ANNEX 14. A report on the misdiagnosis of HIV status

Authors: Johnson C[1], Fonner V[2], Sands A[3], Tsui S[2], Ford N[1], Wong V[4], Obermeyer C[5], Baggaley R[1]

14.1 Purpose and introduction

Between 2010 and 2014 nearly 600 million adults (ages 15+) reportedly received HIV testing services in 122 low- and middle-income countries in this period. In 2014 alone, approximately 150 million children and adults men and women across 129 low- and middle-income countries received HIV testing and their test results.[6] Such scale-up has been possible partly through the use of rapid diagnostic tests (RDTs) which have been critical for expanding HTS, particularly in resource-limited settings where access to laboratory services is limited. RDTs that have been prequalified by the World Health Organization (WHO) have at least a diagnostic and clinical sensitivity of 99% and a specificity of 98% (1). The performance of such RDTs within a WHO recommended testing strategy and a validated national testing algorithm has been shown to be highly reliable (2-4). Despite this, in many countries quality management systems have not kept pace with the scale-up of HTS and have not utilized WHO prequalified RDTs or implemented WHO recommended testing strategies. According to a recent analysis of 48 country HIV testing policies, only 20% of testing strategies followed current WHO recommendations (5).

There are a growing number of reports indicating poor quality HIV testing, some of which result in the misdiagnosis of HIV status. Misdiagnosis of HIV status refers to an incorrectly reported outcome of a testing process, i.e. incorrectly identifying someone who is HIV-infected as HIV-uninfected or vice versa. According to audits in some settings there a substantial level of false-positive diagnoses (6, 7). Misdiagnosis of HIV can result from limitations of the assay itself, operator-related factors, facility-related factors such as inappropriate storage of reagents and selection of inappropriate testing strategies, and suboptimal national testing algorithms. Workload can also contribute to misdiagnosis – either through high client volume resulting in overburden, with low client volumes resulting in poor proficiency; inadequate training and lack of supportive supervision can further contribute to operator error (7, 8). In many countries errors that occur may be missed and go without correction due to the lack of comprehensive quality management systems or adequate external quality assessment.

The consequences of such misdiagnoses are serious. A false-negative HIV status is a failure to correctly diagnosis a client, link them to prevention, treatment and care services; and to also prevent ongoing HIV transmission. A false-positive HIV diagnosis can result in the unnecessary initiation of life-long antiretroviral therapy (ART) and the potential for stigma, discrimination, criminalization; it limits the full offer of interventions to prevent HIV infection; it may also harm community and family relationships, and if the results are subsequently found to be incorrect may reduce the credibility of test results and the trust in health services. Incorrectly initiating ART wastes scarce resources, including stocks of medicines, clinic and staff time, and ART monitoring by viral load and other laboratory tests. This is particularly troubling in settings where after receiving a HIV-positive result clients are offered ART immediately regardless of CD4 count or clinical assessment. Left unaddressed, such quality issues and misdiagnoses could undermine the HIV response and make it difficult to reach new global targets, including to *correctly* diagnose 90% of people with HIV by 2020 (9), prevent new HIV infections and provide treatment to people with HIV.

We conducted this systematic literature review to assess the causes and extent of misdiagnosis in different settings.

[1] WHO, HIV/AIDS Department, Geneva, Switzerland
[2] Johns Hopkins University, Bloomberg School of Public Health, Baltimore, USA
[3] WHO, Essential Medicines and Health Products, Geneva, Switzerland
[4] USAID, Global Health Bureau, Office of HIV/AIDS, Washington, DC
[5] American University of Beirut, Beirut, Lebanon
[6] Global AIDS Response Progress Reporting (GARPR) (WHO, UNAIDS, UNICEF) 6 July 2015.

14.2. Methods

We systematically searched for peer-reviewed articles published from January 1990 to July 2014 using a pre-defined set of search terms in the several electronic databases (see Table 14.1A): International AIDS Conference (IAC) and IAS conference abstracts were searched from July 2001 through July 2013. At the time of the search, only the most recent CROI conference (2014) database was searched because abstracts from past conferences were not electronically available. For ASLM, only 2012 conference abstracts were searched because at the time the 2014 meeting had not yet occurred and previous conferences were not electronically available. We also searched reference lists to identify additional relevant literature. This process was repeated iteratively until no new citations were identified. Experts were also contacted to identify additional articles and gray literature reports. No geographic restrictions were placed on the search, but the review was limited to documents in English.

Table 14.1A Search strategy and search terms

Search strategy
1. Database searching: PubMed, CINAHL, and EMBASE will be searching from January 1, 1990-July 9 2014 to identify relevant scientific articles. 2. Conference abstracts: International AIDS Conference (IAC), Conference on HIV Pathogenesis, Treatment, and Prevention (IAS), and Conference on Retroviruses and Opportunistic Infections (CROI) websites to identify conference abstracts. The IAC and IAS conference abstracts will be searched from 2001-present. For CROI, only the most recent conference (2014) will be searched as past conferences are inaccessible. 3. Government and NGO websites: Médecins Sans Frontières, USAID, UNAIDS, WHO, CDC, AMREF, SEAD, ASLM, FHI360, CHAI, JHPIEGO, PSI and ICAP
Search Terms
1. (misclassif* OR error* OR "false positive" OR "false positives" OR "false negative" OR "false negatives" OR misdiagnos*) AND (test* or algorithm* or screen* OR diagnos*) AND (rapid OR "point of care" OR POC OR RDT) AND ("HIV" OR "Human Immunodeficiency Virus") For searching electronic databases related to proficiency testing and external quality assessment of HIV testing algorithms involving RDTs, the following set of search terms will be used: ("proficiency testing" OR "external quality assessment") AND (rapid OR "point of care" OR POC OR RDT) AND (HIV OR "Human Immunodeficiency Virus") 2. For conference abstract searching, more simplified search terms will be used due to the basic search capacity of databases. Multiple searches will take place, including searches on "misclassification," "misdiagnosis" and other terms related to misclassification of HIV status. Additionally, abstracts will be searched using terms related to poor quality testing, quality assurance, and proficiency testing failure. 3. For website searches, all available publications will be searched for reports, abstracts, studies, or articles related to HIV status misclassification, factors related to potential misclassification, and quality of HIV testing occurring in programmatic settings.

Studies were eligible for inclusion if they included rates of HIV status misdiagnosis using algorithms involving at least two RDTs, reported factors related to potential misdiagnosis, or described issues related to the quality of HIV testing being implemented in programmatic settings (facility, laboratory or community-based) as well as sites reporting on results of external quality assessment schemes, such as proficiency testing. Potential causes of

misdiagnosis were extracted from studies using the following categories: test kits, operator error, testing strategy or algorithm, population characteristics, and testing sites.

The primary summary of measure was the count of studies which reported one of the following errors: (1) clerical/technical error (defined as operator error in documenting and reporting information essential to a correct test result), (2) user error (defined as operator error collecting specimen, performing an HIV RDT or interpreting the test result), (3) cross-reactivity (defined as characteristics of the HIV RDT, of the population or of the geographic location that are related to incorrect test results); (4) incorrect/suboptimal testing strategy (defined as errors related to the order in which specific RDTs are used, also known as a testing strategy); and (5) poor management and supervision (defined as lack of active quality management systems). Secondary summary measures include: HIV-positive misdiagnosis rates (number of false-positive tests reported over the number of HIV-positive tests reported and retested to confirm HIV status reports), (2) HIV-negative misdiagnosis rates (number of false-negative tests reported over the number of HIV-negative tests reported); and (3) sensitivity and specificity of the reporting testing algorithm (see full Tables 4a-4f and Table 5).

We followed PRISMA reporting standards (see annex). The full electronic search strategy is available in the review protocol (see annex). Initial titles were screened by one investigator (VF) to determine eligibility. A second screening was then carried out by two reviewers (VF and ST). Final inclusion was determined by CJ, AS, and RB. All differences were resolved through consensus with VF, AS, CJ and RB. Data from all sources were extracted and placed into standardized forms by one reviewer (VF). A second reviewer (ST) verified data extracted. CJ and AS independently reviewed and finalized data for inclusion.

14.3. Results

Forty-two studies that addressed and reported factors potentially related to misdiagnosis were identified and included in this review (see Figure 14.1A, Table 14.2A and Table 14.3A). Studies were carried out in Brazil (n=2) (10, 11), Cameroon (n=2) (12, 13), Cambodia (n=1) (14), the Democratic Republic of Congo (DRC) (n=1) (6), Ethiopia (n=2) (15, 16), Honduras (n=1) (17), Haiti (n=1), India (n=1), Kenya (n=1) (18), Malawi (n=2)(19, 20), Mozambique (n=1) (21), Nigeria (n=2) (22, 23), South Africa (n=3) (24-26), Tanzania (n=3) (27-29), Uganda (n=4) (30-33), United Kingdom (n=1), USA (n=2), Zimbabwe (n=1) (34), and multi-country studies (n=13) (3, 8, 35-44). Sample sizes varied in unit of measure, including clients (n=20 range: 106 035 to 3), specimens (n=14, range: 9 419 to 34), health workers performing HIV testing (n=4 range: 3 835 to 39 personnel), sites where HIV testing was performed (n=8 range: 602 to 38 sites) and six studies reported more than one unit of measure (see Supplementary Tables 3a-f and Table 14.4A). Forty-two studies occurred in a facility-based setting; the other two studies took place in the workplace (n=1) (34) in a field-based mobile-testing service (n=1) (32).

Misdiagnosis of HIV-positive status

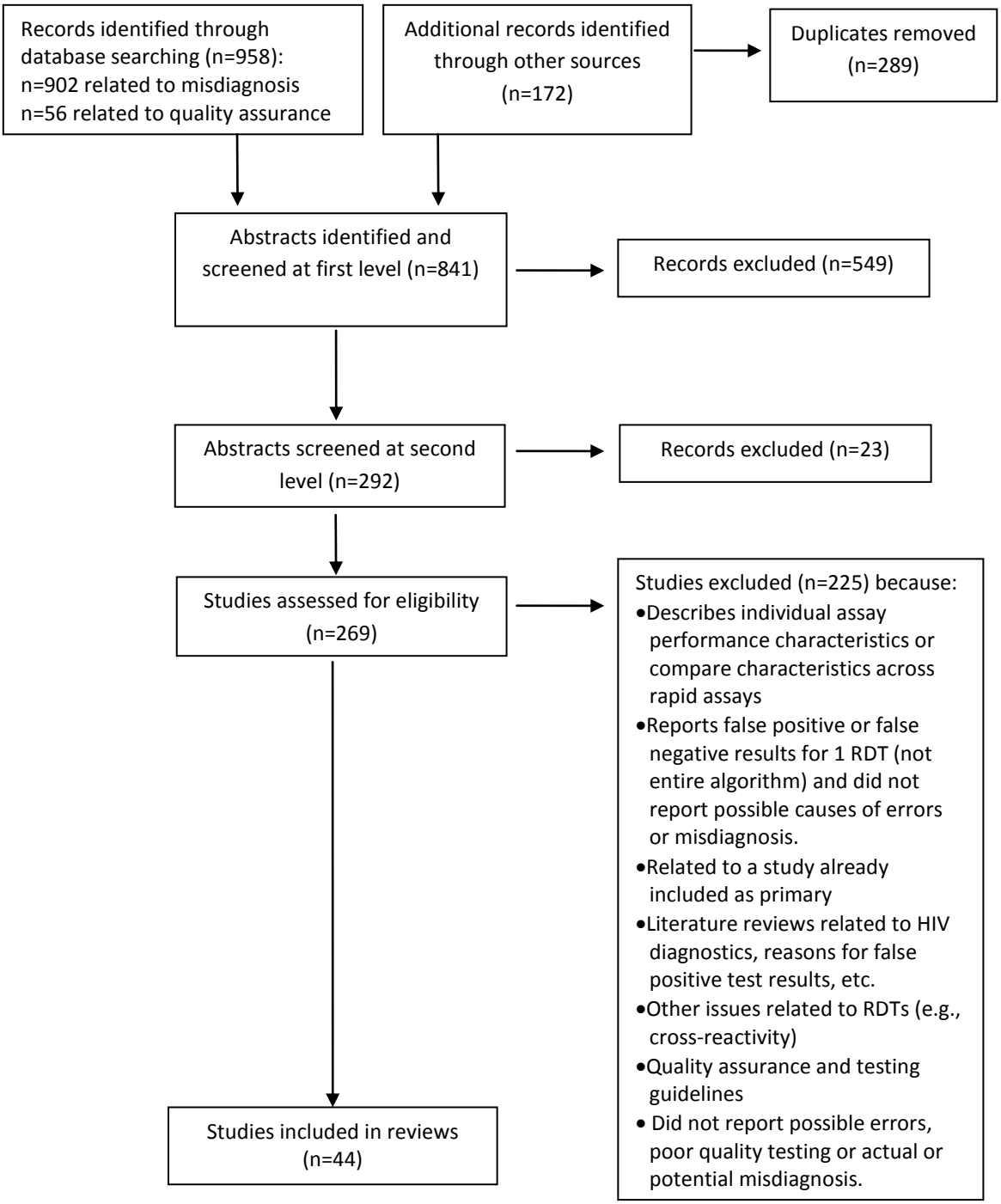
Two studies, which evaluated the same HTS programme, conducted a retrospective audit to assess misdiagnosis of HIV status among individuals in HIV care and treatment programmes (6, 35). The first study, conducted in DRC, Burundi and Ethiopia reported multiple reasons which may have caused misdiagnosis including providers not following standard operating procedures, high workload and stress on testing providers, damaged test kits (i.e. "ruptured test kits") and incorrect diagnosed using the result of a single assay (RDT or enzyme immunosassay (EIA)) rather than a full recommended testing strategy (35). Overall the rate of false-positive diagnoses was 2.6% in Burundi, 4.8% DRC—which was 10.3% in the first 5 months of retesting with a parallel testing strategy—and 4.7% in Ethiopia (35). The second study, conducted in DRC by the same group but over a different time period, reported a false-positive diagnosis rate of 10.5% in the first five months of re-testing and average of 10.3% misdiagnosis rate by the study end (6). The testing strategy used in this study consisted of two HIV RDTs conducted in parallel (6). Potential causes of misdiagnosis reported include the over-interpretation of weak reactive test results and possible cross reactivity due to HLA class II antigens directly modulating the immune response, on-with specific polyclonal B-lymphocyte antibodies (6).

Misdiagnosis HIV-negative status

Three studies reported rates of missed HIV infection where testing algorithms incorrectly identified HIV-positive individuals as HIV-negative (24, 27, 34). One study in South Africa, reported that 2% of patients who were later found to be HIV-positive (and not recently infected) had previously been diagnosed HIV-negative or HIV-

inconclusive (with discrepant HIV test results) (24). The reported cause of these misdiagnoses was the testing strategy: the study used both serial and parallel two-assay testing strategies and the testing algorithm (24). A second study, from Zimbabwe, reported 81% (547/674) of participants who were tested for HIV were HIV-negative; one of which was false-negative (34). It was reported that the cause of this false-negative was "user error" attributed to "slowly reactive" specimens both times the RDT was performed (34). The third study, from Tanzania, reported three false-negative cases, including a single case of a woman who was experiencing symptoms associated with HIV and several opportunistic infections and who tested HIV-negative but was in fact HIV-positive (27). The cause of this misdiagnosis, reported by the authors, was late-stage HIV infection which can result in low levels of antibodies (27).

Figure 14.1A Study selection process



Comparing HIV surveillance data to programme data

Two studies reported on differences between data from on-site prevention of mother-to-child transmission (PMTCT) diagnostic testing and data from central-level testing done for antenatal care (ANC) sentinel surveillance sentinel surveillance in antenatal care (21, 41). One study identified substantial discrepancies across nine countries with generalized HIV epidemics—only 4/9 reported more than 90% agreement between results from PMTCT diagnostic testing and ANC sentinel surveillance (41). In Mozambique, on study compared HIV prevalence rates ascertained through different testing methods on the same population— testing as part of PMTCT programme compared with ANC sentinel surveillance data collected simultaneously (21). Over a two year period, it was reported that HIV prevalence was 14.4% from PMTCT testing and 15.2% from surveillance testing. In both studies the discrepancies were attributed to a multiplicity of factors, including: human error performing an EIA (used for surveillance) or HIV RDTs (use for PMTCT) (21), errors in interpreting test results (41), use of an inappropriate and/or two different testing strategies and algorithms in PMTCT and ANC surveillance (41), problems with stock management (use of exposed or expired test kits) (41) and poor record keeping (21, 41).

Table 14.2A. Classification of included studies (n=44)

Category	Study	Location
Rates of HIV misdiagnosis	Shanks et al. 2014 (35)	DRC, Burundi, Ethiopia
	Klarkowski et al. 2009 (6)	DRC
Rates of HIV-1/HIV-2 type misdiagnosis	Tchounga et al 2014 (37)	Burkina Faso, Cote d'Ivoire, Mali
Evaluation of sensitivity and specificity of RDTs	Aghokeng et al. 2009 (12)	Cameroon
	da Costa et al. 2007 (10)	Brazil
	Eller et al., 2007 (31)	Uganda
	Galiwango et al. 2013 (13)	Cameroon
	Klarkowski et al. 2013 (36)	CAR, Congo-B, DRC, Ethiopia, Haiti, India, Ivory Coast, Myanmar, Uganda, Zimbabwe
	Martin et al. 2011(45)	USA
	Mehra et al. 2014 (46)	Northern India
	Stetler et al. 1998 (17)	Honduras
	Viani et al. 2013 (44)	USA and Mexico
Evaluation of RDT-based testing algorithms in research settings	Baveewo et al. 2012 (30)	Uganda
	Crucitti et al. 2011 (42)	Benin; India; South Africa; Uganda; Karnataka, India
	Gray et al. 2007 (32)	Rakai, Uganda
	Jentsch et al. 2012 (38)	South Africa, Tanzania, Uganda, and Zambia
Comparing surveillance testing strategies data to diagnosis (RDT-based) testing strategy	Young et al. 2013 (21)	Mozambique
	CDC unpublished, 2014 (41)	low- and low-middle income countries (countries not specified)
Potential misdiagnosis in low-income countries	Government of Malawi Ministry of Health 2014 (19)	Malawi
	Government of Malawi Ministry of Health 2014 (20)	Malawi
Focus on misdiagnosis of HIV-negative serostatus	Bassett et al. 2011(24)	South Africa
	Kahemeles et al. 2008 (27)	Tanzania
	Matambo et al. 2006 (34)	Harare, Zimbabwe
EQA (proficiency testing) results	Louis et al. 2013 (47)	Haiti
	Manyazewal et al. 2013 (15)	Ethiopia
	Tegbaru et al. 2009 (16)	Ethiopia
	Sushi et al. 2011 (48)	India
Quality issues	Benzaken et al. 2014 (11)	Brazil

Category	Study	Location
discovered from EQA	Cham et al. 2012 (39)	30 countries in Africa
	Plate et al. 2007 (3)	11 countries in Africa
	Iwe et al. 2011 (22)	Nigeria
	Kalou et al. 2012 (8)	Uganda and Tanzania
	Kitheka et al. 2012 (18)	Kenya
	Lali et al. 2010 (33)	Uganda
	Ocheng et al. 2012 (28)	Tanzania
Factors related to potential misdiagnosis	Boeras et al. 2011 (40)	Lusaka, Zambia and Rwanda
	Granade et al. 2004 (49)	USA
	Kanal et al. 2005 (14)	Cambodia
	Learmonth et al. 2008 (43)	26 countries
	Sacks et al. 2012 (50)	UK
	Wolpaw et al. 2010 (25)	South Africa
General quality issues from sites conducting HTS	Adebayo et al. 2012 (23)	Nigeria
	Mashauri et al. 2007 (29)	Tanzania
	SEAD 2010 (26)	South Africa

Programme data

In Malawi, following the retesting of people diagnosed HIV-positive before they enrolled in care or treatment, 7.7% (534/7,007) could not be conclusively diagnosed (19). A follow-up study also reported 4.3% (346/8,017) of individuals previously diagnosed HIV-positive could not be conclusively diagnosed (20). At the time of both studies it was reported that many sites where ART is initiated did implement retesting routinely (19, 20).

Results from proficiency testing/external quality assessment schemes

Three articles (15, 16, 47) reported results from external quality assessment (EQA) schemes (also known as proficiency testing) that compared with the assigned value (expected reference results). One study, from Haiti, reported that between across a six year period 92.0% (242/263) of laboratories participated in proficiency testing and that concordance ranged from 88% to 90% in 2006 and 2011 respectively. Declines in concordance of results were noted in 2008 and 2010 and were linked to health worker shortages that resulted from natural disasters (47). Two studies from Ethiopia reported good performance within proficiency testing schemes (15, 16). Overall agreement was high (between 95% and 100%)(15, 16). However, one study reported the concordance of results was lower in health centres than in district hospitals when HIV RDTs using capillary whole blood specimens were tested, 87.5% (14/16) 93.8% (15/16) respectively (15). Other errors that caused incorrect results include the use of a tiebreaker strategy (15), health workers having difficulty interpreting weakly reactive test lines and stock-outs of reagents (16).

Clerical errors

Fourteen studies reported on clerical errors (8, 10, 14, 17, 18, 21, 22, 31, 33, 38, 41, 46, 48, 49). These did not always lead to misdiagnosis but report that errors did occur, such as poor recordkeeping including writing reports (14), data reporting problems, labelling and transcription mistakes (48) and specimen mix-ups could have result in misdiagnosis of HIV status. Poor record keeping, according to one study, resulted in nearly 30% of errors resulting in incorrect test status (33). Clerical errors were not always clearly defined in the studies reviewed, but noted as possible causes of incorrect test results and therefore HIV status (10).

User error

User error was reported as a potential cause of misdiagnosis in 21 studies (6, 10, 11, 13, 15, 16, 18, 26, 30, 32, 34, 35, 39, 42, 43, 45, 47-50). User errors were defined as human errors by the operator such as incorrectly performing the test procedure or incorrectly interpreting test results, incorrect buffer volume, not following reading times, inaccurate reading time stated in the instructions for use and non-adherence to standard operating procedures (6, 13, 16, 30, 32, 39, 42, 43, 48, 50). In some provinces in South Africa, it was reported that the average turnaround time for an HIV testing visit was four minutes whereas the first assay of the testing algorithm had a recommended minimum reading time of 15 minutes (26); other errors identified include users having difficulty performing the RDT

(48, 49)—particularly specimen collection (10, 15), the incorrect interpretation of results (6, 13, 16, 18, 30, 32, 39, 42, 43, 45, 47, 50), and the use of a single assay testing strategy to issue a HIV-positive diagnosis (47).

Eleven studies reported that users had difficulty interpreting weak reactive test results (i.e. faint test lines/bars) (6, 13, 16, 30, 32, 35, 39, 42, 43, 45, 50). It was noted that some providers "over-interpreted" faint lines and were more likely to incorrectly assign an HIV reactive test result and therefore issue an incorrect HIV-positive diagnosis. A study which collected specimens from 26 countries used photographs of test lines/bands from RDTs to test the proficiency of laboratory technicians for reading and interpretation of test results and found that specimens with very weak dilutions of specimens containing HIV-1/2 antibodies were less accurately reported (43). A study from Uganda also found that the majority of false-reactive test results came from specimens that provided weak test lines/bands on one or more of the RDTs used (32). A study from the UK that assessed the visual depiction of false-reactive and true positive readings reported that most false-reactive specimens had a fainter test line/band compared to true-positive specimens (50). Furthermore, both studies assessing misdiagnosis speculated that reported faint test lines as reactive on one or both of the RDTs contributed to the misdiagnosis of HIV infections (6, 35).

Table 14.3A. Reported factors related to poor quality HIV testing & potential misdiagnosis (n=44)[7]

Category	#
Clerical/technical errors (e.g. mis-labelling, poor recordkeeping, clerical mistakes) (8, 10, 14, 17, 18, 21, 22, 31, 33, 38, 41, 46, 48, 49)	14
User error (e.g. errors performing RDT or interpreting results, misapplication of buffer, inaccurate reading time and other human errors) (6, 10, 11, 13, 15, 16, 18, 26, 30, 32, 34, 35, 39, 42, 43, 45, 47-50)	21
Cross-reactivity (e.g. antibodies from inter-current infection, environmental exposure to test components, HIV subtype, or late-stage AIDS) (6, 11, 12, 27, 32, 36, 40, 42)	8
Incorrect / suboptimal testing strategy or algorithm (e.g. tiebreaker testing strategy) (3, 6, 8, 10, 12, 13, 18-20, 23, 24, 30, 31, 36-42, 44, 45)	22
Poor management and supervision (work load stress, staff shortages, lack of training, poor adherence to testing strategy or testing algorithm, substandard operating procedures, testing in window period) (8, 14, 16, 18-20, 23, 25, 26, 28, 29, 32, 33, 35, 39, 41, 42, 46, 47, 49)	20

Suboptimal testing strategy

A testing strategy generically describes a testing sequence for a specific objective, taking into consideration the presumed HIV prevalence in the population being tested. Twenty-two reported using a suboptimal testing strategy that differed from WHO recommendations (3, 6, 8, 10, 12, 13, 18-20, 23, 24, 30, 31, 36-42, 44, 45). Five studies used a parallel testing strategy (13, 19, 20, 38, 44). Nine studies reported using a testing strategy known as a "tiebreaker" where a third assay is used to resolve discrepant results and rule in HIV infection (3, 10, 13, 30-32, 38, 40, 42). One study reported using a high prevalence testing strategy in what was most likely a low prevalence setting (i.e. where HIV prevalence of the population tested was under 5%) (45). Test results were mostly concordant (368/369) between results from two consecutively reactive HIV RDTs with results confirmed using a Western blot; however there were 32 discordant results and two false-positives using the RDT-based algorithm alone (one client which was confirmed HIV-negative using Western blot and one client who received on the RDT-based algorithm and refused Western blot) (45).

Seven studies reported a substantial level of false-positive test results and possible misdiagnoses as a result of using a tiebreaker (10, 30-32, 38, 40, 42). In Brazil, a tiebreaker was used three times, two of which led to a false-positive test result (10). Two studies from Uganda that used a tiebreaker report high false-positive rates of 43.7% (30) and 48.2% (32); one of which reported that 95.3% (123/129) of false-positives resulted from specifically using a

[7] This table shows the proportion of included studies that reported errors within these select categories. Note since multiple studies report various factors related to poor quality testing and factors that could be related to potential misdiagnosis, this table does not add up to 100%.

tiebreaker (32). Three multi-country studies report false-positive rates of 8.7% (4/46) (42), 10.2% (55/537) (38) and 45.9% (17/37) (40) when participants were incorrectly diagnosed HIV-positive using a tiebreaker.

Cross-reactivity and characteristics of the individual undergoing testing

Cross-reactivity related issues based on population and individual characteristics were reported in nine studies (6, 11, 12, 27, 32, 36, 40, 42). Results for potential interactions could be biological due to antibodies from inter-current infections, adverse environmental exposure to assay components, HIV subtype, or interactions between test kits in an algorithm. One study suggested cross-reactivity between test kits within an algorithm led to final false-positive result (42).One study hypothesized that cross-reactivity may cause weak reactive test results (32). Six studies (6, 11, 27, 35, 36, 40) reported potential issues with RDTs interacting with particular characteristics of the individual undergoing testing (6, 35, 36) including individuals with low levels of HIV-1/2 antibodies (11, 27)—particularly due to late stage HIV infection (27), and potential cross-reactivity or environmental exposure to test kits components (36, 40).

Poor management and supervision

Twenty studies reported poor management and supervision practices related to poor quality testing and potential misdiagnosis of HIV status (8, 14, 16, 18-20, 23, 25, 26, 28, 29, 32, 33, 35, 39, 41, 42, 46, 47, 49). Of these, 10 studies specifically reported on improper practices for management of supplies (8, 16, 26, 28, 29, 33, 35, 39, 41, 42). Errors that were specifically linked to misdiagnosis included the use of damaged RDT test kits (35); the use of damaged or expired kits (8, 33, 41, 42); stock outs (8, 26, 29, 33, 39, 41); the use of wrong RDT for HIV testing (28); and not using a timer nor the correct reagent with RDT selected (26). According to one study from Tanzania, an assessment of non-laboratory personnel competency identified that a syphilis RDT was being used incorrectly to test for an HIV infection (28).

Errors related to poor management and supervision also include clients receiving HIV testing within the window period (32, 46), HIV testing performed by under-trained or ineligible staff (8, 23, 26, 28, 49), poor quality instructions for use (32), low levels of re-testing before ART initiation (19, 20), poor participation in EQA schemes (39), poor site-level supervision (18), lack of and poor adherence to standard operating procedures (14, 23, 25, 26, 29, 33). Four studies also reported that user and/or clerical errors were linked to high workload and stress (8, 35, 39, 47). An EQA scheme in Nigeria reported that over 40% of laboratory respondents had no training on HIV testing and only 33.3% of sites used an approved national testing algorithm (23). In South Africa, an assessment of 38 facilities reported that compliance to standard operating procedures was very low and that out of all HIV testing events observed (n=265) only 3.4% were performed correctly (26). They reported common causes for incorrectly performing processes included: a deviation of staff roles; lack of necessary equipment, or lack of useable equipment; and inadequate training (26).

14.4. Discussion

This review found limited data on the magnitude of misdiagnosis of HIV status. The review only identified two studies that specifically reported the number of individuals who were actually HIV-negative but who had been misdiagnosed and enrolled in HIV-related care and treatment (6, 35). In both studies substantial loss to follow-up was documented and not all clients diagnosed HIV-positive were retested. Thus, it is possible that these rates are under-estimates and the actual level of misdiagnosis may be even higher than what was reported. Although none of the reviewed studies were able to determine and quantify the exact cause(s) of misdiagnosis, several provided hypotheses that attributed misdiagnosis to various factors. Three key factors emerged in this review: (1) poor management of supplies including the use of inappropriate, damaged or expired test kits, (2) user errors interpreting weak reactive test results and (3) the use of a "tiebreaker" testing strategy.

Poor management of supplies appears in half (10/20) of studies reporting poor management and supervision. To ensure correct results, all staff providing HIV testing must be adequately trained and certified. In several studies this was not the case and ineligible providers were performing HIV testing services (26, 28). Training, including pre-service, in-service and periodic refresher training is important to maintain and improve the quality of services, as well as participation in proficiency testing schemes. In this review the studies that participated in such programmes were able to monitor and track their performance and identify issues, including difficulty interpreting faint lines and health worker shortages that needed to be addressed to improve the quality of HIV testing. In addition to training

personnel and participating in external quality assurance, overarching management and supervision of sites is essential. Several studies reported user and clerical errors were a result of demanding workloads, burnout and high levels of stress (8, 35, 39, 47). Sites should routinely assess and manage their human resource planning to prevent and reduce staff shortages as much as possible.

User error interpreting faint lines and weak reactive test results was also a common challenge identified, often leading to substantial levels of false-positive results. This issue has been reported and evaluated previously. Historically this issue has been well-documented and evaluated in the literature (32, 50); however the challenges and errors with over-interpretation persist. To address this issue specialized training for health workers may be needed, as well as work with manufacturers to improve instructions for use on the interpretation of faint lines and site-level standard operating procedures to guide health workers on how to handle faint lines when they occur. No assay or testing algorithm is perfect and false-positive and false-negative test results do occur and should be anticipated. However, this review highlights errors and practices that compromise the reliability test results; in particular the use of a tiebreaker testing strategy. Based on the studies reviewed, it appeared that a tiebreaker increases the likelihood of false-positive test results and possible misdiagnosis. Although our review did not examine possible cross-reactivity between assays used within an algorithm, this also may have contributed to false-positive results. Since our search was completed, two more studies reported on the use of a tiebreaker strategy. In Ethiopia, a tiebreaker testing strategy resulted in 16 false positive diagnoses (PPV 92.7%, 95% CI: 88.4%-95.8%), compared to a serial testing strategy with 1 false positive diagnosis (PPV of 99.5% (95% CI 97.3%-100%)) (51). In a multi-country study among African adults, using a tiebreaker testing strategy with RDTs resulted in a poor sensitivity within the algorithm, ranging from 8.3%-43% (52).

These results are concerning because out of 48 recent national policies reviewed, 30% recommend using a tiebreaker testing strategy (5). It is important that national programmes and policy makers select and use an appropriate testing strategy and validated testing algorithm. A tiebreaker testing strategy has never been recommended by WHO, but it appears to be common practice in many settings. To ensure that test results and HIV status are reported correctly it is important WHO recommendations emphasize re-testing people diagnosed HIV-positive before they enrol in care and/or treatment. This is increasingly critical as more people are being offered immediate treatment without being clinically assessed.

14.5. Limitations

This review suffers from publication bias as it focused on identifying studies related to the misdiagnosis of HIV status in the published and gray literature, including information from programme evaluations. However, this information is often not publically available. To find additional information to corroborate the concerns regarding misdiagnosis we also reviewed available data external quality assessment (proficiency testing and on-site supervision) and studies reporting sensitivity and specificity HIV testing algorithms which also reported on causes of false-positive and false-negative results. Nevertheless, there are several limitations to this review.

This review was limited to the English language and therefore may have missed reports in other languages. Because of the heterogeneity of studies we were unable to assess the quality of studies more rigorously, as reports were all observational and no the primary outcome for most studies reviewed. Likewise, because the review was designed to identify reports on misdiagnosis it is possible that studies only reporting on errors and quality of HIV testing may have been missed. While some reports of cross-reactivity were identified in this review, we focus on human errors and quality system failures. Other reviews exist and provide more details information on such issues (53). Furthermore, we did not identify or examine reports of incorrect test results or misdiagnosis of HIV status due to testing people exposed to ARV drugs, e.g. ART, pre-exposure prophylaxis or post-exposure prophylaxis.

14.6. Conclusion

The paucity of data regarding the magnitude of misdiagnosis within HIV testing programmes does not make it possible to determine whether misdiagnosis is not reported because it is negligible or because it is not monitored. Nevertheless, our review suggests that there are many potential factors and practices which could contribute to misdiagnosis and insufficient attention is given to quality assurance in many settings.

The consequences of misdiagnoses are serious on an individual and public health level. Ensuring the quality of HIV testing services and that HIV diagnoses are correct is an urgent priority. Any site providing HIV testing must have

quality management systems to minimize the risk of incorrect test results and to identify and correct misdiagnoses if they occur. Without such provisions there is a risk the many successes of national programmes and the HIV response as a whole could be compromised and undermined. With the momentum to increase diagnosis of people with HIV and link them to care and treatment, a parallel push to improve quality and prevent misdiagnosis is essential.

References

1. Annex 6. Diagnostics for HIV diagnosis of the Consolidated Guidelines on HIV testing services. Geneva: World Health Organization; 2015.
2. Molesworth AM, Ndhlovu R, Banda E, Saul J, Ngwira B, Glynn JR, et al. High accuracy of home-based community rapid HIV testing in rural Malawi. J Acquir Immune Defic Syndr. 2010 Dec 15;55(5):625-30.
3. Plate DK, Rapid HIVTEWG. Evaluation and implementation of rapid HIV tests: the experience in 11 African countries. AIDS Res Hum Retroviruses. 2007 Dec;23(12):1491-8.
4. Lyamuya EF, Aboud S, Urassa WK, Sufi J, Mbwana J, Ndugulile F, et al. Evaluation of simple rapid HIV assays and development of national rapid HIV test algorithms in Dar es Salaam, Tanzania. BMC Infectious Diseases. 2009;9.
5. Flynn D, Johnson C, Sands A, Wong V, Baggaley R. An analysis of the role of lay providers in HIV testing and counseling in 48 countries. In press unpublished data.
6. Klarkowski DB, Wazome JM, Lokuge KM, Shanks L, Mills CF, O'Brien DP. The evaluation of a rapid in situ HIV confirmation test in a programme with a high failure rate of the WHO HIV two-test diagnostic algorithm. PLoS ONE. 2009;4(2).
7. Shanks L, Klarkowski D, O'Brien DP. False positive HIV diagnoses in resource limited settings: operational lessons learned for HIV programmes. PLoS One. 2013;8(3):e59906.
8. Kalou M, Mosha F, Ndlovu N, Jackson K, Mwassekage M, Ngoi M, et al. Implementation of a Proficiency Testing Program Using Dried Tube Specimen (DTS) Approach to Improve the Quality of HIV Rapid Testing in Testing Sites in Uganda and Tanzania. First International Conference of the African Society for Laboratory Medicine (ASLM); 2012; Cape Town, South Africa. Abstract No. 69. 2012.
9. 90-90-90 Ambitious treatment targets: writing the final chapter of the AIDS epidemic - a discussion paper. Geneva: Joint United Nations Programme on HIV/AIDS; 2014.
10. da Costa Ferreira Junior O, Ferreira C, Riedel M., Visinoni Widolin, Sasazawa Ito, Westman S, et al. Field evaluation of an HIV rapid test algorithm for the rapid diagnosis of HIV infection in Brazil. HIV Pathogenesis, Treatment and Prevention; 2007; Sydney, Australia. 2007.
11. Benzaken AS, Bazzo ML, Galban E, Pinto ICP, Nogueira CL, Golfetto L, et al. External quality assurance with dried tube specimens (DTS) for point-of-care syphilis and HIV tests: Experience in an indigenous populations screening programme in the Brazilian Amazon. Sexually Transmitted Infections. 2014;90(1):14-8.
12. Aghokeng AF, Mpoudi-Ngole E, Dimodi H, Atem-Tambe A, Tongo M, Butel C, et al. Inaccurate diagnosis of HIV-1 group M and O is a key challenge for ongoing universal access to antiretroviral treatment and HIV prevention in Cameroon. PLoS ONE. 2009;4(11).
13. Galiwango RM, Musoke R, Lubyayi L, Ssekubugu R, Kalibbala S, Ssekweyama V, et al. Evaluation of current rapid HIV test algorithms in Rakai, Uganda. J Virol Methods. 2013 Sep;192(1-2):25-7.
14. Kanal K, Chou TL, Sovann L, Morikawa Y, Mukoyama Y, Kakimoto K. Evaluation of the proficiency of trained non-laboratory health staffs and laboratory technicians using a rapid and simple HIV antibody test. AIDS Research and Therapy. 2005;2(1).
15. Manyazewal T, Paterniti AD, Redfield RR, Marinucci F. Role of secondary level laboratories in strengthening quality at primary level health facilities' laboratories: an innovative approach to ensure accurate HIV, tuberculosis, and malaria test results in resource-limited settings. Diagn Microbiol Infect Dis. 2013 Jan;75(1):55-9.
16. Tegbaru B, Wolday D, Messele T, Meless H, Kassa D, Tesema D, et al. Assessment of the implementation of HIV-rapid test kits at different levels of health institutions in Ethiopia. Ethiop Med J. 2007 Jul;45(3):293-9.
17. Stetler HC, Granade TC, Nunez CA, Meza R, Terrell S, Amador L, et al. Field evaluation of rapid HIV serologic tests for screening and confirming HIV-1 infection in Honduras. AIDS. 1997;11(3):369-75.
18. F. Kitheka, M. Umuro, S. Mwa, editors. Implementing Rapid HIV Proficiency Testing Program in Kenya: Key Findings from Corrective Actions Following the Switch from Facility-Based to Individual-Based Proficiency Testing. First International Conference of the African Society for Laboratory Medicine (ASLM); 2012; Cape Town, South Africa. Oral Poster on Wednesday, Dec. 5 2012.

19. Malawi Go. Integrated HIV program report January - March 2014. In: Health Mo, editor. Malawi; 2014.
20. Malawi Go. Integrated HIV program report April-June 2014. Malawi: Ministry of Health; 2014.
21. Young PW, Mahomed M, Horth RZ, Shiraishi RW, Jani IV. Routine data from prevention of mother-to-child transmission (PMTCT) HIV testing not yet ready for HIV surveillance in Mozambique: A retrospective analysis of matched test results. BMC Infectious Diseases. 2013;13(1).
22. E. Iwe, I. Livinus, E. Omoregbe, B. Okoye, C. Nwuba, A. Mbanefo, et al. Preparation and use of dry tube specimens (DTS) method by laboratorians and non- laboratorians for quality assurance of HIV services (HCT) in MSH supported sites, Niger State, North-Central Nigeria. 6th IAS Conference on HIV Pathogenesis and Treatment; 2011; Rome, Italy. Abstract no. CDD076. 2011.
23. A. Adebayo, A. Olufemi, B. Adekunle, O. Louisa. A Limited survey on HIV-testing knowledge and practice of laboraty personnel at some point-of-care services in Abuja, Federal Capital Territory (FCT), Nigeria. A Limited survey on HIV-testing knowledge and practice of laboraty personnel at some point-of-care services in Abuja, Federal Capital Territory (FCT), Nigeria; 2012; Cape Town, South Africa. Oral Poster on Wednesday, Dec. 5 2012. 2012.
24. Bassett IV, Chetty S, Giddy J, Reddy S, Bishop K, Lu Z, et al. Screening for acute HIV infection in South Africa: Finding acute and chronic disease. HIV Medicine. 2011;12(1):46-53.
25. Wolpaw BJ, Mathews C, Chopra M, Hardie D, de Azevedo V, Jennings K, et al. The failure of routine rapid HIV testing: a case study of improving low sensitivity in the field. BMC health services research. 2010;10:73.
26. SEAD. Analysis of POCT/VCT performed at South African primary health care clinics. 2010 [updated 2010; cited 2014 July 14]; Available from: http://www.sead.co.za/downloads/POCT-clinics-2011.pdf
27. Kahemele N.J., et al. HIV infection diagnostic uncertainity, a field experience. AIDS 2008 - XVII International AIDS Conference; 2008; Mexico City, Mexico. Abstract no. CDB0362. 2008.
28. D. Ocheng, S. Mattaro, L. Msilama, A. Msauka, M. Khama, N. Mshanga, et al. Establishing competence in HIV rapid testing among non-laboratory health professional is key to tasking: preliminary results from ten regions of Tanzania. 19th International AIDS Conference; 2012; Washington, DC. Abstract no. MOPE682. 2012.
29. Mashauri FM, Siza JE, Temu MM, Mngara JT, Kishamawe C, Changalucha JM. Assessment of quality assurance in HIV testing in health facilities in Lake Victoria zone, Tanzania. Tanzania health research bulletin. 2007 May;9(2):110-4.
30. Baveewo S, Kamya MR, Mayanja-Kizza H, Fatch R, Bangsberg DR, Coates T, et al. Potential for false positive HIV test results with the serial rapid HIV testing algorithm. BMC research notes. 2012;5:154.
31. Eller LA, Eller MA, Ouma BJ, Kataaha P, Bagaya BS, Olemukan RL, et al. Large-scale human immunodeficiency virus rapid test evaluation in a low-prevalence ugandan blood bank population. Journal of clinical microbiology. 2007 Oct;45(10):3281-5.
32. Gray RH, Makumbi F, Serwadda D, Lutalo T, Nalugoda F, Opendi P, et al. Limitations of rapid HIV-1 tests during screening for trials in Uganda: diagnostic test accuracy study. 2007.
33. W. Lali, G. Guma, P. Awongo, Z. Akol, H. Kajumbula, P. Namupijja, et al. Challenges of implementation of integrated national laboratory quality improvement in Uganda. AIDS 2010 - XVIII International AIDS Conference 2010; Vienna, Austria. Abstract No. LBPE43 2010.
34. Matambo R, Dauya E, Mutswanga J, Makanza E, Chandiwana S, Mason PR, et al. Voluntary counseling and testing by nurse counselors: What is the role of routine repeated testing after a negative result? Clinical Infectious Diseases. 2006;42(4):569-71.
35. Shanks L, Klarkowski D, O'Brien DP. False Positive HIV Diagnoses in Resource Limited Settings: Operational Lessons Learned for HIV Programmes. PLoS ONE. 2013;8(3).
36. Klarkowski D, Glass K, O'Brien D, Lokuge K, Piriou E, Shanks L. Variation in specificity of HIV rapid diagnostic tests over place and time: An analysis of discordancy data using a bayesian approach. PLoS ONE. 2013;8(11).
37. Tchounga BK, Inwoley A, Coffie PA, Minta D, Messou E, Bado G, et al. Re-testing and misclassification of HIV-2 and HIV-1&2 dually reactive patients among the HIV-2 cohort of The West African Database to evaluate AIDS collaboration. J Int AIDS Soc. 2014;17(1):19064.
38. Jentsch U, Lunga P, Lacey C, Weber J, Cairns J, Pinheiro G, et al. The Implementation and Appraisal of a Novel Confirmatory HIV-1 Testing Algorithm in the Microbicides Development Programme 301 Trial (MDP301). PLoS ONE. 2012;7(9).

39. Cham F, Maleka M, Masango M, Goetsch E, Belabbes EH, Singh B, et al. The World Health Organization African region external quality assessment scheme for anti-HIV serology. African Journal of Laboratory Medicine. 2012;1(1).
40. Boeras DI, Luisi N, Karita E, McKinney S, Sharkey T, Keeling M, et al. Indeterminate and discrepant rapid HIV test results in couples' HIV testing and counselling centres in Africa. Journal of the International AIDS Society. 2011;14(1).
41. CDC. HIV testing for surveillance: Evidence from assessments of routine diagnostic testing in ANC. WHO HIV testing and counselling consultation. In press 2014.
42. Crucitti T, Taylor D, Beelaert G, Fransen K, Van Damme L. Performance of a rapid and simple HIV testing algorithm in a multicenter phase III microbicide clinical trial. Clinical and Vaccine Immunology. 2011;18(9):1480-5.
43. Learmonth KM, McPhee DA, Jardine DK, Walker SK, Aye TT, Dax EM. Assessing proficiency of interpretation of rapid human immunodeficiency virus assays in nonlaboratory settings: Ensuring quality of testing. Journal of Clinical Microbiology. 2008;46(5):1692-7.
44. Viani RM, Araneta MRG, Spector SA. Parallel rapid HIV testing in pregnant women at Tijuana General Hospital, Baja California, Mexico. AIDS Research and Human Retroviruses. 2013;29(3):429-34.
45. Martin EG, Salaru G, Paul SM, Cadoff EM. Use of a rapid HIV testing algorithm to improve linkage to care. Journal of Clinical Virology. 2011;52(SUPPL. 1):S11-S5.
46. Mehra B, Bhattar S, Bhalla P, Rawat D. Rapid Tests versus ELISA for Screening of HIV Infection: Our Experience from a Voluntary Counselling and Testing Facility of a Tertiary Care Centre in North India. Isrn Aids. 2014;2014:296840.
47. Louis FJ, Anselme R, Ndongmo C, Buteau J, Boncy J, Dahourou G, et al. Evaluation of an external quality assessment program for HIV testing in Haiti, 2006-2011. Am J Clin Pathol. 2013 Dec;140(6):867-71.
48. K.M. Sushi, G. Saraminijacob, G. Thatchinamoorthy. External quality assurance scheme in a national reference laboratory for HIV testing in resource limited settings, India. 6th IAS Conference on HIV Pathogenesis and Treatment; 2011; Rome, Italy. Abstract no. CDD032. 2011.
49. Granade TC, Parekh BS, Phillips SK, McDougal JS. Performance of the OraQuick and Hema-Strip rapid HIV antibody detection assays by non-laboratorians. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. 2004 Jul;30(3):229-32.
50. Sacks R., Omodele-Lucien A., Whitbread N., Muir D., Smith A. Rapid HIV testing using Determine HIV 1/2 antibody tests: is there a difference between the visual appearance of true- and false-positive tests? International journal of STD & AIDS. 2012 Sep;23(9):644-6.
51. Shanks L, Siddiqui MR, Kliescikova J, Pearce N, Ariti C, Muluneh L, et al. Evaluation of HIV testing algorithms in Ethiopia: the role of the tie-breaker algorithm and weakly reacting test lines in contributing to a high rate of false positive HIV diagnoses. BMC Infect Dis. 2015;15(1):39.
52. Fogel JM, Piwowar-Manning E, Donohue K, Cummings V, Marzinke MA, Clarke W, et al. Determination of HIV status in African adults with discordant HIV rapid tests. J Acquir Immune Defic Syndr. 2015 Mar 27.
53. Klarkowski D, O'Brien DP, Shanks L, Singh KP. Causes of false-positive HIV rapid diagnostic test results. Expert Review of Anti-Infective Therapy. 2014;12(1):49-62.

Appendix

Table 14.4A-14.9A. Summary table of all studies reporting on factors potentially related to misdiagnosis of HIV status

14.4A. Studies reporting misdiagnosis rates of HIV status and misdiagnosis of HIV type					
Source	**Location**	**Sample size**	**Factors identified potentially relating to misdiagnosis of HIV status**	**Algorithm used**	**Misdiagnosis rates**
Shanks et al., 2014	Bukavu, DRC Kinyinya, Burundi Abdurafi, Ethiopia	DRC: 914 clients diagnosed HIV-positive Burundi: 78 clients diagnosed HIV-positive Ethiopia: 149 clients diagnosed HIV-positive	• Nurses did not follow SOP • Workload and stress on HWs was high • 2 kits were "ruptured" • Some individuals may have been diagnosed on single assay (RDT or EIA)	**Serial algorithm initially used:** Determine HIV-1/2 (Alere Medical, Japan) and Capillus HIV-1/HIV-2 (Trinity Biotech, Ireland) **Parallel algorithm** used by the end of 2004: Determine HIV- 1/2 (Alere Medical, Japan) and Uni-Gold HIV (Trinity Biotech, Ireland)	**DRC:** 6 initial cases + 38 out of 54 suspected false positive diagnoses were found to be HIV-negative (calculated: 4.8% incorrectly classified) **Burundi:** Two (2.6%) individuals were HIV-negative on re-testing. **Ethiopia:** 191 individuals were identified and 149 (78.0%) retested. Seven (4.7%) were found to be false positive on the basis of RDTs, CombFirm and PCR.
Klarkowski et al., 2009	Bukavu, DRC	229 clients with two reactive HIV RDTs	• Parallel testing strategy • Cross-reactivity (because HLA class II antigens directly modulate the immune response, on-specific polyclonal B-lymphocyte antibodies) • Weak reactive results	**Parallel algorithm:** Determine HIV-1/2 (Alere Medical, Japan) as first assay and UniGold HIV (Trinity Biotech, Ireland) as second assay for all reactive specimens. In December 2005, algorithm changed: retesting of all positive results and confirmation testing by ImmunoComb HIV 1&2 CombFirm (Orgenics, Israel) and/or Western blot (WB).	• Dual reactive results: 365 (12.7%) • Tested at laboratory for confirmation: 229 • Negative/indeterminate by CombFirm & WB: 24 • Overall false-positive rate: 10.5% (95% CI 6.6–15.2) Note: Of 229 specimens, 212 (92.6%) were strong, and 17 (7.4%) were weak reactive by RDT. All 17 "weak reactives" were negative or indeterminate by WB and CombFirm, which gives a false-positive rate of 100% (95%CI 80.5–100).
Tchounga et al., 2014	Burkina Faso, Cote d'Ivoire and Mali	547 clients	• RDTs, Western blot and Multispot are not gold standard for discriminatory diagnosis of HIV-1 and HIV-2 • Testing algorithm and strategy for HIV-2 is suboptimal	National algorithms of all three countries used for initial testing (note: article does not provide the actual testing algorithms)	312 patients initially classified as HIV-2: • 267 (85.7%) confirmed as HIV-2 • 16 (5.1%) reclassified as HIV-1 • 9 (2.9%) reclassified as HIV-1&2 • Both tests gave discrepant results for 20 patients (6.4%) 235 patients initially classified as HIV-1&2 dually reactive: • 54 (23.0%) confirmed as dually reactive • 103 (43.8%) reclassified as HIV-1 monoinfected • 33 (14.0%) reclassified as HIV-2 monoinfected Both tests gave discrepant results for 45 patients (19.1%)

Table 14.5A. Evaluation of testing algorithms: sensitivity and specificity

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Misdiagnosis rates
Aghokeng et al., 2009	Yaounde, Cameroon	Scenario 1: 460 specimens Scenario 2: 490 specimens	• Cross-reactivity with HIV-1 group O infection • Poor quality testing algorithm (Determine followed by ImmunoComb II) for confirming HIV-positive results and resolving indeterminate results	**Serial algorithms:** **Algorithm** 1: **First assay:** Retrocheck HIV (Qualpro Diagnostics, India) **Second assay:** SD Bioline HIV 1/2 3.0 (Standard Diagnostics Inc., South Korea) **Algorithm 2:** **First assay:** Determine HIV-1/2 (Alere Medical, Japan) **Second assay**: ImmunoComb II HIV 1&2 Biospot (Orgenics LTD., Israel).	**Algorithm 1**: **Retrocheck/SD Bioline:** Algorithm sensitivity: 94.7% Algorithm specificity: 98.8% **Potential misdiagnosis:** 6 out of 100 persons would be falsely identified as HIV-negative, and 2 out of 100 persons would be falsely identified as HIV-positive. **Algorithm** 2: **Determine/ImmunComb II:** Algorithm sensitivity: 100%, Algorithm specificity: 91.5%
da Costa et al., 2007	Brazil	1 151 clients	• Tiebreaker • Technical mistake • Problems found with operating RDT, including sample coagulation between the time from finger pricking to sample application to the RDT.	**Strategy Unclear:** "HIV diagnosis was made when the result of both screening assays [Determine HIV-1/2 (Alere Medical, Japan) and HIV Rapid Check (NDI-UFES, Brazil)] were concordant or, in the case of discordance, based on the result of the tiebreaker assay [Uni-Gold HIV (Trinity Biotech)]."	26 positive and 1,151 negative specimens were correctly diagnosed by the 3 RDT algorithm Tiebreaker used three times: in two cases the Determine assay was false-positive and in one case it was false-negative
Eller et al., 2007	Uganda	5 252 specimens	• Tiebreaker • Technical errors: possible errors include labeling or transcription mistakes and/or specimen mix ups	**Serial algorithm:** **First assay:** Uni-Gold Recombigen HIV (Trinity Biotech, Ireland) **Second assay:** HIV-1/2 STAT-PAK (Chembio Diagnostic Systems, USA) **Third assay ("tiebreaker"):** OraQuick HIV-1 (Orasure, USA; assembled in Thailand)	Sensitivity: 98.6% (95% CI, 91.4 to 99.9%); 1 false negative results Specificity: 99.9% (95% CI, 99.76 to 99.96%); 5 false positive results.
Galiwango et al., 2013	Rakai, Uganda	2 624 clients (2617 specimens)	• Tiebreaker—discordant results were tested with Uni-Gold (not listed as cause by authors) • Parallel testing strategy (not listed as cause by authors) • Weak reactive results	**Serial and parallel algorithms assessed using:** **First assay:** Determine HIV-1/2 (Alere Medical , Japan) OR HIV 1/2 Stat-Pak® Dipstick (Chembio Diagnostic Systems, USA) **Second assay:** Uni-Gold HIV (Trinity Biotech, Ireland)	• **Determine as a screening test:** Sensitivity: 97.3% (95% CI 96.0–98.3). Specificity: 99.9% (95% CI 99.6–99.99%) **When weak reactives considered reactive: Sensitivity: 97.4 (95% CI 96.1–98.4); false positive rate increased • **Stat-Pak as a screening test:** Sensitivity: 97.4 (95% CI 96.1–98.4) Specificity: 99.7% (95% CI 99.3–99.9%) **When weak reactives considered reactive: Specificity decreased from 99.7% to 96.9%

Table 14.5A. Evaluation of testing algorithms: sensitivity and specificity

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Misdiagnosis rates
Klarkowski et al., 2013	Uganda, Zimbabwe, Myanmar, India, Ivory Coast, Haiti, Ethiopia, DRC, Congo, CAR	106 035 clients, 51 sites	• Cross-reactivity across geographies using different RDTs	**All countries used parallel testing strategy** **CAR:** Determine HIV-1/2, Unigold® **Congo-B:** Determine HIV-1/2, Unigold® **DRC:** Determine HIV-1/2, Unigold® **Ethiopia:** Determine HIV-1/2, Unigold® **Haiti:** Determine HIV-1/2, Unigold® or Capillus® **India:** Determine HIV-1/2, Unigold® or TriDot® **Ivory Coast:** Determine HIV-1/2, Unigold®or Hexagon® **Myanmar:** Determine HIV-1/2, Unigold® or Capillus® **Uganda:** Determine HIV-1/2, Unigold® **Zimbabwe:** Determine HIV-1/2, First Response®or SD Bioline®	**CAR**: Out of 4, 329 patients tested, 143 (3.3%) had discordant results. **Congo**: Out of 438 patients tested, 26 (5.9%) had discordant results. **DRC**: Out of 19, 065 patients tested, 119 (0.6%) had discordant results. **Ethiopia**: Out of 4, 961 patients tested, 363 (7.3%) had discordant results. **Haiti**: Out of 2, 006 patients tested using Determine & Unigold, 20 (1.0%) had discordant results; and out of 4, 458 patients tested with Determine & Capillus 54 (1.2%) had discordant results. **India**: Out of 1, 193 patients tested with Determine & Unigold 19 (1.6%) had discordant results; and out of 4, 266 patients tested with Determine & TriDot, 59 (1.4%) had discordant results. **Ivory Coast**: Out of 3, 386 patients tested using Determine & Unigold, 125 (3.7%) had discordant results; and out of 5, 648 patients tested using Determine & Hexagon 334 (5.9%) had discordant results. **Myanmar**: Out of 14, 796 patients who were tested using Determine & Unigold 331 (2.2%) had discordant results; and out of 25, 024 patients tested using Determine & Capillus 528 (2.1%) had discordant results. **Uganda**: Out of 6, 056 patients tested, 129 (2.1%) had discordant results. **Zimbabwe**: Out of 7, 782 patients tested using Determine & FirstResponse 18 (0.2%) had discordant results; and out of 2, 627 patients tested using Determine & SD Bioline, 11 (0.4%) had discordant results
Martin et al. 2011	USA	787 specimens	• Operator error: under-interpretation of weak reactive • Using high prevalence testing strategy in low prevalence setting (not reported by authors)	**Serial testing algorithm:** **First assay**: Clearview® HIV 1/2 STAT-PAK (or Orasure OraQuick ADVANCE® **Second test**: Unigold (Trinity Biotech Plc, Ireland) All samples reactive on at least one RDT confirmed through Western blot at the New Jersey Public Health and Environmental Laboratories.	426 reactive rapid test results; 394 (92.5%) were reactive by a second rapid test, 32 (7.5%) had a negative second rapid test. WB confirmed 368 of 369 positive RTA results (99.5%) but only 87.0% of all clients with two consecutive reactive RDTs were confirmed HIV-positive. There were a total of 32 discordant cases. Two false positive result occurred: 1 was false positive following two reactive RDTs the Western blot confirmed an individual was not HIV-positive, the other was identified among one client who refused Western blot confirmation.
Mehra et al., 2014	North India	1 672 specimens	• Testing in window period, low antibody titre • Possible technical errors	**Serial testing algorithm:** **First assay:** SD Bioline HIV-1/2 3.0 (SD Biostandard Diagnostics Private Limited, India). **Second assays:** Pareekshak HIV-1/2 Triline card test (Bhat Bio-Tech India Private Limited, Bangalore, Karnataka, India), and Pareekshak HIV 1/2 rapid test kit (Trispot) (Bhat Bio-Tech India Private Limited, Bangalore, Karnataka, India).	The first RDT had missed 9 (22.5%) HIV reactive samples (also confirmed to be positive by western blot) and its sensitivity on comparison with ELISA was 77.5%. SD Bioline HIV-1/2 3.0 registered 5 false positive results (negative by ELISA and western blot) giving a specificity of 99.3%. The negative and positive predictive values of SD Bioline HIV-1/2 3.0 were 98.8% and 86.1%, respectively

Table 14.5A. Evaluation of testing algorithms: sensitivity and specificity

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Misdiagnosis rates
Stetler et al., 1998	Honduras	3 375 clients	• Technical errors	**Strategy unclear:** Retrocell (Abbott Laboratories, USA), Genie 1/2 (Bio-Rad), Multispot(Sanofi Pasteur Diagnostics, France) HIVCHEK 1 + 2 (Ortho Diagnostics, France)	Specimens: HIV-1-positive (n = 37), HIV-1-negative (n= 1118), and inconclusive (n =11) Two sera, positive by WB in Honduras and at CDC but were non-reactive in most simple/rapid tests in Honduras, were found to be reactive by the same simple/rapid assays when retested at CDC. The sensitivity of the three test methods as performed in Honduras (field sensitivity= 99.3%)
Viani et al., 2013	Baja California, Mexico	787 specimens	• Parallel testing strategy (not listed as cause by authors)	**Parallel testing algorithm:** Determine HIV-1/2 (Alere Medical , Japan) and Uni-Gold Recombigen HIV (Trinity Biotech Plc, Ireland)	24/25 pregnant women testing positive by parallel rapid HIV testing had a positive confirmatory western blot and 1/25 (0.03%) was confirmed as false positive Two (0.06%) women had parallel rapid HIV discordant testing results; both tested negative by western blot. Algorithm sensitivity: 100% Algorithm specificity: 99.9%
Baveewo et al., 2012	Kampala, Uganda	3 388 clients	• Tie-breaker • Weak reactives • DNA PCR QL test may result in occasional false-positives but not false-negatives	**Serial algorithm:** **First assay:** Determine (Abbott Laboratories, IL, USA) **Second assay:** STAT-PAK (Chembio Diagnostics) **Third assay ("tie-breaker"):** Uni-Gold Recombigen According to national algorithm, specimens testing positive on Determine, negative on STAT-PAK and positive on Uni-Gold would be reported as positive and not retested, but this study retested specimens	Of 1275 individuals who tested positive on Determine, 984 also tested positive on STAT-PAK, while 291 were negative on STAT-PAK and were retested with Uni-Gold Of 291 retested with Uni-Gold, 262 tested negative and were reported as negative; and 29 tested positive Of these 29 individuals, 14 (48.2%) tested HIV-negative using DNA PCR QL.
Crucitti et al., 2011	Sites in Cotonou, Benin; Chennai, India; Durban, South Africa; Kampala, Uganda; Karnataka, India	1 398 clients (5 734 specimens)	• Tie-breaker (not listed as cause by authors) • Use of deteriorated Determine HIV-1/2 test strips--test strips were not stored properly and were exposed to humidity at this test site • Cross-reactivity with SD Bioline HIV-1/2 3.0 assay had a false reaction leading to a final false-positive result • Weak reactives misinterpretation (over-interpretation)	**Serial algorithm:** **First assay:** Determine HIV-1/2 (Alere Medical, Japan) **Second assay:** SD Bioline HIV-1/2 3.0 (Standard Diagnostics, South Korea) **Third assay ("tie-breaker"):** Uni-Gold HIV (Trinity Biotech, Ireland) The final result scored a specimen as HIV-positive if 2 of 3 assays were reactive. In this case, a second specimen was collected and tested using the same algorithm. The participant was informed of her positive serostatus only if a second positive result was obtained.	46 specimens tested positive according to the sequential testing algorithm, with all but one specimen reactive with both Determine HIV-1/2 and SD Bioline HIV-1/2 3.0 Confirmatory results were available for all 46 women, with 42 (91.3%) confirmed to be HIV-1 positive and 4 (8.7%) confirmed to be HIV-negative

Table 14.5A. Evaluation of testing algorithms: sensitivity and specificity

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Misdiagnosis rates
Gray et al., 2007	Rakai, Uganda	1 517 clients	• Weak reactives • Tie-breaker (Determine, StatPak, Unigold) • Poor quality manufacturer instructions stating "instructions for the Determine test state that any red in the patient's window should be interpreted as positive" • Potential cross-reactivity causing weak bands	**Serial algorithm:** **First assay:** Determine HIV-1/2 (Alere Medical, Japan) **Second assay:** HIV-1/2 Stat-Pak (Chembio Diagnostic Systems, USA). **Third assay ("tie-breaker"):** Uni-Gold Recombinant HIV-1/2 (Trinity Biotech, Ireland)	Algorithm sensitivity: 97.7% (95% CI: 94.1% -99.4%) Algorithm specificity: 90.4% (95% CI: 88.7%-91.9%) 129 /295 positive results from algorithm were false-positives (43.7%) 4/1222 negative results were false-negatives (0.3%) Of sub-set of 125 samples tested positive for HIV with recorded band intensities, 37 were classified as weak reactive bands (5.8%) Exclusion of 37 samples with weak reactive bands improved specificity (99.6%, 98.6% to 100.0%) and reduced rate of false-positive results to 2.3% (2/86)
Jentsch et al., 2012	3 research sites in South Africa; 1 each in Tanzania, Uganda, and Zambia	9 385 clients	• Tie-breaker • Possible errors, method just introduced in Mwanza • Technical errors, misinterpretation of results could not be excluded	**Parallel testing algorithms**: *Durban. South Africa:* Determine HIV 1/2 (Alere Medical, Japan) and OraQuick Advance HIV 1/2 (OraSure Technologies, USA) *JHB, Mtubatuba, South Africa:* Determine HIV 1/2 and Uni-Gold HIV *Zambia:* Determine HIV 1/2 and Genie II HIV-1/HIV-2 (BioRad)	Of 537 participants with positive rapid tests after enrolment, 66 were discordant and 471 were positive on both rapid tests. The stored sera and BC samples from all were tested according to the HIV confirmatory algorithm: • 419 (78%) were confirmed as HIV infected immediately and 63 (12%) were confirmed as HIV infected at screening or enrolment • 55 (10%) were confirmed to be HIV uninfected

Table 14.6A. Comparing surveillance data to programmatic data

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Issues identified with rapid testing related to potential misdiagnosis
CDC unpublished 2014	9 unidentified countries Countries A & C-E used DBS specimen Countries B, F, H & I used Serum specimen Country G used Serum/Plasma	20 563 clients	• User errors • Misinterpretation of results • Poor record keeping/ errors in recording data • Problems with stock management (use of exposed, expired or invalid kits) • Use of incorrect algorithm--algorithm not specific	**Country A** *PMTCT Algorithm:* **First assay:** Determine HIV-1/2 (Alere Medical, Japan) **Second assay:** Unigold HIV (Trinity Biotech, Ireland) *ANC Surveillance Algorithm:* **First assay:** Vironostika (bioMérieux, France) **Second assay:** Murex (DiaSorin, UK) **Country B** *PMTCT Algorithm:* **First assay:** Determine HIV-1/2 (Alere Medical, Japan) **Second assay**: Unigold HIV (Trinity Biotech, Ireland) **Third assay**: ELISA (unspecified) *ANC Surveillance Algorithm:* **First assay** Vironostika (bioMérieux, France) **Second assay:** Murex (DiaSorin, UK) or Enzygnost (Siemens Healthcare Diagnostics) **Country C** *PMTCT Algorithm:*	There were diagnostic positives and surveillance negatives in all countries-most were less than 1% but in countries C, E, and I these were > 1%. There was disagreement with diagnostic negatives and surveillance positives. Countries D, F, G and H reported disagreement between ~75%-89%. The remaining countries reported at least 90% agreement. **Country A** using DBS specimens, PMTCT algorithm had a sensitivity of 90.6% and specificity of 99.1% **Country B** using serum, PMTCT algorithm had a sensitivity of 97.5% and specificity of 99% **Country C** using DBS specimen, PMTCT algorithm had a sensitivity of 93.8% and specificity of 98.5% **Country D** using a DBS specimen, PMTCT algorithm had a sensitivity of 75.9% and specificity of 99.4% **Country E** using DBS specimen, PMTCT algorithm had a sensitivity of 91.2% and specificity of 98.7%. **Country F** using serum specimen, PMTCT algorithm had a sensitivity of 83.2% and specificity of 99.2% **Country G** using serum specimen, PMTCT algorithm has as sensitivity of 79.8% and specificity of 99.3% **Country H** using serum, PMTCT algorithm had a sensitivity of 85.8% and specificity of 99.8% **Country I** using serum, PMTCT algorithm had a sensitivity of 88.8% and specificity of 98.6%. Information on the surveillance and diagnostic algorithms is missing and unclear in some settings.

				Parallel algorithm: UniGold HIV (Trinity Biotech, Ireland) or Determine HIV-1/2 (Alere Medica, Japan) and Colloidal Gold (Shanghai Kehue, China) **Second assay:** OraQuick ADVANCE HIV-1/2 Rapid Antibody Test (Bethlehem, USA) *ANC Surveillance Algorithm:* Vironostika (bioMérieux, France) & Murex (DiaSorin, UK) done in parallel testing strategy **Country D** *PMTCT Algorithms:* **First assay** Determine HIV-1/2 (Alere, Japan) **Second assay:** SD Bioline HIV-1/2 3.0 (Standard Diagnostics, South Korea) or UniGold HIV_(Trinity Biotech, Ireland) *ANC Surveillance Algorithm:* **First assay:** Vironostika (bioMérieux, France) **Second assay:** Murex (DiaSorin, UK) **Country E** *PMTCT Algorithms:* **First assay:** Determine HIV-1/2 (Alere, Japan) **Second assay**: SD Bioline HIV-1/2 3.0, (Standard	

				Diagnostics, South Korea) or Insti HIV-1/2 (Biolytical Laboratories) *ANC Surveillance Algorithm:* **First assay:** HIV EIA (AniLabsystems, Finland) **Second assay:** Enzygnost (Siemens Healthcare Diagnostics, Germany) Confirm all reactive test results on Western blot **Country F** *PMTCT Algorithms:* **First assay**: Determine HIV-1/2 (Alere Medical, Japan) **Second assay**. Immuno-Comb (Orgenics, Israel) *ANC Surveillance Algorithm:* **First assay:** Murex or Enzygnost **Second assay:** Immuno-Comb (Orgenics, Israel) or Hexagon (Humna GmbH, Germany) Confirm all reactive test results on Western blot **Country G** *PMTCT Algorithm:* **First assay**: Determine HIV-1/2 (Alere Medical, Japan) **Second assay**: Hexagon (Human GmbH) or OraQuick ADVANCE HIV-1/2 (Bethlehem, USA) *ANC Surveillance Algorithm:*	

				First assay: Determine HIV-1/2 (Alere Medical, Japan) **Second assay:** Hexagon (Human GmbH) or OraQuick ADVANCE HIV-1/2(Bethlehem, USA) **Country H** *PMTCT Algorithm:* **First assay**: Determine HIV-1/2 (Alere Medical, Japan) or OraQuick ADVANCE HIV-1/2 (Bethlehem,USA) **Second assay**: Colloidal Gold (Shanghai Kehua, China) *ANC Surveillance Algorithm:* **First assay:** Murex (DiaSorin, UK) **Second assay:** Collodial Gold (Shanghai Kehua, China) Confirm all reactive test results on Western blot **Country I** *PMTCT Algorithm:* **First assay**: Determine HIV-1/2 (Alere Medical , Japan) **Second assay**: UniGold HIV (Trinity Biotech, Ireland) or Clearview® HIV 1/2 STAT-PAK (Chembio Diagnostics Systems, USA) *ANC Surveillance Algorithm:* **First assay:** Architect HIV	

				Ag/Ab Combo Assay **Second assay:** DXI Analyzer	
Young et al., 2013	Mozambique	Not specified	• Testing errors • Data reporting problems	**<u>Serial testing algorithm</u>** *PMTCT Algorithm:* **First assay:** Determine HIV-1/2 (Alere Medical, Japan). **Second assay:** Uni-Gold HIV (Trinity Biotech, Ireland) **Discrepant results** are considered indeterminate and clients are asked to return after one month for retesting *ANC Surveillance Algorithm:* For those who consented, dried blood spot specimens were collected for centralized ANC surveillance testing in addition to routine PMTCT Testing	HIV prevalence from routine PMTCT testing was 14.4% versus 15.2% from surveillance testing (relative difference -5.1%; absolute difference -0.78%). Positive percent agreement (PPA) of PMTCT versus surveillance tests was 88.5% (95% CI: 85.7-91.3%), with 19 sites having PPA below 90%; Negative percent agreement (NPA) was 98.9% (CI: 98.5- 99.2%). No significant difference found among 3 regions (North, Center and South), however both PPA and NPA were significantly higher in 2009 than 2007 ($p < 0.05$).

Table14.7A. Misdiagnosis results from low-income countries

Source	Location		Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Issues identified with rapid testing related to potential misdiagnosis
Government of Malawi 2014	Malawi	7 007 clients diagnosed HIV-positive	• Re-testing before ART initiation not happening in most facilities • Parallel testing strategy (not listed as cause by authors)	**Parallel algorithm:** **First & second assays:** Determine HIV-1/2 (Alere Medical, Japan) and Uni-Gold (Trinity Bio-Tech, Ireland)	7.7% (n=534/7,007) of people previously diagnosed HIV-positive who were re-tested did not have concordant positive test results—and may have been misclassified as HIV-positive.
Government of Malawi 2014	Malawi	8 017 clients diagnosed HIV-positive	• Re-testing before ART initiation not happening in most facilities • Parallel testing strategy (not listed as cause by authors)	**Parallel algorithm:** **First & second assays:** Determine HIV-1/2 (Alere, Tokyo, Japan) and Uni-Gold HIV (Trinity BioTech, Ireland)	4.3% (346/8,017) of people previously diagnosed HIV-positive who were re-tested did not have concordant positive test results—and may have been misclassified as HIV-positive.

Table 14.8A. Studies focusing on false-negative results for rapid testing

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used	Potential misdiagnosis rates OR Sensitivity and specificity of algorithm
Bassett et al., 2011	Durban, South Africa	949 clients	• Parallel testing strategy	**Period of using a serial testing strategy** (March–August 2007), followed by a period of **parallel testing strategy** (September–November 2007). **RDTs used included:** • Determine HIV 1/2 (Alere Medical, Japan), • SmartCheck HIV 1&2 (World Diagnostic Inc., USA), • Sensa Tri-line HIV 1/2/0 (Hitech Healthcare Ltd, China), and • SD Bioline HIV-1/2 3.0 (Standard Diagnostics Inc., South Korea)	22 patients had negative or discordant rapid tests but positive qualitative HIV RNA testing. 10/22 patients were found to have chronic HIV Infection. These 10 patients had 'false negative' rapid HIV tests (10 chronically infected out of 976 with negative rapid tests; 1.0% false negative; 95% CI 0.6–1.9%). 12 patients with discrepant rapid HIV tests had positive qualitative HIV RNA tests. 10/ 12 patients were found to have chronic HIV infection with positive EIA and/ or WB (10 chronically infected out of 18 with discordant rapid tests; 56% false negative; 95% CI: 34–76%). 20/994 patients (2.0%; 95% CI 1.3–3.1%) with negative or discrepant rapid HIV test results were confirmed to have chronic HIV infection. Standard rapid HIV testing missed 2% of patients who had chronic HIV infection
Kahemele et al. 2008	Tanzania	3 clients	• Cross-reactivity, patient with late stage AIDS	Not specified	Noted three cases of missed HIV infection using a serial algorithm, including one case where a women was showing symptoms of HIV infection
Matambo et al., 2006	Harare, Zimbabwe	674 clients	• Operator error: repeat testing using same specimen "slow reactive"	**Parallel algorithm:** Determine HIV-1/2 (Alere Medical, Japan) and Unigold HIV (Trinity Biotech) Note: No further information on tests provided	1 false negative identified due to operator error (risk, 0.26%; 95% CI, 0.0065%– 1.4%).

Table 14.9A. Proficiency testing (PT)/external quality assessment (EQA) results

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	Algorithm used for PT/EQA	PT and/or EQA Results
Louis et al., 2013	Haiti	263 laboratories	• Inaccuracy in interpreting Capillus results and use of a single test to diagnose HIV • Staff turnover following the two major natural disasters	**Serial testing algorithm** adopted in 2006: **First assay:** Determine HIV1/2 (Abbott Laboratories, USA) or OraQuick ADVANCE Rapid HIV-1/2 (OraSure Technologies, USA) **Second assay**: Capillus HIV-1/HIV-2 (Trinity Biotech Plc, Ireland) Note: In 2011, the national algorithm was reviewed, and Capillus was replaced by a rapid immune-chromatographic assay, HIV (1+2) Antibody (Colloidal Gold) (KHB; Shanghai Kehua Bio-engineering, China)	From 2006 through 2011, 242 out of 263 unique laboratories enrolled in the PT program reported results at least once; response rates across laboratories ranged from 86% to 98% Between 2006 and 2011, results of proficiency testing were 93% concordant with reference results Increasing PT scores were observed from 2006 through 2011, with the lowest score (87.7%) observed in 2008 and the highest score (97.8%) in 2011
Manyazewal et al., 2013	Ethiopia	64 specimens	• Error in sampling techniques • Use of tiebreaker (not reported as cause by authors)	**Serial testing algorithm:** **First assay:** HIV (1+2) Antibody (Colloidal Gold) (KHB Shanghai Kehua Bio-engineering Co Ltd) **Second assay**: HIV 1/2 STAT-PAK (Chembio Diagnostics) **Third assay ("tie-breaker"):** Uni-Gold HIV (Trinity Biotech, Ireland)	Overall agreement of 95.3% (61/64) for HIV testing with RDTs and 100% (32/32) agreement when plasma samples were tested at health centres or district hospitals When whole blood specimens were tested, agreement was 87.5% (14/16) and 93.8% (15/16) for samples prepared by health centres and district hospitals
Sushi et al. 2011	India	9 419 specimens	• Mislabeling, • Contamination of specimens • RDT not performed correctly	All the samples were tested using HIV rapid test (Combaids) and positives alone were tested using Tridot and EIA Comb. (note: no further explanation of testing algorithm is provided)	Out of 9419 samples tested for QC, 9371(99.49%) reported correct results and 48(0.50%) discordant results. 26/48 samples (54.16%) were false positive, 22/48 (45.8%) false negative. For proficiency testing 91.8% reported test results. 645 (97.13%) reported correct results and 19 (2.86%) incorrect results. 7/ 19 (36.84%) were false positive and 12/19 (63.15%) false negative.
Tegbaru et al., 2009	Ethiopia	44 health institution laboratories	• Difficulty with weak reactives • Test kit & reagent shortages/ stock-out	**Serial algorithm**: **First assay:** Determine (test specifics not provided) **Second assay:** Capillus (test specifics not provided)	44 (100%) of the assessed laboratories correctly identified negative specimens 37/39 (94.8% conducted among labs that carried out confirmatory tests and reported results) labs correctly identified positive samples 43.6% and 38.7% reporting of false negatives when identifying weak reactives in first and second samples, respectively. 22 laboratories (50%) cited frequent shortages of test kits

Table 14.10A. Factors related to potential misdiagnosis, including quality issues identified during testing and during external quality assessment schemes (EQAS)

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	HIV tests used and setting were testing occurred	Methods for assessing quality and additional results
Benzaken et al., 2014	Brazil	1 608 specimens	• False negatives were in samples of very low antibody concentration	Standard Diagnostics (TR Syphilis 3.0-SD-Bioline) Note: this was the only POC HIV test mentioned in the article	All healthcare workers participating in the community-screening were trained. Used HIV and syphilis DTS panels developed by the reference laboratory, containing samples with negative and positive results at different antibody concentrations, for both infections. DTS panels were distributed to health workers in community for reconstitution and testing using HIV and syphilis RDTs. Results of testing were sent to the reference laboratory for marking and remedial action taken where necessary.
Cham et al., 2012	30 countries in Africa	16 proficiency testing panels, 49 laboratories	• Lack of human resources • Frequent stock outs of test kits, reagents and consumables for routine HIV testing • Faint bans / weak reactives reported • Some sites used one RDT to diagnose HIV • Poor uptake of EQA schemes by laboratories	8/49 laboratories confirmed that they use WHO Strategy II and III, for HIV prevalence testing with serial or parallel testing algorithms for diagnostic purposes. Three laboratories reported using WHO Strategy II with serial testing algorithm, whereas 21 laboratories reported using the serial testing algorithm only. Five laboratories reported using WHO testing strategy III, which differs from strategy II as it employs a third assay platform based on different antigen preparations and/or different test principles from assay platforms used in screening tests.	The EQA schemes (EQAS), established in March 2002, aim to assess the quality of anti-HIV-1 and HIV-2 serological testing for interested laboratories. Additionally, the REQAS allow comparison of testing facilities, in addition to evaluating the quality of serological testing using EIAs and RDTs. Participation in the EQAS is voluntary and at no cost to laboratories, thus allowing participation of laboratories in at least one of the EQAS components. The EQAS involved the distribution of proficiency testing (PT) panels to participating laboratories from 2002 to 2010. During the survey period, this included 16 distributions of PT panels to 49 laboratories in 30 countries. Laboratories were requested to complete a questionnaire on their use of internal quality control (IQC) materials. The purpose of the questionnaire was to determine whether IQC was routinely performed and if it was, whether laboratories were using test kit controls supplied in the test kit or 'in-house' produced controls.
Plate et al., 2007	Multiple countries in Africa	Not specified	• Tie-breaker	Not applicable	We identified 11 countries with a GAP country program in which evaluations of commercially available HIV RDTs had been conducted within the past 5 years. Collaborators in each country completed a standardized questionnaire about evaluation methods, test performance, algorithm development, implementation, and measures used to ensure quality of testing.
Iwe et al., 2011	Nigeria	122 specimens, 61 sites	• Typographical errors	Not reported	122 DTS panels were prepared in one of the district hospitals by 4 laboratory scientists and piloted in 61 Posts. DTS were distributed, 2 each to 61 posts (6 laboratories and 55 other points mostly manned by non-laboratory personnel. SOPs for rehydration and testing were used.

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	HIV tests used and setting were testing occurred	Methods for assessing quality and additional results
Kalou et al., 2012	Uganda and Tanzania	200 sites	• Stock outs / use of expired kits • Under-trained staff • Deviation from the testing algorithm • DTS reconstitution procedure and clerical errors due to high workload	Not reported	In 2010, the African Field Epidemiology Network, in collaboration with the Centers for Disease Control and Prevention (CDC) and the Ministries of Health (MOH) in Uganda and Tanzania piloted a DTS based PT programmes for HIV RDTs in 200 selected sites in each country. To ensure readiness of selected sites, training on DTS reconstitution and country HIV testing algorithms was provided to HIV testers. PT panels were distributed quarterly through local courier or postal services and results analyzed based on a 100% satisfactory performance score. Feedback was provided to all sites and supportive visits were conducted to ensure corrective measures were applied to facilities with performance score lower than 100%.
Kitheka et al., 2012	Kenya	365 individuals providing HIV testing	• Failure to adhere to written procedures and approved national algorithm • Transcription errors • Poor site level supervision and lack of job aides and timers • Test result interpretation challenges	Not reported	Between March-April 2012, teams from NHRL targeted 365 individual HIV testers countrywide with incorrect PT results. A standardized checklist was developed for recording testers' responses and the auditing teams' direct observation. Site visits conducted aimed at identifying weaknesses, challenges and gaps resulting in poor performance while providing on-site mentorship and corrective actions. Feedback was provided directly to testing personnel and facility management.
Lali et al., 2010	Uganda	122 laboratories	• Stock-out / expiry accounting for 11 (32.4%), • Technician factor 7 (20.5%), • Poor documentation 10 (29.4%) • Lack of standard operating 4 (11.7%) • Poor procedures 2 (5.8%)	Not reported	National reference laboratories for HIV, TB, and the public health prepared and validated the proficiency testing that were centrally distributed. A central team carried quality audit in the participating laboratories to resolve performance problems. National quality assessment database was developed at the Central Public Health Laboratories where data was analyzed using Epi-Info version 3.4.3. Performance indicators used were for laboratories with a score of 100%, sensitivity, specificity, predictive values positive and percentage mean scores. Factors causing poor performance were analyzed using Pareto Chart Analysis
Ocheng et al., 2012	Tanzania	100 proficiency testing samples, 3 835 individuals providing HIV testing, and 602 sites	• Ineligible providers performing HIV testing without license or training • Incompetent staff • Issuing invalid results as negatives • Using wrong RDT	Not applicable	Competency assessment of non-laboratory personnel was conducted throughout all 26 regions of Tanzania. Tester's biodata was recorded and the testers randomly selected and tested two samples from a panel of 100 proficiency testing samples. Results were interpreted and reported according to the National HIV testing algorithm. Assessors observed and scored performance using competency assessment tools. Feedback was provided on the same day and individuals who performed poorly received retraining on quality assurance procedures.

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	HIV tests used and setting were testing occurred	Methods for assessing quality and additional results
Boeras et al., 2011	Lusaka, Zambia and Kigali, Rwanda	45 820 clients (tested as couples)	• Tie-breaker • Cross-reactivity with antibodies from intercurrent infection with other pathogens or environmental exposure to test kit components, such as bovine products	Sequentially up to three RDTs. The four assays used included: First test: Determine HIV-1/HIV-2 (Abbott Laboratories, Abbott Park, IL) or First Response® HIV Card Test 1-2.0 (Premier Medical Corporation Ltd., Colonia, NJ) for screening, Second test: Capillus HIV-1/HIV-2 (Trinity Biotech, Ireland) and Uni-Gold™ HIV (Trinity Biotech, Ireland) for confirmatory and/or tie-breaker testing. At the follow-up visit, all three RDTs were performed again (screening, confirmatory and tie-breaker tests). If indeterminate/ discrepant results persisted longer than 2 months, HIV-1 RNA RT-PCR (Amplicor HIV-1 Monitor Test, v1.5, standard version, Roche Diagnostics, Indianapolis, IN) was performed. HIV-RNA viral load under 400 copies/mL (the lower limit of detection of the test) was considered "Uninfected" and a HIV-RNA viral load of >2000 copies/mL was considered "Infected".	Research performed at couples' voluntary counselling and testing) centres by Rwanda-Zambia HIV Research Group (RZHRG). 45,820 individuals tested as couples, 2.3% (4.1% of couples) had at least one discrepant or indeterminate rapid result. These individuals/couples were asked to retest monthly. A total of 65% of those individuals had follow-up testing. 661 clients returned for retesting to resolve their HIV status. There were 4/265 false-negative results and 17/37 false-positive results. Also a total of 68/214 received discrepant results again and could not have their HIV status resolved. Note: Information on testers not provided (although results were read by "trained technicians")
Granade et al., 2004	Atlanta, GA USA	153 clients	• Mislabeling & data entry errors • Training did not demonstrate how to perform RDT • The proper insertion of the hema-strip® test device into the buffer vial remained problematic despite the demonstration	OraQuick (OraSure Technologies Inc., Bethlehem, PA), and Hema-Strip (Chem-Bio Diagnostic Systems Inc., Medford,NY)	Testing sessions were conducted in small groups (not more than five members). Each group was given a brief summary of the goals of the project; an overview of the laboratory layout, specimen handling, and basic lab safety; proper clean-up and disposal procedures; and personal protective equipment (lab coat, face shield, and latex gloves). All participants received a short, written step-by-step procedure derived from the product insert which included diagrams and photographs; the complete product insert; a data recording form; and a post-test evaluation form to solicit their opinions of the experience. One-half of the participants received a short demonstration (<2 min per assay) of how to perform each. The rest were asked to rely solely on the written instructions.
Kanal et al., 2005	Cambodia	563 specimens	• Technical errors writing reports in laboratory	Determine (Abbot Laboratories) HIV1/2 test kits	In this study, those conducting HIV testing services were trained by experienced laboratory technicians in our centre on HIV testing using Determine (Abbot Laboratories) HIV1/2 test kits through a half-day training course, which consisted of use of a pipette, how to process whole blood samples, and how to read test result. The trained counsellors were midwives working for ANC and delivery ward in our centre without any experience on laboratory works. Laboratory technicians routinely did the same test and returned reports of the test results to counsellors. Reports by the counsellors and the laboratory technicians were compared, and discordant reports in two groups were re-tested with the same rapid test kit using the same blood sample.

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	HIV tests used and setting were testing occurred	Methods for assessing quality and additional results
Learmoth et al., 2008	192 laboratories from 26 countries	148 health workers	• Difficulty interpreting weak reactives	All laboratories participated in Serology EQAS of the National Serology Reference Laboratory, Australia The formats used for HIV testing included two rapid immunochromatographic assays, Abbott Determine HIV-1/2 (Abbott Laboratories, Abbott Park, IL) and Trinity Biotech Plc Uni-Gold HIV (Trinity Biotech Plc., Wicklow, Ireland); one rapid immunofiltration assay, Bio-Rad Multispot HIV-1/HIV-2 (Bio-Rad Laboratories, Inc., Redmond, WA); one rapid latex agglutination assay, Trinity Biotech Plc Capillus HIV-1/HIV-2 (Trinity Biotech Plc., Wicklow, Ireland); and one semirapid particle agglutination assay, Fujirebio Inc. Serodia HIV (Fujirebio Diagnostics, Inc., Tokyo, Japan).	Photographed rapid HIV test results were used in a modified external quality assessment scheme to assess the interpretation proficiency and, therefore, to assess the feasibility of using this method as a basis for a quality assessment program for non-laboratory based testing. Participants (n= 148), both experienced and inexperienced in the performance and interpretation of rapid HIV testing, interpreted the photographed results of five rapid HIV assays. These were scored according to the degree of technical discordance. Error scores were grouped according to each participant's technical experience.
Sacks et al., 2012	London, UK	39 laboratory personnel	• Weaker pigment uptake compared to true-positive POCT (false-reactives visually differ from true-reactives)	Determine™ HIV1/2 antibody POCT (Alere Limited, Stockport, UK)	Research performed at Jefferiss Wing, St Mary's Hospital, a hospital-based genitourinary medicine clinic. Seventeen false- and 17 true-positive serum samples were randomized into pairs, comprising one false- and one true-positive sample. Two independent readers identified each POCT as negative or positive and compared line strength between pairs. Six further readers graded line strength, 0-5, from POCT photographs.
Sushi et al., 2011	India		• Mislabeling, • Contamination of specimens • RDT not performed correctly	All the samples were tested using HIV rapid test (Combaids) and positives alone were tested using Tridot and EIA Comb. (note: no further explanation of testing algorithm is provided)	Out of 9419 samples tested for QC, 9371(99.49%) reported correct results and 48(0.50%) discordant results. 26/48 samples (54.16%) were false positive, 22/48 (45.8%) false negative. For proficiency testing 91.8% reported test results. 645 (97.13%) reported correct results and 19 (2.86%) incorrect results. 7/ 19 (36.84%) were false positive and 12/19 (63.15%) false negative.
Wolpaw et al., 2010	Cape Town, South Africa	Not specified	• Poor adherence to the recommended testing protocol	Not reported	The researchers and local health administrators collaborated to investigate the cause of the poor test performance and make necessary corrections. The clinic changed the brand of rapid test being used and later introduced quality improvement measures. Observations were made of the clinic staff as they administered rapid HIV tests to real patients. Estimated testing sensitivity was calculated as the number of rapid HIV test positive individuals detected by the clinic divided by this number plus the number of PCR positive, highly reactive 3rd generation ELISA patients identified among those who were rapid test negative at the clinic.

Source	Location	Sample size	Factors identified potentially relating to misdiagnosis of HIV status	HIV tests used and setting were testing occurred	Methods for assessing quality and additional results
Adebayo et al., 2012	Nigeria	39 laboratory personnel	• Poor to no training, no SOPS • 16 different combinations of algorithms involving 14 different RDT brands were observed with 13 of the 39 (33.3%) facilities applying the approved algorithm	Not applicable	The first conscious efforts of Nigeria Government at providing a national HIV-testing strategy occurred in 2006. Five non-cold-chain-dependent HIV-rapid test kits (RTKs) were recognized and combined pairwise, in parallel, to make eight different interim algorithms. This was adopted by stakeholders in early 2007. In late 2008, a survey of some health facilities was carried out in Abuja, Nigeria to evaluate the extent of compliance with these algorithms and some other laboratory quality essentials by HIV-testing laboratory personnel. Methods: 39 health facilities (15 public and 24 private) were randomly selected within AMAC. Structured questionnaire bordering on knowledge, skill, RTK supply sources and compliance with National Interim RTK Algorithms was self-administered to an HIV testing laboratory personnel in each site. Data analysis was done using epi-info.
Mashauri et al. 2007	Lake Victoria region of Tanzania	89 sites	• Few facilities with SOPS • Kits not readily available in some facilities	Not reported	89 health facilities (29 hospitals, 34 health centres, 9 dispensaries and 17 voluntary and counselling testing centres) were surveyed.
SEAD, 2010	South Africa	38 sites	• SOPS not complied with, inconsistent use of quality assurance • Lack of staff training, deviations of staff, intentional errors • Incorrect completion of registers and incubation time of tests • Stock-outs and absence of a timer, test strips and buffer • Incorrect specimen collection	Not reported	38 analysis sites selected, 28 were primary health clinics (PHC), and 10 were comprehensive community health centers

Table 14.11A. PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	4
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	4-5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	1-3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	(in protocol)
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4-5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	N/A
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4-5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A

Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	(Fig 1)
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	(Table 14.1A, Tables 14.4A-9A, & Table 14.4A)
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	(Fig 14.2A)
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	13-14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	14
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14