Screening for Asymptomatic Bacteriuria in Pregnancy: Systematic Review & Meta-analysis

October 13, 2017

University of Alberta Evidence Review and Synthesis Centre Team:

Aireen Wingert, Jennifer Pillay, Robin Featherstone, Michelle Gates, Meghan Sebastianski, Kassi Shave, Ben Vandermeer, Lisa Hartling

University of Alberta, Edmonton, Canada

Clinical Experts:

Anne Biringer, MD, CCFP, FCFP

Fiona Smaill, MBChB, FRCPC

CTFPHC Asymptomatic Bacteriuria Working Group:
Ainsley Moore (Chair), Roland Grad, Stéphane Groulx, Kevin Pottie, Brett Thombs, and Marcello Tonelli

PHAC Global Health & Guidance Division Staff:
Marion Doull, Alejandra Jaramillo Garcia, Susan Courage, Nicki Sims-Jones

Suggested citation: Wingert, A., Pillay, J., Featherstone, R., Gates, M., Sebastianski, M., Shave, K., Vandermeer B., Hartling, L. (October 2017). *Screening for Asymptomatic Bacteriuria in Pregnancy: Systematic Review and Meta-analysis.* Evidence Review and Synthesis Centre, University of Alberta: Edmonton, Alberta.

REVISION HISTORY

Section	Date	Description/Changes	Reason for Change
Analytical Framework and Staged Approach	December 9, 2016*	Added "direct evidence" to Stage 1 and "indirect evidence for KQ1" to Stage 2 and 3 to clarify that KQ4 (treatment) and KQ5 (test accuracy) are encompassed in KQ1 (screening effectiveness), but each component on its own does not provide direct evidence on benefits/harms of screening, and therefore are included as indirect evidence.	To clarify that evidence on each of, treatment and test accuracy, are indirect evidence of screening effectiveness.
Figure 1. Analytical Framework	December 9, 2016*	Added treatment and no treatment to each arm of ASB+ and ASB-.	To identify that, although unlikely, there is the possibility of a study with treatment and no treatment arms for both patients in the ASB-positive and ASB-negative group.
Eligibility Criteria	December 9, 2016*	Added examples of different screening tests (e.g., dipstick vs. Griess test) to illustrate the difference between KQ1a (screening vs. no screening) and KQ1b (different screening tests or algorithms such as frequency of testing or testing criteria).	To clarify the difference between KQ1a and KQ1b.
Eligibility Criteria	December 9, 2016*	Added "maternal and neonatal" to Harms to include both categories are included in serious and non-serious AEs. Added neonatal thrush to list of non-serious AEs as an example of a potential neonatal harm.	To clarify that harms to both mother and neonate are included.
Eligibility Criteria	December 9, 2016*	Revised setting to any primary care or clinical setting which provides obstetric/antenatal care to pregnant women.	To avoid precluding care provided in other settings (e.g., obstetric office/clinic).
Eligibility Criteria	January 27, 2017**	Revised perinatal mortality to ≥ 20 weeks of gestation.	To capture all perinatal mortality reported, including stillbirths which are reported using different criteria among studies.
Eligibility Criteria	February 24, 2017***	Revised PICOTS for KQ1a to examine benefits and harms of a screening program compared to no-screening program, i.e. a screening test as the intervention was removed. Case-control study (Friedman 2012) was excluded for KQ1a.	A screening program is differentiated from a screening test, such that in the former screening would be intended for all women in the intervention group with a majority, but not all, receiving a screening test. This resembles a typical screening trial.
GRADE Assessments	October 6, 2017***	Revised GRADE assessments in KQ4 to no longer downgrade for indirectness due to use of evidence on treatment to infer knowledge about screening interventions; this is considered "linked" evidence. This revision did not lead to any changes to the overall GRADE evaluations or conclusions of our review.	To align with GRADE guidance for interpreting linked evidence and consider the body of evidence (for KQ4) independently from that for KQ1. The "linked" evidence will be considered as such by the CTFPHC when creating the Evidence to Decision framework for their guideline.

*Revision prior to final study selection and extraction

**Revision prior to data extraction and analyses; Canadian Task Force on Preventive Health Care (CTFPHC) members were blinded to all study reports

***Revision post-hoc, after data extraction and analyses; CTFPHC members were not blinded to study details

Summary

Purpose: This review was produced for the Canadian Task Force on Preventive Health Care (CTFPHC) to inform their recommendations on screening for asymptomatic bacteriuria (ASB) in pregnancy.

Review Approach: Following CTFPHC methods, a staged approach was used based on the quality of evidence when applying Grading of Recommendations Assessment, Development and Evaluation (GRADE) methods. The quality of evidence was determined for outcomes rated by the CTFPHC, using input from consultations with Canadian women, as important or critical for decision-making. A reduction in the following outcomes would favor screening: maternal mortality, maternal sepsis, pyelonephritis, perinatal mortality, spontaneous abortion, neonatal sepsis, preterm delivery, low birthweight, and serious harms (e.g., fetal abnormalities after antibiotic treatment). Stage 1 examined screening effectiveness on the benefits and harms of any screening program compared with no screening and benefits and harms of different screening methods/algorithms (e.g., detection methods, timing and collection; test for cure after treatment of women found to have significant bacteriuria). Women's valuation ("weighing") of the benefits and harms of screening was also examined during this stage, with evidence on outcome valuation related to any antibiotic use in pregnancy considered. Evidence on screening effectiveness was very low quality, therefore we did not review studies on cost-effectiveness of screening programs which would have relied on similar evidence. Based on very low quality evidence from stage 1, stage 2 employing a "linked" evidence approach was undertaken to examine the effectiveness of antibiotic treatment for pregnant women with ASB. Since evidence from stages 1 and 2 considered screening programs and treatment based on the use of urine culture (gold standard), but not point-of-care (rapid) methods, we did not conduct an evidence review of the accuracy of point-of-care screening methods.

Data Sources: Comprehensive searches were conducted in bibliographic databases most relevant for each key question. For evidence on screening effectiveness, we searched MEDLINE (1946-) via Ovid; Embase (1974-) via Ovid; Cochrane Library; CINAHL (1937-present) via EBSCOhost; and PubMed via NCBI Entrez on June 15, 2016 (update searches ran on September 6, 2017). For evidence on women's outcome valuation, we modified the search (ran on July 4, 2016; update searches ran on September 5, 2017) to include relevant terms and added the database PsycINFO; a search for evidence on cost-effectiveness was not conducted. For evidence on treatment, we searched on October 14, 2016 for systematic reviews using PubMed (1946-) via NCBI Entrez, the Cochrane Database of Systematic Reviews (inception-) and the Database of Abstracts of Reviews of Effects (DARE) (inception-2013) via Wiley Cochrane Library. The authors of the included systematic review on treatment, published in 2015, provided us with the results of their recent search update (using Cochrane Pregnancy and Childbirth Group's Trials Register) in November 2016 and October 2017. We also searched for grey literature and additional studies through internet-based searches, electronic libraries, trial registries, conference proceedings, and contact with experts.

Study Selection: Two reviewers independently screened titles and abstracts of citations from all database searches. Full texts of studies that were classified as "include/unsure" by either reviewer were retrieved and screened independently by two reviewers using a standard form with explicit inclusion and exclusion criteria. Disagreements were resolved through consensus or consultation with a third reviewer. For each key question, the flow of literature and reasons for full-text exclusion are recorded in a PRISMA Flow

Chart. For evidence related to treatment effectiveness, systematic reviews were assessed for eligibility based on having conducted a search strategy in more than one database, whether the selection criteria were reported, and whether the population, intervention, comparator, timing, and setting (PICOTS) criteria closely matched ours.

Data Abstraction: For evidence in stage 1, one reviewer independently extracted data, and another reviewer verified all data from each included study on its study design; country of origin, sample size, characteristics of the patients, interventions, and comparator(s); clinical setting; and, outcomes of interest. Authors of included studies were contacted for clarification of study details and outcome data as necessary. For treatment evidence in stage 2, we extracted data from the systemic review on its selection criteria (PICOTS) and its included studies, as well as from an additional trial captured by the review authors' search update. We verified data from the systematic review, and also examined the primary studies for additional participant characteristics and outcome details relevant to the current review. A narrative summary with accompanying tables is reported for all studies. Two reviewers independently assessed the methodological quality of each included study with the following tools: Newcastle-Ottawa Quality Assessment Scale for observational studies, the Center for Evidence-based Management appraisal tool for cross-sectional studies, and the Cochrane Risk of Bias tool for trials. Disagreements on data extraction or methodological quality assessments were resolved through consensus or consultation with a third reviewer.

Analysis & Interpretation: We performed meta-analyses for the dichotomous outcomes in the evidence for screening and treatment, using the DerSimonian and Laird random effects model with Mantel-Haenszel method, and report relative risks (RR) with corresponding 95% confidence intervals (CIs). For outcomes having statistically significant effects, we calculated absolute risk reduction (ARR), and number needed to screen (NNS) or number needed to treat (NNT) based on the control group event rates and RR. Where there were at least two studies per category for a variable, we performed subgroup analyses as planned for clinical (patient and intervention) characteristics of interest. We conducted sensitivity analyses for methodological issues (e.g., risk of bias) when substantial heterogeneity was found in meta-analysis. We examined funnel plots and conducted Egger's test to detect small-study bias when there were at least eight studies in a meta-analysis. When data were not pooled, we provided a narrative summary of findings. Two reviewers independently assessed the quality of the body of evidence using GRADE methodology, with consensus based on discussion and input from a third reviewer.

Results: Four non-concurrent cohort studies compared outcomes for groups of pregnant women before and after introduction of a screening program for ASB. All studies used a urine culture for screening, with some variability in the collection methods and treatment protocols. Three studies compared screening with no screening; meta-analysis using data from these studies showed a statistically significant reduction in pyelonephritis (RR 0.28; 95% CI 0.15, 0.54; ARR 1.3%; NNS 77, 95% CI 65, 121). No significant differences were found when comparing screening with no screening for other outcomes of perinatal mortality based on two studies (RR 1.21; 95% CI 0.01, 102.93), spontaneous abortion based on one study (RR 0.96; 95% CI 0.41, 2.27), preterm delivery from two studies (RR 8.70; 95% CI 0.32, 240.07), and fetal abnormalities (neonatal serious harm) from one study (RR 1.50; 95% CI 0.25, 8.87). One study compared frequent screening with one-time (first prenatal visit) screening and found no significant difference for pyelonephritis (RR 1.09; 95% CI 0.27, 4.35) or preterm delivery (RR 1.57; 95% CI 1.11, 2.23). No study provided evidence on how women weigh the benefits and harms of screening for ASB; seven studies provided evidence on sentiments on harms only, and reported conflicting opinions about

antibiotic use during pregnancy particularly on teratogenic risks. Fifteen trials examined the effectiveness of antibiotics versus placebo or no antibiotics for women with bacteriuria ($\geq 10^5$ colony-forming units of one organism per mL); only three trials reported that participants were asymptomatic and some trials included high-risk women. Fourteen of the trials reported on outcomes relevant to this review. Meta-analysis from 12 trials found a significant reduction from antibiotic treatment compared with placebo/no treatment in development of pyelonephritis among women with bacteriuria (RR 0.24; 95% CI 0.13, 0.41; ARR 17.6%; NNT 6, 95% CI 5, 7; I^2=60%). One of our planned subgroup analysis, for pyelonephritis based on whether or not a confirmatory (second specimen) culture was used (RR 0.19, 95% CI 0.11, 0.31, I^2=31% versus RR 0.50, 95% CI 0.19, 1.35, I^2=41%), seemed to have some credibility based on visual inspection of the forest plots (indicating possible important difference) and a reduction in heterogeneity within each subgroup; results from testing for a difference between subgroup effects, though, was not statistically significant (p=0.08). Seven studies found that treatment reduced low birth weight (RR 0.63; 95% CI 0.45, 0.90; ARR 4.4%; NNT 23, 95% CI 15, 85). No significant difference between groups was found for all other outcomes: perinatal mortality based on six studies (RR 0.96, 95% CI 0.27, 3.39), spontaneous abortion based on two studies (RR 0.60, 95% CI 0.11, 3.10), neonatal sepsis based on two studies (RR 0.22, 95% CI 0.01, 4.54), preterm delivery based on four studies (RR 0.57, 95% CI 0.21, 1.56), and neonatal harms (fetal abnormalities) from four studies (RR 0.49, 95% CI 0.17, 1.43; no cases of infant hemolytic anemia in one study). No study on screening or treatment reported on maternal mortality, maternal sepsis or maternal harms.

Limitations: Based on our risk of bias tools, the non-concurrent cohort studies examining screening effectiveness were of unclear or low risk of bias; nevertheless, observational studies introduce several potential biases which are not captured in this tool, particularly as related to reporting bias which was suspected for outcomes apart from pyelonephritis. For evidence related to screening effectiveness, studies used a urine culture to detect ASB but the criteria for defining a positive test was not always clear or reported. Many patient and intervention characteristics were not reported, or were inconsistently reported between studies. Outcomes were defined variably among studies. One treatment study only included women who were treated for group B streptococcus based on urine culture, only three of the treatment trials reported that participants were asymptomatic, and four trials included high-risk women. The small sample sizes and event rates for many outcomes led to imprecise effect estimates. Subgroup analyses were few because of the limited reporting on subgroup variables of interest and number of studies contributing to most outcomes; although our findings on pyelonephritis for subgroups based on two (for confirmation) versus one culture specimen appear to have some credibility, these analyses rely on study-level data and are observational (i.e., studies are not randomized) and exploratory in nature. The majority of studies on treatment were published in the 1960s, pre-dating current obstetric practices having, for example, better recognition of risk factors for urinary tract infections and other pregnancy complications, prompt treatment of symptoms, and a broader range of antibiotic options; these factors would suggest a lower control group (baseline) event rate and therefore less absolute benefit in current practice. Much of the evidence came from trials on treatment of bacteriuric women (2-10% of screening population), therefore the results fail to incorporate several effects that would be captured in studies of screening effectiveness (e.g. effects on non-screened women who develop symptoms, or on ASB-negative women; effects from non-adherence to screening protocol). Studies published in languages other than English and French were not included; however, literature suggests language restrictions in systematic reviews of conventional medicine do not appear to bias results of meta-analyses.

Interpretation of Results & Conclusion: This systematic review examined three sets of evidence to inform recommendations on screening for ASB in pregnancy. Using the GRADE approach, we determined the evidence to be of very low quality for most outcomes from observational studies comparing screening programs using urine culture with no screening; as such, we have no or very little certainty in the effect estimates for these outcomes. Moreover, several outcomes were not reported. Similar interpretations are made about the evidence from one study comparing frequent screening with one-time screening. No direct evidence was found on how women weigh the benefits and harms of screening and/or treatment for ASB and how this might affect their decisions to undergo screening. Antibiotic treatment for women having significant bacteriuria likely reduces the incidence of pyelonephritis in these women and the number of their babies born at low birth weight (both of low quality evidence). We are uncertain if the magnitudes of the effect estimates from treatment are true, and about the extent to which we can apply these results to asymptomatic populations. Very low quality evidence from these trials did not allow us to have any certainty about effects from treatment on other maternal and neonatal benefits and for fetal abnormalities and hemolytic anemia; no evidence was found for other serious harms.

PROSPERO Registration #: CRD42016045263

Chapter 1. Introduction

Background & Purpose

Asymptomatic Bacteriuria in Pregnancy

Asymptomatic bacteriuria (ASB)—synonymous with asymptomatic urinary tract infection (UTI)—signifies a significant quantitative count of bacteria in the urine without symptoms of a lower (acute cystitis) or upper urinary tract/kidney (acute pyelonephritis) infection.[1, 2] There is a 2-10% prevalence of ASB in premenopausal, ambulatory women,[1] but due to anatomical and physiological changes (e.g., displaced bladder) to the urinary tract in pregnancy there are theoretical reasons to suspect higher rates of ASB during pregnancy and consequently a greater chance of progression to symptomatic UTI and other pregnancy complications (e.g., pyelonephritis, preterm delivery).[1, 3] Numerous risk factors for ASB in pregnancy have been identified, with low socioeconomic status, higher parity, a history of recurrent UTI, diabetes, and anatomical abnormalities of the urinary tract most cited.[1, 2, 4]

Consequences of Untreated Bacteriuria in Pregnancy and Rationale for Review of Screening
There is a potentially greater risk in pregnant women compared to other populations for ASB developing into pyelonephritis[3] with its associated inflammation of the renal parenchyma, calices and pelvis,[5] although controversy exists. Historical reports pre-1980[6-8] finding that upwards of 40% of pregnant women with ASB developed pyelonephritis lend support for screening and treatment with antibiotics; current estimates of the incidence of pyelonephritis in ASB positive women are hard to locate because of universal acceptance of this practice (e.g., in Canada for more than two decades). Reports of a reduced incidence of pyelonephritis in pregnant women after introduction of routine screening (e.g., 0.3 to 0.57% vs. 1-2%[9]) suggest that these programs have been beneficial.

Recent evidence suggests an association between clinical signs of pyelonephritis and perinatal outcomes. A retrospective cohort study (Wing et al[10]) of women who delivered in hospitals in the United States from 1993 to 2010 found that pyelonephritis was linked to higher risk of maternal respiratory insufficiency, septicemia, renal dysfunction, and anemia. However, controversy exists over the mechanism linking ASB, pyelonephritis, and adverse perinatal outcomes (e.g., whether ASB affects pregnancy and neonatal outcomes solely through pyelonephritis or also other mechanisms such as prostaglandin activation),[2, 4] and therefore also about whether treatment of ASB with antibiotics will reduce the risk of such adverse outcomes. A 2015 Cochrane review[4] of fourteen trials found that antibiotic treatment for ASB in pregnancy may reduce the incidence of pyelonephritis, preterm birth, and low birth weight babies. However, the authors' confidence in the findings was low due to poor quality evidence.

Although the direct link between pyelonephritis and adverse perinatal outcomes may not be easily resolved[4], an examination of whether screening of all pregnant women and treatment for significant bacteriuria is effective is of interest. Knowledge of whether screening and treatment offer as much benefit today, when there is more advanced obstetrical care (e.g., for treating acute pyelonephritis) and awareness of risk factors for pregnancy complications, would be valuable information. Knowing that some risk for harm exists from taking antibiotics during pregnancy, the benefit-to-harm ratio may be less favourable

than historically thought. This review will examine up-to-date evidence on screening for ASB in pregnancy, for reducing the risks of pyelonephritis and neonatal and maternal complications.

Issues to Consider for Screening Tests

Significant bacteriuria is usually defined by the presence of at least 100 x 10^6 colony-forming units (CFU) per litre of urine of a single organism in two consecutive clean-catch specimens (non-Canadian criteria typically report $\geq 10^5$ CFU/mL).[4, 7] Acceptable thresholds and repetitions of testing to confirm bacteriuria in pregnancy may vary in practice. The quantitative urine culture is considered to be the gold standard for accurate detection of ASB. However, it is costlier, more labour intensive and more time-consuming compared with rapid urine screening tests (urinalysis, dipstick nitrite tests) which reportedly are less accurate in identifying people with bacteriuria.[1, 2] A recent systematic review comparing the accuracy of onsite methods (point-of-care tests that are widely available in resource-limited settings) with urine culture, concluded that point-of-care tests were not reliable in detecting pregnant women with ASB.[11] Further, pregnant women have very active urinary sediment which may contribute to issues with test accuracy. There is no consistent recommendation for urine specimen collection in pregnancy (number of specimens, clean-catch with or without perineal cleansing) or optimal timing and frequency of screening tests or follow-up cultures.[2] It is unclear whether available point-of-care methods for ASB are comparable to the current gold standard (urine culture) for identifying bacteriuric patients. The standard urine culture protocol is evolving with the testing of emerging techniques that may improve the detection of the most clinically relevant uropathogens.[12, 13] However, at this time, urine culture is considered the reference standard.

Issues to Consider for Harms of Screening

Patients may have preferences for avoiding harms due to screening with the intention to treat in asymptomatic conditions, particularly when they may otherwise not benefit from the treatment (e.g. in cases where ASB would not lead to complications). Harms from antibiotic treatment need to be considered when making decisions about screening practices for all women with ASB in pregnancy. Some sources have outlined concerns with incidence and reporting on adverse effects of antibiotic treatment for ASB, UTIs, or antibiotic use in general during pregnancy.[2, 4, 14] Some trials evaluating treatment versus no treatment/placebo of ASB in pregnancy have been critiqued for poorly reporting harms,[4] such that making judgments on the net balance of benefits and harms may be difficult. Increasingly, there are concerns about the effect of antibiotics on the human microbiota and the immune system. Antimicrobial resistance has made the selection of an antibiotic for an individual more difficult.[4] Further, a test-for-cure is increasingly more important and more than one type of antibiotic may be required if sensitivity testing is not performed or accurate.

Recommendations in Other Guidelines and Current Practice

Canadian Organizations

The Society of Obstetricians and Gynecologists of Canada (SOGC) recommends screening during pregnancy using routine testing for ASB with a single quantitative culture in any trimester and treating single-strain colony counts of 10^5 CFU/mL (or 10^8 CFU/L) or greater with appropriate antibiotics to prevent adverse outcomes such as pyelonephritis and preterm birth.[15] They support a single quantitative culture in any trimester as sufficient and recommend re-treatment with antibiotic sensitivity testing for

women with recurrent bacteriuria, although they do not make recommendations for timing or frequency of re-testing. As for ASB $\geq 10^5$ CFU/mL, similar recommendations apply when group B streptococcal (GBS) bacteria is detected in the urine; separate recommendations (not relevant for this review) are made for screening and treating GBS (at any colony counts) at time of labor or rupture of membranes for prevention of early-onset neonatal GBS disease.

Guidelines from International Organizations

The U.S. Preventive Services Task Force (USPSTF) 2008 guideline[16] on screening for ASB in adults recommends with high certainty of substantial net benefit that all pregnant women be screened at 12 to 16 weeks of gestation (or first prenatal visit) for ASB using a urine culture, and that treatment with antibiotics significantly reduces the incidence of pyelonephritis and low birthweight. The evidence informing this reaffirmation of the original USPSTF recommendation from 2004 is mainly drawn from a 2007[17] Cochrane review of treatment effectiveness. The American Academy of Family Physicians (AAFP)[18] endorses the recommendations of the USPSTF. The Infectious Diseases Society of America[19] recommends screening for bacteriuria by urine culture for pregnant women in early pregnancy, and treatment if results are positive, with periodic re-testing for recurrent bacteriuria after therapy. The American Academy of Pediatrics (AAP), jointly with the American College of Obstetricians and Gynecologists (ACOG) recommend to screen and to treat significant bacteriuria and then to test for cure.[20]

The UK's National Institute for Health and Care Excellence (NICE) states that women should be offered routine screening for ASB by midstream urine culture early in pregnancy to reduce the risk of developing pyelonephritis.[21]

The Scottish Intercollegiate Guidelines Network (SIGN) recommends that pregnant women be tested for ASB by urine culture at the first antenatal visit and culture-positive patients be treated with an antibiotic.[22]

Current Practice

Several major healthcare organizations in North America (USPSTF, IDSA, ACOG, AAP, AAFP) recommend screening of pregnant women and treating patients who have been confirmed with ASB using antibiotics. In Canada, the current usual practice is to obtain at least one urine sample (with reported variations in practice on timing such as at first prenatal visit), and potentially with subsequent testing if indicated (e.g., if patient presents with symptoms). Urine samples may be tested with a dipstick, for example, to test for protein or glucose, and may also be used to detect leukocytes, blood and/or nitrites; urine testing in pregnancy may be intended for detecting conditions other than for ASB. Furthermore, there appears to be diversity in urine testing for the presence of significant bacteriuria, with respect to how the sample is collected, what is used to detect presence of bacteriuria (e.g., culture most often but perhaps not always), when sample(s) for ASB is/are collected in pregnancy, and if/when confirmatory tests are used. Because of this screening for ASB may consist of several variations in terms of testing methods, timing, and collection, as well as treatment protocols (duration, test-for-cure, threshold of bacteria for treatment).

The goal of this review is to determine the effectiveness of screening for ASB among pregnant women. This evidence synthesis will inform recommendations on screening for ASB made by the Canadian Task

Force on Preventive Healthcare (CTFPHC). As part of the guideline development process, the CTFPHC will also engage patient and organizational stakeholders to gather information on patient preferences and key implementation considerations, such as strategies to help address potential health inequities and any concerns about the acceptability and feasibility of the guideline.

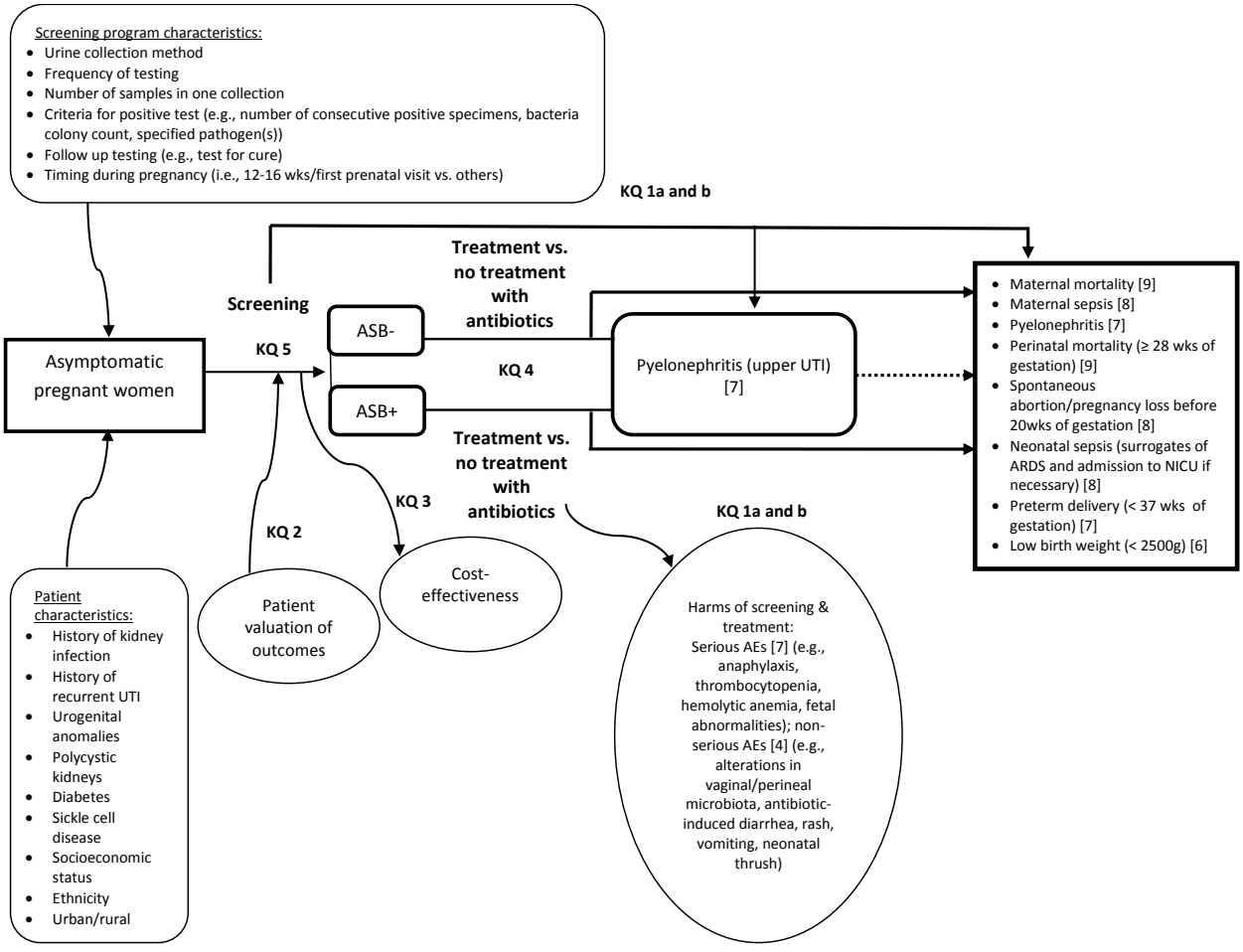
Chapter 2. Methods

An *a priori* protocol was developed following the methods of the Canadian Task Force on Preventive Health Care (CTFPHC)[23] and is registered with the International Prospective Register of Systematic Reviews (PROSPERO) (Registration #CRD42016045263).

Analytic Framework, Review Approach and Key Questions

Figure 1 is an analytical framework that depicts the basic structure used to address the Key Questions (KQs) for evaluating the benefits and harms of screening programs for asymptomatic bacteriuria (ASB) during pregnancy.

Figure 1. Analytical Framework



AEs: adverse events; ARDS: acute respiratory distress syndrome; ASB: asymptomatic bacteriuria; g: grams; KQ: key question; NICU: neonatal intensive care unit; UTI: urinary tract infection; wks: weeks
Note: Patient-important outcome ratings by CTFPHC members and patients as critical [7-9, out of 9] or important [4-6, out of 9] are included.

A staged approach was used based on the availability and quality of the body of evidence. Quality of evidence (classified as high, moderate, low, very low) was assessed using methods developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group (http://www.gradeworkinggroup.org/), whereby high quality evidence relies on precise and consistent effect estimates from studies having few limitations on internal validity (i.e., low bias) and examining directly relevant populations, interventions, comparators, outcomes, timing, and setting (i.e., PICOTS). Decisions made during the evidence review are based on the information needs of the CTFPHC for making a recommendation in favour of or against screening based on the balance of benefits and harms for critical patient-important outcomes.

Stage 1 focused on identifying and using data from studies directly linking screening programs for ASB to patient-important benefits and harms (KQ1). Study designs providing the highest internal validity (i.e., RCTs) for this KQ were preferred with a hierarchy of evidence used after this point if necessary. After RCTs we planned to consider controlled clinical trials (CCTs; defined for this review as *experimental trials without random allocation but where intervention(s) are introduced, standardized, and allocated objectively [e.g., by date of birth, but not using subjective means such as patient or clinician preferences] by investigators and blinding of participants is typically possible*) and then prospective and retrospective controlled observational studies. This stage also included examination of women's valuation of benefits and harms (KQ2) of screening for ASB (and more broadly/indirectly treatment with antibiotics) in pregnancy. The cost-effectiveness of screening for ASB (KQ3) was to be considered only if there was moderate or high quality evidence from KQ1 indicating a favourable benefit-harm ratio. The quality of evidence from cost-effectiveness studies relies on the quality of the data inputs, such that very low quality evidence on effectiveness will not lead to any certainty in the estimates of cost-effectiveness.

If stage 1 did not provide high enough quality of evidence for making a recommendation, the CTFPHC planned to carefully consider pursuing stage 2 with documentation of rationale before proceeding. Stage 2 would commence with examination of effectiveness of treatment (linked evidence) of ASB in pregnancy (KQ4). Moreover, if studies for treatment effectiveness examined the use of point-of-care methods, rather than the current gold standard which is urine culture, an examination of KQ5 on accuracy of these tests would be considered in stage 3. Due to the linked evidence provided by treatment effectiveness (KQ4) and test accuracy (KQ5) for making recommendations on the clinical effectiveness of screening programs for all pregnant women, we would only seek data from study designs offering the greatest potential for high internal validity. That is, for KQ4 (treatment) we planned to focus on RCTs, and for KQ5 (test accuracy) we would exclude case-control designs. Where a high quality systematic review existed examining these evidence linkages to screening effectiveness, we would utilize these when possible.

Key Questions (KQs)

Stage 1 (direct evidence):

Benefits and harms of screening

KQ1a: What are the benefits and harms of screening compared with no screening for asymptomatic bacteriuria in pregnancy? Are there subgroup differences for patient characteristics (e.g., socioeconomic status [SES])?

KQ1b: What are the comparative benefits and harms of screening programs with different screening methods or algorithms for asymptomatic bacteriuria in pregnancy?

Outcome valuation

KQ2a: How do women weigh the benefits and harms of screening and treatment of asymptomatic bacteriuria in pregnancy?

KQ2b: How do women's valuation of benefits and harms of screening and treatment inform their decisions to undergo screening?

Resource use

KQ3: What is the cost-effectiveness of screening for asymptomatic bacteriuria in pregnancy?

Stage 2 (linked evidence):

Treatment

KQ4: What are the benefits and harms of antibiotic treatment compared with placebo or no treatment for asymptomatic bacteriuria in pregnancy?

Stage 3 (linked evidence):

Diagnostic accuracy of screening tests

KQ5: What is the accuracy of point-of-care screening tests compared with urine culture for asymptomatic bacteriuria in pregnancy?

Search Strategy

The literature search strategies were developed and implemented by a research librarian and peer reviewed. Searches were restricted by language to include full texts published in English and French only; literature suggests language restrictions in systematic reviews in conventional medicine do not appear to bias results from meta-analyses.[24, 25] No restrictions were applied to publication dates or study design. Full detailed search strategies for all databases are reported in Appendix 9.

Comprehensive searches were conducted in bibliographic databases most relevant for each KQ. For KQ1 (screening effectiveness), we searched MEDLINE (1946-) via Ovid; Embase (1974-) via Ovid; Cochrane Library; CINAHL (1937-present) via EBSCOhost; and PubMed via NCBI Entrez. For KQ2 (women's outcome valuation), we modified the search to include relevant terms and added PsycINFO as a database. We did not search for studies on cost-effectiveness (KQ3) because of the very low quality evidence for KQ1 (see description of staging approach, p.5). We did not search for studies or reviews on test accuracy (KQ5) because there was no evidence from KQ1 or KQ4 that point-of-care tests may replace urine culture as an accurate screening method. Searches for KQ1 and KQ2 are current to September 2017.

For KQ4 (treatment effectiveness), we conducted a database search for systematic reviews, meta-analyses and health technology assessments to ensure all potentially relevant systematic reviews were identified. We searched PubMed (1946-) via NCBI Entrez, the Cochrane Database of Systematic Reviews (inception-) and the Database of Abstracts of Reviews of Effects (DARE) (inception-2013) via Wiley Cochrane Library on October 14, 2016. Our PubMed search utilized a search filter from the Canadian Agency for Drugs and Therapeutics in Health (CADTH).[26] The authors of the included systematic review on treatment, published in 2015, provided us with the results of their recent search update (using the Cochrane Pregnancy and Childbirth Group's Trials Register)[4] on November 2016 and October 2017.

Grey literature was searched and documented according to CTFPHC methods and included internet-based searches (via adapted CADTH checklists[27]), electronic libraries (e.g., Health Canada Library, Canadian Electronic Library), and trial registries (ClinicalTrials.gov, World Health Organization International Clinical Trials Registry Platform). Based on consultation with clinical experts, the following highly relevant conference proceedings were hand-searched for recent studies not yet published (2014-August 2016): Society of Obstetricians and Gynaecologists of Canada (SOGC, Association of Medical Microbiology and Infectious Disease Canada, ID Week, and American Society for Microbiology meeting (ICAAC). Clinical and content experts identified by the CTFPHC for review of the protocol were invited to identify relevant research reports for consideration. Potentially relevant papers and websites identified by stakeholders and peer reviewers during protocol review were also searched and screened for eligibility: Society for Maternal Fetal Medicine (SMFM), American College of Obstetrics and Gynecology (ACOG) annual meeting, and Infectious Diseases Society of Obstetrics and Gynecology (IDSOG).

Study Selection

Two reviewers independently screened the titles and abstracts of all citations retrieved by the database searches. Full texts of studies that were classified as "include/unsure" were retrieved for review and screened independently by two reviewers using a standard form with explicit inclusion and exclusion criteria. Disagreements were resolved through consensus or consultation with a third reviewer. All decisions for title/abstract screening and full-text review were conducted and documented in DistillerSR.[28] For each KQ, the flow of screening and reasons for full-text exclusion are recorded in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Chart.

We followed methods adopted by the CTFPHC for integrating systematic reviews for KQ4 on treatment (see Appendix 13), where existing systematic review(s) are eligible based on a) searching more than one database, b) reporting selection criteria, and c) using populations, interventions, comparators, timing and

setting (PICOTS) criteria that closely match the current review. The included studies were assessed for eligibility to meet our inclusion criteria, incorporating existing data and extracting additional data as necessary, conducting quality assessments, and performing new meta-analyses and GRADE quality assessments.

Eligibility Criteria

We included studies of asymptomatic women at any stage of pregnancy, including populations where a proportion of women may have symptoms or present with risk factors (e.g., kidney infection, recurrent UTI, diabetes), but are considered to represent a routine prenatal care population. Studies that *exclusively* examined women with risk factors (e.g., high risk for ASB, pyelonephritis, or poor outcomes associated with some conditions such as history of kidney infection, renal anomalies, polycystic kidneys, recurrent UTI, diabetes, or sickle-cell disease) were excluded. Studies that included non-pregnant women were excluded.

The population subgroups of interest included: history of kidney infection, urogenital anomalies, polycystic kidneys, recurrent urinary tract infection (UTI), diabetes, sickle-cell disease, socioeconomic status (SES; i.e., education, income), ethnicity, and urban/rural setting.

For clinical effectiveness of screening (KQ1a) comparing any screening program with no screening, the screening program could include any screening algorithm for ASB (e.g., different screening methods, collection and timing; treatment duration, test-for-cure). For KQ1b comparing different screening programs, programs could differ by screening method (e.g., culture vs. dipstick) or algorithm (e.g., frequency of screening, urine collection methods); studies that compared differing criteria for a positive urine culture (e.g., threshold 10^3 CFU/mL versus 10^5 CFU/mL) were also eligible for inclusion. For the screening studies (KQ1), we did not exclude studies if a treatment protocol was not reported; as part of an overall screening program, it was assumed there was an intent to treat screen-positive cases. For women's outcome valuation (KQ2), any screening program for ASB during pregnancy was eligible for inclusion; we planned to use indirect evidence about antibiotic treatment during pregnancy broadly if needed. For cost-effectiveness (KQ3), we planned to include any screening program compared with no screening or another screening program. For treatment effectiveness (KQ4), any antibiotic treatment for ASB compared with no treatment or placebo was eligible for inclusion. For diagnostic accuracy (KQ5), we planned to include any index (rapid point-of-care) test compared with a urine culture (reference standard) for detecting bacteriuria. For all KQs, studies that included screening or treatment for group B streptococcus (GBS) at any time of pregnancy for any of the outcomes of interest were included. We excluded studies that screened pregnant women for group B streptococcus near delivery or at time of rupture of membranes for the prevention or treatment of chorioamnionitis or neonatal GBS (without other outcomes of interest listed above). Studies that exclusively examined urine tests for screening other conditions (e.g., proteinuria, glycosuria), and non-urine screening tests (e.g., vaginal/rectal swab culture for GBS testing) were excluded.

The screening subgroups of interest included: urine collection methods (e.g., clean-catch and/or midstream; excluding catheter methods/samples), frequency of testing, number of samples (e.g., use of

confirmatory cultures), criteria for a positive test (bacterial colony count, and specified organism(s)), follow-up testing during pregnancy, and timing during pregnancy.

All outcomes were rated independently by members of the CTFPHC and by women, as per the patient engagement activities of an independent group with expertise in knowledge translation from St. Michael's Hospital in Toronto, Ontario. All patient-important outcomes rated as critical (7 to 9 out of 9) or important (4 to 6 out of 9) for decision making were considered for inclusion (See Table 1.0 for list of critical/important outcomes and ratings). From these ratings, the eight outcomes that were rated as critical were included; of three outcomes rated as important, low birth weight (but not hypertension or acute kidney injury) was included because this was conceptualized in older studies to be the same as "preterm birth", which both the CTFPHC members and patients rated as critical. Considering harms separately, when no evidence was initially found for any of the outcomes (serious adverse events [AEs]), we planned to then include non-serious AEs which are considered important but not critical for decision making by the CTFPHC.

For perinatal mortality, we revised the original criteria of ≥28 weeks of gestation (Statistics Canada's[29] definition for perinatal mortality including late fetal deaths [stillbirths ≥28 weeks] and early neonatal deaths [deaths of infants <1 week old]) to ≥20 weeks of gestation to allow for inclusion of data from studies that used slight variations in defining this outcome. For preterm delivery defined as <37 weeks of gestation, we included one study that defined preterm birth as <38 weeks of gestation as this was considered inclusive of our criteria. For low birth weight, we included studies where low birth weight was defined as ≤2500 grams (for live births).

Table 1.0 Outcomes and ratings for KQs 1 (screening effectiveness) and 4 (treatment effectiveness)

Benefits (reduced incidence for all):

1. maternal mortality (9)*
2. maternal sepsis (8)
3. pyelonephritis (7)
4. perinatal mortality (≥ 20 weeks of gestation [e.g., intrauterine demise, stillbirth, early neonatal death]) (9)
5. spontaneous abortion/pregnancy loss < 20 weeks of gestation (8)
6. neonatal sepsis (if not reported will include surrogate outcomes of acute respiratory distress syndrome [ARDS] or admission to neonatal intensive care unit [NICU]) (8)
7. preterm delivery (live fetus passed < 37 weeks of gestation) (7)
8. low birth weight (≤ 2500g) (6)

Harms (maternal and neonatal):

1. serious adverse event(s)** associated with antibiotic treatment, *including but not limited to*: (7)
   a. anaphylaxis,
   b. thrombocytopenia,
   c. hemolytic anemia,
   d. fetal abnormalities; and,
2. non-serious adverse event(s) associated with treatment, *including but not limited to*: (4)
   a. alterations in vaginal/perineal microbiome (e.g., candidiasis, vaginitis),
   b. antibiotic-induced diarrhea,

c. rash,
d. vomiting
e. neonatal thrush

*Bracketed numbers next to each outcome above refer to patient-important outcome ratings by CTFPHC members and patients as critical [7-9, out of 9] or important [4-6, out of 9] are included
**Serious adverse event (experience) or reaction is any untoward medical occurrence that: a) results in death, b) is life-threatening, c) requires in-patient hospitalisation or prolongation of existing hospitalisation, d) results in persistent or significant disability/incapacity, or e) is a congenital anomaly/birth defect (Health Canada, 2011)

Women's outcome valuation (KQ2) included several possible outcomes related to the weighing of benefits and harms of screening and treatment (KQs 1 and 4) and how this may affect their decisions to undergo screening (e.g., relative weight/utilities of benefit and harms; willingness to be screened based on relative value placed on benefits and harms of screening programs or treatment, anxiety).

Cost-effectiveness (KQ3) outcomes would include cost per quality-adjusted life year (QALYs), incremental cost-effectiveness ratios (ICERs), and net benefit (in dollars from cost-benefit studies).

Diagnostic test accuracy (KQ5) outcomes include: sensitivity, specificity, false positives, false negatives, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio.

We included studies conducted in any primary care, or another clinical setting which provides obstetric/antenatal care to pregnant women (e.g., obstetric and hospital outpatient clinics, prisons, remote stations, community health centers, midwifery practice). For KQ3 on cost-effectiveness we planned to limit studies to those conducted using data relevant to Canada, thus within countries having a Very High Human Development Index.[30]

For KQ1 (screening effectiveness), eligible study designs included RCTs, CCTs and controlled observational studies (i.e., prospective and retrospective cohort, non-concurrent cohort (two or more groups identified on basis of common features at different time points), case-control, controlled before-after). For KQ2 (outcome valuation), we included any study where women were asked to balance the benefits and harms of screening and treatment for ASB and/or state their willingness to be screened and treated based on information or reflection on benefits and harms; surveys, experimental designs (e.g., contingent valuation), and qualitative research were eligible examples. For KQ3 (cost-effectiveness), we planned to look at any study comparing effects and costs (e.g., cost-effectiveness, cost-utility, cost-benefit), including modelling of effects and/or costs. For KQ4 (treatment), we planned to include RCTs. For KQ5 (test accuracy), we planned to use prospective and retrospective studies where a consecutive or random sample of participants receive both the index test(s) and reference standard, or where participants are randomized to different index tests but all receive the reference standard; assessment would generally be performed in a cross-sectional manner. For KQs 4 (treatment effectiveness) and 5 (test accuracy), we planned to use existing high-quality systematic review(s) if found.

For all KQs, case reports and case series (i.e., group of patient selected based on particular outcome) were excluded as were non-primary research (e.g. editorials, commentaries, opinion pieces). Conference

abstracts were not considered eligible for inclusion, but were planned to be used to identify full study reports and to assess the quality of evidence in relation to potential publication and reporting biases.

For all KQs, studies were included if they were published in English or French. No date restrictions were applied to publications. The inclusion and exclusion criteria for all KQs are detailed in Tables 1.1-1.5.

Table 1.1 - KQ1a, b: Benefits and harms of screening

Population	Asymptomatic pregnant women at any stage of pregnancy who are not at high risk for bacteriuria. Patient subgroups: women with kidney infection, urogenital anomalies, polycystic kidneys, recurrent urinary tract infection [UTI]), diabetes, sickle cell disease, socioeconomic status, ethnicity, urban/rural Exclude: studies *exclusively* including women with conditions that place them at substantially higher than average risk of bacteriuria (kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI, diabetes, and sickle cell disease), or with symptoms of UTI
Interventions	Any screening program, whereby there is an intent (i.e., clinical algorithm) for all pregnant women to receive a screening test with follow-up of screen-positive cases Screening subgroups/algorithms, including: urine collection method, frequency of testing, number of samples in one collection, criteria for a positive test (including number of consecutive positive specimens, bacterial colony count, and specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine screening is done for other conditions (e.g., proteinuria, glycosuria, Chlamydia), non-urine screening test (e.g., vaginal/rectal swab culture for group B streptococcus (GBS) testing)
Comparator	KQ1a: No screening program (but may include indicated testing and/or treatment upon development of symptoms) KQ1b: A different screening test or algorithm (see intervention subgroups)
Outcomes	*Benefits (reduced incidence for all)*: 1. maternal mortality (9)* 2. maternal sepsis (8) 3. pyelonephritis (7) 4. perinatal mortality (≥ 20 weeks' gestation [e.g., intrauterine demise, stillbirth, early neonatal death]) (9) 5. spontaneous abortion/pregnancy loss before 20 weeks' gestation (8) 6. neonatal sepsis (if not reported will include surrogate outcomes of acute respiratory distress syndrome [ARDS] or admission to neonatal intensive care unit [NICU]) (8) 7. preterm delivery (live fetus passed < 37 week's gestation) (7) 8. low birth weight (< 2500g) (6) *Harms (maternal and neonatal):* 1. serious adverse event(s)**associated with antibiotic treatment, *including but not limited to*: (7) a. anaphylaxis, b. thrombocytopenia, c. hemolytic anemia, d. fetal abnormalities; and, 2. non-serious adverse event(s) associated with treatment, *including but not limited to*: (4) a. alterations in vaginal/perineal microbiome (e.g., candidiasis, vaginitis), b. antibiotic-induced diarrhea, c. rash, d. vomiting e. neonatal thrush Exclude: screening for GBS near delivery or at time of rupture of membranes for the prevention or treatment of chorioamnionitis or neonatal GBS (without other outcomes of interest in list above)
Study Designs	Staged: RCTs, CCTs, controlled observational (i.e., prospective and retrospective cohorts, case-control, controlled before-after)
Language	English and French
Setting	Any primary care and clinical setting providing obstetric/antenatal care to pregnant women (e.g., obstetric and hospital outpatient clinics, prisons, remote stations, community centers, midwifery practices)

Timeframe	No publication date limits

*Bracketed numbers next to each outcome above refer to patient-important outcome ratings by CTFPHC members and patients as critical [7-9, out of 9] or important [4-6, out of 9] are included
ARDS: acute respiratory distress syndrome; CCT: controlled clinical trial; GBS: group B streptococcus; KQ: key question; NICU: neonatal intensive care unit; RCT: randomized controlled trial; UTI urinary tract infection
**Serious adverse event (experience) or reaction is any untoward medical occurrence that: a) results in death, b) is life-threatening, c) requires in-patient hospitalisation or prolongation of existing hospitalisation, d) results in persistent or significant disability/incapacity, or e) is a congenital anomaly/birth defect (Health Canada, 2011)

Table 1.2 - KQ2: Outcome valuation

Population	Asymptomatic pregnant women at any stage of pregnancy who are not at high risk for bacteriuria; will also accept asymptomatic women who are not pregnant if necessary Patient subgroups: women with kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI), diabetes, sickle cell disease, socioeconomic status, ethnicity, urban/rural Exclude: studies *exclusively* including women with conditions that place them at substantially higher than average risk of bacteriuria (kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI, diabetes, and sickle cell disease), or with symptoms of UTI
Interventions/Index Test	Any screening program or test, and any antibiotic; will accept studies on treatment for any bacterial condition in pregnancy Screening subgroups/algorithms, including: urine collection method, frequency of testing, criteria for a positive test (including number of consecutive positive specimens, bacteria colony count, and specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine *screening* is done for other conditions (e.g., proteinuria, glycosuria), non-urine screening test (e.g., vaginal/rectal swab culture for GBS testing)
Comparator/Reference Standard	Not applicable
Outcomes	Several possible outcomes (e.g., relative weight/utilities of benefit and harms; willingness to be screened based on relative value placed on benefits and harms of screening programs or treatment)
Study Designs	Qualitative, mixed methods, surveys/cross-sectional
Language	English and French
Setting	Any primary care and clinical setting providing obstetric/antenatal care to pregnant women (e.g., obstetric and hospital outpatient clinics, prisons, remote stations, community centers, midwifery practices)
Time frame	No publication date limits

GBS: group B streptococcus; KQ: key question; UTI: urinary tract infection

Table 1.3 - KQ3: Cost-effectiveness of screening

Population	Asymptomatic pregnant women at any stage of pregnancy who are not at high risk for bacteriuria. Patient subgroups: women with kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI, diabetes, sickle cell disease, socioeconomic status, ethnicity, urban/rural Exclude: studies *exclusively* including women with conditions that place them at substantially higher than average risk of bacteriuria (kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI, diabetes, and sickle cell disease), or with symptoms of UTI
Interventions/Index Test	Any screening program Screening subgroups/algorithms, including: urine collection method, frequency of testing, number of samples in one collection, criteria for a positive test (including number of consecutive positive specimens, bacterial colony count, specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine screening is done for other conditions (e.g., proteinuria, glycosuria, Chlamydia), non-urine screening test (e.g., vaginal/rectal swab culture for GBS testing)
Comparator/Reference Standard	No screening (but may include indicated testing and/or treatment upon development of symptoms), or a different screening algorithm (see intervention subgroups)
Outcomes	Cost per quality-adjusted life-years (cost per QALY), incremental cost-effectiveness ratio (ICER), net

	benefit/cost
Study Designs	Economic evaluations
Language	English and French
Setting	Any primary care and clinical setting providing obstetric/antenatal care to pregnant women (e.g., obstetric and hospital outpatient clinics, prisons, remote stations, community centers, midwifery practices); limited to countries rated as having very high Human Development Index[30]
Time frame	No publication date limits

GBS: group B streptococcus; ICER: incremental cost-effectiveness ratio; KQ: key question; QALY: quality-adjusted life-years; UTI: urinary tract infection

Table 1.4 - KQ4: Benefits and harms of treatment

Population	Asymptomatic pregnant women at any stage of pregnancy who are not at high risk for bacteriuria. Patient subgroups: women with kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI], diabetes, sickle cell disease, socioeconomic status, ethnicity, urban/rural Exclude: studies *exclusively* including women with conditions that place them at substantially higher than average risk of bacteriuria (kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI, diabetes, and sickle cell disease), or with symptoms of UTI
Interventions/Index Test	Any antibiotic Screening subgroups/algorithms, including: urine collection method, frequency of testing, number of samples in one collection, criteria for a positive test (including number of consecutive positive specimens, bacterial colony count, and specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine screening is done for other conditions (e.g., proteinuria, glycosuria, Chlamydia), non-urine screening test (e.g., vaginal/rectal swab culture for GBS testing)
Comparator/Reference Standard	No treatment or placebo
Outcomes*	*Benefits (reduced incidence for all)*: 1. maternal mortality (9)* 2. maternal sepsis (8) 3. pyelonephritis (7) 4. perinatal mortality (≥ 20 weeks' gestation [e.g., intrauterine demise, stillbirth, early neonatal death]) (9) 5. spontaneous abortion/pregnancy loss before 20 weeks' gestation (8) 6. neonatal sepsis (if not reported will include surrogate outcomes of ARDS or admission to NICU) (8) 7. preterm delivery (live fetus passed < 37 week's gestation) (7) 8. low birth weight (< 2500g) (6) *Harms (maternal and neonatal):* 1. serious adverse event(s)** associated with antibiotic treatment, *including but not limited to*: (7) a. anaphylaxis, b. thrombocytopenia, c. hemolytic anemia, d. fetal abnormalities; and, 2. non-serious adverse event(s) associated with treatment, *including but not limited to*: (4) a. alterations in vaginal/perineal microbiome (e.g., candidiasis, vaginitis), b. antibiotic-induced diarrhea, c. rash, d. vomiting e. neonatal thrush Exclude: screening for group B streptococcus near delivery or at time of rupture of membranes for the prevention or treatment of chorioamnionitis or neonatal GBS (without other outcomes of interest listed above)
Study Designs	RCTs (or systematic review(s))
Language	English and French
Setting	Any primary care and clinical setting providing obstetric/antenatal care to pregnant women (e.g., obstetric or hospital outpatient clinics, prisons, remote stations, community centers, midwifery practices)

Time frame	No publication date limits

*Bracketed numbers next to each outcome above refer to patient-important outcome ratings by CTFPHC members and patients as critical [7-9, out of 9] or important [4-6, out of 9] are included
ARDS: acute respiratory distress syndrome; GBS: group B streptococcus; KQ: key question; NICU: neonatal intensive care unit; RCT: randomized controlled trial; UTI: urinary tract infection
**Serious adverse event (experience) or reaction is any untoward medical occurrence that: a) results in death, b) is life-threatening, c) requires in-patient hospitalisation or prolongation of existing hospitalisation, d) results in persistent or significant disability/incapacity, or e) is a congenital anomaly/birth defect (Health Canada, 2011)

Table 1.5 - KQ5: Accuracy of screening tests

Population	Asymptomatic pregnant women at any stage of pregnancy who are not at high risk for bacteriuria. Patient subgroups: women with kidney infection, urogenital anomalies, polycystic kidneys, recurrent UTI), diabetes, sickle cell disease, socioeconomic status, ethnicity, urban/rural Exclude: studies *exclusively* including women with conditions that place them at substantially higher than average risk of bacteriuria (kidney infection, urogenital anomalies, polycystic kidneys, recurrent urinary UTI, diabetes, and sickle cell disease), or with symptoms of UTI
Interventions/Index Test	Any index test (rapid point-of care tests) Screening subgroups/algorithm, including: urine collection method, frequency of testing, number of samples in one collection, criteria for a positive test (including number of consecutive positive specimens, bacterial colony count, and specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine screening is done for other conditions (e.g., proteinuria, glycosuria, Chlamydia), non-urine screening test (e.g., vaginal/rectal swab culture for GBS testing)
Comparator/Reference Standard	A urine culture Screening subgroups/algorithm, including: urine collection method, frequency of testing, number of samples in one collection, criteria for a positive test (including number of consecutive positive specimens, bacterial colony count, and specified organism(s)), follow-up testing during pregnancy, timing during pregnancy Exclude: urine screening is done for other conditions (e.g., proteinuria, glycosuria, Chlamydia), non-urine screening test (e.g., vaginal/rectal swab culture for GBS testing)
Outcomes	Sensitivity, specificity, false positives, true positive, false negatives, true negatives, positive and negative likelihood ratios, prevalence/pre-test probability (true positive + false positive)/total number of people)
Study Designs	Prospective and retrospective studies where a consecutive or random sample of participants receive both the index test(s) and the reference standard, or where participants are randomized to different index tests but all receive the reference standard, and assessment in a cross-sectional manner Exclude: case-control studies and studies with longitudinal assessment of the reference standard
Language	English and French

GBS: group B streptococcus; KQ: key question; UTI: urinary tract infection

Data abstraction and risk of bias assessments

One reviewer independently extracted data, and another reviewer verified all data from each included study on study design, country of origin, sample size, population and subgroup(s), intervention and comparator, setting, and outcomes of interest. Disagreements were resolved through discussion or a third reviewer. For each KQ, a narrative summary with accompanying tables was produced for all studies including design, country of origin, setting, populations and subgroups, tests, treatment and comparators, and outcome measures. For studies with multiple publications, we extracted data from the primary source and added data reported in associated publications as applicable. We contacted authors of included studies via email (with follow-up as necessary) for clarification of study details (i.e., interventions, outcomes and numerical data). For KQ4 (treatment), we extracted data from the eligible systematic review on its scope

(PICOTS), and for the individual studies with specifics related to the population (size and characteristics), outcomes evaluated (including definitions and timing of assessment), and risk of bias (ROB) (by domain/construct). We conducted data verification on 10% of included studies for quality assurance, and also examined the primary studies for additional participant characteristics and outcome details relevant to the current review.

When using individual studies, we recorded intention-to-treat results whenever possible. For dichotomous outcomes, we reported counts or proportions, and sample size, by study arm. For dichotomous data on harms, each adverse event (AE) was counted as if it represented a unique individual. Only numerical data for AEs were extracted; no assumptions were made on lack or presence of an AE when this was not reported. For patient and intervention subgroups (see Tables 1.1-1.5), we collected data for performing our own subgroup analyses (e.g., stratified analysis, meta-regression) based on study-level data.

Two reviewers independently assessed the ROB of each included study (KQs 1, 2 and 4), with disagreements resolved through discussion or third-party consultation to reach consensus. The results for each study and across studies were reported for each domain and for an overall quality score. For KQ1 (screening effectiveness), all controlled observational studies were appraised using the Newcastle-Ottawa Quality Assessment Scale.[31] The scale comprises eight items that evaluate three domains: sample selection, comparability of cohorts, and assessment of outcomes. Each that is adequately addressed is awarded one star (up to two stars may be awarded for comparability), and the overall score is calculated by tallying the stars. We considered a total score of 7 to 9 to indicate low ROB; 4 to 6 to indicate unclear ROB; and, 3 or lower to indicate high ROB. We included a separate assessment for reporting bias due to suspected selective outcome reporting. For KQ2 (outcome valuation), all cross-sectional studies were appraised using the tool developed by the Center for Evidence-based Management.[32] For KQ4 (treatment effectiveness), all RCTs and CCTs were appraised using the Cochrane Risk of Bias tool.[33] This tool consists of six domains (sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and "other" sources of bias [e.g., baseline imbalances between groups]) and a categorization of the overall ROB. The overall assessment is based on the responses to individual domains. If one or more individual domains were assessed as having a high ROB, the overall score was rated as high ROB. The overall ROB was considered low only if all components were rated as having a low ROB. The ROB for all other studies was rated as unclear. Information was collected for each study on the source of funding. We conducted assessments on 10% of included studies in the Cochrane review[4] and found that our methods for assessing ROB were somewhat different: we considered objective outcomes to be at lower ROB than subjective outcomes when assessing the blinding domains; we considered 10-30% loss to follow-up as unclear ROB (rather than high ROB) for incomplete reporting if there appeared to be no imbalances between groups or reasons were provided; we used a default of low ROB (rather than unclear ROB) for selective reporting when this was not detected or not highly suspected; we revised the "other bias" domain to low if no additional bias was detected apart from comparability between groups at baseline; and, we updated the overall ROB to align with guidance (see Chapter 8, Section 8.5.d in the *Cochrane Handbook for Systematic Reviews of Interventions*[34]) as stated above. To assist with outcome reporting bias assessments, we searched for study protocols and considered reporting by similar studies included.

Data Analysis

Key Question 1 (screening effectiveness)

For pair-wise meta-analysis in KQ1 (screening effectiveness), we employed a random effects model. For dichotomous outcomes, we reported relative risks (RR) using the DerSimonian and Laird random effects model with Mantel-Haenszel method with corresponding 95% CIs. The decision to pool studies was not based on the statistical heterogeneity (I^2 statistics were reported), but rather on interpretation of the clinical and methodological differences between studies. We conducted sensitivity analyses when substantial heterogeneity was found and if appropriate (e.g., in the presence of studies with outlying effect sizes, for studies rated as high ROB in some domains such as incomplete outcome data [<80 percent] or lack of allocation concealment). Heterogeneity was also examined for important patient and intervention variables (Tables 1.1-1.5). We did not perform a funnel plot test or Egger's test to detect small-study bias as there were fewer than eight studies included for KQ1.

Key Question 2 (outcome valuation)

For KQ2 (outcome valuation) results were narratively described, summarizing themes across studies.

Key Question 4 (treatment effectiveness)

For stage 2 examining evidence related to treatment effectiveness, meta-analyses were recalculated with the addition of one new study identified in the search update as well as for subgroups of interest (e.g., for population risk factors, screening/treatment characteristics, ROB and study design). Although we intended to restrict primary studies to RCTs in our protocol, we included and pooled results from CCTs in order to integrate all study designs (RCTs and CCTs) included in the Cochrane systematic review. Where there were at least eight studies in a meta-analysis, we assessed small-study bias both visually using the funnel plot and quantitatively using Egger's test.[35]

Key Question 5 (test accuracy)

For this report, we did not examine evidence related to diagnostic test accuracy; this may be conducted and produced in a separate report upon consideration by the CTFPHC based on guideline recommendations following the evidence review for KQs 1 to 4.

Subgroup Analyses

We performed subgroup analyses using study-level data, when possible, using Cochran's Q (α=0.05) to detect statistical heterogeneity and the I^2 statistic was used to quantify the magnitude of statistical heterogeneity between studies. When determining whether entire studies fell into a particular subgroup category (e.g., recurrent UTI), we considered ≥80% of the study population meeting the criteria as sufficient. We planned to conduct regression analyses for categorical variables (e.g., history of recurrent UTI), when there were at least three studies for each category level; however, this was not performed due to limited study reporting, variations in size and heterogeneity of effect sizes, and/or insufficient number of studies for each category comprising a subgroup. We did perform some stratification of meta-analyses based on our planned key subgroups, and a minimum of 2 of the following criteria was used to determine

credibility of subgroup investigations: a) visual inspection of forest plot showing a meaningful difference between effect estimates (e.g., clinical decision making on the intervention would differ for each subgroup), b) a reduction in the heterogeneity (I^2) for each subgroup from the original meta-analysis, and c) a statistically significant between-group test for differences.

For outcomes that showed significant effects, we calculated absolute risk reduction (ARR) and number needed to screen (NNS) or number needed to treat (NNT), as applicable, and reported these in a table summarizing overview of results. The values for NNS or NNT were calculated using the absolute numbers presented in the GRADE tables estimated using the control group event rate and RR with the 95% confidence interval (CI) obtained from the meta-analysis (see Chapter 12, Section 12.5.4.2 in the *Cochrane Handbook for Systematic Reviews of Interventions*[34]).

Analyses were performed using Review Manager Version 5.3 and GRADEpro software packages. Whenever studies did not provide data for pooling, the results were described narratively.

Assessment of Overall Quality of Evidence Using GRADE

Two reviewers independently assessed the quality of the body of evidence or confidence in the effect for each outcome of interest using the GRADE methodology.[36, 37] Discrepancies were resolved through discussion or consultation with a third reviewer, to reach consensus. Assessments were entered into the GRADEpro software[38] and summarized in GRADE Evidence Profiles (EP) and Summary of Findings (SOF) tables.[39] Footnotes to the tables provided explanations for all decisions.

The general approach is outlined here although methods align with GRADE guidance.[36, 37] For evidence on benefits and harms of screening (KQ1) and treatment (KQ4), as a starting point the quality was assigned as high for evidence from RCTs and low for evidence from observational studies. Thereafter, we examined and potentially downgraded the quality based on five core domains: study limitations/ROB, inconsistency, indirectness, imprecision, and publication/reporting bias.

For the *study limitations (ROB)* domain RCTs and CCTs were downgraded one or two levels depending on the proportion of trials (e.g., one very large trial may outweigh two very small trials) assessed as having high ROB for the particular outcome under consideration.[40] Evidence from observational studies was downgraded when most studies had moderate or high ROB. For *inconsistency* we assessed the magnitude of the effects of the included studies (e.g., inconsistent when lack of overlap in 95% CIs for some studies).[41] *Indirectness* of the evidence was based on evaluating the relevance of the studies' PICOTS compared to those of the current review. We assessed *imprecision* on the basis of Optimal Information Size (OIS) and a relative risk of under 0.75 to over 1.25.[42] If the OIS criterion was met and the pooled 95% CI excluded no effect (i.e. CI around RR excluded 1.0), we did not downgrade for imprecision. If the OIS criterion was met and the 95% CI crossed no effect, we downgraded for precision if one or more of the limits of the CI crossed a RR of 0.75 or 1.25 (indicating a possibly important benefit or harm), which suggested lower certainty of no effect. *Reporting bias* (suspected or undetected) was evaluated with respect to publication bias. When considering the need to balance patient-important benefits and harms for making a screening recommendation, the CTFPHC may choose to use a different

approach than ours to assess the imprecision domain, taking into account different baseline risks applicable to specific outcomes and applying clinically meaningful decision thresholds.

Interpretation of Results

We chose to use standard wording to describe our interpretations of the findings and quality of evidence. For findings supported by high, moderate, low, and very low quality evidence (for which we have similar confidence in the results) we use "will", "probably/likely", "may/appears to", and "not known/very uncertain", respectively, in our textual descriptions of the results.

Chapter 3. Results

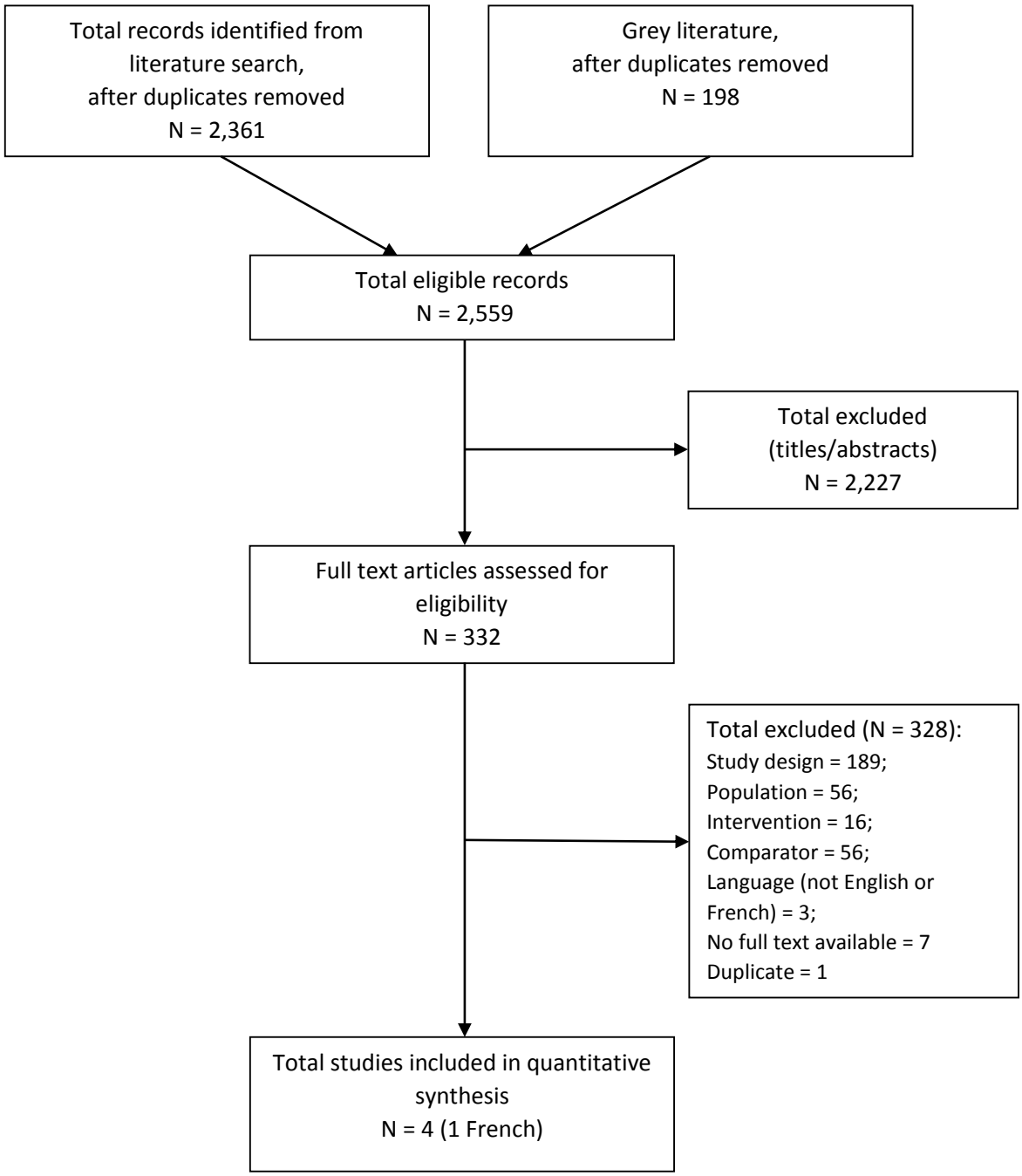
Summary of Studies for Review

Key Questions (KQs) 1a, b (screening effectiveness)

The total number of records identified from the literature search, including grey literature, was 2,559. After screening of titles and abstracts, 2,227 were excluded. Of the 332 papers that underwent full text screening, 327 were initially excluded resulting in five studies for inclusion. However, a post-hoc decision was made based on input from the Canadian Task Force on Preventive Health Care (CTFPHC) to clarify criteria for the intervention/comparator to examine screening programs, thereby removing one case-control study (Friedman et al,[249] Appendix 10) where the exposure may have been defined by whether or not the women received a screening culture, rather than by whether or not there was the intent to do so; the results reflect an intent to screen all women with some for some reason not receiving the culture, which would also be reflected in the screening arm in other studies.

The search results and study flow and selection are presented in Figure 2.

Figure 2. PRISMA flow diagram of study selection for KQ1 (screening effectiveness)



Characteristics of included studies relevant to KQ1 are summarized in Appendix 1. Detailed study information is reported in Appendix 3.

A total of four studies (7,611 women) examined screening effectiveness for asymptomatic bacteriuria (ASB). One study[43] was published in French. All four studies were non-concurrent cohort studies, comparing outcomes for women before and after introduction of a screening program. The studies were each conducted in France,[43] Spain,[44] Turkey,[45] and the United States (US).[46] None of the studies[43-46] provided details on funding. One enrolled women at a hospital[43], one at a hospital-based midwifery clinic[46] and two at an obstetrics clinic.[44, 45]

Among the two studies reporting on the proportion of women with gestational diabetes, Rhode et al[46] reported a relatively high rate of gestational diabetes in their group receiving frequent screening (9% [81 out of 933]) compared with that receiving screening at first prenatal visit only (4% [42 out of 1019]), and another[45] reported approximately 3% of women with gestational diabetes mellitus range (3.8% [7 out of 186] in the screening group compared with 2.7% [5 out of 186] in the no-screening group).

Two studies[44, 45] reported gestational age criteria for including women in their study, one at <25 weeks of gestation and the other at <32 weeks of gestation. Two of the four studies[43, 45] specified criteria ($\geq 10^5$ CFU/mL) as positive for ASB while this was not reported in the other two studies.[44, 46] The study by Gérard et al[43] compared outcomes for women in the 10-month period (March to December 1978), when women were only tested if they had clinical signs, before introducing a screening program (January to October 1979) where women were screened at multiple intervals (3, 5, 7 and 9 months). The study by Gratacós et al[44] also compared outcomes of women before (January 1987 to December 1990) and after (January 1991 to December 1992) introduction of a screening program for ASB. Rhode et al[46] compared women who were screened at every prenatal visit, before August 15, 2002 ("routine screening group") with women who were screened at the first prenatal visit only ("indicated screening group"). The study by Uncu et al[45] compared pregnant women who delivered in the clinic but were not screened for ASB (prior to June 1998), with outcomes of women who were routinely screened for ASB (June 1998 to January 1999).

With regard to treatment protocols, two studies[43, 44] reported treating screen-positive women based on antibiotic sensitivity testing, with one study[43] only specifying treatment was provided at the discretion of the treating physician, and the other study[44] detailing 7 days of antibiotics administered 1 to 2 weeks after a second culture was obtained with additional antibiotics 1 to 4 weeks after treatment, and again prior to delivery (as well as additional antibacterial therapy when repeat cultures were positive for bacteriuria). One study[45] reported treating women with antibiotics for 7 to 10 days followed by 7 days of antibiotics for persistent or recurrent bacteriuria. One study[44] did not specify a treatment protocol. One study[45] reported follow-up of women with cultures one week after treatment (test-of-cure), and another study[44] reported re-testing women with urine cultures twice to determine presence of persistent bacteriuria; two studies[43, 46] did not report whether women were followed up after treatment to determine test-of-cure.

Outcomes were not uniformly defined among studies. Pyelonephritis was defined as "acute pyelonephritis" by two studies[43, 44] with a combination of symptoms including fever, lumbar or flank pain, tenderness in costovertebral angle, dysuria, and at least one positive urine culture. "Pyelonephritis" was not specified by criteria in two studies;[45, 46] however, it was clearly differentiated from "ASB", "cystitis" and "undetermined urinary tract infection (UTI)" in the Rhode study.[46]

Two studies reported on perinatal mortality: Rhode et al[46] used ≥31 weeks of gestation, and Uncu et al[45] defined perinatal mortality as no fetal cardiac activity on ultrasound after 20 weeks of gestation.

Gérard et al[43] reported spontaneous abortion, defined as ≤28 weeks of gestation; since this was distinguished from perinatal mortality, it was included in the analysis for this outcome.

All three studies[43-45] that reported preterm delivery used <37 weeks of gestation as the criterion. For the study by Uncu et al[45], we were unable to confirm with the authors on eligibility of criteria (i.e., whether women were at risk of, or actual cases of, preterm delivery) and the data for preterm delivery; however, removal of the data would not change overall conclusions for this outcome (see Results below for KQ1a).

One study[45] reported harms of screening (fetal abnormalities) without a specific definition.

No study reported on maternal mortality, maternal sepsis, neonatal sepsis or low birthweight.

Most studies were of unclear risk of bias (ROB) (rated 6 out of 9 in 3 cohort studies[43-45]) with one study of low ROB that rated 8 out of 9.[46] None of the studies[43-46] reported that pyelonephritis was not present in pregnant women at the outset of the study. Three studies[43-45] did not demonstrate comparability of baseline characteristics between groups. All of the studies[43, 44, 46] except one[45] were suspected of selective outcome reporting due to lack of reporting for neonatal outcomes (e.g., spontaneous abortion, perinatal mortality, preterm delivery and fetal abnormalities) despite following women to delivery. Methodological quality assessments for studies relevant to KQ1 are summarized in Table; detailed assessments for each study are reported in Appendix 6.

Table 4. Summary of methodological quality[a] - KQ1 a & b (screening effectiveness)

First Author, Year	Selection					Comparability		Outcome				Total Score[a] (max 9)	Selective Outcome Reporting[b]
	Representativeness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome not present at start of study (pyelonephritis/ other outcomes)	Total	Comparability of cohorts	Total	Assessment of outcome	Adequacy of length of follow-up	Adequacy of follow-up of cohorts	Total		
Gérard, 1983[43]	1	1	0	0/1	3	0	0	1	1	1	3	6	Suspected[c]
Gratacós, 1994[44]	1	1	0	0/1	3	0	0	1	1	1	3	6	Suspected[d]
Rhode, 2007[46]	1	1	1	0/1	4	1	1	1	1	1	3	8	Suspected[e]
Uncu, 2002[45]	1	1	1	0/1	4	0	0	1	1	0	2	6	Not suspected[f]

[a]Assessed using the Newcastle-Ottawa Quality Assessment Scale[31]
[b]Assessed due to concern regarding reporting bias in the studies, but assessment not included in the total score
[c]Did not report on fetal abnormalities
[d]Did not report on spontaneous abortion, perinatal mortality, preterm delivery or fetal abnormalities
[e]Did not report on spontaneous abortion, perinatal mortality, or fetal abnormalities
[f]Reported on all outcomes, including fetal death >20 weeks of gestation (eligible for perinatal mortality)

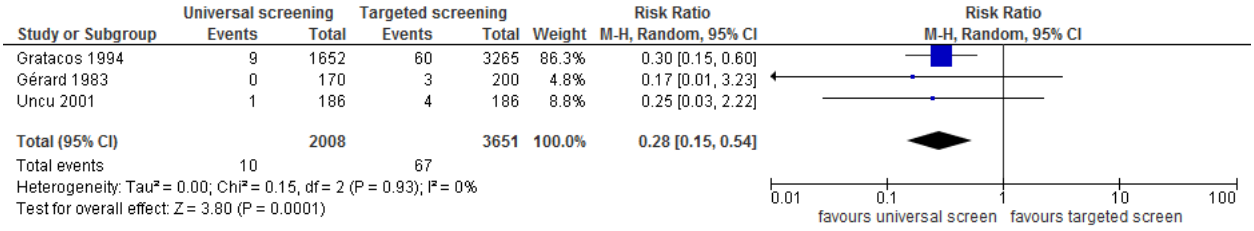
KQ1a: What are the benefits and harms of screening compared with no screening for asymptomatic bacteriuria in pregnancy? Are there subgroup differences for patient characteristics (e.g., socioeconomic status [SES])?

Three studies (non-concurrent cohort) of unclear ROB[43-45] and a combined sample of 5,659 pregnant women addressed the benefits and harms of screening compared with no screening. The results are summarized below; for additional details see Evidence Set 1 for GRADE EP and SOF tables and forest plots.

Pyelonephritis

Three studies[43-45] of unclear ROB (5,659 women) found a statistically significant difference for screening compared to no screening on the outcome of pyelonephritis (RR 0.28; 95% CI 0.15, 0.54; I^2=0%; ARR 1.3%; NNS 77, 95% CI 65, 121). The overall quality for this body of observational evidence was rated as very low due to downgrading for study design and ROB.

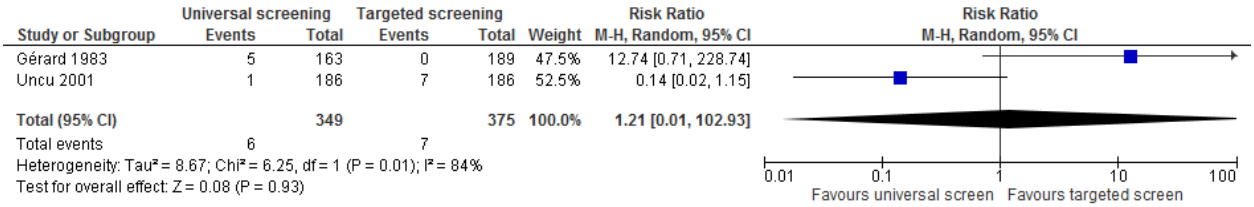
1.1 Pyelonephritis



Perinatal mortality

Two studies[43, 45] (724 women) with unclear ROB but suspected reporting bias[43] found no significant difference (RR 1.21, 95% CI 0.01, 102.93, I^2=84%) in perinatal mortality. The quality of this body of evidence was rated as very low due to downgrading for study design, ROB, inconsistency, indirectness, and imprecision.

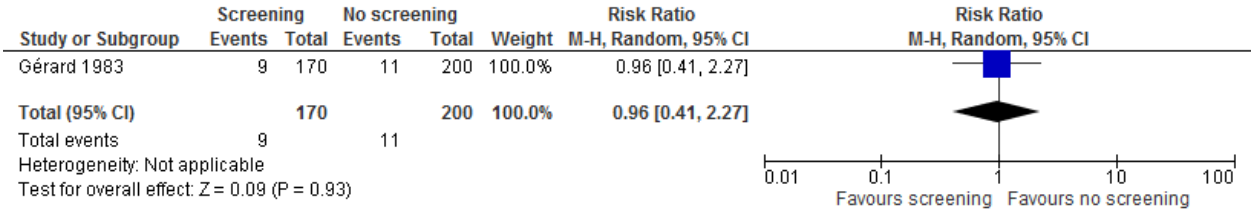
1.2 Perinatal mortality



Spontaneous abortion

One study of 370 women[43] with unclear ROB but suspected reporting bias found no significant difference (RR 0.96, 95% CI 0.41, 2.27) in spontaneous abortion at ≤28 weeks of gestation. This body of evidence was rated as very low due to concerns with study design, ROB, inconsistency and imprecision.

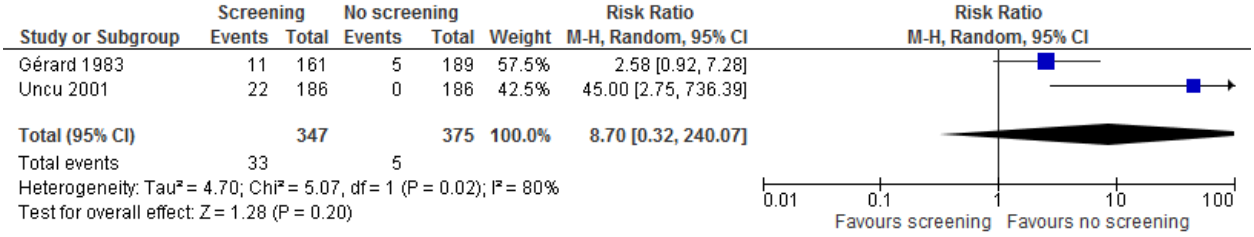
1.3 Spontaneous abortion



Preterm delivery

Two studies[43, 45] (722 women) with unclear ROB but suspected reporting bias[43] found no significant difference (RR 8.70, 95% CI 0.32, 240.07; I^2=80%) in preterm delivery before 37 weeks of gestation; this body of evidence was rated as very low due to downgrading for ROB and imprecision.

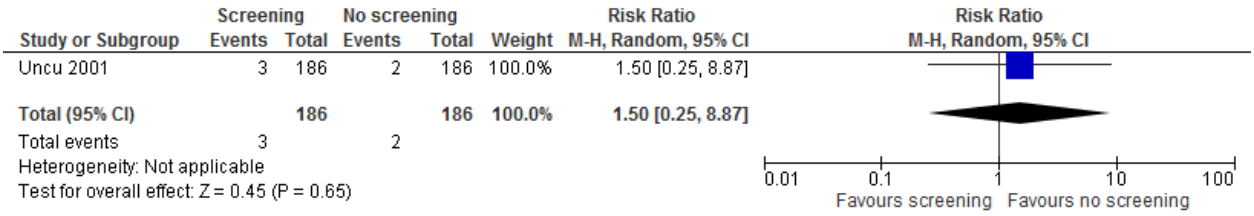
1.4 Preterm delivery



Fetal abnormalities (harm)

One study[45] (372 women) with unclear ROB found no significant difference (RR 1.50, 95% CI 0.25, 8.87) in fetal abnormalities (harm); this body of evidence was rated as very low due to downgrading for study design, ROB, inconsistency, and imprecision.

1.5 Fetal abnormalities (harm)



Subgroup analyses

We did not perform subgroup analyses due to insufficient number of studies contributing to each category comprising our *a priori* subgroups.

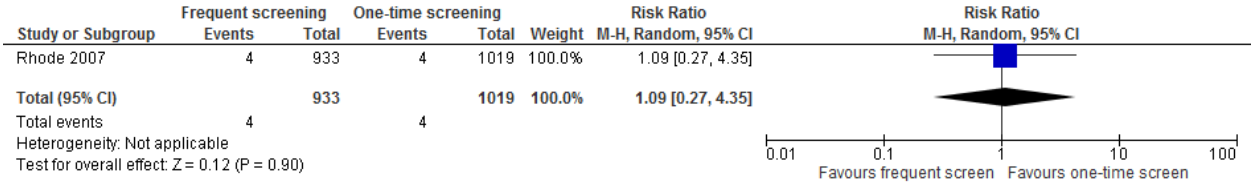
KQ1b. What are the benefits and harms of screening programs using different screening methods or algorithms for asymptomatic bacteriuria in pregnancy?

One non-concurrent cohort study[46] (1,952 women) with low ROB compared screening at all prenatal visits with screening at first prenatal visit only. This study only reported on pyelonephritis and preterm delivery. See Evidence Set 2 for GRADE EP and SOF tables and forest plots.

Pyelonephritis

No significant difference was found for pyelonephritis (RR 1.09; 95% CI 0.27, 4.35); this evidence was rated as very low due to downgrading for study design, ROB, inconsistency, indirectness and imprecision.

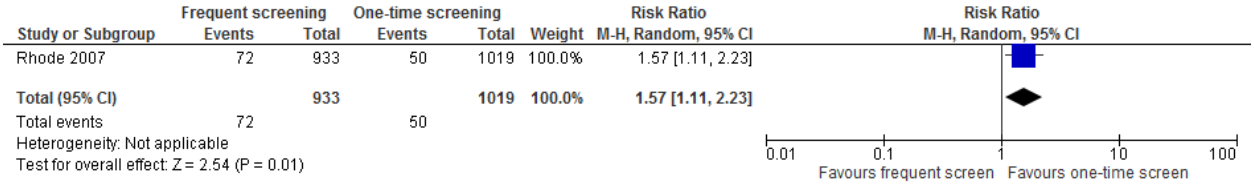
2.1 Pyelonephritis



Preterm delivery

A significant difference was found for preterm delivery (RR 1.57; 95% CI 1.11, 2.23) with more preterm deliveries among the group that was screened at all prenatal visits. The study authors did not present a possible hypothesis to explain this result. This body of evidence was rated as very low due to downgrading for study design, ROB, inconsistency, indirectness, and imprecision.

2.2 Preterm delivery

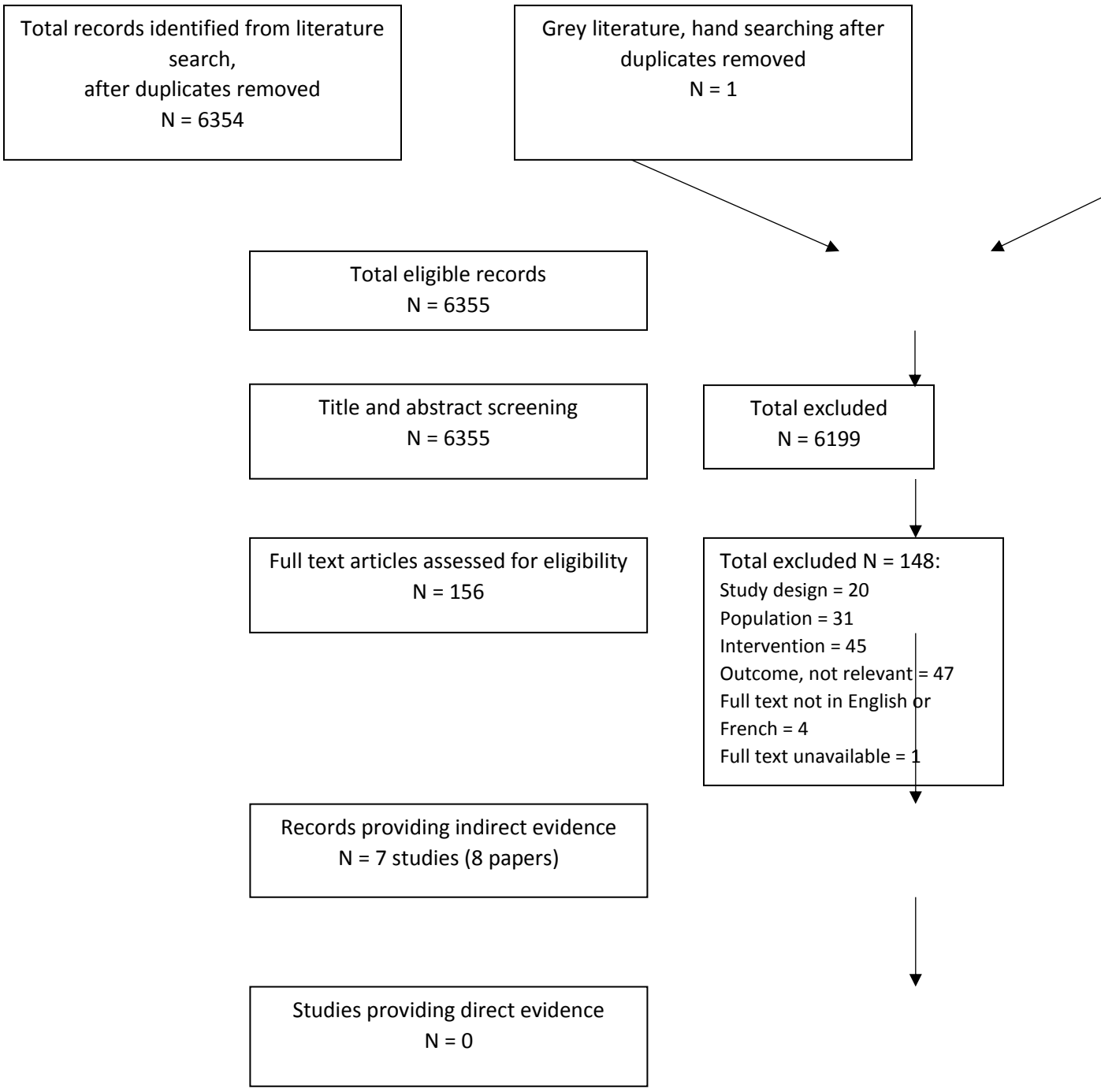


Key Question 2 (outcome valuation)

The total number of records identified from database searching, grey literature and hand searching was 6,355; this included searching included studies from KQ1 (screening effectiveness) and KQ4 (treatment effectiveness) for any information regarding women's valuation of benefits and harms of screening and treatment for ASB. After primary screening of titles and abstracts 6,199 studies were excluded. Of the 156 papers that underwent full text screening 20 were excluded due to study design, 31 due to population, 45 due to the intervention, 47 due to the reported outcomes, 1 was not available in full text and 4 papers did not have full text available in either English or French. No studies were identified that answered the question as to how women weigh the benefits and harms of screening and treatment of asymptomatic bacteriuria in pregnancy or how their valuation of benefits and harms inform their decisions to undergo treatment; however, eight papers (seven studies) focusing only on the harms of antibiotic treatment

(considered indirect evidence) were identified and are analyzed here. The search and study flow and selection are presented in Figure 3.

Figure 3. PRISMA flow diagram of study selection for KQ2 (outcome valuation)



From the eight papers providing indirect evidence, two were from the same study, but reported on different outcomes.[47, 48] Six of the papers were cross-sectional surveys while one paper was a cross-sectional study using a visual analogue scale.[49] One study was a multicenter screening cohort of pregnant women with an embedded randomised controlled trial (RCT) of antibiotic treatment for women with significant bacteriuria; cross-sectional findings from the women eligible for treatment are used for this KQ.[50] The sample sizes ranged from 144 to 4999 participants with three papers[47, 51, 52] including more than 1000 participants. Six papers reported age ranges from 15 to 45 years.[47, 48, 50-53] Three studies provide information on drug utilization opinions,[48, 50, 53] while five papers (four studies) provide information on perceptions of teratogenic risk.[47, 49, 51, 52, 54] Additional characteristics of these papers are included in Appendix 4.

While all seven studies addressed a focused research question and used a sample representative of this study question, their reported sampling methods could potentially introduce bias and only one of the studies[47] fully accounted for confounding factors through statistical analysis. None of the papers reported that their sample size was based on pre-study considerations while only two papers used survey questions that were considered valid and reliable. The summary of methodological quality for KQ2 are reported in Table 5; detailed study quality assessments are reported in Appendix 7.

Table 5. Summary of methodological quality[a] – KQ2 (outcome valuation)

First Author, Year	Did the study address a clearly focused question / issue?	Is the research method (study design) appropriate for answering the research question?	Is the method of selection of the subjects clearly described?	Could the way the sample was obtained introduce bias?	Was the sample of subjects representative of the population to which the findings will be referred?	Was the sample size based on pre-study considerations of statistical power?	Was a satisfactory response rate achieved?	Are the measurements (questionnaires) likely to be valid and reliable?	Was the statistical significance assessed?	Are confidence intervals given for the main results?	Could there be confounding factors that haven't been accounted for?	Can the results be applied to your organization?
Butters, 1990[48]	1	1	1	2	1	2	1	2	1	3	1	1
Kazemier, 2015[50]	2	2	1	3	1	2	2	2	3	3	2	3
Lupattelli, 2014[47]	1	1	1	1	1	2	2	2	1	1	3	1
Mashayekhi, 2009[54]	1	1	1	1	1	2	2	2	1	3	1	1
Nordeng, 2010[51]	1	1	1	1	1	2	2	2	1	3	2	1
Sanz, 2000[49]	1	1	3	2	1	2	2	2	1	3	1	1
Sharma, 2006[53]	1	1	3	2	1	2	2	1	1	3	1	1
Twigg, 2016[52]	1	1	1	1	1	2	2	1	1	3	1	1

[a]Assessed using a tool developed by the Center for Evidence-based Management[32] for cross-sectional studies (surveys)
1=Yes, 2=Can't Tell, 3=No

KQ2a. How do women weigh the benefits and harms of screening and treatment for asymptomatic bacteriuria in pregnancy?

No study directly examined how women weigh the benefits and harms of screening and treatment for asymptomatic bacteriuria.

KQ2b. How do women's valuation of benefits and harms of screening and treatment inform their decisions to undergo screening?

No studies were identified that examined women's valuation of benefits and harms of screening and treatment to inform their decisions to undergo screening.

The included studies herein reported on women's opinions of antibiotic use and their perception of teratogenic risk related to antibiotics or medication for UTIs.

Opinions on Antibiotic Use During Pregnancy

A questionnaire among recently postpartum women (n=514) in Scotland reported that 49% of participants said they would take a doctor-prescribed antibiotic during pregnancy while 48% said they would not; 3% did not respond to the question.[48] In contrast, a study of 395 pregnant women in North India found that 6.2% believed antibiotics should be used in pregnancy while 46.9% felt they should not be used; 46.9% did not provide a response.[53] The cohort study of screening with embedded treatment trial[50] reported over 61% (n=255) of women who screened positive for ASB opted out of participation in the trial because they did not want to receive antibiotics during pregnancy for an asymptomatic condition.

Perception of Teratogenic Risk

One study addressed the risk perception relating to medication treatment for pregnant women with UTIs. In a web-based study in the United Kingdom (UK) (n=1120), Twigg et al[52] reported that women who were taking medication for a UTI perceived the risk of overuse and harm of medication to be lower and the benefits to be higher than women who were not taking medication (Overuse [mean(SD)]: 11.5 (2.8) vs. 12.6 (2.7), p=0.006; Harm [mean(SD)]: 9.3 (2.7) vs. 10.4 (2.9), p=0.014; Benefit [mean(SD)]: 16.3 (2.2) vs. 14.9 (2.3), p<0.001). Nordeng et al[51] also reported a significant difference in mean risk perception scores for penicillin use during pregnancy between those using the drug and those who were not (n=1793; 3.0 vs. 4.3, p<0.001, on a scale of 0 to 10).

Throughout the included studies there were inconsistencies in opinions of the teratogenic risk perception of antibiotics. An internet study of 4,999 pregnant women across 18 countries reported that 96.2% of participants felt penicillin antibiotics posed a teratogenic risk,[47] whereas an Iranian study (n=400) reported that up to 1.3% of pregnant women felt antibiotics, including penicillin, ampicillin, amoxicillin, metronidazole and cephalosporin, were unsafe for the mother, while 31.3% to 36.8% felt these antibiotics were unsafe for the fetus and 4.5% to 10.0% felt these antibiotics were unsafe for both.[54]

One study using a visual analogue scale also revealed differences in perception of teratogenic risk of some antibiotics between pregnant and non-pregnant women. The authors reported that pregnant women (n=81) have a significantly lower perception of the risk of malformations than non-pregnant women (n=63) for erythromycin (38.7% vs. 55.6%, p<0.001) while there was no significant difference in their risk perception of amoxicillin (40.4% vs. 49.3%, p>0.05).[49]

Key Question 3. What is the cost-effectiveness of screening for asymptomatic bacteriuria in pregnancy?

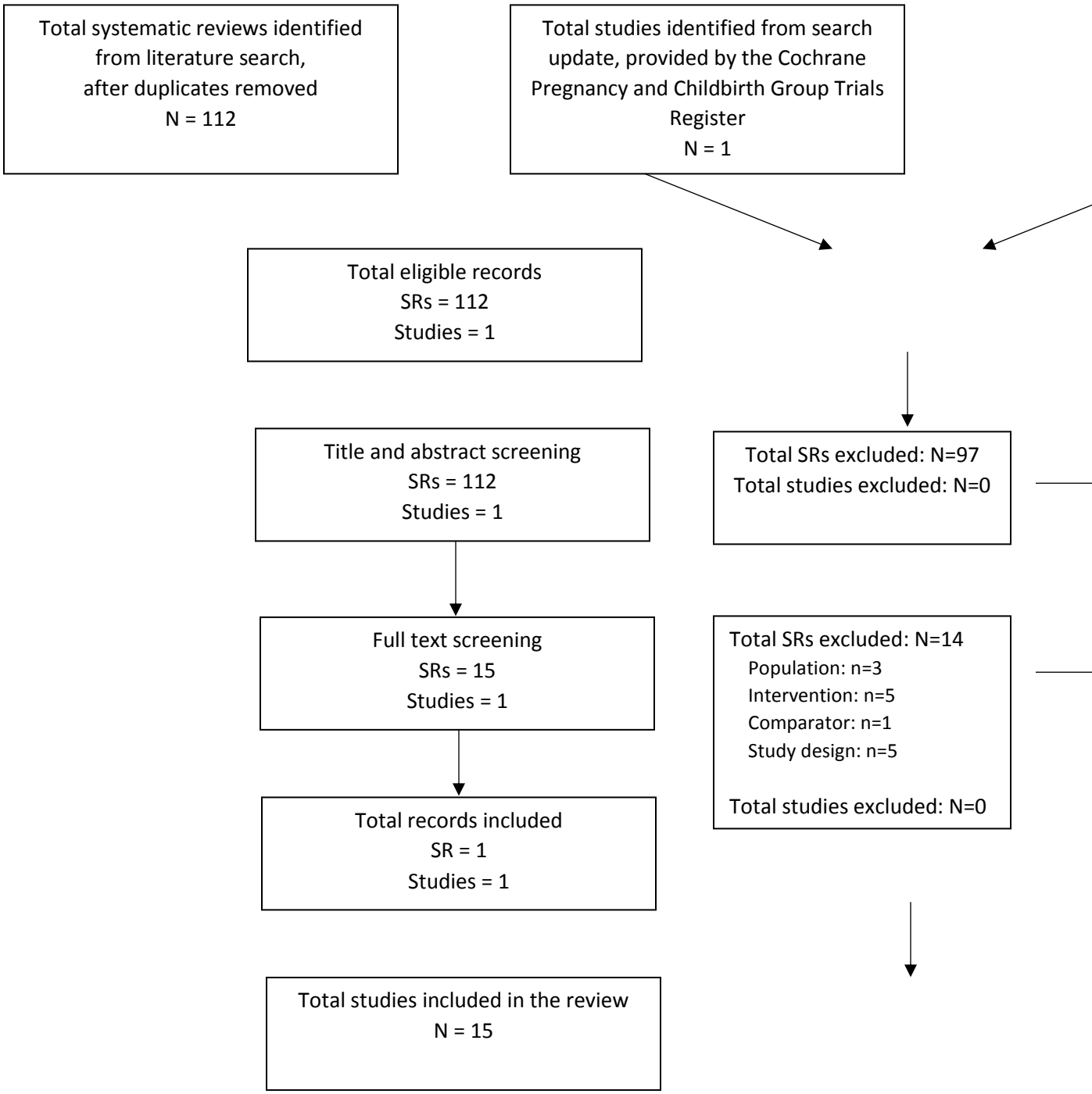
Evidence on screening effectiveness (KQs 1a and 1b) was very low quality, therefore we did not review studies on cost-effectiveness of screening programs which would have relied on similar evidence.

Key Question 4. What are the benefits and harms of antibiotic treatment compared with placebo or no treatment for asymptomatic bacteriuria in pregnancy?

The total number of records identified from the search for systematic reviews was 112. After screening of titles and abstracts, 97 were excluded. Of the 15 reviews that underwent full text screening, 14 were excluded resulting in one review for inclusion. Contact with the information specialist of the Cochrane Pregnancy and Childbirth Group's Trials Register confirmed the identification of only one study

(Kazemier et al[50]) from their ongoing search updates (to October 2017) relevant for this KQ. The systematic review search results and study flow and selection are presented in Figure 4.

Figure 4. PRISMA flow diagram of study selection for KQ4 (treatment effectiveness)



A summary of the study characteristics for KQ4 is reported in Appendix 2; detailed study information is provided in Appendix 5.

Fifteen studies[50, 55-68] (2,869 women) were included that examined treatment effectiveness for bacteriuria. Most of the included studies were published in the 1960s, and one recent (2015) study[50] was included from the search update. One study[56] included in the Cochrane systematic review[4] only reported on the outcome of persistent bacteriuria and therefore was excluded from analysis and the overall body of evidence relevant to the current review's outcomes of interest.

The majority of studies were RCTs, with four controlled clinical trials (CCTs) included.[57, 60, 61, 68] Five studies were conducted in the US,[56, 57, 60, 61, 64] three studies were conducted in the UK[55, 63, 67] and one in Ireland,[58] three studies were conducted in Australia,[62, 59, 68] and one study conducted each in Denmark,[65] Jamaica,[65] and the Netherlands.[50] Seven studies[50, 55-57, 60, 61, 67] were not industry-funded, two studies[66, 68] were industry-funded, four studies[59, 62, 63, 65] were both industry and non-industry funded, and two studies[58, 64] did not report on funding. All pregnant women were enrolled from hospital-based clinics.

The studies varied in reporting of population characteristics. Only four studies[50, 56, 64, 67] (three used in our analysis) specified inclusion criteria for only asymptomatic women. One study[62] included more than 50% (n=117) of patients with radiological renal abnormalities, one study[57] reported previous UTI in its population (36% [n=133] of women in the treatment group and 40% [n=148] in the placebo group), one study[63] reported 23% (of 265 women) with a past history of renal-tract disease, and one study[65] included 18% (9 out of 50) of women with renal abnormalities. One study[50] reported exclusion of women with urogenital anomalies from the study and one study[61] excluded women with chronic renal insufficiency. Two studies[55, 66] reported no differences between groups for socioeconomic status. No study reported whether women were enrolled in an urban or rural setting.

The studies varied in reporting of screening characteristics. Most (n=9) studies[50, 55, 57, 58, 61-63, 67, 68] enrolled women at their first prenatal visit, with one study[59] enrolling women at the second antenatal visit. Five studies[55, 56, 58, 60, 64] followed women until delivery or the postpartum period for outcomes. One study[67] followed women until 10 days post-delivery. Four studies[50, 59, 66, 68] followed women until 6 weeks post-delivery. One study[61] followed women until the post-delivery period but then again 3 to 4 years later. One study[62] followed women until 6 months post-delivery, and one study[65] followed women until 9 months after delivery. Five studies[50, 58, 59, 64, 66] required at least one urine sample to detect bacteriuria, with seven studies[55, 60, 62, 63, 65, 67, 68] requiring confirmation with another sample, and three others[56, 57, 61] requiring three total urine samples. The majority of studies used a routine culture to test for bacteriuria, while two studies[50, 59] used a urine dipslide device.

All studies[50, 56-68] except one[55] treated women with more than 1 dose of antibiotics. Five studies[50, 58, 64, 66, 67] provided up to one week of antibiotics, one study[65] treated women for at least three weeks, one study[63] treated women for at least 30 days, and six studies[56, 59-62, 68] treated women for bacteriuria up to delivery. Most (n=7) studies[50, 57, 58, 60, 63, 64, 67] tested for persistent bacteriuria during pregnancy (with re-treatment as warranted); only one study[59] tested for persistent bacteriuria after delivery and three studies[61, 62, 65] tested for cure during pregnancy and after delivery. The control arm in ten studies[50, 55-57, 60-63, 65, 66] was provided with a placebo; two studies[58, 59] did not provide antibiotics to participants in the control group. Although we would anticipate that studies would treat (initially asymptomatic) women in the control group upon development of symptoms, only three studies[64, 67, 68] reported this.

The outcome reported by the most number of studies (n=12[50, 55, 57-65, 67]) was pyelonephritis. Most studies[50, 55, 57, 59, 61-63, 67] used a combination of two or more of the following symptoms to determine development of pyelonephritis: fever (≥100◦F or ≥38◦C) or pyrexia, nausea, chills or rigours, vomiting, dysuria, frequency of urination, burning during urination, costovertebral tenderness, flank pain, and loin pain and/or tenderness. Three studies[58, 60, 65] did not define criteria for pyelonephritis, and one study[64] used "acute symptoms of cystopyelitis".

Perinatal mortality was variably defined among the studies[50, 57, 61-63, 68] that reported this outcome. Two studies used gestation to define perinatal mortality: >20 weeks[61] and >28 weeks.[62] One study[50] defined perinatal death as stillbirth, death during labor or death within 28 days of life. One study[63] did not define "perinatal mortality". Two studies[57, 68] combined stillbirths with "neonatal death" or "death prior to hospital discharge".

Spontaneous abortion was reported by two studies[59, 68] that did not specify gestational age.

Of the four studies that reported on preterm delivery, three studies[50, 66, 68] used <37 weeks of gestation as the criteria, and the study by Furness et al[59] used <38 weeks of gestation.

Seven studies[50, 55, 57, 61-63, 68] reported low birth weight as ≤2500g or <2500g; Kazemier at al[50] used small for gestational age (SGA) at <10th percentile and <5th percentile, and we combined these data for this study.

Neonatal sepsis was reported by one study[50] as confirmed with culture, and without criteria in another study.[66]

For harms (any serious adverse event (AE)), two studies[50, 57] reported congenital/abnormalities, one study[63] reported fetal abnormalities and one study[59] reported anencephaly. Additionally, Elder et al[57] reported no events of hemolytic anemia for infants ("erythroblastosis fetalis") in either group.

No study reported on maternal mortality, maternal sepsis, or maternal harms (serious AE).

Overall, most of the studies that reported on at least one of the outcomes of interest were assessed as having high ROB, with three studies[62, 63, 66] assessed as having unclear ROB, and only one study[50] assessed at low ROB. The main issues were due to poor reporting of research methods and characteristics of the study population. Groups of studies contributing to each outcome had at least one study with high overall risk. Many studies[55, 57, 59, 62-67] reported their methodological design as "random" without adequate details, with only one study[50] using a computer-generated random assignment of participants. Many (n=9) studies[55, 56, 58, 59, 63-67] did not adequately describe concealment of allocation, with four studies[57, 60, 61, 68] describing allocation by alternation. The Netherlands study[50] used central allocation to support a judgment of low ROB for this domain. Five studies[50, 55, 56, 61, 62] that reported double-blinding were assessed at low ROB for this domain; the remaining ten studies[57-60, 63-68] were assessed as unclear ROB for blinding of participants and personnel due to lack of reporting within the context of objective outcomes. Four studies[50, 55, 56, 62] mentioned blinding of assessors or "double-blind" conditions to support a judgment of low ROB, whereas eleven studies[57-61, 63-68] assessed as unclear ROB did not report blinding of outcome assessors within the context of subjective outcomes. Two studies were assessed at high ROB for incomplete reporting as there were inconsistent data for low birth weight between groups and missing data on pyelonephritis in the treatment group for one study,[55] and no details on dropouts (20 out of 226 women) as well as 17% loss to follow-up for low birth weight and gestational age at delivery in the other

study.[59] Five studies[57, 58, 61, 62, 67] did not provide details on loss to follow-up for pyelonephritis and/or neonatal outcomes; these were assessed at unclear ROB. The majority (n=8) of studies[50, 56, 60, 63-66, 68] reported on details of dropouts, if any. Six studies[55, 56, 58, 64, 65, 67] were assessed at high ROB for selective reporting due to lack of reporting on pyelonephritis and/or neonatal outcomes. Eight studies[57, 59-63, 66, 68] were assessed as having unclear ROB due to lack of protocol and ability to assess selective reporting. As no other bias was identified, all the studies were assessed at low ROB for "other sources of bias".

For the summary of ROB assessments for KQ4, see Table 6; detailed study quality assessments are reported in Appendix 8.

Table 6. Summary of methodological quality[a] - KQ4 (treatment effectiveness)

Study	Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessors	Incomplete Reporting	Selective Reporting	Other Bias*	Overall Risk of Bias**
Brumfitt 1975[55]								
Elder 1966[56]								
Elder 1971[57]								
Foley 1987[58]								
Furness 1975[59]								
Gold 1966[60]								
Kass 1960[61]								
Kazemier 2015[50]								
Kincaid-Smith 1965[62]								
Little 1966[63]								
Mulla 1960[64]								
Pathak 1969[65]								
Thomsen 1987[66]								
Williams 1969[67]								
Wren 1969[68]								

[a]Assessed using the Cochrane Risk of Bias[34] tool

*Assessed as: Low risk of bias if no other sources of bias are identified, High risk of bias if other sources of bias detected such as: participant characteristics (baseline imbalances), study design characteristics (crossover, cluster-randomized, or blocked randomization in trials without blinding); Unclear risk of bias assessment not applicable for this domain.

**Assessed as: Low if all domains are assessed as low, Unclear if at least one domain is assessed as unclear and no domains are assessed as high, or High if at least one domain is assessed as high.

Legend:

- Low risk
- Unclear risk
- High risk

KQ4. What are the benefits and harms of antibiotic treatment compared with placebo or no treatment for asymptomatic bacteriuria in pregnancy?

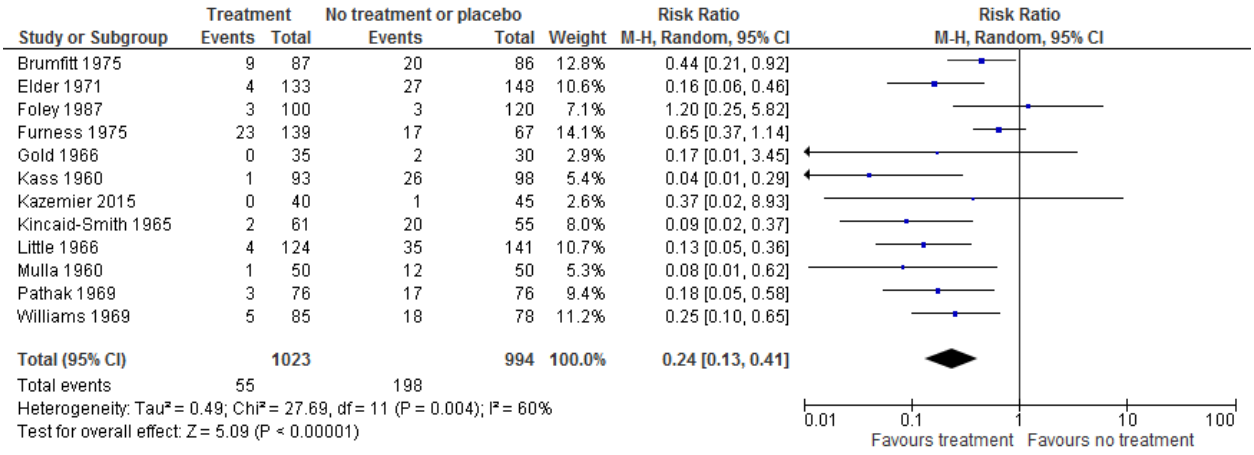
Results are summarized below by outcome. See Evidence Set 3 for GRADE EP and SOF tables and forest plots. We conducted subgroup analyses to explore possible reasons for heterogeneity among studies, whenever sufficient number of studies (e.g., 2 per subgroup if categorical) reported on the *a priori*

subgroups for population and screening characteristics. Sensitivity analyses were also carried out for ROB and study design. Only those subgroup findings that were sufficiently credible, as per criteria outlined in method, in explaining inconsistencies between studies are reported here; results for all other subgroup analyses can be obtained by contacting the review authors.

Pyelonephritis

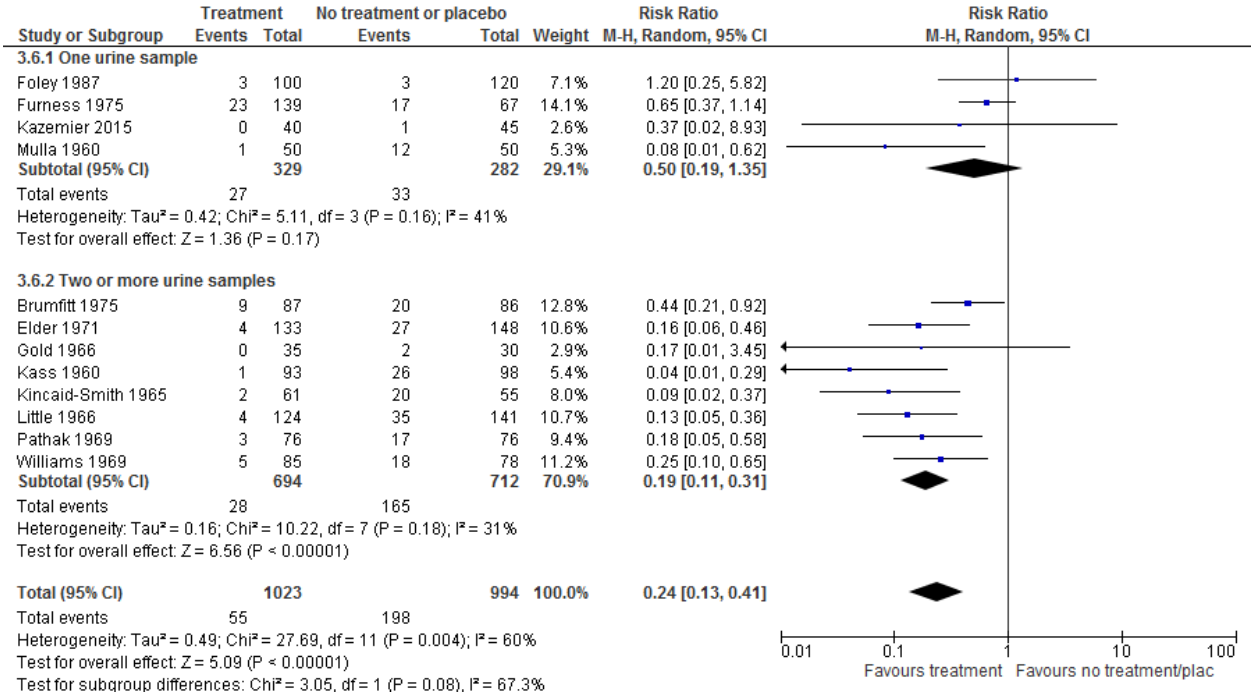
A total of 12 studies[50, 55, 57-65, 67] (2,017 women) with the majority at high ROB examined the effects of antibiotic treatment and found a significant difference in development of pyelonephritis (RR 0.24; 95% CI 0.13, 0.41; I^2=60%; ARR 17.6%; NNT 6, 95% CI 5, 7). Three of the trials clearly stated that only women without symptoms at baseline were included (other trials may have included some symptomatic women); sensitivity analysis by removing the nine trials did not affect the results (3 trials, RR 0.22; 95% CI 0.10, 0.49; I^2=0%). Sensitivity analysis for ROB (removing those studies with overall high risk) and study design (CCTs removed) did not change the results: removal of nine trials did not affect results (1 trials, RR 0.37; 95% CI 0.02, 8.93), and removal of three CCTs did not affect overall results (9 RCTs, RR 0.28; 95% CI 0.16, 0.51; I^2=60%). The quality of this body of evidence was rated as low due to concerns with ROB and indirectness (i.e., majority of studies did not report including exclusively asymptomatic women and some included some high-risk women). We have some certainty that treatment will reduce risk for pyelonephritis but are uncertain about the magnitude of the effect.

3.1 Pyelonephritis (overall)



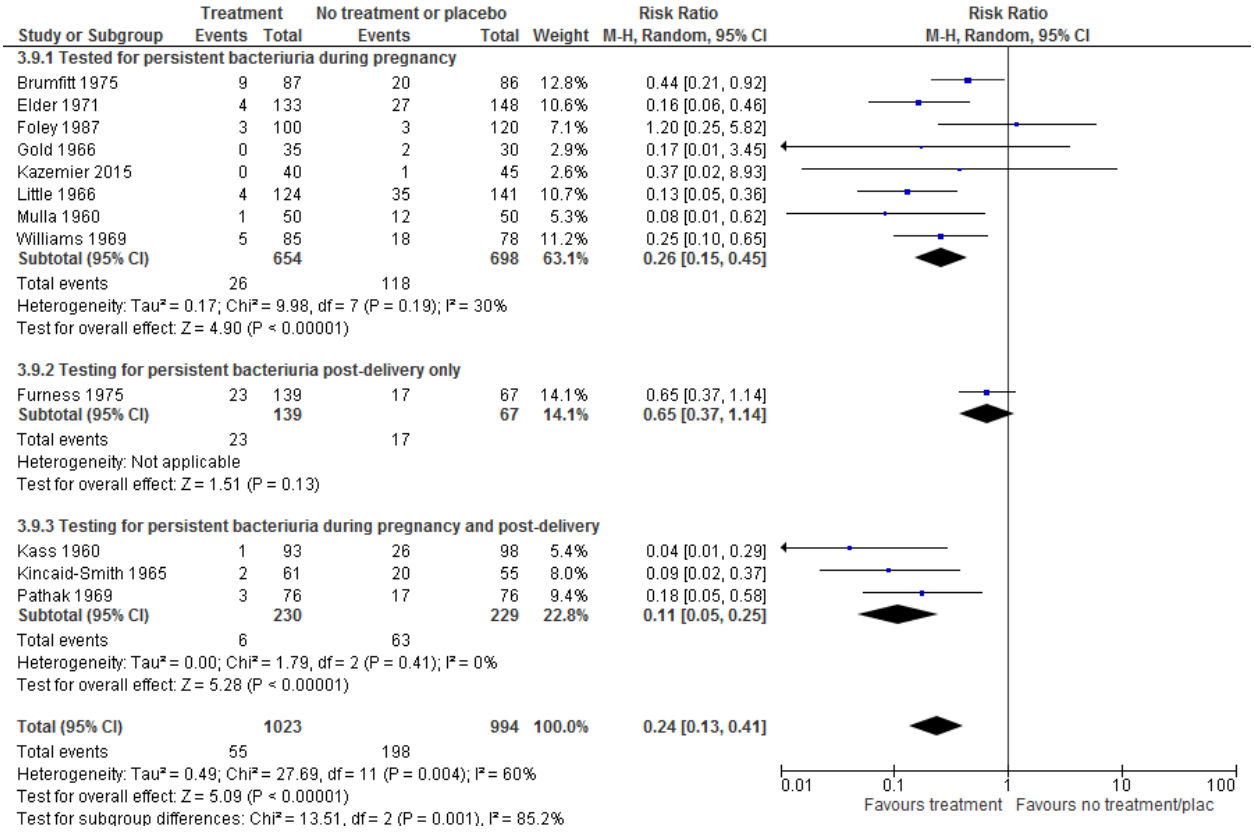
Subgroup analyses considered to have some credibility examined the number of urine samples (e.g. use of confirmatory culture), testing for persistent bacteriuria, and length of follow-up (ES Forest Plots 3.1.1-3.1.3).

3.1.1 Pyelonephritis subgroup: number of urine samples at each screening visit*

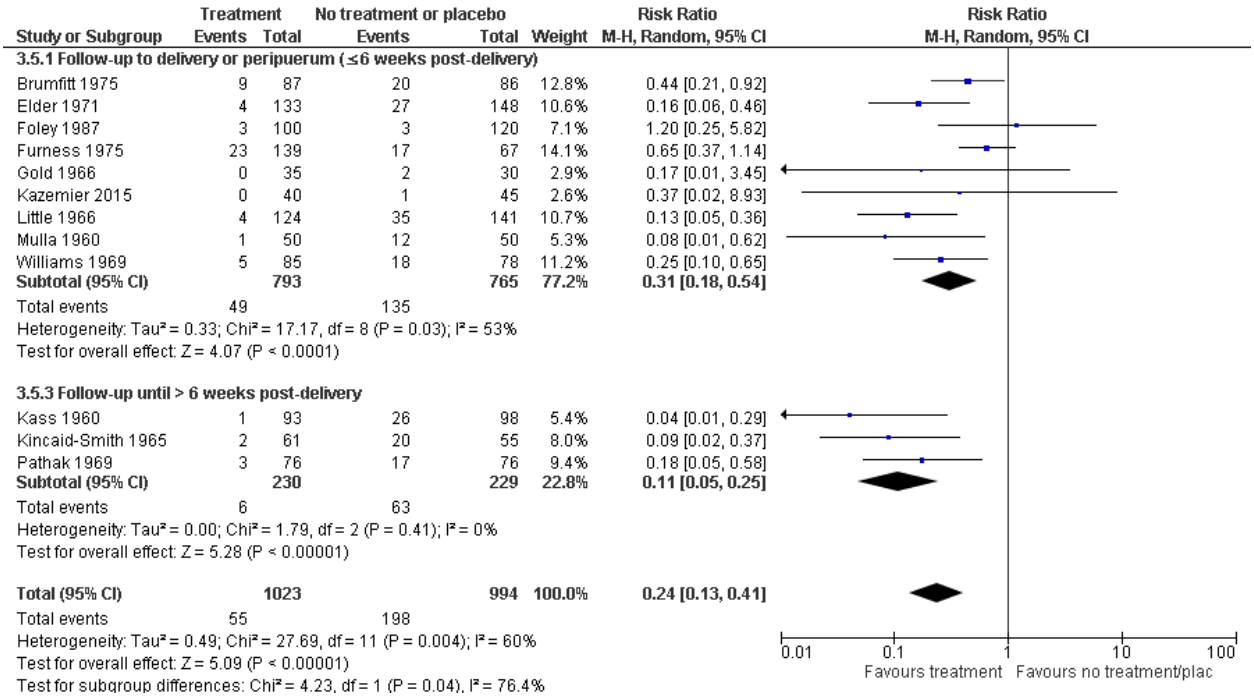


*The additional culture(s) was used to confirm levels of bacteriuria.

3.1.2 Pyelonephritis subgroup: timing of testing for persistent bacteriuria



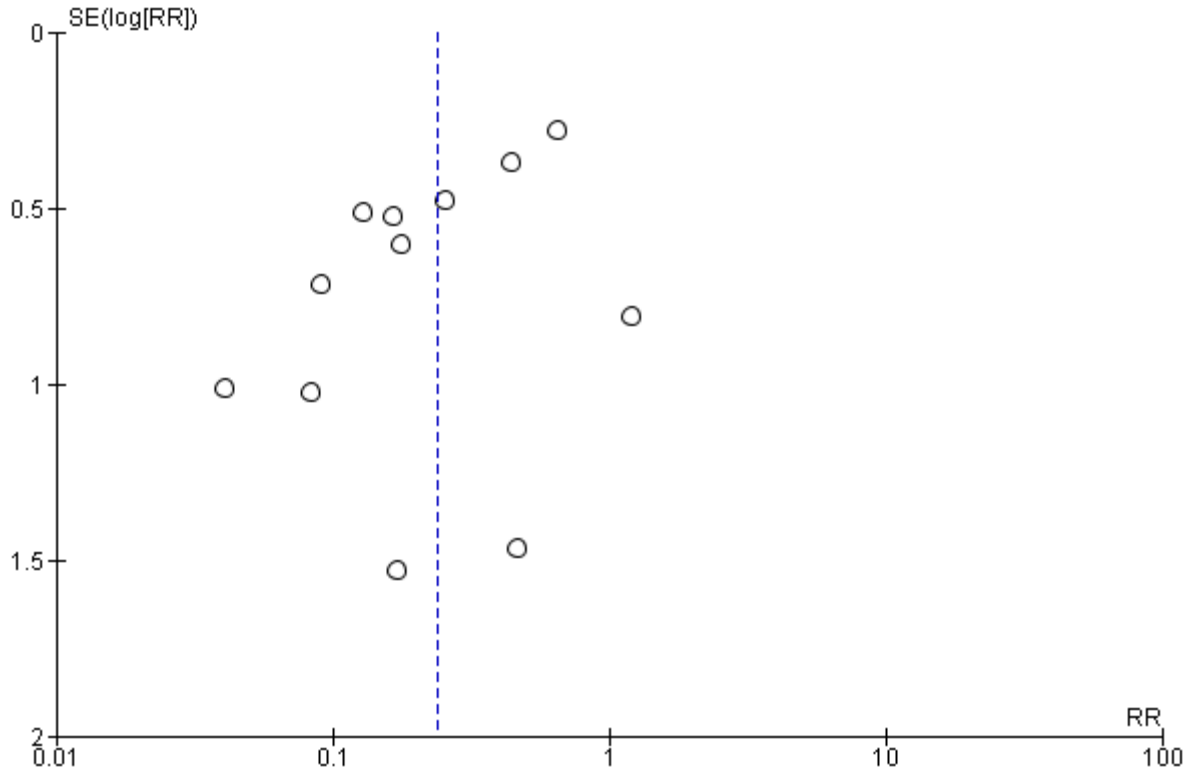
3.1.3 Pyelonephritis subgroup: duration of follow-up



Subgroup analysis for the number of urine samples—studies using one or more additional cultures to confirm ASB compared with just one culture—appeared to explain the heterogeneity among all studies combined (I^2=60%) for the outcome of pyelonephritis (RR 0.19, 95% CI 0.11, 0.31; I^2=31% versus RR 0.50, 95% CI 0.19, 1.35; I^2=41%). The test for subgroup differences did not meet our criteria for statistical significance (p=0.08), but the heterogeneity in each subgroup was reduced and visual inspection of the forest plots suggests there may be an important difference in effect. There was a statistically significant subgroup difference (p=0.001) when testing for persistent bacteriuria was done during pregnancy and after delivery (RR 0.11, 95% CI 0.05, 0.25; I^2=0%) compared with testing during pregnancy (RR 0.26, 95% CI 0.15, 0.45; I^2=30%) or with testing only after delivery (RR 0.65, 95% CI 0.37, 1.14). The test for subgroup differences for duration of follow-up was statistically significant (p=0.04) between studies that followed women beyond six weeks after delivery (RR 0.11, 95% CI 0.05, 0.25; I^2=0%) compared with those that only followed women until delivery or six weeks post-delivery (RR 0.31, 95% CI 0.18, 0.54; I^2=53%).

A funnel plot (Figure 5) was performed to visually assess small-study bias, and appeared symmetrical. The Egger's test was conducted and the result approached significance, but was inconclusive (p=0.065). The twelve studies with small sample sizes limit the ability to detect or exclude the possibility of small-study bias.

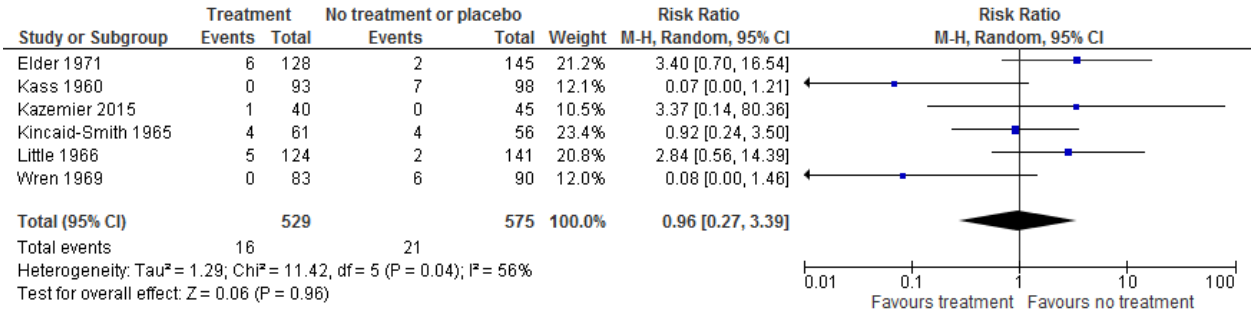
Figure 5. Funnel Plot Asymmetry Test for outcome of pyelonephritis for KQ4 (treatment)



Perinatal mortality

A total of six studies (1,104 women) examined the outcome of perinatal mortality; one study[50] was at low ROB, three studies[57, 61, 68] were at high ROB, and two studies[62, 63] were at unclear ROB. There was no significant difference for antibiotics compared with no treatment on perinatal mortality (RR 0.96, 95% CI 0.27, 3.39; I^2=56%). This body of evidence was rated as very low due to downgrading for ROB, indirectness, and imprecision.

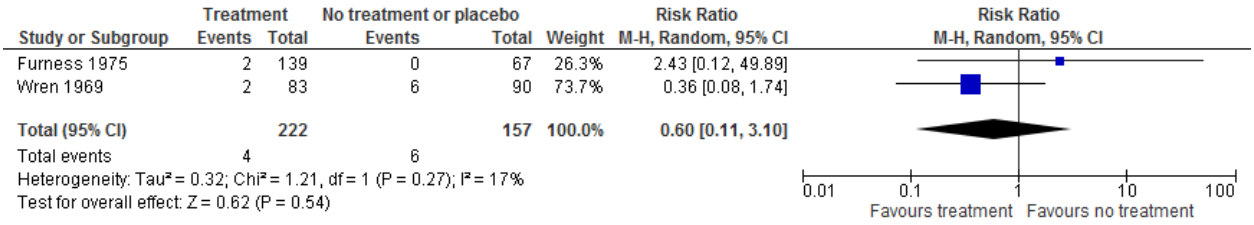
3.2 Perinatal mortality



Spontaneous abortion

Two studies[59, 68] (379 women) with high ROB reported on spontaneous abortion and found no significant difference between groups (RR 0.60, 95% CI 0.11, 3.10; I^2=17%). This body of evidence was downgraded for ROB, indirectness, and imprecision for an overall quality of very low.

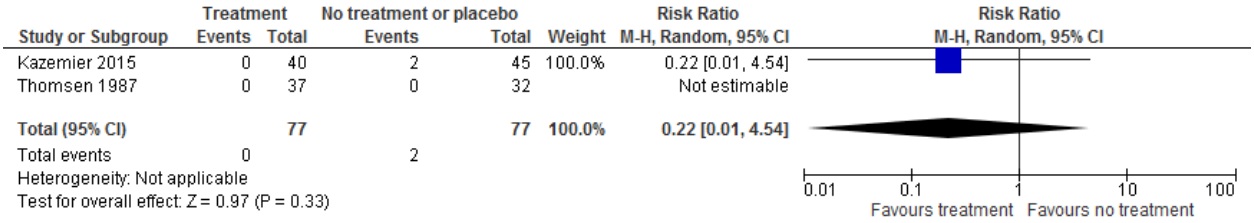
3.3 Spontaneous abortion



Neonatal sepsis

Two studies[50, 66] (154 women) with low ROB reported on neonatal sepsis and there was no statistically significant difference found between groups. Meta-analysis was not conducted due to there being no events in the study by Thomsen. This body of evidence was downgraded for indirectness, and imprecision for an overall quality of very low.

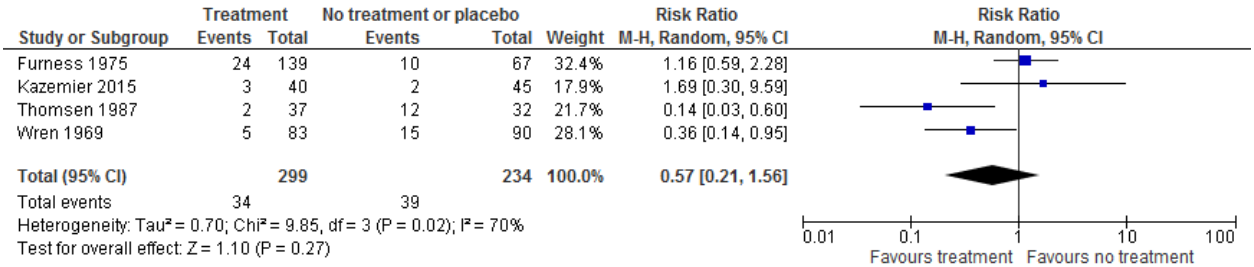
3.4 Neonatal sepsis



Preterm delivery

Two studies[50, 66] with low risk of bias and two studies[59, 68] with high ROB with a combined total of 533 women found no significant difference between antibiotics and no treatment on preterm delivery (RR 0.57, 95% CI 0.21, 1.56; I^2=70%). This body of evidence was rated as very low due to downgrading for ROB, inconsistency, and indirectness.

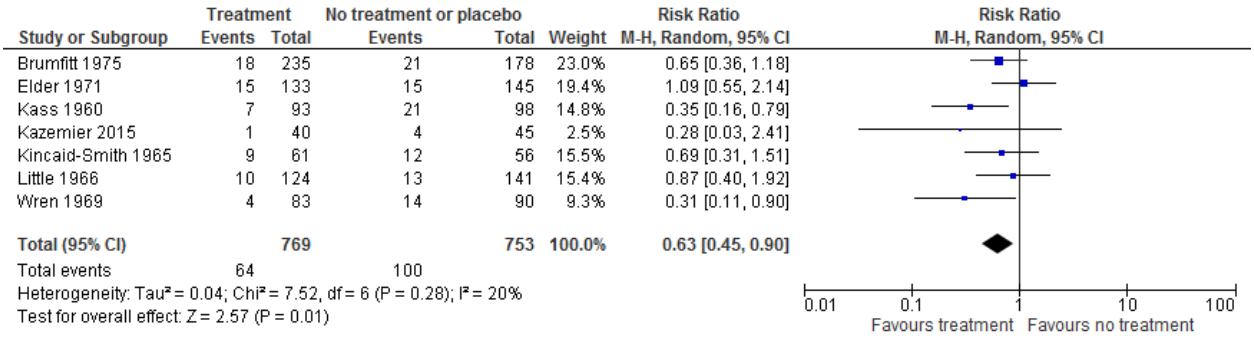
3.5 Preterm delivery



Low birth weight

A total of seven studies (1,522 women) with two studies[50, 63] at low ROB, three[55, 57, 61] at high ROB and one[62] at unclear ROB examined the effect of treatment on low birth weight. There was a statistically significant difference favoring antibiotic treatment (RR 0.63; 95% CI 0.45, 0.90; I^2=20%; ARR 4.4%; NNT 23, 95% CI 15, 85). This body of evidence was rated as low quality due to downgrading for ROB and indirectness. The Optimal Information Size did not quite meet our criteria but we did not have serious concerns to warrant downgrading for this domain.

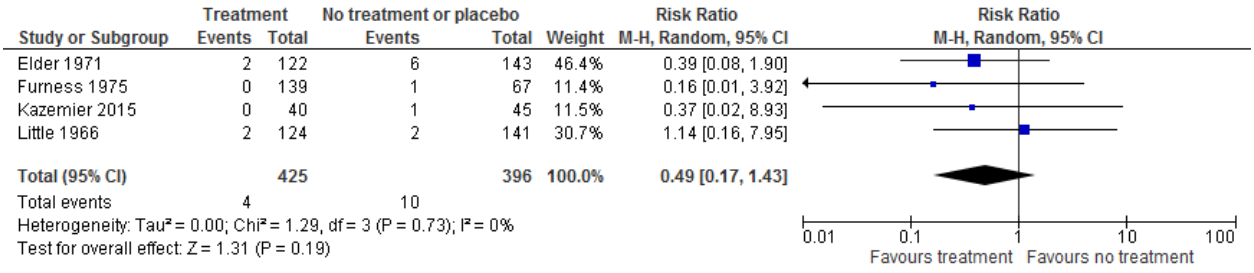
3.6 Low birthweight



Neonatal serious harm: fetal abnormalities

Four studies (821 women) with low ROB in two[50, 63] and high ROB in two[57, 59] examined the effect of antibiotic treatment on fetal abnormalities (harm). There was no statistically significant difference between groups (RR 0.49, 95% CI 0.17, 1.43; I^2=0%). This body of evidence was rated as very low due to downgrading for ROB, indirectness, and imprecision.

3.7 Neonatal serious harm: fetal abnormalities



Neonatal serious harm: hemolytic anemia

One study[57] (265 women) with high ROB reported no cases of hemolytic anemia (harm) in infants for the intervention and control groups; this body of evidence was downgraded for ROB, inconsistency, indirectness, and imprecision for an overall quality of very low.

Chapter 4. Discussion, Applicability and Conclusion

Overview of Findings

KQ1a. What are the benefits and harms of screening compared with no screening for asymptomatic bacteriuria in pregnancy? Are there subgroup differences for patient characteristics (e.g., socioeconomic status [SES])?

Three observational studies of unclear risk of bias (ROB) examined the effectiveness of screening with urine culture compared with no screening for asymptomatic bacteriuria (ASB). Timing, collection methods, and treatment protocols differed between studies. None of the studies reported on several of our critical benefit outcomes (maternal mortality, maternal sepsis, neonatal sepsis, and serious maternal harms) or the important benefit of low birthweight. A significant difference was found for pyelonephritis, from three studies (RR 0.28, 95% CI 0.15, 0.54; I^2=0%, ARR 1.3%; NNT 77, 95% CI 65, 121). No significant differences were found for the remaining critical benefit outcomes (spontaneous abortion, perinatal mortality, preterm delivery). Only one study reported on serious neonatal harms (fetal abnormalities) and found no differences (although the number of events and overall sample were small). The quality of evidence was very low for all outcomes. Based on the available evidence we are very uncertain about the effects of screening compared with no screening on these outcomes.

KQ1b. What are the benefits and harms of screening programs with different screening methods or algorithms for asymptomatic bacteriuria in pregnancy?

One observational study of low ROB compared frequent screening (using chemical reagent strip, lab urinalysis and urine culture for all visits) with one-time screening (using chemical reagent strip, lab urinalysis and urine culture on first visit) and found a significant difference in preterm delivery but no difference in pyelonephritis; no other outcomes were reported. The study found more preterm deliveries among the group with frequent screening (RR 1.57, 95% CI 1.11, 2.23), a finding for which the authors did not comment on or suggest an explanation. The quality of evidence was very low; therefore, we are very uncertain about the effects of frequent screening compared with one-time screening for these outcomes. The study did not report harms related to the different approaches to screening.

KQ2. How do women weigh the benefits and harms of screening and treatment of asymptomatic bacteriuria in pregnancy? How do women's valuation of benefits and harms of screening and treatment inform their decisions to undergo screening?

The evidence for women's outcome valuation was very limited as no studies directly addressed our KQs of weighing benefits versus harms and how this might affect decisions to undergo screening and treatment. Six cross-sectional studies and cross-sectional findings from women being recruited for a treatment trial provided indirect evidence of women's valuation of benefits and harms of screening and/or treatment of ASB; the findings only reflect valuation on harms, not the balance of benefits and harms, and are not specific to the context of ASB. These studies demonstrated varied opinions on antibiotic use during pregnancy, with nearly half of participants from two studies (47-48%) expressing that antibiotics should not be used during pregnancy. The cross-sectional analysis of patients recruited for a randomised controlled trial (RCT) of treatment for ASB found similar results, with 61% of 255 women with ASB not wanting to be treated for an asymptomatic condition. There was some evidence suggesting that women thought penicillin posed a teratogenic risk and that antibiotics were unsafe during pregnancy particularly for the fetus; these risks may be perceived as greater by women who are pregnant. How these attitudes

may inform the women's decisions on whether or not to screen for ASB was not reported, nor were details presented on accuracy or understanding of information regarding potential risks and benefits.

KQ4. What are the benefits and harms of antibiotic treatment compared with placebo or no treatment for asymptomatic bacteriuria in pregnancy?

Fifteen RCTs and controlled clinical trials (CCTs) compared antibiotics with placebo or no treatment for bacteriuria in pregnancy; the majority were assessed as high ROB. No study reported on maternal mortality, maternal sepsis or serious maternal harms. The most frequently reported critical benefit outcome (by 12 studies) was pyelonephritis and overall a significant difference was found showing a large relative reduction (RR 0.24, 95% CI 0.13, 0.41; I^2=60%; ARR 17.6%; NNT 6, 95% CI 5, 7). However, the quality of evidence for this outcome was low because of ROB and indirectness for concerns on applicability to asymptomatic and not-at-high risk populations; we have some certainty that treatment will reduce risk for pyelonephritis but are uncertain about the magnitude of the effect. A significant difference was also found for low birth weight (important benefit outcome) based on seven RCTs with a relative risk of 0.63 (95% CI 0.45, 0.90; I^2=20%; ARR 4.4%; NNT 23, 95% CI 15, 85). The quality of evidence for this outcome was also low for the same reasons as pyelonephritis. No significant differences were found between treatment and placebo/no treatment for spontaneous abortion, perinatal mortality, neonatal sepsis, preterm delivery, harms (fetal abnormalities and hemolytic anemia), and the quality of evidence for these outcomes was very low.

Subgroup analyses for pyelonephritis suggested variation in treatment effects based on several factors including number of urine samples used to confirm ASB, testing for persistent bacteriuria, and length of follow-up. The treatment effect appeared to be larger for studies where women were tested at least twice to confirm bacteriuria and initiate treatment, compared with those only testing women with one sample (and finding no significant difference for this outcome). These findings appear to reflect a reduced accuracy for one versus two-sample screening, whereby some women in these studies of one sample would be false positives and thus not having as much potential to gain from treatment (i.e., unnecessarily treated). Studies where women were tested for persistent bacteriuria (test of cure) during pregnancy or during pregnancy and post-delivery showed a larger treatment effect than those testing only post-delivery (although the latter group was represented by only one study). Finally, length of follow-up (>6 weeks post-delivery vs. ≤6 weeks) showed a greater treatment effect among those followed for more than 6 weeks, although both subgroups benefited from treatment. Relatively higher effects from studies with follow up >6 weeks post-delivery may indicate that some cases of pyelonephritis only occurred in this period rather than during pregnancy. These findings should be considered exploratory as they are based on between-study rather than within-study comparisons (i.e., non-randomized comparisons). Moreover, some of the subgroups contained few studies.

Comparison with other reviews

Similar to our findings, a recent systematic review by Angelescu et al[69] that examined benefits and harms of screening for ASB in pregnancy found no trials on screening effectiveness. The review authors included four RCTs focused on treatment of ASB: the recent study from the Netherlands[50] and three others (Elder et al[56], Mulla[64], and Williams et al[67]). These authors chose to limit their inclusion to studies reporting exclusively on treatment in asymptomatic women. We included studies that likely included some women with symptoms. In addition to other intervention characteristics (e.g., treatment regimen and

adjunct treatments) and outcomes (e.g., lower urinary tract infection (UTI), infant morbidity, very low birth weight <1500g) that were not included in our review, Angelescu et al[69] also examined population (diabetes, history of UTI, sociodemographic data) and screening (e.g., urine collection method, diagnostic procedures and cutoffs) characteristics, and similar outcomes (pyelonephritis, perinatal mortality, early preterm birth <32 weeks of gestation, adverse events) to ours. These review authors concluded that there was no reliable evidence on the benefits and harms of screening to support routine screening for ASB using urine culture in pregnant women.[69]

Applicability

KQ1 (screening vs. no screening): Some of the included studies may not represent a general population of women who are asymptomatic for bacteriuria. Most studies did not provide descriptive information about their populations' risk factors. For KQ1b (frequent vs. one-time screening), the included study setting was a hospital-based midwifery practice providing care to predominantly underserved and Hispanic women (72%) and the population had a relatively high rate of gestational diabetes (4 to 9%). All of the studies included in KQ1a and b used a urine culture to screen for bacteriuria.

KQ2 (outcome valuation): No study directly addressed how women weighed the benefits and harms of screening and treatment for ASB. Some information was available on women's perspectives regarding antibiotic treatment during pregnancy. None of the studies focused on Canadian women. Most studies involved internet surveys to pregnant, antepartum and/or postnatal women.

KQ4 (treatment vs. no treatment): All studies enrolled women from hospital-based clinics, and most enrolled women at their first prenatal visit. Only four studies reported exclusive inclusion of asymptomatic women, and at least four studies included a significant proportion of women that would be considered high-risk for ASB and its sequelae. As women needed to be positive for bacteriuria to be eligible for treatment, this population is not representative of women who undergo screening. Most studies were published in the 1960s; there was only one published post-1990, which was conducted in The Netherlands and published in 2015. The majority of studies used a urine culture to screen for bacteriuria with most using two or more samples, to allow for confirmation of bacteriuria to warrant treatment. Further, the majority tested for persistent bacteriuria during pregnancy and followed women to delivery or six weeks after delivery for outcomes.

Limitations

Methodological limitations were common and heterogeneous across studies. Controlled clinical trials included in the evidence base for treatment may not have allocated participants in an unbiased manner that ensures comparability between groups. Observational study designs do not systematically allocate participants and are therefore at risk of including unknown confounders that may influence outcomes; ratings of low or unclear methodological quality for these studies does not imply that they have comparable validity to RCTs with similar ratings. It is unclear whether poor reporting by many studies is an indicator of true methodological flaws, age of publications, or other potential reasons. Methodological standards for trials have changed over time as empirical evidence becomes available about design features introducing bias; the RCTs examined may have been considered as high quality when conducted although to today's standards this may not be true. The reporting in the observational screening studies did not demonstrate comparability at baseline or determine whether patients were symptomatic or had pyelonephritis when presenting to the study. Moreover, there were concerns with outcome reporting bias

as some pregnancy and neonatal outcomes (i.e., perinatal mortality, spontaneous abortion, neonatal sepsis, preterm delivery, low birth weight, and harms) were not reported among studies despite their relevance and high importance to clinicians and patients. While most studies used a urine culture to detect asymptomatic bacteriuria, criteria for defining a positive test were not always clear or reported. One study only included women positive for group B streptococcus with a lower range criterion for bacteriuria warranting treatment (with many considered contaminated specimens, rather than ASB); it is unclear if these women differ from women positive for other organisms. It is unclear whether the variations in definitions of outcomes have any effect on detection and reporting of outcomes. Early stopping due to low incidence of primary outcomes in the Kazemier study[50] may have biased effects of treatment. The small sample sizes among individual studies and pooled analyses limit the precision of effect size estimates.

Examining evidence on treatment for ASB as linked evidence for benefits and harms of screening programs has limitations. There is a likelihood that the absolute effects from treating bacteriuric women overestimate the effects for the screening population of all pregnant women where an estimated 2-10% will have asymptomatic bacteriuria.[1] Only three studies contributing to the meta-analyses reported study patients as exclusively asymptomatic pregnant women, while the remaining studies did not specify this criterion; a concern is that among women who are treated, effect of benefit may be larger among symptomatic women compared with women who are asymptomatic for bacteriuria.

The mechanisms of pyelonephritis progressing to adverse maternal and neonatal outcomes are unclear. Multiple factors may influence outcomes; for example, preterm birth may be confounded especially in the older studies by issues such as access to contraception and family planning, treatment of other asymptomatic infections such as chlamydia and bacterial vaginosis, and detection and management of pregnancy complications/conditions. With limited reporting of baseline characteristics among studies, it is difficult to make direct associations between specific risk factors and subsequent outcomes.

Screening and treatment practices have evolved since the 1960s when most of the studies began publishing on asymptomatic bacteriuria. Current obstetric practices have, for example, better recognition of risk factors for urinary tract infections and other pregnancy complications, prompt treatment of symptoms, and a broader range of antibiotic options. These factors would suggest a lower control group (baseline) event rate and therefore less absolute benefit in current practice.

As we did not include studies published in languages other than English and French, it is unknown whether we are missing studies that may provide information on screening and treatment of ASB. There is some evidence showing that meta-analyses from systematic reviews in conventional medicine using language restrictions do not appear to be biased.[24, 25]

Future Research

Although the anticipation of a large relative risk reduction for pyelonephritis appears to limit the clinical equipoise necessary to conduct RCTs on screening for ASB, we think there may be sufficient rationale to consider such trials based on: (1) very low quality evidence from screening studies and an appreciation of the linked nature of treatment evidence, particularly considering there are concerns about the methodological quality and the applicability of these old trials to current practice, and (2) some evidence suggesting that the incidence of pyelonephritis in untreated ASB (e.g., 2.5% in recent screening cohort study[50]) may be substantially lower than that reported in historical literature and most of the available

treatment trials (median control group incidence of 23%), such that the absolute number of women who actually benefit from screening may be relatively low. Should such RCTs, or some other design valid for evaluating screening programs, be conducted we strongly encourage investigators to capture data accurately on harms and suitable for conducting a cost-effectiveness analysis, in clearly defined populations and using modern definitions for outcomes. Screening for ASB is not currently performed in all settings,[50] indicating that clinical equipoise exists for enough clinicians to make these trials feasible and informative.

Prior to embarking on designing a trial for screening, but useful in any case, better information is needed to determine whether or not there are important moderating factors for ASB screening, as we attempted to examine in KQ1b. Our subgroup analyses examining moderators of effect, for example based on studies using one urine culture versus at least one additional confirmatory culture, had some credibility but were limited because of the need to rely on between-study effects. Studies directly examining this, and other factors such as different thresholds for treatment when particular organisms are detected, could provide high-quality data and be informative for how to maximize benefit. Enhanced culture protocols (e.g. expanded spectrum) for detecting the most clinically relevant uropathogens are emerging,[12, 13] and if found to consistently provide better detection of these microorganisms than standard urine culture, studies comparing screening programs differing by these methods are encouraged to determine if they also predict how well treatment reduces the risk for pyelonephritis and other pregnancy complications in women without symptoms.

More evidence or information about how women, especially those living in Canada, weigh the benefits and harms of screening (including treatment when screened positive) for ASB in pregnancy would be valuable. Understanding the difficulties in providing patients with results on benefits and harms in easily understood formats (particularly in absolute numbers), and because of low-quality evidence to support such information, it is hard to know how well some forms of additional research (e.g., population surveys) could answer this question. It may be useful to use deliberative processes or focus group research, to facilitate understanding and in-depth considerations on this question. Regardless of whether this information influences recommendations to screen or not for ASB on a population level, this information may be informative to determine whether it is critical to better engage patients in decision-making on their care.

Conclusions

This systematic review examined three sets of evidence to inform recommendations on screening for ASB in pregnancy. Using the GRADE approach, we determined the evidence to be very low quality for most outcomes from observational studies comparing screening programs using urine culture with no screening; as such, we have no or very little certainty in the effect estimates for these outcomes. Moreover, several outcomes were not reported. Similar interpretations are made about the evidence from one study comparing frequent screening with one-time screening with culture. No direct evidence was found on how women weigh the benefits and harms of screening and/or treatment for ASB and how this might affect their decisions to undergo screening. Antibiotic treatment for women having significant bacteriuria may reduce the incidence of pyelonephritis in these women and the number of their babies born at low birth weight. We are uncertain if the magnitudes of the effect estimates from treatment are true. Very low quality evidence from these trials did not allow us to have any certainty in effects on other

maternal and neonatal benefits and for fetal abnormalities and hemolytic anemia; no evidence was found for other serious harms.

Evidence Sets 1 - 3

Evidence Set 1. Table 1.1 GRADE Summary of Findings – KQ1a: Benefits and harms of screening compared to no screening

Screening compared to no screening for asymptomatic bacteriuria in pregnant women

Patient or population: asymptomatic bacteriuria in pregnant women

Setting: Any primary or clinical care setting providing care to pregnant women

Intervention: screening

Comparison: no screening

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with no screening	**Risk with screening**				
Maternal mortality	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal mortality.
Maternal sepsis	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal sepsis.
Pyelonephritis	**Median**		**RR 0.28** (0.15 to 0.54)	5659 (3 observational studies[43, 44, 45])	⨁◯◯◯ VERY LOW [1, a]	We are very uncertain about the effects of screening on pyelonephritis.
	18 per 1,000	**13 fewer per 1,000** (from 8 fewer to 16 fewer)				
Perinatal mortality	**Median**		**RR 1.21** (0.01 to 102.93)	724 (2 observational studies[43, 45])	⨁◯◯◯ VERY LOW [1, b]	We are very uncertain about the effects of screening on perinatal mortality.
	19 per 1,000	**4 more per 1,000** (from 19 fewer to 1,000 more)				
Spontaneous abortion	55 per 1,000	**2 fewer per 1,000** (from 32 fewer to 70 more)	**RR 0.96** (0.41 to 2.27)	370 (1 observational study[43])	⨁◯◯◯ VERY LOW [1, c]	We are very uncertain about the effects of screening on spontaneous abortion.
Neonatal sepsis	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on neonatal sepsis.
Preterm delivery	**Median**		**RR 8.70**	722	⨁◯◯◯	We are very uncertain about the

Screening compared to no screening for asymptomatic bacteriuria in pregnant women

Patient or population: asymptomatic bacteriuria in pregnant women

Setting: Any primary or clinical care setting providing care to pregnant women

Intervention: screening

Comparison: no screening

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with no screening	**Risk with screening**				
	13 per 1,000	**102 more per 1,000** (from 9 fewer to 1,000 more)	(0.32 to 240.07)	(2 observational studies[43, 45])	VERY LOW [1, d]	effects of screening on preterm delivery.
Low birthweight	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on low birthweight.
Maternal serious harm(s)	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal serious harms.
Neonatal serious harm: fetal abnormalities	11 per 1,000	**5 more per 1,000** (from 8 fewer to 85 more)	**RR 1.50** (0.25 to 8.87)	372 (1 observational study[45])	⨁◯◯◯ VERY LOW [1, e]	We are very uncertain about the effects of screening on fetal abnormalities.

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1] The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).

Pyelonephritis [a] → Very Low Quality Evidence: Three non-concurrent cohort studies (Gérard 1983, Gratacós 1994, Uncu 2001) reported this outcome (n=5,659). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics. The optimal information size is met (sample size >5600),

therefore downgrading for **imprecision** is not warranted. There were no serious concerns to warrant downgrading for **inconsistency**, **indirectness**, or **other considerations**.

Perinatal mortality [b] → Very Low Quality Evidence: Two non-concurrent cohort studies (n=724; Gérard 1983, Uncu 2001) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on perinatal mortality. Further downgrading is warranted for **imprecision** due to optimal information size not being met with a small sample size. There were no serious concerns to warrant downgrading for **inconsistency, indirectness** or **other considerations**.

Spontaneous abortion [c] → Very Low Quality Evidence: One non-concurrent cohort study reported this outcome (n=370; Gérard 1983). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on spontaneous abortion. Only one study provided data on spontaneous abortion, so this warrants downgrading for **inconsistency**. Further downgrading for **imprecision** is warranted due to low event rates (total of 20) without optimal information size. There were no serious concerns to warrant downgrading for **indirectness** or **other considerations**.

Preterm delivery [d] → Very Low Quality Evidence: Two non-concurrent cohort studies (n=722; Gérard 1983, Uncu 2001) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on preterm delivery. Further downgrading is warranted for **imprecision** for inadequate sample size and optimal information size not being met (total of 38 events). There were no serious concerns to warrant downgrading for **inconsistency, indirectness** or **other considerations**.

Neonatal serious harm: fetal abnormalities (harm) [e] → Very Low Quality Evidence: One non-concurrent cohort study reported this outcome (n=370; Uncu 2001). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as three studies did not report on fetal abnormalities. Only one study provided data on this outcome so this warrants downgrading for **inconsistency**. Further downgrading for **imprecision** is warranted due to the optimal information size not being met for rare events. There were no serious concerns to warrant downgrading for **indirectness** or **other considerations**.

Evidence Set 1. Table 1.2 GRADE Evidence Profile – KQ1a: Benefits and harms of screening compared to no screening

Question: Screening compared to no screening for asymptomatic bacteriuria in pregnant women

Setting: Any primary or clinical care setting providing care to pregnant women

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision** [1]	**Other considerations**	**screening**	**no screening**	**Relative (95% CI)**	**Absolute (95% CI)**		
Maternal mortality												
0									not estimable		-	CRITICAL
Maternal sepsis												
0									not estimable		-	CRITICAL
Pyelonephritis												
3	observational studies[43, 44, 45]	serious	not serious	not serious	serious	none	10/2008 (0.5%)	1.8%	**RR 0.28** (0.15 to 0.54)	**13 fewer per 1,000** (from 8 fewer to 16 fewer)	⊕◯◯◯ VERY LOW [1, a]	CRITICAL
Perinatal mortality												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision** [1]	**Other considerations**	**screening**	**no screening**	**Relative (95% CI)**	**Absolute (95% CI)**		
2	observational studies[43, 45]	serious	not serious	not serious	serious	none	6/349 (1.7%)	1.9%	**RR 1.21** (0.01 to 102.93)	**4 more per 1,000** (from 19 fewer to 1,000 more)	⨁◯◯◯ VERY LOW [1, b]	CRITICAL
Spontaneous abortion												
1	observational studies[43]	serious	serious	not serious	serious	none	9/170 (5.3%)	11/200 (5.5%)	**RR 0.96** (0.41 to 2.27)	**2 fewer per 1,000** (from 32 fewer to 70 more)	⨁◯◯◯ VERY LOW [1, c]	CRITICAL
Neonatal sepsis												
0									not estimable		-	CRITICAL
Preterm delivery												
2	observational studies[43, 45]	serious	not serious	not serious	serious	none	33/347 (9.5%)	1.3%	**RR 8.70** (0.32 to 240.07)	**102 more per 1,000** (from 9 fewer to 1,000 more)	⨁◯◯◯ VERY LOW [1, d]	CRITICAL
Low birthweight												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision [1]	Other considerations	screening	no screening	Relative (95% CI)	Absolute (95% CI)		
0									not estimable		-	IMPORTANT
Maternal serious harm(s)												
0									not estimable		-	CRITICAL
Neonatal serious harm: fetal abnormalities												
1	observational studies[45]	serious	serious	not serious	serious	none	3/186 (1.6%)	2/186 (1.1%)	**RR 1.50** (0.25 to 8.87)	**5 more per 1,000** (from 8 fewer to 85 more)	⨁◯◯◯ VERY LOW [1, e]	CRITICAL

CI: Confidence interval; **RR:** Risk ratio

[1] **The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).**

Pyelonephritis [a] → Very Low Quality Evidence: Three non-concurrent cohort studies (Gérard 1983, Gratacós 1994, Uncu 2001) reported this outcome (n=5,659). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics. The optimal information size is met (sample size >5600), therefore downgrading for **imprecision** is not warranted. There were no serious concerns to warrant downgrading for **inconsistency**, **indirectness**, or **other considerations**.

Perinatal mortality [b] → Very Low Quality Evidence: Two non-concurrent cohort studies (n=724; Gérard 1983, Uncu 2001) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on perinatal mortality. Further downgrading is warranted for **imprecision** due to optimal information size not being met with a small sample size. There were no serious concerns to warrant downgrading for **inconsistency, indirectness** or **other considerations**.

Spontaneous abortion [c] → Very Low Quality Evidence: One non-concurrent cohort study reported this outcome (n=370; Gérard 1983). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on spontaneous abortion. Only one study provided data on spontaneous abortion, so this warrants downgrading for **inconsistency**. Further downgrading for **imprecision** is warranted due to low event rates (total of 20) without optimal information size. There were no serious concerns to warrant downgrading for **indirectness** or **other considerations**.

Preterm delivery [d] → Very Low Quality Evidence: Two non-concurrent cohort studies (n=722; Gérard 1983, Uncu 2001) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as two studies did not report on preterm delivery. Further downgrading is warranted for **imprecision** for inadequate sample size and optimal information size not being met (total of 38 events). There were no serious concerns to warrant downgrading for **inconsistency, indirectness** or **other considerations**.

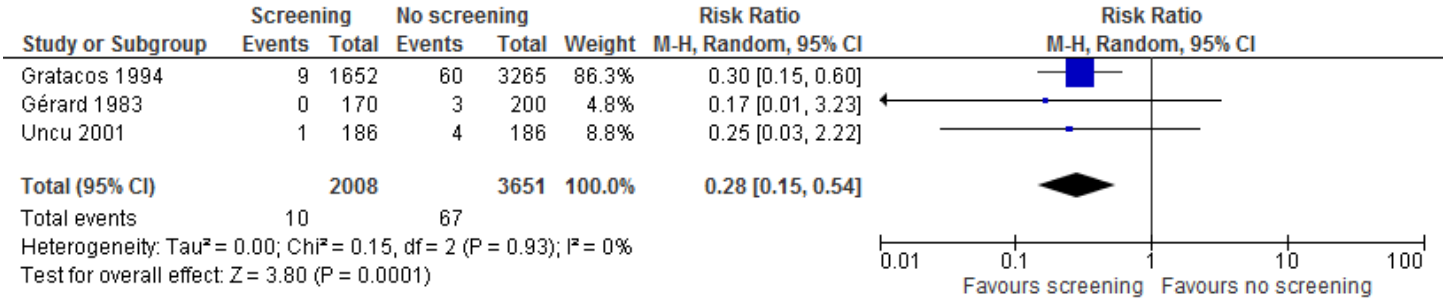
Neonatal serious harm: fetal abnormalities [e] → Very Low Quality Evidence: One non-concurrent cohort study reported this outcome (n=370; Uncu 2001). Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** across studies associated with: 1) no demonstration of comparability between screening and no screening groups, and 2) no adjustment to analyses for risk factors or other patient characteristics, and suspected reporting bias as three studies did not report on fetal abnormalities. Only one study provided data on this outcome so this warrants downgrading for **inconsistency**. Further downgrading for **imprecision** is warranted due to the optimal information size not being met for rare events. There were no serious concerns to warrant downgrading for **indirectness** or **other considerations**.

Evidence Set 1. Forest Plots 1.1-1.5 – KQ1a: Benefits and harms of screening compared to no screening

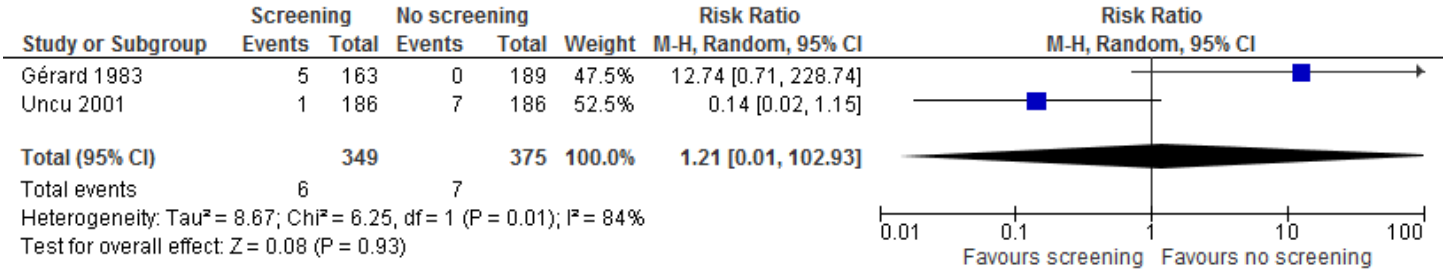
Outcome	No. of studies	No. of participants	Effect size (Risk Ratio; M-H, Random, 95%CI)
1.1 Pyelonephritis	3	5659	0.28 [0.15, 0.54]
1.2 Perinatal mortality >=20 wks GA note: Gérard >=31 wks; Uncu >20 wks	2	724	1.21 [0.01, 102.93]
1.3 Spontaneous abortion <20 wks GA note: 1 study <=28 wks (all occurred 7-21 wks)	1	370	0.96 [0.41, 2.27]
1.4 Preterm delivery <37 wks GA	2	722	8.70 [0.32, 240.07]
1.5 Neonatal serious harm: fetal abnormalities	1	372	1.50 [0.25, 8.87]

CI: confidence interval; GA: gestational age; M-H: Mantel-Haenszel; No.: number; wks: weeks

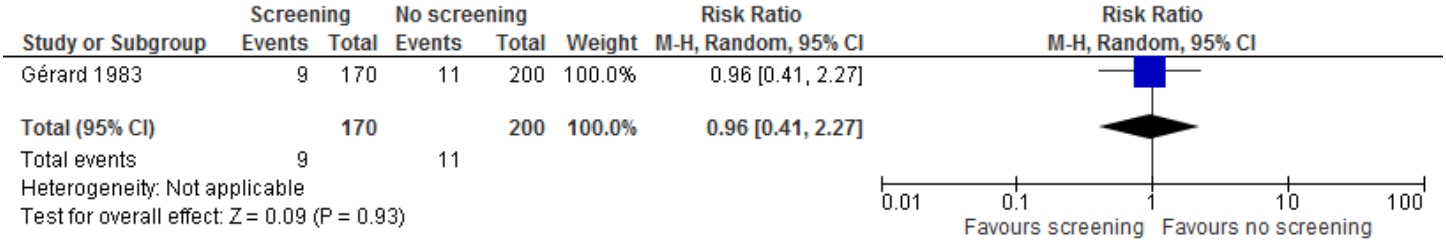
1.1 Pyelonephritis



1.2 Perinatal mortality (>=20 wks GA)



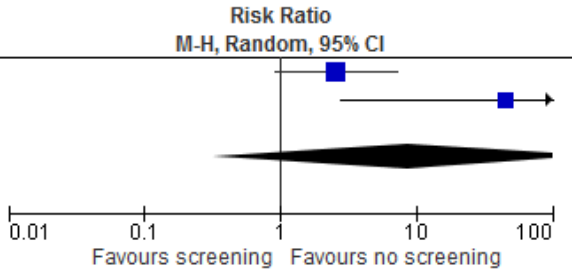
1.3 Spontaneous abortion (<20 wks GA)



1.4 Preterm delivery (<37 wks GA)

Study or Subgroup	Screening Events	Screening Total	No screening Events	No screening Total	Weight	Risk Ratio M-H, Random, 95% CI
Gérard 1983	11	161	5	189	57.5%	2.58 [0.92, 7.28]
Uncu 2001	22	186	0	186	42.5%	45.00 [2.75, 736.39]
Total (95% CI)		**347**		**375**	**100.0%**	**8.70 [0.32, 240.07]**
Total events	33		5			

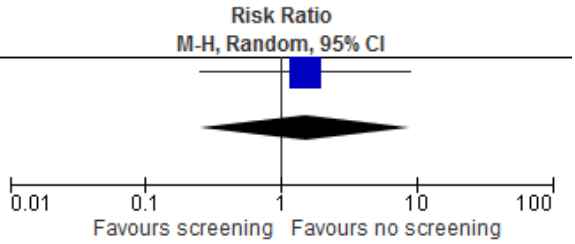
Heterogeneity: Tau² = 4.70; Chi² = 5.07, df = 1 (P = 0.02); I² = 80%
Test for overall effect: Z = 1.28 (P = 0.20)



1.5 Neonatal serious harm: fetal abnormalities

Study or Subgroup	Screening Events	Screening Total	No screening Events	No screening Total	Weight	Risk Ratio M-H, Random, 95% CI
Uncu 2001	3	186	2	186	100.0%	1.50 [0.25, 8.87]
Total (95% CI)		**186**		**186**	**100.0%**	**1.50 [0.25, 8.87]**
Total events	3		2			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.45 (P = 0.65)



Evidence Set 2. Table 2.1 GRADE Summary of Findings - KQ1b: Benefits and harms of frequent screening compared to one-time screening

Frequent screening compared to one-time screening for asymptomatic bacteriuria

Patient or population: asymptomatic bacteriuria

Setting: Any primary clinical care setting providing care to pregnant women

Intervention: frequent screening

Comparison: one-time screening

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with one-time screening	**Risk difference with frequent screening**				
Pyelonephritis	4 per 1,000	**0 fewer per 1,000** (from 3 fewer to 13 more)	**RR 1.09** (0.27 to 4.35)	1952 (1 observational study[46])	⨁◯◯◯ VERY LOW [1, a]	We are very uncertain about the effects of frequent screening compared to one-time screening on pyelonephritis.
Preterm delivery	49 per 1,000	**28 more per 1,000** (from 5 more to 60 more)	**RR 1.57** (1.11 to 2.23)	1952 (1 observational study[46])	⨁◯◯◯ VERY LOW [1, b]	We are very uncertain about the effects of frequent screening compared to one-time screening on preterm delivery.

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1] **The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).**

CI: Confidence interval; **RR:** Risk ratio

Pyelonephritis [a] → Very Low Quality Evidence: One non-concurrent cohort study (n=1952; Rhode 2007) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** associated with: 1) no demonstration that pyelonephritis was not present at start of study, 2) no demonstration of comparability between frequent and one-time screening groups, and 3) no adjustment to analyses to account for risk factors or other patient characteristics. Only one study provided data for this outcome so downgrading is warranted for **inconsistency**. Further downgrading is warranted for **indirectness** as the women

are predominantly medically underserved, Hispanic and receiving care from a midwifery clinic, with a high rate of gestational diabetes (9%). The optimal information size is not met (8 events) with sample size (n=1952), therefore this warrants downgrading for **imprecision**. There were no serious concerns to warrant downgrading for **other considerations**.

Preterm delivery [b] → Very Low Quality Evidence: One non-concurrent cohort study (n=1952; Rhode 2007) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to very serious **risk of bias** associated with: 1) no demonstration of comparability between frequent and one-time screening groups, 2) no adjustment to analyses to account for risk factors or other patient characteristics, and 3) suspected reporting bias among outcomes reported by studies (did not report on spontaneous abortion, perinatal mortality or fetal abnormalities). Only one study provided data for this outcome so downgrading is warranted for **inconsistency**. Further downgrading is warranted for **indirectness** as the women are predominantly medically underserved, Hispanic and receiving care from a midwifery clinic, with a high rate of gestational diabetes (9%). The event rate is low (122 events) without meeting optimal information size, so this is downgraded for **imprecision**. There were no serious concerns to warrant downgrading for **other considerations**.

Evidence Set 2. Table 2.2 GRADE Evidence Profile - KQ1b: Benefits and harms of frequent screening compared to one-time screening

Question: Frequent screening compared to one-time screening for asymptomatic bacteriuria

Setting: Any primary clinical care setting providing care to pregnant women

Bibliography:

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision[1]**	**Other considerations**	**frequent screening**	**one-time screening**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pyelonephritis												
1	observational studies[46]	serious	serious	serious	serious	none	4/933 (0.4%)	4/1019 (0.4%)	**RR 1.09** (0.27 to 4.35)	**0 fewer per 1,000** (from 3 fewer to 13 more)	⨁◯◯◯ VERY LOW [1, a]	CRITICAL
Preterm delivery												
1	observational studies[46]	serious	serious	serious	serious	none	72/933 (7.7%)	50/1019 (4.9%)	**RR 1.57** (1.11 to 2.23)	**28 more per 1,000** (from 5 more to 60 more)	⨁◯◯◯ VERY LOW [1, b]	CRITICAL

CI: Confidence interval; **RR:** Risk ratio

[1] The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).

CI: Confidence interval; **RR:** Risk ratio

Pyelonephritis [a] → Very Low Quality Evidence: One non-concurrent cohort study (n=1952; Rhode 2007) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to serious **risk of bias** associated with: 1) no demonstration that pyelonephritis was not present at start of study, 2) no demonstration of comparability between frequent and one-time screening groups, and 3) no adjustment to analyses to account for risk factors or other patient characteristics. Only one study provided data for this outcome so downgrading is warranted for **inconsistency**. Further downgrading is warranted for **indirectness** as the women are predominantly medically underserved, Hispanic and receiving care from a midwifery clinic, with a high rate of gestational diabetes (9%). The optimal information size is not met (8 events) with sample size (n=1952), therefore this warrants downgrading for **imprecision**. There were no serious concerns to warrant downgrading for **other considerations**.

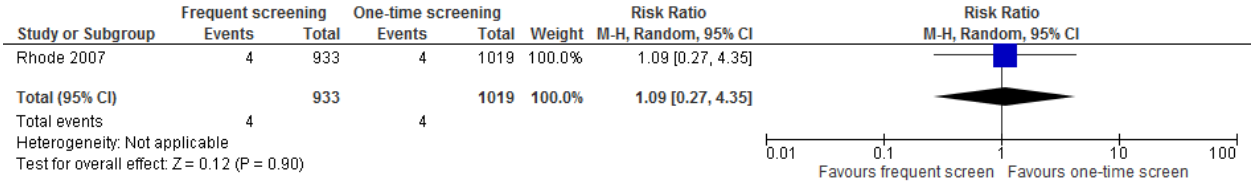
Preterm delivery [b] → Very Low Quality Evidence: One non-concurrent cohort study (n=1952; Rhode 2007) reported this outcome. Quality of evidence is downgraded from high to low due to observational level data. Further downgrading from low to very low is warranted due to very serious **risk of bias** associated with: 1) no demonstration of comparability between frequent and one-time screening groups, 2) no adjustment to analyses to account for risk factors or other patient characteristics, and 3) suspected reporting bias among outcomes reported by studies (did not report on spontaneous abortion, perinatal mortality or fetal abnormalities). Only one study provided data for this outcome so downgrading is warranted for **inconsistency**. Further downgrading is warranted for **indirectness** as the women are predominantly medically underserved, Hispanic and receiving care from a midwifery clinic, with a high rate of gestational diabetes (9%). The event rate is low (122 events) without meeting optimal information size, so this is downgraded for **imprecision**. There were no serious concerns to warrant downgrading for **other considerations**.

Evidence Set 2. Forest Plots 2.1-2.2 - KQ1b: Benefits and harms of frequent screening compared to one-time screening

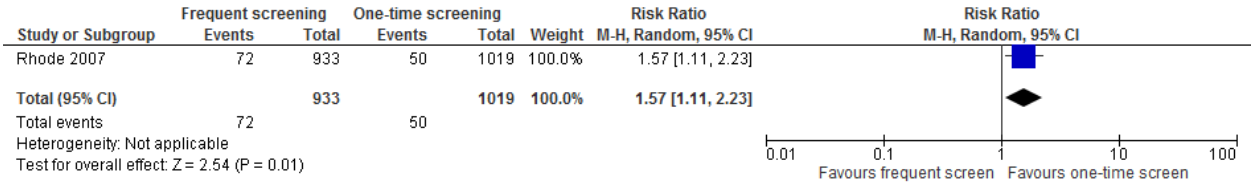
Outcome	No. of studies	No. of participants	Effect size (Risk Ratio; M-H, Random, 95%CI)
2.1 Pyelonephritis	1	1952	1.09 [0.27, 4.35]
2.2 Preterm delivery <37 wks GA	1	1952	1.57 [1.11, 2.23]

CI: confidence interval; GA: gestational age; M-H: Mantel-Haenszel; No.: number; wks: weeks

2.1 Pyelonephritis



2.2 Preterm delivery (<37 wks GA)



Evidence Set 3. Table 3.1 GRADE Summary of Findings – KQ4: Benefits and harms of treatment compared to no treatment

Treatment compared to no treatment for asymptomatic bacteriuria

Patient or population: asymptomatic bacteriuria

Setting: Any primary or clinical care setting providing care to pregnant women

Intervention: treatment

Comparison: no treatment

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with no treatment	**Risk with treatment**				
Maternal mortality	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal mortality.
Maternal sepsis	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal sepsis.
Pyelonephritis	**Median**		**RR 0.24** (0.13 to 0.41)	2017 (12 RCTs[50, 55, 57-65, 67])	⨁⨁◯◯ LOW [1, a]	There may be a reduction in pyelonephritis from treatment.
	232 per 1,000	**176 fewer per 1,000** (from 137 fewer to 202 fewer)				
Perinatal mortality	**Median**		**RR 0.96** (0.27 to 3.39)	1104 (6 RCTs[50, 57, 61-63, 68])	⨁◯◯◯ VERY LOW [1, b]	We are very uncertain about the effects of treatment on perinatal mortality.
	40 per 1,000	**2 fewer per 1,000** (from 29 fewer to 97 more)				
Spontaneous abortion	**Median**		**RR 0.60** (0.11 to 3.10)	379 (2 RCTs[59, 68])	⨁◯◯◯ VERY LOW [1, c]	We are very uncertain about the effects of treatment on spontaneous abortion.
	33 per 1,000	**13 fewer per 1,000** (from 30 fewer to 70 more)				
Neonatal sepsis	**Median**		**RR 0.22** (0.01 to 4.54)	154 (2 RCTs[50, 66])	⨁◯◯◯ VERY LOW [1, d]	We are very uncertain about the effects of treatment on neonatal sepsis.
	22 per 1,000	**17 fewer per 1,000** (from 22 fewer to 79 more)				

Treatment compared to no treatment for asymptomatic bacteriuria

Patient or population: asymptomatic bacteriuria

Setting: Any primary or clinical care setting providing care to pregnant women

Intervention: treatment

Comparison: no treatment

Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with no treatment	**Risk with treatment**				
Preterm delivery	**Median**		**RR 0.57** (0.21 to 1.56)	533 (4 RCTs[50, 59, 66, 68])	⨁◯◯◯ VERY LOW [1, e]	We are very uncertain about the effects of treatment on preterm delivery.
	158 per 1,000	**68 fewer per 1,000** (from 125 fewer to 88 more)				
Low birth weight	**Median**		**RR 0.63** (0.45 to 0.90)	1522 (7 RCTs[50, 55, 57, 61, 62, 63, 68])	⨁◯◯◯ LOW [1, f]	There may be a reduction in low birth weight from treatment.
	118 per 1,000	**44 fewer per 1,000** (from 12 fewer to 65 fewer)				
Maternal serious harm(s)	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	(0 studies)	-	No study reported on maternal serious harms.
Neonatal serious harm: fetal abnormalities	**Median**		**RR 0.49** (0.17 to 1.43)	821 (4 RCTs[50, 57, 59, 63])	⨁◯◯◯ VERY LOW [1, g]	We are very uncertain about the effects of treatment on harms (fetal abnormalities).
	19 per 1,000	**9 fewer per 1,000** (from 15 fewer to 8 more)				
Neonatal serious harm: hemolytic anemia	0 per 1,000	**0 per 1,000** (0 to 0)	not estimable	265 (1 RCT[57])	⨁◯◯◯ VERY LOW [1, h]	We are very uncertain about the effects of treatment on harms (hemolytic anemia).

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

Treatment compared to no treatment for asymptomatic bacteriuria

Patient or population: asymptomatic bacteriuria

Setting: Any primary or clinical care setting providing care to pregnant women

Intervention: treatment

Comparison: no treatment

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with no treatment	Risk with treatment				

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1] **The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).**

CI: Confidence interval; **RR:** Risk ratio

Pyelonephritis, overall [a] → Low Quality Evidence: Twelve trials (Brumfitt 1975, Elder 1971, Foley 1987, Furness 1975, Gold 1966, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Mulla 1960, Pathak 1969, Williams 1969) reported this outcome (n=2,017). Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Gold 1966, Kass 1960), inadequate allocation concealment (Elder 1971, Gold 1966, Kass 1960), and incomplete reporting (Brumfitt 1975, Furness 1975). This body of evidence on treatment effectiveness is downgraded from moderate to low for **indirectness** due to studies that did not explicitly include asymptomatic women (only 3 studies included exclusively asymptomatic women; Kazemier 2015, Mulla 1960, and Williams 1969), and studies that included high-risk women (Elder 1971, Kincaid-Smith 1965, Little 1966, and Pathak 1969). The optimal information size criterion is met (control group event rate=20%; total number of events=253) with an adequate sample size (n=2,017), and the confidence interval (0.13 to 0.41) indicates there may be important benefit; therefore, downgrading is not warranted for **imprecision**. There were no concerns with **inconsistency** or **other considerations** to warrant further downgrading.

Perinatal mortality [b] → Very Low Quality Evidence: Six trials (n=1,104; Elder 1971, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Kass 1960, Wren 1969), and inadequate allocation concealment (Elder 1971, Kass 1960). This body of evidence on treatment effectiveness is downgraded for **indirectness** due to studies that did not explicitly include asymptomatic women as well as studies that included high-risk women. Further downgrading is warranted for **imprecision** due to the samples size not being met for optimal information size criterion (37 events). There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Spontaneous abortion [c]→ Very Low Quality Evidence: Two trials (n=379; Furness 1975, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Wren 1969), inadequate allocation concealment (Wren 1969) and incomplete reporting (Furness 1975). Further downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women. The sample size is inadequate with optimal information size not met (10 events) to warrant downgrading twice from low to very low for **imprecision**. There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Neonatal sepsis [d] → Very Low Quality Evidence: Two trials (n=154; Kazemier 2015, Thomsen 1987) reported this outcome. Quality of evidence is downgraded for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women. The sample size (<2000) is

not met with only 2 events to warrant downgrading twice for **imprecision**. There were no concerns to warrant downgrading for **risk of bias, inconsistency** or **other considerations**.

Preterm delivery [e] → Very Low Quality Evidence: Four trials (n=533; Furness 1975, Kazemier 2015, Thomsen 1987, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate for **risk of bias** associated with use of alternation for sequence generation (Wren 1969), inadequate allocation concealment (Wren 1969), and incomplete reporting (Furness 1975). There is substantial heterogeneity (I^2=70%) with point estimates on both sides of the line of no effect to warrant downgrading for **inconsistency**. Downgrading from moderate to low for **indirectness** is warranted due to studies that did not explicitly include exclusively asymptomatic women. There were no concerns to warrant downgrading for **imprecision or other considerations**.

Low birth weight [f] → Low Quality Evidence: Seven trials (n=1,522; Brumfitt 1975, Elder 1971, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate for serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Kass 1960, Wren 1969), inadequate allocation concealment (Elder 1971, Kass 1960, Wren 1969), and incomplete reporting (Brumfitt 1975). Further downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women as well as studies that included high-risk women. The optimal information size was not quite met (<2000 patients and <200 events), but we did not think the concerns were serious enough to downgrade for this outcome for **imprecision.** There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Neonatal serious harm: fetal abnormalities [g] → Very Low Quality Evidence: Four trials (n=821; Elder 1971, Furness 1975, Kazemier 2015, Little 1966) reported this outcome. Quality of evidence is downgraded from high to moderate for serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971), inadequate allocation concealment (Elder 1971), and incomplete reporting (Furness 1975). Downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women as well as studies that included high-risk women. Further downgrading from low to very low for **imprecision** is warranted due to optimal information size (sample size of 821) not being met for rare events. There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Neonatal serious harm: hemolytic anemia [h] → Very Low Quality Evidence: One trial (n=265; Elder 1971) reported this outcome. Quality of evidence is downgraded from high to moderate for **risk of bias** associated with use of alternation for sequence generation and inadequate allocation concealment. Only one study provided data for this outcome so downgrading from moderate to low for **inconsistency** is warranted. Further downgrading from low to very low is warranted for **indirectness** due the study not explicitly including exclusively asymptomatic women as well as studies that included high-risk women. Due to optimal information size (sample size of 265) not being met for rare events, downgrading twice is warranted for **imprecision**. There were no concerns to warrant downgrading for **other considerations**.

Evidence Set 3. Table 3.1 GRADE Evidence Profile – KQ4: Benefits and harms of treatment compared to no treatment

Question: Treatment compared to no treatment for asymptomatic bacteriuria

Setting: Any primary or clinical care setting providing care to pregnant women

Bibliography:

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision** [1]	**Other considerations**	**treatment**	**no treatment**	**Relative (95% CI)**	**Absolute (95% CI)**		
Maternal mortality												
0									not estimable		-	CRITICAL
Maternal sepsis												
0									not estimable		-	CRITICAL
Pyelonephritis												
12	randomised trials[50, 55, 57-65, 67]	serious	not serious	serious	not serious	none	55/1023 (5.4%)	23.2%	**RR 0.24** (0.13 to 0.41)	**176 fewer per 1,000** (from 137 fewer to 202 fewer)	⊕⊕◯◯ LOW [1, a]	CRITICAL
Perinatal mortality												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision** [1]	**Other considerations**	**treatment**	**no treatment**	**Relative (95% CI)**	**Absolute (95% CI)**		
6	randomised trials[50, 57, 61-63, 68]	serious	not serious	serious	serious	none	16/529 (3.0%)	4.0%	**RR 0.96** (0.27 to 3.39)	**2 fewer per 1,000** (from 29 fewer to 97 more)	⨁◯◯◯ VERY LOW [1, b]	CRITICAL
Spontaneous abortion												
2	randomised trials[59, 68]	serious	not serious	serious	very serious	none	4/222 (1.8%)	3.3%	**RR 0.60** (0.11 to 3.10)	**13 fewer per 1,000** (from 30 fewer to 70 more)	⨁◯◯◯ VERY LOW [1, c]	CRITICAL
Neonatal sepsis												
2	randomised trials[50, 66]	not serious	not serious	serious	very serious	none	0/77 (0.0%)	2.2%	**RR 0.22** (0.01 to 4.54)	**17 fewer per 1,000** (from 22 fewer to 79 more)	⨁◯◯◯ VERY LOW [1, d]	CRITICAL
Preterm delivery												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision** [1]	**Other considerations**	**treatment**	**no treatment**	**Relative (95% CI)**	**Absolute (95% CI)**		
4	randomised trials[50, 59, 66, 68]	serious	serious	not serious	very serious	none	34/299 (11.4%)	15.8%	**RR 0.57** (0.21 to 1.56)	**68 fewer per 1,000** (from 125 fewer to 88 more)	⨁◯◯◯ VERY LOW [1, e]	CRITICAL
Low birth weight												
7	randomised trials[50, 55, 57, 61, 62, 63, 68]	serious	not serious	serious	not serious	none	64/769 (8.3%)	11.8%	**RR 0.63** (0.45 to 0.90)	**44 fewer per 1,000** (from 12 fewer to 65 fewer)	⨁⨁◯◯ LOW [1, f]	IMPORTANT
Maternal serious harm(s)												
0									not estimable		-	CRITICAL
Neonatal serious harm: fetal abnormalities												
4	randomised trials[50, 57, 59, 63]	serious	not serious	serious	very serious	none	4/425 (0.9%)	1.9%	**RR 0.49** (0.17 to 1.43)	**9 fewer per 1,000** (from 15 fewer to 8 more)	⨁◯◯◯ VERY LOW [1, g]	CRITICAL
Neonatal serious harm: hemolytic anemia												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision [1]	Other considerations	treatment	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials[57]	serious	serious	serious	very serious	none	0/122 (0.0%)	0/143 (0.0%)	not estimable		⨁◯◯◯ VERY LOW [1, h]	CRITICAL

[1] The imprecision domain is assessed using GRADE guidance[42] relevant for systematic reviews as follows: when optimal information size (OIS) criterion is met, and the 95% confidence interval overlaps no effect, consideration of important benefit or important harm will be assessed using a relative risk of 1.0 (0.75 to 1.25).

CI: Confidence interval; **RR:** Risk ratio

Pyelonephritis, overall [a] → Low Quality Evidence: Twelve trials (Brumfitt 1975, Elder 1971, Foley 1987, Furness 1975, Gold 1966, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Mulla 1960, Pathak 1969, Williams 1969) reported this outcome (n=2,017). Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Gold 1966, Kass 1960), inadequate allocation concealment (Elder 1971, Gold 1966, Kass 1960), and incomplete reporting (Brumfitt 1975, Furness 1975). This body of evidence on treatment effectiveness is downgraded from moderate to low for **indirectness** due to studies that did not explicitly include asymptomatic women (only 3 studies included exclusively asymptomatic women; Kazemier 2015, Mulla 1960, and Williams 1969), and studies that included high-risk women (Elder 1971, Kincaid-Smith 1965, Little 1966, and Pathak 1969). The optimal information size criterion is met (control group event rate=20%; total number of events=253) with an adequate sample size (n=2,017), and the confidence interval (0.13 to 0.41) indicates there may be important benefit; therefore, downgrading is not warranted for **imprecision**. There were no concerns with **inconsistency** or **other considerations** to warrant further downgrading.

Perinatal mortality [b] → Very Low Quality Evidence: Six trials (n=1,104; Elder 1971, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Kass 1960, Wren 1969), and inadequate allocation concealment (Elder 1971, Kass 1960). This body of evidence on treatment effectiveness is downgraded for **indirectness** due to studies that did not explicitly include asymptomatic women as well as studies that included high-risk women. Further downgrading is warranted for **imprecision** due to the samples size not being met for optimal information size criterion (37 events). There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Spontaneous abortion [c]→ Very Low Quality Evidence: Two trials (n=379; Furness 1975, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate due to serious **risk of bias** associated with use of alternation for sequence generation (Wren 1969), inadequate allocation concealment (Wren 1969) and incomplete reporting (Furness 1975). Further downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women. The sample size is inadequate with optimal information size not met (10 events) to warrant downgrading twice from low to very low for **imprecision**. There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Neonatal sepsis [d] → Very Low Quality Evidence: Two trials (n=154; Kazemier 2015, Thomsen 1987) reported this outcome. Quality of evidence is downgraded for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women. The sample size (<2000) is not met with only 2 events to warrant downgrading twice for **imprecision**. There were no concerns to warrant downgrading for **risk of bias, inconsistency** or **other considerations**.

Preterm delivery [e] → Very Low Quality Evidence: Four trials (n=533; Furness 1975, Kazemier 2015, Thomsen 1987, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate for **risk of bias** associated with use of alternation for sequence generation (Wren 1969), inadequate allocation concealment (Wren 1969), and incomplete reporting (Furness 1975). There is substantial heterogeneity (I^2=70%) with point estimates on both sides of the line of no effect to warrant downgrading for **inconsistency**. Downgrading from moderate to low for **indirectness** is warranted due to studies that did not explicitly include exclusively asymptomatic women. There were no concerns to warrant downgrading for **imprecision or other considerations**.

Low birth weight [f] → Low Quality Evidence: Seven trials (n=1,522; Brumfitt 1975, Elder 1971, Kass 1960, Kazemier 2015, Kincaid-Smith 1965, Little 1966, Wren 1969) reported this outcome. Quality of evidence is downgraded from high to moderate for serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971, Kass 1960, Wren 1969), inadequate allocation concealment (Elder 1971, Kass 1960, Wren 1969), and incomplete reporting (Brumfitt 1975). Further downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women as well as studies that included high-risk women. The optimal information size was not quite met (<2000 patients and <200 events), but we did not think the concerns were serious enough to downgrade for this outcome for **imprecision.** There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

Neonatal serious harm: fetal abnormalities [g] → Very Low Quality Evidence: Four trials (n=821; Elder 1971, Furness 1975, Kazemier 2015, Little 1966) reported this outcome. Quality of evidence is downgraded from high to moderate for serious **risk of bias** associated with use of alternation for sequence generation (Elder 1971), inadequate allocation concealment (Elder 1971), and incomplete reporting (Furness 1975). Downgrading from moderate to low is warranted for **indirectness** due to studies that did not explicitly include exclusively asymptomatic women as well as studies that included high-risk women. Further downgrading from low to very low for **imprecision** is warranted due to optimal information size (sample size of 821) not being met for rare events. There were no concerns to warrant downgrading for **inconsistency** or **other considerations**.

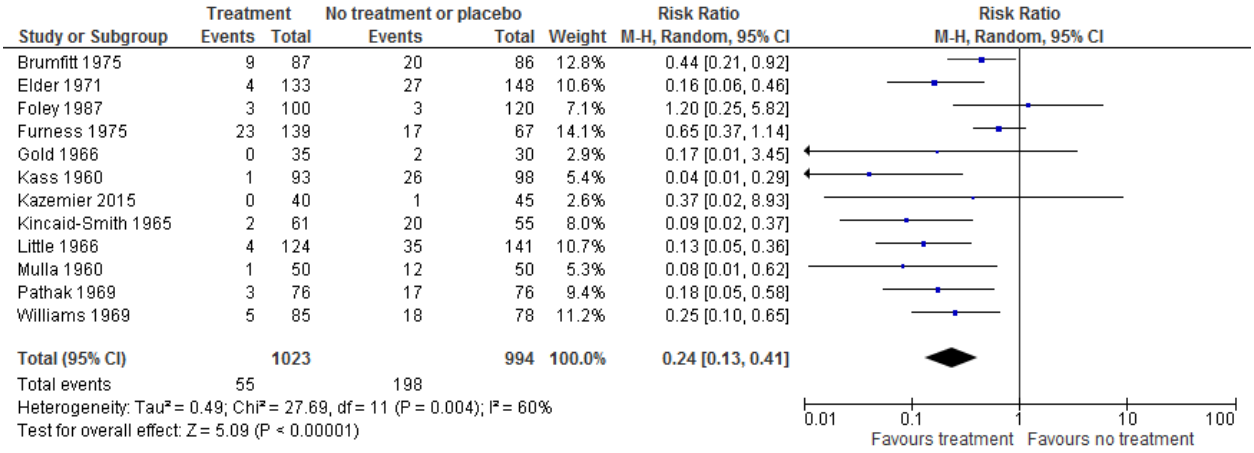
Neonatal serious harm: hemolytic anemia [h] → Very Low Quality Evidence: One trial (n=265; Elder 1971) reported this outcome. Quality of evidence is downgraded from high to moderate for **risk of bias** associated with use of alternation for sequence generation and inadequate allocation concealment. Only one study provided data for this outcome so downgrading from moderate to low for **inconsistency** is warranted. Further downgrading from low to very low is warranted for **indirectness** due the study not explicitly including exclusively asymptomatic women as well as studies that included high-risk women. Due to optimal information size (sample size of 265) not being met for rare events, downgrading twice is warranted for **imprecision**. There were no concerns to warrant downgrading for **other considerations**.

Evidence Set 3: Forest Plots 3.1-3.8 - KQ4: Benefits and harms of treatment compared to no treatment

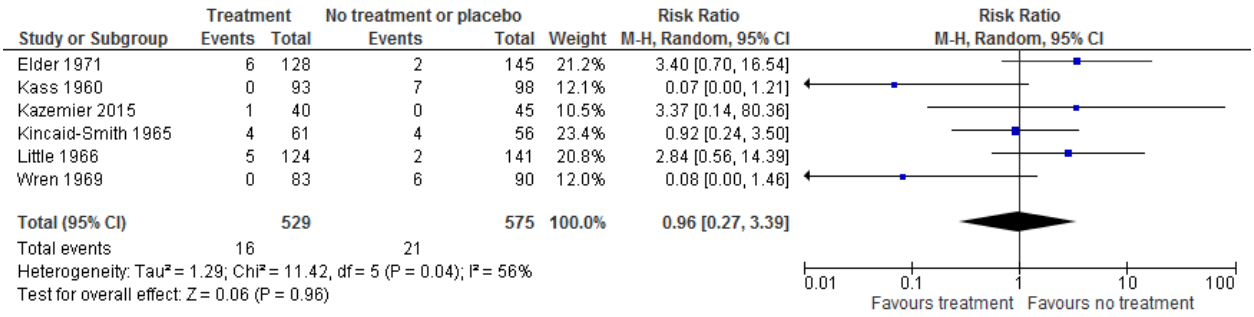
Outcome	No. of studies	No. of participants	Effect size (Risk Ratio; M-H, Random, 95%CI)
3.1 Pyelonephritis	12	2017	0.24 [0.13, 0.41]
3.2 Perinatal mortality (≥20 wks, including intrauterine demise, stillbirth, early neonatal death)	6	1104	0.96 [0.27, 3.39]
3.3 Spontaneous abortion (<20 wks)	2	379	0.60 [0.11, 3.10]
3.4 Neonatal sepsis	2	154	0.22 [0.01, 4.54]
3.5 Preterm delivery (<38 wks)	4	533	0.57 [0.21, 1.56]
3.6 Low birth weight (≤2500g; SGA <10th percentile & <5th percentile)	7	1522	0.63 [0.45, 0.90]
3.7 Neonatal serious harm: fetal abnormalities	4	821	0.49 [0.17, 1.43]
3.8 Neonatal serious harm: hemolytic anemia	1	265	Not estimable

CI: confidence interval; g: grams; M-H: Mantel-Haenszel; No.: number; SGA: small for gestational age; wks: weeks

3.1 Pyelonephritis



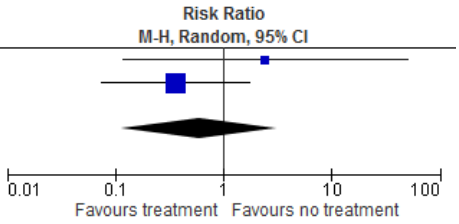
3.2 Perinatal mortality



3.3 Spontaneous abortion

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Furness 1975	2	139	0	67	26.3%	2.43 [0.12, 49.89]
Wren 1969	2	83	6	90	73.7%	0.36 [0.08, 1.74]
Total (95% CI)		**222**		**157**	**100.0%**	**0.60 [0.11, 3.10]**
Total events	4		6			

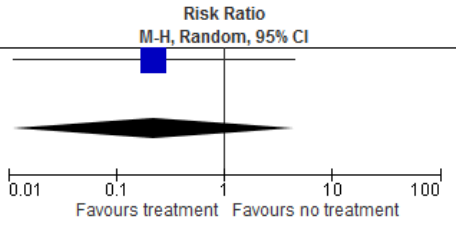
Heterogeneity: Tau² = 0.32; Chi² = 1.21, df = 1 (P = 0.27); I² = 17%
Test for overall effect: Z = 0.62 (P = 0.54)



3.4 Neonatal sepsis

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Kazemier 2015	0	40	2	45	100.0%	0.22 [0.01, 4.54]
Thomsen 1987	0	37	0	32		Not estimable
Total (95% CI)		**77**		**77**	**100.0%**	**0.22 [0.01, 4.54]**
Total events	0		2			

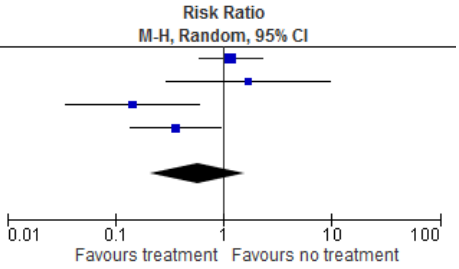
Heterogeneity: Not applicable
Test for overall effect: Z = 0.97 (P = 0.33)



3.5 Preterm delivery

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Furness 1975	24	139	10	67	32.4%	1.16 [0.59, 2.28]
Kazemier 2015	3	40	2	45	17.9%	1.69 [0.30, 9.59]
Thomsen 1987	2	37	12	32	21.7%	0.14 [0.03, 0.60]
Wren 1969	5	83	15	90	28.1%	0.36 [0.14, 0.95]
Total (95% CI)		**299**		**234**	**100.0%**	**0.57 [0.21, 1.56]**
Total events	34		39			

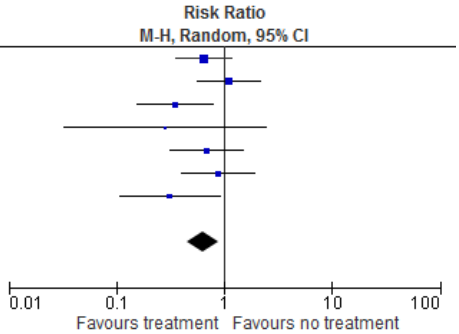
Heterogeneity: Tau² = 0.70; Chi² = 9.85, df = 3 (P = 0.02); I² = 70%
Test for overall effect: Z = 1.10 (P = 0.27)



3.6 Low birthweight

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Brumfitt 1975	18	235	21	178	23.0%	0.65 [0.36, 1.18]
Elder 1971	15	133	15	145	19.4%	1.09 [0.55, 2.14]
Kass 1960	7	93	21	98	14.8%	0.35 [0.16, 0.79]
Kazemier 2015	1	40	4	45	2.5%	0.28 [0.03, 2.41]
Kincaid-Smith 1965	9	61	12	56	15.5%	0.69 [0.31, 1.51]
Little 1966	10	124	13	141	15.4%	0.87 [0.40, 1.92]
Wren 1969	4	83	14	90	9.3%	0.31 [0.11, 0.90]
Total (95% CI)		**769**		**753**	**100.0%**	**0.63 [0.45, 0.90]**
Total events	64		100			

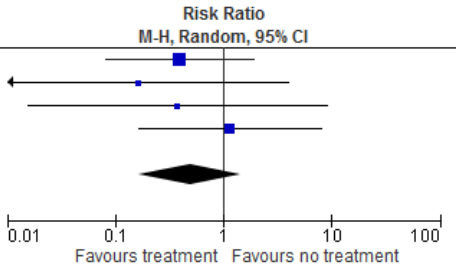
Heterogeneity: Tau² = 0.04; Chi² = 7.52, df = 6 (P = 0.28); I² = 20%
Test for overall effect: Z = 2.57 (P = 0.01)



3.7 Neonatal serious harm: fetal abnormalities

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Elder 1971	2	122	6	143	46.4%	0.39 [0.08, 1.90]
Furness 1975	0	139	1	67	11.4%	0.16 [0.01, 3.92]
Kazemier 2015	0	40	1	45	11.5%	0.37 [0.02, 8.93]
Little 1966	2	124	2	141	30.7%	1.14 [0.16, 7.95]
Total (95% CI)		**425**		**396**	**100.0%**	**0.49 [0.17, 1.43]**
Total events	4		10			

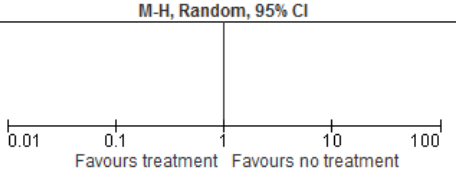
Heterogeneity: Tau² = 0.00; Chi² = 1.29, df = 3 (P = 0.73); I² = 0%
Test for overall effect: Z = 1.31 (P = 0.19)



3.8 Neonatal serious harm: hemolytic anemia

Study or Subgroup	Treatment Events	Total	No treatment or placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Elder 1971	0	122	0	143		Not estimable
Total (95% CI)		**122**		**143**		**Not estimable**
Total events	0		0			

Heterogeneity: Not applicable
Test for overall effect: Not applicable

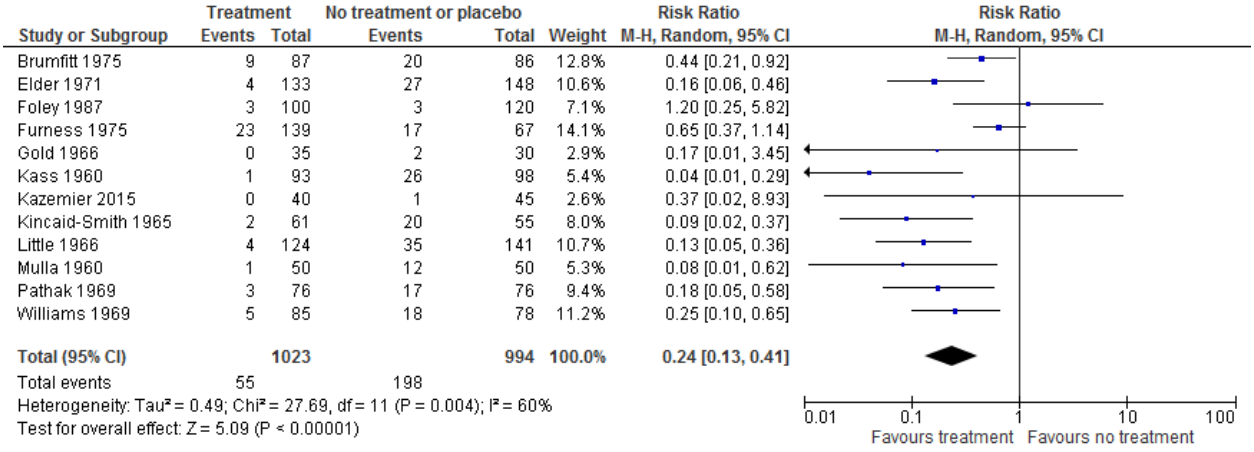


Evidence Set 3. Forest Plots for Subgroup Analyses 3.1.1-3.1.4 – KQ4: Benefits and harms of treatment compared to no treatment

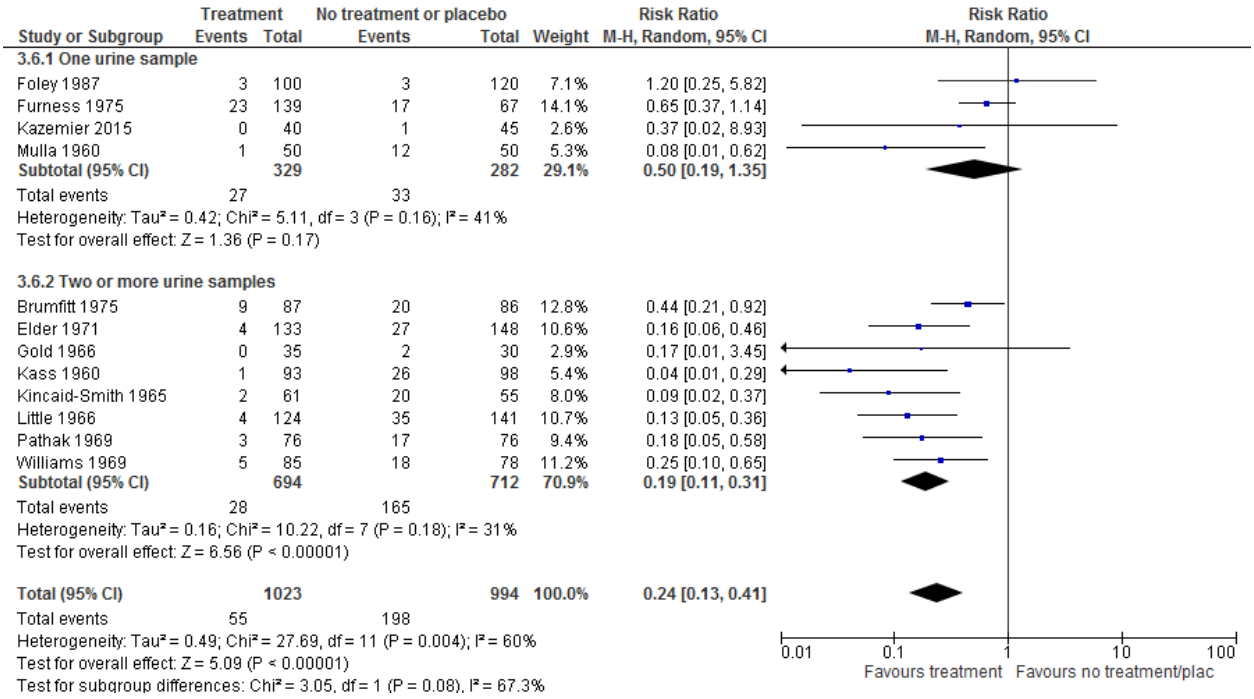
Outcome	No. of studies	No. of participants	Effect size (Risk Ratio; M-H, Random, 95%CI)
3.1 Pyelonephritis (overall)	**12**	**2017**	**0.24 [0.13, 0.41]**
3.1.1 Subgroup analysis: no. of urine samples before confirming bacteriuria and giving treatment			
One urine sample	4	611	0.50 [0.19, 1.35]
Two or more urine samples	8	1406	0.19 [0.11, 0.31]
3.1.2 Subgroup analysis: testing for persistent bacteriuria			
Tested for persistent bacteriuria during pregnancy	8	1352	0.26 [0.15, 0.45]
Testing for persistent bacteriuria post-delivery only	1	206	0.65 [0.37, 1.14]
Testing for persistent bacteriuria during pregnancy and post-delivery	3	459	0.11 [0.05, 0.25]
3.1.3 Subgroup analysis: follow-up			
Follow-up until delivery or puerperium (≤6 wks post-delivery)	9	1558	0.31 [0.18, 0.54]
Follow-up until >6 wks post-delivery	3	459	0.11 [0.05, 0.25]

CI: confidence interval; M-H: Mantel-Haenszel; No.: number; wks: weeks

3.1 Pyelonephritis (overall)

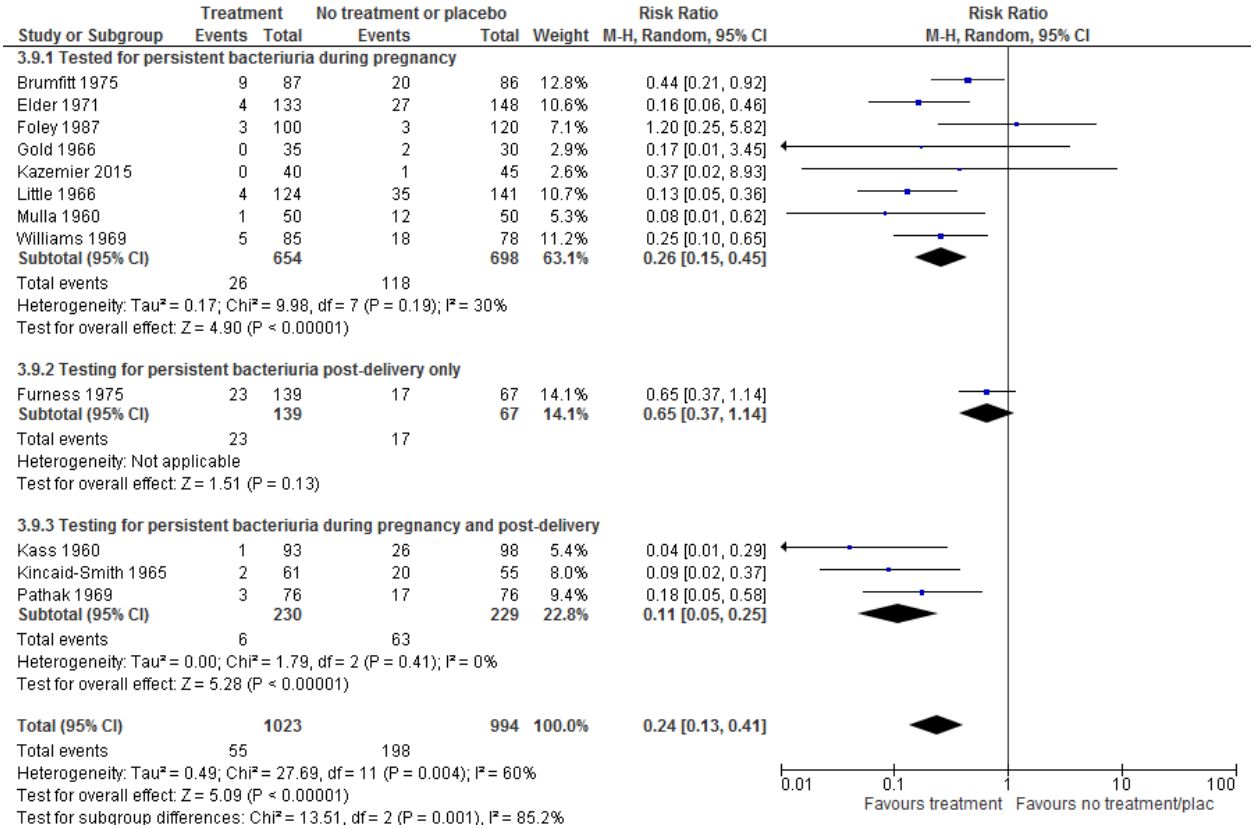


3.1.1 Pyelonephritis subgroup: number of urine samples at each screening visit*

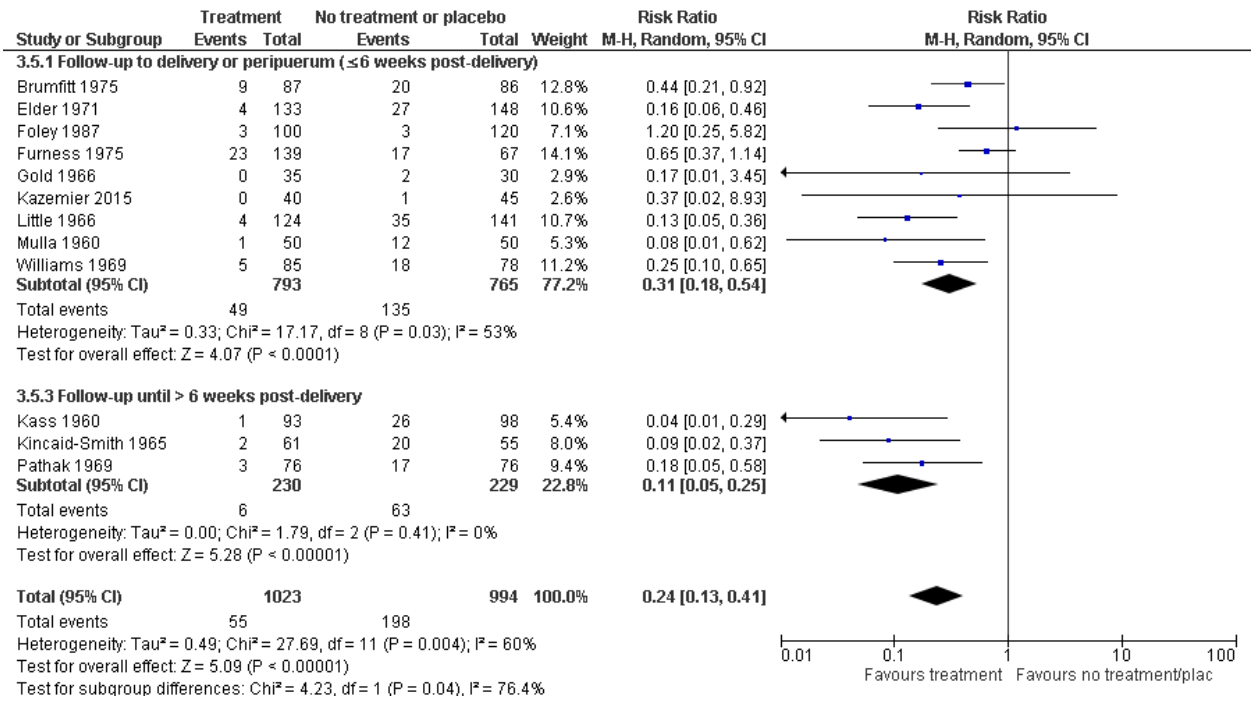


*The additional culture(s) was used to confirm levels of bacteriuria.

3.1.2 Pyelonephritis subgroup: timing of testing for persistent bacteriuria



3.1.3 Pyelonephritis subgroup: duration of follow-up



Author Contributions

LH, AW, JP, RF, MS, KS and BV critically reviewed and contributed to drafts of the report. AW, JP, MG, MS and KS conducted screening, quality assessments, data extraction and analysis. AW, JP, LH, MS, KS and BV contributed to interpretation of results. All of the authors approved the final version of this report.

Acknowledgements

We would also like to acknowledge Tara Landry for peer reviewing the search strategy, Lynn Hampson for conducting search updates for treatment trials, Allison Gates for assistance with quality assessment, and data extraction and verification for the French study, Hamza Jafri for assistance with screening, Megan Nuspl for assistance with screening grey literature, Tara MacGregor for set-up and assistance with DistillerSR, and Evans Kwak and MacKinna Hauff for assistance with article retrieval. We would like to thank Dr. Konstanze Angelescu, Dr. Brenda Kazemier, Mary Ann Rhode, and Dr. Yeşim Uncu for their assistance in providing study details.

External Reviewers

We are grateful to the following reviewers for their feedback on this evidence review, as well as those individuals who preferred to remain anonymous: Drs. Anne Biringer, Linda Brubaker, Venu Jain, Jennifer Marandola, Deborah Money, Lindsay Nicolle, and Fiona Smaill.

Declaration of funding

Funding for the Evidence Review is provided by the Public Health Agency of Canada. The views of the funding bodies have not influenced the content of the guideline; competing interests have been recorded and addressed. The views expressed in this article are those of the authors and do not necessarily represent those of the Public Health Agency of Canada.

List of Abbreviations

AAFP	American Academy of Family Physicians
AAP	American Academy of Physicians
ACOG	American College of Obstetricians and Gynecologists
AE	Adverse Event
AMSTAR	Assessing the Methodological Quality of Systematic Reviews
ARDS	Acute Respiratory Distress Syndrome
ARR	Absolute Risk Reduction
ASB	Asymptomatic Bacteriuria
CADTH	Canadian Agency for Drugs and Technologies in Health
CCT	Controlled Clinical Trial
CFU	Colony-Forming Units
CI	Confidence Interval
CTFPHC	Canadian Task Force on Preventive Healthcare
DARE	Database of Abstracts of Reviews of Effects
EP	Evidence Profile
ES	Evidence Set
GBS	Group B Streptococcus
GRADE	Grading of Assessment, Development and Evaluation
ICER	Incremental Cost-Effectiveness Ratio
IDSA	Infectious Diseases Society of America
IDSOG	Infectious Diseases Society of Obstetrics and Gynecology
KQ	Key Question
NICE	National Institute for Health and Care Excellence
NICU	Neonatal Intensive Care Unit
NNS	Number Needed to Screen
NNT	Number Needed to Treat
NOS	Newcastle-Ottawa Quality Assessment Scale
OIS	Optimal Information Size

OR	Odds Ratio
PHAC	Public Health Agency of Canada
PICOTS	Population, Intervention, Comparator, Timing and Setting
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
QALY	Quality-Adjusted Life-Year
RCT	Randomized Clinical Trial
ROB	Risk of Bias
RR	Risk Ratio
SGA	Small for Gestational Age
SIGN	Scottish Intercollegiate Guidelines Network
SMFM	Society for Maternal-Fetal Medicine
SOF	Summary of Findings
SOGC	Society of Obstetricians and Gynaecologists of Canada
USPSTF	US Preventive Services Task Force
UTI	Urinary Tract Infection

References

1. Glaser AP, Schaeffer AJ. Urinary tract infection and bacteriuria in pregnancy. The Urologic Clinics of North America 2015;42(4);547-60.

2. Schnarr J, Smaill F. Asymptomatic bacteriuria and symptomatic urinary tract infections in pregnancy. European Journal of Clinical Investigation 2008;38(Suppl 2);50-7.

3. Ipe DS, Sundac L, Benjamin WH Jr., Moore KH, Ulett GC. Asymptomatic bacteriuria: Prevalence rates of causal microorganisms, etiology of infection in different patient populations, and recent advances in molecular detection. FEMS Microbiology Letters 2013;346(1);1-10.

4. Smaill FM, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. The Cochrane Database of Systematic Reviews 2015;8(CD000490).

5. Connolly A, Thorp JM Jr. Urinary tract infections in pregnancy. The Urologic Clinics of North America 1999;26(4);779-87.

6. Harris RE. The significance of eradication of bacteriuria during pregnancy. Obstetrics and Gynecology 1979;53(1);71-3.

7. Kass EH. Bacteriuria and the pathogenesis of pyelonephritis. Laboratory Investigation: A Journal of Technical Methods and Pathology 1960;9;110-6.

8. Sweet RL. Bacteriuria and pyelonephritis during pregnancy. Seminars in Perinatology 1977;1(1);25-40.

9. Patterson TF, Andriole VT. Bacteriuria in pregnancy. Infectious Disease Clinics of North America 1987;1(4);807-22.

10. Wing DA, Fassett MJ, Getahun D. Acute pyelonephritis in pregnancy: An 18-year retrospective analysis. American Journal of Obstetrics and Gynecology 2014;210(3);219.e1-6.

11. Rogozinska E, Formina S, Zamora J, Mignini L, Khan KS. Accuracy of onsite tests to detect asymptomatic bacteriuria in pregnancy: A systematic review and meta-analysis. Obstetrics and Gynecology 2016;128(3);495-503.

12. Hilt EE, McKinley K, Pearce MM, et al. Urine is not sterile: Use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. Journal of Clinical Microbiology 2014;52(3);871-6.

13. Price TK, Dune T, Hilt EE, et al. The clinical urine culture: Enhanced techniques improve detection of clinically relevant microorganisms. Journal of Clinical Microbiology 2016;54(5);1216-22.

14. Guinto VT, De Guia B, Festin MR, Dowswell T. Different antibiotic regimens for treating asymptomatic bacteriuria in pregnancy. The Cochrane Database of Systematic Reviews 2010;9(CD007855).

15. Allen VM, Yudin MH, Bouchard C, et al. Management of group B streptococcal bacteriuria in pregnancy. Journal of Obstetrics and Gynaecology Canada 2012;34(5);482-6.

16. United States Preventive Services Task Force. Screening for asymptomatic bacteriuria in adults: U.S. Preventive Services Task Force reaffirmation recommendation statement. Annals of Internal Medicine 2008;149(1);43-7.

17. Smaill F, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. The Cochrane Database of Systematic Reviews 2007;2(CD000490).

18. American Academy of Family Physicians. Summary of recommendations for clinical preventive services. Leawood, KS; 2015.

19. Nicolle LE, Bradley S, Colgan R, et al. Infectious Diseases Society of America guidelines for the diagnosis and treatment of asymptomatic bacteriuria in adults. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 2005;40(5);643-54.

20. American Academy of Pediatrics, American College of Obstetricians and Gynecologists. Guidelines for perinatal care, 7th edition. American Academy of Pediatrics; 2012.

21. NICE (National Institute for Health and Care Excellence). [Internet]. Antenatal care for uncomplicated pregnancies. Clincial guideline 62. [cited: January 28, 2017] Available from: https://www.nice.org.uk/guidance/cg62

22. SIGN (Scottish Intercollegiate Guidelines Network). [Internet]. Sign 88. Management of suspected bacterial urinary tract infections in adults. [cited: January 28, 2017] Available from: http://www.sign.ac.uk/guidelines/fulltext/88/index.html

23. Canadian Task Force on Preventive Health Care. [Internet]. Canadian Task Force on Preventive Health Care: Procedure manual. [cited: January 28, 2017] Available from: http://canadiantaskforce.ca/methods/

24. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than english in systematic reviews. Health Technology Assessment 2003;7(41);1-90.

25. Morrison A, Polisena J, Husereau D, et al. The effect of english-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. International Journal of Technology Assessment in Health Care 2012;28(2);138-44.

26. Canadian Agency for Drugs and Technologies in Health. [Internet]. Systematic reviews/meta-analysis/health technology assessment search filter. [cited: January 27, 2017]. Available from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#syst

27. Canadian Agency for Drugs and Technologies in Health. Grey matters: A practical tool for searching health-related grey literature. Ottawa, ON; 2015.

28. Evidence Partners Inc. [Internet]. Distiller (DistillerSR systematic review software program). [cited: January 26, 2017]. Available from: http://systematic-review.net/

29. Statistics Canada. [Internet]. Table 102-0508 - perinatal mortality (number and rates) and components, Canada, provinces and territories, annual, CANSIM (database). [cited: January 27, 2017]. Available from: http://www5.statcan.gc.ca/cansim/a05?lang=eng&id=01020508

30. Malik K, Human Development Report 2013 Team. Human development report 2013. The rise of the south: Human progress in a diverse world. New York, NY; 2013.

31. Wells GA, Shea B, O'Connell D, et al. [Internet]. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [cited: January 28, 2017] Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

32. Centre for Evidence-based Management (CEBMa). [Internet]. Critical appraisal of a cross-sectional study (survey). [cited: January 28, 2017]. Available from: https://www.cebma.org/resources-and-tools/what-is-critical-appraisal/

33. Higgins JP, Green S. [Internet]. The Cochrane handbook for systematic reviews of interventions. Section 8: Assessing risk of bias in included studies. [cited: January 28, 2017] Available from: http://handbook.cochrane.org

34. Higgins JP, Green S. [Internet]. The Cochrane handbook for systematic reviews of interventions. Table 8.5.d: Criteria for judging risk of bias in the 'Risk of bias' assessment tool. [cited: January 28, 2017] Available from: http://handbook.cochrane.org

35. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. British Medical Journal (Clinical Research Ed) 1997;315(7109);629-34.

36. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. Journal of Clinical Epidemiology 2011;64(4);380-2.

37. Hsu J, Brozek JL, Terracciano L, et al. Application of GRADE: Making evidence-based recommendations about diagnostic tests in clinical practice guidelines. Implementation Science 2011;6;62.

38. McMaster University. [Internet]. GRADEpro (GRADEpro GDT computer software). Available from: http://www.gradepro.org

39. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology 2011;64(4);383-94.

40. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). Journal of Clinical Epidemiology 2011;64(4);407-15.

41. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. Journal of Clinical Epidemiology 2011;64(12);1294-302.

42. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 6. Rating the quality of evidence--imprecision. Journal of Clinical Epidemiology 2011;64(12);1283-93.

43. Gérard J, Blazquez G, Mounac M. Importance of systematic research of urinary infection in pregnant women and the cost of its detection. Proposal for a practical approach. Journal de Gynécologie, Obstétrique et Biologie de la Reproduction (Paris) 1983;12(3);243-51.

44. Gratacós E, Torres PJ, Vila J, Alonso PL, Cararach V. Screening and treatment of asymptomatic bacteriuria in pregnancy prevent pyelonephritis. Journal of Infectious Diseases 1994;169(6);1390-2.

45. Uncu Y, Uncu G, Esmer A, Bilgel N. Should asymptomatic bacteriuria be screened in pregnancy? Clinical and Experimental Obstetrics and Gynecology 2002;29(4);281-5.

46. Rhode MA, Shapiro H, Jones OW 3rd. Indicated vs. routine prenatal urine chemical reagent strip testing. Journal of Reproductive Medicine 2007;52(3);214-9.

47. Lupattelli A, Picinardi M, Einarson A, Nordeng H. Health literacy and its association with perception of teratogenic risks and health behavior during pregnancy. Patient Education and Counseling 2014;96(2);171-8.

48. Butters L, Howie CA. Awareness among pregnant women of the effect on the fetus of commonly used drugs. Midwifery 1990;6(3);146-54.

49. Sanz E, Gomez-Lopez T, Martinez-Quintas MJ. Perception of teratogenic risk of common medicines. European Journal of Obstetrics, Gynecology and Reproductive Biology 2001;95(1);127-31.

50. Kazemier BM, Koningstein FN, Schneeberger C, et al. Maternal and neonatal consequences of treated and untreated asymptomatic bacteriuria in pregnancy: A prospective cohort study with an embedded randomised controlled trial. The Lancet Infectious Diseases 2015;15(11);1324-33.

51. Nordeng H, Ystrom E, Einarson A. Perception of risk regarding the use of medications and other exposures during pregnancy. European Journal of Clinical Pharmacology 2010;66(2);207-14.

52. Twigg MJL, A., Nordeng, H. Women's beliefs about medication use during their pregnancy: A UK perspective. International Journal of Clinical Pharmacy 2016;38;968-76.

53. Sharma R, Kapoor B, Verma U. Drug utilization pattern during pregnancy in north India. Indian Journal of Medical Sciences 2006;60(7);277-87. PMID:2006369645.

54. Mashayekhi SO, Dilmaghanizadeh M, Fardiazar Z, Bamdad-Moghadam R, Ghandforoush-Sattari F. Study of awareness among pregnant women of the effects of drugs on the fetus and mother in Iran. Health Policy 2009;91(1);89-93.

55. Brumfitt W. The effects of bacteriuria in pregnancy on maternal and fetal health. Kidney International 1975;4(Suppl 4);113.

56. Elder H, Santamarina B, Smith S, Kass E. Use of sulfasymazine in the treatment of bacteriuria of pregnancy. Antimicrobial Agents and Chemotherapy 1966;6;142-8.

57. Elder H, Santamarina B, Smith S, Kass E. The natural history of asymptomatic bacteriuria during pregnancy: The effect of tetracycline on the clinical course and the outcome of pregnancy. American Journal of Obstetrics and Gynecology 1971;111(3);441-62.

58. Foley M, Farquharson R, Stronge J. Is screening for bacteriuria in pregnancy worthwhile? British Medical Journal (Clinical Research Ed) 1987;295(6592);270.

59. Furness E, McDonald P, Beasley N. Urinary antiseptics in asymptomatic bacteriuria of pregnancy. The New Zealand Medical Journal 1975;81(539);417-9.

60. Gold EM, Traub FB, Daichman I, Terris M. Asymptomatic bacteriuria during pregnancy. Obstetrics and Gynecology 1966;27(2);206-9.

61. Kass E. The role of asymptomatic bacteriuria in the pathogenesis of pyelonephritis. In: Quinn EL, Kass EH (editors). Biology of pyelonephritis. Boston: Little Brown and Company 1960;399-412.

62. Kincaid-Smith P BM. Bacteriuria in pregnancy. The Lancet 1965;1;1382-7.

63. Little P. The incidence of urinary infection in 5000 pregnant women. The Lancet 1966;288(7470);925-8.

64. Mulla N. Bacteriuria in pregnancy. Obstetrics and Gynecology 1960;16(1);89-92.

65. Pathak U, Tang K, Wifflams L, Stuart K. Bacteriuria of pregnancy: Results of treatment. The Journal of Infectious Diseases 1969;120(1);91-103.

66. Thomsen A, Mo/rup L, Hansen KB. Antibiotic elimination of group-B streptococci in urine in prevention of preterm labour. The Lancet 1987;329(8533);591-3.

67. Williams G, Campbell H, Davies K. Urinary concentrating ability in women with asymptomatic bacteriuria in pregnancy. British Medical Journal 1969;3(5664);212-5.

68. Wren B. Subclinical renal infection and prematurity. The Medical Journal of Australia 1969;2(12);596-600.

69. Angelescu K, Nussbaumer-Streit B, Sieben W, Scheibler F, Gartlehner G. Benefits and harms of screening for and treatment of asymptomatic bacteriuria in pregnancy: A systematic review. BMC Pregnancy and Childbirth 2016;16(1);336.

Appendices

1. Summary of Study Characteristics - KQ1a & b
2. Summary of Characteristics of SR and Studies - KQ4
3. Characteristics of Included Studies – KQ1a & b
4. Characteristics of Included Studies – KQ2
5. Characteristics of Included Studies – KQ4
6. Quality Assessments for Included Studies – KQ1a & b
7. Quality Assessments for Included Studies – KQ2
8. Quality Assessments for Included Studies – KQ4
9. Search Strategies, Grey Literature
10. Excluded Studies – KQ1a & b
11. Excluded Studies – KQ2
12. Excluded Systematic Reviews – KQ4
13. Methods for Integrating Systematic Reviews