Spot protein–creatinine ratio and spot albumin–creatinine ratio in the assessment of pre-eclampsia: a diagnostic accuracy study with decision-analytic model-based economic evaluation and acceptability analysis

Jason Waugh,[1]* Richard Hooper,[2] Edmund Lamb,[3] Stephen Robson,[4] Andrew Shennan,[5] Fiona Milne,[6] Christopher Price,[7] Shakila Thangaratinam,[2] Vladislav Berdunov[2] and Jenn Bingham[8]

[1]The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[2]Centre for Primary Care and Public Health, Queen Mary University of London, London, UK
[3]East Kent Hospitals University NHS Foundation Trust, Canterbury, UK
[4]Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
[5]Women's Academic Health Centre, King's College London, London, UK
[6]Lay representative
[7]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[8]Newcastle Clinical Trials Unit, Newcastle University, Newcastle upon Tyne, UK

*Corresponding author jason.waugh@nuth.nhs.uk

Declared competing interests of authors: Edward Lamb reports that he was a member of the development groups for both the National Institute for Health and Care Excellence chronic kidney disease and Kidney Disease Improving Global Outcomes clinical guidelines, which have considered the relative accuracies of proteinuria and albuminuria testing.

Published October 2017
DOI: 10.3310/hta21610

Scientific summary

SPCR and SACR in assessment of pre-eclampsia

Health Technology Assessment 2017; Vol. 21: No. 61
DOI: 10.3310/hta21610

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

Scientific summary

Background

Pre-eclampsia (PE) is a multisystem disorder of pregnancy associated with new raised blood pressure (BP) and proteinuria. It remains the second leading cause of direct maternal deaths in the UK and causes 20% of all stillbirths. The 2010 National Institute for Health and Care Excellence (NICE) guideline for the management of hypertensive disorders of pregnancy differentiates PE from gestational hypertension by the presence of new significant proteinuria. This differentiation is critical in terms of monitoring and management, cost and outcomes. NICE has acknowledged the lack of evidence relating both to the diagnosis of significant proteinuria in pregnant women and the prognostic value of various urinary protein thresholds. The aim of the study was to evaluate the various methods of measuring proteinuria in pregnant women with suspected PE both in the diagnosis of significant proteinuria and in the prediction of clinically significant outcomes and cost-effectiveness.

Objectives

The primary objective of the study was to evaluate the accuracy of quantitative assessments of spot protein–creatinine ratio (SPCR) and spot albumin–creatinine ratio (SACR) at different thresholds in predicting severe PE compared with 24-hour urine protein measurement in pregnant women with hypertension and suspected proteinuria.

The secondary objectives, after amendment, were to:

- investigate differences between laboratory assay methods for 24-hour proteinuria estimation
- evaluate the accuracy of quantitative assessments of SPCR and SACR at different thresholds in predicting adverse perinatal outcomes
- estimate the diagnostic utility of SPCR or SACR tests as a potential replacement for 24-hour protein estimation by developing a decision-analytic model
- assess the cost-effectiveness using incremental cost per quality-adjusted life-year (QALY) gained by the mother and baby in each test compared with standard practice.

Methods

Study conduct

As part of routine antenatal care to assess for PE, pregnant women have their BP measured and a urine sample checked for protein using a urine dipstick. Potential participants were identified from women with suspected PE, who were subsequently referred to a hospital maternity assessment unit, delivery suite or outpatient department for repeat assessments.

The inclusion criteria for the study were pregnant women aged $\geq$ 16 years, who were at > 20 weeks' gestation with new hypertension (systolic BP of $\geq$ 140 mmHg and/or diastolic BP of $\geq$ 90 mmHg) and a trace or more proteinuria on an automated dipstick urinalysis, and who were able to give written informed consent.

Women with pre-existing renal disease (proteinuria before 20 weeks), pre-gestational diabetes or chronic hypertension were excluded.

All eligible women were given a patient information sheet and invited to participate by a research midwife.

Eligible women who consented to take part in the study had provided a routine spot urine sample for protein analysis (the recruitment sample). This sample of urine was used to provide five 1-ml aliquots, which were frozen and stored for secondary analysis. The remainder was sent to a local laboratory for quantitative assessment of SPCR.

Participants were asked to collect urine for 24 hours in a collection container, as either an inpatient or an outpatient depending on their clinical management plan. An aliquot of the 24-hour sample was frozen and stored for secondary analysis.

A further spot sample of urine was taken and stored for analysis immediately before delivery.

Six weeks after giving birth, women were followed up and the pregnancy diagnosis was confirmed. Outcome data were collected from hospital records.

Index tests

There were four index tests on a spot sample of urine: (1) SPCR test (conducted at the local laboratory), (2) SPCR test [conducted at the central laboratory using the benzethonium chloride (BZC) assay], (3) SPCR test [conducted at the central laboratory using the pyrogallol red (PGR) assay] and (4) SACR test (conducted at the central laboratory using an automated chemistry analyser). The comparator tests on 24-hour urine collection were a central test using the BZC assay and a central test using the PGR assay.

Reference standards

The primary reference standard was the NICE definition of severe PE: PE with severe hypertension, symptoms, biochemical and/or haematological impairment.

- PE is defined as new hypertension after 20 weeks' gestation (systolic BP of $\geq$ 140 mmHg and/or diastolic BP of $\geq$ 90 mmHg) and significant proteinuria ($\geq$ 300 mg from 24-hour urine collection using the central laboratory's BZC assay).
- Severe hypertension is defined as systolic BP of $\geq$ 160 mmHg or diastolic BP of $\geq$ 110 mmHg.
- Severe features included at least one of severe headache, visual disturbances, problems with vision, severe pain just below the ribs or vomiting, papilloedema, signs of clonus (three or more beats), liver tenderness, HELLP (Haemolysis, Elevated Liver enzymes, Low Platelets) syndrome, platelet levels below 100×10^9/l, abnormal liver enzyme levels (alanine aminotransferase or aspartate aminotransferase levels of > 70 U/l).

A second reference standard was a clinician diagnosis of severe PE, which was defined as those instances in which a woman was treated with magnesium sulphate or put on a severe PE protocol. The use of such a protocol is a core compliance standard for clinical risk assessment followed by all participating units.

Adverse perinatal outcome (composite identified by Delphi survey of clinicians) was defined as one or more of perinatal or infant mortality, bronchopulmonary dysplasia, necrotising enterocolitis or grade III/IV intraventricular haemorrhage. The definition used for bronchopulmonary dysplasia was oxygen dependence at 36 weeks of postmenstrual age (gestational age at delivery plus chronological age of baby).

For the economic analysis, the diagnostic accuracy of each test, derived from the statistical analysis of the clinical study, was combined with the intervention cost and length of stay in maternal and neonatal units obtained from study sites and, cost and utility assumptions in the hypertension in pregnancy models published in NICE clinical guideline (CG) number 107 (NICE. *Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy.* CG107. London: NICE; 2010). The study estimated the incremental cost per QALY in deterministic and probabilistic sensitivity analyses.

Analysis

Diagnostic performance of the index tests against the reference standards were compared using prespecified thresholds of 30 mg/mmol for SPCR and 2 mg/mmol for SACR. Sensitivities, specificities, positive likelihood ratios and negative likelihood ratios (LR–s) and area under receiver operating characteristic (ROC) curves were calculated. For the economic evaluation, a modelling approach was used to estimate the incremental cost per quality-adjusted life-year (QALY) gained of the test, comparing comparator and other index tests against the local laboratory's SPCR test (standard care comparator). The analysis was based on the clinician diagnosis of severe PE as the reference standard.

A rule-out test for severe PE would require a LR– of ≤ 0.1 and a sensitivity of $\geq 90\%$ with a high specificity.

Sample size

To demonstrate that a quantitative assessment of SPCR or SACR at a given cut-off point could rule out severe PE, the original sample size required 240 women with severe PE. Assuming that 5% of recruited women would develop severe PE, the recruitment target was set at 3000 women with new hypertension and suspected proteinuria.

Interim reports indicated that the prevalence of severe PE might be higher; the Trial Steering Committee recommended that the proportion of women with severe PE be estimated from the first 500 participants. The laboratory assessments (NICE definition) were not available for all participants, so a surrogate definition was agreed (the clinician diagnosis of severe PE). Using this, the prevalence was estimated to be 78 out of 500 (15.5%) and a new recruitment target was calculated of 1790 women.

Results

Population

In total, 1823 women were recruited and, of these, 959 had no missing test data for all four test index assays and were available for primary analysis. There were 475 women with PE (NICE definition), of which 417 out of 475 (88%) had severe PE.

Using the NICE definition of severe PE, the prevalence in this group was 43% (417/959), the majority of whom (339/417) had severe hypertension. In total, 23% of women delivered prior to 36 completed weeks of gestation (preterm) and 6% of babies had severe perinatal morbidity. There were no maternal deaths in this group, but there were still 30 cases of eclampsia.

Using the surrogate clinician diagnosis of severe PE, the prevalence rate was 20% (193/959) of women, of whom the majority (149/193) both followed a severe PE protocol and received magnesium sulphate. Only 32% of those with severe PE (NICE definition) had a clinician diagnosis of severe PE, and 8% of those with a NICE definition of severe PE and 15% with a clinical definition had an adverse perinatal outcome.

Primary objective analysis

The diagnostic performance of the four assays testing the spot urine sample at recruitment against the reference standard of severe PE (NICE definition) was similar. The three SPCR tests had sensitivities in excess of 90% at prespecified thresholds, with poor specificities and LR–s of ≥ 0.1. The central laboratory's SACR test had a significantly higher sensitivity of 99% [95% confidence interval (CI) 98% to 100%] and lower specificity. The ROC curves, when overlain, were similar (area under ROC curve between 0.87 and 0.89); the area under the central laboratory's SACR curve was significantly greater than the local laboratory's SPCR curve ($p = 0.004$).

Using the severe PE (the clinician diagnosis) as reference standard, the three SPCR tests had sensitivities below 90% at the prespecified thresholds. The central laboratory's SACR test had significantly higher sensitivity (97%, 95% CI 93% to 99%) than the local laboratory's SPCR test, but with a significantly lower

specificity (16%, 95% CI 14% to 19%). When overlain, the ROC curves for the assays were similar, and demonstrated a much poorer diagnostic performance than with the NICE definition of severe PE. The negative predictive values of each of the three SPCR tests was 92–93% (98% for the SACR test), so a negative result from one of these tests brings the risk of developing severe PE down to 7–8% (2% for the SACR test), compared with the pre-test risk of 43%.

Comparing the central laboratory's PGR and BZC assay methods to diagnose significant proteinuria from the 24-hour urine sample, 55% and 50%, respectively, of samples were identified as having significant proteinuria (≥ 300 mg/24 hours). Consequently, the proportion of women categorised as having severe PE (NICE definition) was greater using the PGR assay (48%) than using the BZC assay (43%). Using the data from the PGR assay in the reference standard, the diagnostic performance of the index tests was similar: the three SPCR assays had sensitivities above 90% with low specificities. LR–s were slightly higher than those found using the BZC assay in the proteinuria component of the definition.

There is a non-linear relationship when comparing the two assays using a Bland–Altman plot. Up to values of around 1200 mg/l the PGR assay value is typically higher, whereas from 1200 to 2000 mg/l the PGR assay value is typically lower.

In terms of adverse perinatal outcomes, the diagnostic performance of the four assays testing the spot urine sample at recruitment against the reference standard of adverse perinatal outcomes was similar. The three SPCR tests had sensitivities below 80% at prespecified thresholds, with poor specificities. The central laboratory's SACR test had a significantly higher sensitivity of 94% (95% CI 84% to 98%) and lower specificity of 14% (95% CI 12% to 16%), with poor likelihood ratios. The ROC curves, when overlain, were similar, and all demonstrated poor diagnostic performance.

Economic evaluation

The standard care comparator (local SPCR test) generated 52.39 maternal and neonatal QALYs, at a lifetime cost of £6621. Across the six tests, the differences in cost and QALYs were small; the incremental cost-effectiveness analysis found that there was a difference of only 0.04 QALYs between the best and worst performing test. Cost differences were greater, with urine dipstick testing alone generating £66 more than the local laboratory's SPCR test. The central laboratory's SACR test was the most effective strategy (52.42 QALYs gained).

Conclusions

From the results of this study we can draw the following conclusions:

1. The SACR test has marginally better diagnostic performance characteristics than the SPCR test when predicting severe PE according to the NICE CG107 definition.
2. All four tests could be used as rule-out tests for the NICE definition of severe PE, in that if we get a negative test result the odds of severe PE are considerably reduced (by 10-fold or more) compared with the odds in the target population as a whole before we do the test.
3. The collection of a 24-hour urine sample confers no additional value over a spot urine sample to quantify proteinuria in women with hypertension in pregnancy.
4. The threshold level of SPCR that performs best for the prediction of severe PE is the current threshold of 30 mg/mmol.
5. The threshold level of SACR that is equivalent to this in terms of clinical performance is 8 mg/mmol.
6. Urine dipstick testing without laboratory testing of SPCR or SACR performs poorly as a predictive strategy for either severe PE or adverse perinatal outcome.
7. The measurement of 'maximum proteinuria' or a 'rise in proteinuria' confers no advantage in the prediction of severe PE or adverse perinatal outcomes in hypertensive pregnancies.

8. Biochemical assays for proteinuria are not universal and as such some assays (e.g. PGR assay) will overdiagnose PE compared with the BZC assay by at least 5%.
9. The SACR test was deemed to have a 100% probability of being the most cost-effective option at the standard willingness-to-pay threshold of £20,000–30,000 per QALY recommended by NICE.
10. Clinicians continue to plan care (and interpret results) in an inconsistent manner when caring for women with severe PE.
11. The non-uniform application of NICE CG107 severe PE management algorithms and definitions by clinicians does not confer any additional benefits in reducing the number of women who will have adverse perinatal outcome.

Implications for clinical practice

- Evidence would suggest that all proteinuria assessments should be performed by an initial dipstick screening test read on an automated dipstick reader, and for all non-negative tests a SACR test should be considered. In the absence of a SACR test being available, a SPCR test could be used as an alternative.
- Given the intermethod/laboratory variability of protein assays demonstrated, all 'protein' tests should be viewed with extreme caution and consideration should be given to a urine albumin assay being employed for future definitions.
- Clinically significant proteinuria should remain defined at a level of 8 mg/mmol as measured by the SACR test or 30 mg/mmol as measured by the SPCR test.
- The evidence from this clinical study does not support the recommendation of 24-hour urine sample collection in hypertensive pregnant women.
- Once confirmed to be > 30 mg/mmol, no further proteinuria measurements are required during hypertensive pregnancy.
- Whenever possible, proteinuria measurements should be by a single (accepted as standard) assay. When this is not a BZC assay, clinicians should be aware that some assays can overdiagnose proteinuria and, hence, PE.
- In the presence of hypertension in pregnancy without proteinuria, the progression to severe maternal disease is unlikely (7% for SPCR and 2% for SACR). Adverse maternal and perinatal outcomes do occur and all women with gestational hypertension should be closely monitored.

Trial registration

This trial is registered as ISRCTN82607486.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

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 HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. 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 *Health Technology Assessment* journal

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

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/65/02. The contractual start date was in September 2012. The draft report began editorial review in September 2016 and was accepted for publication in March 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). 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, NETSCC, the HTA programme or the Department of Health. 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 HTA programme or the Department of Health.

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