Accuracy of on-site tests to detect asymptomatic bacteriuria in pregnancy: a systematic review and meta-analysis

Ewelina Rogozińska MSc,[1,2] Sandra Formina MD,[3] Javier Zamora Ph.D.,[1,2,4] Luciano Mignini MD, [3] Khalid S. Khan MSc,[1,2]

Affiliations

1. Women's Health Research Unit, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
2. Multidisciplinary Evidence Synthesis Hub (mEsh), Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
3. Centro Rosarino de Estudios Perinatales, Rosario, Argentina
4. Clinical Biostatistics Unit, Hospital Ramon y Cajal (IRYCIS) and CIBER Epidemiology and Public Health, Madrid, Spain

Corresponding author: Ewelina Rogozinska, Women's Health Research Unit, Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, tel: +44 20 7882 5881, e-mail: e.a.rogozinska@qmul.ac.uk

Funding The World Health Organization and the European Union funding to the EBM Connect Collaboration through Framework Programme, Marie Curie Actions, International Staff Exchange Scheme (Proposal no. 101377; Grant Agreement no. 247613).

Short Title: On-site tests for asymptomatic bacteriuria

Acknowledgments: The authors would like to acknowledge the assistance of the following advisors from the WHO Department of Reproductive Health and Research (A. Metin Gülmezoglu and Özge Tunçalp). We are thankful Dr. Caoimhe NicFhogartaigh, Microbiology Consultant at The Barts Health NHS Trust in London for her microbiological advice.

Conflict of interest: All authors have completed the Unified Competing Interest form an (available on request from the corresponding author) and declared that we have no conflicts of

interest. The views expressed are solely those of the authors and do not necessarily reflect the decisions or stated policy of the World Health Organization.

Precis: Sensitivity of the on-site tests to detect asymptomatic bacteriuria in pregnancy varies, however, their specificity to rule in the infection is high.

Abstract

Objective: The main objective of this systematic review of the literature was to determine the accuracy of on-site tests that require fewer resources to detect asymptomatic bacteriuria among pregnant women.

Data source: We searched Medline, Embase, Web of Science, Scopus, and Latin-American Literature (LILACS) from inception until June 2015 without language restrictions.

Methods of Study Selection: Two independent reviewers selected studies that recruited asymptomatic pregnant women to evaluate the accuracy of on-site tests in detecting the presence of bacteria in the urine using urine culture as a reference standard.

Tabulation, Integration, and Results: Women's characteristics, study design, urine sample collection and handling were extracted along with the test accuracy data. Where possible, we pooled the data using a bivariate, hierarchical random effects model. Of 1,360 screened references, 27 papers (13,641 women) with test accuracy data on nine tests met the inclusion criteria. The most commonly evaluated test was urine dipstick. The pooled sensitivity and specificity of nitrites detected by dipstick to detect asymptomatic bacteriuria were 0.55 (95% CI 0.42 to 0.67) and 0.99 (95% CI 0.98 to 0.99), respectively. Griess test to detect nitrites had a sensitivity of 0.65 (95% CI 0.50 to 0.78) and specificity of 0.99 (95% CI 0.98 to 1·00). Dipslide with gram staining had a pooled sensitivity of 0.86 (95% CI 0.80 to 0.91) and specificity of 0.97 (95% CI 0.93 to 0.99).

Conclusions: The sensitivity of evaluated on-site tests to exclude bacterial urinary infection varies, however, their specificity to rule in disease is high.

Registration number: PROSPERO No. CRD42015027905

Keywords: test accuracy, asymptomatic bacteriuria, pregnancy, on-site test

Introduction

Asymptomatic bacteriuria, a common urinary tract infection, varies in prevalence by factors such as age, gender, or level of sexual activity. The prevalence of the infection in pregnancy ranges from 2–15% of whom 20–40% progress to symptomatic urinary infections (UTI).[1] Pregnant women with undetected asymptomatic bacteriuria are more likely to deliver prematurely[2] or low-birth-weight infants, and have a 20- 30-fold increased risk of developing pyelonephritis compared with those without the infection.[3]

Although some bodies recommend a routine urine culture screening in early pregnancy[4, 5], it is an expensive, cumbersome, and time-consuming test (taking 24 to 48 hours to obtain results) that requires access to laboratory facilities. There is a wide range of tests requiring fewer resources and minimal training,[6] of which the most commonly used to detect the presence of bacteria instantly in the urine is a dipstick. Available evidence synthesis on their accuracy in pregnancy is limited in range of evaluated test[7], and methodological strength.[6, 8]

We bridge the above gap through a systematic and comprehensive evaluation of a wide range of on-site tests used to detect bacteriuria compared against urine culture as a reference standard in asymptomatic pregnant women taking into account potential sources of heterogeneity.

Methods

The review was conducted prospectively guided by a pre-defined protocol (PROSPERO No. CRD42015027905). We followed current standards of evidence synthesis for test accuracy[9-11] and reported findings in compliance with guidelines.[12]

Sources

We searched major databases such as Medline, Embase, Web of Science, Scopus, and a specialized database of Latin-American literature (LILACS) for studies published from database inception to August 2014, with no language restrictions. The search was updated to June 2015 and was supplemented by a hand search of the references from the included publications. The search strategy combined terms such as: 'Pregnancy', 'Antenatal', 'Gestation', 'asymptomatic bacteriuria' and 'Urinary Tract Infections' and applied a filter for test accuracy studies (for details see Appendix 1).[13] The ClinicalTrials.gov register database was screened to identify any recently completed studies.

Study selection

Two independent reviewers (ER and SF) screened references and full-text of previously selected articles. The consensus on the eligibility of evaluated publications was reached through discussion, or consultation with a third reviewer (KSK). We looked for studies reporting the accuracy of any on-site tests to detect asymptomatic bacteriuria among pregnant women without symptoms of urinary tract infections or not on antibiotic treatment. The reference standard had to be a urine culture, and asymptomatic bacteriuria had to be defined as equal, or more than 10^5 Colony Forming Units of a single organism per mL of urine.[8] Test accuracy had to be reported in a way allowing construction of 2 x 2 tables. We excluded studies with a case-control design and were reference standard was not reported or used a different definition of bacteriuria than specified above as this design and variation in reference standard were associated with bias.[14]

Data were extracted independently by ER and SF on to a piloted sheet. We collected authors' details, year of publication, country, women's characteristics, gestational age at testing; urine collection method, storage, and handling. The data were tabulated, crossed checked and in the

case of discrepancies discussed between the reviewers. The studies were grouped according to country income (low-, low-middle, upper-middle) using the World Bank classification.[15] The risk of bias and applicability of included studies were assessed by two independent reviewers (ER and SF) using the QUADAS-2 tool[10] tailored for this review. Study quality was assessed for selection of participants, implementation of the index test and the reference standard, and patient flow. Studies with low risk of bias used a suitable spectrum of participants, recruited in consecutive or random manner; all participants were tested using the same reference standard, and the majority of the study population was included in analyses. Any disagreements over quality assessment were resolved by a third reviewer (KSK). We did not assess publication bias due to limitations of available methods.[16, 17]

We calculated test accuracy estimates (sensitivity, specificity, and likelihood ratios for positive and negative test result) with 95% confidence intervals (CIs). Heterogeneity was investigated visually on forest plots with sensitivity and specificity estimates (with 95% confidence intervals) for individual studies. The impact of quality of study design, the reliability of population description, and sample collection and storage was explored through sensitivity analyses. All analyses were conducted using STATA version 12.1.[18] If less than required number of data points was available, we pooled accuracy of sensitivity and specificity, and likelihood ratios using univariate model using *metaprop* and *metan* commands, respectively. Where a higher number of studies was available, we pooled the accuracy parameters using bivariate, random effects model as implemented in *metandi*[19] and *midas*[20] commands . Posttest probabilities were calculated using following formula: $O = p1/(1 - p1)$, $p2 = O * L$, $p = p2/(1 + p2)$, where p1 pretest probability, O pretest odds, p2 posttest odds, L likelihood ratio, p posttest probability.[21]

Results

Out of the 1,360 references, 39 examining 27 types of index tests appeared initially to meet the inclusion criteria (Figure 1). After exclusion of tests not suitable for use in the asymptomatic population, we were left with 27 studies with nine index tests. List of all identified tests and reasons for study exclusion can be found in Appendix 2. Selected tests were: dipstick with only nitrites marker as positive, dipstick with nitrites or leucocytes as positive, urine analysis with bacteria count, dipslide with gram stain, Uricult (Orion Diagnostica, Espoo, Finland), Microstix-3 (Bayer Schering Pharma, Berlin, Germany), Griess test to detect nitrites, chlorhexidine reaction, and uriscreen catalase test. Reference of the included studies can be accessed in Appendix 3.

The majority of identified studies were conducted in low-middle (11 studies) or upper-middle (five studies) income countries; ten in high-income countries and only one in a low-income country. The studies were published between 1981 and 2015; ten studies were published before the year 2000, nine between 2000 and 2010 and remaining eight in the last five years. The majority (19/27) of included studies contributed to evidence synthesis accuracy data of only one test (Table 1) with urine dipstick as the most commonly reported test. Urine was mostly collected through clean catch midstream technique and as a random voided or first-morning sample in 56% of studies (15/27). Use of sterile containers was mentioned in ten out of 27 studies. More details on urine sample collection, handling and storage, and the details of urine culture incubation can be found in Appendix 4.

The overall quality of included studies was moderate (Figure 2). Twelve out of 27 studies gave a proper description of patients' selection with the remaining not giving enough details to assess this methodological aspect of the study. There was no concern for risk of bias due to

153 index test implementation in over 80% of the studies (22/27). Similarly, for the reference
154 standard except two studies, the performance of the urine culture was classified as high risk of
155 bias. Flow and timing were described with sufficient details in one-third of studies (9/27). The
156 high concern over the applicability of findings was due to the type of the reference standard in
157 five studies and the index test in one case. The main concern in the case of the reference
158 standard was the use of a double urine culture to confirm the diagnosis of bacterial infection.
159
160 Twenty-one studies (9,491 women) reported accuracy data for the detection of nitrites using
161 urine dipstick and eight for the combination of positive nitrites or leukocytes (5,940 women).
162 The average prevalence of asymptomatic bacteriuria in these studies were 0.08 (95% CI 0.06 to
163 0.10). The pooled sensitivity of urine dipstick for positive nitrites in detecting infection was
164 0.55 (95% CI 0.42 to 0.67) with specificity 0.99 (95% CI 0.98 to 0.99). The pooled sensitivity
165 of positive nitrites or leukocytes was 0.73 (95% CI 0.59 to 0.83) with specificity 0.89 (95% CI
166 0.79 to 0.94). For both tests, the accuracy parameters were heterogeneous with greater
167 variability in sensitivity than specificity (Figure 3), 95% prediction contour was visibly wider
168 for the combined markers (Appendix 5). The likelihood ratio of the positive test result for the
169 urine dipstick test using only nitrites marker was 54.1 (95% CI 26.5 to 266.21.
170
171 One study each contributed data on the specificity and sensitivity of chlorhexidine reaction and
172 uriscreen catalase tests. The sensitivity of the former was 1.00 (95% 0.65 to 1.00) and
173 specificity (0.54, 95% CI 0.46 to 0.62) (Table 2). Use of Griess test to detect the presence of
174 nitrites was reported in two studies (728 women). The sensitivity of the test was comparable to
175 Uriscreen catalase test 0.65 (95% CI 0.50 to 0.78) with a specificity of 0.99 (95% CI 0.98 to
176 1.00). The likelihood ratio of the positive test result was 56·6 (95% CI 12.6 to 255.1). Only one
177 study reported the accuracy of the microscopic technique with the bacterial count in a

centrifuged urine sample with a clearly defined threshold of more than 20 bacteria per High Power Field (HPF). The sensitivity and specificity were 0.78 (95% CI 0.45 to 0.94) and 0.92 (95% CI 0.88 to 0.94), respectively.

Accuracy data of three dipslide-based tests included evaluation of Uricult (two studies), Microstix-3 (one study) and a generic dipslide method with gram stain dyeing and threshold of one or more bacteria per Oil Immersed Field (OIF) (six studies). Uricult had a sensitivity of 0.92 (95% CI 0.69 to 1.00) and specificity 0.85 (95% CI 0.24 to 1.00). The dipslide with gram staining on uncentrifuged urine had sensitivity and specificity of 0.86 (95% CI 0.80 to 0.91) and 0.97 (95% CI 0.93 to 0.99) respectively (Figure 3). The likelihood ratio of the positive test result was 30.2 (95% CI 11.9 to 76.6).

Sensitivity analysis was possible for dipstick with nitrites only as a marker, dipstick with nitrites or leukocytes and dipslide with gram staining. In all three cases, we explored the impact of population description and use of the sterile containers for urine storage. Neither of the factors changed the summary accuracy of the dipslide with gram staining. Analysis limited to studies with a clearly described population (asymptomatic women or not taking antibiotics) showed a marginal reduction in sensitivity (by 4%) for urine dipstick with positive leukocyte or nitrites marker. The pooled sensitivity of urine dipstick (nitrites with or without leukocytes) limited to studies providing details of urine container's sterility, presented a minimal increase in parameter precision. Findings from studies with low risk of bias and studies where the type of urine sample was not properly described had a minimal impact on the sensitivity the dipstick test with no change in the value of the pooled specificity.

Discussion

Out of 27 types of index tests identified in the literature, nine were suitable for use in the asymptomatic population. Three of them (urine dipstick, Griess test and dipslide with gram staining) had values of likelihood ratios for positive test result indicative of their usefulness (values > 10) in detecting asymptomatic bacteriuria during antenatal care. All test were minor to moderate usefulness to rule out the infection (likelihood ratios for the negative result between 0.5–0.1).

This systematic review is a comprehensive and robust synthesis of accuracy data concerning on-site tests to detect asymptomatic bacteriuria during antenatal care. Prospectively registered protocol with pre-specified population, reference standard, and definition of the outcome informed study selection, data extraction, and analysis. On all stages of the review process, we followed current guidelines and standards.[11] The literature search in electronic databases restricted to test accuracy studies due to pragmatic reasons was supplemented by manual reference check. The publication bias due to limitations of available statistical methods[16, 17] was not investigated in this review. However, we did undertake an extensive exploration of the heterogeneity between estimates of tests accuracy in individual studies.

The main limitation of this review was poor reporting in individual studies and paucity of data. The quality assessment was hindered by insufficient reporting of characteristics or recruited women, their flow through the study and timing between the use of index test and reference standard. Empirical evidence showed that test accuracy estimates can be affected by flaws in study design and its conduct.[14] The estimates of test accuracy for four included tests were based on data from single studies with small sample sizes.[22-24] This makes the parameters less reliable (wide confidence intervals) and more prone to chance findings. In order to compare the accuracy of all identified tests, we used the univariate model to pool sensitivity and specificity

estimates when less than four studies were available. Even though this approach does not account for correlation between two parameters as in the bivariate model, the findings should be fairly similar.[25] Despite these limitations our findings merit consideration as the most robust and current evidence synthesis.

The prevalence of asymptomatic bacteriuria in included studies ranged from 2 – 23% which overlaps with previously reported range[1]. The likelihood ratios of the positive test result for the urine dipstick test (only nitrites), Griess test and generic dip slide with gram staining (bacterial count > 1/OIF) were indicative of tests usefulness in ruling in asymptomatic bacteriuria.[21] The likelihood ratio of a positive result with Dipslide Uricult due to wide confidence intervals cannot be considered reliable. However, its likelihood ratio for the negative result was the only one indicative its usefulness to rule out the infection (< 0.1). Likelihood ratios can be used to help adapt the results of the findings to individual situation basing on Bayes' theorem.[26] With pretest probability derived from identified studies we calculate the posttest probability of having the infection with a positive and negative test result (Table 2). Two out of nine evaluated tests (urine dipstick with positive nitrites and Griess test) increased the probability from 8·0% to above 80.0% in case of positive result, and both reduced it by half in case of a negative result. Need for training and access to basic laboratory facilities might make Griess test and Gram staining less attractive than urine dipstick in resource-limited settings.

Undetected and subsequently not treated asymptomatic bacteriuria is linked to pyelonephritis and other complications.[3] Antibiotic treatment seem to reduce the risk of pyelonephritis in pregnancy and undesired pregnancy outcomes (preterm birth and low birth weight). Women incorrectly classified as positive (false positive) may be exposed to an unnecessary course of antibiotics with not well documented adverse effects.[27] In light of lack of robust evaluation of

harms and increasing antimicrobial resistance, it is crucial to correctly identify women who will truly benefit from the treatment.[8]

All identified on-site tests when positive increased posttest probability of detecting asymptomatic bacteriuria during the antenatal period. Urine dipstick, Griess test and dipslide with gram staining are most useful point-of-care options for ruling in the infection. Future research should aim to support the clinical decision-making on the management of asymptomatic pregnant women when access to urine culture is limited.

Contributors

ER selected eligible texts, data extraction form, extracted data, wrote the protocol, cleaned and analyzed the data, drafted and revised the manuscript. SF selected eligible texts, extracted data, and revised the paper. KSK, LM resolved discrepancies between reviewers and revised the draft paper. ER did statistical analysis, supervised by JZ. All authors contributed to the drafts and final version of the manuscript.

ER have full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Figures

Figure 1 Flow diagram describing selection of studies and tests

Figure 2 Study quality assessment using QUADAS-2 tool

Figure 3 Overview of sensitivity and specificity of tests to detect asymptomatic bacteriuria in pregnancy

References

1. Te West NIM, K.H. Urinary tract infection in gynaecology and obstetrics. *Obs Gyn Rep Med* 2014; 24(11): 321-5.
2. Sheiner E, Mazor-Drey E, Levy A. Asymptomatic bacteriuria during pregnancy. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* 2009; 22(5): 423-7.
3. Kincaid-Smith P, Bullen M. BACTERIURIA IN PREGNANCY. *Lancet (London, England)* 1965; 1(7382): 395-9.
4. Council AHMA. Clinical Practice Guidelines: Antenatal Care – Module 1. Canberra, Australia: Australian Government Department of Health and Ageing; 2012.
5. Excellence NIfHaC. Antenatal care for uncomplicated pregnancies (CG62). In: Health TNCCfWsaCs, editor. NICE guidelines: National Institute for Health and Care Excellence; 2008.
6. Lumbiganon P, Laopaiboon M, Thinkhamrop J. Screening and treating asymptomatic bacteriuria in pregnancy. *Current opinion in obstetrics & gynecology* 2010; 22(2): 95-9.
7. Deville WL, Yzermans JC, van Duijn NP, Bezemer PD, van der Windt DA, Bouter LM. The urine dipstick test useful to rule out infections. A meta-analysis of the accuracy. *BMC urology* 2004; 4: 4.
8. Colgan R, Nicolle LE, McGlone A, Hooton TM. Asymptomatic bacteriuria in adults. *American family physician* 2006; 74(6): 985-90.
9. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann InternMed* 2008; 149(12): 889-97.
10. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann InternMed* 2011; 155(8): 529-36.

11. Deeks JJB, P.; Gatsonis, C. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. 2010.
12. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj* 2009; 339: b2700.
13. van der Weijden T, CJ IJ, Dinant GJ, van Duijn NP, de Vet R, Buntinx F. Identifying relevant diagnostic studies in MEDLINE. The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. *Family practice* 1997; 14(3): 204-8.
14. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *Jama* 1999; 282(11): 1061-6.
15. Group TWB. Country and Lending Groups. 2015 2015. http://data.worldbank.org/about/country-and-lending-groups (accessed 2/5/2016 2016).
16. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; 58(9): 882-93.
17. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002; 31(1): 88-95.
18. StataCorp. Stata Statistical Software. 12.1 ed: College Station, TX: StataCorp LP. 2015.; 2015.
19. Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *The Stata Journal* 2009; 9(2): 211-29.
20. Dwamena BAS, R.; Carlos, R.C. Midas: meta-analysis of diagnostic accuracy studies. 2010.
21. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ (Clinical research ed)* 2004; 329(7458): 168-9.

22. Archbald FJ, Verma U, Tejani NA. Screening for asymptomatic bacteriuria with microstix. *Journal of Reproductive Medicine for the Obstetrician and Gynecologist* 1984; 29(4): 272-4.
23. Okusanya B, Aigere E, Eigbefoh J, Okome G, Gigi C. Is a chlorhexidine reaction test better than dipsticks to detect asymptomatic bacteriuria in pregnancy? *Journal of Obstetrics and Gynaecology* 2014; 34(1): 21-4.
24. Teppa RJ, Roberts JM. The uriscreen test to detect significant asymptomatic bacteriuria during pregnancy. *Journal of the Society for Gynecologic Investigation* 2005; 12(1): 50-3.
25. Simel DL, Bossuyt PM. Differences between univariate and bivariate models for summarizing diagnostic accuracy may not be large. *J Clin Epidemiol* 2009; 62(12): 1292-300.
26. Fagan TJ. Letter: Nomogram for Bayes theorem. *The New England journal of medicine* 1975; 293(5): 257.
27. Smaill FM, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. *The Cochrane database of systematic reviews* 2015; 8: Cd000490.