International Journal of Infectious Diseases 118 (2022) 183–193

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Review

The diagnostic accuracy of pooled testing from multiple individuals for the detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae*: a systematic review☆

Yangqi Xu[1,*,**], Lily Aboud[1,2,*], Eric P.F. Chow[1,3,6], Maeve B. Mello[4], Teodora Wi[4], Rachel Baggaley[4], Christopher K. Fairley[1,3], Rosanna Peeling[5], Jason J. Ong[1,3,5]

[1] *Central Clinical School, Monash University, Melbourne, Australia*
[2] *College of Medicine and Dentistry, James Cook University, Townsville, Australia*
[3] *Melbourne Sexual Health Centre, Alfred Health, Melbourne, Australia*
[4] *Global HIV, Hepatitis and STI Programmes, World Health Organization, Geneva, Switzerland*
[5] *Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom*
[6] *Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia*

ARTICLE INFO

Article history:
Received 9 December 2021
Revised 25 February 2022
Accepted 7 March 2022

ABSTRACT

Objectives: Molecular testing for *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (NG) is costly. Therefore, we appraised the evidence regarding pooling samples from multiple individuals to test for CT/NG.

Methods: In this systematic review, we searched 5 databases (2000-2021). Studies were included if they contained primary data describing pooled testing. We calculated the pooled sensitivities and specificities for CT and NG using a bivariate mixed-effects logistic regression model.

Results: We included 22 studies: most were conducted in high-income countries (81.8%, 18 of 22), among women (73.3%, 17 of 22), and pooled urine samples (63.6%, 14 of 22). Eighteen studies provided 25 estimates for the meta-analysis of diagnostic accuracy, with data from 6,913 pooled specimens. The pooled sensitivity for CT was 98.4% (95% confidence intervals [CI]: 96.8-99.2%, I^2=77.5, p<0.001), and pooled specificity was 99.9% (95% CI: 99.6-100.0%, I^2=62.6, p<0.001). Only 2 studies reported pooled testing for NG, and both reported similarly high sensitivity and specificity as for CT. Sixteen studies provided data on the cost of pooling, reporting cost-savings ranging from 39%-90%.

Conclusions: Pooled testing from multiple individuals for CT is highly sensitive and specific compared with individual testing. This approach has the potential to reduce the cost of screening in populations for which single anatomic site screening is recommended.

INTRODUCTION

Chlamydia trachomatis (CT) and *Neisseria gonorrhoeae* (NG) are common bacterial sexually transmitted pathogens with a significantly associated disease burden. In 2020, approximately 128 million chlamydia and 82 million gonorrhoea cases were newly acquired (World Health Organization 2021). Untreated sexually transmitted infections (STIs) can lead to reproductive morbidity and infertility in women, vertical transmission to neonates, and increased risk of acquiring HIV. The rise of antimicrobial resistance in STIs, particularly for gonorrhoea, underscores the necessity for aetiological diagnosis to optimise effective and early STI management.

As CT/NG infections are often asymptomatic, early detection relies on regular screening for those at risk. However, current molecular-based diagnostics are relatively expensive and remain inaccessible for many resource-limited settings (World Health Organization. Laboratory diagnosis of STIs, including human immunodeficiency virus). One strategy to improve access to molecular testing for CT/NG includes pooling specimens to reduce costs

☆ The study protocol is registered in PROSPERO, an international database of prospectively registered systematic reviews (CRD42021240793).

** Corresponding author: Jason J. Ong, 580 Swanston Street, Carlton, Victoria 3053, Australia, Phone: +613 9341 6200.
E-mail addresses: Jason.ong@monash.edu, Jason.Ong@lshtm.ac.uk (Y. Xu).
* Equal first-co-authors

https://doi.org/10.1016/j.ijid.2022.03.009

and improve efficiency. This strategy was previously implemented for pathogens such as *Mycobacterium tuberculosis* (Cuevas et al., 2021) and SARS-CoV-2 (Burdett et al., 2021). Indeed, pooled testing increases the number of people tested with the same budget. However, as each positive pool will require retesting to identify the positive sample(s), the cost-savings inherent to this strategy depend on the background prevalence of the pathogen and the number of samples pooled for testing. Several mathematical formulas such as those by Kacena et al. (1998) and Peeling et al. (1998b) have been constructed to estimate how likely a pool is to be positive given a selected population disease prevalence and pool size.

We have previously conducted a systematic review demonstrating the diagnostic accuracy of pooling urine, anorectal and oropharyngeal from a single individual to detect CT/NG among populations at higher risk of infection (Aboud et al.). It demonstrated that multisite pooled testing was a highly sensitive and specific method, with an associated cost-saving benefit and opportunity to increase screening coverage and detect more infections that otherwise would go unnoticed. However, there has not been a systematic review to evaluate pooled testing from multiple individuals. This systematic review and meta-analysis aim to critically appraise the existing evidence regarding the diagnostic accuracy and estimated cost-savings of single anatomic site pooled testing from multiple individuals for the screening of CT/NG.

METHODS

SEARCH STRATEGY AND SELECTION CRITERIA

Inclusion and Exclusion Criteria

Studies had to contain primary data that assessed at least 1 primary or secondary outcome: the diagnostic accuracy of the single anatomic site pooled testing (index test) from multiple individuals evaluated against a single sample testing (reference standard), resource use, clients or provider acceptability and impact on health equity. Deduplicated studies or studies with no relevance to the outcomes of interest or no primary data were excluded. Measures of diagnostic accuracy included sensitivity, specificity, or provision of true-positive, false-positive, true-negative and/or false-negative values.

Search Strategy

We searched for articles published between January 1, 2000, and February 4, 2021, limited to English in 5 databases: Medline, Embase, CINAHL, CABI Global Health, Web of Science. The search strategies looked for information on single anatomic site pooled testing for CT or NG from multiple individuals. Further details of the search strategy are provided in the Appendix.

Study Selection

Titles and abstracts were reviewed using Covidence by 2 researchers (LA, YX) independently, and any conflict was resolved by a third researcher (JO). The selection process is summarised in the PRISMA study flow diagram (Figure 1).

DATA ANALYSIS

Data Extraction

Two researchers (LA, YX) independently extracted data, with a third researcher (JO) resolving any conflicts. We used an electronic data extraction sheet in Excel to extract information from each study, including the author, publication year, study year, country, study type, study population, sample size, study settings, aims, method of pooling, pooling results (true positive, false positive, true negative, false negative) evaluated against the reference standard, resource use, acceptability, impact on health equity, benefits and harms and subsequent actions post results of pooled testing.

Risk of Bias Assessment

Included studies were evaluated using the QUADAS-2 checklist (Whiting et al., 2011) by 2 researchers (YX and LA). We assessed the certainty of the evidence using the GRADE (Schunemann et al., 2020).

Data Analysis

Descriptive statistics were used to summarise the characteristics of included studies. We used a generalised linear mixed model approach to bivariate meta-analysis of sensitivity and specificity (Chu and Cole 2006) in STATA version 17.0 (StataCorp. 2019. *Stata Statistical Software: Release 17*. College Station, Texas: StataCorp LLC). In the presence of zero events, we used a 0.5 continuity correction to enable parameter estimation. Statistical heterogeneity between studies was assessed with the I^2 statistic. Random-effects meta-regression models were conducted to explore study-level factors to explain the heterogeneity observed. Deek's test (Deeks JJ, 2021) was used to evaluate publication bias.

We reported the pooled sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratios. The positive likelihood ratio expresses how many times more likely people with the condition receive a positive test result than those who do not have the condition. In contrast, the negative likelihood ratio expresses how likely it is that people with the condition will receive a negative test result than those who do not have the condition. The inverse of the negative likelihood ratio (1/LR-) can be compared with the positive likelihood ratio to indicate whether the positive or negative test result has a greater impact on the odds of disease. We also present the summary receiver operating characteristic curve from the hierarchical summary receiver operating characteristic model, the prediction region (i.e. for the forecast of the true sensitivity and specificity in a future study). Plotting the summary operating point and its confidence region allowed us to display the trade-off between sensitivity and specificity graphically. Forest plots were used to show within-study estimates and confidence intervals for sensitivity and specificity separately. We report our findings using the PRISMA checklist. The systematic review was conducted with the guidance of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.

ROLE OF THE FUNDING SOURCE

The World Health Organization funded the study and helped with the study design, analysis, interpretation of data, writing a report from this study, and the decision to submit the paper for publication.

RESULTS

We identified a total of 7,814 records using our search strategies, 88 full texts were examined, and 22 articles were eligible and included in the analysis (Figure 1).

Study Characteristics (Table 1)

Most studies were conducted in high-income countries (HIC) as per the fiscal year classification (81.8%, 18 of 22) (Figure 2). Women were the most frequently studied population (73.3%, 17 of 22), and first void urine was the most commonly used sample (63.6%, 14

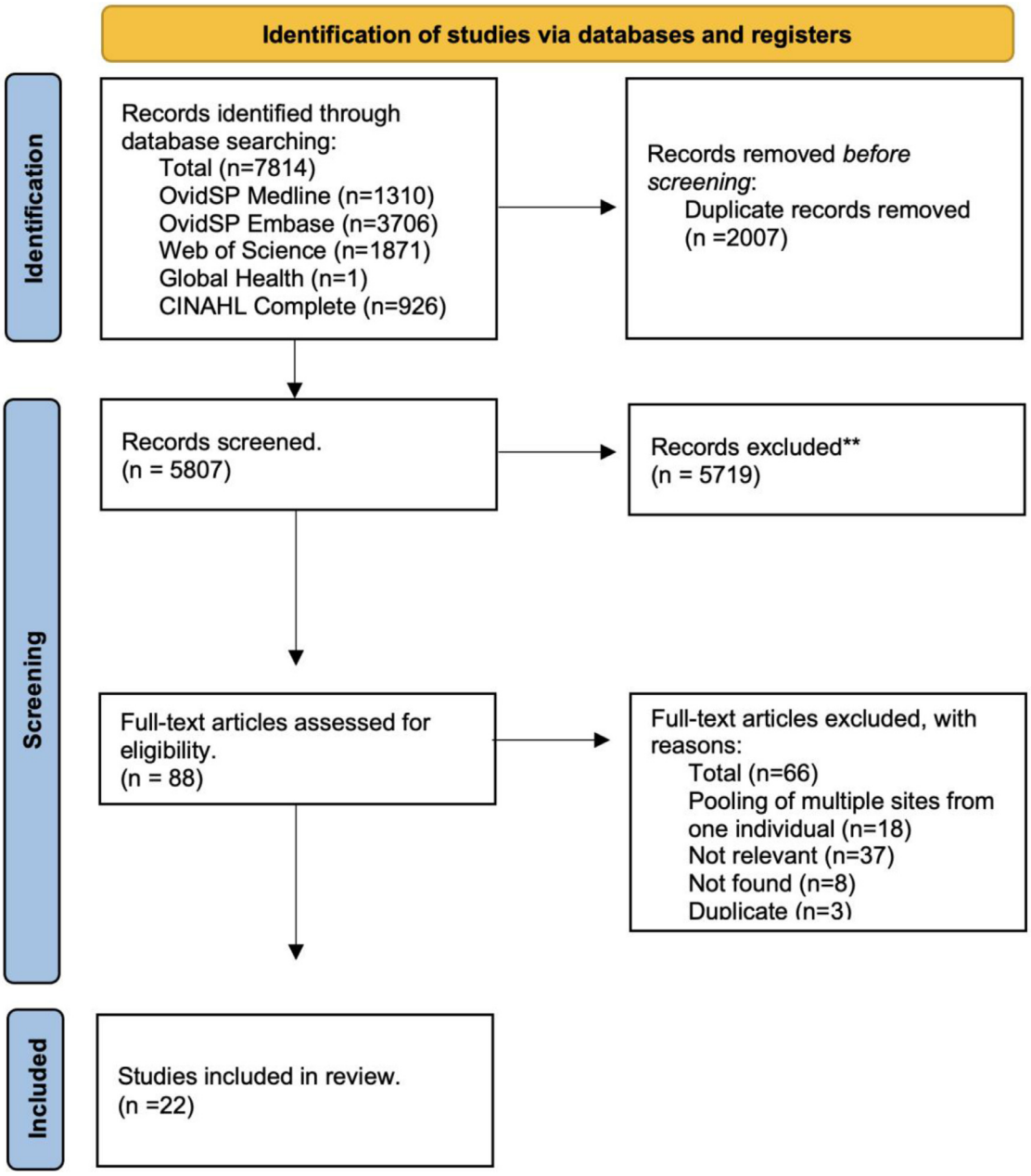

Figure 1. PRISMA Flowchart

of 22), followed by endocervical swabs (36.4%, 8 of 22). The meta-analysis included 18 of 22 studies that reported the diagnostic accuracy of pooled testing. Sixteen of 22 studies discussed the cost-saving aspects of pooling from multiple individuals.

Diagnostic accuracy of pooled testing for chlamydia

Eighteen studies provided 25 estimates for the meta-analysis with data from 6,913 pooled specimens Figure 3. shows that the pooled sensitivity was 98.4% (95% confidence intervals [CI]: 96.8-99.2, I^2=77.5, p<0.001), and pooled specificity was 99.9% (95% CI: 99.6-100.0, I^2=62.6, p<0.001). The diagnostic odds ratio was 82,642 (11,478-595,014), the positive likelihood ratio was 1,296 (228-7,352), the negative likelihood ratio was 0.02 (0.01-0.03), and the inverse negative likelihood ratio was 64 (31-131). Supplementary Figure 1 shows the receiver operating curve, demonstrating the high accuracy of pooling specimens from multiple individuals. Publication bias was likely (p<0.01, Supplementary Figure 2). Supplementary Table 1 summarises the meta-regression results showing lower pooled specificity in studies published after 2010, but no other impact on the accuracy of pooled testing regarding country-income level, study population, pool size, or sample type Figure 3. is the Forest plot of the sensitivity and specificity of pooled testing for chlamydia. Supplementary Table 2 demonstrates the impact on positive and negative predictive values when the background prevalence of chlamydia changes. Table 2 provides the consequences of pooled testing for chlamydia.

Diagnostic accuracy of pooled testing for gonorrhoea

Only 2 studies provided data for the diagnostic accuracy of pooled testing for gonorrhoea (Altwegg et al., 2007; Lindan et al., 2005). One study using 231 specimens from Switzerland reported a sensitivity of 100% and specificity of 97.3% (Altwegg et al., 2007).

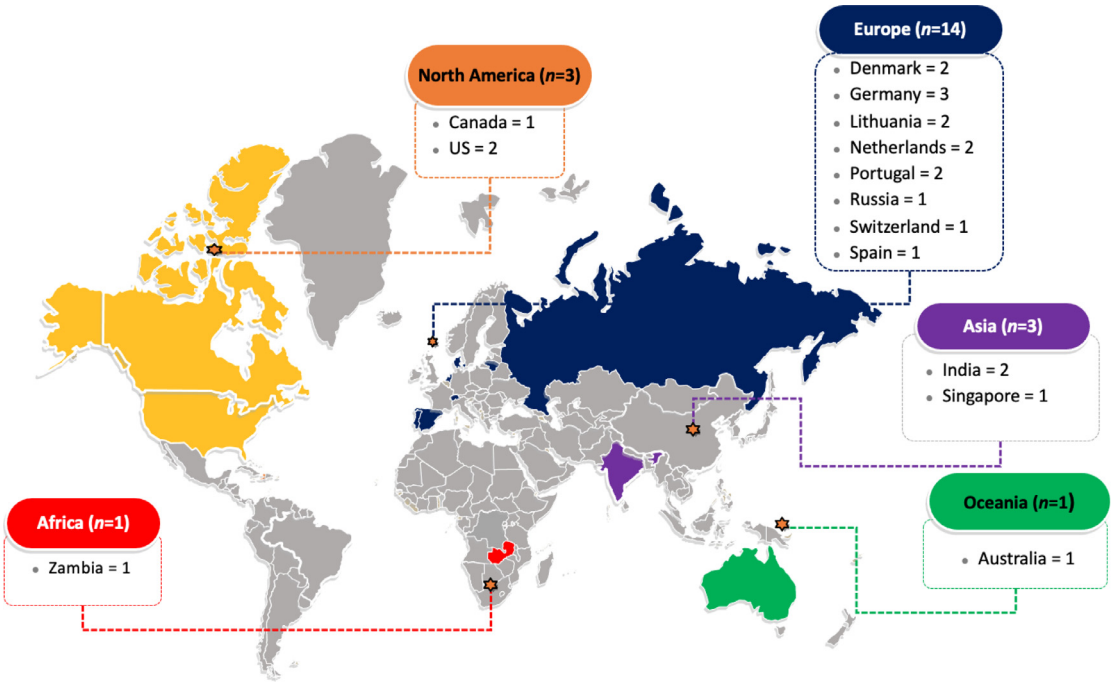

Figure 2. Countries of studies with an evaluation of pooled testing for CT and NG (N=22).

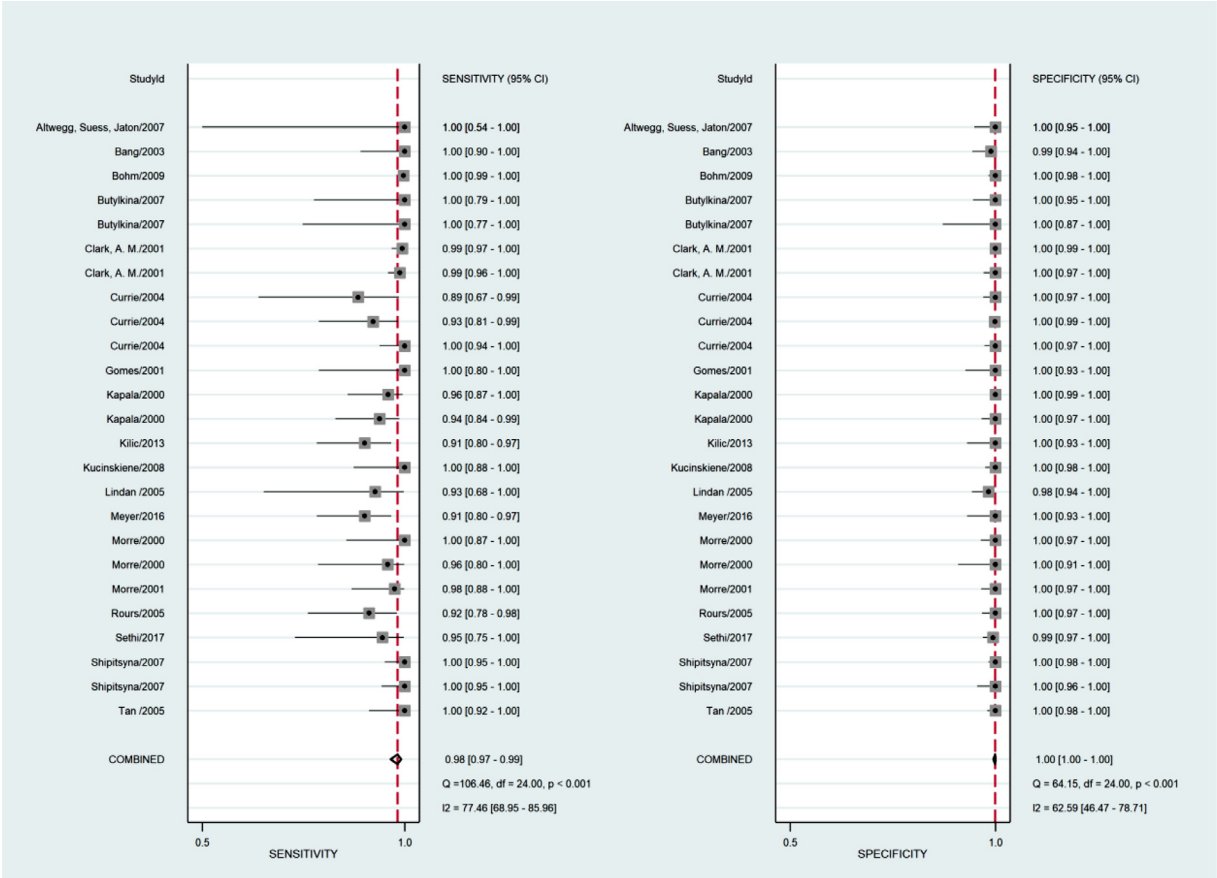

Figure 3. Forest plot of the sensitivity and specificity of pooled testing for chlamydia

Table 1
Characteristics of 22 included studies.

	Total (*N*=22)
Country income level[†]	**n (%)**
High	18 (81.8)
Middle	4 (18.2)
Low	0 (0)
Settings[*]	
Primary care	7 (31.8)
Youth health centres	2 (9.1)
Hospital	1 (4.5)
Community outpatient clinic	2 (9.1)
STI clinic	4 (18.2)
Not specified	10 (45.5)
Secondary school/college	2(9.1)
Populations[*]	
Women	17 (77.3)
Female sex workers	2 (9.1)
Men	7 (31.8)
Not specified	1 (4.5)
Samples used in pooling[*]	
First void urine	14 (63.6)
Endocervical	8 (36.4)
Vaginal	2 (9.1)
Urethral	2 (9.1)
Not specified	2 (9.1)
Outcomes addressed	
Diagnostic accuracy of pooled testing	18 (81.8)
Resource use	16 (72.7)
Acceptability	1 (4.5)
Health equity	5 (22.7)
Actions of pooled sample results	15 (68.2)

[*] Some studies contained more than 1 population/setting/sample type.
[†] As per the New World Bank current 2021 fiscal year.[19]

The other study of 690 men from India reported a sensitivity of 97.3% and specificity of 99.1% (Lindan et al., 2005).

Risk of bias assessment

The risk of bias assessment is presented in Supplementary Figure 3 and Supplementary Table 3. Most studies scored "high" for risk of bias in the patient selection criterion as the nature of the study designs and context infers an automatically high risk of selection bias (i.e. patients were not randomised or recruited consecutively). However, the population selected should have no significant effect on the sensitivity and specificity of nucleic acid amplification test (NAAT). In the study by Gomes et al., a pool of 4 or 8 was created by mixing negative samples with 1 known positive sample, which may lead to biased interpretations of equivocal results. In the study by Clark et al. (2001), an individual reading with a sample-to-cut-off ratio (S/CO) greater than 1 was considered positive, whereas a S/CO of more than 0.2 was considered positive in pooling. Similarly, in the study by Kapala et al. (2000), the S/CO was lowered by 0.2 for pooled testing. Studies could be prone to yield more false positives given a lower cut-off. About half of studies (11 of 21) only retested individuals in positive pools, which led to a detection bias where false negatives in pooled samples would not be detected. Other studies (3 of 21) tested all samples individually before pooling samples to assess for congruency, where any discrepant results in pooled testing were repeated. For example, the study by Tan et al. randomly selected 200 negative samples to retest in addition to all positive samples (Tan and Chan 2005). In the previously mentioned cases, not all negative pools were retested to ensure these were true negatives. This may lead to an overestimation of specificity.

Method of pooling

Table 3 summarises data from 21 studies that provided information on the methodology used in pooled testing. Fifteen studies investigated pooled testing of urine samples from multiple individuals. The most common pooled sample size was pooling from 5 individuals (15 of 21, 71.4%). Two studies compared the diagnostic accuracy between pooling urine samples by 5 and 10 individuals, where pooled testing by 5 demonstrated a slightly greater sensitivity of 99.5% and 100%, respectively, whilst the sensitivity of pooling by 10 was 98.9% and 96.4%, respectively (Clark et al., 2001; Morre et al., 2000). The specificity did not change with pooling by 5 or 10. We did not find a consistent volume of urine added into the pooled sample; the amount varied from 12.5 μL (Butylkina et al., 2007) to 1000 μL (Bohm et al., 2009). Further research is required regarding the ideal urine volume to be used for each pooled sample and the potential for increased urine volumes in sample dilution. Eight studies used endocervical swabs for pooling; among 7 studies that provided the amount of specimen used in pooling, most (n=5) utilised 100 μL per person. Eight studies reported the proportion of inhibited samples in pooled testing (i.e. pooled samples were initially false negatives because of inhibitors in the pool). The proportion of pooled samples that were inhibited varied between 0% and 10%, but pooled testing had a lower proportion of inhibition than individually tested samples.

Table 2
Consequences of pooled testing for chlamydia
Pooled sensitivity: 0.984 (95% CI: 0.968 to 0.992) | **Pooled specificity**: 0.999 (95% CI: 0.996 to 1.0).

Test result	Number of results per 1000 patients tested (95% CI)		Number of pools (studies)	Certainty of the Evidence (GRADE)
	Prevalence 5%	Prevalence 10%		
Cases correctly identified with chlamydia and treated	**49** (48 to 50)	**98** (97 to 99)	6913 (18)	⊕⊕◯◯ LOW [*-§]
Cases of chlamydia missed	**1** (0 to 2)	**2** (1 to 3)		
Cases correctly identified without chlamydia and not treated	**949** (946 to 950)	**899** (896 to 900)	6913 (18)	⊕⊕◯◯ LOW [*-§]
Cases unnecessarily treated for Chlamydia	**1** (0 to 4)	**1** (0 to 4)		

CI = Confidence interval
Explanations
[*] Selection bias noted in 12 studies - participants not enrolled in a randomized or consecutive fashion.[†]Two studies had different cut-offs for individual and pooled screenings. Lower sample-to-cut-off ratio was used to deem a sample positive in pooled testing.[‡]Only positive pools were retested individually in 10 studies. There is potential for unidentified false-negative samples in negative pools.
[§] Deek's test for publication bias ($p < 0.01$)

Table 3
Study characteristics, methods of pooling, reported sensitivity and specificity of pooled testing for chlamydia and gonorrhoea.

Altwegg[11]	Switzerland	N/A	231	Unspecified swabs and FVU pooled separately	3	N/A	50 μL urine, 25 μL swab	Roche Cobas Amplicor	Roche Cobas Amplicor	CT: 100% (6/6)NG: 100% (3/3)	CT: 100% (71/71)NG: 97.3% (72/74)	2.7% (2/73)
Bang	Denmark	9.5	388 females, 104 males	Urethral, endocervical, conjunctival pooled separately	3 or 4	N/A	100 μL urine	Roche Cobas Amplicor CT test	Roche Cobas Amplicor CT test	100 (35/35)	98.9% (97/98)	N/A
Bohm	Germany	5.0	135,799 females	FVU	5	N/A	1000 uL urine	Rotor-Gene 6000^{TM} real-time rotary analyser	Rotor-Gene 6000^{TM} real-time rotary analyser	99.8% (1721/1725)	100% (36400/36400)	0.02%
Butylkina	Lithuania	4.4	410 males military recruits	FVU	5 or 10	Within 24-48 h	25 ul in pool of 5, 12.5 ul in pool of 10	Digene Hybrid Capture II CT/NG Test	Digene Hybrid Capture II CT/NG Test	Pool by 5: 100% (16/16) Pool by 10: 100% (14/14)	Pool by 5: 100% (66/66) Pool by 10: 100% (27/27)	N/A
Clark	US	N/A	3170 females	Endocervical swabs	5 or 10	Upon arrival in laboratory	20 uL in pool of 5, 10 uL in pool of 10	Abbott LCx	Abbott LCx	Pool by 5: 99.5% (187/188) Pool by 10: 98.9% (186/188)	Pool by 5: 100% (446/446) Pool by 10: 100% (129/129)	N/A
Currie	Australia	4.5	715 vaginal swabs, 885 endocervical swabs, 1,000 urine samples.	Vaginal, endocervical, FVU separately	5	Once thawed in laboratory	100 uL each	Roche Cobas	Roche Coba	Vaginal 89.5% (17/19) Endocervical 92.8 (39/42) FVU 100% (63/63)	Vaginal 100% (124/124) Endocervical 99.9% (842/843) FVU 100% (137/137)	Vaginal 4.2%, (6/143), Endocervical 5.8% (10/173) FVU 21/200 (10.5%)
Gomes	Portugal	5.2	330 females	FVU	5	N/A	N/A	Amplicor PCR	Known +ve samples were tested by Roche amplicor	100% (17/17)	100% (49/49)	0%
Gomes	Portugal	14.0	264 females and males	FVU	4 or 8	N/A	40 uL	AMP-CT-TMA/Gen-Probe assay	AMP-CT-TMA/Gen-Probe assay	Pool by 4: 94.3% (33/35) Pool by 8 86.5% (32/37)		N/A
Kapala	Canada	4.1	1288 females	Endocervical	4 or 8	Within 48 hr	100 uL	Abbott LCx	Abbott LCx	Pool by 4: 96.2% (51/53) Pool by 8: 94.3% (50/53)	Pool by 4: 100% (269/269) Pool by 8: 100% (108/108)	N/A
Kilic	Germany	2.1	1649 females	FVU	5	N/A	N/A	PelvoCheck CT/NG	Roche Cobas Amplicor, Abbott Real Time CT/GC assay	90.9% (50/55)	100% (52/52)	N/A
Kucinskiene	Lithuania	5.6	533 High school-aged women	Vaginal	3	N/A	25 uL	Digene Hybrid Capture II CT/GC Test	Digene Hybrid Capture II CT/GC Test	100% (30/30)	100% (147/147)	N/A

(*continued on next page*)

Table 3 (*continued*)

Altwegg[11]	Switzerland	N/A	231	Unspecified swabs and FVU pooled separately	3	N/A	50 μL urine, 25 μL swab	Roche Cobas Amplicor	Roche Cobas Amplicor	CT: 100% (6/6)NG: 100% (3/3)	CT: 100% (71/71)NG: 97.3% (72/74)	2.7% (2/73)
Lewis	US	N/A	2787 males and females	FVU, endocervical, urethral separately	4	N/A	100- uL	Gen-probe Aptima Combo 2 assay		N/A	N/A	N/A
Lindan	India	CT: 2.2 NG: 5.4	690 males	FVU	5	N/A	10 uL	GeneAmp PCR System 9600	GeneAmp PCR System 9600	CT: 93.3% (14/15) NG: 97.3% (36/37)	CT: 98.4% (122/124) NG: 99.1% (106/107)	0% (0/138) in pools, 1.7% (12/690) in individual testing
Lopez-Corbeto	Spain	7.0	1032 16 -25 yo females and males	FVU	3	Upon arrival in laboratory	400 uL	Anyplex II STI-7 detection assay	N/A			N/A
Meyer	Germany	2.1	1650 females in total, 535 pooled	FVU	5	N/A	200 uL	PelvoCheck CT/NG	Divided into three parts for testing: PelvoCheck CT/NG test, COBAS TaqMan CT Test V.2.0 and Abbott RealTime CT/NG	90.9% (50/55)	100% (52/52)	N/A
Morre	Denmark	4.0	650 asymp-tomatic male military recruits	FVU	5 or 10	N/A	100 uL in pool of 5, 50 uL in pool of 10	Roche Amplicor	Roche Amplicor	Pool by 5: 100% (26/26) Pool by 10: 96.1% (25/26)	Pool by 5: 100% (104/104) Pool by 10: 100% (39/39)	0% (0/650) in pools, 0.5% in individual testing (3/650)
Morre	Netherlands	4.0	500 females	Endocervical	5	N/A	50 uL	N/A	N/A	98% (43/44)	100% (106/106)	N/A
Rours	Netherlands	6.4	750 pregnant women	FVU	5	N/A	200 uL	Cobas Amplicor	Cobas Amplicor	92% (34/47)	100% (113/113)	0.7% in pools and 4.9% in individual testing
Sethi	India	N/A	1000 pregnant women	FVU	5	Within 7 days of collection	10 uL	Roche Amplicor	Roche Amplicor	95% (19/20)	99.4% (179/180)	N/A
Shipitsyna	Russia	6.1	1500 asymp-tomatic females	Endocervical	5 or 10	Within 1-3 days	100 uL	Lytech PCR	Lytech PCR	Pool by 5: 100% (80/80) Pool by 10: 100% (69/69)	Pool by 5: 100% (220/220) Pool by 10: 100% (81/81)	N/A
Tan	Singapore	4.1	1200 female sex workers	Endocervical	5	Within 48 hrs	100uL	Roche Cobas Amplicor	Roche Cobas Amplicor	100% (44/44)	100% (192/192)	0% in pools, 1.5% - 2.3% in individual testing

CT = *Chlamydia trachomatis;* FN = false negative; FP = false positive; FVU = first void urine; N/A = not applicable; NG = *Neisseria gonorrhoeae;* PCR = polymerase chain reaction; TN = true negative; TP= true positive.

Resource use

Sixteen studies provided data on the costs of pooled testing (Table 4). Stratified by pool size, the pooling of vaginal swabs by 3 demonstrated a cost reduction of 85% to estimate population prevalence. Further retesting of positive pools for individual diagnosis showed a cost reduction of 70% (Kucinskiene et al., 2008). The pooling of urine samples by 3 demonstrated a 33% decrease in reagent cost as per Lopez-Corbeto et al. (Lopez-Corbeto et al., 2020). The pooling of endocervical swabs by 4 showed a reduction in cost by 47% for the diagnosis of individual positive cases as per Kapala et al. (2000), with a 50% decrease in technician time and 60% decrease in reagent cost as per Lewis et al. The pooling of urine by 5 demonstrated overall cost-savings for diagnosis of positive cases ranging between 39% (Currie et al., 2004) and 62% (Rours et al., 2005), and up to an 80% (Morre et al., 2000) cost reduction for determination of population-based prevalence. Pooling endocervical swabs by 5 demonstrated a 77%-80% reduction in the number of tests (Morre et al., 2001a; Shipitsyna et al., 2007) a 53% total reduction in cost for diagnosis of positive samples (Shipitsyna et al., 2007). Pooling endocervical swabs were demonstrated to decrease technician time by 50%, reagent cost by 55%, and an overall total cost reduction of 63% for the diagnosis of individual cases (Kapala et al., 2000). Pooling endocervical swabs by 10 demonstrated a 47% decrease in reagent cost (Clark et al., 2001) and an overall cost reduction of 44% for diagnosing individual cases (Shipitsyna et al., 2007). Pooling of urine by 10 showed an overall cost reduction of 90% for estimation of population prevalence (Butylkina et al., 2007; Morre et al., 2000) and 54% (Morre et al., 2000) and 56% (Butylkina et al., 2007) cost reduction for diagnosing individual cases. It must be noted that each study assumed a different disease prevalence, which ranged from 4% to 6.4%.

Action post positive pooled test

Eleven studies retested individual samples in positive pools to identify false-positive samples, whereas 3 studies retested discrepant samples compared with individual testing. In Currie's study (Currie et al., 2004), in addition to testing positive samples, specimens in pools containing a negative internal control were diluted and retested individually to determine the presence of false negatives. In the study of Gomes et al. (2002), equivocal samples were reprocessed and retested, yet the positive or negative pools were not retested, potentially increasing the risk of having more false positive and false negatives. The absence of retesting negative pools for the false-negative samples could overestimate the accuracy of pooled testing. These studies did not discuss the treatment options post pooled testing or whether patients were required to return to the clinic to provide specimens for a confirmation test. Possible disadvantages proposed by Morre et al. (2000) were linked to the laboratory burden of deconvoluting and retesting pools, yet the degree of the burden depends on the background CT/NG prevalence. It is possible to retest positive pooled samples within 5 days as urine samples can stay stable during this period without DNA degradation at room temperature. (Morre, 1999) Otherwise, urine specimens could be stored at 4°C, which may require extra storage space (Morre, 1999).

Economic Modelling

Using data from a cross-sectional study conducted in a cohort of HIV-negative women in Zambia, Connolly et al. describe a pooling algorithm and formation of a risk stratification checklist to stratify and guide decisions to pool (Connolly et al., 2020a; Connolly et al., 2020b). Based on identifying factors associated with testing positive for CT/NG through logistic regression modelling, their checklist stratified populations by prevalence and recommended the optimal pool size within each stratum, thus maximising cost-savings. Pooling according to the algorithm results in a 30% cost reduction compared with individual testing and a 52% reduction if combined with syndromic management and presumptive treatment. We identified an economic model developed by van Valkengoed et al. (2001), which was extended by Morre et al. (2002) regarding improving the performance of pooled testing. Morre et al. estimated that pooling urine by 5 would reduce net costs per averted major outcome (pelvic inflammatory disease, chronic pelvic pain, ectopic pregnancy, infertility, and neonatal pneumonia) in asymptomatic women by 57%. When high-performance testing (with 98.8% sensitivity and 99.9% specificity) is assumed, pooled testing would decrease cost by 67%.

Impact on health equity

Health equity aims to allow each person the opportunity to attain his or her full health potential, regardless of socially determined circumstances. An increase in accessibility to low-cost STI testing contributes towards this goal. Butylkina's study discussed that for lower-risk populations providing pooled testing can be economical within large scale screening programs, which could redirect the cost-savings towards screening other STIs, monitor attitudes and knowledge regarding STIs and/or improve health literacy (Butylkina et al., 2007). Likewise, Kucinskiene et al. (2008) emphasise that screening should be incorporated with subsequent examinations, counselling, and testing for other STIs to achieve the maximal benefits. Hence, pooling offers a lower cost screening program that can realise additional benefits previously mentioned through increasing accessibility and diagnostic capacity, especially in resource-limited settings (Tan and Chan 2005).

Impact of dilution and inhibition in different pool sizes

Pooling dilutes the bacterial load and has the potential for a higher false-negative rate. Gomes et al. showed that placing positive samples with samples containing inhibitors in the pool (e.g. urate, phosphate, nitrites) can also mask positive samples, which concurs with the finding in Kapala's study. (Gomes et al., 2002; Kapala et al., 2000). Morre et al. (2001b) offered a potential solution to reduce inhibition by reducing the volume of samples from cervical swabs to 5 uL. In contrast, the dilution effect in the pool can counteract the effect of inhibitors present in the positive samples. This was demonstrated in Gomes' study (Gomes et al., 2002), where 3 positive samples were only detected in the pool of 4 and 8 because of the dilution of inhibitors. The inhibition rate decreased from 1.7% to 0% with pooling of first void urine in 5, demonstrated by Lindan et al. (2005). However, there was little evidence suggesting a standardised optimal dilution ratio. Kapala's study showed that inhibition activity was still detectable after a 1:4 dilution, and Currie's study found the activity of inhibitors in the urine pool of 3 (Currie et al., 2004; Kapala et al., 2000). In summary, dilution in pooled testing can lead to both favourable and unfavourable outcomes. A gap in the present knowledge includes a dilution ratio that balances offsetting the inhibitor effect and maintaining a detectable bacterial load.

Impact of different assays

The majority (16 of 22) of the studies were performed between 2000 and 2010; hence some of the nucleic acid-based assays are nowadays obsolete, including Abbot LCx, AMP-CT-TMA and Roche Amplicor, which were used across 13 studies. For instance, Abbot LCx is now replaced with Abbot Realtime CT/NG assay, according

Table 4
Cost-savings of pooled testing according to number of samples pooled.

STUDY AUTHOR	CURRENCY (YEAR IF AVAILABLE)	SAMPLES USED IN POOLING	NUMBER OF SAMPLES IN ONE POOL	ASSUMED PREVALENCE FOR CT (%)	DECREASE IN REAGENT COST (%)	DECREASE IN TECHNICIAN TIME (%)	REDUCTION IN NUMBER OF TESTS (%)	COST-SAVING FOR ESTIMATION OF POPULATION PREVALENCE (%)*	COST-SAVING FOR DIAGNOSIS OF INDIVIDUAL CASES (%) †
KUCINSKIENE	EUR (2004)	Vaginal swabs	3	5.6	-	-	-	85	70
LOPEZ-CORBETO	EUR (2006)	FVU	3	7	33	-	-	-	-
KAPALA	-	Endocervical swabs	4	4	60	50	-	-	-
LEWIS	USD	FVU, endocervical swabs, urethral swabs	4	-	-	-	-	-	47
BUTYLKINA	-	FVU	5	4.4	-	-	-	80	60
CLARK	USD	Endocervical swabs	5	-	54	-	-	43	-
CURRIE	AUD (2003)	FVU, Vaginal swabs, endocervical swabs separately	5	4.5	43	26	60	-	39
GOMES	EUR	FVU	5	5.2	-	-	-	-	52
LINDAN	-	FVU	5	CT: 2.2 NG: 5.4	50	-	-	-	-
MORRE	EUR	FVU	5	4	-	-	-	80	61
MORRE	EUR	Endocervical swabs	5	4	-	-	77	-	-
ROURS	EUR	FVU	5	6.4	-	-	-	-	62
SETHI	-	FVU	5	-	-	-	-	70	-
SHIPITSYNA	USD	Endocervical swabs	5	6.1	-	-	80	73	53
KAPALA	-	Endocervical swabs	8	4	55	50	-	-	63
BUTYLKINA	-	FVU	10	4.4	-	-	-	90	56
CLARK	USD	Endocervical swabs	10	-	47	-	-	-	-
MORRE	EUR	FVU	10	4	-	-	-	90	54
SHIPITSYNA	USD	Endocervical swabs	10	6.1	-	-	90	54	44

* Population-based screening without further testing of positive pools.
† The subsequent retesting of pools to diagnose individual positive cases.

to the current Food and Drug Administration approved testing devices (Nucleic Acid-Based Tests 2021). Butylkina et al. (2007) and Kucinskiene et al. (2008) used the Hybrid Capture assay, which is an outdated assay with no nucleic acid amplification and can have lower sensitivity as compared with a nucleic acid amplification test (Quint et al., 2007). We performed a sensitivity analysis by comparing currently available assays with obsolete assays and found no difference in sensitivity (p=0.962) or specificity (p=0.590).

DISCUSSION

Our systematic review and meta-analysis appraised the evidence of the diagnostic accuracy and cost-savings of pooled testing from multiple individuals for CT/NG screening. Studies were mainly from HIC and used urine samples from women. Pooled testing maintains high sensitivity and specificity compared with individual testing whilst enabling an economical way to improve the CT/NG screening coverage in lower-risk populations.

The magnitude of cost-savings is dependent on the pool size and background CT/NG prevalence, with the current evidence suggesting that pooled testing is more suitable for low-risk populations. A higher background STI prevalence requires increased retesting of positive pools, which reduces the cost-effectiveness of pooled testing. Populations such as men who have sex with men and some female sex workers have a higher rate of extragenital infections, requiring triple anatomic site testing (Chan et al., 2016). Previously, Peeling et al. (1998a) demonstrated cost-savings when the background STI prevalence was under 20%. The study of Kucinskiene et al. (2008) demonstrated cost-savings of up to 85% using pooled testing among high school female students with a 5.6% chlamydia prevalence. The studies included in this review had an average chlamydia prevalence between 4% and 6% (Rowley et al., 2019). There was no consensus regarding optimal pool size, with no apparent patterns in cost-savings across all studies (Table 4). Connolly examined the relationship between costs and pool sizes. With a pool size of 4 in the low scoring category (low-prevalence of 7.7%), a minimum cost of US$9.4 per sample was achieved, whilst in the mid scoring category with a prevalence of 15.8%, the minimum cost of US$13.3 was achieved in pools of 3. (Connolly et al., 2020b). In this study (Connolly et al., 2020b), the cost was calculated using a predetermined formula that contained the number of pools, the background prevalence and the cost of each GeneXpert cartridge. Therefore, the total cost per sample also would change with the method used for testing. Further research is needed to identify the optimal background prevalence cut-off and pool size for the utility of pooled testing in other settings to maximise cost-effectiveness.

There are other practical considerations for pooled testing. First, there must be consideration of the clinic flow, including how to incorporate time for self-sampling. Connolly et al. recommend a clinic visit structure (which also applies to point-of-care testing) where visits begin with specimen collection and testing, which runs parallel to clients waiting time and clinical counselling regarding risk reduction (Connolly et al., 2020b). Second, laboratories must be prepared to have the capacity to retest individual samples from positive pools in a manner that does not extend the time to treatment (Morre et al., 2000). Minimising the time interval between screening and treatment reduces negative health consequences for the individual and reduces onward transmission (Connolly et al., 2020a; Connolly et al., 2020b). Laboratory quality controls should ensure efficient transport of samples, reduce potential contamination, staff training, and the appropriate facilities for storage of individual samples, which may require retesting. Third, a reduction in technician time was only demonstrated by 2 studies conducted in 2000 and 2004 when the automation of testing was not popular. In these studies, the retesting of positive pools only was the largest contributor to a decrease in laboratory technician time. However, with the automation of sample transfer in modern laboratories, pooling has the potential to increase the laboratory technician's time if performed manually. Further, there may be concerns by laboratory technicians regarding the potential for mistakes and the confusion of samples if a clear protocol of implementing pooled testing is not in place. Thus, future studies must measure the impact on technician time and acceptability of this method by laboratory technicians to evaluate the feasibility of implementing pooled testing accurately.

Fourth, because of the necessity of retesting positive individuals, clients' acceptance of a potential delay in receiving test results may affect the acceptability of pooled testing (Butylkina et al., 2007). Further implementation research will be needed to verify the benefits and potential harms of pooled testing in various settings, especially in assessing the impact of increasing automation of testing, any treatment delays or additional costs incurred with the retesting of individual samples after a positive pool and the legality, in some countries, of reporting results from pooled testing as the samples were not processed following the manufacturers' directions.

The strength of this study is that we systematically reviewed the current evidence for pooled testing for chlamydia and gonorrhoea. However, our findings should be read in light of several limitations. First, we found that studies did not consistently use a standardised pooling method, nucleic based assay system, pool size, or specimen volume. However, it is reassuring that despite different methodologies used, we identified consistently high sensitivity and specificity. Second, most studies came from HIC for women using urine samples to test for chlamydia. Thus, more research is needed in low- and middle-income countries where pooled testing is likely to have the most significant impact in reducing costs and improving testing coverage. Given that the laboratory environment and facilities in low- and middle-income countries might not be as advanced as those in HIC (e.g. to prevent cross-contamination), the potential reduction in the accuracy of pooled testing and the need to ensure quality laboratory facilities could reduce the benefits of this pooling strategy. Thus, more research in low- and middle-income countries is needed to determine the feasibility of implementing pooled testing. Third, we found that most studies did not retest the negative pools, which could overestimate the specificity. Future studies will need to address this issue. Fourth, we only found 2 studies related to pooled testing for gonorrhoea; therefore, our review provides stronger evidence for pooled testing for CT than for NG. Although the 2 studies for NG reported high sensitivity and specificity, there remains some uncertainty regarding the diagnostic accuracy of pooled testing for gonorrhoea. Furthermore, several assays included in this review may no longer be available. Therefore, we recommend that before pooled testing is implemented, there should be a context-specific evaluation of this approach, including the acceptability (by patients and providers), feasibility (impact on clinic flow, laboratory technician time, costs) and performance of the locally used molecular assay (including optimal volume used). Last, the sensitivity of first-pass urine might be lower than vaginal swabs (Van Der Pol et al., 2019), but we only found 2 studies that evaluated the pooled sensitivity of chlamydia using vaginal swabs. Therefore, even though we demonstrate high accuracy of pooled urine samples, future studies should confirm the accuracy of pooled vaginal swabs.

Our systematic review and meta-analysis found that single anatomic site pooled testing for CT is highly sensitive and specific compared with individual testing. This approach can reduce screening costs in low-prevalence populations for which single anatomic site screening is recommended. This can facilitate higher numbers of people being screened, and with prompt treatment,

both reduce adverse health consequences for the individual and breaks the chain of transmission to uninfected partners.

CONFLICTS OF INTERESTS

All authors declare they do not have any conflicts of interest.

Funding

This study is funded by World Health Organization through a grant from the Bill and Melinda Gates Foundation. JJO and EPFC are supported by the Australian National Health and Medical Research Council (NHMRC) Emerging Leadership Investigator Grant (GNT1193955, GNT1172873). CKF is supported by an Australian NHMRC Leadership Investigator Grant (GNT1172900).

CONTRIBUTIONS

JJO, EPFC, TW, and MBM designed the research study. LA and YX conducted the screening and data extraction. LA, YX, JJO, and EPFC analysed the data and wrote the first draft of the manuscript. All authors have read and approved the final manuscript.

ETHICAL APPROVAL

No ethical approvals were needed for this study as it was a review of literature without the direct involvement of patients.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijid.2022.03.009.

REFERENCES

Altwegg M, Suess O, Jaton K, et al. Specimen pooling for the detection of Chlamydia trachomatis and Neisseria gonorrhoeae in urogenital specimens. International Journal of Antimicrobial Agents 2007;29:S239–40. doi:10.1016/s0924-8579(07)70760-4.

Bohm I, Groning A, Sommer B, et al. A German Chlamydia trachomatis screening program employing semi-automated real-time PCR: results and perspectives. Journal of Clinical Virology 2009;46 Suppl 3:S27–32.

Burdett A, Toumazou C, Sahoo R, et al. Pooled sputum to optimise the efficiency and utility of rapid, point-of-care molecular SARS-CoV-2 testing. BMC Infect Dis 2021;21:665. doi:10.1186/s12879-021-06316-z.

Butylkina R, Jusevicite V, Kasparaviciene G, et al. Pooling of urine specimens allows accurate and cost-effective genetic detection of Chlamydia trachomatis in Lithuania and other low-resource countries. Scandinavian Journal of Infectious Diseases 2007;39:209–12. doi:10.1080/00365540600978914.

Chan PA, Robinette A, Montgomery M et al. (2016) Extragenital Infections Caused by*Chlamydia trachomatis*and*Neisseria gonorrhoeae*: A Review of the Literature Infectious Diseases in Obstetrics and Gynecology 2016:5758387 doi:10.1155/2016/5758387

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol 2006;59:1331–2 author reply 1332-1333. doi:10.1016/j.jclinepi.2006.06.011.

Clark AM, Steece R, Crouse K, et al. Multisite pooling study using ligase chain reaction in screening for genital Chlamydia trachomatis infections. Sexually Transmitted Diseases 2001;28:565–8.

Connolly S, Kilembe W, Inambao M, et al. A Population-Specific Optimized GeneXpert Pooling Algorithm for Chlamydia trachomatis and Neisseria gonorrhoeae To Reduce Cost of Molecular Sexually Transmitted Infection Screening in Resource-Limited. Settings Journal of Clinical Microbiology 2020a;58:24.

Connolly S, Kilembe W, Inambao M, et al. A population-specific optimized genexpert pooling algorithm for chlamydia trachomatis and neisseria gonorrhoeae to reduce cost of molecular sexually transmitted infection screening in resource-limited settings. Journal of Clinical Microbiology 2020b;58(9) no pagination.

Cuevas LE, Santos VS, Lima S, et al. Systematic Review of Pooling Sputum as an Efficient Method for Xpert MTB/RIF Tuberculosis Testing during the COVID-19. Pandemic Emerg Infect Dis 2021;27:719–27. doi:10.3201/eid2703.204090.

Currie MJ, McNiven M, Yee T, et al. Pooling of clinical specimens prior to testing for Chlamydia trachomatis by PCR is accurate and cost saving. Journal of Clinical Microbiology 2004;42:4866–7.

Deeks JJ HJ, Altman DG, editors. Chapter 10: Analysing data and undertaking meta–analyses. Cochrane Handbook for Systematic Reviews of Interventions 2021 version 6.2 (updated February 2021) edn.

Gomes JP, Viegas S, Paulino A, et al. Sensitivity evaluation of the Gen-Probe AMP-CT assay by pooling urine samples for the screening of Chlamydia trachomatis urogenital infection. International Journal of STD & AIDS 2002;13:540–542.

Kacena KA, Quinn SB, Howell MR, et al. Pooling urine samples for ligase chain reaction screening for genital Chlamydia trachomatis infection in asymptomatic women. Journal of clinical microbiology: JCM 1998;36:481–5. doi:10.1128/JCM.36.2.481-485.1998.

Kapala J, Copes D, Sproston A, et al. Pooling cervical swabs and testing by ligase chain reaction are accurate and cost-saving strategies for diagnosis of Chlamydia trachomatis. Journal of Clinical Microbiology 2000;38:2480–3.

Kucinskiene V, Juseviciute V, Valiukeviciene S, et al. Home sampling and pooling of vaginal samples are effective tools for genetic screening of Chlamydia trachomatis among high school female students in. Lithuania Scandinavian Journal of Infectious Diseases 2008;40:88–93.

Lindan C, Mathur M, Kumta S, et al. Utility of pooled urine specimens for detection of Chlamydia trachomatis and Neisseria gonorrhoeae in men attending public sexually transmitted infection clinics in Mumbai, India. by PCR Journal of Clinical Microbiology 2005;43:1674–7.

Lopez-Corbeto E, Gonzalez V, Lugo R, et al. Pooling of urine samples for molecular detection of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium as a screening strategy among young adults in Catalonia Enfermedades Infecciosas. y. Microbiologia Clinica 2020;38:65–71.

Morre SA, Meijer C, Munk C, et al. Pooling of urine specimens for detection of asymptomatic Chlamydia trachomatis infections by PCR in a low-prevalence population: Cost-saving strategy for epidemiological studies and screening programs. Journal of Clinical Microbiology 2000;38:1679–80.

Morre SA, van Dijk R, Meijer C, et al. Pooling cervical swabs for detection of Chlamydia trachomatis by PCR: Sensitivity, dilution, inhibition, and cost-saving aspects. Journal of Clinical Microbiology 2001a;39:2375–6.

Morre SA, Van Dijk R, Meijer CJLM, et al. Pooling cervical swabs for detection of Chlamydia trachomatis by PCR: Sensitivity, dilution, inhibition, and cost-saving aspects. [7]Journal of Clinical Microbiology 2001b;39:2375–6.

Morre SA, Van Valkengoed IGM, De Jong A, Boeke AJP, Van Eijk JTM, Meijer CJLM, Van den Brule AJC. Mailed, Home-Obtained Urine Specimens: a Reliable Screening Approach for Detecting Asymptomatic Chlamydia trachomatis. Infections JOURNAL OF CLINICAL MICROBIOLOGY 1999;4:976. doi:10.1128/JCM.37.4.976-980.1999.

Morre SA, Welte R, Postma MJ. Major improvements in cost effectiveness of screening women for Chlamydia trachomatis using pooled urine specimens and high performance testing. [4]Sexually Transmitted Infections 2002;78:74–75.

Nucleic Acid Based Tests. (2021) U.S. Food and Drug Administration. https://www.fda.gov/medical-devices/in-vitro-diagnostics/nucleic-acid-based-tests#microbial. Accessed 07/02/2022 2022.

Peeling RW, Toye B, Jessamine P, et al. Pooling of urine specimens for PCR testing: a cost saving strategy for Chlamydia trachomatis control programmes. Sex Transm Infect 1998a;74:66–70. doi:10.1136/sti.74.1.66.

Peeling RW, Toye B, Jessamine P, et al. Pooling of urine specimens for PCR testing: a cost saving strategy for Chlamydia trachomatis control programmes. Sexually transmitted infections 1998b;74:66–70. doi:10.1136/sti.74.1.66.

Quint K, Porras C, Safaeian M, et al. Evaluation of a novel PCR-based assay for detection and identification of Chlamydia trachomatis serovars in cervical specimens. J Clin Microbiol 2007;45:3986–91. doi:10.1128/jcm.01155-07.

Rours G, Verkooyen RP, Willemse HFM, et al. Use of pooled urine samples and automated DNA isolation to achieve improved sensitivity and cost-effectiveness of large-scale testing for Chlamydia trachomatis in pregnant women. Journal of Clinical Microbiology 2005;43:4684–90. doi:10.1128/jcm.43.9.4684-4690.2005.

Rowley J, Vander Hoorn S, Korenromp E, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates. 2016 Bull World Health Organ 2019;97:548–62. doi:10.2471/blt.18.228486.

Schunemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. J Clin Epidemiol 2020;122:129–41. doi:10.1016/j.jclinepi.2019.12.020.

Shipitsyna E, Shalepo K, Savicheva A, et al. Pooling samples: The key to sensitive, specific and cost-effective genetic diagnosis of Chiamydia trachomatis in low-resource countries. Acta Dermato-Venereologica 2007;87:140–3.

Tan HH, Chan R. Use of polymerase chain reaction on pooled cervical swabs to detect Chlamydia trachomatis infections in female sex workers in Singapore. Singapore Medical Journal 2005;46:215–18.

Van Der Pol B, Fife K, Taylor SN, et al. Evaluation of the Performance of the Cobas CT/NG Test for Use on the Cobas 6800/8800 Systems for Detection of Chlamydia trachomatis and Neisseria gonorrhoeae in Male and Female Urogenital Samples. J Clin Microbiol 2019;57. doi:10.1128/JCM.01996-18.

van Valkengoed IG, Postma MJ, Morré SA, et al. Cost effectiveness analysis of a population based screening programme for asymptomatic Chlamydia trachomatis infections in women by means of home obtained urine specimens. Sexually transmitted infections 2001;77:276–82. doi:10.1136/sti.77.4.276.

Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36. doi:10.7326/0003-4819-155-8-201110180-00009.

World Health Organization. Laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus; 2021 https://www.who.int/medical_devices/diagnostics/selection_in-vitro/selection_in-vitro-meetings/00007_01_WHO_Laboratory_Manual_STIs.pdf Accessed 7th August.