Mini-combined test compared with NICE guidelines for early risk-assessment for pre-eclampsia: the SPREE diagnostic accuracy study

Liona C Poon,[1] David Wright,[2] Steve Thornton,[3] Ranjit Akolekar,[4] Peter Brocklehurst[5] and Kypros H Nicolaides[1*]

[1]Fetal Medicine Research Institute, King's College Hospital, King's College London, London, UK
[2]Institute of Health Services Research, University of Exeter, Exeter, UK
[3]Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[4]Department of Obstetrics and Gynaecology, Medway Maritime Hospital, Kent, UK
[5]Comprehensive Clinical Trials Unit, University College London, London, UK

*Corresponding author kypros@fetalmedicine.com

Declared competing interests of authors: Steve Thornton reports grants from the National Institute for Health Research (NIHR), the British Heart Foundation (BHF) and Barts Charity during the conduct of the study. Outside the submitted work, he was a member of the NIHR Efficacy and Mechanism Evaluation Editorial Board (2016–19), a trustee of Wellbeing of Women until 2019 (London, UK), MedCity (London, UK) and the William Harvey Research Institute (London, UK). He provides consultancy advice for commercial organisations [i.e. GlaxoSmithKline plc (Brentford, UK), Hologic, Inc. (Bedford, MA, USA), and Ferring Pharmaceuticals (Saint-Prex, Switzerland)]. He is a non-executive director of the NHS Trust (2016 to present). Ranjit Akolekar and Kypros H Nicolaides are trustees of the Fetal Medicine Foundation. Peter Brocklehurst reports grants and personal fees from the Medical Research Council (London, UK), grants from the NIHR Services and Delivery Research programme, the NIHR Health Technology Assessment (HTA) programme and the Wellcome Trust (London, UK) outside the submitted work. He was chairperson of the NIHR HTA Maternal, Neonatal and Child Health Panel and the Medical Research Council Methodology Research Programme Panel between 2014 and 2018.

Published November 2020
DOI: 10.3310/eme07080

Scientific summary

The SPREE diagnostic accuracy study

Efficacy and Mechanism Evaluation 2020; Vol. 7: No. 8
DOI: 10.3310/eme07080

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Pre-eclampsia, which affects 2–3% of pregnancies, is a major cause of maternal and perinatal morbidity and mortality. Pre-eclampsia can be subdivided into preterm pre-eclampsia, with delivery before 37 weeks' gestation, and term pre-eclampsia. Preterm pre-eclampsia is associated with a higher incidence of adverse short- and long-term outcomes. Recent evidence suggests that the risk of preterm pre-eclampsia can be substantially reduced by use of aspirin, given from 11–14 to 36 weeks' gestation in high-risk women. Identification of the high-risk group is based on maternal characteristics and medical history, as defined by the National Institute for Health and Care Excellence guidelines, but the performance of such an approach and the uptake of aspirin have not been evaluated by prospective studies. An alternative approach to screening for pre-eclampsia, which allows estimation of individual patient-specific risks of pre-eclampsia requiring delivery before a specified gestation, is to use a survival time model for gestational age at delivery with pre-eclampsia. Implementation using Bayes' theorem allows an a priori distribution of gestational age at delivery with pre-eclampsia, obtained from maternal characteristics and medical history, to be combined with the results of various biophysical and biochemical measurements at different stages in pregnancy (i.e. the competing risk model). Extensive research in the last decade has led to the identification of four potentially useful biomarkers at 11–13 weeks' gestation: (1) mean arterial pressure, (2) uterine artery pulsatility index, (3) serum pregnancy-associated plasma protein-A and (4) serum placental growth factor.

Objectives

The primary aim of the study was to compare the performance of first trimester screening for pre-eclampsia by the competing risk model with that of the current National Institute for Health and Care Excellence guidelines. The objectives were to:

1. finalise the algorithm used in the competing risk model
2. evaluate the performance of the new method compared with that recommended by the National Institute for Health and Care Excellence
3. provide the predictive performance of the new method
4. explore the possibility of carrying out first-stage screening in the whole population by some of the components of the complete test and reserving the rest for a subgroup of the population selected on the basis of the risk derived from first-stage screening
5. explore the potential value of another biomarker, inhibin A, in predicting pre-eclampsia
6. examine the performance of the new method in the prediction of small for gestational age neonates.

Methods

This was a prospective multicentre observational study in seven NHS maternity hospitals in England, between April and December 2016. Women aged > 18 years with a singleton pregnancy and a live fetus at the 11- to 13-week scan were included in the study. Women who were unconscious or severely ill at the time of recruitment, those with learning difficulties or serious mental illness and those with major fetal abnormality identified at the 11- to 13-week scan were excluded from the study.

Participants who provided written informed consent had recordings of maternal characteristics, medical history and measurements of mean arterial pressure, uterine artery pulsatility index, placental growth

factor and pregnancy-associated plasma protein-A at 11–13 weeks' gestation. The decision concerning administration of aspirin was made by the attending clinicians in accordance with routine standard of care at each site and the information was recorded in the research database both at the time of screening and during collection of data on pregnancy outcome.

Pre-eclampsia was defined according to the criteria of the International Society for the Study of Hypertension in Pregnancy or the American College of Obstetricians and Gynecologists. The performance of screening for pre-eclampsia by the competing risk model was compared with that of the National Institute for Health and Care Excellence method. Some patients received aspirin and in the comparison of sensitivities appropriate adjustments were made. Primary comparison was detection rate of National Institute for Health and Care Excellence method compared with a mini-combined test (including maternal factors, mean arterial pressure and pregnancy-associated plasma protein-A) in the prediction of all pre-eclampsia, for the same screen-positive rate determined by the National Institute for Health and Care Excellence method using McNemar's test. Key secondary comparisons were detection rate of risk assessment recommended by the National Institute for Health and Care Excellence guidelines compared with screening by the competing risk model with three combinations of markers [i.e. (1) maternal factors, mean arterial pressure and pregnancy-associated plasma protein-A, (2) maternal factors, mean arterial pressure and placental growth factor, and (3) maternal factors, mean arterial pressure, uterine artery pulsatility index and placental growth factor or triple test] in the prediction of preterm pre-eclampsia.

We proposed to recruit 16,850 women and on the assumption of a 5% no follow-up rate there would be 16,000 women for evaluation. On the extreme assumption that 90% of the National Institute for Health and Care Excellence screened-positive patients and 10% of the National Institute for Health and Care Excellence screened-negative patients would be treated with aspirin and that aspirin reduces the incidence of all pre-eclampsia by 50%, the power to detect a 10% difference in detection rate between the National Institute for Health and Care Excellence method and the mini-combined test in the prediction of all pre-eclampsia at the one-sided 2.5% level would be $> 80\%$.

We performed the following additional studies. First, we examined the calibration. Second, we evaluated a two-stage screening approach in the detection of preterm pre-eclampsia and early pre-eclampsia at an overall screen-positive rate of 10% and 20%, respectively, from a policy in which first-stage screening of the whole population is carried out by some of the components of the triple test and second-stage screening is carried out by the full triple test on women selected on the basis of results from first-stage screening. Third, we conducted a case–control study and a screening study to evaluate the potential value of inhibin A in improving the performance of screening provided by the other biomarkers. Last, we estimated the proportion of small for gestational age neonates born at ≥ 37, < 37 and < 32 weeks' gestation with a first-trimester combined risk for preterm pre-eclampsia (calculated by the competing risk model through the triple test) of > 1 in 100.

Results

Screening for pre-eclampsia was carried out in 17,051 women and outcome data were obtained from 16,747 women. Pre-eclampsia developed in 473 (2.8%) pregnancies, including 142 (0.8%) cases of preterm pre-eclampsia. Aspirin was taken by 400 (23.2%) women in the National Institute for Health and Care Excellence screen-positive group and 349 (2.3%) in the National Institute for Health and Care Excellence screen-negative group. The screen-positive rate by the National Institute for Health and Care Excellence method was 10.3% and the detection rate for all pre-eclampsia was 30.4% (95% confidence interval 26.3% to 34.6%). In screening by the competing risk model using maternal factors, mean arterial pressure and pregnancy-associated plasma protein-A, the detection rate of all pre-eclampsia was 42.5% (95% confidence interval 38.0% to 46.9%) and the difference in detection rate between the two methods was 12.1% (95% confidence interval 7.9% to 16.2%). After adjustment for the effect of aspirin

(i.e. a 30% reduction in rate of all pre-eclampsia) in those receiving this drug, the detection rate of the National Institute for Health and Care Excellence method was 31.5% (95% confidence interval 27.3% to 35.7%), that of the competing risk model was 42.8% (95% confidence interval 38.3% to 47.2%) and the difference between the two methods was 11.3% (95% confidence interval 7.1% to 15.5%). The detection rate of the National Institute for Health and Care Excellence method for preterm pre-eclampsia was 40.8% (95% confidence interval 32.8% to 48.9%), which was lower than that of the competing risk model using maternal factors, mean arterial pressure and pregnancy-associated plasma protein-A (53.5%, 95% confidence interval 45.3% to 61.7%), the competing risk model using maternal factors, mean arterial pressure and placental growth factor (69.0%, 95% confidence interval 61.4% to 76.6%) and the competing risk model using the triple test (82.4%, 95% confidence interval 76.1% to 88.7%). After adjustment for the effect of aspirin (i.e. a 60% reduction in rate of preterm pre-eclampsia) in those receiving this drug, the difference in detection rate of the three competing risk models from the National Institute for Health and Care Excellence method were 10.5% (95% confidence interval 2.3% to 18.8%), 24.0% (95% confidence interval 14.3% to 33.7%) and 35.1% (95% confidence interval 25.1% to 45.0%), respectively.

Calibration of risks for the incidence of pre-eclampsia was good and the calibration slope was very close to 1.0. In the two-stage screening a similar screen-positive rate and detection rate was achieved at substantially lower costs than with carrying out screening with all biomarkers in the whole population. If the method of first-stage screening is maternal factors, then measurement of biomarkers can be reserved for only 70% of the population, and if some of the biomarkers are included in first-stage screening then the need for the complete triple test can be reduced to 30–40% of the population. With regard to the inhibin A, although this biomarker improved the prediction of preterm pre-eclampsia provided by maternal factors alone (i.e. a detection rate of 49% vs. 60%), it did not improve the prediction provided by biomarkers that included placental growth factor. In relation to small for gestational age neonates, the competing risk model using the triple test at a risk cut-off point of 1 in 100 identified 46% of cases of preterm small for gestational age neonates and 56% of cases of early small for gestational age.

Conclusions

The screening programme for pre-eclampsia study has demonstrated that risk assessment for pre-eclampsia by current National Institute for Health and Care Excellence guidelines identifies approximately 30% of women who would develop pre-eclampsia and about 40% of those who will develop preterm pre-eclampsia, at a screen-positive rate of 10%. Compliance with the National Institute for Health and Care Excellence recommendation that women at high risk for pre-eclampsia should be treated with aspirin from the first trimester to the end of pregnancy was only 23%. Such low compliance may at least in part be attributed to the generally held belief, based on the results of a meta-analysis in 2007, that aspirin reduces the risk of pre-eclampsia by about 10% (Askie LM, Duley L, Henderson-Smart DJ, Stewart LA, PARIS Collaborative Group. Antiplatelet agents for prevention of pre-eclampsia: a meta-analysis of individual patient data. *Lancet* 2007;**369**:1791–8). The performance of screening by the competing risk model that combines maternal factors with biomarkers was superior to that of risk assessment by National Institute for Health and Care Excellence guidelines. At the same screen-positive rate as for the National Institute for Health and Care Excellence method, the detection rate for all pre-eclampsia in screening by maternal factors, mean arterial pressure and serum pregnancy-associated plasma protein-A was 42.5% and the detection rate for preterm pre-eclampsia by the triple test was 82.4%.

The strengths of the study are that (1) the study was a prospective examination of a large number of pregnant women in several maternity units covering a wide spectrum of demographic and racial characteristics; (2) > 90% of patients attending for routine care agreed to participate in the study, measurement of all biomarkers was recorded in all cases and complete follow-up was obtained from > 98% of patients; and (3) there was consistency in data collection through training of all investigators, regular University College London Comprehensive Clinical Trials Unit monitoring, and external validation

and quality assurance of biomarker measurements. A potential limitation of the study is lack of formal health economic assessment. This was beyond the scope of this study, but it is currently being carried out.

The performance of screening for preterm pre-eclampsia by the competing risk model utilising the triple test observed in this study is compatible with that reported in several previous studies of singleton pregnancies at 11–13 weeks' gestation. In four studies involving a combined total of 129,044 pregnancies, the detection rate of preterm pre-eclampsia was consistently approximately 75% at a screen-positive rate of 10% (Wright D, Tan MY, O'Gorman N, Poon LC, Syngelaki A, Wright A, Nicolaides KH. Predictive performance of the competing risk model in screening for preeclampsia. *Am J Obstet Gynecol* 2019;**220**:199.e1–199.e13). None of these studies found evidence that pregnancy-associated plasma protein-A improved screening achieved by the triple test.

Recent evidence suggests that first-trimester risk assessment should focus on prediction of preterm pre-eclampsia. Aspirin is considerably more effective than previously thought in reducing the risk of preterm pre-eclampsia, provided the daily dose of the drug is $\geq$ 100 mg and the gestational age at onset of therapy is < 16 weeks (Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet*. 2010;**376**:631–44). Against this background, there are ongoing debates about prediction and prevention of preterm pre-eclampsia centred on two questions: (1) whether or not aspirin should be recommended for all women or to a subpopulation of those women predicted to be at increased risk of developing pre-eclampsia and (2) if a strategy of prediction and prevention is to be used, what method should be used for prediction.

The arguments in favour of recommending aspirin to all women are that it avoids the need for prediction and the whole population benefits from the prophylactic treatment with aspirin. Arguments against this are that (1) compliance is likely to be worse when aspirin is applied to the whole population than when recommended to a subpopulation selected, and counselled, based on risk and (2) there is a need to balance the benefit from aspirin in prevention of preterm pre-eclampsia with potential harm due to haemorrhagic and other adverse effects. Assuming all women took aspirin, an incidence of preterm pre-eclampsia of 0.8% and a relative risk reduction of 60%, 208 women would be exposed to aspirin for every case of preterm pre-eclampsia prevented. Using risk stratification with maternal factors, mean arterial pressure, uterine artery pulsatility index and placental growth factor with the same screen-positive rate as National Institute for Health and Care Excellence, 16 women would be exposed to aspirin to prevent one case compared with 30 women using the National Institute for Health and Care Excellence guidelines.

Regarding the method of prediction, the debate centres around screening performance, costs and practical issues of implementation.

The main focus of this report has been on the detection rate achieved from the competing risk model compared with that of the National Institute for Health and Care Excellence method. For the same screen-positive rate as National Institute for Health and Care Excellence method, the detection rate for preterm pre-eclampsia achieved by combining maternal factors with mean arterial pressure, uterine artery pulsatility index and placental growth factor is 79.6% (95% confidence interval 72.7% to 86.5%) compared with 44.1% (95% confidence interval 35.7% to 52.6%) using the National Institute for Health and Care Excellence method. Using these estimates, with an incidence of preterm pre-eclampsia of 0.8%, the positive predictive values are 1 in 16 compared with 1 in 29 for the competing risk model and National Institute for Health and Care Excellence method, respectively. Among women who screen negative, the proportions with preterm pre-eclampsia (i.e. 1 – negative predictive value) are 1 in 550 and 1 in 200 for the competing risk model and National Institute for Health and Care Excellence method, respectively.

The main argument against the use of risk algorithms, such as the competing risk model, is that they are too complex to use in practice. Simple methods, such as the National Institute for Health and Care Excellence criteria or cut-off points applied to biomarker measurements or their ratios, should be preferred because they are easy to implement in practice. Indeed, the essential features of our approach of using Bayes' theorem to update likelihoods from biomarker multiple of the median values to update a prior based on maternal factors have been used for many decades in screening for aneuploidies. These algorithms have been built into commercial software used extensively in practice. The commercial software suppliers have implemented the competing risk model for pre-eclampsia screening into their software systems.

Regarding the approaches based on application of cut-off points to individual markers or ratios of different markers, the following points need to be considered. First, they do not provide individualised risks for decision-making. Second, their performance is inferior to approaches based on probability theory to make optimal use of the available information. Last, because biomarkers are affected by covariates such as ethnicity, they are likely to be inequitable in the way they perform across different groups within the population.

In clinical implementation of the competing risk model, recording maternal characteristics, measurement of blood pressure and hospital attendance at 11–13 weeks' gestation for ultrasound examination are an integral part of routine antenatal care. Measurement of uterine artery pulsatility index can be carried out by the same sonographers and machines used for the routine scan at 11–13 weeks' gestation; however, the sonographers will require training to carry out this test and the measurement would add 2–3 minutes to the current 20–30 minutes used for the scan. Serum placental growth factor can be measured in the same blood sample and by the same automated platforms that are currently used for measurement of pregnancy-associated plasma protein-A, as part of routine clinical practice in screening for trisomies in all maternity hospitals in England; however, there is an additional cost for the reagents.

In conclusion, the screening programme for pre-eclampsia study has demonstrated that the performance of first trimester screening for pre-eclampsia by a combination of maternal factors and biomarkers is superior to that achieved by the risk assessment method recommended by the current National Institute for Health and Care Excellence guidelines and it also predicts a high proportion of small for gestational age neonates.

Future research

Future research should focus on prospective evaluation of (1) the Bayes' theorem-based model in populations who are dissimilar to the research population described in this study; (2) the proposed two-stage screening approach and identification of potential biomarkers for further improvement of the competing risk model; and (3) implementation in different clinical set-ups and economic assessment of the new method of screening.

Trial registration

This trial is registered as ISRCTN83611527.

Funding

This project was funded by the Efficacy and Mechanism Evaluation programme, a Medical Research Council and National Institute for Health Research (NIHR) partnership. This will be published in full in *Efficacy and Mechanism Evaluation*; Vol. 7, No. 8. See the NIHR Journals Library website for further project information.

Efficacy and Mechanism Evaluation

ISSN 2050-4365 (Print)

ISSN 2050-4373 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full EME archive is freely available to view online at www.journalslibrary.nihr.ac.uk/eme. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Efficacy and Mechanism Evaluation* journal

Reports are published in *Efficacy and Mechanism Evaluation* (EME) if (1) they have resulted from work for the EME programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

EME programme

The Efficacy and Mechanism Evaluation (EME) programme funds ambitious studies evaluating interventions that have the potential to make a step-change in the promotion of health, treatment of disease and improvement of rehabilitation or long-term care. Within these studies, EME supports research to improve the understanding of the mechanisms of both diseases and treatments.

The programme supports translational research into a wide range of new or repurposed interventions. These may include diagnostic or prognostic tests and decision-making tools, therapeutics or psychological treatments, medical devices, and public health initiatives delivered in the NHS.

The EME programme supports clinical trials and studies with other robust designs, which test the efficacy of interventions, and which may use clinical or well-validated surrogate outcomes. It only supports studies in man and where there is adequate proof of concept. The programme encourages hypothesis-driven mechanistic studies, integrated within the efficacy study, that explore the mechanisms of action of the intervention or the disease, the cause of differing responses, or improve the understanding of adverse effects. It funds similar mechanistic studies linked to studies funded by any NIHR programme.

The EME programme is funded by the Medical Research Council (MRC) and the National Institute for Health Research (NIHR), with contributions from the Chief Scientist Office (CSO) in Scotland and National Institute for Social Care and Health Research (NISCHR) in Wales and the Health and Social Care Research and Development (HSC R&D), Public Health Agency in Northern Ireland.

This report

The research reported in this issue of the journal was funded by the EME programme as project number 14/01/02. The contractual start date was in February 2016. The final report began editorial review in September 2018 and was accepted for publication in June 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The EME editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the EME programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the EME programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).