Received: 11 June 2021 | Revised: 13 September 2021 | Accepted: 7 November 2021

DOI: 10.1111/aogs.14295

ORIGINAL RESEARCH ARTICLE

Evaluating cut-off levels for progesterone, β human chorionic gonadotropin and β human chorionic gonadotropin ratio to exclude pregnancy viability in women with a pregnancy of unknown location: A prospective multicenter cohort study

Shabnam Bobdiwala[1] | Christopher Kyriacou[1] | Evangelia Christodoulou[2,3] | Jessica Farren[4] | Nicola Mitchell-Jones[5] | Maya Al-Memar[1] | Francis Ayim[6] | Baljinder Chohan[7] | Emma Kirk[8] | Osama Abughazza[9] | Bramara Guruwadahyarhalli[5] | Sharmistha Guha[5] | Veluppillai Vathanan[7] | Debbie Gould[4] | Catriona Stalder[1] | Dirk Timmerman[2,10] | Ben Van Calster[2] | Tom Bourne[1,2]

[1]Department of Obstetrics and Gynaecology, Tommy's National Centre for Miscarriage Research, Queen Charlotte's & Chelsea Hospital, Imperial College London, London, UK

[2]Department of Development & Regeneration, KU Leuven, Leuven, Belgium

[3]Department of Cancer Epidemiology, DKFZ, Heidelberg, Germany

[4]Department of Gynaecology, St Mary's Hospital, London, UK

[5]Department of Gynaecology, Chelsea and Westminster NHS Trust, London, UK

[6]Department of Gynaecology, Hillingdon Hospital NHS Trust, London, UK

[7]Department of Gynaecology, Wexham Park Hospital, London, UK

[8]Department of Gynaecology, Royal Free NHS Foundation Trust, London, UK

[9]Department of Gynaecology, Royal Surrey County Hospital, Guildford, UK

[10]Department of Gynecology, University Hospital Leuven, Leuven, Belgium

Correspondence
Tom Bourne, Tommy's National Centre for Miscarriage Research, Queen Charlotte's & Chelsea Hospital, Imperial College, London, UK.
Email: t.bourne@imperial.ac.uk

Funding information
SB is supported by the National Institute of Health Research Collaboration for Leadership in Applied Health Research & Care, NorthWest London (grant RDIP033). CK is supported by the Imperial Health Charity (grant RFPrD1920/116). EC, DT, and BVC are supported by Research

Abstract

Introduction: There is no global agreement on how to best determine pregnancy of unknown location viability and location using biomarkers. Measurements of progesterone and β human chorionic gonadotropin (βhCG) are still used in clinical practice to exclude the possibility of a viable intrauterine pregnancy (VIUP). We evaluate the predictive value of progesterone, βhCG, and βhCG ratio cut-off levels to exclude a VIUP in women with a pregnancy of unknown location.

Material and methods: This was a secondary analysis of prospective multicenter study data of consecutive women with a pregnancy of unknown location between

Abbreviations: βhCG, β human chorionic gonadotropin; CI, confidence interval; EP, ectopic pregnancy; FPUL, failed pregnancy of unknown location; IUP, intrauterine pregnancy; NVIUP, non-viable intrauterine pregnancy; PPUL, persisting pregnancy of unknown location; PUL, pregnancy of unknown location; TVS, transvaginal ultrasonography; VIUP, viable intrauterine pregnancy.

Shabnam Bobdiwala and Christopher Kyriacou are Joint first authors.

Foundation—Flanders grant G0B4716N and Internal Funds KU Leuven grant C24/15/037. DT is Fundamental Clinical Researcher of Research Foundation—Flanders. TB is supported by the NIHR Imperial Biomedical Research Center based at Imperial College Healthcare NHS Trust and Imperial College London, grant number IS-BRC-1215-20013.

January 2015 and 2017 collected from dedicated early pregnancy assessment units of eight hospitals. Single progesterone and serial βhCG measurements were taken. Women were followed up until final pregnancy outcome between 11 and 14 weeks of gestation was confirmed using transvaginal ultrasonography: (1) VIUP, (2) non-viable intrauterine pregnancy or failed pregnancy of unknown location, and (3) ectopic pregnancy or persisting pregnancy of unknown location. The predictive value of cut-off levels for ruling out VIUP were evaluated across a range of values likely to be encountered clinically for progesterone, βhCG, and βhCG ratio.

Results: Data from 2507 of 3272 (76.6%) women were suitable for analysis. All had data for βhCG levels, 2248 (89.7%) had progesterone levels, and 1809 (72.2%) had βhCG ratio. The likelihood of viability falls with the progesterone level. Although the median progesterone level associated with viability was 59 nmol/L, VIUP were identified with levels as low as 5 nmol/L. No single βhCG cut-off reliably ruled out the presence of viability with certainty, even when the level was more than 3000 IU/L, there were 39/358 (11%) women who had a VIUP. The probability of viability decreases with the βhCG ratio. Although the median βhCG ratio associated with viability was 2.26, VIUP were identified with ratios as low as 1.02. A progesterone level below 2 nmol/L and βhCG ratio below 0.87 were unlikely to be associated with viability but were not definitive when considering multiple imputation.

Conclusions: Cut-off levels for βhCG, βhCG ratio, and progesterone are not safe to be used clinically to exclude viability in early pregnancy. Although βhCG ratio and progesterone have slightly better performance in comparison, single βhCG used in this manner is highly unreliable.

KEYWORDS

early pregnancy complications, ectopic pregnancy, pregnancy, pregnancy of unknown location, reproductive endocrinology, ultrasound

1 | INTRODUCTION

Pregnancy of unknown location (PUL) is an early pregnancy classification defined as when a woman has a positive pregnancy test, but a pregnancy cannot be visualized on transvaginal ultrasonography (TVS). The final pregnancy outcome can be a viable intrauterine pregnancy (VIUP), non-viable intrauterine pregnancy (NVIUP), failed PUL (FPUL), persisting PUL (PPUL), or an ectopic pregnancy (EP).[1] The last of these can have life-threatening consequences, such as rupture causing intra-abdominal hemorrhage.[2]

Measurements of serum hormone levels of β human chorionic gonadotropin (βhCG) and progesterone are currently used clinically to indicate likely pregnancy viability and location in the PUL population.[3] Their use in the management of PUL and EP is well documented.[3–5]

There is no global agreement on how to best determine PUL viability and location using serum biomarkers. Despite this, βhCG and progesterone cut-off levels are commonly used in clinical practice to exclude the possibility of a VIUP before the use of methotrexate or uterine cavity instrumentation.[3–15]

Key message

βhCG, βhCG ratio and progesterone cut-off levels are not safe to be used clinically to exclude viability in early pregnancy. Although βhCG ratio and progesterone have slightly better performance in comparison, single βhCG used in this manner is highly unreliable.

Literature and guidance in several countries including Brazil, the USA and France state that intrauterine pregnancy (IUP) should be reliably visualized on TVS when βhCG discriminatory levels are either more than 1500 to 2500 IU/L, more than 3500 IU/L, or more than 3510 IU/L. The presumption is that an EP is likely to be present if an IUP cannot be visualized.[9,16–19] Pregnancies with progesterone levels below 5–10 nmol/L are also classified as non-viable according to some authors.[17,19]

Serial βhCG measurements are used to calculate a βhCG ratio (βhCG at 48 hours divided by βhCG at 0 hours). Various ratios,

including less than 0.85 and less than 1.5 have been used to define non-viability, with the presumption that a VIUP does not exist once the ratio is below a defined level.[1,3,5,8,16–20] Cut-off levels for βhCG ratio are then used to guide the need for medical or surgical intervention, particularly in the event of a possible EP.[21]

UK guidelines for the management of PUL state a βhCG ratio greater than 1.63 is likely associated with an IUP, a ratio less than 0.5 is likely associated with a failing pregnancy, and women with ratios in between are in need of prompt clinical review given the risk of ectopic pregnancy.[22] Guidelines in the USA define non-viability using βhCG ratios of of less than 1.33 to 1.53, depending on initial βhCG.[16,18] French guidance states that a βhCG ratio up to 0.85 when the initial βhCG is less than 2000 IU/L is not associated with a VIUP.[19] This level of heterogeneity between guidelines is concerning when they may be used to determine the viability of a wanted pregnancy. The result is that despite evidence to show they are unsafe and that mathematical models exist that perform significantly better, many units continue to use single biomarker cut-off values clinically.[1,5,23,24]

We aimed to evaluate the value of βhCG, progesterone, and βhCG ratio cut-off levels in excluding the possibility of a VIUP as an outcome in women classified with a PUL.

2 | MATERIAL AND METHODS

2.1 | Design and settings

This was a secondary analysis of a prospective multicenter study of consecutive women classified with a PUL at their initial Early Pregnancy Assessment Unit visit, carried out between January 2015 and January 2017 in four university teaching hospitals and four district general hospitals. The university teaching hospitals included Queen Charlotte's and Chelsea Hospital, St Mary's Hospital, Chelsea and Westminster Hospital, and West Middlesex University Hospital. The district general hospitals included Hillingdon Hospital, North Middlesex Hospital, Wexham Park Hospital, and Royal Surrey Hospital.

The primary aim of this study was to evaluate the use of a triage protocol for PUL routinely used in clinical practice to identify high-risk outcomes (PPUL and EP). The protocol has a two-step approach: first, women with low progesterone values (ie, ≤2 nmol/L) were considered at low risk (and therefore at high chance of an outcome of FPUL). The remaining women returned 48 hours later to obtain a βhCG ratio and apply the M6 risk prediction model (using initial progesterone, initial βhCG, and βhCG ratio as part of a logistic regression algorithm).[23–25] If the model estimated that the EP risk was 5% or more, women were classified as high risk.

We reported the study according to updated Standards for Reporting Diagnostic accuracy studies (STARD) guidance given the relevance of reporting VIUP diagnostic accuracy for given cut-off values. This enabled a structured approach to our work and analysis.[26]

2.2 | Core outcome sets and public involvement

A core outcome set was not used and on review of the core outcomes in the women's and newborns' health database, a relevant core outcome set does not yet exist. Women were not involved in the design of the study.

2.3 | Population

Women were included in the original study if they were classified with a PUL following their first visit to an Early Pregnancy Assessment Unit where a TVS was performed and were clinically well, hemodynamically stable, and suitable for outpatient management. Early Pregnancy Assessment Units are outpatient facilities and so most women are hemodynamically stable, even though they may present with a degree of bleeding and pain. Women were excluded if they did not initially meet the classification for a PUL (eg, pregnancies of uncertain viability; diagnosis of EP at the first scan) or were unsuitable for outpatient management. For the secondary analysis reported in this paper, we additionally excluded women with a presenting βhCG of 25 IU/L or less, the level below which a urine pregnancy test would be negative.

2.4 | Screening and index tests

At the initial Early Pregnancy Assessment Unit visit, following a questionnaire assessing subjective symptoms (mainly abdominal pain and bleeding), a TVS was performed by an appropriately trained healthcare professional. Women were classified as having a PUL according to definitions published in an earlier review.[2]

The index tests were: (1) serum levels of initial βhCG, (2) serum levels of initial progesterone, and (3) the βhCG ratio. These were measured using validated, automated laboratory immunoassays in each center by trained technicians who had no knowledge of the women.

The assay platforms used were variable, dependent on those chosen by each individual hospital. Although limited by acceptable bias and variation of data from group laboratory means, assay calibration is assured by following rigid internal and external quality control checks in order that results may be interpreted in a similar manner clinically. Each hospital is subscribed to an external quality assurance scheme, who define their own acceptance criteria for bias and variation, ratified by a national quality assurance advisory panel.[27,28] This forms part of the overall quality management system in UK laboratories, which includes internal quality control, audit, document control, staff training, and competency. Although there will be differences in standardization of methods (calibration) and in antibody pairs used by different immunoassay manufacturers in each laboratory, external quality assurance maintains a high standard for obtaining reproducible results, with accreditation services ensuring individual laboratory compliance against internationally recognized quality standards.[29,30]

We defined categories for analysis in order to capture and evaluate commonly used cut-off levels defining non-viability and assessed the univariate predicted probability of each outcome using data for each index test in a continuous manner. For initial βhCG we subdivided 0–5000 IU/L into 250-IU/L groupings. For initial progesterone we subdivided 0–20 nmol/L into 1-nmol/L categories. We assessed βhCG ratios of 0–4 in 0.2 increments. Test performance can therefore be easily derived for any chosen level of progesterone, βhCG, or βhCG ratio. By plotting the cut-off levels in graphical form, we did not limit the project to only commonly used values, allowing probability at any level to be read with ease.[1,3,5–11,13–15]

2.5 | Main outcome measure

The primary outcome measure was the predictive value of cut-off levels in excluding a viable pregnancy.

2.6 | Reference standard

Final outcomes were categorized into one of three groups: (1) VIUP (where an embryo with visible cardiac activity was seen at initial follow up and is still present at the time of the dating scan at 11–14 weeks of gestation), (2) NVIUP (where an IUP seen on TVS had miscarried by the time of the dating scan) or FPUL (where βhCG levels reduced and resolved spontaneously without the visualization of a pregnancy on TVS), and (3) EP (an extrauterine mass seen on TVS) or PPUL (where TVS did not reveal the pregnancy location when more than three βhCG levels taken over 48-hour intervals remained static with a difference of 15% or less each time).[1,31,32] Further TVS and serum βhCG levels were the only other investigations performed between the index tests and reference standard. No interventions influenced the outcome.

2.7 | Statistical analyses

No set sample size was required for this descriptive, secondary analysis. The availability of 2507 women (76.6% of cohort following exclusions) with known outcome represents the largest PUL sample size to date focusing on the performance of cut-off levels, so was considered sufficient.[1,3–12,14,15,20]

We defined the percentage of each outcome within each biomarker category, together with multinomial 95% confidence intervals (CI).[33] We also predicted the final pregnancy outcome using the continuous values for each biomarker, because categorization can be associated with information loss. This was performed via univariable multinomial logistic regression, using restricted cubic splines to model the relation of each biomarker with the outcome. We used five knots with default knot locations for the splines.[34]

The primary analysis was pre-specified as a complete case analysis. For this analysis, we excluded women lost to follow up and women with an IUP where final viability was not recorded. For each biomarker, we excluded women who did not have the required data. For the progesterone analysis, women taking progesterone supplements were also excluded. Those who did not have a 48-hour βhCG sample taken on the second day following the initial βhCG measurement, or whose final pregnancy outcome was known on the second day because TVS was carried out for clinical considerations, were excluded from the βhCG ratio analysis.

We performed a sensitivity analysis based on multiple imputation of biomarker values.[35] In the imputation procedure, we included cases that had been lost to follow up. For the analysis of the imputed data, however, these were excluded. For βhCG ratio, we excluded cases where the final pregnancy outcome was known on the second day. We imputed missing values 100 times, leading to 100 completed data sets. These data sets were analyzed separately before their results were combined. Details on the imputation procedure are given in the Supporting Information (Appendix S1).

Analyses were performed using R version 3.6.1 (www.r-project.org).

2.8 | Ethical approval

This project makes up part of a study approved by the Health Research Authority and Health and Care Research Wales Research Ethics Committee, reference 21/HRA/0260, on January 26, 2021. As these data are collected routinely as part of normal clinical practice and were analyzed in an anonymous fashion, written and verbal consent was not required.

3 | RESULTS

Figure 1 highlights the flow of women through the study. A total of 3272 women were classified as PUL. Six (0.2%) met the exclusion criteria and 367 (11.2%) had an initial βhCG of 25 IU/L or less. Of the remaining 2899 women, 297 (10.2%) were lost to follow up and the final pregnancy outcome was not known (see Supporting Information, Table S1). Of the remaining 2602 women, 95 (3.7%) had an IUP with unknown final viability. This left 2507 women (76.6% of entire cohort following exclusions) for complete case analysis.

All 2507 had an initial βhCG measurement; 2248/2507 (90%) had an initial progesterone level measurement and were not taking progesterone supplements; and 1809/2507 (72.2%) were still classified with a PUL 2 days later and had a second βhCG measurement to calculate the βhCG ratio.

Sensitivity analysis based on multiple imputations for initial βhCG and progesterone included 2602 women (79.5% of entire cohort following exclusions), using the 95 with an IUP of uncertain final viability. In all, 2536 women were included in the βhCG ratio multiple

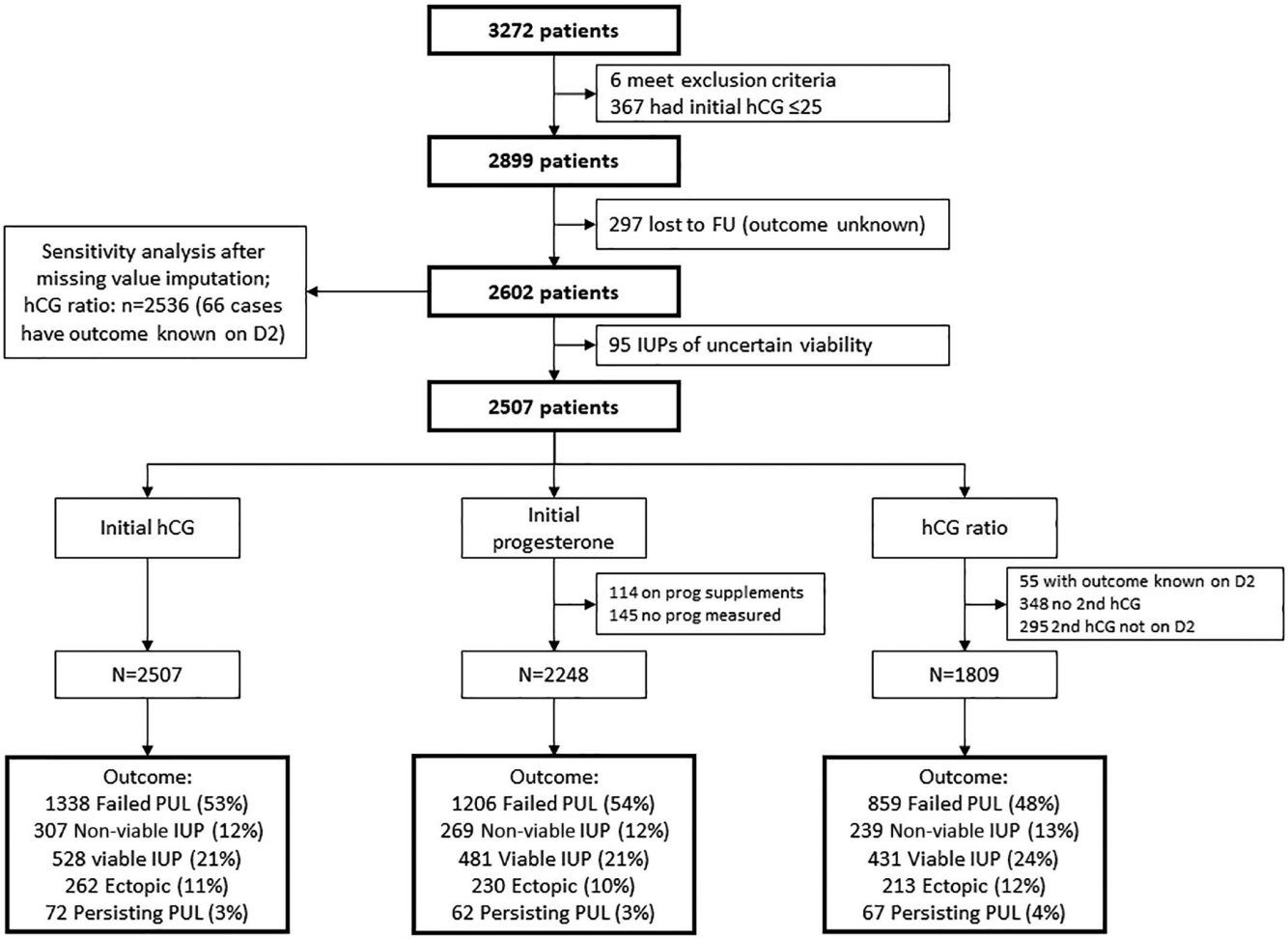


FIGURE 1 Flowchart of recruitment. FPUL, failed pregnancy of unknown location; NVIUP, non-viable intrauterine pregnancy; VIUP, viable intrauterine pregnancy; EP, ectopic pregnancy; hCG, human chorionic gonadotropin ($N = 3272$)

imputation, with 66/2602 (2.5%) excluded because the final pregnancy outcome was known on day two.

3.1 | Complete case analysis

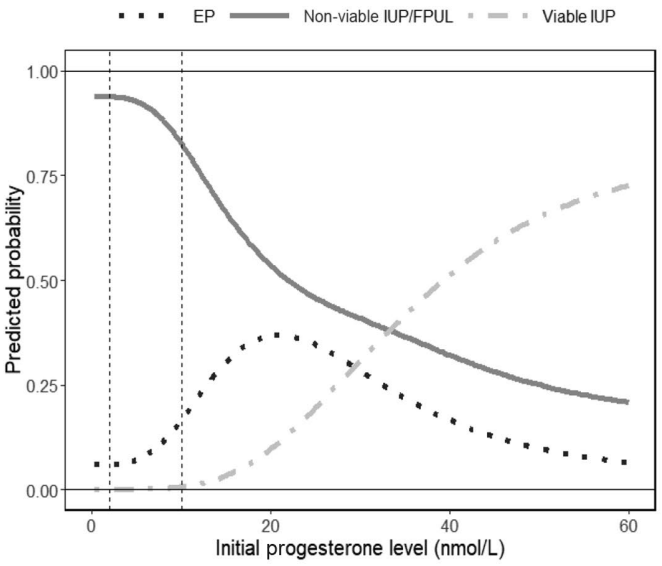
3.1.1 | Single measurements of progesterone

When assessing the predicted probability of each outcome based on continuous progesterone levels, the likelihood of a viable pregnancy increased with the progesterone level. This can be read directly from both Figure 2A and Table 1A. When assessing progesterone levels in 1-nmol/L increments from 0 to 20 nmol/L, predicted VIUP probability increased from 0.001 to 0.097. Cut-offs of 2 and 10 nmol/L are discussed in more detail in Table 2A. Among 327 PUL with an initial progesterone ≤2 nmol/L, none were VIUP (0%, 95% CI 0–1.2). With a progesterone level below 10 nmol/L, the predicted probability of viability was low (0.007) but could not be excluded with certainty. In this data set, 2/1112 (0.2%, 95% CI <0.01 to 1.8) PUL with an initial progesterone of 10 nmol/L or less were VIUP, whereas 1023/1112 (92.0%, 95% CI 90.6–93.6) had an NVIUP or FPUL. Although the median progesterone level associated with viability was 59 nmol/L, VIUP were identified with initial progesterone levels as low as 5 nmol/L (Table 2B).

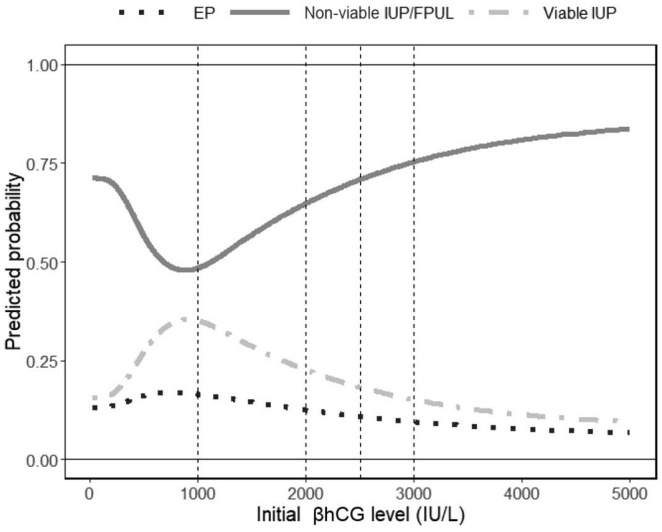
3.1.2 | Single measurements of serum βhCG

No single cut-off reliably ruled out the presence of a VIUP when assessing the predicted probability of each outcome based on continuous βhCG levels. This can be read directly from Figure 2B and Table 1B. Although the most likely outcomes were NVIUP or FPUL when assessing all possible cut-off values of βhCG from 0 to 5000 IU/L in 250-IU/L increments, a VIUP remained a possibility at each level (0.095–0.351). Very high values of initial βhCG were associated with the lowest estimated probability for a VIUP (0.095 when βhCG 5000 IU/L). Commonly used cut-off values between 1000 and 3000 IU/L are discussed in more detail in Table 2A. With each cut-off, viability cannot be excluded (viability ranging from 10.9% to 18.9%). With an initial βhCG greater than 3000 IU/L, 39/358 (10.9%, 95% CI 7.3–14.9) of women with a PUL had an outcome of a VIUP. Although the VIUP median single βhCG from this data set

(A) Continuous progesterone levels (nmol/L) (N=2248);



(B) Continuous βhCG levels (IU/L) (N=2507);



(C) Continuous βhCG ratio levels (N=1809).

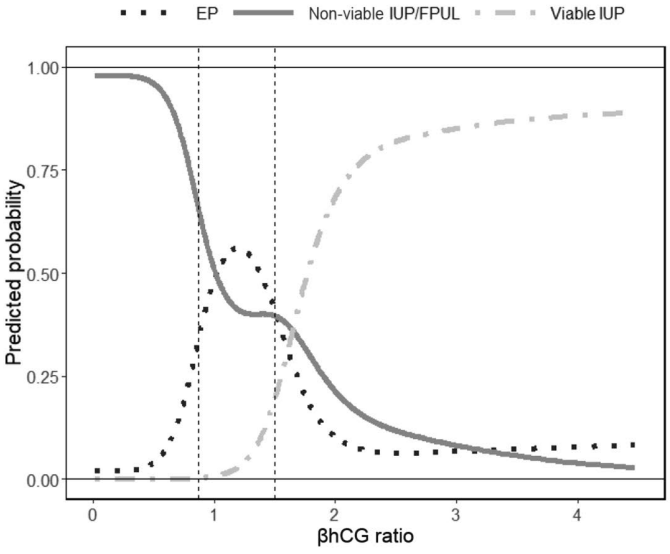


FIGURE 2 Predicting pregnancy of unknown location outcome using univariate predicted probability of each outcome. Combining all outcomes at each biomarker cut-off level makes up a total of 1, with levels for each outcome ranging between 0 and 1. These are based on: (A) continuous progesterone levels (nmol/L) (*N* = 2248); (B) continuous β human chorionic gonadotropin (βhCG) levels (IU/L) (*N* = 2507); (C) continuous βhCG ratio levels (*N* = 1809)

TABLE 1 Predicting pregnancy of unknown location outcome using estimated probability of each outcome. Combining all outcomes in each row makes up a total of 1, with levels for each outcome ranging between 0 and 1. These are based on selected values of: (A) progesterone levels (nmol/L) (*N* = 2248); (B) βhCG levels (IU/L) (*N* = 2507); (C) βhCG ratio levels (*N* = 1809)

(A)

Progesterone cut-offs	EP	Non-viable IUP/FPUL	Viable IUP
0	0.060	0.939	0.001
1	0.060	0.939	0.001
2	0.060	0.939	0.001
3	0.061	0.938	0.001
4	0.065	0.934	0.001
5	0.071	0.928	0.001
6	0.082	0.917	0.002
7	0.096	0.902	0.002
8	0.115	0.882	0.003
9	0.139	0.857	0.005
10	0.166	0.827	0.007
11	0.196	0.794	0.010
12	0.226	0.761	0.014
13	0.254	0.727	0.019
14	0.281	0.694	0.025
15	0.304	0.662	0.033
16	0.325	0.632	0.043
17	0.341	0.605	0.054
18	0.354	0.579	0.067
19	0.363	0.556	0.081
20	0.368	0.535	0.097

(B)

βhCG cut-offs	EP	Non-viable IUP/FPUL	Viable IUP
0	0.132	0.712	0.156
250	0.138	0.687	0.175
500	0.162	0.565	0.273
750	0.169	0.488	0.343
1000	0.165	0.485	0.351
1250	0.156	0.523	0.322
1500	0.145	0.569	0.286
1750	0.135	0.611	0.254
2000	0.125	0.648	0.227
2250	0.116	0.681	0.203
2500	0.108	0.709	0.183
2750	0.101	0.733	0.166
3000	0.095	0.754	0.151
3250	0.090	0.771	0.139
3500	0.085	0.786	0.129
3750	0.081	0.799	0.120
4000	0.077	0.720	0.113
4250	0.074	0.727	0.107
4500	0.072	0.826	0.103
4750	0.069	0.832	0.099
5000	0.068	0.837	0.095

(C)

βhCG ratio cut-offs	EP	Non-viable IUP/FPUL	Viable IUP
0.0	0.020	0.979	<0.001
0.2	0.020	0.979	<0.001
0.4	0.026	0.974	<0.001
0.6	0.073	0.926	<0.001
0.8	0.251	0.748	0.001
1.0	0.485	0.508	0.007
1.2	0.560	0.410	0.03
1.4	0.487	0.401	0.112
1.6	0.318	0.376	0.305
1.8	0.172	0.294	0.533
2.0	0.102	0.212	0.686
2.2	0.073	0.159	0.767
2.4	0.064	0.128	0.808
2.6	0.063	0.109	0.829
2.8	0.065	0.094	0.841
3.0	0.067	0.082	0.851
3.2	0.069	0.071	0.86
3.4	0.072	0.061	0.868
3.6	0.074	0.052	0.874
3.8	0.076	0.045	0.879
4.0	0.078	0.039	0.883

Abbreviations: βhCG, β human chorionic gonadotropin; EP, ectopic pregnancy; FPUL, failed pregnancy of unknown location; IUP, intrauterine pregnancy.

is 597 IU/L, VIUP can present with an initial βhCG levels as high as 105 006 IU/L (Table 2B).

3.1.3 | The performance of the βhCG ratio

The predicted probability of a viable pregnancy increases with the βhCG ratio. This can be read directly from both Figure 2C and Table 1C. As the βhCG ratio increases in 0.2 decrements, so does the probability of viability. The estimated probability of a VIUP is below 0.001 when the βhCG ratio is 0 to 0.8, compared with 0.883

TABLE 2 Summary table of pregnancy of unknown location outcome: (A) for commonly used bhCG, progesterone, and bhCG ratio cut-offs of non-viability. Number (*N*), percentage (%) and confidence intervals (CI) (*N* = 327 to *N* = 1205); (B) Using median values alongside range for initial progesterone, initial bhCG, Day 2 bhCG, and bhCG ratio (if second bhCG at Day 2) by PUL outcome (*N* = 2507)

(A)

List of cut-offs	*N*	EP/PPUL *N* (%, 95% CI)	Non-viable IUP/FPUL *N* (%, 95% CI)	Viable IUP *N* (%, 95% CI)
Initial βhCG >1000 (IU/L)	832	87 (10.5%, 7.5–13.6)	588 (70.6%, 67.6–73.8)	157 (18.9%, 15.9–22.1)
Initial βhCG >2000 (IU/L)	513	43 (8.4%, 5.1–12.0)	399 (77.8%, 74.5–81.4)	71 (13.8%, 10.5–17.5)
Initial βhCG >2500 (IU/L)	428	34 (7.9%, 4.4–11.6)	345 (80.6%, 77.1–84.2)	49 (11.4%, 7.9–15.1)
Initial βhCG >3000 (IU/L)	358	27 (7.5%, 3.9–11.5)	292 (81.6%, 77.9–85.5)	39 (10.9%, 7.3–14.9)
Initial progesterone ≤2 (nmol/L)	327	6 (1.8%, 0.61–3.1)	321 (98.2%, 96.9–99.4)	0 (0.0%, 0.0–1.2)
Initial progesterone ≤10 (nmol/L)	1112	87 (7.8%, 6.4–9.4)	1023 (92.0%, 90.6–93.6)	2 (0.18%, <0.01–1.8)
βhCG ratio <0.87	883	55 (6.2%, 4.8–7.8)	828 (93.8%, 92.3–95.3)	0 (0.0%, 0.0–1.5)
βhCG ratio <1.5	1205	204 (16.9%, 14.9–19.1)	985 (81.7%, 79.7–84.0)	16 (1.3%, <0.01–3.5)

(B)

	EP/PPUL Median (range)	Non-viable IUP/FPUL Median (range)	Viable IUP Median (range)
Initial progesterone (nmol/L)[a]	16.00 (1.00–92.00)	5.00 (0.30–153.00)	59.00 (5.00–219.00)
Missing (%)	17 (0.7%)	102 (4.3%)	26 (1.1%)
Initial serum βhCG (IU/L)	480.00 (31.00–42 520.00)	477.00 (25.70–60 542.00)	597.00 (26.00–105 006.00)
Missing (%)	0	0	0
48-h serum βhCG (IU/L) if at Day 2[b]	507.50 (21.00–44 103.00)	279.50 (3.00–109 568.00)	1280.00 (61.00–30 073.00)
Missing (%)	53 (2.1%)	547 (21.8%)	97 (3.9%)
βhCG ratio if second βhCG at D2[b]	1.19 (0.36–3.92)	0.41 (0.01–4.25)	2.26 (1.02–6.20)
Missing (%)	53 (2.1%)	547 (21.8%)	97 (3.9%)

Abbreviations: bhCG, b human chorionic gonadotropin; EP, ectopic pregnancy; FPUL, failed pregnancy of unknown location; IUP, intrauterine pregnancy; PPUL, persisting pregnancy of unknown location.

[A]This excludes known progesterone levels from cases who were on progesterone supplements.

[B]Corresponds to secondary bhCG measurements taken exactly 2 days after the first (as indicated).

when the βhCG ratio is 4.0. Commonly used cut-off levels of 0.87 and 1.5 are discussed in more detail in Table 2A. In this data set, with a βhCG ratio below 0.87, 0/883 (0%, 95% CI 0%–1.5%) were VIUP, whereas with a βhCG ratio of less than 1.5, 16/1205 (1.3%, 95% CI <0.01%–3.5%) were VIUP. The median βhCG ratio for VIUP was 2.26. However, the lowest βhCG ratio associated with a VIUP in this data set was 1.02 (Table 2B).

3.2 | Sensitivity analysis using multiple imputation

When using multiple imputation of 100 data sets, the predicted probability of a VIUP was slightly higher compared with the complete case analysis results at each biomarker cut-off value, with the probability of a VIUP no longer zero when progesterone was less than 2 nmol/L (0.14%, 95% CI 0.01%–1.4%) or βhCG ratio was less than 0.87 (0.13%, 95% CI 0.02%–1.1%) (see Supporting Information, Tables S2 and S3). However, the trend of viability probability remains the same as with complete case analysis, with very high levels of βhCG, low progesterone and low βhCG ratios associated with a lower likelihood of a VIUP (see Supplementary Information, Figure S1).

4 | DISCUSSION

This study highlights that cut-off values that are still commonly used in clinical practice to define probable non-viability in early pregnancy are not safe. Single βhCG levels are highly unreliable, with cut-off levels for the βhCG ratio and serum progesterone having slightly better performance in comparison.

The main strength of this study was its large multicenter population. One limitation is that 392 women were either lost to follow up or had an IUP of unknown final viability, whereas others did not have a βhCG ratio value because a second βhCG measurement was not taken 2 days later. However, those without a second βhCG reading, as well as those with an IUP of unknown viability, were included in a sensitivity analysis following multiple imputation of missing values. Although this method assumes that missing values are missing at random, conditional on other information in the database or that "any systematic difference between the missing values and the observed values can be explained by differences in observed data", this assumption was considered plausible.[35]

Our findings categorically show that the use of single "discriminatory zone" measurements for serum βhCG have poor diagnostic performance and should not be used in clinical practice to exclude the presence of an intrauterine pregnancy. This is important as some guidelines still state that women classified with a PUL who are found to have a single βhCG measurement at presentation of more than 3000, 2000 or even 1000 IU/L are unlikely to have an intrauterine pregnancy and a presumption may be made that the pregnancy is in the fallopian tube.[16–19] Not surprisingly, this has led to the use of the discriminatory zone remaining part of local guidelines and is still being used by some clinicians.[3,9,11] The risk of this approach is the inadvertent administration of methotrexate to a wanted intrauterine pregnancy.

In our data set, a proportion of women had high initial levels of βhCG. Many of the pregnancies were technically difficult to visualize, leading to a PUL classification. This was a result of the presence of multiple fibroids, diffuse adenomyosis, molar pregnancy, early multiple pregnancy, non-tubal ectopic pregnancy, and likely miscarriages that met PUL criteria on the first scan. As each participating center carries TVS training responsibilities, a small proportion of supervised TVS operators may have been unable to confirm pregnancy location with confidence on the initial scan.

Further, the βhCG acceptance criteria for bias and variation are reported locally as approximately 20%.[27,28] This means that a mean serum βhCG of 1000 IU/L measured in one unit could be reported as 800–1200 IU/L elsewhere. Accordingly using "standardized" discriminatory zones that are not derived from laboratory βhCG values specific to an individual unit is dangerous.

A βhCG ratio below 0.87 or an initial progesterone measurement of 2 nmol/L or less was very unlikely to be associated with viability but was not definitive when taking into account multiple imputations. Although the number of viable pregnancy misclassifications would be low if these cut-off levels were used, it is important to note that these findings are unique to this data set, under the constraints of our population, definition of PUL, and methods of laboratory biomarker processing. As such, these are not generalizable.

In this large data set, PUL with a single βhCG as high as 105 006 IU/L, a single progesterone as low as 5 nmol/L, and a βhCG ratio as low as 1.02 have been associated with viable pregnancies at 11–14 weeks of gestation. Although cut-off levels have been identified that are associated with non-viability in our data set, again these only reflect our specific population and cannot be generalized.[3–19] Indeed, upon literature review, cases of VIUP have been reported with lower progesterone levels (3 nmol/L) and falling serial βhCG values (βhCG ratio <1) that begin as high as 167 343 IU/L, further supporting the argument that cut-off levels in the differentiation of viability in PUL are subjective and unreliable.[15,20,36,37]

Systematic reviews and meta-analyses, as well as previous work performed by our group in differentiating pregnancy location, have highlighted how combining variables in prediction models outperform any variable in isolation.[3,5,24,25] In line with this, the use of isolated cut-off levels to define viability gives an inaccurate impression of diagnostic certainty that cannot be generalized. Caution must therefore be shown when using any cut-off, with awareness of their limited ability to effectively predict final pregnancy outcome in women with a PUL.

5 | CONCLUSION

We have used this large data set of women with a PUL to describe the limitations of defining viability in early pregnancy using cut-off values. Single βhCG cut-off values, which are still commonly used in clinical practice to define non-viability in early pregnancy, are highly unreliable and unsafe. Progesterone and βhCG ratio cut-off levels have slightly better performance in comparison. Great care must be taken to exclude the possibility of a viable pregnancy when contemplating either methotrexate therapy or instrumentation of the uterine cavity. Whereas measurements of both single and serial levels of serum hormone levels in early pregnancy can offer guidance, the entire clinical picture must be considered before intervention. If women are stable and being managed as outpatients, a conservative approach is unlikely to be associated with harm.

CONFLICT OF INTEREST

None.

AUTHOR CONTRIBUTIONS

TB, BVC, SB, and DT participated in the conception and design of the study. SB, JF, NMJ, CK, MAM, FA, BC, EK, OA, BG, SG, VV, CB, DG, and CS acquired data. BVC and EC performed the statistical analysis. SB, CK, BVC, EC, and TB interpreted the results and wrote the initial version of the manuscript. All authors critically revised the manuscript and approved the final version.

ORCID

Shabnam Bobdiwala https://orcid.org/0000-0003-0540-2191
Christopher Kyriacou https://orcid.org/0000-0001-9001-5545
Evangelia Christodoulou https://orcid.org/0000-0001-7900-5952
Jessica Farren https://orcid.org/0000-0002-8082-6499
Nicola Mitchell-Jones https://orcid.org/0000-0003-0299-9586
Dirk Timmerman https://orcid.org/0000-0002-3707-6645
Ben Van Calster https://orcid.org/0000-0003-1613-7450
Tom Bourne https://orcid.org/0000-0003-1421-6059

REFERENCES

1. Bobdiwala S, Al-Memar M, Farren J, Bourne T. Factors to consider in pregnancy of unknown location. *Womens Health*. 2017;13:27-33.
2. Kirk E, Bottomley C, Bourne T. Diagnosing ectopic pregnancy and current concepts in the management of pregnancy of unknown location. *Hum Reprod Update*. 2014;20:250-261.
3. Bobdiwala S, Saso S, Verbakel JY, et al. Diagnostic protocols for the management of pregnancy of unknown location: a systematic review and meta-analysis. *BJOG*. 2019;126:190-198.
4. Condous G, Lu C, Van Huffel SV, Timmerman D, Bourne T. Human chorionic gonadotrophin and progesterone levels in pregnancies of unknown location. *Int J Gynaecol Obstet*. 2004;86:351-357.

5. van Mello NM, Mol F, Opmeer BC, et al. Diagnostic value of serum hCG on the outcome of pregnancy of unknown location: a systematic review and meta-analysis. *Hum Reprod Update*. 2012;18:603-617.
6. 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 Obstet Gynecol*. 2009;33:704-710.
7. 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;118:693-697.
8. 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;113:521-527.
9. Condous G, Kirk E, Lu C, et al. Diagnostic accuracy of varying discriminatory zones for the prediction of ectopic pregnancy in women with a pregnancy of unknown location. *Ultrasound Obstet Gynecol*. 2005;26:770-775.
10. 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 >3000 mIU/mL, but without a gestational sac on ultrasonography. *Acad Emerg Med*. 1997;4:967-971.
11. Guha S, Ayim F, Ludlow J, et al. Triaging pregnancies of unknown location: the performance of protocols based on single serum progesterone or repeated serum hCG levels. *Hum Reprod*. 2014;29:938-945.
12. Silva C, Sammel MD, Zhou L, Gracia C, Hummel AC, Barnhart K. Human chorionic gonadotropin profile for women with ectopic pregnancy. *Obstet Gynecol*. 2006;107:605-610.
13. 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*. 1995;1:258-262.
14. Mol BWJ, Hajenius PJ, Engelsbel S, et al. Serum human chorionic gonadotropin measurement in the diagnosis of ectopic pregnancy when transvaginal sonography is inconclusive. *Fertil Steril*. 1998;70:972-981.
15. Verhaegen J, Gallos ID, van Mello NM, et al. Accuracy of single progesterone test to predict early pregnancy outcome in women with pain or bleeding: meta-analysis of cohort studies. *BMJ*. 2012;345:e6077.
16. Bulletins—Gynecology ACoOaGCoP. ACOG Practice Bulletin No. 193: Tubal Ectopic Pregnancy. *Obstet Gynecol*. 2018;131:e91-e103.
17. Pereira PP, Cabar FR, Gomez Ú, Francisco RPV. Pregnancy of unknown location. *Clinics*. 2019;74:e1111.
18. Medicine PCoASfR. Medical treatment of ectopic pregnancy: a committee opinion. *Fertil Steril*. 2013;100:638-644.
19. Huchon C, Deffieux X, Beucher G, et al. Pregnancy loss: French clinical practice guidelines. *Eur J Obstet Gynecol Reprod Biol*. 2016;201:18-26.
20. Bignardi T, Condous G, Alhamdan D, et al. The hCG ratio can predict the ultimate viability of the intrauterine pregnancies of uncertain viability in the pregnancy of unknown location population. *Hum Reprod*. 2008;23:1964-1967.
21. Helmy S, Bader Y, Pablik E, et al. Cut-off value of initial serum β-hCG level predicting a successful MTX therapy in tubal ectopic pregnancy: a retrospective cohort study. *Eur J Obstet Gynecol Reprod Biol*. 2014;179:175-180.
22. Webster K, Eadon H, Fishburn S, Kumar G, Committee G. Ectopic pregnancy and miscarriage: diagnosis and initial management: summary of updated NICE guidance. *BMJ*. 2019;367:l6283.
23. Bobdiwala S, Christodoulou E, Farren J, et al. Triaging women with pregnancy of unknown location using two-step protocol including M6 model: clinical implementation study. *Ultrasound Obstet Gynecol*. 2020;55:105-114.
24. Christodoulou E, Bobdiwala S, Kyriacou C, et al. External validation of models to predict the outcome of pregnancies of unknown location: a multicentre cohort study. *BJOG*. 2021;128:552-562.
25. Van Calster B, Bobdiwala S, Guha S, et al. 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 Obstet Gynecol*. 2016;48:642-649.
26. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
27. NEQAS U. UK NEQAS. 2021.
28. Pathologists TRCo. National Quality Assurance Advisory Panel (NQAAP). 2021.
29. UKAS. UKAS: United Kingdom Accreditation Service. 2021.
30. (ISO) IOfS. International Organization for Standardization (ISO) 15189:2012(en) Medical Laboratories - requirements for quality and competence. 2021.
31. Van Calster B, Abdallah Y, Guha S, et al. Rationalizing the management of pregnancies of unknown location: temporal and external validation of a risk prediction model on 1962 pregnancies. *Hum Reprod*. 2013;28:609-616.
32. Bobdiwala S, Guha S, Van Calster B, et al. 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;31:1425-1435.
33. Sison C, Glaz J. Simultaneous confidence intervals and sample size determination for multinomial proportions. *J Am Stat Assoc*. 1995;90:366-369.
34. Harrell F. Regression Modelling Strategies 2e: Springer; 2015.
35. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393.
36. Crochet JR, Peavey MC, Price TM, Behera MA. Spontaneous pregnancy reaches viability after low first trimester serum progesterone: a case report. *J Reprod Med*. 2012;57:171-174.
37. Konrad G. First-trimester bleeding with falling HCG: don't assume miscarriage. *Can Fam Physician*. 2007;53:831-832.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Bobdiwala S, Kyriacou C, Christodoulou E, et al. Evaluating cut-off levels for progesterone, β human chorionic gonadotropin and β human chorionic gonadotropin ratio to exclude pregnancy viability in women with a pregnancy of unknown location: A prospective multicenter cohort study. *Acta Obstet Gynecol Scand*. 2022;101:46–55. doi:10.1111/aogs.14295