



CONSOLIDATED GUIDELINES ON

HIV TESTING SERVICES

2019

Web Annex B. GRADE table: should HIV self-testing be offered as an additional HIV testing approach?

ISBN 978-92-4-001174-8 (electronic version)

Consolidated guidelines on HIV testing services, 2019. Web Annex B. GRADE table: should HIV self-testing be offered as an additional HIV testing approach?

ISBN 978-92-4-001174-8 (electronic version)

This publication was originally published under WHO reference number WHO/UCN/HHS/19.41.

© **World Health Organization 2020**

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo>).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition".

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization (<http://www.wipo.int/amc/en/mediation/rules/>).

Suggested citation. Web Annex B. GRADE table: should HIV self-testing be offered as an additional HIV testing approach? In: Consolidated guidelines on HIV testing services, 2019. Geneva: World Health Organization; 2020. Licence: [CC BY-NC-SA 3.0 IGO](https://creativecommons.org/licenses/by-nc-sa/3.0/igo).

Cataloguing-in-Publication (CIP) data. CIP data are available at <http://apps.who.int/iris>.

Sales, rights and licensing. To purchase WHO publications, see <http://apps.who.int/bookorders>. To submit requests for commercial use and queries on rights and licensing, see <http://www.who.int/about/licensing>.

Third-party materials. If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

This publication forms part of the WHO guideline entitled *Consolidated guidelines on HIV testing services, 2019*. It is being made publicly available for transparency purposes and information, in accordance with the *WHO handbook for guideline development*, 2nd edition (2014).

Contents

Table B1. GRADE table for HIVST compared to SOC for HTS.....	4
Table B2. GRADE table for HIVST + linkage intervention compared to HIVST only for linkage to HIV CARE or ART	33

Table B1. GRADE table for HIVST compared to SOC for HTS

Author(s): Muhammad S. Jamil, T. Charles Witzel, Ingrid Wilson

Question: HIVST compared to SOC for HTS

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
Uptake of HIV testing (overall)												
23 ^a	randomised trials ^b	serious ^c	not serious ^d	not serious ^e	not serious	none	16439/24739 (66.4%) ^f	7775/19260 (40.4%) ^f	RR 1.62 (1.43 to 1.84)	250 more per 1,000 (from 174 more to 339 more)	⊕⊕⊕○ MODERATE	CRITICAL
Uptake of HIV testing (general population)												
13 ^g	randomised trials ^h	very serious ⁱ	not serious ^j	not serious ^k	not serious	none	12870/19308 (66.7%) ^l	5812/14523 (40.0%) ^l	RR 2.09 (1.69 to 2.58)	436 more per 1,000 (from 276 more to 632 more)	⊕⊕○○ LOW	CRITICAL
Uptake of HIV testing (key populations)												
10 ^m	randomised trials ⁿ	serious ^o	not serious ^p	not serious ^q	not serious	none	3569/5431 (65.7%) ^r	1963/4737 (41.4%) ^r	RR 1.28 (1.13 to 1.45)	116 more per 1,000 (from 54 more to 186 more)	⊕⊕⊕○ MODERATE	CRITICAL
Uptake of HIV testing (men who have sex with men)												
7 ^s	randomised trials ^t	serious ^u	not serious ^v	not serious	not serious	none	2427/3404 (71.3%) ^w	1432/3393 (42.2%) ^w	RR 1.37 (1.19 to 1.57)	156 more per 1,000 (from 80 more to 241 more)	⊕⊕⊕○ MODERATE	CRITICAL
Uptake of HIV testing (female sex workers)												
3 ^x	randomised trials ^y	serious ^z	not serious ^{aa}	not serious	not serious	none	1142/2027 (56.3%) ^{ab}	531/1344 (39.5%) ^{ab}	RR 1.16 (0.98 to 1.36)	63 more per 1,000 (from 8 fewer to 142 more)	⊕⊕⊕○ MODERATE	CRITICAL

Uptake of HIV testing (men)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
17 ^{ac}	randomised trials ^{ad}	serious ^{ae}	not serious ^{af}	not serious	not serious	none	9308/12954 (71.9%) ^{ag}	3981/9632 (41.3%) ^{ag}	RR 1.81 (1.56 to 2.11)	335 more per 1,000 (from 231 more to 459 more)	⊕⊕⊕○ MODERATE	CRITICAL

Uptake of HIV testing (women)

6 ^{ah}	randomised trials ^{ai}	serious ^{aj}	not serious ^{ak}	not serious	not serious	none	2772/4415 (62.8%) ^{al}	1578/3820 (41.3%) ^{al}	RR 1.45 (1.13 to 1.86)	186 more per 1,000 (from 54 more to 355 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	---	------------------	----------

Uptake of HIV testing (young people 15-24 years)

5 ^{am}	randomised trials ^{an}	very serious ^{ao}	not serious ^{ap}	not serious	not serious	none	1006/2355 (42.7%) ^{aq}	700/2057 (34.0%) ^{aq}	RR 2.10 (1.36 to 3.23)	374 more per 1,000 (from 123 more to 759 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	---------------------------------	--------------------------------	----------------------------------	--	-------------	----------

Uptake of HIV testing (measurement time-point: ≤6 months)

17 ^{ar}	randomised trials ^{as}	serious ^{at}	not serious ^{au}	not serious	not serious	none	11814/17488 (67.6%) ^{av}	4058/12267 (33.1%) ^{av}	RR 1.79 (1.51 to 2.13)	261 more per 1,000 (from 169 more to 374 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-----------------------------------	----------------------------------	----------------------------------	--	------------------	----------

Uptake of HIV testing (measurement time-point: >6 months)

6 ^{aw}	randomised trials ^{ax}	very serious ^{ay}	not serious ^{az}	not serious	not serious	none	4625/7251 (63.8%) ^{ba}	3717/6993 (53.2%) ^{ba}	RR 1.27 (1.10 to 1.48)	144 more per 1,000 (from 53 more to 255 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	---	-------------	----------

Uptake of HIV testing (online and mail distribution)

5 ^{bb}	randomised trials ^{bc}	serious ^{bd}	not serious ^{be}	not serious	not serious	none	2161/3106 (69.6%) ^{bf}	1218/3099 (39.3%) ^{bf}	RR 1.47 (1.30 to 1.66)	185 more per 1,000 (from 118 more to 259 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	--	------------------	----------

Uptake of HIV testing (facility-based distribution)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
6 ^{bg}	randomised trials ^{bh}	serious ^{bi}	not serious ^{bj}	not serious	not serious	none	903/1155 (78.2%) ^{bk}	520/810 (64.2%) ^{bk}	RR 1.28 (1.01 to 1.61)	180 more per 1,000 (from 6 more to 392 more)	⊕⊕⊕○ MODERATE	CRITICAL

Uptake of HIV testing (secondary distribution - women to male partners)

4 ^{bl}	randomised trials ^{bm}	serious ^{bn}	not serious ^{bo}	not serious	not serious	none	4477/5737 (78.0%) ^{bp}	840/2578 (32.6%) ^{bp}	RR 2.63 (1.81 to 3.82)	531 more per 1,000 (from 264 more to 919 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	---------------------------------	--------------------------------	----------------------------------	--	------------------	----------

Uptake of HIV testing services (secondary distribution - HIV-positive clients to partners)

2 ^{bq}	randomised trials ^{br}	very serious ^{bs}	not serious ^{bt}	not serious	not serious	none	507/823 (61.6%) ^{bu}	120/369 (32.5%) ^{bu}	RR 2.03 (1.01 to 4.09)	335 more per 1,000 (from 3 more to 1,000 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	-------------------------------	-------------------------------	----------------------------------	--	-------------	----------

Uptake of HIV testing (secondary distribution - peers)

2 ^{bv}	randomised trials ^{bw}	serious ^{bx}	not serious ^{by}	not serious	not serious	none	555/612 (90.7%) ^{bz}	244/324 (75.3%) ^{bz}	RR 1.19 (0.97 to 1.47)	143 more per 1,000 (from 23 fewer to 354 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	-------------------------------	----------------------------------	--	------------------	----------

Uptake of HIV testing (community- or home-based distribution)

3 ^{ca}	randomised trials ^{cb}	very serious ^{cc}	not serious ^{cd}	not serious	not serious	none	6530/9559 (68.3%) ^{ce}	4419/8516 (51.9%) ^{ce}	RR 1.43 (0.95 to 2.13)	223 more per 1,000 (from 26 fewer to 586 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	--	-------------	----------

Uptake of HIV testing (HIVST at facilities)

4 ^{cf}	randomised trials ^{cg}	very serious ^{ch}	not serious ^{ci}	not serious	not serious	none	1344/3747 (35.9%) ^{cj}	414/3564 (11.6%) ^{cj}	RR 2.14 (1.22 to 3.74)	132 more per 1,000 (from 26 more to 318 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	---------------------------------	--------------------------------	----------------------------------	---	-------------	----------

Uptake of HIV testing (no or basic support)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
4 ^{ck}	randomised trials ^{cl}	very serious ^{cm}	not serious ^{cn}	not serious	not serious	none	268/377 (71.1%) ^{co}	157/355 (44.2%) ^{co}	RR 1.60 (1.13 to 2.28)	265 more per 1,000 (from 57 more to 566 more)	⊕⊕○○ LOW	CRITICAL

Uptake of HIV testing (IFU enhancement, video or study hotline)

8 ^{cp}	randomised trials ^{cq}	serious ^{cr}	not serious ^{cs}	not serious	not serious	none	6985/9448 (73.9%) ^{ct}	2096/5848 (35.8%) ^{ct}	RR 2.02 (1.65 to 2.47)	366 more per 1,000 (from 233 more to 527 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	--	------------------	----------

Uptake of HIV testing (group demonstration)

3 ^{cu}	randomised trials ^{cv}	serious ^{cw}	serious ^{cx}	not serious	not serious	none	2124/3374 (63.0%) ^{cy}	736/2599 (28.3%) ^{cy}	RR 1.26 (1.01 to 1.58)	74 more per 1,000 (from 3 more to 164 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	-----------------------	-----------------------	-------------	-------------	------	---------------------------------	--------------------------------	----------------------------------	---	-------------	----------

Uptake of HIV testing (in-person demonstration or training)

4 ^{cz}	randomised trials ^{da}	very serious ^{db}	not serious ^{dc}	not serious	not serious	none	6626/9675 (68.5%) ^{dd}	4511/8630 (52.3%) ^{dd}	RR 1.31 (0.93 to 1.86)	162 more per 1,000 (from 37 fewer to 450 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	---------------------------------	---------------------------------	----------------------------------	--	-------------	----------

Uptake of HIV testing (in-person observation or supervision)

3 ^{de}	randomised trials ^{df}	serious ^{dg}	not serious ^{dh}	not serious	not serious	none	1650/243 (679.0%) ^{di}	166/1613 (10.3%) ^{di}	RR 1.69 (1.07 to 2.67)	71 more per 1,000 (from 7 more to 172 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	---------------------------------	--------------------------------	----------------------------------	---	------------------	----------

Uptake of HIV testing (virtual real-time support or supervision)

1	randomised trials	serious ^{dj}	serious ^{dk}	not serious	not serious	none	193/215 (89.8%) ^{dl}	109/215 (50.7%) ^{dl}	RR 1.77 (1.54 to 2.04)	390 more per 1,000 (from 274 more to 527 more)	⊕⊕○○ LOW	CRITICAL
---	-------------------	-----------------------	-----------------------	-------------	-------------	------	-------------------------------	-------------------------------	----------------------------------	--	-------------	----------

Certainty assessment							N° of patients		Effect		Certainty	Importance
N° of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		

Frequency of HIV testing (men who have sex with men)

3 ^{dm}	randomised trials ^{dn}	serious ^{do}	not serious ^{dp}	not serious	not serious	none	Among 3 trials 1242 participants in the HIVST arm reported 6351 HIV tests over 12 - 15 months compared to 1222 participants reporting 2107 HIV tests in the SOC arm during the same period. This represents 2.6 additional tests per participant during follow-up period (mean difference: 2.6, 95% CI: 1.2 - 4.0)		⊕⊕⊕○	MODERATE	IMPORTANT
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--	--	------	----------	-----------

HIV positivity rate among randomized (overall)

16 ^{dq}	randomised trials ^{dr}	serious ^{ds}	not serious ^{dt}	not serious	not serious	none	388/16636 (2.3%) ^{du}	143/10330 (1.4%) ^{du}	RR 1.06 (0.76 to 1.48)	1 more per 1,000 (from 3 fewer to 7 more)	⊕⊕⊕○	MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	--------------------------------	----------------------------------	---	------	----------	----------

HIV positivity rate among tested (overall)

17 ^{dv}	randomised trials ^{dw}	serious ^{dx}	not serious ^{dy}	not serious	not serious	none	444/13824 (3.2%) ^{dz}	357/9689 (3.7%) ^{dz}	RR 0.97 (0.74 to 1.27) ^{ea}	1 fewer per 1,000 (from 10 fewer to 10 more)	⊕⊕⊕○	MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	-------------------------------	--	--	------	----------	----------

HIV positivity rate among tested (general population)

8 ^{eb}	randomised trials ^{ec}	serious ^{ed}	not serious ^{ee}	not serious	not serious	none	227/11088 (2.0%) ^{ef}	243/7932 (3.1%) ^{ef}	RR 0.81 (0.45 to 1.47)	6 fewer per 1,000 (from 17 fewer to 14 more)	⊕⊕⊕○	MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	-------------------------------	----------------------------------	--	------	----------	----------

HIV positivity rate among tested (key populations)

9 ^{eg}	randomised trials ^{eh}	serious ^{ei}	not serious ^{ej}	not serious ^{ek}	not serious	none	217/2736 (7.9%) ^{el}	114/1757 (6.5%) ^{el}	RR 0.99 (0.75 to 1.30)	1 fewer per 1,000 (from 16 fewer to 19 more)	⊕⊕⊕○	MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	---------------------------	-------------	------	-------------------------------	-------------------------------	----------------------------------	--	------	----------	----------

HIV positivity rate among tested (men who have sex with men)

6 ^{em}	randomised trials ^{en}	serious ^{ed}	not serious ^{eo}	not serious	not serious	none	34/1538 (2.2%) ^{ep}	14/1152 (1.2%) ^{ep}	RR 1.59 (0.87 to 2.89)	7 more per 1,000 (from 2 fewer to 23 more)	⊕⊕⊕○	MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	------------------------------	----------------------------------	--	------	----------	----------

HIV positivity rate among tested (female sex workers)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
3 ^{eq}	randomised trials ^{er}	serious ^{es}	not serious ^{et}	not serious	not serious	none	183/1198 (15.3%) ^{eu}	100/605 (16.5%) ^{eu}	RR 0.87 (0.63 to 1.21)	21 fewer per 1,000 (from 61 fewer to 35 more)	⊕⊕⊕○ MODERATE	CRITICAL

HIV positivity rate among tested (men)

13 ^{ev}	randomised trials ^{ew}	serious ^{ex}	not serious ^{ey}	not serious	not serious	none	103/3675 (2.8%) ^{ez}	28/1647 (1.7%) ^{ez}	RR 1.13 (0.74 to 1.72)	2 more per 1,000 (from 4 fewer to 12 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	------------------------------	----------------------------------	--	------------------	----------

HIV positivity rate among tested (women)

4 ^{fa}	randomised trials ^{fb}	serious ^{fc}	not serious ^{fd}	not serious	not serious	none	193/1648 (11.7%) ^{fe}	102/710 (14.4%) ^{fe}	RR 0.88 (0.65 to 1.19)	17 fewer per 1,000 (from 50 fewer to 27 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	-------------------------------	----------------------------------	---	------------------	----------

HIV positivity rate among tested (young people 15-24 years)

4 ^{ff}	randomised trials ^{fg}	very serious ^{fh}	not serious ^{fi}	not serious	not serious	none	93/6011 (1.5%) ^{fj}	220/6973 (3.2%) ^{fj}	RR 0.67 (0.33 to 1.34)	10 fewer per 1,000 (from 21 fewer to 11 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	------------------------------	-------------------------------	----------------------------------	---	-------------	----------

HIV positivity rate among tested (measurement time-point: ≤6 months)

13 ^{fk}	randomised trials ^{fl}	serious ^{fm}	not serious ^{fn}	not serious	not serious	none	412/12577 (3.3%) ^{fo}	344/8833 (3.9%) ^{fo}	RR 0.91 (0.70 to 1.20)	4 fewer per 1,000 (from 12 fewer to 8 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	-------------------------------	----------------------------------	---	------------------	----------

HIV positivity rate among tested (measurement time-point: >6 months)

4 ^{fp}	randomised trials ^{fq}	serious ^{fr}	not serious ^{fs}	not serious	not serious	none	32/1247 (2.6%) ^{ft}	13/856 (1.5%) ^{ft}	RR 1.60 (0.85 to 2.98)	9 more per 1,000 (from 2 fewer to 30 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	-----------------------------	----------------------------------	--	------------------	----------

HIV positivity rate among tested (online and mail distribution)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
4 ^{fu}	randomised trials ^{fv}	very serious ^{fw}	not serious ^{fx}	not serious	not serious	none	27/1270 (2.1%) ^{fy}	12/931 (1.3%) ^{fy}	RR 1.45 (0.75 to 2.78)	6 more per 1,000 (from 3 fewer to 23 more)	⊕⊕○○ LOW	CRITICAL

HIV positivity rate among tested (facility-based distribution)

4 ^{fz}	randomised trials ^{ga}	serious ^{gb}	not serious ^{gc}	not serious	not serious	none	97/828 (11.7%) ^{gd}	52/502 (10.4%) ^{gd}	RR 1.02 (0.57 to 1.85)	2 more per 1,000 (from 45 fewer to 88 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	------------------------------	----------------------------------	---	------------------	----------

HIV positivity rate among tested (secondary distribution - women to male partners)

3 ^{ge}	randomised trials ^{gf}	serious ^{gg}	not serious ^{gh}	not serious	not serious	none	65/4155 (1.6%) ^{gi}	8/734 (1.1%) ^{gi}	RR 0.58 (0.18 to 1.87)	5 fewer per 1,000 (from 9 fewer to 9 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	----------------------------	----------------------------------	--	------------------	----------

HIV positivity rate among tested (secondary distribution - HIV-positive clients to partners)

2 ^{gj}	randomised trials ^{gk}	very serious ^{gl}	not serious ^{gm}	not serious	not serious	none	74/360 (20.6%) ^{gn}	13/99 (13.1%) ^{gn}	RR 1.42 (0.74 to 2.71)	55 more per 1,000 (from 34 fewer to 225 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	------------------------------	-----------------------------	----------------------------------	---	-------------	----------

HIV positivity rate among tested (secondary distribution - peers)

2 ^{go}	randomised trials ^{gp}	serious ^{gq}	not serious ^{gr}	not serious	not serious	none	88/567 (15.5%) ^{gs}	50/281 (17.8%) ^{gs}	RR 0.85 (0.55 to 1.32)	27 fewer per 1,000 (from 80 fewer to 57 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	------------------------------	----------------------------------	---	------------------	----------

HIV positivity rate among tested (community- or home-based distribution)

1	randomised trials ^{gt}	serious ^{gu}	serious ^{gv}	not serious	not serious	none	56/5353 (1.0%) ^{gw}	214/6728 (3.2%) ^{gw}	RR 0.64 (0.18 to 2.22)	11 fewer per 1,000 (from 26 fewer to 39 more)	⊕⊕○○ LOW	CRITICAL
---	---------------------------------	-----------------------	-----------------------	-------------	-------------	------	------------------------------	-------------------------------	----------------------------------	---	-------------	----------

HIV positivity rate among tested (HIVST at facilities)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
4 ^{gx}	randomised trials ^{gy}	very serious ^{gz}	not serious ^{ha}	not serious	serious ^{hb}	none	37/1291 (2.9%) ^{hc}	8/414 (1.9%) ^{hc}	RR 1.18 (0.29 to 4.74)	3 more per 1,000 (from 14 fewer to 72 more)	⊕○○○ VERY LOW	CRITICAL

Linkage to additional or confirmatory HIV testing (HIVST arm only)

7 ^{hd}	randomised trials ^{he}	serious ^{hf}	serious ^{hg}	not serious	not serious	none	A pooled estimated from 7 studies showed 65% (52% - 78%) of those with a reactive HIVST result (n=497) had confirmatory HIV testing (n=332).				⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	-----------------------	-----------------------	-------------	-------------	------	--	--	--	--	-------------	----------

Linkage to ART initiation or HIV care among randomized (Overall)

12 ^{hh}	randomised trials ^{hi}	serious ^{hj}	not serious ^{hk}	not serious	not serious	none	217/12440 (1.7%) ^{hl}	79/8013 (1.0%) ^{hl}	RR 1.07 (0.67 to 1.72)	1 more per 1,000 (from 3 fewer to 7 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	--------------------------------	------------------------------	----------------------------------	---	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (Overall)

11 ^{hm}	randomised trials ^{hn}	serious ^{ho}	not serious ^{hp}	not serious	not serious	none	217/403 (53.8%) ^{hq}	79/149 (53.0%) ^{hq}	RR 0.94 (0.81 to 1.09)	32 fewer per 1,000 (from 101 fewer to 48 more)	⊕⊕⊕○ MODERATE	CRITICAL
------------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	------------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (general population)

6 ^{hr}	randomised trials ^{hs}	serious ^{ht}	not serious ^{hu}	not serious	not serious	none	166/217 (76.5%) ^{hv}	37/47 (78.7%) ^{hv}	RR 0.95 (0.81 to 1.09)	39 fewer per 1,000 (from 150 fewer to 71 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	-----------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (key populations)

5 ^{hw}	randomised trials ^{hx}	serious ^{hy}	not serious ^{hz}	not serious ^{ia}	not serious	none	51/186 (27.4%) ^{ib}	42/102 (41.2%) ^{ib}	RR 0.83 (0.58 to 1.18)	70 fewer per 1,000 (from 173 fewer to 74 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	---------------------------	-------------	------	------------------------------	------------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (men who have sex with men)

Certainty assessment							N° of patients		Effect		Certainty	Importance
N° of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials ^{ic}	serious ^{id}	not serious ^{ie}	not serious	not serious	none	8/8 (100.0%) ^{if}	2/2 (100.0%) ^{if}	RR 1.14 (0.67 to 1.97)	140 more per 1,000 (from 330 fewer to 970 more)	⊕⊕⊕○ MODERATE	CRITICAL

Linkage to ART initiation or HIV care among HIV positive (female sex workers)

2 ^{ig}	randomised trials ^{ih}	serious ⁱⁱ	not serious ^{ij}	not serious	not serious	none	43/178 (24.2%) ^{ik}	40/100 (40.0%) ^{ik}	RR 0.66 (0.41 to 1.04)	136 fewer per 1,000 (from 236 fewer to 16 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	------------------------------	----------------------------------	---	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (men)

8 ^{il}	randomised trials ^{im}	serious ⁱⁿ	not serious ^{io}	not serious	not serious	none	83/115 (72.2%) ^{ip}	15/17 (88.2%) ^{ip}	RR 0.88 (0.65 to 1.20)	106 fewer per 1,000 (from 309 fewer to 176 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	-----------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (women)

4 ^{iq}	randomised trials ^{ir}	serious ^{is}	not serious ^{it}	not serious	not serious	none	52/189 (27.5%) ^{iu}	41/100 (41.0%) ^{iu}	RR 0.72 (0.47 to 1.11)	115 fewer per 1,000 (from 217 fewer to 45 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	------------------------------	------------------------------	----------------------------------	---	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (young people 15-24 years)

2 ^{iv}	randomised trials ^{iw}	very serious ^{ix}	not serious ^{iy}	not serious	not serious	none	4/10 (40.0%) ^{iz}	1/1 (100.0%) ^{iz}	RR 0.59 (0.21 to 1.65)	410 fewer per 1,000 (from 790 fewer to 650 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	----------------------------	---------------------------	-------------	-------------	------	----------------------------	----------------------------	----------------------------------	--	-------------	----------

Linkage to ART initiation or HIV care among HIV positive (measurement time-point: ≤6 months)

8 ^{ja}	randomised trials ^{jb}	serious ^{jc}	not serious ^{jd}	not serious	not serious	none	177/353 (50.1%) ^{je}	64/127 (50.4%) ^{je}	RR 0.88 (0.71 to 1.09)	60 fewer per 1,000 (from 146 fewer to 45 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	------------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (measurement time-point: >6 months)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
3 ^{jf}	randomised trials ^{jj}	serious ^{jh}	not serious ^{ji}	not serious	not serious	none	40/50 (80.0%) ^{jj}	15/22 (68.2%) ^{jj}	RR 0.98 (0.79 to 1.22)	14 fewer per 1,000 (from 143 fewer to 150 more)	⊕⊕⊕○ MODERATE	CRITICAL

Linkage to ART initiation or HIV care among HIV positive (no linkage support)

9 ^{jk}	randomised trials ^{jl}	serious ^{jm}	not serious ^{jn}	not serious	not serious	none	101/274 (36.9%) ^{jo}	59/122 (48.4%) ^{jo}	RR 0.82 (0.62 to 1.07)	87 fewer per 1,000 (from 184 fewer to 34 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-------------------------------	------------------------------	----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (phone reminder or follow up)

1 ^{jp}	randomised trials ^{jq}	serious ^{jr}	serious ^{js}	not serious	not serious	none	2/3 (66.7%) ^{jt}	1/1 (100.0%) ^{jt}	RR 0.83 (0.28 to 2.51)	170 fewer per 1,000 (from 720 fewer to 1,000 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	-----------------------	-----------------------	-------------	-------------	------	---------------------------	----------------------------	----------------------------------	--	-------------	----------

Linkage to ART initiation or HIV care among HIV positive (home visit or in-person referral)

1 ^{ju}	randomised trials ^{jv}	serious ^{jw}	serious ^{jx}	not serious	not serious	none	33/43 (76.7%) ^{jy}	13/20 (65.0%) ^{jy}	RR 0.96 (0.76 to 1.21)	26 fewer per 1,000 (from 156 fewer to 136 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	-----------------------	-----------------------	-------------	-------------	------	-----------------------------	-----------------------------	----------------------------------	---	-------------	----------

Linkage to ART initiation or HIV care among HIV positive (virtual real-time support)

1	randomised trials ^{jz}	serious ^{ka}	serious ^{kb}	not serious	serious ^{kc}	none	1/1 (100.0%) ^{kd}	0/0	RR 1.50 (0.18 to 12.46)	2 fewer per 1,000 (from 12 fewer to 0 fewer)	⊕○○○ VERY LOW	CRITICAL
---	---------------------------------	-----------------------	-----------------------	-------------	-----------------------	------	----------------------------	-----	-----------------------------------	--	------------------	----------

Linkage to ART initiation or HIV care among HIV positive (financial incentive)

2 ^{ke}	randomised trials ^{kf}	serious ^{kg}	not serious ^{kh}	not serious	not serious	none	80/82 (97.6%) ^{ki}	6/6 (100.0%) ^{ki}	RR 1.09 (0.87 to 1.38)	90 more per 1,000 (from 130 fewer to 380 more)	⊕⊕⊕○ MODERATE	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	-------------	-------------	------	-----------------------------	----------------------------	----------------------------------	--	------------------	----------

Misuse of HIVST kits - coercion (HIVST arm only)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST	SOC	Relative (95% CI)	Absolute (95% CI)		
2 ^{kj}	randomised trials ^{kk}	serious ^{kl}	not serious ^{km}	not serious ^{kn}	serious ^{ko}	none	One trial reported 4 instances of coercion or forced to test among 13267 participants. The other trial reported 0 instances of coercion to test or disclose results among 1063 participants.				⊕⊕○○ LOW	CRITICAL

Adverse events among those randomized

6 ^{kp}	randomised trials ^{kq}	serious ^{kr}	not serious ^{ks}	not serious ^{kt}	serious ^{ku}	none	12/5502 (0.2%) ^{kv}	1/3124 (0.0%) ^{kv}	RR 1.89 (0.54 to 6.54)	0 fewer per 1,000 (from 0 fewer to 2 more)	⊕⊕○○ LOW	CRITICAL
-----------------	---------------------------------	-----------------------	---------------------------	---------------------------	-----------------------	------	------------------------------	-----------------------------	----------------------------------	--	-------------	----------

Condomless anal sex with male partners among those randomized

4 ^{kw}	randomised trials ^{kx}	serious ^{ky}	not serious ^{kz}	not serious ^{la}	not serious	none	416/1544 (26.9%) ^{lb}	528/1890 (27.9%) ^{lb}	RR 1.09 (0.95 to 1.24)	25 more per 1,000 (from 14 fewer to 67 more)	⊕⊕⊕○ MODERATE	IMPORTANT
-----------------	---------------------------------	-----------------------	---------------------------	---------------------------	-------------	------	--------------------------------	--------------------------------	----------------------------------	--	------------------	-----------

Condomless anal sex with male partners among those who completed follow-up

4 ^{lc}	randomised trials ^{ld}	serious ^{le}	not serious ^{lf}	not serious ^{lg}	not serious	none	416/1482 (28.1%) ^{lh}	528/1747 (30.2%) ^{lh}	RR 1.02 (0.91 to 1.15)	6 more per 1,000 (from 27 fewer to 45 more)	⊕⊕⊕○ MODERATE	IMPORTANT
-----------------	---------------------------------	-----------------------	---------------------------	---------------------------	-------------	------	--------------------------------	--------------------------------	----------------------------------	---	------------------	-----------

CI: Confidence interval; RR: Risk ratio

Explanations

a. Meta-analysis: Choko, 2019a and Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting); Choko, 2019b had two study groups (ANC women and ART clients) and these were presented separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

b. 14 individual randomized trials, 9 cluster randomized trials.

c. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 14 trials and attrition bias in 3 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm; Patel, 2018: 36% LTFU overall, 44% in the intervention and 27% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

- d. There was high statistical heterogeneity (Heterogeneity: $Tau^2 = 0.091$; $Chi^2 = 409.12$, $df = 25$, $p < 0.01$; $I^2 = 94\%$, 92% - 95%). Sub-group analyses by population, study design, measure time-point, and distribution method did not fully explain heterogeneity. Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.
- e. We did not downgrade for indirectness.
- f. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- g. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined. Choko, 2019b had two study groups (ANC women and ART clients) and these were presented separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- h. 7 individual randomized trials and 6 cluster randomized trials
- i. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 10 trials and attrition bias in one trial (Patel, 2018: 36% LTFU overall, 44% in the intervention and 27% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts. 9 of 13 trials had more than three high risk or unclear risk of bias domains.
- j. There was a high statistical heterogeneity (Heterogeneity: $Tau^2 = 0.133$; $Chi^2 = 213.31$, $df = 13$, $p < 0.01$; $I^2 = 94\%$, 91% - 96%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.
- k. We did not downgrade for indirectness but noted that all but one trial were conducted in Africa (6 in Malawi, 4 in Kenya, one in Zambia, one in South Africa, one in the US). This is expected most countries with generalized epidemics are in Africa.
- l. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- m. Meta-analysis: Wray, 2018 had two intervention arms which both included HIVST with similar HIVST kit distribution method, therefore both arms were combined. Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- n. 7 individual randomized trials and 3 cluster randomized trials
- o. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 5 trials and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- p. There was high statistical heterogeneity (Heterogeneity: $Tau^2 = 0.040$; $Chi^2 = 108.69$, $df = 11$, $p < 0.01$; $I^2 = 90\%$, 846% - 93%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.
- q. We did not downgrade but noted that only men who have sex with men, transgender women and female sex workers were represented in included studies and results should be viewed with caution when applying to other key populations.
- r. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- s. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. in these trials. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- t. 6 individual randomized trials and one cluster randomized trial.

u. We downgraded once. due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 trials and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

v. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.028$; $\text{Chi}^2 = 45.50$, $\text{df} = 6$, $p < 0.01$; $I^2 = 87\%$, 75% - 93%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

w. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

x. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

y. One individual randomized trial and 2 cluster randomized trials.

z. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 2 trials and unclear risk in one trial. One trial had unclear risk associated with random sequence generation.

aa. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.052$; $\text{Chi}^2 = 39.90$, $\text{df} = 4$, $p < 0.01$; $I^2 = 90\%$, 79% - 95%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

ab. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ac. 7 trials were conducted among MSM, 4 were among male partners of ANC women, 2 were among male truck drivers, and for the remaining sub-analysis among men was included. Meta-analysis: Choko, 2019a and Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms and presented ANC data only. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ad. 11 individual randomized trials and 6 cluster randomized trials.

ae. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 11 trials and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

af. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.084$; $\text{Chi}^2 = 240.11$, $\text{df} = 17$, $p < 0.01$; $I^2 = 93\%$, 91% - 95%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

ag. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ah. Three trials were conducted among female sex workers, one was among young women, and for two trials sub-analysis among women was included. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ai. 2 individual randomized trials and 4 cluster randomized trials.

aj. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 3 trials or unclear detection bias in 2 trials. One cluster randomized trial was subject to recruitment bias. Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

ak. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.110$; $\text{Chi}^2 = 62.95$, $\text{df} = 7$, $p < 0.01$; $I^2 = 89\%$, 80% - 94%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

al. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

am. Age groups 15-24 years were included. One trial was conducted among young women, one among young MSM, and for three sub-analysis among young people was included. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

an. 2 individual randomized trials and 3 cluster randomized trials.

ao. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 4 trials or unclear risk of detection bias in one trial, and attrition bias in one trial (Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts. Each of the trials had more than three high risk or unclear risk of bias domains.

ap. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.230$; $\text{Chi}^2 = 82.25$, $\text{df} = 4$, $p < 0.01$; $I^2 = 95\%$, 91% - 97%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

aq. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ar. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting); Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

as. 10 individual randomized trials and 7 cluster randomized trials.

at. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 11 trials or unclear risk of detection bias in 3 trials, and attrition bias in 2 trials (Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm; Patel, 2018: 36% LTFU overall, 44% in the intervention and 27% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

au. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.131$; $\text{Chi}^2 = 351.02$, $\text{df} = 20$, $p < 0.01$; $I^2 = 95\%$, 93% - 96%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

av. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

aw. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ax. 4 individual randomized trials and 2 cluster randomized trials.

ay. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 4 trials or unclear risk of detection bias in one trial, and attrition bias in one trial (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm). Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts. 4 of 6 trials had more than three high risk or unclear risk of bias domains.

az. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.029$; $\text{Chi}^2 = 42.71$, $\text{df} = 5$, $p < 0.01$; $I^2 = 88\%$, 77% - 94%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

ba. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

bb. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

bc. 4 individual randomized trials and one cluster randomized trial.

bd. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 trials and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

be. There was high moderate statistical heterogeneity (Heterogeneity: $Tau^2 = 0.012$; $Chi^2 = 12.51$, $df = 4$, $p = 0.01$; $I^2 = 68\%$, 17% - 88%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

bf. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

bg. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, in this analysis comparison of one arm (facility-based HIVST distribution) with standard of care was included. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

bh. 4 individual randomized trials and 2 cluster randomized trials.

bi. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 trials or unclear risk for detection bias in 2 trials, and attrition bias in one trial (Patel, 2018: 36% LTFU overall, 44% in the intervention and 27% in the control arm). There was unclear risk of bias for random sequence generation for Patel, 2018 and Pettifor, 2018.

bj. There was high statistical heterogeneity (Heterogeneity: $Tau^2 = 0.068$; $Chi^2 = 60.88$, $df = 5$, $p < 0.01$; $I^2 = 92\%$, 85% - 96%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

bk. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

bl. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Choko, 2019b had two study groups (ANC women and ART clients), only ANC women group was included in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

bm. 2 individual randomized trials and 2 cluster randomized trials.

bn. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in all trials. One trial had unclear risk of attrition bias and two cluster randomized trials were subject to recruitment bias.

bo. There was high statistical heterogeneity (Heterogeneity: $Tau^2 = 0.133$; $Chi^2 = 53.54$, $df = 3$, $p < 0.01$; $I^2 = 94\%$, 89% - 97%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

bp. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

bq. Meta-analysis: Choko, 2019b had two study groups (ANC women and ART clients), but only ART group was included in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

br. One individual randomized trial and one cluster randomized trial.

bs. We downgraded twice. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in both trials. Choko, 2019b cluster randomized trial was subject to recruitment bias. Each of the trials had more than three high risk or unclear risk of bias domains.

bt. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.211$; $\text{Chi}^2 = 5.45$, $\text{df} = 1$, $p = 0.02$; $I^2 = 82\%$). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

bu. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

bv. Meta-analysis: Both trials (Chanda, 2017 and Ortblad, 2017) had more than one intervention arm which included HIVST but HIVST kit distribution method was different, only one arm which involved direct distribution of HIVST kits to peers compared to standard of care is included in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

bw. 2 cluster randomized trials.

bx. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in both trials.

by. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.018$; $\text{Chi}^2 = 5.02$, $\text{df} = 1$, $p = 0.03$; $I^2 = 80\%$). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

bz. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ca. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

cb. 3 cluster randomized trials.

cc. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in three trials. Some other risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts. All trials had more than three high risk or unclear risk of bias domains.

cd. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.122$; $\text{Chi}^2 = 52.78$, $\text{df} = 3$, $p < 0.01$; $I^2 = 96\%$, $92\% - 98\%$). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

ce. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

cf. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

cg. 2 cluster randomized trials.

ch. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in one trial or unclear detection bias in 2 trials. Three of 4 trials had unclear risk of bias for random sequence generation. Three of 4 trials had more than three high risk or unclear risk of bias domains.

ci. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.274$; $\text{Chi}^2 = 29.10$, $\text{df} = 3$, $p < 0.01$; $I^2 = 90\%$, $76\% - 95\%$). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

cj. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ck. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

cl. 4 individual randomized trials.

cm. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 or unclear detection bias in one trial and attrition bias in 2 trials (Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm; Patel, 2018: 36% LTFU overall, 44% in the intervention and 27% in the control arm). Each of the trials had more than three high risk or unclear risk of bias domains.

cn. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.090$; $\text{Chi}^2 = 18.22$, $\text{df} = 3$, $p < 0.01$; $I^2 = 84\%$, 58% - 94%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

co. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

cp. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined. Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

cq. 5 individual randomized trials and 3 cluster randomized trials.

cr. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 6 trials and attrition bias in one trial (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm). Two cluster randomized trials were subject to recruitment bias. Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

cs. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.083$; $\text{Chi}^2 = 126.61$, $\text{df} = 8$, $p < 0.01$; $I^2 = 94\%$, 90% - 96%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

ct. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

cu. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

cv. 3 cluster randomized trials.

cw. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in all trials. One cluster randomized trial was subject to recruitment bias. One cluster randomized trial had unclear risk of bias for selection bias (allocation concealment), selective reporting, and baseline cluster imbalance.

cx. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.054$; $\text{Chi}^2 = 39.74$, $\text{df} = 4$, $p < 0.01$; $I^2 = 90\%$, 79% - 95%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

cy. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

cz. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

da. One individual randomized trial and 3 cluster randomized trials.

db. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 trials or unclear detection bias in one trial. Three of 4 trials had more than three high risk or unclear risk of bias domains.

dc. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.121$; $\text{Chi}^2 = 85.06$, $\text{df} = 3$, $p < 0.01$; $I^2 = 96\%$, 94% - 98%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

dd. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

de. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

df. 3 individual randomized trials.

dg. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and unclear risk for detection bias (self-reported or non-validated outcomes) in 2 trials. Two trials had unclear risk for random sequence generation (selection bias).

dh. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.123$; $\text{Chi}^2 = 10.49$, $\text{df} = 2$, $p < 0.01$; $I^2 = 81\%$, 40% - 94%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

di. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

dj. We downgraded once due to potential for performance bias due to lack of blinding.

dk. Single trial, we downgraded once as inconsistency cannot be evaluated.

dl. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

dm. Three trials were included in meta-analysis (Jamil, 2017; Katz, 2018; MacGowan, 2017). All trials were conducted among men who have sex with men in high income countries (Australian and the United States).

dn. 3 individual randomized trials.

do. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in one trial or unclear risk of detection bias in one trial, and attrition bias in one trial (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm).

dp. There was high statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 1.435$; $\text{Chi}^2 = 63.07$, $\text{df} = 2$, $p < 0.01$; $I^2 = 97\%$, 94% - 98%). Study effects from individual RCTs were consistently beneficial and no difference was observed in other critical outcomes. The GDG determined that downgrading for inconsistency was not necessary. We did not downgrade.

dq. Meta-analysis: Choko, 2019a and Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting); Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

dr. 11 individual randomized and 6 were cluster randomized trials.

ds. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 6 trials or unclear risk of detection bias in 4 trials, and attrition bias in two trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Three cluster randomized trials were subject to potential recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

dt. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.117$; $\text{Chi}^2 = 24.26$, $\text{df} = 18$, $p = 0.15$; $I^2 = 26\%$, 0% - 57%). We did not downgrade but noted that outcome definition was not consistent across trials: 8 trials defined positivity as confirmed HIV diagnosis; the remaining trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

du. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

dv. Meta-analysis: Choko, 2019a and Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting); Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

dw. 11 individual randomized and 6 were cluster randomized trials.

dx. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 6 trials or unclear risk of detection bias in 4 trials, and attrition bias in two trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Three cluster randomized trials were subject to potential recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

dy. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0.048$; $Chi^2 = 22.14$, $df = 19$, $p = 0.28$; $I^2 = 14\%$, $0\% - 49\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 8 trials defined positivity as confirmed HIV diagnosis; the remaining trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear. We also noted that one trial (Nichols, 2019) had significantly lower HIV positivity among those tested in the intervention arm. This trial focused on young people in a setting with high testing coverage.

dz. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ea. RR for HIV positivity among randomized: 1.06 (0.76 - 1.48).

eb. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined; Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ec. 4 individual randomized trials and 4 cluster randomized trials.

ed. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 2 trials and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). One trial had unclear risk of bias for allocation concealment (selection bias).

ee. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0.283$; $Chi^2 = 13.12$, $df = 8$, $p = 0.11$; $I^2 = 39\%$, $0\% - 72\%$). We did not downgrade but noted that the outcome definition was not consistent across trials: 4 of 9 trials defined positivity as confirmed HIV positive diagnosis; the remaining 5 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

ef. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

eg. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

eh. 7 individual randomized trials and 2 cluster randomized trials.

ei. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 4 trials or one unclear risk for detection bias, and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Two trials had unclear risk of selection bias (one for random sequence generation and one for allocation concealment).

ej. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0$; $Chi^2 = 8.93$, $df = 10$, $p = 0.54$; $I^2 = 0\%$, $0\% - 55\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 4 of 9 trials defined positivity as confirmed HIV positive diagnosis; the remaining 5 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

ek. We did not downgrade but noted that only men who have sex with men, transgender women and female sex workers were represented in included studies and results should be viewed with caution when applying to other key populations.

el. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

em. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

en. 6 individual randomized trials.

eo. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 1.62$, $\text{df} = 5$, $p = 0.90$; $I^2 = 0\%$, $0\% - 22\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 4 of 6 trials defined positivity as confirmed HIV positive diagnosis; the remaining 5 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

ep. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

eq. Three 3 arm RCTs contributed to this analysis. Kelvin 2019 included two SoC arms, one enhanced and one standard. Enhanced SoC was excluded and standard SoC used in the comparison with the intervention arm. Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST: intervention arms were not pooled as differences in the intervention package appeared likely to impact positivity.

er. 2 individual randomized and one cluster randomized trials.

es. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 2 trials or unclear risk of detection bias in one trial. One trial had unclear risk of bias for random sequence generation (selection bias).

et. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.008$; $\text{Chi}^2 = 4.24$, $\text{df} = 4$, $p = 0.35$; $I^2 = 6\%$, $0\% - 80\%$). We did not downgrade.

eu. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ev. 6 trials were conducted among MSM, 3 were among male partners of ANC women, 2 were among male truck drivers, and for the remaining sub-analysis among men was included. Meta-analysis: Choko, 2019a and Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined in these trials; Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but included ANC group only in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ew. 10 individual randomized and 3 cluster randomized trials.

ex. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 4 trials or unclear risk of detection bias in 2 trials, and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

ey. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 8.62$, $\text{df} = 12$, $p = 0.73$; $I^2 = 0\%$, $0\% - 40\%$). We did not downgrade but noted that the outcome definition was not consistent across trials: 7 of 13 trials defined positivity as confirmed HIV positive diagnosis; the remaining 7 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

ez. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

fa. 3 trials were conducted among female sex workers and for the remaining one sub-analysis among women was included. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

fb. One individual randomized trial and 3 cluster randomized trials.

fc. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 3 trials or unclear risk of detection bias in one trial. One cluster randomized trial was subject to recruitment bias. Two trials had unclear risk of selection bias (one random sequence generation and one allocation concealment).

fd. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 4.38$, $\text{df} = 5$, $p = 0.47$; $I^2 = 0\%$, $0\% - 71\%$). We did not downgrade.

fe. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ff. Two trials were conducted among young people and other two presented sub-analysis among young people. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

fg. 2 individual randomized trials and two cluster randomized trial.

fh. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in three trials or unclear risk of detection bias in one trial, and attrition bias in one trial (Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). One cluster randomized trial was subject to recruitment bias. All trials had unclear risk of selection bias. 3 of 4 trials had more than three high risk or unclear risk of bias domains.

fi. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.074$; $\text{Chi}^2 = 3.45$, $\text{df} = 3$, $p = 0.33$; $I^2 = 13\%$, $0\% - 87\%$). We did not downgrade.

fj. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

fk. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting); Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but presented sub-groups separately. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

fl. 7 individual randomized trials and 6 cluster randomized trials.

fm. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 5 trials or unclear risk of detection bias in 4 trials, and attrition bias in 2 trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

fn. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.056$; $\text{Chi}^2 = 18.08$, $\text{df} = 15$, $p = 0.26$; $I^2 = 17\%$, $0\% - 54\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 5 of 14 trials defined positivity as confirmed HIV positive diagnosis; the remaining 9 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

fo. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

fp. Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

fq. 4 individual randomized trials.

fr. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in one trial and attrition bias in one trial (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm).

fs. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 1.20$, $\text{df} = 3$, $p = 0.75$; $I^2 = 0\%$, $0\% - 62\%$). We did not downgrade but noted outcome definition was not consistent across trials: confirmed HIV positive diagnosis in three trials and self-reported positivity in one trial.

ft. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

fu. Meta-analysis: Wray, 2018 had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore intervention arms were combined.

fv. 4 individual randomized trials.

fw. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 2 trials and attrition bias in two trials (MacGowan, 2017: 27.1% LTFU in the intervention arm and 28.5% in the control arm; Merchant, 2018: 38.4% LTFU overall, 26% in the intervention 50% in the control arm). 3 of 4 trials had more than three high risk or unclear risk of bias domains.

fx. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 0.88$, $\text{df} = 3$, $p = 0.83$; $I^2 = 0\%$, $0\% - 48\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 2 of 4 trials defined positivity as confirmed HIV positive diagnosis; the remaining 2 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

fy. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

fz. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting) and only HIVST arm with facility-based HIVST distribution is included in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

ga. 2 individual randomized and 2 cluster randomized trials.

gb. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 2 trials.

gc. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0.107$; $Chi^2 = 4.25$, $df = 3$, $p = 0.24$; $I^2 = 29\%$, $0\% - 74\%$). We did not downgrade but noted that outcome definition was not consistent across trials: 2 of 4 trials defined positivity as confirmed HIV positive diagnosis; the remaining 2 trials reported reactive HIVST results as positive, relied on self-reported positivity or the definition was unclear.

gd. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ge. Meta-analysis: Choko, 2019a had more than one intervention arm which included HIVST with similar HIVST kit distribution method, therefore all intervention arms were combined; Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but only included ANC group in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

gf. One individual randomized and 2 cluster randomized trials.

gg. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and unclear risk of attrition bias in one trial. Two cluster randomized trials were subject to recruitment bias and one unclear risk for loss of clusters and not reporting cluster-adjusted analysis (we adjusted for clustering in meta-analysis).

gh. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0$; $Chi^2 = 1.60$, $df = 2$, $p = 0.45$; $I^2 = 0\%$, $0\% - 87\%$). We did not downgrade.

gi. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

gj. Meta-analysis: Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but only included ART group in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

gk. One individual randomized and one cluster randomized trial.

gl. We downgraded twice. This was due to potential for performance bias (lack of blinding) in both trials and detection bias (self-reported or non-validated outcomes) in one trial. One trial unclear risk of selection bias (random sequence generation and allocation concealment). One cluster randomized trials were subject to recruitment bias. Each of the trials had more than three high risk or unclear risk of bias domains.

gm. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0.061$; $Chi^2 = 1.36$, $df = 1$, $p = 0.24$; $I^2 = 26\%$). We did not downgrade.

gn. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

go. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined and only HIVST arm with direct distribution to peers was included in this analysis (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

gp. 2 cluster randomized trials.

gq. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in both trials.

gr. There was low statistical heterogeneity (Heterogeneity: $Tau^2 = 0$; $Chi^2 = 0.27$, $df = 1$, $p = 0.61$; $I^2 = 0\%$). We did not downgrade.

- gs. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- gt. 1 cluster randomized trials.
- gu. We downgraded once. This was due to potential for performance bias (lack of blinding) and unclear risk of detection bias (self-reported or non-validated outcomes).
- gv. Single trial - we downgraded once.
- gw. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- gx. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- gy. 3 individual randomized and one cluster randomized trial.
- gz. We downgraded twice. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 2 trials or unclear risk of detection bias in 2 trials. Each of the trials had more than three high risk or unclear risk of bias domains.
- ha. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.667$; $\text{Chi}^2 = 4.39$, $\text{df} = 3$, $p = 0.22$; $I^2 = 32\%$, $0\% - 76\%$). We did not downgrade.
- hb. We downgraded once because confidence intervals around pooled effect estimate are wide, and likely due to small number of events.
- hc. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- hd. This outcome is only relevant for intervention (HIVST) arm and reported in 9 trials. Results from 7 trials were pooled and reported as pooled percentage. The remaining two trials did not report data usable for pooled analysis.
- he. Three individual randomized trials and 6 cluster randomized trials.
- hf. We downgraded once. This was due to potential for performance bias (lack of blinding) in 8 trials, detection bias (self-reported or non-validated outcomes) in 3 trials and attrition bias in one trials. Two cluster randomized trials were subject to recruitment bias and one loss of clusters. Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- hg. There was high statistical heterogeneity (Heterogeneity: $I^2 = 77\%$, $p < 0.01$). The measurement time-point varied (range: 2 weeks - 5 months) and point estimate from individual studies also varied (25% - 76%). We downgraded once.
- hh. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined arms with similar linkage interventions; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we did not combine arms; Shahmanesh, 2019 had two different interventions - we did not combine arms; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST with different HIVST kit distribution method but no linkage intervention, we combined arms (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- hi. 5 individual randomized and 7 cluster randomized trials.
- hj. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 5 trials or unclear risk of detection bias in one trial, and unclear attrition bias in one trial. Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- hk. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 13.74$, $\text{df} = 15$, $p = 0.56$; $I^2 = 0\%$, $0\% - 47\%$). Sub-group analysis showed heterogeneity was explained by population type. We did not downgrade but noted that outcome definition was not consistent across trials: 7 trials defined linkage as ART initiation whereas other 5 defined it as linkage to any HIV care.
- hl. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

hm. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined arms with similar linkage interventions; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we did not combine arms; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST with different HIVST kit distribution method but no linkage intervention, we combined arms (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

hn. 5 individual randomized and 6 cluster randomized trials.

ho. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in 5 trials or unclear risk of detection bias in one trial, and unclear attrition bias in one trial. Three cluster randomized trials were subject to recruitment bias. Several risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

hp. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.011$; $\text{Chi}^2 = 15.71$, $\text{df} = 13$, $p = 0.29$; $I^2 = 15\%$, $0\% - 53\%$). Sub-group analysis showed heterogeneity was explained by population type. We did not downgrade but noted that outcome definition was not consistent across trials: 7 trials defined linkage as ART initiation whereas other 5 defined it as linkage to any HIV care.

hq. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

hr. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined arms with similar linkage interventions; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we did not combine arms (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

hs. 2 individual randomized trials and 4 cluster randomized trials

ht. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in one trial or unclear risk of detection bias in one trial. Three cluster randomized trials were subject to recruitment bias. Three trials had unclear risk of selection bias (one random sequence generation and 3 allocation concealment) and some other risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.

hu. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.018$; $\text{Chi}^2 = 11.29$, $\text{df} = 8$, $p = 0.19$; $I^2 = 29\%$, $0\% - 67\%$). We did not downgrade but noted that the outcome definition was not consistent across trials: some trials defined linkage as ART initiation whereas others defined it as linkage to any HIV care.

hv. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

hw. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.

hx. 3 individual randomized trials and two cluster randomized trials.

hy. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 2 trials.

hz. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 3.24$, $\text{df} = 5$, $p = 0.52$; $I^2 = 0\%$, $0\% - 74\%$). We did not downgrade but noted that the outcome definition was not consistent across trials: some trials defined linkage as ART initiation whereas others defined it as linkage to any HIV care.

ia. We did not downgrade but noted that trials were conducted among men who have sex with men, transgender women and female sex workers, so results should be viewed with caution when applying to other key populations.

ib. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

ic. 3 individual randomized trials.

id. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials.

- ie. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 0.27$, $\text{df} = 2$, $p = 0.87$; $I^2 = 0\%$, $0\% - 24\%$). We did not downgrade.
- if. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- ig. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- ih. 2 cluster randomized trials.
- ii. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in both trials.
- ij. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 0.63$, $\text{df} = 1$, $p = 0.43$; $I^2 = 0\%$). We did not downgrade.
- ik. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- il. 3 studies were conducted among MSM, 3 among male partners of women, and for 2 sub-group analysis for men was included. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined all intervention arms for this analysis; Choko, 2019b had two intervention arms which included HIVST and had two sub-groups (ANC women and ART clients) - we combined HIVST arms but only included ANC group in this analysis. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trial.
- im. 5 individual randomized trials and 3 cluster randomized trials.
- in. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, and detection bias (self-reported or non-validated outcomes) in one trial or unclear risk of detection bias in one trial. Three cluster randomized trials were subject to recruitment bias. Some other risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts. Each of the trials had more than three high risk or unclear risk of bias domains.
- io. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.045$; $\text{Chi}^2 = 9.30$, $\text{df} = 7$, $p = 0.23$; $I^2 = 25\%$, $0\% - 66\%$). We did not downgrade.
- ip. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- iq. 2 trials were conducted among female sex workers and for 2 trials sub-group analysis among women was included. Meta-analysis: Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- ir. One individual randomized and two cluster randomized trials.
- is. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 2 trials or unclear risk of detection bias in one trial. Two trials had unclear risk of selection bias (one random sequence generation, two allocation concealment) and some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- it. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 2.48$, $\text{df} = 3$, $p = 0.48$; $I^2 = 0\%$, $0\% - 81\%$). We did not downgrade.
- iu. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- iv. For both trials sub-analysis among young people was included in the analysis. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- iw. One individual randomized and once cluster randomized trial.
- ix. We downgraded twice. This was due to potential for performance bias (lack of blinding) in both trials and unclear risk of selection bias (one random sequence generation and two allocation concealment). One cluster randomized trial was subject to recruitment bias. Each of the trials had more than three high risk or unclear risk of bias domains.

- iy. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 0.26$, $\text{df} = 1$, $p = 0.61$; $I^2 = 0\%$). We did not downgrade.
- iz. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- ja. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined arms with similar linkage interventions; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we did not combine arms; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- jb. 3 individual randomized and 5 cluster randomized trials.
- jc. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, and detection bias (self-reported or non-validated outcomes) in 4 trials or unclear risk of detection bias in one trial. Three cluster randomized trials were subject to recruitment bias. Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- jd. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.034$; $\text{Chi}^2 = 17.74$, $\text{df} = 10$, $p = 0.14$; $I^2 = 32\%$, $0\% - 67\%$). We did not downgrade.
- je. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- jf. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- jg. 2 individual randomized trials and one cluster randomized trial.
- jh. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and unclear risk of allocation concealment (selection bias) in one trial.
- ji. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 0.46$, $\text{df} = 2$, $p = 0.79$; $I^2 = 0\%$, $0\% - 55\%$). We did not downgrade.
- jj. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- jk. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we only included the arm with no linkage intervention in this analysis; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we only included the arm with no linkage intervention in this analysis; Chanda, 2017 and Ortblad, 2017 had more than one intervention arm which included HIVST but HIVST kit distribution method was different, therefore intervention arms were not combined (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- jl. 4 individual randomized trials and 5 cluster randomized trials.
- jm. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials and detection bias (self-reported or non-validated outcomes) in 3 trials or unclear risk of detection bias in one trial. Three cluster randomized trials were subject to recruitment bias. Two trials were unclear risk of selection bias (one random sequence generation and two allocation concealment) and some other risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- jn. There was moderate statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.058$; $\text{Chi}^2 = 13.03$, $\text{df} = 8$, $p = 0.11$; $I^2 = 39\%$, $0\% - 72\%$). Heterogeneity can be explained by difference in study populations. We did not downgrade.
- jo. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- jp. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we only included the arm with phone reminder for linkage. Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- jq. One cluster randomized trial.
- jr. We downgraded once. This was due to potential for performance bias (lack of blinding) and high risk of cluster recruitment bias.

- js. Single trial, we downgraded once as inconsistency cannot be evaluated.
- jt. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- ju. Meta-analysis: Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- jv. 1 cluster randomized trials.
- jw. We downgraded once. This was due to potential for performance bias (lack of blinding) in both trials and unclear risk of allocation concealment (selection bias) in one trial.
- jx. Single trial - we downgraded once as inconsistency cannot be evaluated.
- jy. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- jz. One individual randomized trial.
- ka. We downgraded once. This was due to potential for performance bias (lack of blinding).
- kb. Single trial, we downgraded once as inconsistency cannot be evaluated.
- kc. Confidence intervals around point estimate are wide. We downgraded once.
- kd. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- ke. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with different linkage interventions - we combined arms with financial linkage interventions and included in the analysis; Choko, 2019b had two intervention arms which included HIVST and different linkage intervention - we did not combine arms and included the arm with financial linkage intervention in this analysis (control arm sample was adjusted to prevent double counting). Cluster-adjusted analysis was included in meta-analysis for cluster randomized trials.
- kf. 2 cluster randomized trials.
- kg. We downgraded once. This was due to potential for performance bias (lack of blinding) in both trials and unclear risk of detection bias (self-reported or non-validated outcomes) in one trial. Both cluster randomized trials were subject to recruitment bias. One trial had unclear risk of bias for attrition or loss of clusters and unadjusted cluster analysis (we included cluster adjusted estimates).
- kh. There was low statistical heterogeneity (Heterogeneity: $\tau^2 = 0$; $\text{Chi}^2 = 0.08$, $df = 1$, $p = 0.77$; $I^2 = 0\%$). We did not downgrade.
- ki. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.
- kj. Two trials reported this outcome which is only relevant to the intervention (HIVST) arm.
- kk. Two cluster randomized trials.
- kl. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in both trials. One cluster randomized trials was subject to recruitment bias. Some risk of bias domains were unclear risk due to lack of information from unpublished reports or conference abstracts.
- km. The outcomes were not pooled therefore inconsistency cannot be evaluated. Similar outcomes were reported from both included trials.
- kn. We did not downgrade but noted that one trial involved facility-based HIVST and one trial involved home-based HIVST distribution by community workers. Both trials were among general population. The results should be viewed with caution when applying to other populations.
- ko. The results are based on very few events so caution is needed when interpreting. We downgraded once.

kp. Two trials were conducted involved HIVST distribution by women to their male partners (Choko, 2019a; Masters, 2016), one involved HIVST distribution by HIV-positive clients to their partners (Dovel, 2019), one involved HIVST distribution by both women to their male partners and HIV-positive clients to their partners (Choko, 2019b), and two involved distribution of HIVST kits or coupons by female sex workers to their peers (Chanda, 2017, Ortblad, 2017).

kq. 2 individual randomized and 4 cluster randomized trials.

kr. We downgraded once. This was due to potential for performance bias (lack of blinding) and detection bias (self-reported or non-validated outcomes) in all trials. One cluster randomized trials was subject to recruitment bias and one trial had unclear risk of selection bias (random sequence generation and allocation concealment).

ks. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0$; $\text{Chi}^2 = 2.96$, $\text{df} = 5$, $p = 0.71$; $I^2 = 0\%$, $0\% - 57\%$). We did not downgrade but noted that type of adverse events varied across trials (intimate partner violence: Chanda, 2017, Choko, 2019a, Masters, 2016, Ortblad, 2017; temporary relationship breakdown: Choko, 2019b; verbal abuse: Dovel, 2019).

kt. All trials involved secondary distribution to partners (in Mulubwa, 2019 a subset of all participants distributed to their partners) or peers so right population of interest for this outcome.

ku. Confidence intervals around pooled effect size were large, likely due to very few events. We downgraded once.

kv. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

kx. All trials were among men who have sex with men.

ky. 3 individual randomized and one cluster randomized trial.

ky. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in all trials, and unclear risk of selection bias (allocation concealment) in one trial.

kz. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.002$; $\text{Chi}^2 = 3.39$, $\text{df} = 3$, $p = 0.34$; $I^2 = 11\%$, $0\% - 86\%$). We did not downgrade but noted that outcome definitions varied across trials (Jamil, 2017: condomless anal intercourse with casual male partner(s) during 12 months follow-up; Katz, 2018: non-concordant condomless anal intercourse with male partner(s) in the past 3 months measured at 15 months; Tang, 2018: condomless anal sex with male partner(s) in the past 3 months measured at 6 months (mid-point during follow-up); Wang, 2017: condomless anal intercourse with male partner(s) in the past 3 months measured at 6 months).

la. All trials were among men who have sex with men so results should be viewed with caution when applying to other populations.

lb. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

lc. All trials were among men who have sex with men.

ld. 3 individual randomized and one cluster randomized trial.

le. We downgraded once. This was due to potential for performance bias (lack of blinding) in all trials, detection bias (self-reported or non-validated outcomes) in all trials, and unclear risk of selection bias (allocation concealment) in one trial.

lf. There was low statistical heterogeneity (Heterogeneity: $\text{Tau}^2 = 0.0003$; $\text{Chi}^2 = 3.06$, $\text{df} = 3$, $p = 0.38$; $I^2 = 2\%$, $0\% - 85\%$). We did not downgrade but noted that outcome definitions varied across trials (Jamil, 2017: condomless anal intercourse with casual male partner(s) during 12 months follow-up; Katz, 2018: non-concordant condomless anal intercourse with male partner(s) in the past 3 months measured at 15 months; Tang, 2018: condomless anal sex with male partner(s) in the past 3 months measured at 6 months (mid-point during follow-up); Wang, 2017: condomless anal intercourse with male partner(s) in the past 3 months measured at 6 months).

lg. All trials were among men who have sex with men so results should be viewed with caution when applying to other populations.

lh. Original sample size, does not represent effective cluster adjusted sample size in the meta-analysis.

Table B2. GRADE table for HIVST + linkage intervention compared to HIVST only for linkage to HIV CARE or ART

Author(s): Muhammad S. Jamil

Question: HIVST + linkage intervention compared to HIVST only for linkage to HIV care or ART

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	HIVST + linkage intervention	HIVST only	Relative (95% CI)	Absolute (95% CI)		
Linkage to HIV care or ART initiation												
5 ^a	randomised trials ^b	serious ^c	serious ^d	not serious	not serious	none	308/632 (48.7%)	117/375 (31.2%)	RR 1.38 (1.02 to 1.87)	119 more per 1,000 (from 6 more to 271 more)	⊕⊕○○ LOW	CRITICAL

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

Explanations

a. Meta-analysis: Choko, 2019a had multiple intervention arms which included HIVST with financial incentive for linkage - we combined these arms (HIVST only arm sample was adjusted to prevent double counting); Choko, 2019b: HIVST + financial incentive for linkage vs. HIVST only; MacPherson, 2014: home-based HIVST + optional home ART initiation vs. HIVST + standard linkage; Nichols, 2019: HIVST arm had two linkage interventions (HIVST + self-referral and HIVST + in-person referral/escort); we compared HIVST + in-person referral/escort vs. HIVST + self-referral; Sibanda, 2019 had two intervention arms (HIVST + fixed provider financial incentive and HIVST + fixed provider incentive and conditional incentive per linkage); we compared HIVST + fixed and conditional incentive vs. HIVST + fixed incentive. Cluster-adjusted analysis was included in meta-analysis cluster randomized trials.

b. 5 cluster randomized trials.

c. We downgraded once. This was due to potential for performance bias (lack of blinding) in 4 trials or unclear risk of performance bias in one trial, high risk of detection bias (self-reported or non-validated outcomes) in one trial. Four cluster randomized trials were subject to recruitment bias (high or unclear risk). Some risk of bias domains were unclear due to limited information from conference abstracts.

d. There was high statistical heterogeneity (Heterogeneity: Tau² = 0.097; Chi² = 21.75, df = 5, p <0.01; I² = 77%, 49% - 90%). We downgraded once and noted that study design, type of linkage intervention and comparison group varied across studies.

9789240011748

