Accepted: 22 December 2021 | Published Online 13 February 2022

DOI: 10.1111/1471-0528.17096

BJOG An International Journal of Obstetrics and Gynaecology

RESEARCH ARTICLE

Epidemiology

Serum PlGF compared with PAPP-A in first trimester screening for preterm pre-eclampsia: Adjusting for the effect of aspirin treatment

David Wright[1] | **Min Yi Tan**[2] | **Neil O'Gorman**[2] | **Argyro Syngelaki**[2] | **Kypros H. Nicolaides**[2]

[1]Institute of Health Research, University of Exeter, Exeter, UK

[2]Harris Birthright Research Centre for Fetal Medicine, King's College Hospital, London, UK

Correspondence
KH Nicolaides, Harris Birthright Research Centre for Fetal Medicine, Fetal Medicine Research Institute, King's College Hospital, Denmark Hill, London SE5 8BB, UK.
Email: kypros@fetalmedicine.com

Funding information
Fetal Medicine Foundation; National Institute for Health Research Efficacy and Mechanism Evaluation (NIHR EME) Programme, Grant/Award Number: 14/01/02; European Union Seventh Framework Program, Grant/Award Number: FP7-HEALTH-2013-INNOVATION-2

Abstract

Objective: To compare the predictive performance for preterm-pre-eclampsia (PE) in first-trimester screening by serum placental growth factor (PlGF) versus pregnancy associated plasma protein-A (PAPP-A), in combination with maternal risk factors, mean arterial pressure (MAP) and uterine artery pulsatility index (UtA-PI), after adjustment for the effect of aspirin in women receiving this treatment.

Design: Non-intervention multicentre screening studies for PE in singleton pregnancies.

Setting: Maternity hospitals.

Population: Two independent prospective studies of 8775 and 16 451 women with singleton pregnancies attending for routine assessment at 11^{+0}–13^{+6} weeks' gestation.

Methods: The competing risks model was used to estimate patient-specific risks of delivery with PE at <37 weeks' gestation based on maternal risk factors and combinations with MAP, UtA-PI and either PlGF or PAPP-A. McNemar's test was used to compare the detection rate (DR) of preterm-PE of screening utilising PlGF versus PAPP-A, after adjustments for the effects of aspirin.

Main outcome measure: Predictive performance for preterm-PE.

Results: In the combined data of 25 226 women, including 678 (2.7%) who developed PE, there were 194(0.8%) with preterm-PE. Addition of PlGF improved the DR of preterm-PE, at 10% screen positive rate, by 18.4% (95% CI 12.2–24.6) in screening by maternal risk factors, by 19.9% (95% CI 13.6–26.2) in screening by maternal factors and MAP, and by 7.0% (95% CI 2.3–11.6) in screening by maternal factors, MAP and UtA-PI. PAPP-A did not significantly improve the DR provided by any combination of biomarkers.

Conclusion: The predictive performance of first trimester PlGF for preterm-PE is superior to that of PAPP-A.

KEYWORDS
aspirin, ASPRE, competing risk model, first trimester screening, mean arterial pressure, placental growth factor, pre-eclampsia, pregnancy associated plasma protein-a, pyramid of pregnancy care, SPREE, uterine artery Doppler

Linked article: This article is commented on by Howard Cuckle, pp. 1318 in this issue. To view this minicommentary visit https://doi.org/10.1111/1471-0528.17099.

This article includes Author Insights, a video abstract available at https://vimeo.com/bjogabstracts/authorinsights17096

Tweetable abstract: The predictive performance for preterm pre-eclampsia of first trimester PlGF is superior to that of PAPP-A.

1 | INTRODUCTION

First trimester assessment of risk for preterm pre-eclampsia (PE) is beneficial because treatment of the high-risk group with aspirin (150 mg/day from 11–14 weeks' gestation to 36 weeks) reduces the rate of preterm-PE with delivery at <37 weeks by about 60%.[1,2] The method of identifying the high-risk group was the competing risks model, which combines maternal risk factors and biomarkers.[3-5] Development of the competing risks model based on maternal characteristics and medical history was derived from the study of 120 492 singleton pregnancies undergoing screening at 11–13 weeks' gestation; factors contributing to the risk of PE included age, weight, height, racial origin, parity, personal and family history of PE, chronic hypertension, diabetes mellitus, systemic lupus erythematosus or antiphospholipid syndrome, method of conception and interpregnancy interval.[4] In a subsequent first-trimester screening study of 35 948 singleton pregnancies, we extended the model to include uterine artery pulsatility index (UtA-PI), mean arterial pressure (MAP), serum pregnancy associated plasma protein-A (PAPP-A) and placental growth factor (PlGF) multiple of the median (MoM) values.[5] The detection rate of preterm-PE, at a 10% false-positive rate, increased from about 50% in screening by maternal risk factors alone, to 70% with the addition of MAP and UtA-PI, and then to 75% with the further addition of PlGF; PAPP-A improved the prediction of maternal risk factors by 4%, but had no additive effect to the combination of MAP and UtA-PI or MAP, UtA-PI and PlGF.[5] The predictive performance of the competing risks model was subsequently validated in two multicentre, non-intervention studies in 8775 and 16 451 women with singleton pregnancies; in both studies there was a very high discrimination between affected and unaffected pregnancies and good agreement between the predicted risks and observed incidence of PE.[6-8]

We recently reported results from a study comparing the performance of screening with PAPP-A and PlGF in 57 131 pregnancies including 452 (0.8%) with preterm-PE.[9] When used in combination with maternal risk factors, MAP and UtA-PI, with a screen positive rate of 10%, the detection rate for preterm-PE with PlGF was superior to that with PAPP-A (74.1% versus 67.1%), increasing the detection rate by (7.1%; 95% CI 3.8–10.6%; P = 0.0001). The study was considered to be non-interventional because women and their medical team were not given the risks from screening and they did not receive aspirin on the basis of such risks.[5,6,7,9] However, the medical teams would have been aware of the PAPP-A results and may have recommended aspirin if the level of PAPP-A was low.[10] This would have introduced bias against the predictive performance of PAPP-A and may in part have contributed to our finding that the predictive performance for preterm-PE of PAPP-A was poor in comparison with PlGF.

The objective of this study is to compare the predictive performance for preterm-PE in first-trimester screening by serum PlGF with that using PAPP-A, in combinations with maternal risk factors, MAP and UtA-PI, after adjustment for the effect of aspirin in patients who received this medication.

2 | METHODS

2.1 | Study populations

This study is based on data from 25 226 women including 194 (0.8%) who delivered preterm with PE. These data were taken from two previously reported prospective non-interventional studies of screening for pre-eclampsia at 11^{+0} to 13^{+6} weeks.[4,5] The studies were chosen because they satisfied the following criteria:

- risk information was not provided to the women or their obstetricians;
- there was no planned intervention with aspirin;
- aspirin treatment was ascertained as fully as possible;
- the risks were computed prospectively using the algorithm developed and parameterised independently.[4,5]

These criteria make this an independent prospective evaluation, with relatively low intake of aspirin, but that can be adjusted for the effect of aspirin treatment.

The first study, referred to as SQS (screening quality study), was comprised of 8775 singleton pregnancies undergoing first-trimester screening for PE, using the competing risks model,[3] in 12 maternity hospitals in England, Spain, Belgium, Italy and Greece, between February and September 2015.[6] This study was carried out before ASPRE (Combined Multimarker Screening and Randomized Patient Treatment with Aspirin for Evidence-Based Preeclampsia Prevention) trial [1] and was primarily designed to examine the feasibility of multicentre screening and establish methods for quality assurance of the biomarkers. The results from screening were not made available to the patients or their obstetricians. The second study, referred to as SPREE, was a multicentre cohort study in 16 451 women carried out in seven National Health Service maternity hospitals in England, between April and December 2016.[7] This study was specifically designed to examine the performance of screening by the competing risks model in comparison with that of the method advocated by the National Institute for Health and Clinical Excellence (NICE);[11] the results from screening by the competing risks model were not made available to the patients or their obstetricians.

In both studies, women with singleton pregnancies in the participating hospitals had a routine examination at 11^{+0}–13^{+6} weeks' gestation. This visit included, first, recording of

maternal characteristics and medical history,[4] second, measurement of the left and right UtA-PI by transabdominal colour Doppler ultrasound and calculation of the mean PI,[12] thirdly, measurement of MAP by validated automated devices and standardised protocol,[13] and fourthly, measurement of serum concentration of PlGF and PAPP-A (DELFIA Xpress system, PerkinElmer Life and Analytical Sciences, Waltham, MA,USA or BRAHMS KRYPTOR analyser, ThermoFisher Scientific, Hennigsdorf, Germany). The measurements of MAP were carried out by healthcare assistants or sonographers who had received specific training for this purpose and measurements of UtA-PI were performed by doctors or sonographers who had obtained the Fetal Medicine Foundation Certificate of Competence in Doppler ultrasound. In both studies, quality control was applied on a monthly basis to achieve consistency of measurement of biomarkers across different hospitals throughout the duration of the study. The distribution of measurements of MAP and UtA-PI were reported to the coordinator who provided feedback and, if necessary, retraining of the personnel with large deviations from the expected values. Similarly, the laboratories were provided with diagnostics for PlGF and PAPP-A measurements so that appropriate corrective actions could be undertaken. Gestational age was determined from the fetal crown–rump length.[14] The women gave written informed consent to participate in the studies, which were approved by the relevant research ethics committee in each participating hospital.

The inclusion criteria were singleton pregnancy undergoing first trimester combined screening for PE and subsequently delivering a morphologically normal live birth or stillbirth at ≥24 weeks' gestation. We excluded pregnancies with aneuploidies and major fetal abnormalities and those ending in termination, miscarriage or fetal death before 24 weeks. There was no patient involvement in the design of the study.

2.2 | Study funding

This study was supported by grants from the Fetal Medicine Foundation (UK Charity No: 1037116). Reagents and equipment for the measurement of serum placental growth factor were provided free of charge by Roche Diagnostics and by PerkinElmer Life and Analytical Sciences. These bodies had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

2.3 | Definition of pre-eclampsia

Data on pregnancy outcome were collected from the hospital maternity records or the women's general medical practitioners. The obstetric records of all women with chronic hypertension or pregnancy-associated hypertension were examined to determine the diagnosis of PE. This was based on the finding of new onset hypertension (systolic blood pressure of ≥140 mm Hg or diastolic blood pressure of ≥90 mmHg on at least two occasions 4 hours apart developing after 20 weeks' gestation in previously normotensive women) or chronic hypertension and at least one of the following: proteinuria (≥300 mg/24 h or protein to creatinine ratio ≥ 30 mg/mmol or ≥2 + on dipstick testing), renal insufficiency with serum creatinine >97 μmol/l in the absence of underlying renal disease, hepatic dysfunction with blood concentration of transaminases more than twice the upper limit of normal (≥65 IU/l for our laboratory), thrombocytopenia (platelet count <100 000/microl), neurological complications (e.g. cerebral or visual symptoms) or pulmonary oedema.[15]

2.4 | Outcome measure

The outcome measure was the predictive performance for preterm-PE of PlGF in comparison with PAPP-A, after adjustments for the effects of aspirin in women receiving this treatment, assessed as areas under the receiver operation characteristic (ROC) curves and detection rate at 10% screen positive rate.

2.5 | Statistical methods

Risks of delivery with PE at <37 weeks' gestation were calculated using the competing risks approach.[3] In this approach, every woman has a personalised distribution of gestational age at delivery with PE, and whether she develops PE or not before a specified gestational age depends on competition between delivery before or after development of PE. The personalised distribution comes from the application of Bayes theorem to combine a prior distribution, determined from maternal demographic characteristics and medical history with likelihoods from biomarkers. At 11–13 weeks' gestation, useful biomarkers for subsequent development of PE are MAP, UtA-PI, PlGF and PAPP-A. The measured values for these biomarkers are expressed as multiples of the median (MoM) after adjustment for gestational age, weight, race, method of conception, medical conditions, elements from the obstetric history associated with the individual on whom they are measured, and the instrument used for measurement. The posterior distribution of gestational age at delivery with PE is obtained using Bayes theorem by multiplying the prior probability density from maternal factors by the likelihood function from biomarker MoM values.

Risks were computed using the same pre-specified algorithm with pre-specified parameters.[5] McNemar's test was used to compare differences in detection rates between screening with and without PlGF and PAPP-A, for fixed screen positive rates of 10%. Areas under the ROC curves were compared using the DeLong test.[16]

Prevention of PE by treatment with aspirin in a high-risk group converts outcomes that would, without aspirin, be true positives into false positives and biases the assessment of screening performance. When aspirin treatment is given in those with low PAPP-A MoM values, this would bias comparisons against

TABLE 1 Maternal and pregnancy characteristics of the study populations

Variables	SQS[6] (*n* = 8775)	SPREE[7] (*n* = 16 451)
Maternal age in years, median (IQR)	31.5 (27.3–35.0)	31.5 (27.4–35.1)*
Maternal weight in kg, median (IQR)	66.5 (59.0–77.0)	67.0 (59.2–78.0)*
Maternal height in cm, median (IQR)	164.5 (160.0–169.0)	165.0 (160.0–169.0)*
Body mass index in kg/m^2, median (IQR)	24.5 (21.9–28.4)	24.7 (22.0–28.7)*
Gestational age in weeks, median (IQR)	12.7 (12.3–13.1)	12.9 (12.4–13.3)*
Racial origin		
White, *n* (%)	6883 (78.4)	11 922 (72.5)
Black, *n* (%)	1090 (12.4)	2337 (14.2)
South Asian, *n* (%)	462 (5.3)	1361 (8.3)
East Asian, *n* (%)	154 (1.8)	407 (2.5)
Mixed, *n* (%)	186 (2.1)	424 (2.6)
Conception		
Natural	8483 (96.7)	15 765 (95.8)
In vitro fertilisation	227 (2.6)	561 (3.4)
Assisted by use of ovulation drugs	64 (0.7)	125 (0.8)
Medical history		
Chronic hypertension	100 (1.1)	137 (0.8)
Diabetes mellitus type 1	31 (0.4)	46 (0.3)
Diabetes mellitus type 2	37 (0.4)	71 (0.4)
SLE/APS	19 (0.2)	39 (0.2)
Cigarette smokers, *n* (%)	732 (8.3)	1105 (6.7)
Family history of pre-eclampsia, (*n*, %)	339 (3.9)	535 (3.3)
Parity		
Nulliparous, *n* (%)	4127 (47.0)	7587 (46.1)
Parous with no previous PE, *n* (%)	4459 (50.8)	8483 (51.6)
Parous with previous PE, *n* (%)	189 (2.2)	381 (2.3)
Aspirin treatment	196 (2.2)	870 (5.3)
Pre-eclampsia		
Total, *n* (%)	239 (2.7)	439 (2.7)
Delivery <37 weeks, *n* (%)	59 (0.7)	135 (0.8)

APS, antiphospholipid syndrome; IQR, interquartile range; SLE, systemic lupus erythematosus.

screening tests including PAPP-A. To remove this bias, 10 datasets were generated in which cases of preterm-PE prevented by aspirin were replaced by cases. These without-aspirin datasets were produced by simulating outcomes for women who received aspirin in the original dataset and delivered without preterm-PE. For those who were treated with aspirin and did not have preterm-PE, the without-aspirin outcome, either preterm-PE or not, was simulated from a probability model using the risk of preterm-PE to determine the outcome probability. This process of imputation was implemented with Markov chain Monte Carlo methods using a model in which the incidence of PE that would have occurred, had it not been for the effect of treatment, was determined from a logistic regression model dependent on the logit transformation of risk using all four biomarkers (MAP, UtA-PI, PlGF and PAPP-A). The imputation model assumed that aspirin reduced the incidence of preterm-PE by a pre-specified probability of 0.62, as found in ASPRE.[1] Estimates from the 10 without-aspirin datasets were pooled using Rubin's Rules.[17]

The WINBUGS software was used for multiple imputation of preterm-PE that were prevented by treatment with aspirin (Appendix S1).[18] R software was used for other analyses.[19] The package pROC was used for the ROC curve analysis.[20] The MICE package was used for pooling estimates from the multiple imputations.[21]

3 | RESULTS

3.1 | Study population

Characteristics of the study populations are summarised in Table 1. In SPREE,[7] 5.3% (870/16451) of women reported that they took aspirin during pregnancy, but

only 18 (0.1%) took 150 mg/day starting from <16 weeks' gestation and continuing to 36 weeks or delivery, which is the dose and duration of treatment shown in ASPRE substantially to reduce the rate of preterm-PE.[1] In SQS,[6] 2.2% (196/8775) of women took aspirin but in all cases the daily dose was 75 mg. In the SPREE, aspirin intake was associated with low PAPP-A MoM values; among the 800 women with PAPP-A MoM values <0.4, 13.9% took aspirin as compared with 4.8% in those with PAPP-A levels >0.4 MoM ($P < 0.0001$).[7] There was no such evidence in the SQS population; among the 427 women with PAPP-A MoM values <0.4, 2.2% took aspirin as compared with 2.3% in those with PAPP-A MoM values of ≥0.4.[6]

3.2 | Predictive performance for preterm-PE

The predictive performance for preterm-PE of first trimester screening by PlGF versus PAPP-A in combinations with maternal risk factors, MAP and UtA-PI is compared in Tables 2 and 3 and Tables S1 and S2, and illustrated in Figures 1 and 2 and Figures S1–S4. For all marker combinations there is strong evidence of superiority of PlGF over PAPP-A, for both the original data and the data adjusted for aspirin treatment.

The areas under the ROC curves for preterm-PE were significantly higher when PlGF was used, rather than PAPP-A, in combination with maternal risk factors, maternal risk factors plus MAP, maternal risk factors plus UtA-PI, and maternal risk factors plus MAP plus UtA-PI, both with and without adjustment for the effect of aspirin (Table 2, Figures S1–S4).

Similarly, the detection rate of preterm-PE, at 10% screen positive rate, in screening by any combination of biomarkers involving PlGF was higher than that of screening with PAPP-A (Table 3). The increase in detection rate from adding PlGF or PAPP-A to combinations of maternal risk factors and biomarkers is shown in Figure 1 and Table 3. There is good agreement between the SPREE and SQS datasets. The benefit from the addition of PAPP-A is only about 5% and the upper limits on the confidence intervals suggest that increases in detection rates above 10% can be ruled out. In contrast, addition of PlGF increases detection rates of screening by maternal factors or maternal factors plus MAP by about 20% and the lower confidence limits suggest that increases less than 5% can be ruled out (Table 3, Figure 1).

Figure 2 and Table 3 show the differences in detection rates of preterm-PE from the addition of PlGF versus the addition of PAPP-A to maternal risk factors and combinations with MAP and UtA-PI. Again, there is a high degree of consistency between the two datasets and in the combined data, PlGF is superior to PAPP-A for any combination of maternal risk factors with MAP and UtA-PI.

In the adjustments for aspirin use we considered only those who received aspirin and assumed the risk reduction to be 62%. In Tables S1 and S2 we provide data on the extreme assumption that all women with PAPP-A <0.4 MoM were given aspirin and that aspirin is 100% effective in preventing preterm-PE; again, PlGF was found to be superior to PAPP-A for any combination of maternal risk factors with MAP and UtA-PI. The estimated increase in detection rate of preterm-PE, at 10% screen positive rate with addition of PlGF was about 18% in screening by maternal risk factors or maternal risk factors plus MAP, about 10% screening by maternal risk factors plus UtA-PI, and 8% in screening by maternal risk factors plus MAP plus UtA-PI. In contrast, the estimated increase in detection rate of preterm-PE, at 10% screen positive rate, with addition of PAPP-A was about 8% in screening by maternal risk factors, 6% in screening by maternal risk factors plus MAP, and 5% in screening by maternal risk factors plus UtA-PI; there was no significant increase in detection rate with addition of PAPP-A to maternal risk factors plus MAP plus UtA-PI.

4 | DISCUSSION

4.1 | Main findings

This study on a total of 25 226 women, including 194 (0.8%) who developed preterm-PE, was derived from two prospective, multicentre screening studies in which the estimated risks of preterm-PE were not provided to the women or their obstetricians and where there was good ascertainment of aspirin intake, which was only 4.2%. There are four main findings:

- in first-trimester screening for preterm-PE by maternal risk factors, PAPP-A improved the detection rate, at 10% screen positive rate, by only about 5%, but in screening by combination of maternal risk factors with MAP and UtA-PI addition of PAPP-A did not significantly improve the detection rate;
- screening by PlGF improved the detection rate of screening by maternal risk factors and maternal risk factors plus MAP by about 20% and by a combination of maternal risk factors, MAP and UtA-PI by about 7%;
- the areas under the ROC curves for prediction of preterm-PE by any combination of maternal risk factors with MAP and UtA-PI was significantly higher with the addition of PlGF than PAPP-A;
- the superior performance of screening by PlGF than PAPP-A persisted after adjustment for intake of aspirin, even at the extreme assumption that all women with PAPP-A <0.4 MoM received aspirin and that aspirin was 100% effective in preventing preterm-PE.

We have previously reported that prevention of preterm-PE by aspirin necessitates that the onset of treatment is <16 weeks' gestation and is continued to 36 weeks or delivery and that the minimum daily dose of the drug is 100 mg;[1,2] in our study, 4.2% (1066/25 226) of the women

TABLE 2 Increases in predictive performance for preterm pre-eclampsia, assessed by areas under the curve with 95% confidence interval, from the addition of PlGF and PAPP-A and comparison of predictive performance with use of PlGF versus PAPP-A for the original data (top) and with multiple imputation of events prevented by treatment with aspirin (bottom)

			Addition of PlGF		Addition of PAPP-A		Base+PLGF versus base +PAPP-A
Base test	**Population**	**AUC Base test**	**AUC Base +PLGF**	***P*-values[a]**	**AUC Base+PAPP-A**	***P*-values[a]**	***P*-values[a]**
No Imputation							
History	SPREE, *n* = 135	0.7761 (0.7357–0.8165)	0.8643 (0.8348–0.8939)	<0.0001	0.8009 (0.7640–0.8378)	0.010	<0.0001
	SQS, *n* = 59	0.7600 (0.6994–0.8207)	0.8739 (0.8304–0.9174)	<0.0001	0.7844 (0.7243–0.8446)	0.043	<0.0001
	Combined, *n* = 194	0.7716 (0.7380–0.8051)	0.8667 (0.8423–0.8912)	<0.0001	0.7964 (0.7650–0.8279)	0.001	<0.0001
History + MAP	SPREE, *n* = 135	0.8457 (0.8171–0.8743)	0.8995 (0.8774–0.9216)	<0.0001	0.8576 (0.8292–0.8860)	0.101	0.0001
	SQS, *n* = 59	0.8258 (0.7799–0.8716)	0.9000 (0.8653–0.9347)	<0.0001	0.8358 (0.7890–0.8827)	0.279	<0.0001
	Combined, *n* = 194	0.8397 (0.8155–0.8639)	0.8993 (0.8806–0.9179)	<0.0001	0.8511 (0.8268–0.8754)	0.048	<0.0001
History + UtA-PI	SPREE, *n* = 135	0.8761 (0.8479–0.9044)	0.9047 (0.8807–0.9286)	0.0004	0.8809 (0.8527–0.9091)	0.348	0.008
	SQS, *n* = 59	0.8549 (0.8026–0.9072)	0.9013 (0.8589–0.9437)	0.0004	0.8598 (0.8070–0.9126)	0.468	0.0007
	Combined, *n* = 194	0.8700 (0.8448–0.8951)	0.9033 (0.8824–0.9243)	<0.0001	0.8748 (0.8496–0.9000)	0.233	0.0001
History + MAP + UtA-PI	SPREE, *n* = 135	0.9196 (0.8985–0.9407)	0.9318 (0.9130–0.9506)	0.022	0.9207 (0.8994–0.9420)	0.745	0.054
	SQS, n = 59	0.8969 (0.8623–0.9315)	0.9254 (0.8942–0.9565)	0.003	0.8968 (0.8600–0.9336)	0.996	0.002
	Combined, *n* = 194	0.9128 (0.8948–0.9308)	0.9297 (0.9136–0.9458)	0.0004	0.9136 (0.8951–0.9322)	0.769	0.001
With imputation							
History	SPREE, *n* = 158.7	0.7942 (0.7557–0.8327)	0.8760 (0.8482–0.9037)	<0.0001	0.8184 (0.7834–0.8533)	0.006	<0.0001
	SQS, *n* = 63.2	0.7751 (0.7159–0.8344)	0.8804 (0.8388–0.9219)	<0.0001	0.7976 (0.7393–0.8559)	0.047	<0.0001
	Combined, *n* = 221.9	0.7890 (0.7573–0.8208)	0.8768 (0.8539–0.8998)	<0.0001	0.8129 (0.7832–0.8425)	0.0007	<0.0001
History + MAP	SPREE, *n* = 158.7	0.8585 (0.8313–0.8857)	0.9082 (0.8874–0.9289)	<0.0001	0.8701 (0.8435–0.8968)	0.075	0.0001
	SQS, *n* = 63.2	0.8368 (0.7922–0.8814)	0.9056 (0.8725–0.9387)	<0.0001	0.8461 (0.8009–0.8913)	0.286	0.0001
	Combined, *n* = 221.9	0.8522 (0.8292–0.8752)	0.9070 (0.8896–0.9245)	<0.0001	0.8632 (0.8404–0.8861)	0.036	<0.0001
History + UtA-PI	SPREE, *n* = 158.7	0.8851 (0.8586–0.9116)	0.9123 (0.8899–0.9346)	0.0002	0.8902 (0.8639–0.9165)	0.268	0.007
	SQS, *n* = 63.2	0.8636 (0.8137–0.9136)	0.9059 (0.8657–0.9461)	0.0007	0.868 (0.8177–0.9183)	0.490	0.001
	Combined, *n* = 221.9	0.8790 (0.8556–0.9025)	0.9101 (0.8905–0.9297)	<0.0001	0.884 (0.8605–0.9075)	0.185	0.0001
History + MAP + UtA-PI	SPREE, *n* = 158.7	0.9253 (0.9055–0.9451)	0.9372 (0.9197–0.9547)	0.016	0.9268 (0.907–0.9466)	0.630	0.048
	SQS, *n* = 63.2	0.9032 (0.8701–0.9364)	0.9292 (0.8996–0.9587)	0.005	0.9031 (0.868–0.9382)	0.975	0.003
	Combined, *n* = 221.9	0.9189 (0.9019–0.9358)	0.9347 (0.9196–0.9497)	0.0003	0.92 (0.9027–0.9373)	0.674	0.001

Abbreviations: MAP, mean arterial pressure; PAPP-A, pregnancy-associated plasma protein-A; PlGF, placental growth factor; UtA-PI, uterine artery pulsatility index.

[a]McNemar's test.

(Continues)

TABLE 3 Increases in detection rate for preterm pre-eclampsia with 95% confidence intervals at a 10% screen positive rate from the addition of PlGF and PAPP-A and comparison of predictive performance with use of PlGF versus PAPP-A for the original data (top) and with multiple imputation of events prevented by treatment with aspirin (bottom)

			Addition of PlGF			Addition of PAPP-A			PlGF versus PAPP-A	
Base test	Population	DR Base test	DR Base + PLGF	Increase over Base	*P*[a]	Base + PAPP-A	Increase over Base	*P*[a]	Difference In DR	*P*[a]
No imputation										
History	SPREE–*n* = 135	41.5 (33.4–49.7)	59.2 (51.1–67.2)	17.6 (10.0–25.2)	<0.0001	45.8 (37.6–54.0)	4.2 (−0.9 to 9.3)	0.181	13.4 (6.0–20.7)	0.001
	SQS, *n* = 59	40.7 (28.1–53.2)	61.0 (48.6–73.5)	20.3 (10.1–30.6)	0.002	45.8 (33.1–58.5)	5.1 (−2.2 to 12.4)	0.371	15.3 (3–27.5)	0.039
	Combined, *n* = 194	41.3 (34.5–48.1)	59.7 (52.9–66.5)	18.4 (12.2–24.6)	<0.0001	45.8 (38.9–52.7)	4.5 (0.3 –8.7)	0.067	13.9 (7.6–20.3)	0.0001
History + MAP	SPREE, *n* = 135	49.3 (41.1–57.5)	69.0 (61.4–76.6)	19.7 (12.1–27.3)	<0.0001	52.8 (44.6–61.0)	3.5 (−2.1 to 9.2)	0.332	16.2 (8.7–23.7)	0.0001
	SQS, *n* = 59	44.1 (31.4–56.7)	64.4 (52.2–76.6)	20.3 (9.0–31.6)	0.003	49.2 (36.4–61.9)	5.1 (−0.5 to 10.7)	0.248	15.3 (3.0–27.5)	0.039
	Combined, *n* = 194	47.8 (40.9–54.7)	67.7 (61.2–74.1)	19.9 (13.6–26.2)	<0.0001	51.7 (44.8–58.6)	4.0 (−0.3 to 8.3)	0.118	15.9 (9.5–22.3)	<0.0001
History + UtA-PI	SPREE, *n* = 135	62.0 (54.0–70.0)	71.1 (63.7–78.6)	9.2 (2.4–15.9)	0.016	63.4 (55.5–71.3)	1.4 (−3 to 5.8)	0.752	7.7 (1.0–14.5)	0.046
	SQS, *n* = 59	59.3 (46.8–71.9)	72.9 (61.5–84.2)	13.6 (4.8–22.3)	0.013	66.1 (54.0–78.2)	6.8 (−3.6 to 17.1)	0.343	6.8 (−3.6 to 17.1)	0.343
	Combined, *n* = 194	61.2 (54.5–67.9)	71.6 (65.4–77.9)	10.4 (5.0–15.9)	0.0005	64.2 (57.6–70.8)	3.0 (−1.4 to 7.3)	0.264	7.5 (1.8–13.1)	0.018
History + MAP + UtA-PI	SPREE, *n* = 135	73.9 (66.7–81.2)	82.4 (76.1–88.7)	8.5 (2.8–14.1)	0.010	76.1 (69.0–83.1)	2.1 (−3.2 to 7.4)	0.606	6.3 (0.7–11.9)	0.052
	SQS, *n* = 59	71.2 (59.6–82.7)	74.6 (63.5–85.7)	3.4 (−4.7–11.5)	0.683	66.1 (54.0–78.2)	−5.1 (−10.7 to 0.5)	0.248	8.5 (−1.3 to 18.2)	0.182
	Combined, *n* = 194	73.1 (67.0–79.3)	80.1 (74.6–85.6)	7.0 (2.3–11.6)	0.008	73.1 (67.0–79.3)	0.0 (−4.1 to 4.1)	1.000	7.0 (2.1–11.8)	0.011
With imputation										
History	SPREE, *n* = 158.7	45.7 (37.3–54.1)	62.7 (55.1–70.4)	17.0 (9.6–24.4)	<0.0001	50.5 (42.4–58.5)	4.7 (−0.4 to 9.8)	0.069	12.3 (5.5–19.1)	0.0004
	SQS, *n* = 63.2	44.4 (31.9–57)	62.8 (50.8–74.8)	18.4 (8.2–28.6)	0.0004	49.2 (36.6–61.8)	4.7 (−2.1 to 11.6)	0.174	13.6 (1.7–25.5)	0.025
	Combined, *n* = 221.9	45.4 (38.6–52.2)	62.8 (56.3–69.2)	17.4 (11.4–23.4)	<0.0001	50.1 (43.5–56.8)	4.7 (0.6–8.9)	0.025	12.7 (6.8–18.6)	<0.0001
History + MAP	SPREE, *n* = 158.7	53.4 (45.2–61.6)	71.7 (64.4–79.0)	18.3 (11.0–25.5)	<0.0001	56.8 (48.8–64.7)	3.3 (−2 to 8.7)	0.220	14.9 (8–21.8)	<0.0001
	SQS, *n* = 63.2	47.8 (35.2–60.3)	66.4 (54.7–78.2)	18.7 (7.8–29.6)	0.0008	52.5 (40.0–65.0)	4.7 (−0.5 to 10.0)	0.076	13.9 (2.2–25.7)	0.020
	Combined, *n* = 221.9	51.8 (45.1–58.6)	70.2 (64.0–76.4)	18.4 (12.4–24.4)	<0.0001	55.6 (48.9–62.2)	3.7 (−0.4 to 7.8)	0.074	14.6 (8.7–20.6)	<0.0001
History + UtA-PI	SPREE, *n* = 158.7	64.6 (56.9–72.2)	73.3 (66.3–80.4)	8.8 (2.4–15.2)	0.007	66.3 (58.8–73.8)	1.7 (−2.6 to 6.0)	0.437	7.1 (0.8–13.3)	0.026
	SQS, *n* = 63.2	61.7 (49.5–73.9)	73.7 (62.7–84.8)	12.0 (3.3–20.8)	0.007	68.0 (56.3–79.7)	6.3 (−3.4 to 16.0)	0.200	5.7 (−4.4 to 15.8)	0.267
	Combined, *n* = 221.9	63.8 (57.3–70.2)	73.5 (67.6–79.3)	9.7 (4.5–14.9)	0.0002	66.8 (60.5–73)	3 (−1.1 to 7.2)	0.153	6.7 (1.4–11.9)	0.013
History + MAP + UtA-PI	SPREE, *n* = 158.7	75.8 (69–82.6)	83.7 (77.8–89.6)	7.9 (2.5–13.2)	0.004	77.7 (71.1–84.4)	2.0 (−3.0 to 6.9)	0.436	5.9 (0.7–11.2)	0.027
	SQS, *n* = 63.2	72.9 (61.8–84)	75.6 (64.9–86.3)	2.7 (−5.2–10.6)	0.504	68.2 (56.5–79.9)	−4.7 (−10.0 to 0.5)	0.076	7.4 (−2 to 16.9)	0.122
	Combined, *n* = 221.9	75 (69.2, 80.8)	81.4 (76.2, 86.6)	6.4 (2.0, 10.8)	0.005	75.0 (69.2, 80.8)	0 (−3.8 to 3.9)	0.982	6.4 (1.8–10.9)	0.006

MAP, mean arterial pressure; PAPP-A, pregnancy-associated plasma protein-A; PlGF, placental growth factor; UtA-PI, uterine artery pulsatility index.

[a]McNemar's test.

BJOG An International Journal of Obstetrics and Gynaecology

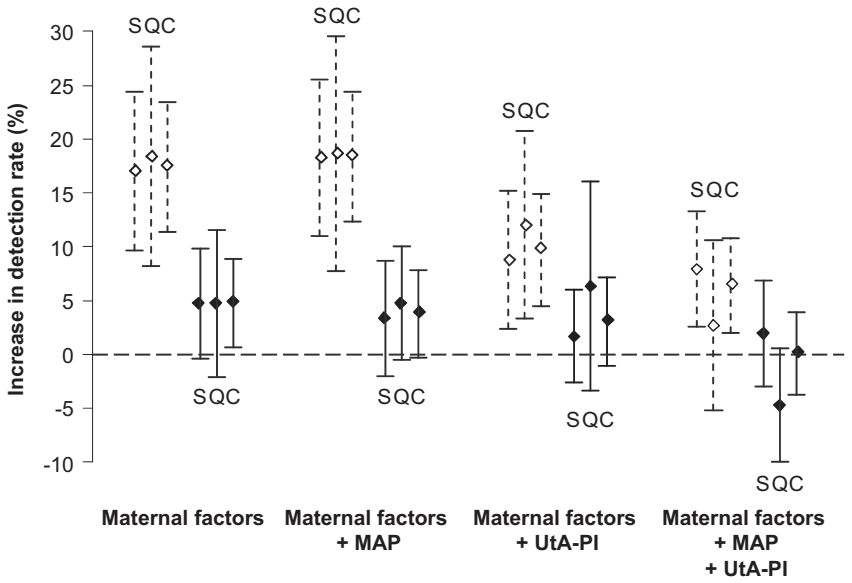

FIGURE 1 Increases in detection rate for preterm pre-eclampsia with 95% confidence intervals at a 10% screen positive rate from the addition of PAPP-A (full line) and PlGF (interrupted line) for the original data (left) and with multiple imputation of events prevented by treatment with aspirin (right). Datasets SPREE (S), SQS (Q) and combined (C)

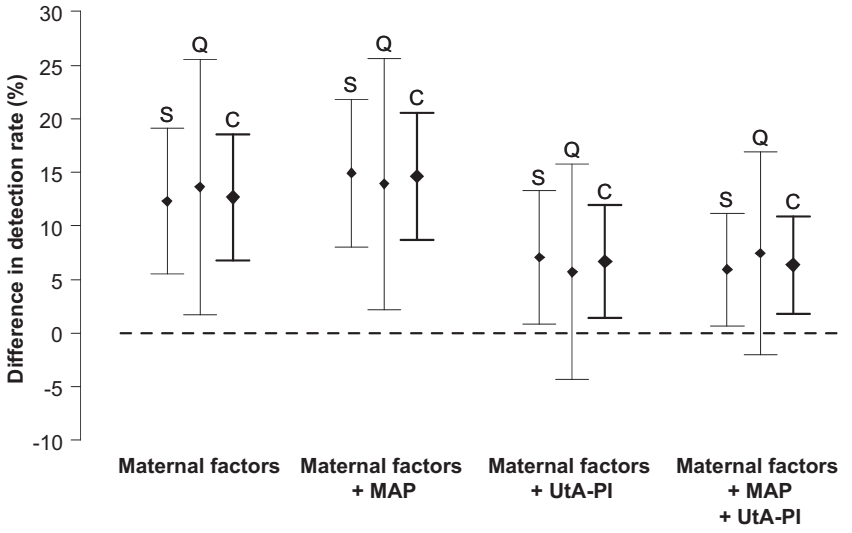

FIGURE 2 Differences in detection rate (PLGF- PAPP-A) for preterm pre-eclampsia with 95% confidence intervals at a 10% screen positive rate from the addition of PlGF versus PAPP-A for the original data (left) and with multiple imputation of events prevented by treatment with aspirin (right). Datasets SPREE (S), SQS (Q) and combined (C)

received aspirin but only 18 (0.07%) fulfilled the above criteria. Despite this, in our adjustments for aspirin use we assumed that in all cases the dose was adequate and that the onset and duration of therapy were appropriate, to avoid any potential criticism of bias against the predictive performance of PAPP-A.

4.2 | Strengths and limitations

The strengths of the study include prospective examination of a large number of pregnant women in several maternity units covering a wide spectrum of demographic and racial characteristics in different European countries and therefore the results are likely to be generalisable. In both SQS and SPREE, measurement of all biomarkers was recorded in all cases and complete follow-up was obtained from >98% of participants. Consistency in data collection was maintained throughout the study period by ensuring adequate training for all investigators based on standardised protocols, regular monitoring and quality assurance of biomarker measurements. An important strength of the study is the adjustment for treatment with aspirin. In this study, as shown in Tables 2 and 3, there was a minimal impact of such adjustments because the number of patients treated with aspirin was small in SQS and SPREE.

In comparative studies investigating the usefulness of biomarkers in the prediction of uncommon events, such as preterm-PE, it is essential that there is a sufficient number of participants. If the study populations are small, substantial differences in performance may not reach statistical significance. Inference should be made on the basis of confidence intervals rather than *P*-values. Our study was a post hoc analysis of data from two previously reported studies that were individually not powered for the comparisons presented here. For example, as shown in Table 3

in the unadjusted results comparing the additive effect of PlGF and that of PAPP-A on the predictive performance for preterm-PE of screening by maternal risk factors, MAP and UtA-PI, there was no significant difference between the two biomarkers in either SQS or SPREE and significant difference was observed only when the results from the two studies were combined.

4.3 | Comparison with results of previous studies

The findings of this study are consistent with those reported previously in 57 131 singleton pregnancies, including 452 (0.8%) that developed preterm-PE.[9] In this previously reported study, the detection rate of preterm-PE, at 10% screen positive rate, by a combination of maternal risk factors, MAP and UtA-PI was improved by the addition of PlGF from 66.8% to 74.1%, whereas the addition of PAPP-A showed little or no benefit.[9] There is some overlap between this and the current study with 12 030 women from King's College Hospital in London or the Medway Maritime Hospital in Kent UK included in both studies comprising 21.1% of the 57 131 of our previously reported study[9]. In the current study PAPP-A showed little evidence of substantial benefit.

Our results are also consistent with those of the study of 35 948 singleton pregnancies used for the development of the competing risks model utilising biomarkers; PAPP-A improved the detection rate of preterm-PE, at 10% false positive rate, provided by maternal risk factors by only 4%, compared with 16% by PlGF, and had little effect on the predictive performance of screening by maternal risk factors plus MAP or maternal risk factors plus MAP and UtA-PI.[5]

A retrospective nested cohort study compared first trimester PAPP-A and PlGF in 30 pregnancies that developed preterm-PE and 754 pregnancies without PE and/or birth of small-for-gestational-age neonates; the detection rate of preterm-PE, at 10% screen positive rate, using a combination of maternal risk factors, MAP and UtA-PI was not significantly different after addition of PlGF (51.7%, 95% CI 32.5–70.6) than after addition of PAPP-A (46.7%, 95% CI 28.3–65.7).[22] Although the authors reached the erroneous conclusion of equivalence of PlGF and PAPP-A, the failure to achieve statistical significance was the consequence of the small size of the study.

A multicentre observational study in China in 10 899 women with singleton pregnancy, including 117 (1.1%) who developed preterm-PE, reported that the detection rate at 10% false positive rate of preterm-PE in screening by maternal risk factors, MAP, UtA-PI and PAPP-A was only 65%, but this was superior to the detection of 56% when PAPP-A was replaced with PlGF.[23] One possible explanation for the poor performance of PlGF may be related to the method of processing the samples; in the 13 participating hospitals, blood samples were centrifuged and ≥2 ml were stored at −20°C for up to 1 week before transportation to a central laboratory where the samples were stored at −70°C and were then analysed within the subsequent 2 weeks. The authors acknowledged that they could not quantify the extent to which differences in room temperature during sample preparation between hospitals or differences in time of day or season might have affected PlGF measurement. In that study, screening performance was generally worse than we have found. Moreover, the distributional properties of PlGF were different, with larger standard deviations for log10 MoM values for PlGF than we have found (0.256[23] versus 0.177[5]).

4.4 | Interpretation of results and implications for clinical practice

First trimester assessment of risk for preterm-PE and treatment of the high-risk group with aspirin reduces the rate of early PE by about 90% and preterm-PE by about 60%.[1] Effective assessment of risk is provided by a combination of maternal risk factors, MAP, UtA-PI and PlGF; SPREE had demonstrated that with this method the predictive performance for preterm-PE is twice as high as that achieved by the risk scoring system recommended by NICE guidelines.[7,11]

Serum PAPP-A is routinely used in first-trimester screening for fetal trisomies[24,25] and it could therefore be argued that this metabolite rather than PlGF should be also used in screening for preterm-PE. However, PAPP-A is a relatively poor biomarker for preterm-PE and is only of substantial benefit when used without UtA-PI, in which case, for a 10% screen positive rate, it improves the detection rate by about 5%, whereas addition of PlGF improves the detection rate by about 20%. When used in combination including UtA-PI, PlGF increases detection rate by around 5%.

Fifteen years ago, effective first-trimester screening for fetal trisomies was implemented in all maternity hospitals in the UK within a few months of the appropriate decision being taken by the National Screening Committee and NICE.[26] The same infrastructure can now be used to expand the aims of first-trimester screening to include identification of women at high-risk of developing preterm-PE and substantially reducing such risk through the prophylactic use of the appropriate dose of aspirin.

5 | CONCLUSION

Previous large studies of first trimester multi-marker screening for preterm-PE show that inclusion of PlGF as a marker has a better performance than inclusion of PAPP-A. These are non-intervention studies (the preterm-PE risk was not reported) but some women took aspirin anyway and it is possible that their decision was based on the PAPP-A result which was not 'blind'. Such a bias would favour PlGF. In this study we apply a statistical method called 'imputation' which overcomes the potential bias. This approach allows comparison of the genuine contributions of serum PlGF and PAPP-A to the prediction of preterm-PE. We found that in first-trimester

prediction for preterm-PE, serum PlGF is a useful biomarker, whereas PAPP-A is a relatively poor biomarker.

CONFLICT OF INTERESTS

None declared. Completed disclosure of interest forms are available to view online as supporting information.

AUTHOR CONTRIBUTIONS

KN and DW conceived and designed the study and wrote the paper. DW provided methodological and statistical expertise and conducted the statistical analysis, NO and AS coordinated the SQS study, MYT and AS coordinated the SPREE study. All authors critically reviewed and approved the final version of the manuscript.

ETHICS APPROVAL

For the SQS study, a favourable ethical opinion was obtained from London-Fulham Research Ethics Committee, reference number 13/LO/1479, on 12 November 2013. For the SPREE study, a favourable ethical opinion was obtained from London–Surrey Borders Research Ethics Committee, reference number 15/LO/216, on 22 December 2015.

DATA AVAILABILITY STATEMENT

Research data are not shared.

REFERENCES

1. Rolnik DL, Wright D, Poon LC, O'Gorman N, Syngelaki A, de Paco MC, et al. Aspirin versus placebo in pregnancies at high risk for preterm preeclampsia. N Engl J Med. 2017;377(7):613–22.
2. Roberge S, Bujold E, Nicolaides KH. Aspirin for the prevention of preterm and term preeclampsia: systematic review and metaanalysis. Am J Obstet Gynecol. 2018;218(3):287–93.e1.
3. Wright D, Wright A, Nicolaides KH. The competing risk approach for prediction of preeclampsia. Am J Obstet Gynecol. 2020;223(1):12–23.e7.
4. Wright D, Syngelaki A, Akolekar R, Poon LC, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. Am J Obstet Gynecol. 2015;213(1):62 e1–10.
5. O'Gorman N, Wright D, Syngelaki A, Akolekar R, Wright A, Poon LC, et al. Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. Am J Obstet Gynecol. 2016;214(1):103.e1–103.e12.
6. O'Gorman N, Wright D, Poon LC, Rolnik DL, Syngelaki A, Wright A, et al. Accuracy of competing-risks model in screening for pre-eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. Ultrasound Obstet Gynecol. 2017;49(6):751–5.
7. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, et al. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. Ultrasound Obstet Gynecol. 2018;51(6):743–50.
8. Wright D, Tan MY, O'Gorman N, Poon LC, Syngelaki A, Wright A, et al. Predictive performance of the competing risk model in screening for preeclampsia. Am J Obstet Gynecol. 2019;220(2):199.e1–e13.
9. Mazer Zumaeta A, Wright A, Syngelaki A, Maritsa VA, Da Silva AB, Nicolaides KH. Screening for pre-eclampsia at 11-13 weeks' gestation: use of pregnancy-associated plasma protein-A, placental growth factor or both. Ultrasound Obstet Gynecol. 2020;56(3):400–7.
10. RCOG. Small-for-Gestational-Age Fetus, Investigation and Management (Green-top Guideline No. 31); 2013. Available from: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg31/
11. National Collaborating Centre for Women's and Children's Health (UK). Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy. London: RCOG Press; 2010.
12. Plasencia W, Maiz N, Bonino S, Kaihura C, Nicolaides KH. Uterine artery Doppler at 11+0 to 13+6 weeks in the prediction of pre-eclampsia. Ultrasound Obstet Gynecol. 2007;30(5):742–9.
13. Poon LC, Zymeri NA, Zamprakou A, Syngelaki A, Nicolaides KH. Protocol for measurement of mean arterial pressure at 11-13 weeks' gestation. Fetal Diagn Ther. 2012;31(1):42–8.
14. Robinson HP, Fleming JE. A critical evaluation of sonar crown rump length measurements. Br J Obstet Gynaecol. 1975;82(9):702–10.
15. American College of Obstetricians and Gynecologists; Task Force on Hypertension in Pregnancy. Gestational hypertension and pre-eclampsia. ACOG Practice Bulletin No. 202. American College of Obstetricians and Gynecologists. Obstet Gynecol. 2019;133(1):e1–25.
16. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.
17. Rubin DB. Multiple imputation after 18+ years. J Am Stat Assoc. 1996;91(3):473–89.
18. Lunn DJ, Thomas A, Best N, et al. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000;10(4):325–7.
19. R Development Core Team. R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.
20. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12(1):77–84.
21. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3):1–67.
22. Noël L, Guy GP, Jones S, Forenc K, Buck E, Papageorghiou AT, et al. Routine first-trimester combined screening for pre-eclampsia: pregnancy-associated plasma protein-A or placental growth factor? Ultrasound Obstet Gynecol. 2021;58(4):540–5.
23. Hu J, Gao J, Liu J, Meng H, Hao N, Song Y, et al. Prospective evaluation of first-trimester screening strategy for preterm pre-eclampsia and its clinical applicability in China. Ultrasound Obstet Gynecol. 2021;58(4):529–39.
24. Spencer K, Souter V, Tul N, Snijders R, Nicolaides KH. A screening program for trisomy 21 at 10-14 weeks using fetal nuchal translucency, maternal serum free beta-human chorionic gonadotropin and pregnancy-associated plasma protein-A. Ultrasound Obstet Gynecol. 1999;13(4):231–7.
25. Nicolaides KH. Screening for fetal aneuploidies at 11 to 13 weeks. Prenat Diagn. 2011;31(1):7–15.
26. UK National Screening Committee. Fetal anomaly screening programme - screening for down's syndrome: UKNSC policy recommendations 2007–2010: model of best practice. London: Department of Health; 2008.

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: Wright D, Tan MY, O'Gorman N, Syngelaki A, Nicolaides KH. Serum PlGF compared with PAPP-A in first trimester screening for preterm pre-eclampsia: Adjusting for the effect of aspirin treatment. BJOG: Int J Obstet Gy. 2022;129:1308–1317. https://doi.org/10.1111/1471-0528.17096