

# Prognostic indicators of severe disease in women with late preterm preeclampsia to guide decision making on timing of delivery: development and validation of prognostic models

Kate Duhig<sup>1</sup>, Paul Seed<sup>1</sup>, Anna Placzek<sup>2</sup>, Jenie Sparkes<sup>1</sup>, Eleanor Hendy<sup>1</sup>, Carolyn Gill<sup>1</sup>, Anna Brockbank<sup>1</sup>, Andrew Shennan<sup>1</sup>, Shakila Thangaratinam<sup>3</sup>, and Lucy Chappell<sup>4</sup>

<sup>1</sup>Kings College London

<sup>2</sup>University of Oxford

<sup>3</sup>WHO Collaborating Centre for Women's Health

<sup>4</sup>King's College London

May 6, 2020

## Abstract

**Objective:** to establish a prognostic model informing optimal timing of delivery in women with late preterm preeclampsia. **Design:** development and validation of a prognostic model **Setting:** prospective cohort study, nested in the PHOENIX trial, in 36 maternity units across England and Wales. **Population:** women with late preterm pre-eclampsia (34+0-36+6 weeks' gestation) **Methods:** prospective recruitment of women in whom blood samples for Placental Growth Factor (PlGF) and soluble fms-like tyrosine kinase-1 (sFlt-1) testing was obtained, alongside clinical data, for use within the 'Prediction of complications in early-onset pre-eclampsia' (PREP)-S model. Candidate variables were compared using standard methods (sensitivity, specificity, Receiver Operator Curve areas). Estimated probability of early delivery from PREP-S was compared to actual event rates by calibration. **Main Outcome Measures:** clinically indicated need for delivery for pre-eclampsia within seven days. **Results:** PlGF testing had high sensitivity (97.9%) for delivery within seven days, but negative predictive value was only 71.4%, with low specificity (8.4%). The area under the curve for PREP-S was 0.64 (standard error (SE) 0.03), for PlGF was 0.60 (SE 0.03), and 0.65 (0.03), and 0.64 (0.03) for PREP-S in combination with PlGF and sFlt-1:PlGF, respectively. **Conclusions:** PlGF-based testing does not add to clinical assessment to determine need for delivery in late preterm pre-eclampsia. Existing models developed in women with early onset pre-eclampsia to predict complications cannot be used to predict clinically indicated need for delivery in women with late preterm pre-eclampsia. **Funding:** NIHR HTA Monitoring Add on Studies Programme (reference 15/59/06). **Keywords:** placental growth factor, preeclampsia, prognosis

## Introduction

Pre-eclampsia affects around 2-3% of all pregnancies,(1) and is associated with potential serious complications for the woman and baby, including multiple maternal organ dysfunction (severe hypertension, renal and liver impairment, abnormal clotting and stroke/ seizures) and fetal morbidity and mortality. Once diagnosed, progression of the syndrome can be unpredictable, and decisions around timing of delivery needs to take into account evolving maternal complications and perinatal morbidity. We have recently completed the multicentre PHOENIX trial, in which we demonstrated that in women with late preterm pre-eclampsia, planned delivery reduces maternal morbidity, whilst increasing neonatal unit admissions (principally for prematurity as the indication), though with no difference in neonatal morbidity (including respiratory distress), compared with expectant management.(2) Of women in this gestational age window (34 to 37 weeks of pregnancy) managed expectantly, over half required delivery for clinical indications before they reached 37 weeks' gestation, and pregnancy was prolonged (compared to planned delivery) by three days only.

Current parameters advised by national guidelines for indicating need for delivery in pre-eclampsia are relatively blunt: e.g. uncontrolled severe maternal hypertension, abnormal maternal haematological/ biochemical indices or fetal compromise on ultrasound or cardiotocography.(3) Novel prognostic models and blood biomarkers for determination of need for delivery in pregnancies with pre-eclampsia are now emerging,(4, 5) but their applicability to contemporaneous populations of women with confirmed late preterm pre-eclampsia needs further evaluation and validation. Existing clinical models, such as those from the Prediction of complications in early-onset pre-eclampsia (PREP) study accurately predict the risk of complications in women with early onset pre-eclampsia before 34 weeks' gestation (PREP-S) and have now been included in UK national guidelines.(3) If accurate, these models and blood markers in women with late preterm pre-eclampsia could potentially enhance the ability of clinicians and women to determine who is at greatest risk of need for delivery, enabling timely surveillance and decisions around use of antenatal corticosteroids or place of care.

The aim of the study was to establish a prognostic model to inform optimal timing of delivery in women with late preterm pre-eclampsia ( $34^{+0}$  to  $36^{+6}$  weeks' gestation), comparing novel candidate biomarkers including plasma placental growth factor (PlGF), and soluble fms-like tyrosine kinase-1 (sFlt-1):PlGF ratio with clinical and routinely collected blood/ urinary parameters to determine clinically indicated need for delivery for pre-eclampsia (or its related complications) within seven days of assessment.

## Methods

We undertook a prospective observational cohort study between February 2016 and December 2018, nested in 36 maternity units participating in the PHOENIX trial (2) in women with late preterm pre-eclampsia. The PHOENIX trial was a multicentre randomised controlled trial, in which 901 women from 46 units across England and Wales with preterm pre-eclampsia from  $34^{+0}$  to  $36^{+6}$  weeks' gestation were randomly allocated to planned delivery or expectant management. Results of the PHOENIX study have been reported separately.(2)

Women were eligible for this nested PEACOCK study if they were between  $34^{+0}$  and  $36^{+6}$  weeks' gestation, with a diagnosis of pre-eclampsia (as defined by the International Society for the Study of Hypertension in Pregnancy),(6) with a singleton or dichorionic diamniotic twin pregnancy and at least one viable fetus. Women were aged 18 years or over and gave written informed consent for participation. Exclusion criteria included a decision to deliver within the next 48 hours. All women eligible for the PHOENIX trial were eligible for participation in the PEACOCK study, whether they agreed or declined randomisation to the main PHOENIX trial. The study was approved by the South Central— Hampshire B Research Ethics Committee (no 13/SC/0645).

Women were approached individually and asked to provide both plasma (EDTA) and serum blood sample at the time of recruitment, which were processed within four hours of sampling. Samples were centrifuged at 1400g for 10 minutes, and the separated supernatant aliquoted and stored at  $-80^{\circ}\text{C}$ . Samples were shipped back to the coordinating centre, thawed, and processed after completion of all participants in the study on an electronic Triage instrument for PlGF (Quidel Cardiovascular Inc: San Diego, CA), and for sFlt-1 and PlGF on the automated Cobas Elecsys assay (Roche Diagnostics, GmbH, Mannheim, Germany) according to the manufacturer's instructions. The readings were concealed from the clinical team involved in the woman's care and all laboratory staff were masked to clinical outcomes. Definitions and outcomes were pre-specified in the study protocol (version 4.0). Outcomes were collected until the primary hospital discharge of the woman and infant.

### Development of the original PREP-S model

PREP-S is a prediction model that was developed and validated in early onset preeclampsia before 34 weeks' gestation, from 53 maternity units across the United Kingdom.(4) The primary outcome for the PREP-S study was maternal complications of preeclampsia that included maternal death, neurological, hepatic, cardiorespiratory, renal or haematological complications, or delivery before 34 weeks' gestation. All candidate predictors identified in the development of the PREP-S as predictor variables were collected in order to determine the performance of PREP-S within our cohort of women with late preterm preeclampsia.

## Clinical predictor variables

PlGF and sFlt-1 concentrations at enrolment were evaluated as predictor variables. PREP-S predictor variables were measured at study entry. PREP-S consisted of the following predictor variables, collected at diagnosis: maternal age (years), gestational age (weeks), exaggerated tendon reflexes, medical history (two or more of the following conditions: chronic hypertension, renal disease, previous history of pre-eclampsia, autoimmune disease and diabetes mellitus), systolic blood pressure (mmHg, highest over 6 hours), abnormal oxygen saturation (<95% on air), platelet count ( $\times 10^9/\text{dl}$ ), alanine aminotransferase (IU/L), serum urea (mmol/L), serum creatinine ( $\mu\text{mol/L}$ ), urine protein creatinine ratio (mg/mmol), treatment with oral/parenteral antihypertensives, treatment with magnesium sulphate. These were combined using the model equation published:(7)

$$S_{(t)} = S_0(t) \exp((\beta_1 * X_1 + [\dots] + \beta_n * X_n))$$

$$S_{(t)} = S_0(t) \exp(-0.031 * \text{maternal age} + 1.514 * (\text{Log}(\text{GA at diagnosis}/10))^{-2} - 0.8345136) + 5.707 * (\text{Log}(\text{GA at diagnosis}/10))^{-2} * \ln(\text{Log}(\text{GA at diagnosis}/10)) - 0.0652155) + 0.122 (\text{exaggerated tendon reflexes}) - 0.169 (\text{one pre-existing medical condition}) - 0.384 (\text{two or more pre-existing medical conditions}) + 0.016 * \text{systolic blood pressure} + 0.797 (\text{oxygen saturation} < 94\% \text{ on air}) - 0.002 * \text{platelet count} + 0.126 * \log(\text{alanine amino transferase}) + 0.605 * \log(\text{serum urea})^2 - 0.144 * \log(\text{serum urea})^3 + 0.265 * \log(\text{serum creatinine}) + 0.080 * \log(\text{protein creatinine ratio}) + 0.176$$

(baseline treatment with any antihypertensive) + 1.066 (baseline treatment with magnesium sulfate))

$S_0(t)$  – baseline survival adjusted for optimism at time  $t$

$$S_0(48 \text{ hrs}) = 0.99142, S_0(72 \text{ hrs}) = 0.98542, S_0(1 \text{ week}) = 0.96492,$$

$$S_0(1 \text{ month}) = 0.87377$$

## Outcomes

The primary outcome for the PEACOCK study was clinically indicated need for delivery for pre-eclampsia (or delivery for related conditions such as eclampsia or HELLP syndrome) within seven days of assessment. Secondary outcomes included clinically indicated need for delivery for pre-eclampsia within 48 hours of assessment and within 14 days of assessment, perinatal deaths and neonatal unit admission. On analysis of the main trial,(2) it became clear that neonatal unit admissions did not directly reflect neonatal morbidity (as intended), but rather clinician behaviour. In the PHOENIX trial,(2) the proportions of infants with confirmed morbidity diagnoses were similar, but admission for the indication of prematurity was higher in the planned delivery group, and it was not possible to mask trial allocation to the attending clinicians. It was therefore decided that neonatal unit admission could not be used in this cohort as a surrogate marker of neonatal morbidity and further analysis of this secondary outcome was not undertaken. There were no perinatal deaths in the trial, and further analysis of this secondary outcome did not proceed.

## Sample Size

It has been recommended that external validation of a prognostic model should ideally involve a minimum of 100 informative events.(8) We estimated that the primary outcome (delivery within seven days due to clinical indication) would occur in around 40% of women receiving expectant management, based on our previous work and other literature. The sample size for estimation of the sensitivity (within 7%) and specificity (within 7%), assuming a sensitivity of 0.90, specificity 0.70, and 95% confidence intervals (2-tailed) required 120 women with the primary outcome (and 180 without) in the expectant management arm, giving a minimum of 10 events per candidate variable. We estimated that two-thirds of the 500 women recruited to PEACOCK would receive expectant management (the group on which the model will be validated). We therefore expected 134 primary outcome events (500 x 68% x 40%).

## Statistical Analysis

The validation sample for the primary analysis (for delivery within 7 days), and secondary analysis evaluating clinically indicated need for delivery for pre-eclampsia within 14 days of assessment was restricted to women in PEACOCK who underwent expectant management: that is, women recruited to PHOENIX (and also enrolled in PEACOCK) who were randomised to the expectant management arm and women who declined the PHOENIX trial and who were recruited to the PEACOCK study only who underwent the usual care strategy of expectant management. An additional analysis was conducted for evaluating clinically indicated need for delivery for pre-eclampsia within 48 hours of assessment which additionally included the PEACOCK women randomised to the planned delivery arm in the PHOENIX trial.

The stages of analysis were as follows: external validation of the PREP-S model, limited updating of the PREP-S model by recalibration, assessment of the model performance of the updated PREP-S model, assessment of the predictive performance of PIGF (Quidel test) and sFlt-1:PIGF (Roche test), comparison of PIGF, sFlt-1:PIGF and PREP-S, assessment of the addition of PIGF and sFlt-1:PIGF to the PREP-S model. The performance of the models was assessed by calibration and discrimination. Model discrimination was assessed primarily using ROC areas, and calibration was assessed and reported graphically using calibration plots and estimated calibration slopes. The recalibrations were additionally reported graphically, with actual event rates compared to predicted rates for specified risk groups. Assessment of the model performance in relation to the primary and secondary outcomes were determined using ROC areas. Test performance of PIGF and sFlt-1:PIGF was evaluated with sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios. When PIGF was assessed as a single predictor, we used a PIGF cut-off of <100 pg/ml. This was based on the evidence that in those presenting <35 weeks' gestation, PIGF <100 pg/ml has a high diagnostic accuracy (0.96; 95% confidence interval, 0.89–0.99) and negative predictive value (0.98; 0.93–0.995) of determining preeclampsia requiring delivery in 14 days.(9) We have previously reported that a PIGF threshold of <100 pg/mL predicted preeclampsia requiring delivery within 14 days or before 37 weeks' gestation (whichever was sooner) with sensitivity and negative predictive values similar to diagnostic accuracy estimates obtained by using a <5th centile cut-off.(9) When sFlt-1:PIGF ratio was reported, the threshold was >38. Kaplan-Meier survival curves of the time from test to delivery were determined, stratified by four categories of risk determined by the PREP-S model. Assessment of PREP-S, PIGF and sFlt-1:PIGF and the combined model was conducted on the primary outcome and all secondary maternal outcomes.

### Missing Data

In line with the approach used in the original PREP study, missing variables were handled as follows: where alanine aminotransferase was not available, aspartate aminotransferase was used instead (like for like); oxygen saturation was assumed to be normal if not recorded in clinical care; no women had exaggerated tendon reflexes, which was imputed as no, as such women were ineligible for the PHOENIX trial; urinary protein:creatinine ratio was derived from 24 hr urinary protein excretion where there was sufficient data to derive a conversion factor; missing values for serum urea were replaced by a value derived from serum creatinine by linear regression (serum urea = 0.053883\*serum creatinine + 0.7874831; for numbers derived from linear regression, as described, in those with sufficient data, the correlation between the measurements was 0.5434 for 264 observations).

### Funding

This study was supported by the National Institute for Health Research (NIHR) HTA Monitoring Add on Studies Programme (project reference 15/59/06), and National Institute for Health Research Professorship (Chappell RP-2014-05-019).

### Results

Between 27 April 2016 and 24 December 2018, we recruited 501 women to the PEACOCK study, across 36 maternity units in England and Wales (Figure S1). There were no statistically or clinically relevant differences across the participants who received expectant management as usual care outside the trial (n=182) and within the PHOENIX trial allocation arm (n=159) (Table 1, Table S1). Women in the two PHOENIX trial

allocation arms (presented here as women in planned delivery group included for secondary analysis) were, as expected, balanced for baseline characteristics (Table S2).

There were similar outcomes in women randomised to expectant management group and those participating in the non-randomised expectant management group (Table 2), whilst outcomes in the planned delivery and randomised expectant management groups (Table S3) reflect those in the larger PHOENIX trial, with earlier gestation at delivery, as expected.(2) In women managed expectantly, 211 of 341 (61.9%) women delivered within seven days; in these women, 81 (50.9%) of those randomised and 95 (52.2%) of those non-randomised (i.e. outside the trial) had indicated delivery due to clinical concerns for maternal or fetal wellbeing. There were no perinatal deaths in the study.

For the PlGF test by the Quidel assay, the test performance for PlGF in determining need for delivery within seven days at low (<100 pg/mL) and very low (PlGF <12 pg/mL) is shown in Table S4. The test performance for PlGF <100 pg/ml in determining need for delivery within seven days had sensitivity of 97.9% (94.8-99.4%) for delivery within seven days; the negative predictive value was 71.4% (41.9-91.6%) and the specificity was 8.4% (4.1-14.9%). Similar test performance statistics for determining need for delivery within 14 days are shown in Table S5 and within two days in Table S6.

For the sFlt-1:PlGF ratio by the Roche assay, the test performance for sFlt-1/PlGF [?]<sup>83</sup> in determining need for delivery within seven days had sensitivity of 91.4 % (86.3-95.1%) for delivery within seven days; the negative predictive value was 60.5% (43.4-76.0%) and the specificity was 20.9% (13.7-29.7%). Similar test performance statistics for determining need for delivery within 14 days are shown in Table S5 and within two days in Table S6.

For evaluation of the PREP-S prognostic model in this cohort, baseline predictor variables were assessed in the PEACOCK study cohort and compared with the original PREP-S cohort (Table S7). There were important differences between the two cohorts, particularly relating to gestation at enrolment (as different inclusion criteria were used), definitions used for (and therefore incidence of) adverse maternal outcomes.

The Receiver Operator Curve (ROC) areas for PlGF (Quidel), sFlt-1:PlGF ratio (Roche) and PREP-S are shown in Table 3 (and Figure 1), with consideration of the PREP-S model for a dichotomised endpoint (delivery within seven days). PREP-S was also assessed in combination with PlGF (Quidel), and with the sFlt-1:PlGF ratio (Roche), treating PREP-S as a single predictor.(4) The corresponding ROC areas for the clinical prediction model (PREP-S), PlGF (Quidel) and sFlt-1:PlGF (Roche) in this cohort in determining need for delivery within seven days was 0.64 (standard error (SE) 0.03), 0.60 (SE 0.03), and 0.63 (SE 0.03) respectively. The ROC area for PREP-S in combination with PlGF (Quidel) in determining need for delivery within seven days was 0.65 (SE 0.03), and PREP-S in combination with sFlt-1:PlGF (Roche) was 0.64 (SE 0.03). PREP-S (when used to determine a binary outcome), and PlGF have limited clinical applicability in this cohort in determining need for delivery within seven days.

Performance of the PREP-S model, PlGF (Quidel), and sFlt-1:PlGF ratio (Roche) was similar in determining delivery in 14 days and in 2 days in this cohort (Table S8 and Figures S2 and S3), with these predictors having limited clinical applicability in this setting.

The Kaplan-Meier time to delivery estimates for women in the expectant management groups, stratified by four PREP-S risk categories (as observed) are shown in Figure S4 and the recalibrated estimates in Figure S5.

Calibration of the PREP-S model is shown in Table S9, with calibration in the large and of the slope assessed for predicting delivery for preeclampsia within seven days. Calibration of the PREP-S model in this cohort was less good than that achieved in the original PREP-S cohorts. Overall, approximately the same number of women had the outcome as predicted by the model (expected value 0; calculated value -0.13; not significantly different). However, calibration of the slope was 0.375 (expected value 1.0), suggesting that the difference between adverse outcome event rates between low and high risk groups was not as great as the PREP-S model suggested, with PREP-S consistently over-predicting the adverse event rate in the higher risk groups.

Recalibration of the model slightly improved the calibration of the PREP-S probabilities (Tables S10-S12), so that the notional probabilities were slightly closer to the actual event rate in the various subgroups.

Calibrations plots are shown for PREP-S for delivery within seven days (Figure S6, Table S10), two days (Figure S7 Table S11) and 14 days (Figure S8, Table S12), used as a prognostic model to determine time to delivery within a certain number of days as a binary outcome, not as a time-to-survival model. Evaluation of other thresholds for PIGF did not substantially improve test performance (Table S13) over and above the pre-specified thresholds.

## Discussion

### Main Findings

In this group of women with late preterm pre-eclampsia, PIGF and sFlt-1:PIGF measurement, and the PREP-S clinical prognostic model are not likely to add to the current clinical assessment to help plan care for these women around timing of delivery. Although PIGF testing had high sensitivity (97.9%) for delivery within seven days, the negative predictive value was only 71.4% and the specificity was low (8.4%). The areas under the curve for the clinical prediction model (PREP-S), PIGF and sFlt-1:PIGF in this cohort in determining need for delivery within seven days were all lower than 0.7, below the threshold deemed clinically useful.<sup>(10)</sup>

### Strengths and Limitations

The PEACOCK study was nested within a larger trial (PHOENIX) which evaluated timing of delivery in women with late preterm pre-eclampsia. We were necessarily constrained by the design of the PHOENIX trial such that we studied women of a different gestational age window (34 up to 37 weeks) compared to the original PREP study (up to 34 weeks' gestation), and we chose a different, binary outcome (clinically indicated need for delivery by seven days) and initial statistical analysis (presenting ROC areas). We originally chose measurement of PIGF concentrations as a potential predictor, based on our other work describing strong test performance of PIGF concentrations in women with suspected pre-eclampsia. However, from our study reported here, the distribution of PIGF concentrations in women with confirmed pre-eclampsia was clearly very different to those presenting with suspected disease, with a high proportion of women (over 90%) having low or very low PIGF values. Although sensitivity of the test remained high, specificity, predictive values and likelihood ratios were all sub-optimal, and the area under the curves for both PIGF and sFlt-1:PIGF in determining need for delivery in seven days was too low to be clinically useful.

PIGF and sFlt-1:PIGF biomarkers are considered reasonably 'upstream' in the pathophysiological process of the development of pre-eclampsia. The low overall prognostic performance in this group may be because the need for delivery from pre-eclampsia within seven days is associated with a variety of multi-organ, end-stage clinical parameters, and therefore an 'upstream' biomarkers such as PIGF or sFlt-1:PIGF are unable to discriminate which individuals are at higher risk. In addition, clinicians act upon early signs of impending clinical deterioration (such as abnormal liver transaminases) in order to avoid severe hepatic dysfunction (as used within the original PREP-S study, in women with pre-eclampsia prior to 34 weeks' gestation). Treatment paradox (e.g. decision for delivery based on early derangement of liver transaminases) could impact on the performance of prognostic markers or models, as women will have the primary outcome (clinically indicated need for delivery within seven days) without necessarily going on to develop severe maternal adverse outcomes. Although our chosen primary outcome (need for delivery for pre-eclampsia within seven days) acts as a surrogate to represent clinician concern of substantial fetal or maternal compromise, the suboptimal performance of PIGF and sFlt-1:PIGF for predicting delivery in this group may also reflect the complex, multi-pathological nature of this endpoint, and that a single biomarker is unable to determine both fetal and maternal compromise which has considerably different pathology (albeit the same clinical end point of early delivery). Whilst PIGF measurements have shown considerable potential as a diagnostic adjunct in women with suspected disease, (11) and the distribution of low and very low PIGF concentrations in the PEACOCK cohort confirms that we had participating women with placental dysfunction, this test does not appear to have strong prognostic value (for need for delivery) in this setting.

The PREP-S model was developed in early onset pre-eclampsia population (prior to 34 weeks), while the PEACOCK population was those with late preterm pre-eclampsia (34 to 37 weeks). The underlying contributions from maternal and placental pathophysiology may vary across these two groups, and hence the model cannot automatically be transported for use in the different population. Importantly, clinicians are likely to have lower threshold for delivery in women with late preterm pre-eclampsia than early onset pre-eclampsia since the risk of prematurity related complications is lower for births after than before 34 weeks' gestation. While the PREP-S model has consistently shown accurate performance both in the development dataset, and in two separate validation datasets of early onset pre-eclampsia,<sup>(4)</sup> we found that the model cannot be transported to a late preterm pre-eclampsia population to predict a different outcome.

At the time of conception of this study, there were a number of studies suggesting strong test performance for angiogenic factors measured in pregnancy, but the majority of the studies focused on women with suspected pre-eclampsia and the role of measurement in confirmed pre-eclampsia was under-explored. One early study by Verlohren and colleagues (12) assessed sFlt-1:PIGF in 95 women with pre-eclampsia after 34 weeks' gestation and compared duration of remaining pregnancy between women in the upper and lowest quartiles of sFlt-1:PIGF (but did not report other test performance statistics for this outcome). They reported that women with pre-eclampsia with a sFlt-1:PIGF in the upper quartile had a significantly reduced duration of pregnancy. However, a more recent study by Lou and colleagues (13) found that in women with pre-eclampsia after 34 weeks' gestation, there was no significant difference in sFlt-1:PIGF between those who delivered within seven days compared to those who delivered later. Meler and colleagues (14) similarly concluded that the predictive role of a low PIGF concentration in predicting maternal complications in early onset pre-eclampsia was limited because of both its low specificity and low positive predictive value.

### Interpretation

PIGF and sFlt-1:PIGF testing, and the PREP-S prediction model cannot be recommended to help plan care for late preterm pre-eclampsia regarding timing of delivery. A high proportion of women in this cohort already had low PIGF concentrations at the time of confirmed diagnosis, reducing the ability of PIGF measurement to further predict adverse outcomes. This is important and timely information given the current NHS-wide adoption of PIGF based testing as a diagnostic adjunct in the assessment of women with suspected pre-eclampsia, a different population to that studied here. Despite the confirmed diagnostic utility of PIGF in women with suspected pre-eclampsia, PIGF and sFlt-1:PIGF do not appear to have a role in assisting clinicians in determining timing of delivery in women with established preterm pre-eclampsia. The PREP-S model both alone, and in combination with PIGF, and in combination with sFlt-1:PIGF appears to have only limited clinical applicability for determining which women would require delivery in seven days from pre-eclampsia in women with late preterm pre-eclampsia.

### Conclusions

In women with late preterm pre-eclampsia, PIGF and sFlt-1:PIGF measurements are not likely to add to the current clinical assessment to help plan care regarding timing of delivery. Existing models developed in women with early onset pre-eclampsia to predict complications cannot be used to predict clinically indicated need for delivery in women with late preterm pre-eclampsia.

### Disclosure of Interests

The authors have no interests to disclose

### Contribution to Authorship

LCC and AS were involved in the study conception and in securing funding for the study. Clinical study data analysis was undertaken by KD, LCC and PS. The manuscript was written by KD and LCC, with assistance from PS, AP, JS, CG, AB, EH, AS, ST and AS. All authors approved the final version of the manuscript.

### Ethical Approval

The study was approved by South Central – Hampshire B Research Ethics Committee (13/SC/0645), March 2016.

#### Funding Source

The study was funded by the HTA Monitoring Add on Studies Programme (ID 15/59/06), National Institute for Health Research Professorship (Chappell RP-2014-05-019), and by the National Institute for Health Research Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. PTS is partly funded by Tommy’s and by CLAHRC South London (NIHR). The funders had no involvement in the study design, collection and analysis of data, data interpretation report writing or the decision to submit the article for publication.

#### References

1. 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 in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology*. 2018;51(6):743-50.
2. Chappell LC, Brocklehurst P, Green ME, Hunter R, Hardy P, Juszczak E, et al. Planned early delivery or expectant management for late preterm pre-eclampsia (PHOENIX): a randomised controlled trial. *Lancet (London, England)*. 2019;394(10204):1181-90.
3. National Institute for Health and Care Excellence. *Hypertension in Pregnancy: Diagnosis and Management*. 2019.
4. Thangaratinam S, Allotey J, Marlin N, Dodds J, Cheong-See F, von Dadelszen P, et al. Prediction of complications in early-onset pre-eclampsia (PREP): development and external multinational validation of prognostic models. *BMC medicine*. 2017;15(1):68.
5. von Dadelszen P, Payne B, Li J, Ansermino JM, Broughton Pipkin F, Cote AM, et al. Prediction of adverse maternal outcomes in pre-eclampsia: development and validation of the fullPIERS model. *Lancet (London, England)*. 2011;377(9761):219-27.
6. Tranquilli AL, Dekker G, Magee L, Roberts J, Sibai BM, Steyn W, et al. The classification, diagnosis and management of the hypertensive disorders of pregnancy: A revised statement from the ISSHP. *Pregnancy hypertension*. 2014;4(2):97-104.
7. Thangaratinam S, Allotey J, Marlin N, Mol BW, Von Dadelszen P, Ganzevoort W, et al. Development and validation of Prediction models for Risks of complications in Early-onset Pre-eclampsia (PREP): a prospective cohort study. *Health technology assessment (Winchester, England)*. 2017;21(18):1-100.
8. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Statistics in medicine*. 2016;35(2):214-26.
9. Chappell LC, Duckworth S, Seed PT, Griffin M, Myers J, Mackillop L, et al. Diagnostic accuracy of placental growth factor in women with suspected preeclampsia: a prospective multicenter study. *Circulation*. 2013;128(19):2121-31.
10. DW Hosmer SL. *Applied Logistic Regression*. 2nd Edition ed. New York, NY: John Wiley and Sons; 2000.
11. Duhig KE, Myers J, Seed PT, Sparkes J, Lowe J, Hunter RM, et al. Placental growth factor testing to assess women with suspected pre-eclampsia: a multicentre, pragmatic, stepped-wedge cluster-randomised controlled trial. *Lancet (London, England)*. 2019;393(10183):1807-18.
12. Verlohren S, Herraiz I, Lapaire O, Schlembach D, Moertl M, Zeisler H, et al. The sFlt-1/PIGF ratio in different types of hypertensive pregnancy disorders and its prognostic potential in preeclamptic patients. *American journal of obstetrics and gynecology*. 2012;206(1):58.e1-8.

13. Lou WZ, Jiang F, Hu J, Chen XX, Song YN, Zhou XY, et al. Maternal Serum Angiogenic Factor sFlt-1 to PlGF Ratio in Preeclampsia: A Useful Marker for Differential Diagnosis and Prognosis Evaluation in Chinese Women. *Disease markers*. 2019;2019:6270187.

14. Meler E, Scazzocchio E, Peguero A, Triunfo S, Gratacos E, Figueras F. Role of maternal plasma levels of placental growth factor for the prediction of maternal complications in preeclampsia according to the gestational age at onset. *Prenatal diagnosis*. 2014;34(7):706-10.

Table 1: Maternal demographics and baseline characteristics at enrolment (by expectant management groups)

Variable	Non-Randomised (Expectant Management) N= 182	Randomised (Expectant Management) N= 159	Comparison	All (Expectant Management) N= 341
Mean Maternal Age (years)	32.7 (5.3)	31.1 (6.1)	MD 0.8 (0.2 to 1.5)	31.9 (5.7)
Non-white ethnicity n (%)	60 (33.0%)	44 (27.8%)	RR 1.18 (0.85-1.64)	104 (30.6%)
Multiparous, n (%)	90 (49.5)	83 (52.2)	RR 0.95 (0.77 to 1.17)	173 (50.7)
Body mass index, Kg/m <sup>2</sup>	30.2 (6.5)	30.1 (7.9)	MD 0.1 (-0.7 to 0.8)	30.2 (7.2)
Maternal History of Pre-eclampsia n (%)	28 (31.1%)	29 (34.9%)	RR 0.84 (0.53 to 1.35)	57 (32.9%)
Chronic Hypertension n (%)	20 (11.0%)	25 (15.7%)	RR 0.70 (0.40 to 1.21)	45 (13.2%)
Chronic Kidney Disease n (%)	3 (1.6%)	2 (1.3%)	RR 1.31 (0.22 to 7.74)	5 (1.5%)
Aspirin use	75 (41.4%)	68 (42.8%)	RR 0.97 (0.75 to 1.24)	43 (42.1%)
Gestation at enrolment (weeks) Mean (SD)	35.4 (0.86)	35.5 (0.89)	MD -0.1 (-0.2 to 0.04)	35.4 (0.88)
Maternal BP 48 hours prior to enrolment (mmHg) (SD)	153 (15) 94 (10)	155 (16) 95 (11)	MD -0.8 (-2.4 to 0.8) MD -0.3 (-1.4 to 0.8)	154 (15) 94 (10)
Highest urinary PCR Number with measurement Mean (SD)	177 145 (238)	156 189 (337)	MD -22 (-54 to 10)	333 166 (289)
Suspected FGR on ultrasound	22 (12.1%)	25 (15.7%)	RR 0.77 (0.45 to 1.31)	47/341 (13.8%)

Variable	Non-Randomised (Expectant Management) N= 182	Randomised (Expectant Management) N= 159	Comparison	All (Expectant Management) N= 341
<b>Maternal PlGF (pg/mL) Number with measurement Mean (SD) Median (IQR)</b>	178 37.1 (134.1) 12.0 (12.0-18.6)	157 43.6 (146.7) 12.6 (12.0-24.1)	MD -3.3 (-18.5 to 11.9)	335 40.16 (140.00) 12.0 (12.0-20.6)
<b>PlGF [?] 100 pg/mL</b>	9 (5.1%)	7 (4.5%)	RR 1.60 (0.64 to 4.02)	18 (5.8%)
<b>PlGF 12-100 pg/mL</b>	67 (43.5%)	75 (47.8%)	RR 0.91 (0.71 to 1.16)	142 (45.7%)
<b>PlGF &lt;12 pg/mL</b>	76 (49.4%)	75 (47.8%)	RR 1.03 (0.82 to 1.30)	151 (48.6%)

SD: standard deviation. BP: blood pressure. PCR: protein:creatinine ratio. FGR: fetal growth restriction. PlGF: Placental growth factor. IQR: Interquartile Range. MD: Mean Difference. RR: Risk Ratio.

Table 2: Maternal and perinatal clinical characteristics at delivery (by expectant management groups).

Variable	Non-Randomised (Expectant Management) N= 182	Randomised (Expectant Management) N= 159	Comparison	All (Expectant Management) N= 341
<b>Gestation at delivery (weeks) Mean (SD)</b>	36.4 (1.05)	36.5 (1.00)	MD -0.02 (-0.1 to -0.1)	36.48 (1.03)
<b>Preterm delivery &lt;37 weeks</b>	101 (55.5%)	84 (52.8%)	RR 1.05 (0.86 to 1.28)	185 (54.3%)
<b>Delivery within 7 days</b>	108 (59.3%)	103 (64.8%)	-	211 (61.9%)
<b>Delivery within 2 days</b>	28 (15.4%)	29 (18.2%)	-	57 (16.7%)
<b>Delivery within 14 days</b>	158 (86.8%)	141 (88.7%)	-	299 (87.7%)
<b>Antenatal SBP &gt;160 mmHg</b>	106 (58.9%)	98 (61.6%)	-	204 (60.2%)
<b>Postpartum SBP [?]160 mmHg</b>	67 (37.9%)	61 (38.4%)	-	128 (38.1%)
<b>Onset of labour</b>	11 (6.0%) 110	7 (4.4%) 101	-	18 (5.3%) 211
Spontaneous	(60.4%) 61 (33.5%)	(63.5%) 50 (31.4%)		(61.9%) 111 (32.6%)
Induced Prelabour	0 (0.0%)	1 (0.6%)		1 (0.3%)
Caesarean section				
PROM and augmentation				

Variable	Non-Randomised (Expectant Management) N= 182	Randomised (Expectant Management) N= 159	Comparison	All (Expectant Management) N= 341
<b>Required indicated delivery</b>	95 (52.2%)	81 (50.9%)	-	176 (51.6%)
<b>Infant Birthweight (grams) Mean (SD)</b>	2489 (558)	2513 (520)	MD -12 (-70 to 45)	2500 (556)
<b>INTERGROWTH Centile Mean (SD)</b>	32.2 (31)	31.6 (23)	MD 0.3 (-2.8 to 3.4)	31.89 (29.61)
<b>INTERGROWTH SGA &lt;10<sup>th</sup> centile n (%)</b>	69 (36.5%)	50 (29.2%)	RR 1.18 (0.95 to 1.47)	119 (33.1%)
<b>INTERGROWTH SGA &lt;3rd centile n (%)</b>	30 (15.9%)	12 (7.0%)	RR 1.58 (1.12 to 2.24)	42 (11.7%)

SD: standard deviation. SBP: systolic blood pressure. SGA: Small for Gestational Age. MD: Mean Difference. RR: Risk Ratio.

Table 3 - ROC areas with standard errors (SE) and 95% confidence intervals (95% CI) for PREP-S, PlGF, sFlt-1 and sFlt-1:PlGF in determining delivery in 7 days singly, and in combination.

Delivery in 7 days	ROC Area (SE) (95% CI)	Harrells C (95% CI)	Comparison (vs. PREP-S alone)
PREP-S alone	0.64 (0.03) (0.58-0.71)	0.61 (0.57-0.64)	-
PlGF alone	0.60 (0.03) (0.54-0.66)	-	p=0.314
sFlt-1:PlGF	0.63 (0.03) (0.56 – 0.70)		p=0.893
PREP-S + PlGF	0.65 (0.03) (0.58-0.71)		p=0.776
PREP-S + sFlt-1:PlGF	0.64 (0.03) (0.58-0.71)		P=0.119

PREP: Prediction of complications in early-onset pre-eclampsia study. PlGF: Placental growth factor. sFlt-1: soluble fms-like tyrosine kinase-1

#### Figure Legends

Figure 1 - ROC areas for determining need for delivery within seven days. PlGF: Placental growth factor. sFlt-1: soluble fms-like tyrosine kinase-1. PREP: Prediction of complications in early-onset pre-eclampsia study.

#### Hosted file

Figure 1.docx available at <https://authorea.com/users/303667/articles/448755-prognostic-indicators-of-severe-disease-in-women-with-late-preterm-preeclampsia-to-guide-decision-making-on-timing-of-delivery-development-and-validation-of-prognostic-models>