relationship between cervical mucus interleukin-8 concentrations and vaginal bacteria in pregnancy. American journal of reproductive immunology (New York, NY : 1989). 2004;52(2):106-12.
164. Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini ME, Ghebremichael MS, Nusbaum C, Huttenhower C, Virgin HW, Ndung'u T, Dong KL, Walker BD, Fichorova RN, Kwon DS. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. Immunity. 2015;42(5):965-76.
165. Nold C, Anton L, Brown A, Elovitz M. Inflammation promotes a cytokine response and disrupts the cervical epithelial barrier: a possible mechanism of premature cervical remodeling and preterm birth. Am J Obstet Gynecol. 2012;206(3):208 e1-7.
166. Kindinger LM, MacIntyre DA, Lee YS, Marchesi JR, Smith A, McDonald JA, Terzidou V, Cook JR, Lees C, Israfil-Bayli F, Faiza Y, Toozs-Hobson P, Slack M, Cacciatore S, Holmes E, Nicholson JK, Teoh TG, Bennett PR. Relationship between vaginal microbial dysbiosis, inflammation, and pregnancy outcomes in cervical cerclage. Sci Transl Med. 2016;8(350):350ra102.
167. Kindinger LM, Bennett PR, Lee YS, Marchesi JR, Smith A, Cacciatore S, Holmes E, Nicholson JK, Teoh TG, MacIntyre DA. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. Microbiome. 2017;5(1):6.
168. Brown RG, Marchesi JR, Lee YS, Smith A, Lehne B, Kindinger LM, Terzidou V, Holmes E, Nicholson JK, Bennett PR, MacIntyre DA. Vaginal dysbiosis increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by erythromycin. BMC Med. 2018;16(1):9.
169. Haque MM, Merchant M, Kumar PN, Dutta A, Mande SS. First-trimester vaginal microbiome diversity: A potential indicator of preterm delivery risk. Sci Rep. 2017;7(1):16145.
170. Petricevic L, Domig KJ, Nierscher FJ, Sandhofer MJ, Fidesser M, Krondorfer I, Husslein P, Kneifel W, Kiss H. Characterisation of the vaginal Lactobacillus microbiota associated with preterm delivery. Sci Rep. 2014;4:5136.
171. Baldwin EA, Walther-Antonio M, MacLean AM, Gohl DM, Beckman KB, Chen J, White B, Creedon DJ, Chia N. Persistent microbial dysbiosis in preterm premature rupture of membranes from onset until delivery. PeerJ. 2015;3:e1398.

172. Giakoumelou S, Wheelhouse N, Cuschieri K, Entrican G, Howie SE, Horne AW. The role of infection in miscarriage. Hum Reprod Update. 2016;22(1):116-33.
173. Brunham RC, Gottlieb SL, Paavonen J. Pelvic inflammatory disease. The New England journal of medicine. 2015;372(21):2039-48.
174. Kosseim M, Brunham RC. Fallopian tube obstruction as a sequela to Chlamydia trachomatis infection. Eur J Clin Microbiol. 1986;5(5):584-90.
175. Brunham RC, Peeling R, Maclean I, Kosseim ML, Paraskevas M. Chlamydia trachomatis-associated ectopic pregnancy: serologic and histologic correlates. J Infect Dis. 1992;165(6):1076-81.
176. Lamont RF, Nhan-Chang CL, Sobel JD, Workowski K, Conde-Agudelo A, Romero R. Treatment of abnormal vaginal flora in early pregnancy with clindamycin for the prevention of spontaneous preterm birth: a systematic review and metaanalysis. Am J Obstet Gynecol. 2011;205(3):177-90.
177. Higham JM, O'Brien PM, Shaw RW. Assessment of menstrual blood loss using a pictorial chart. British journal of obstetrics and gynaecology. 1990;97(8):734-9.
178. Janssen CA, Scholten PC, Heintz AP. A simple visual assessment technique to discriminate between menorrhagia and normal menstrual blood loss. Obstet Gynecol. 1995;85(6):977-82.
179. Bottomley C, Van Belle V, Kirk E, Van Huffel S, Timmerman D, Bourne T. Accurate prediction of pregnancy viability by means of a simple scoring system. Human reproduction (Oxford, England). 2013;28(1):68-76.
180. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken). 2011;63 Suppl 11:S240-52.
181. Bobdiwala S, Saso S, Verbakel JY, Al-Memar M, Van Calster B, Timmerman D, Bourne T. Diagnostic protocols for the management of pregnancy of unknown location: a systematic review and meta-analysis. BJOG. 2019;126(2):190-8.
182. Ankum WM, Hajenius PJ. Discriminatory zone for human chorionic gonadotropin--level of certainty? Fertil Steril. 1995;63(3):683-4.
183. Hajenius PJ, Mol BW, Ankum WM, van der Veen F, Bossuyt PM, Lammes FB. Suspected ectopic pregnancy: expectant management in patients with negative sonographic findings and low serum hCG concentrations. Early pregnancy : biology and medicine : the official journal of the Society for the Investigation of Early Pregnancy. 1995;1(4):258-62.
184. Kaplan BC, Dart RG, Moskos M, Kuligowska E, Chun B, Adel Hamid M, Northern K, Schmidt J, Kharwadkar A. Ectopic pregnancy: prospective study with improved diagnostic accuracy. Ann Emerg Med. 1996;28(1):10-7.
185. Dart R, Kaplan B, Ortiz L, Cloherty J, Lavoie T. Normal intrauterine pregnancy is unlikely in emergency department patients with either menstrual days > 38 days or beta-hCG > 3,000 mIU/mL, but without a gestational sac on ultrasonography. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine. 1997;4(10):967-71.
186. Dart R, Howard K. Subclassification of indeterminate pelvic ultrasonograms: stratifying the risk of ectopic pregnancy. Academic emergency

medicine : official journal of the Society for Academic Emergency Medicine. 1998;5(4):313-9.
187. Mol BW, Hajenius PJ, Engelsbel S, Ankum WM, Van der Veen F, Hemrika DJ, Bossuyt PM. Serum human chorionic gonadotropin measurement in the diagnosis of ectopic pregnancy when transvaginal sonography is inconclusive. Fertil Steril. 1998;70(5):972-81.
188. Dart RG, Mitterando J, Dart LM. Rate of change of serial beta-human chorionic gonadotropin values as a predictor of ectopic pregnancy in patients with indeterminate transvaginal ultrasound findings. Ann Emerg Med. 1999;34(6):703-10.
189. Dart RG, Burke G, Dart L. Subclassification of indeterminate pelvic ultrasonography: prospective evaluation of the risk of ectopic pregnancy. Ann Emerg Med. 2002;39(4):382-8.
190. Kirk E, Condous G, Haider Z, Lu C, Van Huffel S, Timmerman D, Bourne T. The practical application of a mathematical model to predict the outcome of pregnancies of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2006;27(3):311-5.
191. Bignardi T, Condous G, Kirk E, Van Calster B, Van Huffel S, Timmerman D, Bourne T. Viability of intrauterine pregnancy in women with pregnancy of unknown location: prediction using human chorionic gonadotropin ratio vs. progesterone. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2010;35(6):656-61.
192. Bignardi T, Condous G, Alhamdan D, Kirk E, Van Calster B, Van Huffel S, Timmerman D, Bourne T. The hCG ratio can predict the ultimate viability of the intrauterine pregnancies of uncertain viability in the pregnancy of unknown location population. Human reproduction (Oxford, England). 2008;23(9):1964-7.
193. Rivera V, Nguyen PH, Sit A. Change in quantitative human chorionic gonadotropin after manual vacuum aspiration in women with pregnancy of unknown location. Am J Obstet Gynecol. 2009;200(5):e56-9.
194. Barnhart KT, Sammel MD, Appleby D, Rausch M, Molinaro T, Van Calster B, Kirk E, Condous G, Van Huffel S, Timmerman D, Bourne T. Does a prediction model for pregnancy of unknown location developed in the UK validate on a US population? Human reproduction (Oxford, England). 2010;25(10):2434-40.
195. Shaunik A, Kulp J, Appleby DH, Sammel MD, Barnhart KT. Utility of dilation and curettage in the diagnosis of pregnancy of unknown location. Am J Obstet Gynecol. 2011;204(2):130.e1-6.
196. Ko JK, Cheung VY. Time to revisit the human chorionic gonadotropin discriminatory level in the management of pregnancy of unknown location. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine. 2014;33(3):465-71.
197. Morse CB, Sammel MD, Shaunik A, Allen-Taylor L, Oberfoell NL, Takacs P, Chung K, Barnhart KT. Performance of human chorionic gonadotropin curves in women at risk for ectopic pregnancy: exceptions to the rules. Fertil Steril. 2012;97(1):101-6.e2.
198. Zee J, Sammel MD, Chung K, Takacs P, Bourne T, Barnhart KT. Ectopic pregnancy prediction in women with a pregnancy of unknown location: data

beyond 48 h are necessary. Human reproduction (Oxford, England). 2014;29(3):441-7.
199. Bobdiwala S, Guha S, Van Calster B, Ayim F, Mitchell-Jones N, Al-Memar M, Mitchell H, Stalder C, Bottomley C, Kothari A, Timmerman D, Bourne T. The clinical performance of the M4 decision support model to triage women with a pregnancy of unknown location as at low or high risk of complications. Human reproduction (Oxford, England). 2016;31(7):1425-35.
200. El Bishry G, Ganta S. The role of single serum progesterone measurement in conjunction with beta hCG in the management of suspected ectopic pregnancy. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2008;28(4):413-7.
201. Nowacek GE, Meyer WR, McMahon MJ, Thorp JR, Wells SR. Diagnostic value of cervical fetal fibronectin in detecting extrauterine pregnancy. Fertil Steril. 1999;72(2):302-4.
202. Condous G, Kirk E, Syed A, Van Calster B, Van Huffel S, Timmerman D, Bourne T. Do levels of serum cancer antigen 125 and creatine kinase predict the outcome in pregnancies of unknown location? Human reproduction (Oxford, England). 2005;20(12):3348-54.
203. Banerjee S, Aslam N, Zosmer N, Woelfer B, Jurkovic D. The expectant management of women with early pregnancy of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 1999;14(4):231-6.
204. Moschos E, Twickler DM. Endometrial thickness predicts intrauterine pregnancy in patients with pregnancy of unknown location. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2008;32(7):929-34.
205. Ellaithy M, Abdelaziz A, Hassan MF. Outcome prediction in pregnancies of unknown location using endometrial thickness measurement: is this of real clinical value? European journal of obstetrics, gynecology, and reproductive biology. 2013;168(1):68-74.
206. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, Group Q-. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
207. Steinhauser S, Schumacher M, Rucker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol. 2016;16(1):97.
208. Mavrelos D, Memtsa M, Helmy S, Derdelis G, Jauniaux E, Jurkovic D. beta-hCG resolution times during expectant management of tubal ectopic pregnancies. BMC women's health. 2015;15:43.
209. Banerjee S, Aslam N, Woelfer B, Lawrence A, Elson J, Jurkovic D. Expectant management of early pregnancies of unknown location: a prospective evaluation of methods to predict spontaneous resolution of pregnancy. BJOG. 2001;108(2):158-63.
210. Gevaert O, De Smet F, Kirk E, Van Calster B, Bourne T, Van Huffel S, Moreau Y, Timmerman D, De Moor B, Condous G. Predicting the outcome of pregnancies of unknown location: Bayesian networks with expert prior information compared to logistic regression. Human reproduction (Oxford, England). 2006;21(7):1824-31.

211. Van Calster B, Van Hoorde K, Froyman W, Kaijser J, Wynants L, Landolfo C, Anthoulakis C, Vergote I, Bourne T, Timmerman D. Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors. Facts Views Vis Obgyn. 2015;7(1):32-41.
212. Froyman W, Timmerman D. Methods of Assessing Ovarian Masses: International Ovarian Tumor Analysis Approach. Obstet Gynecol Clin North Am. 2019;46(4):625-41.
213. Pauker SG, Kassirer JP. Therapeutic decision making: a cost-benefit analysis. The New England journal of medicine. 1975;293(5):229-34.
214. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. Eur J Clin Invest. 2012;42(2):216-28.
215. Vickers AJ. Prediction models: revolutionary in principle, but do they do more good than harm? J Clin Oncol. 2011;29(22):2951-2.
216. (MHRA). Medicines and Healthcare products Regulatory Agency. The Good Clinical Practice Guide. TSO (The Stationery Office). 2014.
217. Kirk E, Condous G, Bourne T. Pregnancies of unknown location. Best Pract Res Clin Obstet Gynaecol. 2009;23(4):493-9.
218. Kirk E, Daemen A, Papageorghiou AT, Bottomley C, Condous G, De Moor B, Timmerman D, Bourne T. Why are some ectopic pregnancies characterized as pregnancies of unknown location at the initial transvaginal ultrasound examination? Acta obstetricia et gynecologica Scandinavica. 2008;87(11):1150-4.
219. Butts SF, Guo W, Cary MS, Chung K, Takacs P, Sammel MD, Barnhart KT. Predicting the decline in human chorionic gonadotropin in a resolving pregnancy of unknown location. Obstet Gynecol. 2013;122(2 Pt 1):337-43.
220. Silva C, Sammel MD, Zhou L, Gracia C, Hummel AC, Barnhart K. Human chorionic gonadotropin profile for women with ectopic pregnancy. Obstet Gynecol. 2006;107(3):605-10.
221. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ (Clinical research ed). 2009;338:b2393.
222. Farquhar CM. Ectopic pregnancy. Lancet. 2005;366(9485):583-91.
223. Mastromarino P, Di Pietro M, Schiavoni G, Nardis C, Gentile M, Sessa R. Effects of vaginal lactobacilli in Chlamydia trachomatis infection. Int J Med Microbiol. 2014;304(5-6):654-61.
224. Stoyancheva G, Marzotto M, Dellaglio F, Torriani S. Bacteriocin production and gene sequencing analysis from vaginal Lactobacillus strains. Arch Microbiol. 2014;196(9):645-53.
225. Rizzo A, Fiorentino M, Buommino E, Donnarumma G, Losacco A, Bevilacqua N. Lactobacillus crispatus mediates anti-inflammatory cytokine interleukin-10 induction in response to Chlamydia trachomatis infection in vitro. Int J Med Microbiol. 2015;305(8):815-27.
226. Nardini P, Nahui Palomino RA, Parolin C, Laghi L, Foschi C, Cevenini R, Vitali B, Marangoni A. Lactobacillus crispatus inhibits the infectivity of Chlamydia trachomatis elementary bodies, in vitro study. Sci Rep. 2016;6:29024.
227. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, Sun CL, Goltsman DS, Wong RJ, Shaw G, Stevenson DK, Holmes SP, Relman DA.

Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci U S A. 2015;112(35):11060-5.
228. Mitra A, MacIntyre DA, Lee YS, Smith A, Marchesi JR, Lehne B, Bhatia R, Lyons D, Paraskevaidis E, Li JV, Holmes E, Nicholson JK, Bennett PR, Kyrgiou M. Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. Sci Rep. 2015;5:16865.
229. Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, Pullen KM, El-Sayed YY, Taslimi MM, Batzoglou S, Ronaghi M. Bacterial flora-typing with targeted, chip-based Pyrosequencing. BMC Microbiol. 2007;7:108.
230. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013;79(17):5112-20.
231. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261-7.
232. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1.
233. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics. 2014;30(21):3123-4.
234. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 2015;43(W1):W566-70.
235. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12(6):R60.
236. Price MJ, Ades AE, Soldan K, Welton NJ, Macleod J, Simms I, DeAngelis D, Turner KM, Horner PJ. The natural history of Chlamydia trachomatis infection in women: a multi-parameter evidence synthesis. Health Technol Assess. 2016;20(22):1-250.
237. Kacerovsky M, Vrbacky F, Kutova R, Pliskova L, Andrys C, Musilova I, Menon R, Lamont R, Nekvindova J. Cervical microbiota in women with preterm prelabor rupture of membranes. PLoS One. 2015;10(5):e0126884.
238. Han YW, Shen T, Chung P, Buhimschi IA, Buhimschi CS. Uncultivated bacteria as etiologic agents of intra-amniotic inflammation leading to preterm birth. J Clin Microbiol. 2009;47(1):38-47.
239. Bottomley C, Bourne T. Diagnosing miscarriage. Best Pract Res Clin Obstet Gynaecol. 2009;23(4):463-77.

APPENDICES

Appendix 1: Early pregnancy clerking proforma

IMPERIAL COLLEGE NHS TRUST
EARLY PREGNANCY AND ACUTE GYNAECOLOGY UNIT

Hospital No: ☐

Surname: _____ First Name: _____

Date of Birth _____ Age: _____ Partners age _____

Ethnicity: Please circle

White British	White Irish	White Other	Asian Indian	Asian British	Asian all/other
Black Caribbean	Black British Caribbean	Black African	Black British African	Black Other	Mixed All Other
Asian Bangladeshi	Mixed White Asian	Asian Pakistani	Asian Pakistani British	Chinese	All Other

Marital status: Married Co-habiting Separated Divorced Single Widowed

Occupation: _____

Address: _____

Post Code: _____

Telephone No: (Home) _____ Mobile: _____

Name and number of Partner/ Relative/ Friend: _____

GP Name and Address: _____

Are you a UK Resident: YES / NO Date of entry to the UK_____

How were you referred to us, please circle: GP A&E WARD SELF

SECTION 1 - CURRENT PREGNANCY

First date of last menstrual period: __/___/__
Please mark on the line below how certain you are of this date:

Uncertain _____ Certain

Are your periods: Regular / Irregular

How many days are there between each period? _____ days.

Have you had a positive urine pregnancy test? YES/NO
On what date was your first positive pregnancy test? ___/___/___
Please mark on the line below how certain you are of this date:
Uncertain _____ Certain

Is this pregnancy: Natural conception / assisted conception (IVF/ICSI/IUI)
Date of embryo transfer (if applicable) ____/___/____
Number of embryos transferred (if applicable) _____

What was your most recent method of contraception? _____
When did you stop this? _____

SECTION 2 – PREVIOUS PREGNANCIES
IF THIS IS YOUR FIRST PREGNANCY GO TO SECTION 2

Please provide details of any miscarriages, ectopic pregnancies or terminations you have had:

Month/ Year	How many weeks pregnant were you?	Outcome (miscarriage/ectopic)	Management	If you had surgery, what did you have?
		Miscarriage/ ectopic / termination	Natural/tablets or injection/surgery	Surgical management of miscarriage/keyhole surgery /open surgery
		Miscarriage/ ectopic / termination	Natural/tablets or injection/surgery	Surgical management of miscarriage/keyhole surgery /open surgery
		Miscarriage/ ectopic / termination	Natural/tablets or injection/surgery	Surgical management of miscarriage/keyhole surgery /open surgery
		Miscarriage/ ectopic / termination	Natural/tablets or injection/surgery	Surgical management of miscarriage/keyhole surgery /open surgery

Please provide details of any other pregnancies you have had:

Month/ Year	Length of pregnancy in weeks	Mode of delivery (vaginal delivery/caesarean section)	Indication	Birth weight (kg)
		Vaginal delivery/elective caesarean / emergency caesarean		
		Vaginal delivery/elective caesarean / emergency caesarean		
		Vaginal delivery/elective caesarean / emergency caesarean		
		Vaginal delivery/elective caesarean / emergency caesarean		

If you have had a previous pregnancy, please tick if you have had any of the following complications:

Preterm labour (before 37 weeks)	☐	Pre-eclampsia	☐
Placental abruption	☐	Waters breaking before 37 weeks	☐
Small baby (weighing less than 2.5kg)	☐	Stillbirth	☐

SECTION 3 - MEDICAL HISTORY

Do you take any regular medications? YES / NO If yes, what? _____
Do you take long-term steroids, aspirin or heparin injections? YES / NO

Have you taken any of these medications in your last menstrual cycle or the early stages of this pregnancy?
Emergency Pill / Clomid / Progesterone/ antibiotics

Do you have any medical problems / diseases? YES / NO
If yes, please circle if you have any of the following:
Hypertension / Diabetes / recent viral illness / kidney disease / heart disease / easy bruising/ bleeding during dental procedures/ bleeding or clotting problems/other

Have you had any previous surgery? YES / NO
If yes, what? _____
Have you had surgery to any of the following? Uterus/tubes/ ovaries/cervix/appendix

Have you had any previous pelvic infections / sexual transmitted disease? YES / NO
If yes, were you told you had Chlamydia / gonorrhoea / bacterial vaginosis / other

Are you taking folic acid? YES/NO

Are you a smoker? YES/NO Have you been smoking in this pregnancy? YES/NO
If yes, how many cigarettes do you smoke a day? _____

Have you been taking any recreational drugs in pregnancy? YES/NO
If yes, which drugs have you taken? _____

Have you been drinking alcohol in pregnancy? YES/NO
If yes, how many units do you consume per week? _____units/week

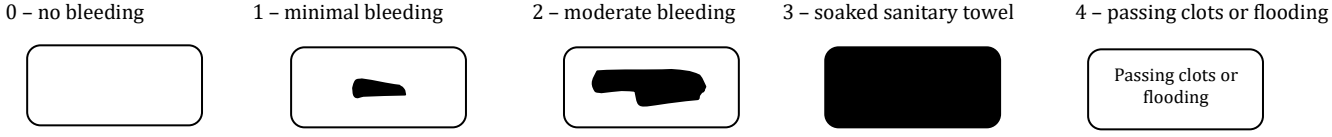
Have you had sexual intercourse in last 48 hours? YES/NO

SECTION 4 - Reason for referral today:

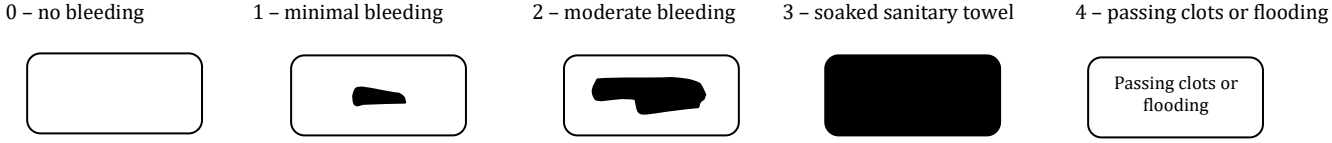
Bleeding ☐ How many days ago did the bleeding start? _____

Please circle one of the pictures below to describe your bleeding today and at its worst.

Today



Worst Bleeding



Pelvic pain ☐ How many days ago did the pain start? _____

Is the pain Left / Right / Both sides?

Please mark on the line below with a cross to indicate the level of your pain today

No pain ——————————————————— Worst pain

Please mark on the line below with a cross to indicate the level of your pain at its worst

No pain ——————————————————— Worst pain

Vomiting ☐

Please circle the box that corresponds to your vomiting

In the last 12 hours, how many hours have you felt nauseated or sick to your stomach?				
>6hrs (5pts)	4-6hrs (4pts)	2-3hrs (3pts)	<1hr (2pts)	Not at all (1pt)
In the last 12 hours, how many times have you vomited?				
>7 (5pts)	5-6 (4pts)	3-4 (3pts)	1-2 (2pts)	None (1pt)
In the last 12 hours, how many times have you had retching or dry heaves without bringing anything up?				
>7 (5pts)	5-6 (4pts)	3-4 (3pts)	1-2 (2pts)	None (1pt)

Other Symptoms: In the last one week have you had any of the following?

Dizziness	Yes / No	Breast pain	Yes / No
Loss of consciousness	Yes / No	Pain in chest and upper abdomen	Yes / No
Shoulder tip pain	Yes / No	Diarrhoea more than 3 times in the in last 24 hours	Yes / No
Pain when you pass urine	Yes / No	Pressure around the back passage (anus)	Yes / No

Have you had any scans already in this pregnancy? _____

If yes, what were you told it showed? _____
Where was the scan performed? _____

Please empty your bladder if you are less than 12 weeks pregnant as it is most likely that you will need an internal (vaginal) scan. This is not painful and is safe in pregnancy. In addition we will be taking your weight and height.

Thank you for your cooperation

Weight ____ Height____ BMI____

Appendix 2: Baseline characteristics (according to STARD criteria) of the 43 studies included in a systematic review and meta-analysis on the test performance of protocols used in the management of women with a pregnancy of unknown location (PUL). A total of 27,493 PUL were recruited and 23,802 PUL analysed in these studies. Reproduced from supplementary material published by Bobdiwala et al[181].

Author (year)	Study type & design	Study dates & location	Study groups, n	PUL definition	Index test	Final outcomes, PUL only, n
Ankum (1995)	Prosp	07/1989-01/1992 ED, TH, USA	Total, 85 Patients analysed: 85	TVS showing no evidence of IUP or EP	Single hCG <1500 hCG ratio	IUP, 19; EP, 20; FPUL/TR, 46
Hajenius (1995)	Prosp	07/1989-12/1994 ED, TH, Netherlands	Total, 265 Patients analysed: 265	No visible pregnancy on TVS	Single hCG >1000	IUP, 77; EP, 68 FPUL/ TR, 120
Kaplan (1996)	Prosp	08/1991-09/1992 ED, TH, USA	Total, 481 Patients analysed: 439	TVS +/- TAS showing no evidence of IUP or EP	Single hCG ≤1000	PUL: 72 VIUP, 11; NVIUP, 46; EP, 15
Dart (1997)	Retrosp	08/1991-12/1994 ED, TH, USA	Total, 248 Patients analysed: 194	TVS indeterminate	Single hCG >3000 Menstrual days >38	VIUP, 21; NVIUP, 146; EP, 27
Dart & Howard (1998)	Retrosp	08/1991-12/1994 ED, TH, USA	Total, 248 Patients analysed: 228	TVS showing no evidence of IUP or EP	Single hCG ≤1000	PUL: 228 IUP, 196; EP, 32
Mol (1998)	Prosp Multi-centre	09/1993-04/1996 x2 TH, Netherlands	Total, 354 Patients analysed: 262	No GS visualized within or outside the uterine cavity	hCG >1000; hCG >1500; hCG >2000	VIUP, 66; FPUL, 133; EP, 63
Banerjee (1999)	Prosp	08/1997-03/2008 EPAU, TH, UK	Total, 135 Patients analysed: 127	No evidence of IUP or EUP on TVS	LR model based on progesterone, bleeding score	FPUL, 64; IUP, 45; EP, 18
Dart (1999a)	Retrosp	08/1991-08/1997 ED, TH, USA	Total, 256 Patients analysed: 224		Single hCG ≤1000	VIUP, 23; NVIUP, 141; EP, 45

Dart (1999b)	Retrosp	08/1991-08/1998 ED, TH, USA	Total, 729 Patients analysed: 307	TVS shows no evidence of IUP or EP	hCG increase <66% hCG decrease <50%	VIUP, 53; NVIUP, 221; EP, 33
Nowacek (1999)	Prosp	Time period: NS EPAU/ IVF clinic/ ED, TH, USA	Total, 46 Patients analysed: 46	High risk for extrauterine pregnancy	FFN >50 ng/ml	EP, 12; IUP, 34
Banerjee (2001)	Prosp	Time period: NS EPAU, TH, UK	Total, 113 Patients analysed: 104	No clear US evidence of IUP, RPOC or EP	hCG <295; prog <20; ET <11mm; LR model 70%	FPUL, 72; NVIUP, 2; EP, 7l VIUP, 23
Dart (2002)	Prosp	01/2005-08/2000 ED, TH, USA	Total, 780 Patients analysed: 635	TVS classified as indeterminate	hCG <1000	NVIUP/FPUL, 476; IUP, 106; EP, 46; TOP, 7
Condous (2004)	Prosp	06/2001-12/2002 EPAU, TH, UK	Total, 388 Patients analysed: 388; 189 (training set); 199 (test set)	TVS: no signs of IUP, EUP or RPOC	LR models: M1; M2; M3	Training: IUP, 63; EP, 20; FPUL, 102 Test: IUP, 75; EP, 12; FPUL, 109
Condous (2005a)	Prosp	06/2001-04/2003 EPAU, TH, UK	Total, 569 Patients analysed: 527	TVS: no signs of IUP, EUP or RPOC	Single hCG >1000 hCG >1500 hCG >2000	FPUL, 300; IUP, 181; EP, 46
Condous (2005b)	Prosp	06/2001-04/2003 EPAU, TH, UK	Total, 560 Patients analysed: 518; 200 (training set); 318 (test set)	TVS: no IUP, EUP or RPOC	P ≤10 or hCG ≤25 or P >50	Training: IUP, 67; FPUL, 111; EP, 22 Test: IUP, 114; FPUL, 189; EP, 15
Condous (2005c)	Prosp	03/2002-07/2003 EPAU, TH, UK	Total, 403 Patients analysed: 297	No IUP or EP on TVS	CA125; CK	FPUL, 153; IUP, 116; EP, 28
Condous (2006)	Prosp	06/2001-10/2004 EPAU, TH, UK	Total, 1003 Patients analysed: 503	No signs of IUP, EP or RPOC on TVS	Uterine curettage if hCG (0h) ≥2000 + P (0h <45 or hCG (0h) <2000 + hCG ratio <1.15	FPUL, 533; IUP 379; EP, 91

Geveart (2006)	Prosp	06/2001-10/2004 EPAU, TH, UK	Total, 1003 Patients analysed: 798; 599 (model building); 199 (validation)	Inconclusive scan	Bayesian network parameter prior model >0.13; Structure & parameter prior model >0.06	FPUL, 460; IUP, 330; EP, 66
Kirk (2006)	Prosp	08/2003-09/2004 EPAU, TH, UK	Total, 474 Patients analysed: 357	No evidence of IUP or EUP on TVS	LR model + prediction by nurse practitioner	FPUL, 162; IUP, 166; EP, 29
Condous (2007a)	Prosp	03/2002-07/2003 EPAU, TH, UK	Total, 376 Patients analysed: 370; 197 (training set); 173 (test set)	TVS: no signs of IUP, EUP or RPOC	M5 model	Training: FPUL, 109; IUP, 76; EP, 12 Test: FPUL, 94; IUP, 64; EP, 15
Condous (2007b)	Prosp	03/2002-07/2003 EPAU, TH, UK	Total, 376 Patients analysed: 370; 197 (training set); 173 (test set)	TVS: no signs of IUP, EUP or RPOC	M4 model	Training: FPUL, 109; IUP, 76; EP, 12 Test: FPUL, 94; IUP, 64; EP, 15
Condous (2007c)	Prosp	07/2003-10/2004 EPAU, TH, UK	Total, 439 Patients analysed: 420	TVS: no signs of IUP, EUP or RPOC	hCG ratio: <0.87; ≤0.79 P <20	FPUL, 219; IUP, 167; EP, 34
Florio (2007)	Prosp	11/2004-09/2005 TH, Italy	Total, 536 Patients analysed: 536	No clear ultrasound evidence of IUP, RPOC or EP	hCG; P; Activin A	NVIUP, 305; VIUP, 155; EP, 76
Kirk (2007)	Prosp	02/2005-02/2006 EPAU, TH, UK	Total, 534 Patients analysed: 363	No evidence of IUP or EP on TVS	M4 model	FPUL, 229; IUP, 111; EP, 23
Bignardi (2008)	Prosp	06/2001-10/2004 EPAU, TH, UK	Total, 1003 Patients analysed: 940	No sign of IUP, EUP or RPOC	hCG ratio	FPUL, 533; VIUP, 276; NVIUP, 58; EP, 73
El Bishry (2008)	Retrosp	Time period: NS TH, UK	Total, 126 Patients analysed: 126	Inconclusive scan	P <16 & hCG >25; P=16–80 & hCG >25; P >80 & hCG <1000; P >80 & hCG >1000	IUP, 90; EP, 36

Moschos (2008)	Retrosp	07/2005-06/2006 TH, USA	Total, 591 Patients analysed: 559	No evidence of IUP or EP on TVS	ET	IUP, 41; EP, 43; FPUL, 475
Day (2009)	Prosp	01/2004-12/2006 EPAU, TH, UK	Total, 1410 Patients analysed: 1110	Pregnancy cannot be identified on TVS	P ≤20; P ≤10; P ≤10 & hCG <450	FPUL, 649; IUP, 376; EP, 85
Rivera (2009)	Prosp	07/2006-06/2008 Radiology dept, TH, Canada	Total, 23 Patients analysed: 22	No evidence on US of an IUP or EP	hCG decrease <50% after MVA	IUP, 13; EP, 9; LFU, 1
Barnhart (2010)	Retrosp	07/2003-10/2004 EPAU, TH, UK 02/2003-09/2007 ED, TH, USA	Total, 1579 UK patients analysed: 431 USA patients analysed: 604 Adj. USA patients: 544	Empty uterus + no IUP or EUP on TVS	M4 model	UK: EP, 26; IUP, 177; FPUL, 229 USA: EP, 96; IUP, 157; FPUL, 351 Adj. USA: EP, 104; IUP, 138; FPUL, 302
Bignardi (2010)	Prosp	06/2001-10/2004 EPAU, TH, UK	Total, 1003 Patients analysed: 940	No sign of IUP, EUP or RPOC	hCG ratio	FPUL, 533; VIUP, 276; NVIUP, 58; EP, 73
Chung (2011)	Retrosp	01/2004-12/2007 TH, USA	Total, 387 Patients analysed: 321	No clear evidence of IUP or EUP on TVS	Uterine curettage	NS
Cordina (2011)	Prosp	07/2007-09/2008 EPAU, TH, UK	Total, 252 Patients analysed: 227	No clear evidence of IUP or EUP on TVS	P ≤10	FPUL, 221; NVIUP, 1; EP, 5
Shaunik (2011)	Retrosp	12/2003-07/2007 ED, TH, USA	Total, 173 Patients analysed: 173	Non-diagnostic TVS	hCG <2000	NVIUP, 66; EP, 107
Majeed (2012)	Prosp	11/2008-12/2009 DGH, Denmark	Total, 105 Patients analysed: 104	No clear US evidence of IUP or EUP	hCG ratio <0.8; P <20; inhibin A <30pg/ml	VIUP, 37; NVIUP, 8; EP, 7; FPUL, 52

Morse (2012)	Retrosp Multi-centre (3)	10/2007-10/2009 ED, TH, USA	Total, 1180 Patients analysed: 1005	No signs of an IUP or EUP on TVS	hCG rise >71%; hCG rise >53%; hCG rise >35%	EP, 179; VIUP, 259; NVIUP, 567
Ellaithy (2013)	Prosp Multi-centre (3)	07/2009-01/2012 EPU, TH, Saudi Arabia	Total, 872 Patients analysed: 872	No evidence of an IUP or EUP on TVS	ET≥10.0mm; P ≥50; LR model I; LR model II	IUP, 78; EP/ PPUL, 97; FPUL, 697
Van Calster (2013)	Prosp Multi-centre (2)	Model development: 07/2003-02/2007; x2 EPAU, TH External validation: 04/2009-04/2011; x3 EPAU, TH; x1 EPAU, DGH, UK	Total, 2184 Patients analysed: 1962	No evidence of an IUP or EUP on TVS	M4 model	FPUL, 1104; IUP, 655; EP, 203
Guha (2014)	Retrosp	02/2005-08/2012 EPAU, TH, UK & Australia	Total, 1271 Patients analysed: 1271 (348 via MI)	No evidence of an IUP or EUP on TVS	P ≤10; hCG ratio; M4 model	FPUL, 774; IUP, 364; EP, 133
Ko (2014)	Retrosp	07/2007-06/2012 EPAU, TH, Hong Kong	Total, 113 Patients analysed: 113	Absence of an IUP or EUP on TVS	hCG >1000	EP, 22; VIUP, 23; NVIUP, 19; FPUL, 45
Zee (2014)	Retrosp Multi-centre (2)	2007-2009 ED, TH, USA	Total, 646 Patients analysed: 527	No signs of an IUP or EUP on TVS	hCG rise >66%; hCG rise >53%; hCG rise >35%	EP, 137; IUP, 390
Bobdiwala (2016)	Prosp Multi-centre (3)	08/2012-12/2013 EPAU, x2 TH, x1 DGH, UK	Total, 1022 Patients analysed: 835	TVS definitions according to a consensus paper (Barnhart et al, 2011)	M4 model	FPUL, 443; IUP, 298; EP, 94
Van Calster (2016)	Prosp Multi-centre (2)	07/2003-02/2007 04/2009-12/2013 EPAU, TH, UK	Total, 3152 Patients analysed: 2753; 1449 (development dataset); 1304 (validation dataset)	Using standardized protocol referenced in paper	Two-step protocol: [1] P ≤2; [2] M6 model	FPUL, 1450; IUP, 1002; EP, 301

Key:
CRL, crown rump length; D&C, dilatation and curettage; ED, emergency department; EP, ectopic pregnancy; ET, endometrial thickness; EUP, extra-uterine pregnancy; EPAU, early pregnancy assessment unit; FH, fetal heart; GS, gestational sac; hCG, human chorionic gonadotrophin; IU, intra-uterine; IUGS, intra-uterine gestational sac; IUP, intra-uterine pregnancy; LR, logistic regression; MI, multiple imputation; MVA, manual vacuum aspiration; NS, not specified; NVIUP, non-viable intra-uterine pregnancy; P, progesterone; Prosp, prospective, PUL, pregnancy of unknown location; Retrosp, retrospective, RPOC, retained products of conception; TAS, transabdominal scan; TH, teaching hospital; TR, trophoblast regression; TVS, transvaginal ultrasound scan; VIUP, viable intra-uterine pregnancy. All values for hCG are in IU/L and those for progesterone in nmol/l.

Appendix 3: Inclusion and exclusion criteria of each of the 43 studies included in the systematic review. Reproduced from supplementary material published by Bobdiwala et al[181].

Study ID	Author (year)	Inclusion criteria	Exclusion criteria
1	Ankum (1995)	1,2,3,9	NS
2	Hajenius (1995)	1,2,3,5	NS
3	Kaplan (1996)	1,2,3,4	A,B
4	Dart (1997)	1,2,3,5	E,F,W
5	Dart & Howard (1998)	1,2,3,5	C,D,E,F
6	Mol (1998)	1,2,5,7	V,W
7	Banerjee (1999)	5,6	K,N,O,P,Q,X
8	Dart (1999a)	1,2,3,5	D,G
9	Dart (1999b)	1,2,3,5	D
10	Nowacek (1999)	1,2,5,13	NS
11	Banerjee (2001)	1,2,3,5,6	D
12	Dart (2002)	1,2,3,5	D,E,F,M,N,O
13	Condous (2004)	1,2,5,6	K, N, O, P, Q
14	Condous (2005a)	3,5	K,N,O,P,Q
15	Condous (2005b)	1,2,5,6	N,O, P, Q
16	Condous (2005c)	NS	NS
17	Condous (2006)	6	NS
18	Geveart (2006)	5	T
19	Kirk (2006)	1,2,5,6,8	K,N,O,P,Q
20	Condous (2007a)	1,2,5	K,N,O,P,Q
21	Condous (2007b)	1,2,5	K,N,O,P,Q
22	Condous (2007c)	1,2,5	K,N,O,P,Q
23	Florio (2007)	5	K,O,Q,X
24	Kirk (2007)	1,2,5,6	K,N,O,P,Q,U
25	Bignardi (2008)	1,2,5,6	N,O, P, Q
26	El Bishry (2008)	5	N
27	Moschos (2008)	2,3,5,6	K,N,O,P
28	Day (2009)	3,6	K,N,O,P,Q
29	Rivera (2009)	6,10	N,O,L,AA
30	Barnhart (2010)	1,2,5	K,N,O,P,Q
31	Bignardi (2010)	1,2,5,6	N,O, P, Q
32	Chung (2011)	7	H
33	Cordina (2011)	1,2,5,6,8	NS
34	Shaunik (2011)	1,2,6	NS
35	Majeed (2012)	1,2,6	H, K,O,Q,S
36	Morse (2012)	5,11,12	W
37	Ellaithy (2013)	3,5	K,N,O,Q,X
38	Van Calster (2013)	6	K,Q,R
39	Guha (2014)	1,2,5,6	K.R
40	Ko (2014)	1,2,3	A,I,J,K,L,M
41	Zee (2014)	5,11,12	W

42	Bobdiwala (2016)	5,6	H,N,O,P,K
43	Van Calster (2016)	1,2,6	Y,Z

Key:

1, abdominal pain; 2, vaginal bleeding; 3, positive serum/urine hCG finding; 4, uterine size equivalent to <12/40 on pelvic examination; 5, inconclusive TVS/ no evidence of an intrauterine or extrauterine gestation on TVS; 6, all clinically stable women classified as having a PUL; 7, had D&C performed; 8, no US evidence of intraperitoneal bleeding; 9, clinical suspicion of EP; 10, hCG >2000; 11, at least 2 hCG values at least 1 day apart; 12, pregnancy outcome known after follow-up; 13, asymptomatic infertility patients at risk for extrauterine pregnancy

A, prior US documentation of an IUP; B, previous enrolment in the study; C, refusal of admission; D, lost to follow-up; E, recent delivery of a baby; F, recent D&C; G, endometrial thickness not identified; H, molar pregnancy; I, hCG<1000 IU/L, J, interval >12 hours between TVS and hCG blood test, K, haemodynamically unstable; L, heavy vaginal bleeding; M, history of passing a tissue mass vaginally; N, sac-like structure within the uterine cavity; O, an adnexal mass believed to be an EP; P, US suggestive of incomplete miscarriage; Q, significant haemoperitoneum; R, acute abdomen; S, severe abdominal pain; T, persistent PUL; U, hCG>10,000IU/L; V, IVF pregnancy; W, missing data; X, POC visualised on speculum examination/ sample given thought to represent POC; Y, on progesterone supplementation; Z, initial hCG <25 IU/L; AA, unable to tolerate outpatient MVA

Appendix 4: The test performance of various management protocols in women with a pregnancy of unknown location (PUL). Reproduced from supplementary material published by Bobdiwala et al[181].

EP = ectopic pregnancy; IUP = intrauterine pregnancy, FPUL = failed pregnancy of unknown location, hCG = human chorionic gonadotrophin, TP = true positive, FP = false positive, FN = false negative, TN = true negative, FFN = fetal fibronectin, AUC = area under the curve, LR = logistic regression, ET = endometrial thickness, CA125 = cancer antigen 125.

Author (year)	PUL population (n)	PUL management protocol used	PUL outcome		
			EP	IUP	FPUL
Single hCG					
Ankum (1995)	85	hCG <1500 IU/L	TP 18 FP 0 FN 2 TN 65	TP 19 FP 0 FN 0 TN 66	TP 46 FP 0 FN 2 TN 37
Hajenius (1995)	265	hCG >1000 IU/L	TP 61 FP 4 FN 7 TN 193	TP 77 FP 123 FN 0 TN 65	TP 116 FP 84 FN 4 TN 61
Kaplan (1996)	72	hCG ≤ 1000 mIU/ml	TP 10 FP 15 FN 5 TN 42	TP 15 FP 10 FN 42 TN 5	NS
Dart (1997)	194	Single hCG >3000 IU/L	TP 7 FP 74 FN 20 TN 93	TP 78 FP 7 FN 89 TN 20	NC
Dart & Howard (1998)	220	hCG ≤ 1000 miu/ml	TP 13 FP 47 FN 16 TN 144	NS	NS
Mol (1998)	262	hCG <1000 IU/l	TP 27 FP 28 FN 36 TN 171	TP 13 FP 9 FN 53 TN 187	TP 118 FP 89 FN 15 TN 40
		1,000 ≤hCG <1,500	TP 2 FP 20 FN 61 TN 179	TP 13 FP 9 FN 53 TN 187	TP 7 FP 15 FN 126 TN 114
		1,500 ≤hCG <2,000	TP 1 FP 5 FN 62 TN 194	TP 0 FP 6 FN 66 TN 190	TP 5 FP 1 FN 128 TN 128
		hCG ≥2000 IU/l	TP 24	TP 0	TP 24

			FP 3 FN 39 TN 196	FP 27 FN 66 TN 169	FP 3 FN 130 TN 105
Banerjee (1999)	127	hCG <285 IU/L (note < or > not specified by authors, presumed to be <)	NC	NC	TP 39 FP 23 FN 25 TN 40
Dart (1999a)	209	hCG ≤ 1000 miu/ml	TP 29 FP 94 FN 16 TN 70	TP 94 FP 29 FN 70 TN 16	NS
Banerjee (2001)	104	hCG <295 IU/L (note < or > not specified by authors, presumed to be <)			TP 30 FN 12 FP 42 TN 20 AUC 0.47
Dart (2002)	635	hCG <1000 miu/ml	TP 38 FP 208 FN 8 TN 359	NC	NC
		hCG >1000 IU/L	TP 10 FP 61 FN 36 TN 420	NC	NC
Condous (2005a)	527	hCG >1500 IU/L	TP 7 FP 32 FN 39 TN 449	NC	NC
		hCG >2000 IU/L	TP 5 FP 23 FN 41 TN 458	NC	NC
Florio (2007)	536	hCG <658IU/L	TP 57 FP 110 FN 19 TN 350	NC	NC
Chung (2011)	321	Uterine curettage if no visible IUP with hCG >2000 mIU/mL	TP 72 FP 54 FN 163 TN 32	TP 32 FP 163 FN 54 TN 72	NS
Shaunik (2011)	173	hCG <2000 mIU/mL	TP 93 FP 40 FN 14 TN 26	TP 32 FP 136 FN 54 TN 72	NS
Ko (2014)	113	hCG >1000 IU/L	TP 22 FP 91 FN nc	TP 42 FP 71 FN nc	NC

			TN nc	TN nc	
hCG ratio					
Dart (1999b)	307	hCG increase <66%	n=33 TP 17 FP 46 FN 16 TN 228	n=274 TP 46 FP 17 FN 228 TN 16	NS
		hCG decrease <50%	TP 8 FP 75 FN 25 TN 199	TP 75 FP 8 FN 199 TN 25	NS
		M1 model	Training set: TP 15 FP 19 FN 5 TN 146 Test set: TP 10 FP 25 FN 2 TN 159	Training set: TP 54 FP 2 FN 9 TN 120 Test set: TP 61 FP 3 FN 14 TN 118	Training set: TP 91 FP 4 FN 11 TN 79 Test set: TP 95 FP 2 FN 14 TN 85
Condous (2004)	189 (training set); 199 (test set)	M1 model (AUC)	Training set: 0.823 Test set: 0.914	Training set: 0.985 Test set: 0.963	Training set: 0.961 Test set: 0.973
		M2 model (AUC)	Training set: 0.867 Test set: 0.728	Training set: 0.960 Test set: 0.958	Training set: 0.943 Test set: 0.986
		M3 model (AUC)	Training set: 0.902 Test set: 0.834	Training set: 0.993 Test set: 0.959	Training set: 0.980 Test set: 0.975
Condous (2007c)	420	hCG ratio: <0.87	NC	NC	TP 189 FP 6 FN 30 TN 195

		hCG ratio: ≤0.79	NC	NC	TP 180 FP 4 FN 39 TN 197
		0.87 - <1.66	TP 51 FP 83 FN 22 TN 784	TP 34 FP 100 FN 300 TN 506	TP 49 FP 85 FN 484 TN 322
		<1.66	TP 62 FP 559 FN 11 TN 308	TP 34 FP 587 FN 300 TN 19	TP 525 FP 96 FN 8 TN 311
Bignardi (2008)	940	<2.00	TP 67 FP 627 FN 6 TN 240	TP 102 FP 592 FN 232 TN 14 NVIUP only: TP 39 FP 655 FN 19 TN 227	TP 525 FP 169 FN 8 TN 238
		≥2.00	TP 6 FP 240 FN 67 TN 627	TP 232 FP 14 FN 102 TN 592 VIUP only: TP 213 FP 33 FN 63 TN 631	TP 8 FP 238 FN 525 TN 169
Majeed (2012)	104	hCG ratio <0.8	NC	NC	TP 38 FP 1 FN 14 TN 51
Guha (2014)	1271	hCG ratio 0.87-1.66	TP 85 FP 129 FN 48 TN 1009	TP 52 FP 162 FN 312 TN 745	TP 77 FP 137 FN 697 TN 360
Morse (2012)	1005	hCG rise >71%	n=179 TP 165 FP 307 FN 14 TN 519	n=259 TP 130 FP 14 FN 49 TN 732	NC

			n=179	n=259	NC
		hCG rise >53%	TP 163 FP 276 FN 16 TN 550	TP 148 FP 21 FN 31 TN 725	
			n=179	n=259	NC
		hCG rise >35%	TP 149 FP 241 FN 30 TN 585	TP 239 FP 45 FN 20 TN 701	
			n=137	NC	NC
		hCG rise >66%	TP 110 FP 138 FN 27 TN 252		
			n=137	NC	NC
Zee (2014)	527	hCG rise >53%	TP 104 FP 127 FN 33 TN 263		
			n=137	NC	NC
		hCG rise >35%	TP 87 FP 114 FN 50 TN 276		
Progesterone					
Banerjee (1999)	127	Progesterone cut-off of 10 nmol/l	NC	NC	TP 48 FP 3 FN 16 TN 60
Banerjee (2001)	104	Progesterone <20 nmol/l			TP 67 FN 5 FP 2 TN 30 AUC 0.95
Florio (2007)	536	Progesterone <5.01 ng/ml	TP 65 FP 136 FN 11 TN 304	NC	NC
Condous (2007b)	420	Progesterone <20nmol/l	NC	NC	TP 191 FP 20 FN 28 TN 181

El Bishry (2008)	126	Progesterone <16 nmol/l	TP 16 FP 36 FN 20 TN 54	TP 36 FP 16 FN 54 TN 20	NC
		Progesterone 16-80 nmol/l	TP 19 FP 45 FN 17 TN 45	TP 45 FP 19 FN 45 TN 17	NC
		Progesterone >80 nmol/l	TP 1 FP 9 FN 35 TN 81	TP 9 FP 1 FN 81 TN 35	NC
Day (2009)	1110	Progesterone ≤20nmol/l	TP 28 FP 637 FN 57 TN 388	TP 51 FP 614 FN 325 TN 120	TP 586 FP 79 FN 63 TN 382
		Progesterone ≤10nmol/l	TP 12 FP 512 FN 73 TN 513	TP 30 FP 494 FN 346 TN 240	TP 482 FP 42 FN 167 TN 419
Bignardi (2010)		hCG ratio vs. progesterone		hCG ratio AUC = 0.756; initial progesterone AUC = 0.678 (95% CI, 0.603–0.753)	
Cordina (2011)	227	Progesterone ≤10nmol/l	TP 5 FP 222 FN NS TN NS	TP 1 FP 226 FN NS TN NS	TP 221 FP 6 FN NS TN NS
Majeed (2012)	104	Progesterone ≤20nmol/l	NC	NC	TP 50 FP 3 FN 2 TN 49
Ellaithy (2013)	872	Progesterone ≥50nmol/l	NC	TP 74 FP 84 FN 4 TN 710	NC
Guha (2014)	1271	Progesterone >10nmol/l	TP 99 FP 518 FN 34 TN 620	TP 345 FP 272 FN 19 TN 635	TP 173 FP 444 FN 601 TN 53
		Progesterone >16nmol/l	TP 74 FP 419 FN 59 TN 719	TP 334 FP 159 FN 30 TN 748	TP 85 FP 408 FN 689 TN 89
		Progesterone >20nmol/l	TP 62 FP 392 FN 71 TN 746	TP 324 FP 130 FN 43 TN774	TP 68 FP 386 FN 706 TN 111

Van Calster (2016)	Development dataset: 1449	Progesterone ≤2nmol/l	TP 160 FP 1089 FN 3 TN 197	NC	NC
		Progesterone ≤10nmol/l	TP 130 FP 683 FN 33 TN 603	NC	NC
	Validation dataset: 1304	Progesterone ≤2nmol/l	TP 134 FP 960 FN 4 TN 206	NC	NC
		Progesterone ≤10nmol/l	TP 108 FP 625 FN 30 TN 541	NC	NC
hCG and progesterone					
Condous (2005b)	Training set: 200 Test set: 318	Progesterone ≤10 nmol/l OR hCG of ≤25 IU/l OR Progesterone >50 nmol/l	NC	NC	TRAINING SET: TP 18 FP 41 FN 4 TN 137 TEST SET: TP 5 FP 48 FN 10 TN 255
El Bishry (2008)	126	Progesterone <16 nmol/l and hCG >25 IU/l	TP 16 FP 36 FN 20 TN 54	TP 36 FP 16 FN 54 TN 20	NC
		Progesterone 16–80 nmol/l and hCG >25 IU/l	TP 19 FP 45 FN 17 TN 45	TP 45 FP 19 FN 45 TN 17	NC
		Progesterone >80 nmol/l and hCG <1000 IU/l	TP 0 FP 1 FN 36 TN 89	TP 1 FP 0 FN 89 TN 36	NC
		Progesterone >80 nmol/l and hCG >1000 IU/l	TP 1 FP 9 FN 35 TN 82	TP 8 FP 1 FN 82 TN 35	NC
LR models					
Banerjee (1999)	127	LR model based on progesterone, bleeding score and ET	NC	NC	TP 59 FP 5 FN 5 TN 58
Banerjee (2001)	104	Cut-off of 70%			TP 59 FN 13 FP 3

					TN 29
		Hahlin's model			AUC 0.93 TP 63 FN 1 FP 9 TN 31
Geveart (2006)	Total, 1003	Parameter prior model >0.13	NS	TP 17 FP 40 FN 5 TN 195	AUC 0.92 NS
	599 (model building); 199 (validation)	Structure an parameter prior model >0.06	NS	TP 17 FP 47 FN 5 TN 188	NS
Kirk (2006)	357	LR model based on hCG ratio (M1 model) Cost value: 1:1:4	TP 22 FP 50 FN 7 TN 278	TP 147 FP 5 FN 19 TN 186	TP 127 FP 6 FN 35 TN 189
		LR model based on hCG ratio (M1 model) Cost value: 1:1:5	TP 24 FP 61 FN 5 TN 267	TP 144 FP 5 FN 22 TN 186	TP 119 FP 4 FN 43 TN 191
		LR model based on hCG ratio (M1 model) Cost value: 1:1:10	TP 25 FP 103 FN 4 TN 225	TP 128 FP 4 FN 38 TN 187	TP 94 FP 3 FN 68 TN 192
		M4 model (AUC, test set)	0.900	0.974	0.978
		M4 model (Test dataset)	TP 12 FP 18 FN 3 TN 140	TP 55 FP 4 FN 9 TN 105	TP 82 FP 2 FN 12 TN 77
Condous (2007a)	Training set: 197 Test set: 173	M5 model (AUC)	Training set: 0.938 Test set: 0.912	Training set: 0.979 Test set: 0.979	Training set: 0.994 Test set: 0.979
		M5 model (Training dataset)	TP 11 FP 20 FN 1 TN165	TP 56 FP 2 FN 20 TN 119	TP 106 FP 3 FN 2 TN 86
		M5 model (Test dataset)	TP 11 FP 16 FN 3 TN 142	TP 55 FP 3 FN 9 TN 105	TP 85 FP 2 FN 9 TN 76
		M1 model (AUC, test dataset)	0.842	0.969	0.965
Condous (2007b)	Training set: 197 Test set: 173	M1 model (test dataset)	TP 11 FP 20 FN 4 TN 138	TP 55 FP 5 FN 9 TN 104	TP 81 FP 1 FN 13 TN 78
		M4 model (AUC,	0.941	0.976	0.991

		training set)			
		M4 model (training set)	TP 11 FP 16 FN 1 TN 169	TP 60 FP 2 FN 16 TN 119	TP 106 FP 2 FN 3 TN 86
Condous (2007c)	420	Cut-off of 70%	NC	NC	TP 171 FP 23 FN 48 TN 178
Kirk (2007)	: 363	M4 model	TP 17 FP 38 FN 6 TN 302	TP 82 FP 4 FN 29 TN 248 VIUP only: N=40	TP 220 FP 2 FN 9 TN 132
Moschos (2008)	559	ET in a LR model Cut-off of 90%		TP 23 FN 17 FP 36 TN 482	
Day (2009)	1110	Progesterone ≤10nmol/l & hCG <450IU/L	TP 8 FP 367 FN 77 TN 658	TP 12 FP 363 FN 364 TN 371	TP 355 FP 20 FN 294 TN 441
Barnhart (2010)	UK population 431	M4 model	TP 21 FP 45 FN 5 TN 360	TP 155 FP 7 FN 22 TN 247	TP 200 FP 3 FN 28 TN 200
	USA population 604		TP 47 FP 64 FN 49 TN 444	TP 132 FP 32 FN 25 TN 415	TP 286 FP 43 FN 65 TN 210
Ellaithy (2013)	872	LR model I	NC	TP 77 FP 16 FN 1 TN 778	NC
		LR model II	NC	TP 78 FP 16 FN 0 TN 778	NC
Van Calster (2013)	1962	M4 model (EP: EP risk >5%; IUP: risk EP <5% and risk IUP > risk FPUL; FPUL: risk EP <5% and risk FPUL > risk IUP)	TP 175 FP 375 FN 28 TN 1384	TP 427 FP 28 FN 228 TN 1279	TP 931 FP 26 FN 173 TN 832
		M4 model (EP risk >3%)	TP 180 FP 449 FN 23 TN 1310	TP 384 FP 26 FN 268 TN 1286	TP 902 FP 23 FN 202 TN 835
Guha (2014)	1271	M4 model (EP: EP risk >5%; IUP: risk EP <5% and risk IUP > risk	TP 112 FP 239 FN 21 TN 899	TP 233 FP 687 FN 131 TN 220	TP 666 FP 254 FN 108 TN 243

Bobdiwala (2016)	835	FPUL; FPUL: risk EP <5% and risk FPUL > risk IUP) M4 model (EP: EP risk >5%; IUP: risk EP <5% and risk IUP > risk FPUL; FPUL: risk EP <5% and risk FPUL > risk IUP)	TP 77 FP 173 FN 17 TN 568	TP 160 FP 470 FN 138 TN67	TP 393 FP 237 FN 50 TN155
		M6P model	TP 154 FP 415 FN 9 TN 871	NC	NC
		M6NP model	TP 152 FP 535 FN 11 TN 751	NC	NC
	Development dataset: 1449	Two-step protocol: [1] progesterone cut-off of ≤2 nmol/l [2] M6P model	TP 151 FP 402 FN 12 TN 884	NC	NC
		Two-step protocol: [1] progesterone cut-off of ≤2 nmol/l [2] M6NP model	TP 149 FP 503 FN 14 TN 783	NC	NC
Van Calster (2016)		M6P model	TP 131 FP 384 FN 7 TN 782	NC	NC
		M6NP model	TP 128 FP 460 FN 10 TN 706	NC	NC
	Validation dataset: 1304	Two-step protocol: [1] progesterone cut-off of ≤2 nmol/l [2] M6P model	TP 127 FP 367 FN 11 TN 799	NC	NC
		Two-step protocol: [1] progesterone cut-off of ≤2 nmol/l [2] M6NP model	TP 124 FP 426 FN 14 TN 740	NC	NC
		M4 model	TP 112 FP 271 FN 26 TN 895	NC	NC

Endometrial thickness (ET)					
Banerjee (1999)	127	ET cut-off of 7.8mm	NC	NC	TP 33 FP 59 FN 31 TN 4
Dart (1999a)	209	ET <8mm	n=45 TP 34 FP 98 FN 11 TN 66 NS	N=164 TP 98 FP 34 FN 66 TN 11 NS	NS
Banerjee (2001)	104	ET<11.0mm			TP 44 FN 28 FP 11 TN 21 AUC 0.67
Moschos (2008)	559	ET, no cut-off given	NS	VIUP only: n=40 TP 3 FN 38 FP 23 TN 495	NS
Ellaithy (2013)	872	ET≥10.0mm	NC	TP 73 FP 520 FN 5 TN 274	NC
Other biomarkers					
Nowacek (1999)	46	FFN >50 ng/ml	TP 5 FP 12 FN 7 TN 22	TP 12 FP 5 FN 22 TN 7	NS
Condous (2005c)	Training set: 162 Test set: 135	CA125 ratio in a LR model (AUC)	Training set: 0.478 Test set: 0.560	Training set: 0.703 Test set: 0.768	Training set: 0.697 Test set: 0.782
Florio (2007)	536	Activin A <0.37 ng/ml	TP 75 FP 2 FN 1 TN 458	NC	NC
Majeed (2012)	104	Progesterone ≤20nmol/l & Inhibin A ≤30pg/ml	NC	NC	TP 50 FP 1 FN 2 TN 49
		Inhibin A ≤30pg/ml	NC	NC	TP 51 FP 23 FN 1 TN 29
		Progesterone ≤20nmol/l & hCG ratio ≥0.8	NC	NC	TP 45 FP 3 FN 7

				TN 49
	Progesterone ≤20nmol/l & Inhibin A ≤30pg/ml & hCG ratio ≥0.8	NC	NC	TP 45 FP 1 FN 7 TN 50

Key: NS, not specified; NC, not calculable

Appendix 5: Development of the $M6_P$ and $M6_{NP}$ models

Reproduced from supplementary material published by Van Calster/ Bobdiwala et al[40].

NB Professor Ben Van Calster, statistician at KU Leuven, developed the M6 models. The details of this process are outlined below.

Coefficients for the M6 models with ($M6_P$ model) or without ($M6_{NP}$ model) the incorporation of initial progesterone as a predictor as obtained on all data used in this study (n=2753).

	Model with Progesterone ($M6_P$)		Model without Progesterone ($M6_{NP}$)	
	FPUL vs EP	IUP vs EP	FPUL vs EP	IUP vs EP
Intercept	3.3265	-5.0661	2.5506	-3.2842
Log(hCG0)	-0.3477	0.3813	-0.4242	0.4072
Log(Prog0)	-0.4501	0.5452		
Log(hCG ratio)	-5.6713	-5.2825	-2.9502	1.9238
Log(hCG ratio) square	1.0781	1.3498	2.1765	2.8952
Log(hCG ratio)*Log(Prog0)	1.0529	2.1392		

IUP = intrauterine pregnancy; hCG = human chorionic gonadotrophin; EP = ectopic pregnancy; FPUL = failed pregnancy of unknown location; Prog = Progesterone

The linear predictors z_1 and z_2 for $M6_P$ are:

$$z_1 = log\left(\frac{P(FPUL)}{P(EP)}\right)$$

$$= 3.3265$$

$$-0.3477 \times log(initial\ hCG)$$

$$-0.4501 \times log(initial\ progesterone)$$

$$-5.6713 \times log(hCG\ ratio)$$

$$+1.0781 \times [log(hCG\ ratio)]^2$$

$$+1.0529 \times log(hCG\ ratio) \times log(initial\ progesterone)$$

$$z_2 = log\left(\frac{P(IUP)}{P(EP)}\right)$$

$$= -5.0661$$

$$+0.3813 \times log(initial\ hCG)$$

$$+0.5452 \times log(initial\ progesterone)$$

$$-5.2825 \times log(hCG\ ratio)$$

$$+1.3498 \times [log(hCG\ ratio)]^2$$

$$+2.1392 \times log(hCG\ ratio) \times log(initial\ progesterone)$$

The probabilities of FPUL, IUP, and EP are then computed as:

$$P(FPUL) = \frac{exp(z_1)}{1 + exp(z_1) + exp(z_2)}$$

$$P(IUP) = \frac{exp(z_2)}{1 + exp(z_1) + exp(z_2)}$$

$$P(EP) = \frac{1}{1 + exp(z_1) + exp(z_2)}$$

The linear predictors z_1 and z_2 for $M6_{NP}$ are

$$z_1 = log\left(\frac{P(FPUL)}{P(EP)}\right)$$

$$= 2.5506$$

$$-0.4242 \times log(initial\ hCG)$$

$$-2.9502 \times log(hCG\ ratio)$$

$$+2.1765 \times [log(hCG\ ratio)]^2$$

$$z_2 = log\left(\frac{P(IUP)}{P(EP)}\right)$$

$$= -3.2842$$

$$+0.4072 \times log(initial\ hCG)$$

$$+1.9238 \times log(hCG\ ratio)$$

$$+2.8952 \times [log(hCG\ ratio)]^2$$

The probabilities of FPUL, IUP, and EP are then computed as:

$$P(FPUL) = \frac{exp(z_1)}{1 + exp(z_1) + exp(z_2)}$$

$$P(IUP) = \frac{exp(z_2)}{1 + exp(z_1) + exp(z_2)}$$

$$P(EP) = \frac{1}{1 + exp(z_1) + exp(z_2)}$$

hCG is measured in IU/L, progesterone in nmol/l.

Appendix 6: data on West Middlesex University Hospital (WMUH), a unit that used a progesterone cut-off of ≤10nmol/l and was analysed separately

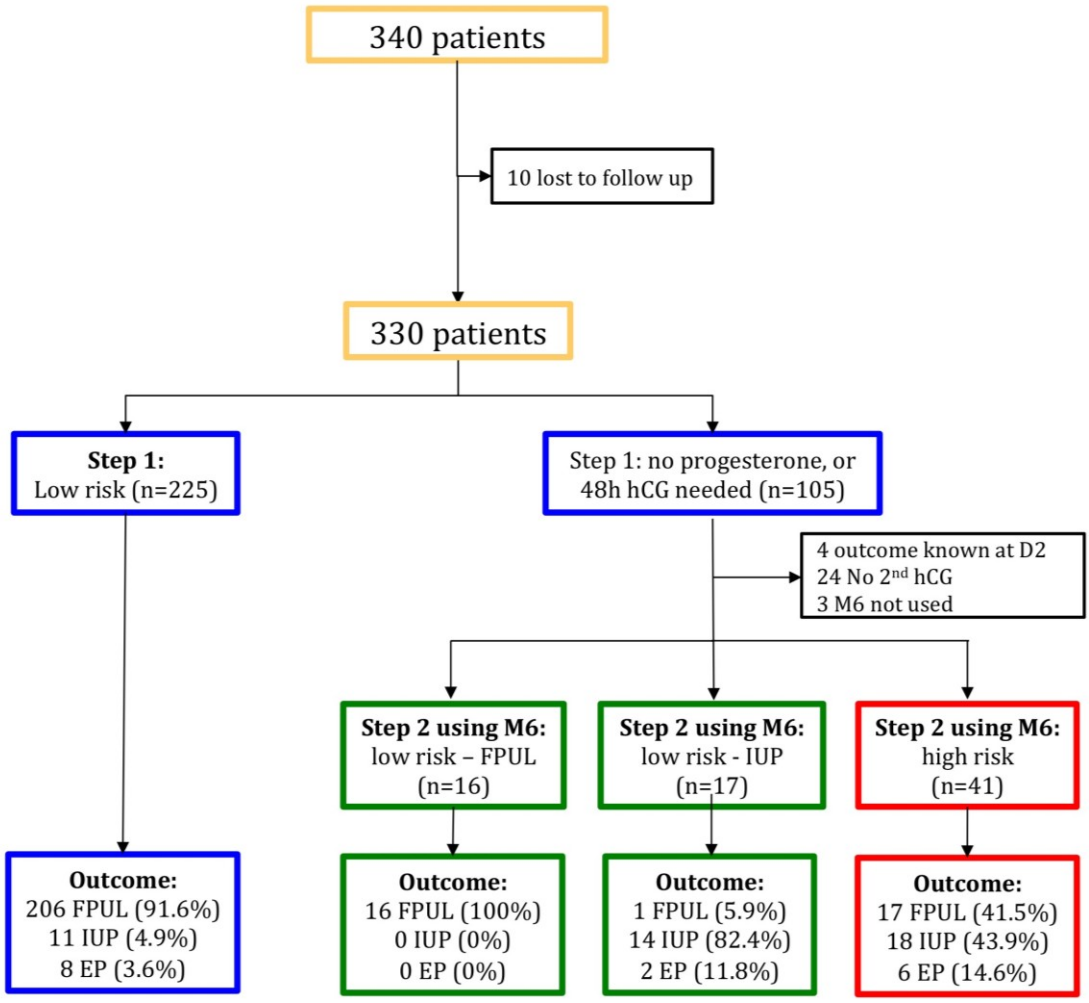


Figure A6: Flowchart for West Middlesex University Hospital patients. FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; EP = ectopic pregnancy; hCG = human chorionic gonadotrophin; prog = progesterone; M6NP = M6 model no progesterone

Final classification		Outcome			LFU
	N	FPUL N (%)	IUP N (%)	EP N (%)	N (%)
Low-risk					
at Step 1	225 (68%)	206 (82%)	11 (17%)	8 (50%)	1 (10%)
at Step 2	33 (10%)	17 (7%)	14 (22%)	2 (13%)	4 (40%)
Probable FPUL	16	16	0	0	0
Probable IUP	17	1	14	2	4
High risk	41 (12%)	17 (7%)	18 (28%)	6 (38%)	2 (20%)
Outcome known at day 2	4 (1%)	0 (0%)	4 (6%)	0 (0%)	0 (0%)
Not classified	27 (8%)	10 (4%)	17 (27%)	0 (0%)	3 (30%)
Total	**330**	**250**	**64**	**16**	**10**

Table A6i: Triage results for West Middlesex University Hospital. FPUL = failed pregnancy of unknown location; IUP = intrauterine pregnancy; EP = ectopic pregnancy; LFU = lost to follow-up

Adverse event	N
Model related	
EP classified as at low risk in step 1	8/16
EP classified as at low risk in step 2	2/16
EP not classified (no 2nd bloods, technical error)	0/16
Clinician related	
Negative laparoscopies	
Incorrectly given MTX	
Patient related	
Unscheduled admission with pain	
Unscheduled admission with heavy PVB	

Table A6ii: Adverse events at West Middlesex University Hospital. EP = ectopic pregnancy; MTX = methotrexate; PVB = per vaginal bleeding

Appendix 7: Vaginal microbiome in PUL – Standard Operating Procedure (SOP)

Vaginal swabs are taken using the Elution-swab system (Copan). Once a swab sample is collected, it should be placed immediately into the ESwab transport tube where it comes into contact with the transport medium. Swab specimens should be transported upright directly to the laboratory on ice, preferably within 1 hour of collection to be stored at -80°C if not processed immediately.

Thaw swabs on ice and vortex vigorously to re-suspend the cells (make sure the swab is in the solution)

Also thaw pre-aliquotted carrier RNA and lysis buffer enzymes. Defrost on ice

Set incubator to 37°C

Label (lids) all 2ml tubes

Use a filter to aliquot the PBS into a smaller tube

Aliquot 70% EtOH into a tube for cleaning of forceps

Remove swab from tube and place on bench. Cut and keep the end of the tube. Push the swab manually into the solution in the tube.

Carefully transfer sponge in tube into a 5ml syringe using a1000ul filter pipette tip and the swab solution can be squeezed by pushing the plunger into a 2ml centrifuge tube (you should get 500ul solution – you can add 100-200ul of PBS to the sponge to extract more fluid/ DNA if you are getting a low volume). Fill in the log and comment on swab (bloody appearance/ volume of additional PBS required (appendix 2))

Centrifuge at 10,000g for 10 mins (2000rpm for 5 mins for urine sample). There should always be an even number of tubes in the centrifuge directly opposite one another and all centrifuge tube hinges should face the outside.

Make the lysis buffer (these are quantities for 1 swab):
50ul lysozyme (10mg/ml)
6ul mutanolysin (25,000U/ml)
3ul lysostaphin (4000U/ml)
41ul TE (Trys 10mmol/ EDTA 50 mmol)
150ul PBS

So for 24 samples (with one additional sample to account for pipetting errors):
1250ul lysozyme
150ul mutanolysin
75ul lysostaphin
1025ul TE
3750ul PBS

Pipette off the swab solution and discard, leaving the 'pellet' (solid sediment) behind at the bottom of the centrifuge tube. Do not disturb the pellet

Re-suspend the pellet in 250ul lysis buffer.

Vortex samples to remix pellet

Clean floater with 70% EtOH. Put samples in the floater. Incubate at 37°C for 1 hour.

Switch off incubator once the samples have been removed.

Pulse spin (3 secs) the tube (Jencons PLS machine) to bring down the condensation

Add 100mg of bleached and rinsed 0.1mm diameter zirconia/silica beads (this doesn't need to be exact and is approx. 1x1 cm area on Yoonis' spatula. It further disrupts bacterial cells to release DNA). Glass beads have to be acid washed, autoclaved and then dried before use – this will usually have already been done.

Oscillate at 2000rpm for 2 mins using the Micro Dismembrator (Sartorius). This will take 4 samples at one time (Rufus machine = 12 samples at one time, settings = oscillator: 50; mins: 2)

Centrifuge briefly (pulse for 15-20 sec in the Jencons PLS)

Label (lids) new 2ml centrifuge tubes

Add 20ul proteinase K to each new 2ml tube (this is stored at room temperature & breaks down proteins so they don't interfere with the lysate column

Transfer lysate to a fresh 2ml tube (use p200 Gilson, set it to 190ul & tilt centrifuge tube as really don't want to get any beads in your lysate. OK to leave a bit of lysate behind to avoid getting beads). Each lysate samples should equal 200ul. Add PBS if it is less (volume is more important than concentration here)

QiAMP protocol from Qiagen page 65, J1:

Prepare the buffer VXL by adding 25ul carrier RNA to 2640ul buffer VXL (this should be enough for 25 samples)

Add 100ul Buffer VXL with carrier RNA to each sample

Mix by pulse vortexing for 10 secs and then pulse centrifuge (Jencons)

Incubate at room temp (20 – 25°C) for 15 mins

Add 350ul buffer ACB to the sample. Pulse vortex for 10 secs and briefly centrifuge (Jencons) to remove drops from inside the lid

Label (lids) QIAamp Mini spin column

Put all 670ul of lysate into the QIAamp Mini spin column (held in a 2ml collection tube) without wetting the rim.

Close the cap and centrifuge at 6000g (8000rpm) for 1 min (first spin).

Place new 2ml collection tubes (no lids) on rack.

Place each QIAamp Mini spin column into the new, clean 2ml collection tube. Discard the old 2ml tube containing the filtrate.

Carefully open the QIAamp Mini spin column and add 600ul Buffer AW1 without wetting the rim.

Close the cap and centrifuge at 6000g (8000rpm) for 1 min (second spin).

Place new 2ml collection tubes (no lids) on rack.

Place each QIAamp Mini spin column into the new, clean 2ml collection tube. Discard the old 2ml tube containing the filtrate.

Carefully open the QIAamp Mini spin column and add 600ul Buffer AW2 without wetting the rim.

Close the cap and centrifuge at 6000g (8000rpm) for 1 min (third spin).

Place new 2ml collection tubes (no lids) on rack.

Place each QIAamp Mini spin column into the new, clean 2ml collection tube. Discard the old 2ml tube containing the filtrate.

Close the cap and centrifuge at 20,000g (14,000rpm) for 2 mins (this is full speed, the max speed for the lab machine = 13,300rpm which is fine).

Label lids and tube of 1.5ml eppendorfs. Cut off lids with scissors and store on a clean petri dish (clean with 70% EtOH if needed). Place on rack.

Place each QIAamp Mini spin column into the eppendorph. Discard the old 2ml tube containing the filtrate.

Add 100ul Buffer AVE (elution buffer into the CENTRE of the column with a p200 Gilson).

Incubate at room temperature for 1 min.

Centrifuge at 20,000g (14,000rpm) for 1 min.

Discard COLUMN this time. Keep the fluid in the eppendorph (make sure you keep the chopped off pre-labelled lids to put back on)

Keep on ice from now onwards if doing PCR straightaway or otherwise store at -20C

The eluted DNA is then analysed by nanodrop and rtPCR is carried out using 16rRNA primer to confirm the success of bacterial DNA extraction.

Master mix for PCR (per sample):
10ul of 5 times polymerase buffer (vortex before use)
2ul forward primer, at a concentration of 10pmol/ul (vortex before use)
(or 0.2ul at a concentration of 100pmol/ul)
2ul reverse primer, at a concentration of 10pmol/ul (vortex before use)
(or 0.2ul at a concentration of 100pmol/ul)
1ul 10mmol DNTP (don't vortex but shake manually)
0.25ul TAC Polymerase

This needs to be made up to 45ul master mix/sample – so add 29.75ul sterile water
if using 2ul of primers OR 33.35ul sterile water if using 0.2ul primers
(NB you will later add 5ul template DNA (your extracted sample))

Mix master mix with pipette, do not vortex

If there are 10 samples in your batch, multiply all values by 11 (1 extra for errors)

e.g. for 27 samples (23 patient samples; 1 negative control; 1 positive control; 1 pure water, 1 extra for pipetting errors) the master mix for PCR will be:
270ul of 5 times polymerase buffer (vortex before use)
54ul forward primer, at a concentration of 10pmol/ul (vortex before use)
(or 5.4ul at a concentration of 100pmol/ul)
54ul reverse primer, at a concentration of 10pmol/ul (vortex before use)
(or 5.4ul at a concentration of 100pmol/ul)
27ul 10mmol DNTP (don't vortex but shake manually)
6.75ul TAC Polymerase
803.25ul sterile water (if using 2ul of primers) OR 900.45ul sterile water (if using 0.2ul of primers)

NB stock concentrations of the primers are always made up to 100pmol/ul. To get a concentration of 10pmol/ul, mix 50ul of stock primer to 450ul of sterile water (vortex to mix). These can be stored under UV light to denature any contaminant DNA.

Label small PCR tubes

Add 45ul master mix to each small PCR tube and then add 5ul template DNA (use p200 Gilson as there are no filter tips for the p100 Gilson in the lab)

Put these small PCR tubes into the PCR machine (avoid the holes on the edges as lid may pick up the tubes):

'programme'
'advanced mode'
'Yooni'
'bacteria', press 'OK'
'save & run'
choose the block the samples are in (usually 2A), press 'OK'
then press 'view' and 'status all' so no-one can accidentally stop your run!

This will take 2 ½ hours. If it is done overnight, the PCR machine will store the samples at 4°C until they are collected

PCR (day 2)

Remove samples from the PCR machine

Quick pulse mini-centrifuge to settle condensation

Place samples in 4°C fridge whilst preparing the gel

Set up the gel cast system with comb

Make 150ml gel: weigh out 3g agarose gel powder. Mix with 200ml 1 times TBE (trysbonic acid + EDTA) running buffer in a cone beaker. This makes a concentration of 1.5%. Agarose gel is not toxic like Acrilomide gel. The smaller the molecular weight of the product you are trying to detect, the higher the concentration of gel you will need.

NB the lab stores 10 times TBE, so may need to dilute: 900ml ionized water, 100ml TBE (use pipette gun/ measuring cylinder). Mix contents (use labcling film cover)

Bung the cone beaker with paper towel.

Heat solution in the microwave (2 mins on medium high). Heat until the agarose powder has completely dissolved – usually takes 3.5 mins. May need to stop/start to settle bubbles.

Cool this under a tap. It should not be poured hot onto the gel cast system

Add 16ul SybrSafe dye (x10,000 conc) to the 200ml Agarose solution (Sybr Safe is usually kept at room temp). Swirl beaker to mix dye.

Pour gel so the gel cast system is half full. Remove any large bubbles with a filter tip (small bubbles are OK). Discard excess gel before it sets in the beaker into the normal lab waste bin, not in the sink as it will block the sink

Get samples from the fridge – each of which should have 50ul of lysate

Add 8ul six times loading dye to each sample (this dilutes the loading dye to times 1). Shake the dye prior to adding it. Mix each sample with a pipette. The samples can be mini-centrifuged to mix in the dye. The loading dye 'weighs down' the DNA

Get the electrophoresis tank and wash it in ionized water

Check the agarose gel – it goes opaque when it is set. Take it out of the gel cast system and place into the electrophoresis tank

Pour over the rest of the 1 times TBE buffer to immerse the gel completely

Put comb on the negative side (black button). Slowly pull out the comb. Use your little and ring fingers to stabilize the gel.

Clean the comb immediately so the gel does not stick in the normal tap and leave to dry.

The Hyperladder dye is your control ('base marker'). Add 10ul to the far left side and 5ul to the far right side. All samples will go in between. Leave a gap between the Hyperladder and other samples and the NC (negative control) and other samples. 25ul of each extracted sample should be added.

Attach the tank lid. Set to 140 volts. Press the running man button to run the gel. If there is an 'error' in the machine, try changing the lid of the tank.

Store samples in the -20°C freezer. Have a sample box log that you log all samples into.