Statistical Analysis Plan

Management of pregnancies of unknown location – a randomised controlled trial of two hCG-based protocols.

2024-01-04

Author

Name/Title: Mattias Molin / Statistician	
Signature:	Date: 20240104

Approvals

Name/Title: Annika Strandell / MD, Associate professor, Sahlgrenska Academy at University of Gothenburg and Sahlgrenska University Hospital	
Signature:	Date: 2024-01-04
Name/Title: Johan Fistouris / MD, Principal Investigator, Sahlgrenska Academy at University of Gothenburg and Sahlgrenska University Hospital	
Signature:	Date: 240104

Table of Contents

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
SAE	Serious Adverse Event
PUL	Pregnancy of Unknown Location
TVU	Transvaginal Ultrasound
EP	Ectopic Pregnancy
IUP	Intrauterine Pregnancy
Failed PUL	Spontaneously resolving pregnancy of unknown location
PPUL	Persisting Pregnancy of Unknown Location
RCT	Randomised Controlled Trial
NICE	National Institute for Health and Care Excellence
M4	Prediction Model M4
hCG	Human Chorionic Gonadotrophin
PP	Per Protocol
HADS	Hospital anxiety and depression scale
SF-36	36-Item Short Form Health Survey
LMP	Last Menstrual Period
CI	Confidence Interval
Mtx	Methotrexate

1 STUDY DETAILS

1.1 Study Objectives

To compare the performance of the M4 prediction model (M4) with the hCG cut-offs recommended by NICE (1, 2), (hereafter referred to as the protocols), when used as part of a clinical management algorithm for women with a pregnancy of unknown location (PUL). The psychological impact on subjects managed as PUL will be studied with a self-reporting scale (HADS) and a questionnaire, which will be presented in a complementary SAP in due time.

1.1.1 Primary objective (Diagnostic performance of the protocols)

To determine the protocols' ability to correctly categorise an ectopic pregnancy (EP) as high risk (sensitivity) and categorise a non-ectopic pregnancy (failed PUL or intrauterine pregnancy (IUP)) as low risk (specificity).

1.1.2 Secondary objective

- To evaluate the safety and effectiveness of the protocols based on clinical outcomes (see clinical variables 3.9.2).
- To investigate if ectopic pregnancies correctly categorised as high risk and non-ectopic pregnancies correctly categorised as low risk benefit in regard of safety and effectiveness in comparison with those incorrectly categorised.

1.2 Study Design

This study is a multicentre randomised comparative diagnostic and clinical impact trial of two hCG-based protocols for PUL management. The study has a superiority design regarding the primary outcome of specificity for non-ectopic pregnancies for M4, given that the primary outcome of sensitivity for ectopic pregnancies is not inferior (non-inferiority design) in comparison with hCG cut-offs recommended by NICE. The diagnostic performance of the protocols will be evaluated on all subjects as a cross-sectional study for the primary analysis as well by randomisation groups in a sensitivity analysis.

1.3 Intervention

Randomisation will be with a 1:1 allocation to either M4 (Study group) or the hCG cut-offs recommended by NICE (Control group). Participants in both groups are categorised as either high or low risk of having an ectopic pregnancy by the allocated protocol which also provides a recommendation (directive decision approach) for the initial follow-up of PUL based on its classification.

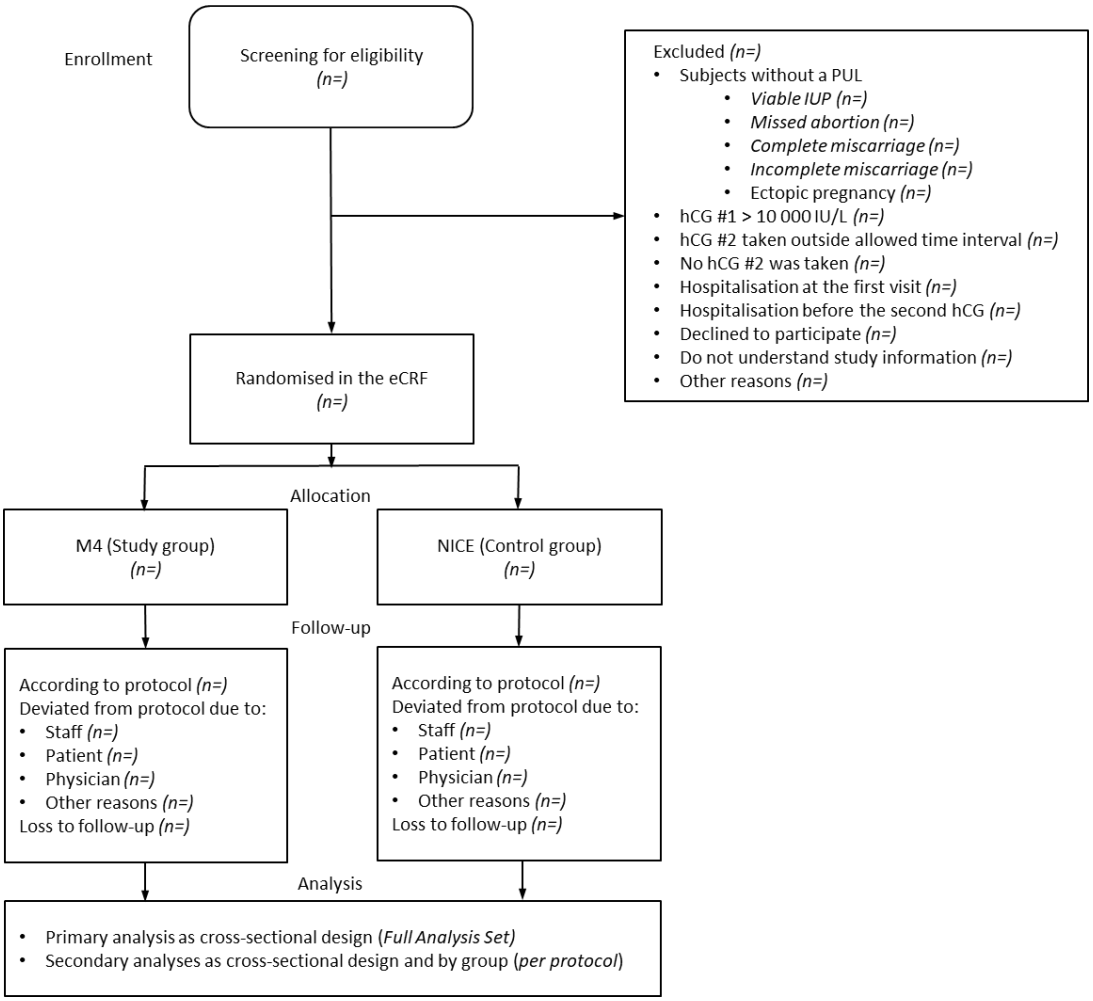
The clinician ultimately decides whether the recommended management is appropriate or not, considering other clinical variables. The following management algorithm of women in both groups is used:

1. A high-risk PUL is scheduled for a re-examination with transvaginal ultrasound (TVU) within 24 hours.

2. A low-risk PUL predicted to be a failed PUL is scheduled for a home urine pregnancy test after two weeks. If the test is positive a re-examination is scheduled within 24 hours.

3. A low-risk PUL predicted to be an IUP is scheduled for a TVU after one week.

A concealed randomisation sequence is generated using an algorithm that maximises the t-test value between Study group (M4) and Control group (NICE). Based on the first hCG the algorithm ensures an equal allocation ratio between groups.



1.4 Sample Size

1.4.1 Original sample size calculation

The sample size calculation was based on data from our previous study where M4 and the hCG cut-offs by NICE had a sensitivity of 89% and 86% for ectopic pregnancy respectively (p=0.42) (3). The specificity for non-ectopic pregnancy (IUP and failed PUL) was 78% and 70% respectively for M4 and the hCG cut-offs (p=0.01) (3).

The primary objective to test the superiority of M4 compared with NICE in specificity for non-ectopic pregnancies with 80% power. A reference specificity of 70% (NICE) is used. Considering eight percentage points to be a clinically important difference in specificity, given β=20%, α=5% and a two-sided test, 495 patients are needed per group to demonstrate superiority.

The primary objective will also test the non-inferiority of M4 to NICE in sensitivity for ectopic pregnancy. A reference sensitivity of 86% (NICE) is used and we would allow a one percentage point lower sensitivity for M4 and that the lower limit of the two-sided 95% CI for the difference between the two protocols should not exceed seven percentage points with a probability of 80% (β=20%), to be a non-inferior difference. Given a significance level (α) of 5% and a two-sided test, 541 patients per randomisation group are needed.

In our previous study five percent of patients were lost to follow-up, while another study reported a 7% loss to follow-up [46]. We have estimated the maximum rate of loss to follow-up to be 10%. For protection against that 10% of patients will be lost to follow-up, a total of approximately 1200 patients are needed.

1.4.2 Cross-sectional sample size calculation

No sample size calculation was performed a priori. A post-hoc power analysis will be performed.

1.4.3 Secondary outcomes

No sample size calculation was performed for secondary outcomes.

2 STUDY POPULATIONS

2.1 Definition of Study Populations

2.1.1 Full Analysis Set

All randomised subjects with a known outcome of PUL (IUP, failed PUL or EP) are included in the Full Analysis Set (FAS).

2.1.2 Per protocol population

All subjects included in FAS that were followed-up according to categorisation in high or low risk and that were managed according to the corresponding algorithm. Subjects deviating from the algorithm are not included in the per-protocol population. The per-protocol population is used for analyses of secondary outcomes.

3 STUDY VARIABLES

3.1 Group variables

Allocated protocol according to randomisation.

- M4 (Study group)
- NICE (Control group)

Each protocol categorises subjects in high-and low-risk of ectopic pregnancy:

- M4 use polynomial logistic regression to estimate the probability (%) of a PUL being a failed PUL, IUP and EP. M4 categorise a PUL as high-risk if the calculated probability of an EP ≥5%. If the calculated risk of an EP is <5% a PUL is categorised as low-risk and either a failed PUL or an IUP is predicted depending on which that has the highest calculated probability.

- For NICE: A PUL with a rise in hCG level of more than 63% or a decline of more than 50% it is categorised as low risk of being an EP, while a change in hCG levels between these cut-offs indicates high risk for EP.

3.2 Screening variables

Data collected for screened but not included subjects will include (not defined in study populations)

- Date of last screening made for a subject not included in the study (yy/mm/dd).
- Age at screening (years)
    - Reason for exclusion (Known pregnancy location (IUP)/Known pregnancy location (EP)/In-patients/Decline to participate in the study/Do not understand oral or written study information/Other reasons)
        - If IUP (with Yolk sac/Foetus with heartbeat/Foetus without heartbeat/Missed abortion/Complete miscarriage/ Incomplete miscarriage/Anembryonic pregnancy)
        - If EP (with Yolk sac/Foetus with heartbeat/Foetus without heartbeat)
            - EP location (fallopian tube/ovary/interstitium/caesarean scar/cervical)
            - EP size (mm)
- Screening site (SUS/SÄS/SkaS)

3.3 Baseline Variables

- Age (years)
- BMI (kg/m^2)
- Smoking (Never smoked/Former smoker/current smoker<10/11-20/21-30/>30)
- Number of prior pregnancies with:
    - miscarriage ≤ 12 weeks
    - miscarriage >12 weeks < 22 weeks
    - legal abortion
    - vaginal delivery
    - caesarean section
    - ectopic pregnancy
- History of infertility (Yes/No)
- Pregnancy conceived with IVF (Yes/No)
- Treated for chlamydia infection (Yes/No)
- Cervical dysplasia (Yes/No)

- Intrauterine device (Yes/No)
- Undergone gynaecological surgery (Yes/No)
    - Ovarian cyst (Yes/No)
    - Pelvic infection (Yes/No)
    - Endometriosis (Yes/No)
    - Ectopic pregnancy (Yes/No)
    - Excision of surgical dysplasia (Yes/No)
    - Uterine fibroids (Yes/No)
    - Polyp (Yes/No)
- Pregnancy weeks according to last menstrual period or embryo transfer date
- Ethnicity (Caucasian/African/Asian/Other)

3.4 First visit data variables

- Presenting symptoms
    - Bleeding (Yes/No)
    - Pain (Yes/No)
    - Pregnancy dating (Yes/No)
    - Previous early pregnancy complication (Yes/No)
    - IUD concerns (Yes/No)
- Self-registered lower abdominal pain (scale 1-10)
- Duration of pain in days
- Primary site of pain (right/left/general)
- Bleeding characteristics
    - Brown discharge (Yes/No)
    - Spotting (Yes/No)
    - As a menstruation (Yes/No)
    - More than a menstruation (Yes/No)
    - With clots (Yes/No)
- Physician seniority (resident/specialist/senior consultant)
- Years of transvaginal ultrasound experience (<1/1-3/>3)
- Type of pregnancy of unknown location at the first transvaginal ultrasound (No sign of IUP or EP/Probable IUP/Probable EP).
    - If Probable IUP
        - intrauterine sac like structure (Yes/No)
            - Size (mm)
        - retained products of conception (Yes/No)

        - Anembryonic pregnancy (Yes/No)
    - If Probable EP
        - inhomogeneous adnexal mass, 'blob' sign (Yes/No)
            - Size (mm)
        - extrauterine sac-like structure, 'doughnut' sign (Yes/No)
            - If Yes (fallopian tube/ovary/interstitium/caesarean scar/cervical)
            - Size (mm)

- Endometrial thickness (mm)
- Fluid in the pouch of Douglas (Yes/No)
- Amount of fluid (mm)
- Echogenic masses suggesting haemoperitoneum (Yes/No)

3.5 HCG related variables

- hCG1-value (IU/L)
- hCG2-value (IU/L)
- Mean of hCG1 and hCG2 (IU/L)
- Ratio of hCG1/hCG2
- Time between the first and second hCG measure (hours)

3.6 Randomisation and follow-up variables

- Randomisation (Yes/No)
    - If No (Hospitalisation before the second hCG/First hCG > 10 000 IU/L/Time between the first and second hCG measurement are less than 44 hours or more than 56 hours/The subject decline further participation or did not show up for a second hCG measurement)
- Categorisation and predicted outcome of PUL by allocated group: M4 or NICE (High risk, EP/Low risk, failed PUL/Low risk, IUP).
- Recommended follow-up by allocated group (Urine pregnancy test after two weeks/Ultrasound examination after one week/TVU examination within 24-48 hours)
- Deviation from recommended follow-up due to:
    - Follow-up visit was scheduled to early or late by the nurse (Yes/No)
    - The physician departed from plan due to concerns regarding other clinical variables (Yes/No)
    - The physician departed from plan due to concerns regarding ultrasound finding (Yes/No)
    - The patient made an unplanned visit due to increased pain or bleeding (Yes/No)

3.7 Treatment

- First line treatment (surgical/medical/expectancy)
- Second line treatment (surgical/medical/expectancy)
- Third line treatment (surgical/medical/expectancy)
    - If surgical (laparoscopic salpingectomy/laparoscopic salpingotomy/no finding/uterine evacuation)
        - Histopathological examination (Yes/No)
            - Confirmed pregnancy tissue (Yes/No)
        - Blood loss (<100/100-300/>300-500/>500-1000/>1000-1500/>1500 ml)
        - Blood transfusion (yes/no)
    - If medical
        - Methotrexate (single dose/repeat dose)
            - Change in hCG levels between days four and seven after injection (%)
        - Misoprostol (single dose/repeat dose)
- Complication
    - Rupture of ectopic pregnancy (Yes/No)
    - Serious side effect of methotrexate (Yes/No)
        - If Yes (Hepatotoxicity/Nephrotoxicity/Pulmonary toxicity/Infection/Myelosuppression)
    - Infection after surgery (Yes/No)
        - If Yes (Tubo-ovarian abscess/Salpingitis/Wound infection)
    - Reoperation due to persistent trophoblast (Yes/No)
    - Reoperation due to bleeding Yes/No)
- Time from randomisation until diagnosis, known outcome of PUL or start of first line treatment (days)
- Time from randomisation until last clinical visit, discharge after surgery or contact by telephone (days)
- Time from diagnosis until resolution of hCG (level < 5.0 IU/L) in ectopic pregnancies and persistent PUL managed expectantly, treated with methotrexate (days) or conservative surgery (salpingotomy)

3.8 Transvaginal ultrasound (TVU) and hCG

- First visit TVU findings (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
    - If Probable IUP
        - intrauterine sac like structure (Yes/No)
            - Size (mm)

        - retained products of conception (Yes/No)
    - If Probable EP
        - inhomogeneous adnexal mass, 'blob' sign (Yes/No)
            - Size (mm)
        - extrauterine sac-like structure, 'doughnut' sign (Yes/No)
            - If Yes (fallopian tube/ovary/interstitium/caesarean scar/cervical)
                - Size (mm)
    - If IUP (Yolk sac/Foetus with heartbeat/ Foetus without heartbeat/ Missed abortion/Complete miscarriage/Incomplete miscarriage/ Anembryonic pregnancy)
        - Size (mm)
    - If EP (with Yolk sac/Foetus with heartbeat/Foetus without heartbeat).
        - EP location (fallopian tube/ovary/interstitium/caesarean scar/cervical)
        - EP size (mm)

- Second visit TVU findings are described with the same variables as for the first visit TVU findings (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Third visit TVU findings are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Fourth visit TVU findings are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Fifth visit TVU findings are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Sixth visit TVU are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Seventh visit TVU are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- Eight visit TVU findings are described with the same variables as for the first visit TVU finding (Pregnancy of unknown location/probable IUP/probable EP/IUP/EP)
- First visit hCG level (IU/L)
- Second visit hCG level (IU/L)
- Third visit hCG level (IU/L)
- Fourth visit hCG level (IU/L)
- Fifth visit hCG level (IU/L)
- Sixth visit hCG level (IU/L)
- Seventh visit hCG level (IU/L)
- Eight visit hCG level (IU/L)
- Ninth visit hCG level (IU/L)

- Tenth visit hCG level (IU/L)
- Eleventh hCG level (IU/L)
- Twelfth visit hCG level (IU/L)
- Number of days between measurement of hCG two and three (needs to be calculated)
- Number of days between measurement of hCG two and three (needs to be calculated)
- Number of days between measurement of hCG three and four (needs to be calculated)
- Number of days between measurement of hCG four and five (needs to be calculated)
- Number of days between measurement of hCG five and six (needs to be calculated)
- Number of days between measurement of hCG six and seven (needs to be calculated)
- Number of days between measurement of hCG eight and nine (needs to be calculated)
- Number of days between measurement of hCG nine and ten (needs to be calculated)
- Number of days between measurement of hCG ten and eleven (needs to be calculated)
- Number of days between measurement of hCG eleven and twelve (needs to be calculated)
- Number of days between measurement of hCG twelve and thirteen (needs to be calculated)
- Number of days between measurement of hCG thirteen and fourteen (needs to be calculated)

3.9 Efficacy Variables

3.9.1 Primary Efficacy Variables

- Ectopic pregnancy (EP) (Yes/No)
    - If yes (Tubal Ampulla/Tubal Isthmus/Interstitium/Ovary/Caesarean scar/cervical)
- Intrauterine pregnancy (IUP) (Yes/No)
    - If yes (Viable IUP/Non-viable IUP /Non-viable IUP (Missed abortion/Anembryonic pregnancy)/Complete miscarriage/Incomplete miscarriage/cervical)
- Failed PUL (Yes/No)
- Persistent PUL, will be incorporated among EP in the analyses (Yes/No)
    - If yes (Treated/Resolved (untreated))
- Risk of ectopic pregnancy (High/Low)

3.9.2 Secondary Efficacy and Safety Variables (clinical outcomes)

- Number of TVU performed
- Number of hCG measurements
- Number of surgical procedures
- Time from randomisation until diagnosis, known final outcome of PUL or start of first line treatment (days)
- Time from randomisation until last clinical visit, hospital stay or contact by telephone (days)
- Time from diagnosis until resolution of hCG in EP and persistent PUL managed expectantly or treated with methotrexate (days) (calculated from the previous two variables by subtracting the second from the first)

Adverse events

- Accidental disruption of a possibly viable IUP (Yes/No)
- Ruptured EP (Yes/No)
- Emergency surgery (Yes/No)
- Adverse effect of Methotrexate (Yes/No)
- Negative laparoscopy (Yes/No)
- Unplanned visits initiated by the patient (Yes/No)
- Perioperative blood loss (ml)
- Blood transfusion (Yes/no)

4 STATISTICAL METHODOLOGY

4.1 General Methodology

Descriptive statistics will be presented by mean, standard deviation (SD), median, Q1 and Q3 (continuous variables) and number and percentage (categorical variables).

All tests will be two-tailed and conducted at the 0.05 significance level. All analyses will be performed by using SAS® v9.4 (Cary, NC).

4.2 Patient Disposition and Data Sets Analysed

The number of subjects screened, included in FAS and per-protocol population. Number of subjects lost to follow-up and reasons for exclusion from per-protocol population. Summary will be totally and by randomised groups and by risk category (seven columns).

4.3 Screening

Summary will be made totally and only for subjects who were not randomised.

4.4 Baseline

Summary will be totally and by randomised groups and by risk category (seven columns as described in general methods above. This will be made for both FAS and per-protocol population.

4.5 First visit data.

Summary will be totally and by randomised groups and by risk category (seven columns as described in general methods above. This will be made for both FAS and per-protocol population.

4.6 HCG relates.

Summary will be totally and by randomised groups and by risk category (seven columns as described in general methods above. This will be for made both FAS and per-protocol population.

4.7 Randomisation and follow-up

Summary will be totally and by randomised groups and by risk category (seven columns) as described in general methods above. This will be made for both FAS and per-protocol population.

4.8 Treatment

Data will be presented for first-, second- and third-line treatments. Summary will be totally and by randomised groups and by risk category (seven columns as described in general methods above. This will be made for both FAS and per-protocol population.

4.9 Efficacy Analyses

4.9.1 Primary Efficacy Analysis

Diagnostic performance will be described with sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), positive and negative likelihood ratio (LR+, LR-). The sensitivity represents ectopic pregnancies (EP) and persistent PUL correctly categorised as high risk of being an EP and the specificity represents non-ectopic pregnancies (IUP and failed PUL) correctly categorised as low risk of being an EP. PPV represents EP among PUL categorised as high risk and NPV represents IUP and failed PUL among PUL categorised as low risk.

The primary analysis will be performed for both M4 and NICE protocols on cross-sectional data for all subjects in the FAS population. To test differences in sensitivity and specificity between M4 and NICE the sign test will be used.

In addition to the cross-sectional analysis above a sensitivity analysis will be conducted to test differences in diagnostic performance measures by randomised groups for the FAS population. An exact 95 % CI will be calculated for the difference in sensitivity between the groups. The same will be produced for the specificity.

Complementary analyses will also be conducted for the per-protocol population.

Summary of all primary variables will also be presented totally and by randomised groups and by risk category (seven columns) as described in general methods above. This will be made for both FAS and per-protocol population.

4.9.2 Secondary Efficacy Analyses

Summary of all secondary variables will also be presented totally and by randomised groups and by risk category (seven columns) as described in general methods above. This will be made for both FAS and per-protocol population.

The main secondary analyses/questions will be evaluated:

1. Comparing the safety and effectiveness based on clinical outcomes between the allocated randomised groups.
2. Comparing the benefit of an ectopic pregnancy or a non-ectopic pregnancy being correctly categorised by a protocol versus those incorrectly categorised based on clinical outcomes.

A t-test will be used for comparison of continuous variables between independent groups. For comparison of proportions between groups a Fisher exact test is used. For comparison of number variables between groups a Mantel-Haenzsel chi-square test is used.

Kaplan-Meier curves are used to report the time from randomisation to the diagnosis and/or start of treatment of an ectopic pregnancy and the last visit in the clinic which includes any visit related to PUL management and for all outcomes of PUL. This will be reported according to allocated randomisation group and for the total cohort.

5 INTERIM ANALYSES

No interim analysis will be made.

6 CHANGES OF ANALYSIS FROM ORIGINAL STUDY PROTOCOL

The primary analysis will be performed for both M4 and NICE protocols on cross-sectional data instead of by randomised groups as stated in the original study protocol. As diagnostic performance is the primary outcome, we reasoned that a cross-sectional design would better utilise the reached sample size and maintain sufficient power, than the originally planned randomised design. Instead, the randomised groups are used in a sensitivity analysis of the diagnostic performance. This change from the original study protocol has been made before lock and analysis of the database. The randomised design is kept for the secondary clinical outcomes.

7 LISTING OF TABLE, FIGURES AND LISTINGS

7.1 Listing of Tables

Table Number	Table Title
14.1.1	Patient Disposition
14.1.2	Screening (non-randomised subjects)
14.1.3.1	Baseline (FAS Population)
14.1.3.2	Baseline (PP Population)
14.1.4.1	First visit data (FAS Population)
14.1.4.2	First visit data (PP Population)
14.1.5.1	hCG related (FAS Population)
14.1.5.2	hCG related (PP Population)
14.1.6.1	Randomisation and follow-up (FAS Population)
14.1.6.2	Randomisation and follow-up (PP Population)
14.1.7.1	Treatment (FAS Population)
14.1.7.2	Treatment (PP Population)
14.1.8.2	Second line treatment by first line treatment (FAS Population)
14.1.8.2	Second line treatment by first line treatment (PP Population)
14.2.1.1	Primary Efficacy Analysis diagnostic accuracy cross-sectional (FAS Population)

14.2.1.2	Primary Efficacy Analysis diagnostic accuracy cross-sectional (PP Population)
14.2.2.1	Primary Efficacy Analysis diagnostic accuracy by randomised groups (FAS Population)
14.2.2.2	Primary Efficacy Analysis diagnostic accuracy by randomised groups (PP Population)
14.2.3.1	Primary Efficacy Analysis descriptives (FAS Population)
14.2.3.2	Primary Efficacy Analysis descriptives (PP Population)
14.2.4.1	Secondary efficacy analysis – clinical outcomes (FAS population)
14.2.4.2	Secondary efficacy analysis – clinical outcomes (PP population)
14.2.5.1	Secondary efficacy analysis – clinical outcomes correctly categorized versus incorrectly categorized (FAS population)
14.2.5.2	Secondary efficacy analysis – clinical outcomes correctly categorized versus incorrectly categorized (PP population)

7.2 Listing of Figures

Decided as needed.

7.3 Listing of Listings

Decided as needed

8 REFERENCES

1. Condous G, Van Calster B, Kirk E, et al. Prediction of ectopic pregnancy in women with a pregnancy of unknown location. Ultrasound Obstet Gynecol 2007; 29:680–7. 10.1002/uog.4015 - DOI - PubMed
2. NICE guideline [NG126], 2019. Available: https://www.nice.org.uk/guidance/ng126 [Accessed 14 Dec 2023].
3. Fistouris J, Bergh C, Strandell A. Classification of pregnancies of unknown location according to four different hCG-based protocols. Hum Reprod 2016; 31:2203–11.doi:10.1093/humrep/dew202

9 APPENDIX