Accuracy of urinalysis dipstick techniques in predicting significant proteinuria in pregnancy

Waugh J J, Clark T J, Divakaran T G, Khan K S, Kilby M D

CRD summary
This review assessed the accuracy of urine dipsticks in predicting significant proteinuria in pregnancy. The authors concluded that the accuracy of dipstick urinalysis is poor and of limited usefulness in ruling in or ruling out pre-eclampsia. The conclusions are supported by the results presented, but should be interpreted with some caution given the differing estimates from the individual studies.

Authors' objectives
To estimate the accuracy of dipstick urinalysis in predicting significant proteinuria in pregnancy.

Searching
MEDLINE and EMBASE were searched from January 1970 to February 2002; the search terms were reported. In addition, the Cochrane Library and relevant specialist registers of the Cochrane Collaboration were reviewed. Recent issues of three obstetric journals were handsearched and the reference lists of all known primary and review articles were screened. Individual experts or groups with an interest in the field and manufacturers of point-of-care urinalysis equipment were contacted for unpublished studies.

Study selection
Study designs of evaluations included in the review
Prospective observational studies or comparative cross-sectional studies in which the results of the index test were compared with a reference standard were eligible for inclusion.

Specific interventions included in the review
Studies of point-of-care tests for urine protein were eligible for inclusion. The tests evaluated in the included studies were colour-change dipsticks and automated dipstick urinalysis.

Reference standard test against which the new test was compared
Studies that included a laboratory assay for urine protein, preferably from a 24-hour sample, as the reference standard were eligible for inclusion. The threshold used to define a positive test result was 300 mg/24 hours or 300 mg/L, the level used in definitions of pre-eclampsia.

Participants included in the review
Studies of pregnant women were eligible for inclusion. These included women with uncomplicated pregnancies, women with hypertension, and pregnancies complicated by co-existing renal disease.

Outcomes assessed in the review
No inclusion criteria in relation to the outcomes were specified, although studies that did not include sufficient data to extract a 2x2 table of test performance were excluded.

How were decisions on the relevance of primary studies made?
Two reviewers independently assessed studies for relevance by scanning titles and abstracts identified through the searches. Full text articles of studies considered potentially relevant were obtained and assessed for inclusion with reference to a checklist. Any disagreements were resolved through consensus or discussion with a third reviewer.

Assessment of study quality
The studies were assessed for methodological quality based on items relating to population (prospective recruitment, population of pregnant women with proteinuria), test (adequate description of the index test, person interpreting
reference standard blinded to the results of the index test) and reference standard (proteinuria estimation from 24-hour urine sample). Two reviewers independently assessed the studies for methodological quality. Any disagreements were resolved through consensus.

**Data extraction**
The data were extracted as 2x2 tables comparing the results from the index test to the reference standard. These data were used to calculate the sensitivity, specificity, positive and negative likelihood ratios (LRs), and pre-and post-test probabilities of disease for each study. A correction factor of 0.5 was used when the data or a study included a zero value, to allow calculation of the LR. Two independent reviewers extracted the data.

**Methods of synthesis**
How were the studies combined?
The LRs were pooled by weighting the log LR for each study by the inverse proportion of its variance. Pooled estimates were used to generate post-test probabilities. Funnel plots were produced using log diagnostic odds ratios to investigate the possibility of publication bias. Estimates of sensitivity were plotted against ’1 minus the specificity’ and a summary receiver operating characteristic (ROC) curve was produced (not shown in the article).

How were differences between studies investigated?
The authors did not report a method for investigating differences between the studies. They stated that a meta-analysis was not possible for all studies because of variation in the study methods and differences in the diagnostic test cut-off levels. They also reported the results of a test for heterogeneity but provided no information on the test used.

**Results of the review**
Seven studies (1,841 women) were included in the review.

Six studies were prospective in design, four of which included consecutive patient samples. Four studies reported blinding of the test results. Follow-up was greater than 90% in all studies. There was significant heterogeneity across the studies (P<0.001). A funnel plot showed no significant heterogeneity.

Visual dipstick urinalysis was evaluated in 6 studies. The positive LRs ranged from 0.54 to 50.12 and the negative LRs from 0.45 to 5.49. For predicting 300 mg/24-hour proteinuria at the 1+ or greater threshold, the pooled positive LR was 3.48 (95% confidence interval, CI: 1.66, 7.27) and the pooled negative LR was 0.6 (95% CI: 0.45, 0.80).

Automated urinalysis (1+, 2+ and 3+ thresholds), other visual dipstick thresholds (2+ or 3+), and a reference standard based on 300 mg/L proteinuria were each evaluated in one or two studies. The results for these studies, which were given in the article, generally showed poor accuracy of the dipstick tests.

A univariate subgroup analysis stratified for items of study quality did not provide an explanation for the observed variation in diagnostic performance.

**Authors’ conclusions**
The accuracy of dipstick urinalysis with a 1+ threshold in the prediction of significant proteinuria is poor and, therefore, of limited usefulness to the clinician. Accuracy may be improved at higher thresholds (greater than 1+ proteinuria), but available data are sparse and of poor methodological quality.

**CRD commentary**
This was a reasonable review of the area. It addressed a focused review question supported by clearly defined inclusion criteria. A thorough literature search, which included attempts to locate unpublished studies, was conducted. Details of the review process were reported, and these showed that appropriate steps were taken to minimise bias. A formal quality assessment was undertaken and the results included in the synthesis of the results. The description of the
methods of analysis was rather limited and the methods used to obtain some of the reported results were not described: e.g. methods used to pool studies, sensitivity analyses undertaken and tests for heterogeneity. The authors stated that an ROC plot with a summary ROC curve was produced, but this was not presented in the paper. Given the extreme heterogeneity in the results of the individual studies, such a curve would have been very helpful in giving a visual indication of this heterogeneity. Also, the heterogeneity in the study results means that it is questionable whether pooling was appropriate; further investigation into possible explanations for the observed heterogeneity would have been helpful. The authors’ conclusions are supported by the pooled results presented, but should be interpreted with caution given the heterogeneity between the studies.

Implications of the review for practice and research
Practice: The authors stated that, with point-of-care urine dipstick analysis, significant proteinuria cannot be accurately detected or excluded with a 1+ threshold and is not recommended for diagnosing pre-eclampsia.

Research: The authors stated that there is an urgent need for research in this area of common obstetric practice.

Bibliographic details

PubMedID
15051572

DOI
10.1097/01.AOG.0000118311.18958.63

Other publications of related interest
This additional published commentary may also be of interest. Walerstein S. Review: point of care dipstick analysis has low accuracy for detecting proteinuria in pregnancy. Evid Based Med 2005;10:25.

Indexing Status
Subject indexing assigned by NLM

MeSH
Female; Humans; Predictive Value of Tests; Pregnancy; Pregnancy Complications /diagnosis; Proteinuria /diagnosis; Reproducibility of Results; Urinalysis

AccessionNumber
12005008167

Date bibliographic record published
30/09/2006

Date abstract record published
30/09/2006

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.