“EA139 POINT OF CARE RAPID TESTS FOR DIAGNOSIS OF SYPHILIS INFECTION”

Nicolás Alejandro Torres Montañez

Universidad Nacional de Colombia
Facultad de Medicina, Departamento Obstetricia y Ginecología
Bogotá DC, Colombia
2017

“EA139 POINT OF CARE RAPID TESTS FOR DIAGNOSIS OF SYPHILIS INFECTION”

Nicolás Alejandro Torres Montañez

Tesis o trabajo de investigación presentado como requisito parcial para optar al título de:

Especialista en Obstetricia y Ginecología

Director (a):

Doctor Carlos Fernando Grillo Ardila

Codirector (a):

Doctor Jairo Amaya Guío
Doctora Edith Ángel Müller

Grupo de Investigación:

Enfermedades infecciosas Departamento Obstetricia y Ginecología

Universidad Nacional de Colombia
Facultad de Medicina, Departamento Obstetricia y Ginecología
Bogotá DC, Colombia
2017

A mi familia, por su apoyo durante todo este proceso

A mis profesores, por su tiempo incondicional, su constante motivación, su persistente empeño y su paciencia.

Resumen

La sífilis es una infección compleja, curable y de transmisión sexual causada por la bacteria Treponema pallidum, que tiene un curso clínico variable. Es causa de enfermedad aguda e importantes consecuencias médicas y psicológicas en una pequeña proporción de la población infectada con la bacteria. La sífilis también es causa de discapacidad a largo plazo y muerte de miles de hombres, mujeres y niños. El diagnóstico de sífilis está basado en la combinación de la historia clínica, presentación de los síntomas y el resultado de las pruebas serológicas. Existen dos tipos de pruebas diagnósticas: treponémicas y no treponémicas. Las pruebas treponémicas son costosas, requieren equipos de laboratorio, el uso continuo de electricidad, personal entrenado y reactivos adecuados, y raramente están disponibles fuera de los laboratorios. La mejoría en el acceso a un diagnóstico apropiado y un tratamiento oportuno puede reducir el impacto de la enfermedad. Aunque algunas infecciones pueden ser manejadas sin la necesidad de prueba diagnósticas (manejo sindrómatico), esto no es apropiado para enfermedades asintomáticas en las cuales una prueba positiva es necesaria para que el tratamiento pueda ser indicado. Este trabajo tiene como objetivo determinar la exactitud diagnóstica de las prueba rápidas para el diagnóstico de la Sífilis, a cualquier edad, en hombres y mujeres no gestantes en edad reproductiva, confirmado mediante la combinación de pruebas treponémicas y no treponémicas como patrón de referencia, así como evaluar la exactitud de estas pruebas de acuerdo al género, tipo de prueba rápida (no treponémica, treponémica), estadio de la infección (primaria, secundaria, latente y tardía) y la infección por VIH, aplicada en condiciones sin recursos de laboratorio.

Ocho estudios (11.783 participantes) cumplieron nuestros criterios de inclusión. Tres estudios fueron patrocinados por la industria farmacéutica. Nuestros hallazgos indican que, con evidencia de baja calidad, las pruebas rápidas para el diagnóstico de sífilis tienen una moderada sensibilidad (83% IC 0.71 to 0.91) y una alta especificidad (98% IC 0.96 to 0.99). Los Forest plots y la curva ROC demostraron un alto grado de heterogeneidad entre los estudios incluidos, la cual fue mayor para la sensibilidad que para la especificidad.

Palabras clave: sífilis, pruebas rápidas, revisión sistemática, Treponema pallidum, diagnóstico.

Abstract

Syphilis is a complex, curable sexually-transmitted infection caused by the Treponema pallidum bacterium that has a variable clinical course. It is a cause of acute illness and serious medical and psychological consequences in a small proportion of people infected with the bacteria. Syphilis is also the cause of long-term disability and death of thousands of men, women and infants. Diagnosis of syphilis is based on a combination of clinical history, symptom presentation, and serologic test results. There are two types of diagnostic tests, treponemal and non-treponemal. The treponemal tests are expensive, laboratory-based, require a continuous supply of electricity, reagents and trained staff, and are rarely available outside of reference laboratories. The improvement in access to appropriate diagnosis and treatment for infectious diseases can reduce the burden of disease. Although some infections can be managed without the need for diagnostic tests, this is not appropriate for asymptomatic infectious diseases in which a positive diagnostic test is needed before treatment can be given. The objective of this review was to determine the diagnostic accuracy of the Point of care rapid tests for the diagnosis of syphilis, at any age, in men and nonpregnant women of reproductive age, as verified with the combination of both reactive non-treponemal and treponemal tests as the reference standard, as well as the behavior according to gender, type of POC test (non-treponemal, treponemal), stage of infection (primary, secondary, latent and late), and HIV infection status, in laboratory free conditions.

Eight studies (11.783 patients) fulfill our inclusion criteria. Three of them were funded by laboratories. We found with low quality evidence that the rapid tests for the diagnosis of syphilis have moderate sensitivity (83% CI 0.71 to 0.91) and high specificity (98% CI 0.96 to 0.99). Forest plots ant the ROC plot demonstrated a high degree of heterogeneity between papers, which was greater for sensitivity than for specificity

Keywords: syphilis, rapid tests, point of care, systematic review, Treponema pallidum, diagnosis, primary care.

TABLE OF CONTENTS

Summary

Background

Syphilis is a complex, curable sexually-transmitted infection caused by the Treponema pallidum bacterium that has a variable clinical course. It is a cause of acute illness and serious medical and psychological consequences in a small proportion of people infected with the bacteria. Syphilis is also the cause of long-term disability and death of thousands of men, women and infants. Diagnosis of syphilis is based on a combination of clinical history, symptom presentation, and serologic test results. There are two types of diagnostic tests, treponemal and non-treponemal. The treponemal tests are expensive, laboratory-based, require a continuous supply of electricity, reagents and trained staff, and are rarely available outside of reference laboratories. The improvement in access to appropriate diagnosis and treatment for infectious diseases can reduce the burden of disease. Although some infections can be managed without the need for diagnostic tests, this is not appropriate for asymptomatic infectious diseases in which a positive diagnostic test is needed before treatment can be given.

Objectives

To determine the diagnostic accuracy of rapid tests at point of care (POC) for detecting syphilis infection, at any stage, in men and non-pregnant women of reproductive age, as verified with the combination of both reactive non-treponemal and treponemal test as the reference standard.

Search methods

We searched the Cochrane Sexually Transmitted Infections Group Specialized Register (18 September 2016), CENTRAL, MEDLINE, Embase, LILACS, World Health Organiazation (WHO), International Trials Registry Platform (ICTRP), Web of Science, System for Information on Grey Literature in Europe, Health Services Research Projects in Progress (HSRProj), the Database of Abstracts of Reviews of Effect (DARE), International Society for Sexually Transmitted Diseases Research (ISSTDR), British Association for Sexual Health and HIV (BASHH), International Congress on Infectious Diseases (ICID), The International Union against Sexually Transmitted Infections (IUSTI), International Society for Infectious Diseases (ISID), International Meeting on Emerging Diseases and Surveillance (IMED), Interscience Conference on Antimicrobial Agents and Chemotherapy (ICAAC), The International Federation of Gynecology and Obstetrics

(FIGO). We also handsearched conference proceedings, contacted trial authors and reviewed the reference lists of retrieved studies.

Selection criteria

We included accuracy or validity studies. Participants in included studies should have been enrolled prospectively and consecutively or through random sampling. Only studies reporting that all participants received the index test and the reference standard and presenting 2 × 2 data will be eligible for inclusion. We included men and non-pregnant women at reproductive age, recruited at primary or secondary care facilities without previous diagnostic testing, who attended an outpatient facility.

Data collection and analysis

Two review authors independently applied inclusion and exclusion criteria in selecting potential titles and abstracts of studies retrieved as a result of the search. Two authors extracted the data and assessed the risk of bias in the included studies. Disagreements were resolved through consensus. We assessed the quality of the evidence using the GRADE approach.

Main results

Eight studies (11.783 patients) fulfill our inclusion criteria. Three of them were funded by laboratories. All included studies were conducted from 1999 to 2015, with a mean prevalence of 10.1%. Eleven rapid test were assessed. Six of the included studies were multicentric, and all the included papers were published in english. Seven of the included studies recruited high risk population (e.g. female sex workers) and one recruited women with low genital tract symptoms. We , with low quality evidence, that the rapid tests for the diagnosis of syphilis have moderate sensitivity (83% CI 0.71 to 0.91) and high specificity (98% CI 0.96 to 0.99). Forest plots ant the ROC plot demonstrated a high degree of heterogeneity between papers, which was greater for sensitivity than for specificity.

Authors` conclusions

Low quality evidence shows that Point of care rapid tests could be an accurate method for the diagnosis of syphilis in high risk population and in those with high probability of following loss. The performance of these rapid tests disagrees with the findings in other systematic reviews and the available studies are low quality. It is needed the publication of more studies applied in general population, with low risk and in laboratory-free

conditions to assess the accuracy of the rapid test in real clinical scenarios and no in controlled conditions.

PLAIN LANGUAGE SUMMARY

POINT OF CARE RAPID TEST FOR DIAGNOSIS OF SYPHILIS YIN MEN AND NONPREGNANT WOMEN

Review question

We assessed the diagnostic accuracy of Point of care rapid tests for the diagnosis of syphilis infection in men and nonpregnant women

Background

Syphilis is a complex, curable sexually-transmitted infection caused by the Treponema pallidum bacterium that has a variable clinical course. It is a cause of acute illness and serious medical and psychological consequences in a small proportion of people infected with the bacteria. Diagnosis of syphilis is based on a combination of clinical history, symptom presentation, and serologic test results. The improvement in access to appropriate diagnosis and treatment for infectious diseases can reduce the burden of disease.

Trial characteristics

Cochrane researchers searched the available literature up to 18 September 2016 an included eight studies with 11.783 participants. The studies included men and nonpregnant women in reproductive age, between 14 to 54 years. Seven studies recruited women as high risk population (female sex workers) and only one recruited men along with women. We assessed a total of eleven rapid test, which were all treponemal tests. Three studies assessed more than one rapid test. All the studies reported the performance of the rapid tests, only one reported the cost-effectiveness data related with the use of the rapid tests and the administration of immediate treatment right after the diagnosis. Three of the included studies were funded by the pharmaceutical companies.

Key results

Compared with the a gold standard composed of a treponemal and a non-treponemal test, the Point of care rapid tests showed a moderate accuracy for the diagnosis of syphilis and might be a reasonable option for the diagnosis in high risk population and those patients with a high probability of following loss.

Quality of the evidence

The quality of the evidence was low due to the population recruited in seven of the eight included studies. This finding could be misunderstood given the fact that the performance of the test will be better in high risk population and in those with high prevalence of the disease.

BACKGROUND

Target condition being diagnosed

Syphilis is a complex, curable sexually-transmitted infection caused by the *Treponema pallidum* bacterium that has a variable clinical course (CDC 2014). It is a cause of acute illness and serious medical and psychological consequences in a small proportion of people infected with the bacteria. Syphillis is also the cause of long-term disability and death of thousands of men, women and infants (Saloojee 2004). The Center for Disease Control (CDC) defines syphilis as: **Syphilis, primary:** A stage of infection characterised by one or more ulcerative lesions (e.g. chancre); **Syphilis, secondary:** a stage of infection characterised by localised or diffuse mucocutaneous lesions (e.g. rash such as non-pruritic macular, papular, or pustular lesions), often with generalised lymphadenopathy; **Syphilis, early latent:** a person without clinical signs or symptoms of syphilis when the initial infection has occurred within the previous 12 months; **Syphilis, late latent:** a person without clinical signs or symptoms of syphilis when the initial infection has occurred >12 months previously; **Syphilis, late, with clinical manifestations; and Syphilis, Con- genital** (CDC 2014 a). Syphilis is the most common disease during the years of peak sexual activity. In 2012, the estimated global prevalence was 0.5% (0.4% to 0.6%) in women aged 15 to 49 years and 0.48% (0.3% to 0.7%) in men. These figures correspond to an estimated 6 million new cases of syphilis (4 to 8 million) each year worldwide. The difference in the male-to-female ratio is attributable to the increased rate of disease among men who have sex with men (Janier 2014).

Current diagnostic strategies

Diagnosis of syphilis is based on a combination of clinical history, symptom presentation, and serologic test results. There are two types of diagnostic tests, treponemal and non-treponemal (PAHO 2015).

The non-treponemal tests such as Venereal Disease Research Laboratory (VDRL) or rapid plasma reagin (RPR) measure the host's response to non-treponemal antigens such as cardiolipin and lecithin released from the damaged host cells, as well as lipoprotein-like material released from the treponema. These non-treponemal tests are generally considered to be sensitive in early syphilis, but their disadvantage being false-positive reactions due to cross-reactivity with autoimmune diseases, collagen diseases and infections such as pian and leprosy. They also have false-negative reactions due their reduced sensitivity in primary syphilis and late latent syphilis, (sensitivity: primary syphilis 78% to 86% compared with dark- field confirmed cases (Creegan 2007)), secondary syphilis 100%, latent syphilis 96% to 98%, late syphilis 71% to 73%; specificity 98% in all stages, compared with *Treponema pallidum* Haemagglutination (TPHA) test (Naidu 2012), and the potential for false- negative results due to prozone phenomenon (Liu 2014). The result of the non- treponemal test is given in titers, with a result of 1:8 titers very suggestive of recent/active infection. The treponemal tests such as fluorescent treponemal antibody absorbed (FTA-ABS), *Treponema pallidum* particle agglutination (TP-PA), enzyme immunoassay (EIA), chemiluminescence immunoassay (CIA) or equivalent serologic methods have high sensitivity for all the stages of disease other than very early primary syphilis (sensitivity 84% to 96%, specificity 98% in all stages). These tests detect human serum/plasma antibodies to *Treponema pallidum* by means of an indirect hemagglutination method. The results of these tests are given in a dichotomy method, being "positive" or "negative" the possible results (Naidu 2012).

Currently, there are two common approaches to the diagnosis of syphilis using serological tests: the traditional algorithm and the reverse algorithm. The CDC recommends the traditional algorithm (CDC 2014). This algorithm begins the screening with a non-treponemal test and confirms a positive result with a treponemal test (CDC 2014; PAHO 2015; WHO 2003). However, the treponemal-based tests remain positive for life and cannot distinguish between recent, active infection and previously treated or old, non-contagious infection (PAHO 2015*).* A reverse testing algorithm utilizes a treponemal primary screening assay followed by a non- treponemal test if the primary treponemal assay is positive. If the secondary, non-treponemal test is reactive, then active syphilis is confirmed (CDC 2008). In the literature, it is considered as active syphilis if a positive treponemal test and a non-treponemal test is reactive with different thresholds, with a non- treponemal test at any titer, more than two dilutions or more than four dilutions (Kay 2014).

The presence of a hard chancre in genitals is suggestive of syphilis, and its treatment is considered in the syndromic approach of genital ulcers, because the treponemal test and the non-treponemal test initially can both be negative. The reactive syphilis test in combination with a typical clinical sign of syphilis such a chancre, skin ulcer, or rash is highly suggestive of the disease (PAHO 2015). In the absence of symptoms, a combination of both reactive non- treponemal and treponemal test indicates the possibility of contagious syphilis infection, and supports the need for treatment of the individual and any sex partners (PAHO 2015).

The treponemal tests are expensive, laboratory-based, require a continuous supply of electricity, reagents and trained staff, and are rarely available outside of reference laboratories. (Peeling 2004) As a result in many countries, treatment is based on non-treponemal tests results, which are cheaper and have more availability, leading to over-treatment of patients, due to false-positive results, besides when the patient has a non-treponemal positive test, the treponemal test is requested for confirmation, and if the patient could not return to the healthcare facility, the opportunity for treatment is lost (Jafari 2013).

Index test(s)

In resource-limited settings, access to screening is limited and the risk of patients lost to follow-up is high. So, syphilis rapid tests to be applied at point of care (POC), which detect antibodies to *Treponema pallidum* antigen or anticardiolipin antibodies, have become popular in those settings due to their advantages: the quickness in giving results, the possibility of giving treatment immediately, are performed with minimal technical training in non-laboratory settings, and detect the disease at the clinical setting (PAHO 2015; Peeling 2004). In addition, as the test results are obtained on the same day, the test is applied, treatment can be provided right away.The risk of over- treatment given by false positives is less dangerous than the secondary risk of not treating the infection, so every patient with positive result must be given treatment (Jafari 2013).

In order to define the ideal characteristics of a rapid and POC test for the detection of sexually transmitted infections, the WHO Sexually Transmitted Diseases Diagnostic Initiative (SDI) established the ASSURED criteria: the test has to be Affordable, Sensitive, Specific, User-friendly, Rapid and Robust, Equipment-free and Delivered to end users.

The setting of these criteria has the potential to make an impact in averting congenital syphilis in primary healthcare settings (Mabey 2012; Peeling 2006; Peeling 2010).

There are two types of POC test: non-treponemal and treponemal. There is a device which offers the ability to provide the antibody detection of both tests. The treponemal tests have two variants: immunochromatographic or agglutinant. The second type of POC, non-treponemal test, is useful because it indicates an active infection. The combined test detects treponemal and non treponemal antibodies. In a high-prevalence setting, the combined test may significantly reduce over-treatment (Greer 2008). Another type of test combines the diagnosis of syphilis and HIV, this test HIV/ syphilis Duo test, is a method of qualitative detection that uses as method the immunochromatography to detect immunoglobulin (Ig) IgG, IgM and IgA antibodies for specific- HIV antigens and the *Treponema pallidum* recombinant antigen in serum (Bristow 2014).

In a recent meta-analysis the sensitivity of treponemal POC tests varied from 74.26% to 90.04% and specificity from 94.15% to 99.58% (Jafari 2013). The combined test has a sensitivity of 98.4% when the standard is a non-treponemal test (RPR) greater than 1: 8 dilutions. but it has been seen that sensitivity falls to 88% when the RPR is negative (Singh 2015). Series that compare syphilis POC tests with the combination of treponemal and non-treponemal tests, have shown a concordance with approximately 90.6%, with a specificity greater than 95% and a sensitivity between 60% to 100% depending on the test (Gaydos 2014). The implementation of rapid testing must be accompanied by quality assurance systems and technical competence (Benzaken 2014).

Clinical pathway

The improvement in access to appropriate diagnosis and treatment for infectious diseases can reduce the burden of disease. Although some infections can be managed syndromically without the need for diagnostic tests, this is not appropriate for asymptomatic infectious diseases in which a positive diagnostic test is needed before treatment can be given (Peeling 2010). In many situations, it is not possible to perform or to wait for the results of laboratory tests, because the delay represents opportunities lost for treatment, which can spread the infection of the STD, therefore diagnosis depends on the availability of POC tests (Peeling 2010). POC tests for STI would provide the ability to offer immediate testing and treatment in a single encounter to mitigate further spread of the disease (Singh 2015).

The POC test for syphilis can be implemented by clinicians, health services or auxiliary personnel in diverse places without laboratory support, under different scenarios regardless of healthcare systems levels, covering a broad patient spectrum, especially in high-risk populations. The clinical pathway for this diagnostic approach can be summarised into two clinical presentations. The first one is the presence of an asymptomatic patient in which the lack of an opportune diagnostic test will be reflected in an absence of an appropriate treatment increasing in consequence, the burden of the condition and serious and deleterious sequelae. Examples for this scenario could be a pregnant woman with syphilis (CDC 2006; Larson 2014), people in developing countries, peripheral health facilities, or remote rural populations that do not have access to laboratory services (Bien 2015; Peeling 2010), people in resource-limited settings (Gaydos 2014), in high-risk populations (sexual workers, men who have sex with men, LGBTI (lesbian, gay, bisexual, transgender and intersex) population, drug addicts) (Gaydos 2014), or in syphilis-control programs for female, male and transgender sex workers (Chen 2012; Gupte 2011).

The second clinical scenario where the POC test could be used is in the case of symptomatic patients. In this population, the syphilis infection can be suspected by unspecific signs of infection, such as hard chancre, inguinal adenopathies, exanthema or condylomata lata. For these patients, the prompt confirmation accompanied by the immediate treatment of the disease could relieve the infection and break the chain of transmission, decreasing the infection rates (Gaydos 2014).

Rationale

Considering the burden and economic costs associated with the condition, even in settings with low prevalence (Larson 2014), it is highly desirable to undertake a critical appraisal of the available evidence of the diagnostic accuracy of the different syphilis rapid POC tests. There is therefore, a need for high-quality systematic reviews to improve the diagnosis of syphilis. This systematic re- view will facilitate the synthesis of the current evidence, and recognise the strengths and weaknesses, address the uncertainty of the current knowledge, and make it possible to assess the effectiveness and safety of this intervention.

Knowing the performance, advantages and limitations of the POC test, could facilitate the decision- making process at individual, organisational, and healthcare systems levels (Larson 2014). The formulation of public health policy focused on early diagnostic and treatment of the infected patient, eliminating the treatment delays and cutting the infection spread, even in clinical settings can help to reduce the burden of syphilis (Singh 2015).

OBJECTIVES

To determine the diagnostic accuracy of rapid tests at point of care (POC) for detecting syphilis infection, at any stage, in men and non-pregnant women of reproductive age, as verified with the combination of both reactive non-treponemal and treponemal test as the reference standard.

Secondary objectives

To assess the accuracy of rapid POC testing according to gender, type of POC test (non-treponemal and treponemal), stage of infection (primary, secondary, latent and late), and HIV infection status.

METHODS

Criteria for considering studies for this review

Types of studies

We included accuracy or validity studies. Participants in included studies should have been enrolled prospectively and consecutively or through random sampling. Only studies reporting that all participants received the index test and the reference standard and presenting 2 × 2 data will be eligible for inclusion. We excluded diagnostic case-control studies because this is not an appropriate design for diagnostic test studies.

Participants

We included men and non-pregnant women at reproductive age, recruited at primary- or secondary-care facilities without previous diagnostic testing, who are attending an outpatient facility. We did not include pregnant women because there is another review with this population.

Index tests

Rapid tests at POC from whole blood, serum or plasma, regard- less of the type of POC test (non- treponemal or treponemal) or technique (Immunochromatographic, agglutination or any other technology).

Target conditions

Active syphilis Infection (primary, secondary, latent and late).

Reference standards

A combination of both reactive non-treponemal (positive at any titer) and treponemal test (positive result) or identification of *Treponema pallidum* by dark-field microscopy.

Search methods for identification of studies

We developed a highly-sensitive, systematic search strategy to identify as many relevant accuracy or validity studies, irrespective of their language, and publication status (published, unpublished, in press, and in progress). We used both electronic searching in bibliographic databases and handsearching, as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

The results of all searches were downloaded and managed using Endnote bibliographic software. Duplicate records of the same study were deleted.

Electronic searches

We contacted with the Information Specialist of the Cochrane Sexually Transmitted Infections (STI) Review Group in order to implement a comprehensive search strategy to capture as many relevant studies as possible in electronic databases. For this purpose, we used a combination of exploded controlled vocabulary (MeSH, Emtree, DeCS) and free-text terms (considering spelling variants, plurals, synonyms, acronyms and abbreviations) for the index tests (point-of-care test, point of care, point of care testing, point of care devices, point of care diagnostic, point of care laboratory, POC, POCT, rapid

test, rapid test device, self testing, self test, patient self testing) and the target condition (syphilis, Treponema pallidum), with field labels, truncation, proximity operators and boolean operators. The search strategies and their results can be found in Appendix 1 (Electronic search strategies). Specifically, we searched the following electronic databases.

• MEDLINE, Ovid platform: inception to present.
• MEDLINE In-Process & Other Non-Indexed Citations, Ovid platform: inception to present. • MEDLINE Daily Update, Ovid platform: inception to present.
• Embase.com: inception to present.
• Cochrane Central Register of Controlled Trials, Ovid platform: inception to present.
• LILACS, iAHx interface: inception to present.

Searching other resources

We attempted to identify additional relevant studies by using of the following methods.
• Searching the Cochrane STI Review Group's Specialized Register, which includes randomised controlled trials (RCTs) and controlled clinical trials (CCTs), from 1944 to 2014, located through the following.
◦ Electronic searching in MEDLINE, EMBASE and CENTRAL.
◦ Online handsearching in those journals not indexed in MEDLINE or Embase, according to the journals' master list of the Cochrane STI Review Group.
• Searching trials registers.
◦ WHO International Clinical Trials Registry Platform
ICTRP portal (http://apps.who.int/trialsearch/): inception to present.
• Searching in Web of Science®: inception to present.
• Searching for grey literature in System for Information on Grey Literature in Europe "OpenGrey" (http:// www.opengrey.eu/): inception to present.
• Searching in Health Services Research Projects in Progress (HSRProj), and the Database of Abstracts of Reviews of Effect (DARE) for additional articles.
• Handsearching of conference proceeding abstracts in the following events.
◦ The International Society for Sexually Transmitted Diseases Research - ISSTDR (http:// www.isstdr.org/): 2007, 2009, 2011, 2013 and 2015.
◦ The British Association for Sexual Health and HIV - BASHH (http://www.bashh.org/): 2014 and 2015.

◦ International Congress on Infectious Diseases - ICID (http://www.isid.org/): 2010, 2012 and 2014.
◦ The International Union against Sexually Transmitted Infections - IUSTI (http://www.iusti.org/): 2011, 2012, 2013, 2014 and 2015.
◦ International Society for Infectious Diseases - ISID (http://www.isid.org/): 2011, 2012, 2013, 2014 and 2015.
◦ International Meeting on Emerging Diseases and
Surveillance - IMED (http://www.isid.org/): 2007, 2009, 2011, 2013 and 2014.
◦ Interscience Conference on Antimicrobial Agents and Chemotherapy - ICAAC (http://www.icaac.org/): 2011, 2012, 2013, 2014 and 2015.
◦ The International Federation of Gynecology and Obstetrics - FIGO (http://www.figo2012.org/ home/): 2009, 2012 and 2015.
• Handsearching within previous systematic reviews and other relevant publications on the same topic.
• Handsearching within reference lists of all relevant studies identified by others methods.

Finally, we will search the citation lists from reviewed articles.

Data collection and analysis Selection of studies

Two review authors (NT, LV) independently applied inclusion and exclusion criteria in selecting potential titles and abstracts of studies retrieved as a result of the search. Disagreements were resolved through consensus or, if required, by consultation with a third review author (MT). We retrieved the full text of a study if we had doubts about whether the study should be included or excluded.

Data extraction and management

We designed a data extraction form. For eligible studies, four review authors (EA, JA, LV, NT) extracted data independently using the form. Discrepancies were resolved through consensus or, if required, by consultation with a third review author (CFG-A). The data extraction form included the following information.

• Methods

◦ Country of the study. Setting.

◦ Basic study design.
◦ Power calculation.
◦ Number of participants and sampling of patients.
◦ Ethical issues.

• Participants

◦ Inclusion and exclusion criteria.

◦ Baseline information on participants: presentation at recruitment and characteristics (e.g. symptoms, presence of risk factors, socio-demographic characteristics and clinical stage).
◦ HIV status.

• Index test

◦ POC specimen: whole blood, serum or plasma,
◦ POC type: single non-treponemal, single treponemal or combined tests, and
◦ POC technology: inmunochromatographic, agglutinant or any other technology. •

Outcomes
◦ True positives, false positives, false negatives, true negatives.
◦ Proportion of patients treated at any time after the diagnosis and the same day of the diagnosis. ◦ Proportion of sexual partners reported as treated.
◦ Proportion of patients with follow-up loss after diagnostic testing.
◦ Acceptability of the test

• Adverse events related to rapid tests (pain, local rash, local infection, numbness).
• Cost- effectiveness.

We collated and presented this information in ’Characteristics of included studies’ tables. We added the data to Review Manager 5.3 (Revman 2014), and two review authors (NT, LV) independently assessed the accuracy of the data. Differences were resolved through consensus or by evaluation by a third review author (MT). When information regarding

any of the above is unclear, we contacted the authors of the original reports to request further details.

Assessment of methodological quality

We assessed the quality of included articles by using a modified version of the Quality Assessment of Diagnostic Accuracy Studies- 2 (QUADAS-2) tool (Whiting 2011). Two review authors (JA, MT) independently performed the quality assessment using the four key domains to assess risk of bias and concerns regarding the applicability to the research question (patient selection, index test, reference standard, and flow-timing domains). We scored the papers as having a ’low’, ’high’ or ’unclear’ risk of bias for each of four domains, and for the patient selection, index test and reference standard applicability. Studies classified as high or unclear risk of bias and/or high concern regarding applicability in at least one domain were regarded as having low methodological design. In the case of disagreements, differences were resolved through consensus or by consultation with a third review author (CFG-A). In the Quadas-2 tool, in the flow- timing domain, we added the following two questions, for the evaluation of the quality.

• Were withdrawals from the study explained?
• Were uninterpretable/intermediate test results reported?

The results were presented in graphics according to each study an as a summary of all the studies.

Statistical analysis and data synthesis

We summarised diagnostic test accuracy by creating a 2 × 2 table for each study based on information retrieved directly from the papers. Each table contained false-positive, false-negative, true- positive, and true-negative rates. Two review authors (CFG-A, JA) independently entered the data into Revman 2014. Discrepancies were resolved by consensus or, if required, by consultation with a third review author.

In the first instance, we analysed in a descriptive way all data retrieved from the included studies. For this purpose and given that results of the POC are reported qualitatively (positive or negative), we presented the results by plotting their sensitivity and specificity (and their 95% confidence intervals) both in forest plots and in a scatter plot in receiver

operating characteristic (ROC) space. For the meta-analysis of diagnostic accuracy measures, we used the bivariate model (Reitsma 2005). For studies with a common threshold, this model takes into account within- study variation and between-study variation and focuses on estimating a summary operating point (i.e. a summary value for sensitivity and specificity). In addition, we estimated the 95% confidence region and the 95% prediction region around the summary operating point. We performed these analyses using the command xtmelogit in STATA, according to the licenses available.

We included a 'Summary of findings' table using the GRADE (Grades of Recommendation, Assessment, Development and Evaluation) approach to diagnostic test accuracy (Hsu 2011), using the template provided in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (DTA)* (Bossuyt 2013). We presented this summary table to provide a more accessible perspective of diagnostic information to healthcare providers and other end users.

Investigations of heterogeneity

We explored heterogeneity initially by performing visual inspection of forest plots of sensitivities and specificities and visual examination of the prediction region. We formally assessed the source of heterogeneity by examining differences in diagnostic accuracy between subgroups of studies. Again, we used the bivariate method to analyse how the summary estimate of sensitivity and specificity varies according to study level covariates. For this purpose, we created a factor variable with N categories and will generate an N-1 dummy that will be entered into the bivariate model to test the effects of covariates on both sensitivity and specificity (Bossuyt 2013).

We defined the sources of heterogeneity a priori and included the following factors: gender, type of population (sex workers, men who have sex with men, drug addicts), types of POC test (non- treponemal or treponemal), stage of infection (primary, secondary, latent and late) and HIV infection status.

Sensitivity analyses

We performed sensitivity analysis for aspects of the review that might affect the results, such as risk of bias associated with the quality of included trials based on an overall 'Risk

of bias' assessment (low versus unclear and high risk of bias) according to QUADAS- 2 patient selection, index test, reference standard, and flow-timing domains.

Assessment of reporting bias

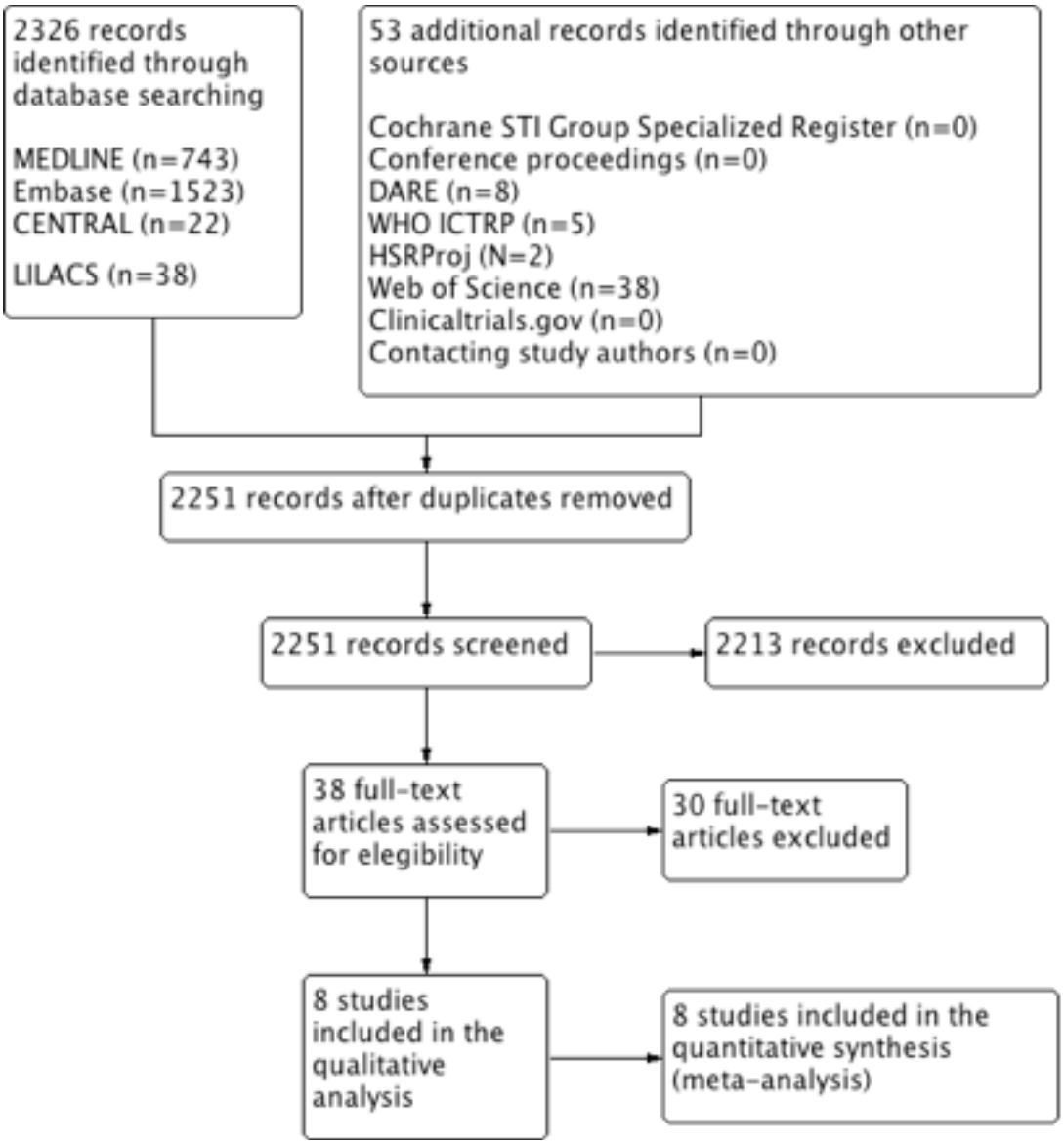
We would investigate publication bias if we found 10 or more studies for inclusion in this systematic review. We investigated reporting bias by using funnel plots. We assessed funnel plot asymmetry visually, and if asymmetry was suggested by a visual assessment, we performed exploratory analysis using Deeks' test to investigate the asymmetry and using diagnostic odds ratio (DOR) as a measure of test accuracy (van Enst 2014).

Result of the search

We searched the available literature up to 18 September in 2016 and retrieved a total of 2.326 references, of which we screened 2251 after we removed duplicates. Of these, we initially screened the full-text articles of 38 references. Eight studies met our inclusion criteria (Benzaken 2008, Black 2016, Campos 2006, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez-Forero 2015, West 2002). We excluded 30 studies.

Basic features of included studies

We have presented the list of the included studies under Characteristics of included studies. One study was conducted in Gambia (West 2002), one in Brazil (Benzaken 2008), one in Colombia (Nuñez-Forero 2015), one in South africa (Black 2016), one in Mexico (Juarez-Figueroa 2007), one in India (Mishra 2010), one in Bangladesh (Nessa 2008) and one in Peru (Campos 2006). None of these studies were conducted in a university hospital. All the included studies were conducted between 1999 (Juarez-Figueroa 2007) and 2015 (Black 2016). Six of the included studies were multicentric, and two used only one center (Juarez-figuera 2007 and Nessa 2008). All trials recruited patients by consecutive sampling, assessing the performance in field settings conditions, and the whole population was received in outpatient services including consultation and specialized reproductive health services. Besides, three studies (Black 2016, Campos 2006, Mishra 2010) recruited patients by the use of a mobile outreach team. Two of the included studies (Nuñez-Forero 2015, Black 2016) calculated the sampling size. All the papers were published in English.



Population

All included studies assessed men or non pregnant women in reproductive age. All the included studies recruited only women, except for one study that allowed the entry of men also (Benzaken 2008). Six studies recruited high risk population, mainly female sex workers (Black 2016, Campos 2006, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008) or their costumers. Another study recruited only women with low genital tract symptoms (Nuñez-Forero 2015) and only one included general population (West 2002). Five of the included studies provided information about age rank for the studied population, which was from 14 to 54 years old. Three studies did not mention age of the included population (Black 2016, Campos 2006, Juarez-Figueroa 2007). Two studies performed a

gynecological examination before the entry to the trial (Nuñez-Forero 2015, Nessa 2008). One study excluded women with hysterectomy or with previous antibiotic treatment in the past few weeks (Nuñez-Forero 2015) and another excluded women with syphilis antecedent (Mishra 2010). The remaining studies did not mentioned specific exclusion criteria. Four studies metioned that the included population was mainly heterosexual, single, with or without permanent sexual partner (Benzaken 2008, Nuñez-Forero 2015, Mishra 2010, Nessa 2008)

Index test

All the studies used a treponemal rapid test as an index test and it was assessed as a screening test. Six studies used whole blood as the source for the index test (Benzaken 2008, Black 2016, Campos 2006, Juárez-Figueroa 2007, Nessa 2008), while two studies used serum for this purpose (Núñez- Forero 2015, West 2002); seven of the eight rapid tests used immunochromatography as technology (Benzaken 2008, Black 2016, Campos 2006, Juárez-Figueroa 2007, Nessa 2008, Núñez-Forero 2015, West 2002), while the remaining study used hemmaglutination for the rapid test (Mishra 2010).

Eight studies reported that the index test was conducted and interpreted by a laboratory technician (Campos 2006, Benzaken 2008, Juarez-Figueroa 2007, Nuñez-Forero 2015, West 2002), nurses (Black 2016) and paramedics (Nessa 2008) with low training level. In one study (Mishra 2010) the sample was taken and analyzed by nurses, physicians and social workers.

Three studies used more than one rapid test (Juarez-Figueroa 2007, Nessa 2008, Nuñez-Forero 2015). One study used a dual test that gave information about the syphilis infection and the HIV serological status (Black 2016). The rapid tests used in the eight studies were: Visitect Syphilis test (Benzaken 2008), SD Bioline HIV/Syphilis test (Black 2016), SD Bioline Syphilis 3.0 RT and ACON Syphilis test (Nuñez-Forero 2015) Determine Syphilis test (Campos 2006, Juarez-Figueroa 2007), Serodia Syphilis test (Juarez-Figueroa 2007), Qualpro Syphicheck WB (Mishra 2010), Immunochromatographic strip (Nessa 2008), Rapid test device (Nessa 2008), ACON Syphilis test (Nuñez-Forero 2015), Rapid Syphilis Test (West 2002). In all the included studies it was considered a positive index test when the device showed two red lines, one in the control area and one in the patient area. A negative result was considered when only one red line was showed in the control area.

Gold Standard

Five studies used TPHA (Campos 2006, Mishra 2010, Nessa 2008, Nuñez-Forero 2015, West 2002), two studies used FTA-ABS (Benzaken 2008, Juarez-Figueroa 2007) and one study used TPPA (Black 2016) as the treponemal test in the Gold Standard.

As non-treponemal test, six studies used RPR (Black 2016, Campos 2006, Mishra 2010, Nessa 2008, Nuñez-Forero 2015, West 2002) and two studies used VDRL (Benzaken 2008, Juarez- Figueroa 2007).

All the included studies reported that the Gold Standard was conducted and interpreted by a trained technician who worked in a certified laboratory (Benzaken 2008, Black 2016, Campos 2006, Mishra 2010, Nessa 2008, Núñez-Forero 2015, West 2002, Juarez-Figueroa 2007).

All the included studies considered that the patients had active syphilis if they had a positive treponemal test with a reactive no treponemal test, irrespective of the level of titers. Besides, four studies reported the performance of the rapid test when it was compared with a positive treponemal test and a reactive non treponemal test with titers above or equal to 1:8 (Black 2016, Campos 2006, Mishra 2010, Nessa 2008).

Outcomes

All the included studies reported the performance of a diagnostic rapid test. One study (Campos 2006) reported the proportion of patients that returned to receive full treatment of the disease. Another study (Mishra 2010) reported the percentage of patients that came back for the follow up 30 days after the diagnosis and who received treatment. One study (Benzaken 2008) reported the costs related to the use of rapid point of care tests and the costs related to the administration of the antibiotic in the same site of the diagnosis.

Of the included studies, three were funded by laboratories as the rapid tests were provided by them (Black 2016, Juárez-Figueroa 2007, West 2002), and two were promoted by academic institutions (Mishra 2010, Campos 2006, Nuñez-Forero 2015).

One of the study authors is included also as an author in this review (Edith Angel-Müller) (Núñez-Forero 2015)

Basic features of the excluded studies

On the basis of full text assessment, we excluded 30 publications, none of which were retrospective design, as it follows: one was a congress meeting that compares the POC syphilis test with the treponemal and non-treponemal test separately (Barnes 2011); another was a study that compared 7 laboratory tests, which does not meet the criteria in our review (Binniker 2011); one was a non prospective study that used stored serum samples with already known syphilis diagnosis (Bristow 2013); there was a letter to the editor that also was excluded (Calderón-Anyosa 2012); five studies did not made a direct comparison between the rapid test and the reference test (Castro 2010, Castro 2010, Castro 2014, Causer 2015, Wang 2007); another two studies didn´t use rapid syphilis tests (Choi 2013, van Dyck 2002); two studies were excluded because they were published congress posters (Dzokto 2013, Núñez-Forero 2013); 6 studies were excluded given the fact that they didn´t used an adequate gold standard according with the inclusion criteria in this review (Mabey 2006, Maln 2015, Tagni 2015, Villazón-Vargas 2009, Zheng 2008, Yin 2012); one study analyses the cost-effectiveness of the rapid syphilis test but does not compare them with the gold standard cost- effectiveness (Rydzak 2008).

Methodological quality of the included studies

We assessed the quality of included articles by using a modified version of the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (Whiting 2011). We performed the quality assessment using the four key domains to assess risk of bias and concerns regarding applicability to the research question (patient selection, index test, reference standard, and flow-timing domains). We scored the papers as having a 'low', 'high' or 'unclear' risk of bias for each of four domains, and for the patient selection, index test and reference standard applicability. Studies classified at high or unclear risk of bias and/or high concern regarding applicability in at least one domain were regarded as having low methodological design.

Domain patient sampling

Seven studies were assessed as high risk of bias for this domain (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015). With the exception of one study (Juarez Figueroa 2007) which did not report the method of sampling, these studies reported a consecutive sampling but were conducted among high risk population (female sex workers or patients presenting lower genital tract symptoms). The remaining study was assessed as low risk bias for this domain (West 2002).

Domain index test

All the included studies were assessed as low risk of bias for this domain (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015, West 2002), describing appropriately the index test and how it was conducted and interpreted, making unlikely the risk of bias.

Domain target condition and reference standard

All the included studies were assessed as low risk of bias for this domain (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015, West 2002), describing appropriately how the reference standard was conducted by trained technicians, nurses or paramedics, and how the results were interpreted in a blind manner regardless of the results of the index test. In one study (Mishra 2010) 2% of the patients agreed to take the rapid test by finger prick but did not agreed to have the venous sampling for the reference standard test, and these patients were not analyzed.

Domain flow and timing

All of the included studies were assessed as low risk of bias for this domain (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015, West 2002), making clear what happened to all patients who entered the study. Nevertheless, two of these studies (Benzaken 2008, Campos 2006) did not explained about withdrawals or exclusions and were unclear if the number of patients recruited match those in the analysis. In two studies (Campos
2006, Mishra 2010, Nuñez-Forero 2015) there was a loss of 2-4% of the participants that was not explained, although we do not consider this as a significant loss.

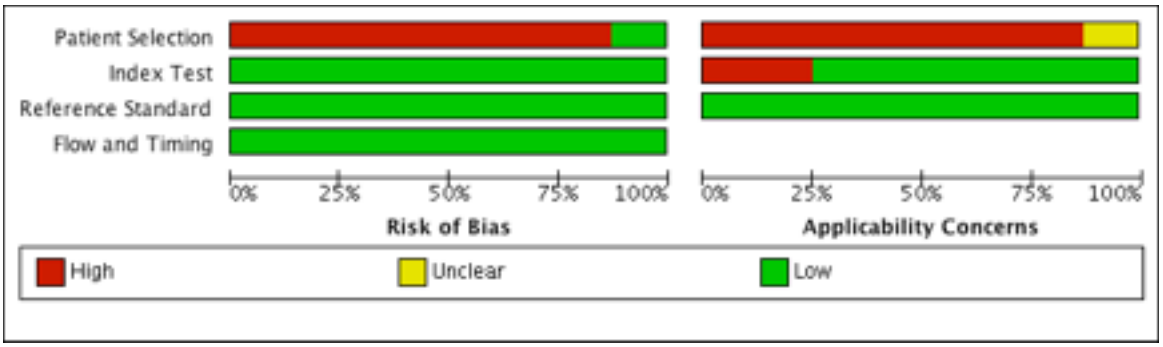
Other bias

Six of the included studies were assessed as low risk for this domain (Benzaken 2008, Black 2016, Campos 2008, Mishra 2010, Nessa 2008, Nuñez Forero 2015), reporting that all the study was conducted without the intervention of the pharmaceutical industry. Two studies (Juarez-Figueroa 2007, West 2002) were assessed as high risk because the index test was donated by the laboratory that owns the test.

Applicability

In seven of the included studies (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015) there was high concern because the included patients did not match the review question in terms of demographic features and presence of differential diagnosis or comorbidity, given the fact that they recruited female sex workers exclusively or patients with low genital tract symptoms. One study was assessed as unclear risk in the patient selection (West 2002), because there was not a clear description of the baseline characteristics of the population, although we considered that the women recruited had a standard risk.

In two studies (Núñez-Forero 2015, West 2002) there was high concern because the index test, its conduct, or its interpretation differs from the review question. They applied the rapid test in serum and not in whole blood. This statement differs from our review question, that looks for the implementation of the rapid tests in low resource settings.
Regarding reference standard and flow and timing applicability, the eight studies were ranked as low concern (Benzaken 2008, Black 2016, Campos 2008, Juarez-Figueroa 2007, Mishra 2010, Nessa 2008, Nuñez Forero 2015, West 2002) because the target condition, as defined by the reference standard, matched our review question.



Findings

For each comparison of the index test with the reference test, we extracted data on the number of true positives, false positives, false negatives and true negatives, in a two by two table. When there were two or more rapid tests evaluated, we included separate data sets, since the tests result was the unit of analysis.

1.0 Primary objective

1.1 POC accuracy for detecting syphilis infection defined as both treponemal and non- treponemal test positives

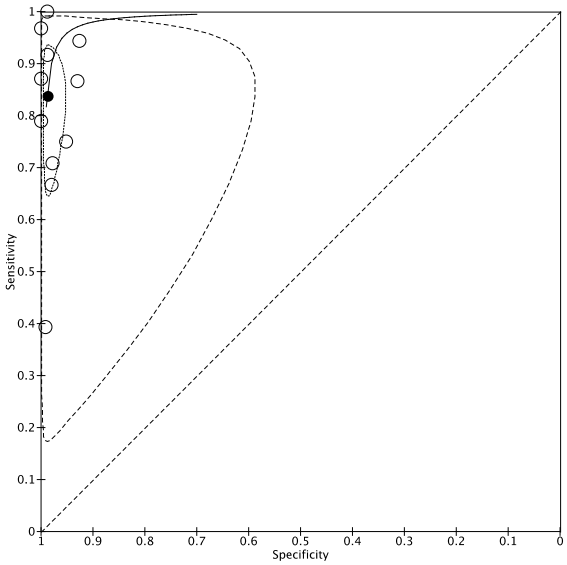
Eight articles, which included a total of 11783 participants (based on data from eleven cohorts described in eight publications), conducted between 1.999 and 2.015, assessed the diagnostic accuracy of POC tests for diagnosis of syphilis infection in men and non-pregnant women, as verified with the combination of both reactive treponemal and non-treponemal test as the reference standard. Of these studies, 1 were conducted in North America (Juarez-Figueroa 2007), 2 in Asia (Mishra 2010, Nessa 2008), 2 in Africa (Black 2016, West 2002) and 3 in South America (Nuñez- Forero 2015, Campos 2006, Benzaken 2008). The median sample size was 1071 (range 198 to 3483), and the reported prevalence of syphilis varied, ranging from 0.9% to 20,4%, while the active syphilis prevalence varied from 2,8% to 15,7%. Nine different POC brands were evaluated: Determine (n=2), Rapid Syphilis Test (n=2), ACON Syphilis Test (n=1), Immunochromatographic strip test (n=1), Qualpro Syphicheck-WB (n=1), SD Bioline 3.0 (n=1), SD Bioline HIV/Syphilis Duo test (n=1), Serodia (n=1) and Visitect Syphilis Test (n=1).

Non of the included studies reported the stage of the syphilis or HIV infection, and the seven studies reported treatment of the patients with the diagnosis of active syphilis. Only one study reported a follow up of 30 days after the diagnosis of the infection (Mishra 2010); 5 studies reported no time interval between the application of the index test and the gold standard (Black 2016, Campos 2006, Mishra 2010, Nessa 2008, Núñez Forero 2015), one study reported an interval of one week (west 2002), while 2 studies did not reported this data (Benzaken 2008, Juárez-Figueroa 2007). Only one study reported the economical analysis (Benzaken 2008), which stated that the cost per case of syphilis was $16.8 for VDRL, $33.2 for low cost and $56.3 for high cost VisiTect Syphilis; the cost per

case of active syphilis was $21.3, $57.5 and $97.6, respectively. In one study (Mishra 2010) 115 female sex workers who tested positive by the POC test received same day treatment, corresponding to $1 dose treatment in 68% (78 out of 115 patients) of women with active syphilis. Also POC guided therapy led to treatment of 33 women who were RPR negative, which corresponds to innapropiate treatment for non infected patients. In another study (Campos 2006), 87% of women who tested positive for the POC test visited the local health centre for treatment, and 64% completed the three dose scheme.

Sensitivities of the tests ranged from 0.39 to 1.00, specificities from 0.93 to 1.00. The mean sensitivity and specificity of all included studies were 0.83 (95% confidence interval (CI) 0.71 to 0.91) and 0.98 (95% CI 0.96 to 0.99). Forest plots and the ROC plot demonstrated a high degree of heterogeneity between papers, which was greater for estimates of sensitivity than for specificity. In conclusion the number of included studies and the large amount of recruited participants, the visual inspection of forest plots of sensitivities, the wide range for the confidence intervals and for the prediction region makes that the estimates did not meet the criteria for triage test.

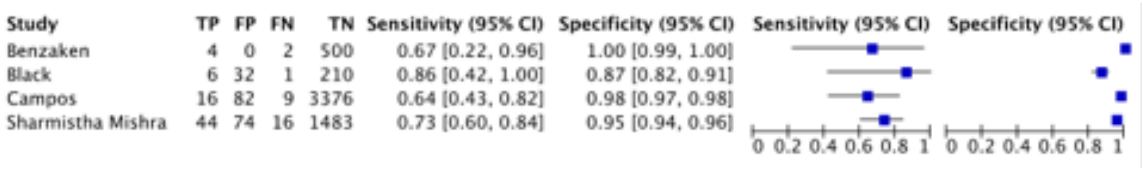
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Benzaken	30	0	8	468	0.79 [0.63, 0.90]	1.00 [0.99, 1.00]
Black	34	4	17	194	0.67 [0.52, 0.79]	0.98 [0.95, 0.99]
Campos	70	28	108	3277	0.39 [0.32, 0.47]	0.99 [0.99, 0.99]
Juarez-Figueroa1	30	0	1	167	0.97 [0.83, 1.00]	1.00 [0.98, 1.00]
Juarez-Figueroa2	27	0	4	167	0.87 [0.70, 0.96]	1.00 [0.98, 1.00]
Khairun Nessa1	134	40	8	502	0.94 [0.89, 0.98]	0.93 [0.90, 0.95]
Khairun Nessa2	123	38	19	504	0.87 [0.80, 0.92]	0.93 [0.91, 0.95]
Nunez-Forero1	11	17	1	1371	0.92 [0.62, 1.00]	0.99 [0.98, 0.99]
Nunez-Forero2	12	17	0	1410	1.00 [0.74, 1.00]	0.99 [0.98, 0.99]
Sharmistha Mishra	85	33	35	1464	0.71 [0.62, 0.79]	0.98 [0.97, 0.98]
West	30	62	10	1223	0.75 [0.59, 0.87]	0.95 [0.94, 0.96]

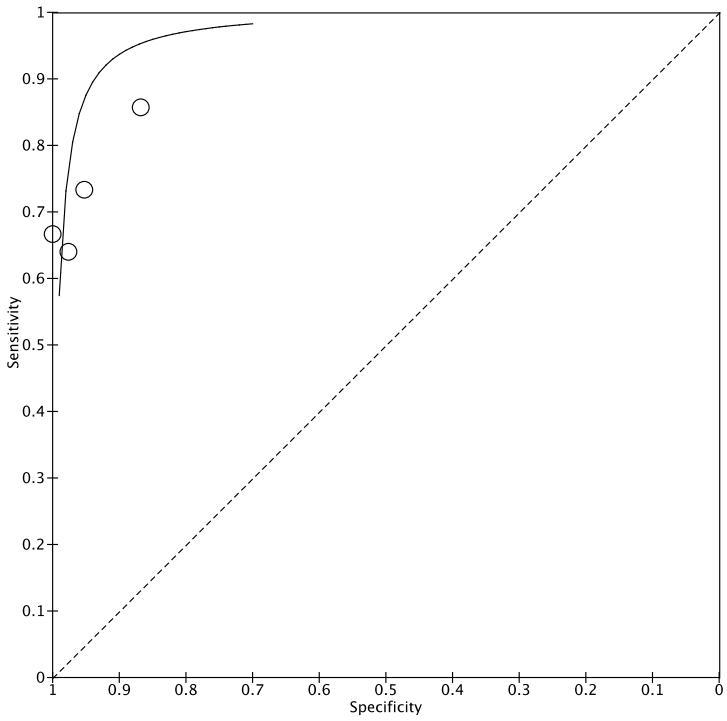


1.2 Secondary objectives

1.2.1 POC accuracy for detecting active syphilis infection defined as a positive treponemal and a non-treponemal test with title ≥1:8.

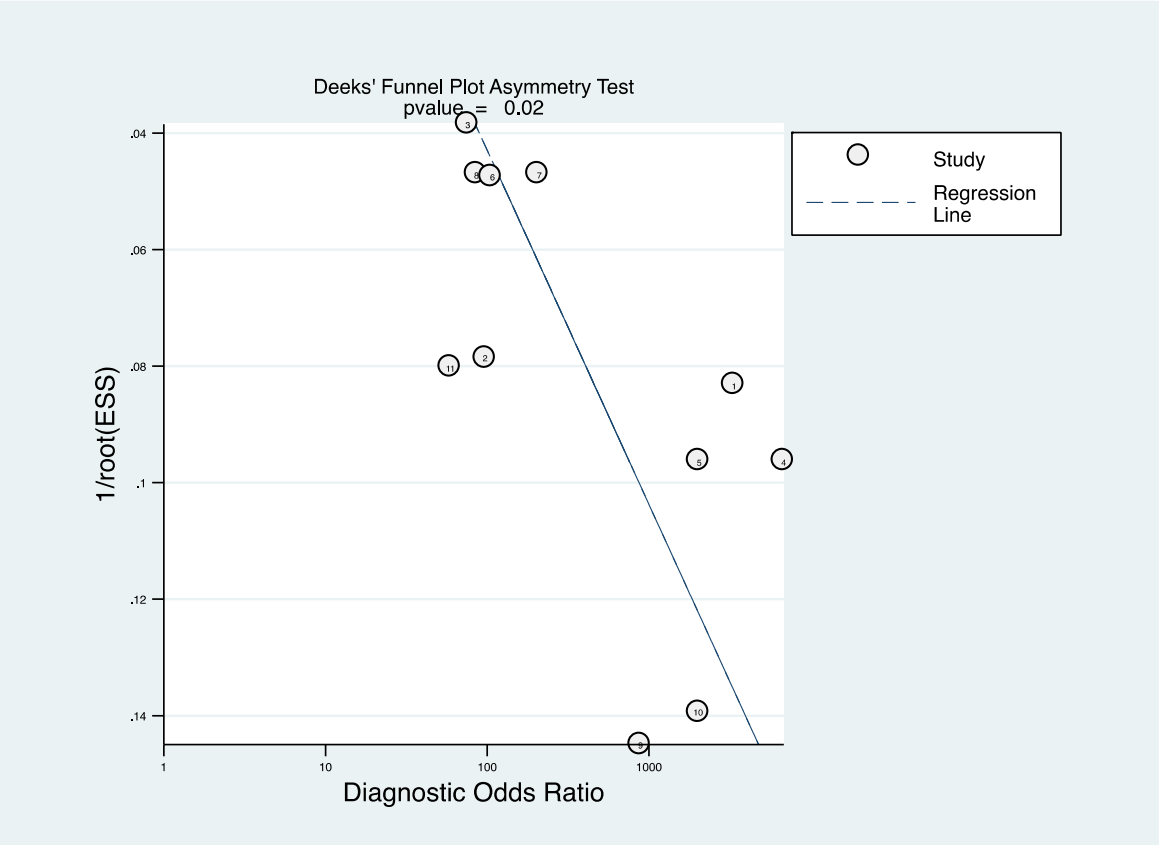
Four studies which included a total of 5855 participants, assessed the diagnostic accuracy of POC tests for diagnosis of active syphilis infection in men and non-pregnant women, as verified with the combination of both reactive treponemal and non-treponemal test with title ≥1:8 as the reference standard (Benzaken 2008, Black 2016, Campos 2006, Mishra 2010). The median sample size was 1463 (range 249 to 3483), and the median prevalence was 10.1% (range 5.1% to 20.5%). Sensitivities of the tests ranged from 0.39 to 0.79, specificities from 0.98 to 1.00. Due to the limited number of studies available, we were unable to estimate the mean sensitivity and specificity. Forest plots and the ROC plot demonstrated some degree of heterogeneity between papers for estimates of sensitivity.





Publication bias

We investigated publication bias because more than ten tests were included in this systematic review. In first instance, we assessed reporting bias through funnel plot visual asymmetry, plotting a measure of effect size against a measure of study precision. This inspection suggested a discreet asymmetrical studies distribution around regression line. For this reason, we proceeded to realize a formal evaluation using Deeks’ test to investigate the asymmetry. The statistically significant *p*- value (0.02) for the slope coefficient suggests asymmetry in the data and a likelihood of publication bias (Deeks 2005).



DISCUSSION

Summary of main results

Eight studies (11.783 participants) met our inclusion criteria for this review, and three of the included studies were funded by the pharmaceutical industry. Eleven rapid tests were assessed in the included studies, six of the studies were multicentric. Four studies were

conducted in America, two in Africa and two in Asia. Most of the population recruited in the studies included was high risk, due to their sexual working status, sexual worker clients or patients with low genital tract symptoms. The median sample size was 1071 (range 198 to 3483), and the reported prevalence of syphilis varied, ranging from 0.9% to 20,4%, while the active syphilis prevalence varied from 2,8% to 15,7%. Sensitivities of the tests ranged from 0.39 to 1.00, specificities from 0.93 to 1.00. The mean sensitivity and specificity of all included studies were 0.83 (95% confidence interval (CI) 0.71 to 0.91) and 0.98 (95% CI 0.96 to 0.99). Forest plots and the ROC plot demonstrated a high degree of heterogeneity between papers, which was greater for estimates of sensitivity than for specificity.

Overall completeness and applicability of evidence

We conducted a comprehensive search to retrieve all published and unpublished RCTs. We were able to evaluate the primary outcomes and some of the secondary outcomes. The applicability of the evidence into the common clinical practice (men and nonpregnant women) is low, due to the high concern that arose from the fact that the rapid tests were applied in high risk population, and even still the sensitivity was no the expected for a screening tests. In this sense, the result of this systematic review shows that, with low quality of evidence, the rapid tests for diagnosis of syphilis might be useful in high risk population and in those with high probability of following loss. This results might not be suitable for the general population in which the disease prevalence is standard or low.

Quality of evidence

We considered seven of the eight included trials as high risk of bias, due to the population recruited for each study, given their job status (e.g. female sex workers) or their clinical presentation (e.g. women with low genital tract symptoms). This situations may lead to a false perception of high performance in the rapid tests, given the fact that as high as the prevalence could present, the performance of the rapid tests would present high as well. Only one study was considered as low risk of bias because the population included were women without any symptoms or any high risk behavior, although this study did not mention the method of recruitment. We considered also a high risk of publication bias, assessed by visual symmetry in the funnel plot.

Potential biases in the review process

Our systematic review has many strengths regarding the review process: we built a clear review question, predefined objectives, inclusion and exclusion criteria; we explained clearly how the studies selection was conducted, and how the search strategy was applied; we performed the extraction of the data by duplicate; we provided a list of the excluded studies and explained why they were excluded; we used a adequate tool for the assessment of biases; we reported the funding in the studies; we used an appropriate method for the statistical analysis; we explained the heterogeneity found and also reported any kind of conflict of interests.

Agreements and disagreements with other studies

There are at least three previous systematic reviews for this topic (Jafari 2013, Gliddon 2017, Swartzendruber 2015) that found a high sensitivity and specificity. Our review assesses the diagnostic accuracy of point of care rapid tests for the diagnosis of syphilis in men and non pregnant women in low resource settings. While our review resulted in a high heterogeneity in sensitivity among the incldued studies, the previous systematic review resulted in a high sensitivity and specificity regardless of the study included. Meanwhile, our review includes patients who were recruited prospectively, most of the systematic reviews before ours included studies that used stored samples with a known previous serologic status. This finding resulted in a high performance of the rapid test and a moderate performance in our review. Our conclusion, about the possibility of using the rapid test as a screening test for the high risk population resides far from the conclusions in the other systematic reviews, which stated that theses rapid tests could be use as a reliable diagnostic test.

Authors` conclusions

Implications for practice

Given the low quality evidence for the main outcome in this Cochrane systematic review, we consider that the point of care rapid tests for the diagnosis of syphilis might be an accurate method for the detection of the infection in high risk population and those patients with high risk of following loss after diagnosis. This conclusion can not be applied in general population with standard risk or with standard prevalence. Also, the performance of these test resides far from the results obtained in other systematic reviews.

Implications for research

Due to the low quality found in the included studies for this review, there are needed more studies conducted among general population with standard risk for the infection, under field, low resource, laboratory-free settings to assess the diagnostic accuracy of this rapid tests.

References to studies included in this review

Benzaken 2008
A S Benzaken, M Sabido, E G Galban, V Pedroza, F Vasquez, A Araujo, R W Peeling, P Mayaud. Field evaluation of the performance and testing costs of a rapid point-of-care test for syphilis in a red-light district of Manaus, Brazil. Sex Transm Infect 2008;84:297–302.

Black 2016
Vivian Black, Brian G Williams, Venessa Maseko, Frans Radebe, Helen V Rees, David A Lewis. Field evaluation of Standard Diagnostics' Bioline HIV/Syphilis Duo test among female sex workers in Johannesburg, South Africa. Sex Transm Infect 2016;92:495–498.

Campos 2006
P E Campos, A L Buffardi, M Chiappe, C Buendıa, P J Garcia, C P Carcamo, G Garnett, P White, K K Holmes. Utility of the Determine Syphilis TP rapid test in commercial sex venues in Peru. Sex Transm Infect 2006;82(Suppl V):v22–v25.

Juarez-Figueroa 2007
Luis Juárez-Figueroa, Felipe Uribe-Salas, Santa García-Cisneros, María Olamendi-Portugal, Carlos J. Conde-Glez. Evaluation of a rapid strip and a particle agglutination tests for syphilis diagnosis. Diagnostic Microbiology and Infectious Disease 59 (2007) 123–126

Mishra 2010
Sharmistha Mishra, Balaji Naik, B Venugopal, Prakash Kudur, Reynold Washington, Marissa Becker, John Kenneth, Krishnamurthy Jayanna, Banadakoppa M Ramesh, Shajy Isac, Marie-Claude Boily, James F Blanchard, Stephen Moses. Syphilis screening among

female sex workers in Bangalore, India: comparison of point-of-care testing and traditional serological approaches. Sex Transm Infect 2010; 86: 193e198.

Nessa 2008

Khairun Nessa, Anadil Alam, Faisal Arif Hasan Chawdhury, Mohsina Huq, Shamsun Nahar, Gazi Salauddin, Shayema Khursheed, Saifur Rahman, Emily Gurley, Robert F Breiman, and Motiur Rahman. Field evaluation of simple rapid tests in the diagnosis of syphilis. International Journal of STD & AIDS 2008; 19: 316–320.

Nuñez-Forero 2015

Lilian Nuñez-Forero, Luisa Moyano-Ariza, Hernando Gaitán-Duarte, Edith Ángel-Müller, Ariel Ruiz-Parra, Patricia González, Andrea Rodríguez, Jorge E Tolosa. Diagnostic accuracy of rapid tests for sexually transmitted infections in symptomatic women. Sex Transm Infect 2015;0:1–5.

West 2002

B West, G Walraven, L Morison, J Brouwers, R Bailey. Performance of the rapid plasma reagin and the rapid syphilis screening tests in the diagnosis of syphilis in field conditions in rural Africa. Sex Transm Infect 2002;78:282–285

REFERENCES

Benzaken 2014

Benzaken AS, Bazzo ML, Galban E, Pinto IC, Nogueira CL, Golfetto L. External quality assurance with dried tube specimens (DTS) for point-of-care syphilis and HIV tests: experience in an indigenous populations screening programme in the Brazilian Amazon. *Sexually Transmitted Infections* 2014;**90**(1):14–8.

Bien 2015

Bien CH, Muessig KE, Lee R, Lo EJ, Yang LG, Yang B, et al. HIV and syphilis testing preferences among men who have sex with men in South China: a qualitative analysis to inform sexual health services. *PLOS One* 2015;**10**(4): e0124161.

Bossuyt 2013

Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten RJ. Chapter 11: Interpreting results and drawing conclusions. In: edition. http://srdta.cochrane.org/. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors(s) editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 0.9*. The Cochrane collaboration, 2013: 1–31.

Bristow 2014

Bristow CC, Adu-Sarkodie Y, Ondondo RO, Bukusi EA, Dagnra CA, Oo KY, et al. Multisite laboratory evaluation of a Dual Human Inmunodeficiency Virus (HIV)/Syphilis point-of-care rapid test for simultaneous detection of HIV and syphilis infection. *Open Forum Infectious Diseases* 2014; **1**(1):OFU015.

CDC 2006

CDC. Department of Health and Human Services. Centers for Disease Control and Prevention Sexually transmitted diseases treatment guidelines. *MMW Morbidity and Mortality Weekly Report* 2006;**55**(RR-11):1-94.

CDC 2008

Centers for Disease Control and Prevention (CDC). Syphilis testing algorithms using treponemal tests for initial screening: four laboratories. New York City, 2005-2006. *MMWR Morbidity and Mortality Weekly Report* 2008;**57** (32):872-5. [PUBMED: 18701877]

CDC 2014

CDC. Centers for Disease Control and Prevention. Sexually transmitted disease surveillance 2012. Atlanta, GA: US Department of Health and Human Services. Available at http://www.cdc.gov/std/ stats12/default.htm 2014:accessed May 15 2015.

CDC 2014 a)

CDC. STD Surveillance Case Definitions-2104. available at: http://www.cdc.gov/std/stats/ casedefinitions-2014.pdf. 2014; Vol. accessed May 2015.

Chen 2012

Chen XS, Yin YP, Shen C, Liu GG, Zhu ZJ, Wei WH, et al. Rapid syphilis testing uptake for female sex workers at sex venues in Southern China: implications for expanding syphilis screening. *PLOS One* 2012;**7**(12):e52579.

Creegan 2007

Creegan L, Bauer HM, Samuel MC, Klausner J, Liska S, Bolan G. An evaluation of the relative sensitivities of the venereal disease research laboratory test and the Treponema pallidum particle agglutination test among patients diagnosed with primary syphilis. *Sexually Transmitted Diseases* 2017;**34**(12):1016–8.

Gaydos 2014

Gaydos C, Hardick J. Point of care diagnostics for sexually transmitted infections: perspectives and advances. *Expert Review of Anti-Infective Therapy* 2014;**12**(6):657–72.

Greer 2008

Greer L, Wendel GD. Rapid diagnostic methods in sexually transmitted infections. *Infectious Disease Clinics of North America* 2008;**22**(4):601–17.

Gupte 2011

Gupte S, Daly C, Agarwal V, Gaikwad SB, George B. Introduction of rapid tests for large-scale syphilis screening among female,male, and transgender sex workers in Mumbai, India. *Sexually Transmitted Diseases* 2011;**38**(6): 499–502.

Higgins 2011

Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Versions 5.1.0 (updated March 2011).The Cochrane Collaboration, 2011. Available from handbook.cochrane.org 2011.

Hsu 2011

Hsu J, Brozek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implementation Science* 2011;**6**(62):1–9.

Jafari 2013

Jafari Y, Peeling RW, Shivkumar S, Claessens C, Joseph L, Pai NP, et al. Are Treponema pallidum specific rapid and point-of-care tests for syphilis accurate enough for screening in resource limited settings? Evidence from a meta- analysis. *PLOS One* 2013;**8**(2):e54695. [DOI: 10.1371/ journal.pone.0054695

Janier 2014

Janier M, Hegyi V, Dupin M, Unemo M, Tiplica GS, Potocnik M, et al. European guideline on the management of syphilis. *Journal of the European Academy of Dermatology and Venereology* 2014;**28**(12):1581–93.

Kay 2014

Kay N, Peeling RW, Mabey DC. State of the of the art syphilis diagnostics: rapid point-of-care tests. *Expert Review of Anti- Infective Therapy* 2014;**12**(1):63–73.

Larson 2014

Larson BA, Lembela-Bwalya D, Bonawitz R, Hammond EE, Thea DM, Herlihy J. Finding a needle in the haystack: the costs and cost-effectiveness of syphilis diagnosis and treatment during pregnancy to prevent congenital syphilis in Kalomo District of Zambia. *PLOS One* 2014;**9**(12): e113868.

Liu 2014

Liu LL, Lin LR, Tong ML, Zhang HL, Huang SJ, Chen YY, et al. Incidence and risk factors for the prozone phenomenon in serologic testing for syphilis in a large cohort. *Clinical Infectious Diseases* 2014;**59**(3):384–9.

Mabey 2012

Mabey DC, Sollis KA, Kelly HA, Benzaken AS, Bitarakwate E, Changalucha J, et al. Point-of-care tests to strengthen health systems and save newborn lives: the case of syphilis. *PLOS Medicine* 2012;**9**(6):e1001233.

Naidu 2012

Naidu NK, Bharucha ZS, Sonawane V, Ahmed I. Comparative study of Treponemal and non- Treponemal test for screening of blood donated at a blood center. *Asian Journal of Transfusion Science* 2012;**6**(1):32–5.

PAHO 2015

Pan American Health Organization. Guidance on Syphilis Testing in Latin America and the Caribbean: Improving Uptake, Interpretation, and Quality of Testing in Different Clinical Settings. Washington, DC. Pan American Health Organization 2015.

Peeling 2004

Peeling RW, Ye H. Diagnostic tools for preventing and managing maternal and congenital syphilis: an overview. *Bulletin of the World Health Organization* 2004;**82**:439-46.

Peeling 2006

Peeling RW, Holmes KK, Mabey D, Ronald A. Rapid tests for sexually transmitted infections (STIs): the way forward. *Sexually Transmitted Infections* 2006;**82**(suppl 5):v1–v6.

Peeling 2010

Peeling RW, Mabey D. Point-of-care tests for diagnosing infections in the developing world. *Clinical Microbiology and Infectious Diseases* 2010;**16**(8):1062–9.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90.

Revman 2014 [Computer program]

Nordic Cochrane Centre. The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre. The Cochrane Collaboration, 2014.

Saloojee 2004

Saloojee H, Velaphi S, Goga Y, Afadapa N, Steen R, Lincetto O. The prevention and management of congenital syphilis: an overview and recommendations. *Bulletin of the World Health Organization* 2004;**82**:424–30.

Singh 2015

Singh AE, Chernesky MA, Morshed M, Wong T. Canadian Public Health Laboratory Network laboratory guidelines for the use of point-of-care tests for the diagnosis of syphilis in Canada. *Canadian Journal of Infectious Diseases & Medical Microbiology* Jan–Feb 2015;**26**(Suppl A):29A– 32A.

van Enst 2014

van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang M. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study.. *BMC Medical Research Methodology* 2014;**14**(70):1–11.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**:529–36.

WHO 2003

WHO. Guidelines for the management of sexually transmitted infections. Geneva. *Available at http:// apps.who.int/iris/bitstream/10665/42782/1/9241546263_ eng.pdf?ua=1* 2003;**1**:pp 98.

APPENDICES

Appendix 1. Search Strategies

MEDLINE and CENTRAL search strategy (OVID platform)

1 exp Point-of-Care Testing/ 2 'point of care'.tw.
3 (bedside adj5 test$).tw.
4 poc.tw.
5 poct.tw.
6 (rapid adj5 test$).tw.
7 (self adj5 test$).tw.
8 (self adj5 collect$).tw.
9 (self adj5 swab$).tw.
10 (syphilis adj5 rapid adj5 test).tw.
11 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10
12 exp Treponema pallidum/
13 treponema$.tw.
14 (spirochaeta adj5 pallid$).tw.
15 pallidum.tw.
16 exp Syphilis/
17 syphili$.tw.
18 lues.tw.
19 chancre.tw.
20 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19
21 11 and 20

Embase search strategy (Embase.com platform)

#1. ’syphilis rapid test’/exp
#2. ’point of care testing’/exp
#3. ’point of care’:ab,ti
#4. (bedside NEAR/5 test*):ab,ti
#5. poc:ab,ti
#6. poct:ab,ti
#7. (rapid NEAR/5 test*):ab,ti
#8. (self NEAR/5 test*):ab,ti
#9. (self NEAR/5 collect*):ab,ti
#10. (self NEAR/5 swab*):ab,ti
#11. ’rapid test’/exp
#12. (syphilis NEAR/5 rapid NEAR/5 test):ab,ti
#13. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12
#14. ’treponema pallidum’/exp
#15. ’treponematosis’/exp
#16. treponema*:ab,ti
#17. (spirochaeta NEAR/5 pallid*):ab,ti
#18. pallidum:ab,ti
#19. ’syphilis’/exp
#20. syphili*:ab,ti
#21. lues:ab,ti
#22. chancre:ab,ti
#23. #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
#24. #13 AND #23
#25. #24 AND [embase]/lim

LILACS Biblioteca Virtual en Salud (BVS), interfaz iAHx

(mh:(Treponema pallidum)) OR (mh:(syphilis)) OR (ab:(syphilis)) OR (ti:(syphilis)) AND (ab: (point of care)) OR (ti:(point of care)) OR (ab:(test*)) OR (ti:(test*)) AND (instance:“regional”) AND (db:(“LILACS”))

DARE Centre for Reviews and Dissemination - University of York Platform

((test OR point of care)) AND ((syphilis OR Treponema pallidum))

Web of Science

(TS=(“point of care”) OR TS=(test*)) AND TS=(syphilis) AND TI=(trial)

WHO International Clinical Trials Registry Platform

point of care AND syphilis

Health Services Research Projects in Progress (HSRProj)

(“point of care”) AND (syphilis)

Sexually Transmitted Infection (STI) Cochrane Review Grouṕ s Specialized Register (CRS)

1 (point of care:AB) AND (INREGISTER) # 2 (point of care:TI) AND (INREGISTER) # 3 (syphilis:AB) AND (INREGISTER)

4 (syphilis:TI) AND (INREGISTER)

5 ((1 or 2) and (3 or 4))

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT Internal sources

• No sources of support supplied **External sources** • Universidad Nacional de Colombia, Colombia.