Self-sampling strategies (with/without digital innovations) in populations at risk of *Chlamydia trachomatis* and *Neisseria gonorrhoeae*: a systematic review and meta-analyses

Fiorella Vialard, Apoorva Anand, Cindy Leung Soo, Anna de Waal, Madison McGuire, Sergio Carmona, Marta Fernández-Suárez, Alice Anne Zwerling, Nitika Pant Pai

**ABSTRACT**

**Background** *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (GC) resulted in over 200 million new sexually transmitted infections last year. Self-sampling strategies alone or combined with digital innovations (ie, online, mobile or computing technologies supporting self-sampling) could improve screening methods. Evidence on all outcomes has not yet been synthesised, so we conducted a systematic review and meta-analysis to address this limitation.

**Methods** We searched three databases (period: 1 January 2000–6 January 2023) for reports on self-sampling for CT/GC testing. Outcomes considered for inclusion were: accuracy, feasibility, patient-centred and impact (ie, changes in linkage to care, first-time testers, uptake, turnaround time or referrals attributable to self-sampling). We used bivariate regression models to meta-analyse accuracy measures from self-sampled CT/GC tests and obtain pooled sensitivity/specificity estimates. We assessed quality with Cochrane Risk of Bias Tool-2, Newcastle–Ottawa Scale and Quality Assessment of Diagnostic Accuracy Studies-2 tool.

**Results** We summarised results from 45 studies reporting self-sampling alone (73.3%; 33 of 45) or combined with digital innovations (26.7%; 12 of 45) conducted in 10 high-income (HICs; n=3) and 8 low/middle-income countries (LMICs; n=11). 95.6% (43 of 45) were observational, while 4.4% (2 of 45) were randomised clinical trials. We noted that pooled sensitivity (n=13) for CT/GC was higher in extragenital self-sampling (>91.6% [86.0%–95.1%]) than in vaginal self-sampling (79.6% [62.1%–90.3%]), while pooled specificity remained high (>99.0% [98.2%–99.5%]). Participants found self-sampling highly acceptable (80.0%–100.0%; n=24), but preference varied (23.1%–83.0%; n=16).

Self-sampling reached 51.0%–70.0% (n=3) of first-time testers and resulted in 89.0%–100.0% (n=3) linkages to care. Digital innovations led to 65.0%–92% engagement and 43.8%–57.1% kit return rates (n=3).

**Quality of studies varied.**

**Discussion** Self-sampling had mixed sensitivity, reached first-time testers and was accepted with high linkages to care. We recommend self-sampling for CT/GC in HICs but additional evaluations in LMICs. Digital innovations impacted engagement and may reduce disease burden in hard-to-reach populations.

**WHAT THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY**

Our review offers evidence of strengths and limitations of self-sampling strategies in diverse populations and income settings, which will be useful for policymakers when implementing screening strategies that can be customised to key populations.

**INTRODUCTION**

**Rationale**

According to the WHO, global annual incidence of common sexually transmitted infections (STIs), *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (GC), is 128 million and 82 million cases, respectively. CT/GC infections can have profound
impacts on physical, social and psychological health of key populations such as men who have sex with men (MSM), sex workers, people who use injection drugs and pregnant women. Untreated infections can cause infertility, pregnancy defects and pelvic inflammatory diseases or increase the risk of acquiring HIV. In addition, antimicrobial resistance is a concern with GC, compromising infection control efforts.

To remain on track with WHO STI elimination targets for 2030 and to address disruptions in STI screening services due to the COVID-19 pandemic, our healthcare systems require an increased use of contactless, affordable, rapid, reliable, point-of-care test (POCT)-based self-sampling (SS) and/or self-testing (ST) technologies. SS-based STI testing can be conducted independently by participants in a clinic setting, at home or in other dedicated locations. SS requires that participants receive their results from healthcare workers, while ST enables them to obtain their results directly. SS has multiple advantages including ensuring confidentiality and convenience of key populations who struggle with lack of access to conventional testing. SS can be combined with digital health innovations to engage these populations.

Digital health innovations are defined by the WHO as online-based, mobile-based or computing technologies that support health interventions. Currently, digital innovations (eg, apps, websites, messenger-based assistants) are in development for ST/SS POCT. Evidence from HIV ST indicates that these tools may offer advantages of increased linkage to care, proportion of first-time testers and partner referrals. Disseminating best practices and adoption of evidence-based digital health interventions are two WHO strategies for STI-endemic countries. To date, very few studies have compared conventional testing for CT/GC with SS-based testing coupled with digital innovations, justifying the rationale for this review. Furthermore, little is known about the effect of digital innovations on SS performance in populations of varying health and digital literacy across diverse income settings.

A systematic review reported higher accuracy of nucleic acid amplification tests (NAATs) for CT detection compared with antigen-based tests but failed to evaluate SS. A systematic review by Lunny et al, published 8 years ago, reported comparable accuracy of self-obtained samples with clinician-obtained samples (n=6100 paired samples). However, extragenital (ie, rectal and pharyngeal) sampling sites, which are more crucial for MSM, were excluded from their meta-analysis. To date, no systematic review has evaluated evidence beyond accuracy such as patient-reported or implementation research outcomes (ie, acceptability/preference, feasibility and impact). To address these limitations and to generate evidence for guideline and product development, we conducted this review.

Objectives
From global data on observational and randomised clinical trial (RCT) studies, we aimed to determine whether SS for CT/GC with or without a digital innovation was accurate, feasible, and impacted patient-reported or implementation research outcomes, compared with conventional testing.

MATERIALS AND METHODS
Search strategy
We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for reporting, Cochrane guidelines for conducting the review (online supplemental document 1) and registered the protocol with PROSPERO (registration number: CRD42021262950).

For the period of 1 January 2000–6 January 2023, four reviewers (MM, AdW, FV, AA) searched three electronic databases (PubMed, Embase and LILACS; search string described in online supplemental document 2) without language restrictions and retrieved full-text studies or conference abstracts. Two authors (AA, FV) independently performed a double screening of all publications (figure 1).

Eligibility criteria
We included publications on SS quantitative accuracy, feasibility, patient-centred and impact outcomes for CT/GC testing with or without digital health interventions.

We excluded editorials, commentaries, reviews, case reports, qualitative, narrative and modelling studies, those in non-English/Spanish language or without SS-specific, CT/GC-specific outcomes.

Data abstraction
Two reviewers (FV, AA) independently abstracted final data from 45 studies. A senior reviewer (NPP) verified data abstraction and resolved disagreements.

Abstracted data included: study design, study location, sample size, study population, digital innovation type, socioeconomic setting, outcome measures and associated metrics.

For diagnostic accuracy, we abstracted 2×2 table values when available and contacted authors for additional data. We narratively synthesised all other outcomes (ie, acceptability, preference, feasibility and impact).

Summary outcome measures and narrative synthesis of results
Table 1 describes abstracted CT/GC outcome measures. We defined digital innovations as assisting with either (1) test procurement via websites or social media; (2) SS conduct via mobile application or online material; and/or (3) communication of results and treatment steps via text messages or online portals.

Data analysis
All statistical analyses were performed in R (V4.1.2 or later). We abstracted true positive, false positive (FP), false negative (FN) and true negative data for CT/GC SS (index test) versus clinician sampling (reference test). We generated sensitivity and specificity forest plots using the Mantel-Haenszel method within the meta package. We assessed heterogeneity using I² as indicator and excluded categories with less than three studies.

We performed random-effects bivariate regression models. We obtained summary values with 95% CIs using the restricted maximum likelihood method within the bivariate-based mada package and generated summary receiver operating characteristic (SROC) curves pooled by pathogen, sampling site and study populations. SROC curves represent summary plots of sensitivity and specificity, with 95% joint intervals in two-dimensional space using diagnostic OR as the outcome to determine overall accuracy. We assessed publication bias using funnel plots and Egger’s test.

Sensitivity analysis
We compared test accuracy in low/middle-income (LMICs) to high-income countries (HICs) and performed sensitivity analyses by removing poor-quality studies.
Quality assessment

Two independent reviewers (FV and AA) performed quality assessment. We used Cochrane Risk of Bias Tool-2, Newcastle–Ottawa Scale (NOS) and Quality Assessment of Diagnostic Accuracy Studies-2 for RCTs, cohort/cross-sectional studies and diagnostic accuracy studies, respectively.13–15 A senior reviewer (NPP) was consulted.

RESULTS

Study selection

Of 122 initial records, 45 studies were included in our synthesis (figure 1). Characteristics and key findings of included studies are summarised in online supplemental tables 1 and 2.

### Table 1  Outcome descriptions for study inclusion and synthesis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>Sensitivity and specificity of index self-sampling as compared with laboratory-confirmed reference standards.65 66</td>
</tr>
<tr>
<td>Preference</td>
<td>Patient preference for an SS option over clinician-based sampling, documented as a proportion over the total number of participants who were offered testing.62</td>
</tr>
<tr>
<td>Feasibility</td>
<td>Feasibility of SS strategies for CT/GC testing, measured by completion rate, test return rate, patient ability to interpret results and reasons for not returning the test.62</td>
</tr>
<tr>
<td>Impact</td>
<td>Net change from baseline in a particular group attributable to an SS strategy for the following metrics: (1) linkage to care; (2) number of first-time testers who used the SS option; (3) partner referrals; (4) turnaround time (TAT) to test results; (5) detection of new CT/GC cases; and (6) test uptake from baseline.62 Metrics were reported as proportions with CIs. TAT was reported as median time with ranges.</td>
</tr>
</tbody>
</table>

CT/GC SS and digital innovations

Of all studies, 26.7% (12 of 45) integrated digital innovations (figure 2C).16–27 These included six with websites to obtain SS kits and/or information on test conduct, combined with notification...
of results via text (13.3%),

five website-based only (11.1%)16 19 22 23 26 and one text-based only (2.2%) strategies.21

Accuracy of SS: meta-analysis

Eighty per cent of studies (36 of 45) detected newly diagnosed positive CT/GC infections and 42.2% of studies (19 of 45) measured test accuracy.16 17 20 22–55 Seventeen studies reported 2×2 table values.22 28 30–33 35 36 38–43 51–53 We excluded four studies where reference test was unclear/not reported.22 31–33 See online supplemental table 3 for case positivity, sensitivity and specificity for CT/GC SS tests.

We generated pooled diagnostic accuracy values (table 2) and SROC curves (figure 3) by sampling site and study population for CT/GC from 13 studies.28 30 32 35 38–43 51 Bivariate regression results revealed overall high specificity and sensitivity for CT and GC with no statistically significant difference between pathogens across included studies (table 2).

Table 2 Sensitivity, specificity and DOR of Chlamydia trachomatis and Neisseria gonorrhoeae self-sampling-based tests with associated 95% CIs, I² metric and p values

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (95% CIs)</th>
<th>P value</th>
<th>Specificity (95% CI)</th>
<th>P value</th>
<th>DOR (95% CI)</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C. trachomatis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>88.2 (81.9–92.6)</td>
<td>Ref</td>
<td>99.3 (98.9–99.5)</td>
<td>Ref</td>
<td>1110 (472–2220)</td>
<td>87.6</td>
</tr>
<tr>
<td>Sampling site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal</td>
<td>79.9 (62.1–90.3)</td>
<td>Ref</td>
<td>96.0 (96.3–99.5)</td>
<td>Ref</td>
<td>479 (124–1290)</td>
<td>84.2</td>
</tr>
<tr>
<td>Pharyngeal</td>
<td>92.5 (86.5–95.9)</td>
<td>0.20</td>
<td>99.7 (99.5–99.8)</td>
<td>0.02</td>
<td>4550 (1710–9880)</td>
<td>0.0</td>
</tr>
<tr>
<td>Rectal</td>
<td>91.3 (87.8–96.2)</td>
<td>0.01</td>
<td>99.4 (98.8–99.7)</td>
<td>0.25</td>
<td>2490 (638–6770)</td>
<td>79.0</td>
</tr>
<tr>
<td>Study population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>85.0 (73.5–92.1)</td>
<td>0.30</td>
<td>99.3 (98.7–99.6)</td>
<td>0.74</td>
<td>895 (262–2260)</td>
<td>87.1</td>
</tr>
<tr>
<td>MSM</td>
<td>91.6 (84.9–95.5)</td>
<td>Ref</td>
<td>99.2 (98.8–99.4)</td>
<td>Ref</td>
<td>1390 (551–2940)</td>
<td>21.0</td>
</tr>
<tr>
<td><strong>N. gonorrhoeae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>88.9 (84.2–92.3)</td>
<td>0.79</td>
<td>99.1 (96.6–99.4)</td>
<td>0.38</td>
<td>937 (459–1710)</td>
<td>78.3</td>
</tr>
<tr>
<td>Sampling site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal</td>
<td>79.9 (68.6–87.9)</td>
<td>Ref</td>
<td>99.1 (98.4–99.5)</td>
<td>Ref</td>
<td>502 (182–1120)</td>
<td>70.8</td>
</tr>
<tr>
<td>Pharyngeal</td>
<td>94.3 (89.6–96.9)</td>
<td>0.003</td>
<td>99.3 (97.1–98.6)</td>
<td>0.099</td>
<td>3370 (422–12600)</td>
<td>0.0</td>
</tr>
<tr>
<td>Rectal</td>
<td>91.6 (86.0–95.1)</td>
<td>0.01</td>
<td>99.2 (98.4–99.6)</td>
<td>0.088</td>
<td>1620 (426–4340)</td>
<td>63.8</td>
</tr>
<tr>
<td>Study population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>83.5 (74.1–89.9)</td>
<td>0.026</td>
<td>99.3 (98.8–99.6)</td>
<td>0.033</td>
<td>838 (298–1890)</td>
<td>73.6</td>
</tr>
<tr>
<td>MSM</td>
<td>92.9 (88.1–95.9)</td>
<td>Ref</td>
<td>97.7 (96.9–98.3)</td>
<td>Ref</td>
<td>573 (307–980)</td>
<td>46.2</td>
</tr>
</tbody>
</table>

Significant values are bolded.

DOR, diagnostic OR; MSM, men who have sex with men.
However, we found statistically significant differences in specificity and sensitivity by sampling site and populations (table 2). For CT, sensitivity estimates (with 95% CI) were 93.1 (87.8–96.1) for rectal to 79.6 (62.1–90.3) for vaginal samples; while specificity remained high: 99.7 (99.5–99.8) for pharyngeal to 99.0 (98.2–99.5) for vaginal samples.

For GC, sensitivity estimates (with 95% CI) were 94.3 (89.6–96.9) for pharyngeal to 79.9 (68.6–87.9) for vaginal samples, with no differences in specificity.

In women compared with MSM, we also found a statistically significant lower sensitivity (83.5 (74.1–89.9) vs 92.9 (88.1–95.9)) but higher specificity (99.3 (98.8–99.6) vs 97.7 (96.9–98.3)). We excluded a single study with SS urine compared with pooled clinician-collected pharyngeal and rectal samples from the regression analysis, with sensitivity of 30.0% (16.0%–48.0%) for CT/GC, because we could not generate a pooled estimate for this different type of sampling.

The funnel plot coefficient slope was −1.30 with a p value of 0.39, indicating low publication bias (online supplemental figure 1).

**SS without digital innovations**

Acceptability for SS was high (80.0%–100.0%; n=21). Participants found SS easy (n=14) and comfortable (ie, painless; n=9), and would recommend it to others (n=3). Only one study in MSM from China reported a lower acceptability for rectal SS (43.6%; 133 of 306) but 100% acceptability for urine SS (306 of 306).

Preference for SS over clinician sampling ranged from 23.1% (823 of 3082) to 83.7% (429 of 512) (n=13). High trust in accuracy and privacy were cited as reasons for preferring SS (n=3). Reasons for not accepting SS (306 of 306). SS completion rate was high (83.8% (341 of 363)–99.0% (346 of 350) in three studies. Participants in a European study were willing to pay between €10 and €20 for SS without digital innovations.

**Quality**

Overall, most (91.7%; 22 of 24) cross-sectional studies (figure 4A) selected a representative sample and assessed exposure correctly but rated poorly in description of non-respondents (unclear for 54.2%; 13 of 24). Cohort studies (figure 4B) had good representativeness (100%; 4 of 4), acceptable comparability between cohorts (75.0%; 3 of 4) and scored poorly on remaining NOs criteria. We assessed one RCT that had low risk of bias for all criteria but performance and detection. Accuracy studies (figure 4C) had low risk for most bias (10–16 of 16; 62.5%–100.0%), except in inclusion criteria (6 of 16; 37.5%). In 62.5% of studies, it was unclear whether results from referral sites were included.

Acceptability of SS with digital innovations was reported by one study. SS was successful in detecting 36.8% (14 of 38) and 17.9% (5 of 28) of new infections and detecting 33.1 CT cases per 100 person-years (51/154.29 years) and 29.9 GC cases per 100 person-years (48/160.77 years). Test uptake was high (85.3%; 110 of 129) in one study offering SS kits via text messaging. In one UK study, 89.0% (93 of 105) of positive participants were linked to care, with 57.0% (60 of 105) accessing their results, consultation and treatment remotely through the website. Turnaround time to treatment was same day to 24 days (median of 1 day; IQR 0–4). In the same study, out of 105, 13 partners were notified and 9 received treatment.

**SS with digital innovations**

Due to heterogeneity in reported metrics and limited number of studies (n=12), we were unable to provide ranges for reported outcomes.

Acceptability of SS with digital innovations (website based only or mixed with text reminders) was reported by participants rating the intervention as: acceptable (93.0%; 1660 of 1785), very satisfactory (95.5%; 407 of 426) or easy to use (89.0%; 336 of 396) in three studies.

Preference for SS and digital innovation was reported by 54.3% (969 of 1785) of participants who indicated they would still test for CT/GC if at-home SS was not available to them, but would do so less often in one study. In two other studies, 51.5% (103 of 200) and 77.0% (307 of 399) of participants preferred this strategy over clinician-based in-hospital testing.

Feasibility of these strategies was reported in six studies. Text-based reminders increased kit return (43.8% (28 of 64)–57.1% (16 of 28) in MSM) and engagement in repeat testing (65.0%; 278 of 427 in women). Website-based questionnaires maintained high completion rates (66.5%, 69 of 105 and 92.9%, 134 of 146; n=2), showing that online-based digital tools can engage individuals with CT/GC testing. One study reported that all returned SS kits were correctly used by participants. Reasons for not accepting SS were: not sexually active, no perceived risk or recently tested for STIs. Cost was also associated with unwillingness to pay for home-based testing (37.0%; 660 of 1785) in one study and led to a lower completion rate (75.0%; 531 of 708) in another.

In one study with online-based ordering, kit return was 56.8% (1948 of 3428).

Impact of SS with digital innovations was reported in five studies. SS was successful in detecting 36.8% (14 of 38) and 17.9% (5 of 28) of new infections and detecting 33.1 CT cases per 100 person-years (51/154.29 years) and 29.9 GC cases per 100 person-years (48/160.77 years). Test uptake was high (85.3%; 110 of 129) in one study offering SS kits via text messaging. In one UK study, 89.0% (93 of 105) of positive participants were linked to care, with 57.0% (60 of 105) accessing their results, consultation and treatment remotely through the website. Turnaround time to treatment was same day to 24 days (median of 1 day; IQR 0–4).

In the same study, out of 105, 13 partners were notified and 9 received treatment.
Sensitivity analysis
We observed that pooled sensitivity and specificity of CT tests and sensitivity of GC tests conducted in LMICs were significantly lower than those from HICs (online supplemental table 4). In addition, when restricting analysis to HICs only, the sensitivity of vaginal SS increased to 86.0% (76.6%–92.1%) for GC and 86.2% (76.0%–92.5%) for CT (online supplemental table 4). Excluding low-quality studies (n=6) resulted in higher pooled sensitivity and specificity than those obtained above (table 2) with statistically significant differences by sampling site but not by study population (online supplemental table 5).

DISCUSSION
Summary of results
Results from our meta-analysis and bivariate regressions indicated that CT/GC diagnostic accuracy values were higher for extragenital SS (ie, pharyngeal and rectal) than for vaginal SS, and that sensitivities for GC were higher for MSM than for women. However, studies conducted in LMICs may have decreased overall sensitivity by sampling site and study population because they reported lower sensitivity compared with those in HICs and excluded extragenital sampling and MSM populations. Reasons for this lower sensitivity may have been due to SS kit type, quality, storage conditions or lack of instructions on proper conduct of vaginal SS. These issues highlight the need for research on SS technologies and integrated digital innovations to improve sensitivity. Specificity was consistently higher than sensitivity, indicating a higher probability of FN than FP results.

Our results are in accordance with a meta-analysis conducted by Zhou et al on POCTs (NAATs or antigen-detection tests) for CT screening that excluded POCT performance with SS. Our results complement a meta-analysis comparing SS with clinician sampling for CT/GC conducted by Lunny et al 8 years ago. However, including data published from 2015 to 2022 allowed inclusion of extragenital sampling in our analysis with stratified results by study population and country type. We found a lower vaginal SS pooled sensitivity for CT (79% overall, 86.0% for high-quality studies only) compared with this previous study (92%). If truly lower in LMICs, test sensitivity could lead to increased probability of FN individuals (2 in 10), which could delay their access to treatment and increase their likelihood of transmitting CT/GC.

Pooled sensitivities and specificities of pharyngeal and rectal sampling reached the WHO-recommended target product profile for CT (sensitivity=90.0% and specificity=98.0%) and GC (sensitivity=90.0% and specificity=90.0%), but caution must be employed when ruling pharyngeal and rectal SS as most accurate due to low study number (n=3), with none being from LMICs.5 We conclude that extragenital SS followed by NAAT for CT/GC detection is as accurate as conventional testing regardless of sampling site for MSM and women in HICs. However, for LMICs, studies are needed to understand the source and variability in sensitivity of vaginal compared with extragenital SS.

Acceptability for SS was very high (80.0%–100%) whether digital innovations were present or not, but did not always translate to patient preference for SS. Studies with lowest preference values (23.1%–50.9%) were conducted in women, MSM and PLWHIV from LMICs, and women, juvenile correction detainees and university students from one HIC.24 34 42 44 52 58 59  It should be noted that patient preference was sometimes dependent on the number of options provided. The preference for each option was reduced when given the choice between three (SS, clinician sampling or no
Furthermore, linkages to care were high overall.\textsuperscript{17,44,50} Higher percentages of repeat testers through internet-based testing were also identified as concerns.\textsuperscript{21,28,32,42,51,52,55} Cost was a major deterrent in studies reporting preference for those facing barriers to access conventional testing.\textsuperscript{20,21,24,34,44,50} and absence of a healthcare professional for test result interpretation were also identified as concerns.\textsuperscript{21}

Digital innovations improved participation and engagement via text-based reminders or website-based questionnaires.\textsuperscript{17–19,21,23,25} Higher percentages of repeat testers through internet-based SS tests and HIV ST with digital innovations have been reported previously.\textsuperscript{6,41}

Finally, impact outcomes are important to consider to ensure greater proportions of people remaining in the STI care cascade.\textsuperscript{6,61} SS was instrumental in reaching first-time testers in key populations, increasing detection and expanding access for those facing barriers to access conventional testing.\textsuperscript{20,21,24,34,44,50}

Furthermore, linkages to care were high overall.\textsuperscript{17,44,50}

**Strengths and limitations**

To our knowledge, this is the first comprehensive systematic review of outcomes beyond diagnostic accuracy (ie, acceptability, preference, feasibility and impact) for all key populations in LMICs who desperately need CT/GC SS services.\textsuperscript{1} This review offers valuable data on effective strategies to advance the WHO’s STI elimination strategy by 2030.\textsuperscript{1} This first meta-analysis includes CT/GC studies with SS compared with clinician-collected sampling. This review offers insights on how to maximise the use of SS coupled with digital innovations that are needed in any pandemic context by offering alternatives to clinic-based testing and limiting physical contact.

Studies evaluating accuracy of SS in populations other than MSM or women, and those that included digital innovations were limited, thereby preventing generalisation. Fewer studies (11 of 45; 24%) were conducted in LMICs despite their higher global STI disease burden.\textsuperscript{3} Use of different references than clinician-collected samples for urine SS constitutes a limitation.

The quality of studies also varied. Lack of data on non-respondents introduced information bias. Lack of confounding data (eg, age, sex, socioeconomic status) limited our ability to assess it. More than half (10 of 16) of the diagnostic accuracy studies did not report test blinding. High-quality RCTs were also limited.\textsuperscript{42,57} Finally, we could not exclude publication bias except for studies with accuracy estimates.

**Implications**

SS methods maintained comparable accuracy with clinician-collected SS samples, with high acceptability, high proportions of first-time testers and high linkages to care.\textsuperscript{7,8,63,64} Thus, these tests will contribute to decrease the disease burden of untreated CT/GC infections. Cost remains a hurdle for implementation of innovative digital diagnostic tools, especially in LMICs. Furthermore, implementing these methods based on characteristics of target populations requires understanding of their preferences. Future direction for research and policy requires consideration of customising SS strategies to specific subpopulations of interest, especially in LMICs where evidence remains thin.

Our results provide support to the WHO’s call to implement innovative STI diagnostic strategies including digital innovations integrated with SS.\textsuperscript{14} They can be accessed by traditionally hard-to-reach populations and address inequities in care exacerbated by the COVID-19 pandemic. In the context of digitisation of healthcare, integrated digital innovations that offer promise to improve SS procedures, reporting, access and linkage could be useful. However, implementation research in robust cohort, quasi-randomised trials or RCTs is warranted to determine their effectiveness.

**CONCLUSION**

Our analysis reveals that high accuracy of SS followed by NAATs was comparable with clinician-sampled tests for extragenital sampling sites, regardless of populations in HICs (ie, women or MSM), and could consequently be recommended for implementation to diagnose CT/GC in similar settings. SS constitutes an innovative opportunity to reduce global STI burden by addressing obstacles associated with conventional testing. SS could be performed by individuals at home or at the clinic, in privacy and anonymity, expanding access to testing. SS, if combined with digital innovations, could potentially enhance access, simplify conduct, interpretation and linkage for key populations. However, impact and effectiveness need to be proven in studies with robust designs. Finally, our comprehensive review could gain credence if policymakers recommend innovative CT/GC screening strategies for key populations worldwide.
permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**
Fiorella Viallard http://orcid.org/0000-0001-8575-436X
Anna de Waal http://orcid.org/0000-0001-6431-4656
Nitika Pant Pai http://orcid.org/0000-0002-4672-0500

**REFERENCES**

12. Mada: meta-analysis of diagnostic accuracy. R package version 0.5.10 [program]. 0.5.10 version, 2020.
33. Shipitsyna E, Khusnutdinova T, Ryzhkova O, et al. P2.021 prevalence of sexually transmitted infections in young people in St. Petersburg, Russia, as
Systematic review


