To err is human, to correct is public health: a systematic review examining poor quality testing and misdiagnosis of HIV status. Journal of the International AIDS Society, 20 (Suppl). pp. 7-18. ISSN 1758-2652 DOI: https://doi.org/10.7448/IAS.20.7.21755

Downloaded from: http://researchonline.lshtm.ac.uk/4328575/

DOI: https://doi.org/10.7448/IAS.20.7.21755

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by-nc-nd/2.5/
To err is human, to correct is public health: a systematic review examining poor quality testing and misdiagnosis of HIV status

Cheryl C. Johnson¹,²,§ Virginia Fonner³, Anita Sands⁴, Nathan Ford¹, Carla Mahklouf Obermeyer⁵, Sharon Tsui⁶, Vincent Wong⁷ and Rachel Baggaley¹

⁶Corresponding author: Cheryl C. Johnson, Department of HIV, World Health Organization, 20 Avenue Appia, Geneva, 1201, Switzerland. (Johnsonc@who.int)

Abstract

Introduction: In accordance with global testing and treatment targets, many countries are seeking ways to reach the “90-90-90” goals, starting with diagnosing 90% of all people with HIV. Quality HIV testing services are needed to enable people with HIV to be diagnosed and linked to treatment as early as possible. It is essential that opportunities to reach people with undiagnosed HIV are not missed, diagnoses are correct and HIV-negative individuals are not inadvertently initiated on lifelong treatment. We conducted this systematic review to assess the magnitude of misdiagnosis and to describe poor HIV testing practices using rapid diagnostic tests.

Methods: We systematically searched peer-reviewed articles, abstracts and grey literature published from 1 January 1990 to 19 April 2017. Studies were included if they used at least two rapid diagnostic tests and reported on HIV misdiagnosis, factors related to potential misdiagnosis or described quality issues and errors related to HIV testing.

Results: Sixty-four studies were included in this review. A small proportion of false positive (median 3.1%, interquartile range (IQR): 0.4-5.2%) and false negative (median: 0.4%, IQR: 0-3.9%) diagnoses were identified. Suboptimal testing strategies were the most common factor in studies reporting misdiagnoses, particularly false positive diagnoses due to using a “tiebreaker” test to resolve discrepant test results. A substantial proportion of false negative diagnoses were related to retesting among people on antiretroviral therapy.

Conclusions: HIV testing errors and poor practices, particularly those resulting in false positive or false negative diagnoses, do occur but are preventable. Efforts to accelerate HIV diagnosis and linkage to treatment should be complemented by efforts to improve the quality of HIV testing services and strengthen the quality management systems, particularly the use of validated testing algorithms and strategies, retesting people diagnosed with HIV before initiating treatment and providing clear messages to people with HIV on treatment on the risk of a “false negative” test result.

Keywords: HIV; HIV testing; misdiagnosis; misclassification; diagnostic error; false positive; healthcare; patient safety

Received 16 December 2016; Accepted 07 March 2017; Published 29 August 2017

Copyright: © 2017 World Health Organization; licensee IAS. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In any reproduction of this article there should not be any suggestion that WHO or this article endorse any specific organisation or products. The use of the WHO logo is not permitted.

Introduction

In the last decade, HIV testing services have been scaled-up substantially. In 2005, it was estimated that only 12% of people who wanted an HIV test were able to access testing; and that only 10% of people with HIV in Africa knew their status [1]. In contrast, between 2010 and 2014, more than 600 million people in 122 low- and middle-income countries received HIV testing [2], and as of 2015, approximately 60% of people with HIV were aware of their status [3]. Such scale-up has been possible through the expansion of provider-initiated testing and counselling and community-based testing programmes, which have routinized HIV testing and extended services to many people.

Rapid diagnostic tests (RDTs) have been instrumental to the scale-up of HIV testing, particularly in resource-limited settings where access to laboratory services is poor. RDTs have been shown to be highly accurate and can often provide a same-day diagnosis when used within a validated testing strategy (i.e. the order in which the tests are performed) and algorithm (i.e. the exact tests used within the testing strategy) according to high (≥5%) and low HIV prevalence (<5%), as recommended by the World Health Organization.
(WHO) [4–6]. Recent reports, however, have shown that HIV testing is not always conducted appropriately [7,8], and in some countries, quality systems have not kept pace with testing scale-up. According to a review of national HIV testing policies, less than 20% of testing strategies were consistent with WHO guidance, and only two included recommendations on retesting prior to the initiation of antiretroviral therapy (ART) [9]. In some cases, poor-quality testing has resulted in incorrect test results and the misdiagnosis of HIV status [10–14].

HIV misdiagnosis refers to any testing event where a diagnosis is missed, inappropriately delayed or incorrect (either false positive or false negative) [15]. Poor-quality HIV testing and misdiagnosis have negative consequences for individuals, families, communities, health workers and health services. False negative diagnoses represent missed opportunities to identify an HIV infection and link people to early treatment. False positive diagnoses may cause social and emotional harm and create mistrust of health workers and the test results they deliver. Without addressing HIV testing quality, new guidance offering same-day treatment to all people diagnosed with HIV [16] could lead to inappropriate ART initiation [11]. Once individuals are on treatment, because ART reduces antibody production and can cause seroreversion, for example, false negative test results, determining a person’s true HIV status can be especially challenging [17,18].

We conducted this systematic review to assess the magnitude of misdiagnosis and to identify and describe poor HIV testing practices using RDTs, including those which may have led to incorrect test results and misdiagnosis.

Methods

We systematically searched for peer-reviewed articles published from 1 January 1990 to 19 April 2017 using a predefined search strategy in the following electronic databases: PubMed, CINAHL and EMBASE. All conferences of the International AIDS Society were searched from July 2001 through July 2016; the most recent Conference on Retroviruses and Opportunistic Infections (2014–2017) database were searched because past conference abstracts were unavailable. Conferences of the African Society of Laboratory Medicine (ASLM) were searched 2012–2016, as well as the ASLM website and other key global health websites (see supplementary information). We searched reference lists to identify additional literature. This process was repeated until no new citations were identified. Experts were also contacted to identify additional reports. No geographic restrictions were placed on the search, but the review was limited to studies published in English.

Studies were eligible if they used at least two RDTs and reported on HIV misdiagnosis, factors related to potential misdiagnosis or described quality issues and related to HIV testing error.

Initial titles were screened by one investigator (VF) to determine eligibility. A second and a third screening was then carried out (VF, ST and CJ). All differences were resolved through consensus. Data from all sources were extracted and placed into standardized forms and verified in duplicate (VF and ST). CJ and NF assessed study quality (see supplementary data).

Potential factors relating to misdiagnosis were extracted from studies using defined categories: (a) clerical error (error in documenting and reporting information essential to a correct status); (b) user error (operator error collecting specimen, performing an HIV RDT or interpreting the result); (c) suboptimal testing strategy (errors related to the order in which specific RDTs are used, also known as a testing strategy); (d) poor management and supervision (lack of active quality management systems); (e) weak reactive results (faint lines appearing on test strips); and (f) additional factors including cross-reactivity, acute/early infection and testing among people on ART.

Other summary measures included: misdiagnosis rates (total number of false positive diagnoses reported over the total number of HIV-positive tests retested and reported using a specific testing algorithm and the total number of false negative diagnoses reported over the number of HIV-negative tests retested and reported using a specific testing algorithm). For studies exclusively among people diagnosed with HIV, reporting on false positive statuses, the total study population was used as the denominator.

For each study, rates of diagnostic error and misdiagnosis and corresponding 95% confidence intervals (CIs) were calculated, using Wilson’s approach, and this was displayed graphically using forest plots [19–21]. All statistical analyses were conducted in STATA v13.0.

Results

Sixty-four studies reporting on misdiagnosis of HIV and factors potentially related to misdiagnosis were included in this review (Figure 1 and Table 1).

Most studies were carried out in Africa (n = 48) [5,7,10–14,16,22–25,29,30,32–34,36,37,39,41,43,44,46,47,49,51,52,54–59,61–65,67,72,74,75] followed by in the Americas (n = 7) [28,31,42,50,53,60,66], Asia (n = 4) [8,35,45,73] and Europe (n = 1) [48]. There were also four multi-country/regional studies [26,27,38,40]. Samples varied by size and unit of measurement, including clients (n = 38 studies, range: 303,010 to 1 clients), specimens (n = 15 studies, range: 9419 to 16 specimens), health workers performing HIV tests (n = 5 studies, range: 3835 to 39 personnel) and sites where HIV testing was performed (n = 12 studies, range: 602 to 4 sites). Nine studies reported more than one unit of measure, and three studies did not specify sample size (see supplementary information). The majority of studies occurred in a facility-based setting; studies carried out in community settings included the workplace (n = 1) [57], home-based testing (n = 2) [14,39] and a mobile setting (n = 1) [32].

Factors related to the quality of HIV testing and potential misdiagnosis

Several factors, including HIV testing errors, were reported frequently (n = 131 times) across all included studies (see Table 2).
Thirty-seven studies reported using a suboptimal testing strategy that differed from the WHO recommendations [5,8,11–14,16,22–30,32–34,36–39,42–44,49,51,53,59,62,64–66,68,72,75]. Suboptimal testing strategies included using a highly specific first-line test and highly sensitive second-line test [14,33,39,55], using a single RDT for HIV-positive diagnoses [11,66,72], using a high prevalence testing strategy in a low prevalence setting [16,49], using a parallel testing algorithms and a tiebreaker testing strategy (where a third assay is used to resolve discrepant test results and rule in HIV infection) [5,12,13,16,24,25,27–30,34,36,37,68]. User errors, incorrectly performing the test procedure or incorrectly interpreting results, defined as human errors, were reported in 25 studies [7,8,11,14,26–28,31,33,34,37,40,42,46,52,54,57,60,65–68,70,72,73]. Errors identified included users having difficulty with specimen collection [14,28,68], performing RDTs [31,73], interpreting test results [10,24,27,30,32,40,42,48,62,65,66,74], reading test results too early [7] and not using the correct reagents/buffer [7].

Twenty-one studies reported inadequate management and supervision [7,8,11,26,27,31,41,43,46,52,59,62,64–67,69,71,72,74]. Of these, 10 studies reported issues with management of supplies [7,11,26,27,62,64,67,69,72,74], including stock-outs [7,26,62,64,67,69], the use of damaged or expired RDTs [26,27,64,67] and inappropriate RDTs (i.e. syphilis RDTs) for HIV testing [72]. Other factors related to poor management and supervision included testing within the...
Table 1. Classification of included studies (n = 64)

<table>
<thead>
<tr>
<th>Category</th>
<th>Study</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Potential HIV misdiagnosis and related factors</strong></td>
<td>Aghokeng et al. [22]</td>
<td>Cameroon</td>
</tr>
<tr>
<td></td>
<td>Baltazar et al. [23]</td>
<td>Mozambique</td>
</tr>
<tr>
<td></td>
<td>Baveewo et al. [24]</td>
<td>Uganda</td>
</tr>
<tr>
<td></td>
<td>Bock et al. [14]</td>
<td>South Africa and Zambia</td>
</tr>
<tr>
<td></td>
<td>Boerars et al. [25]</td>
<td>Zambia and Rwanda</td>
</tr>
<tr>
<td></td>
<td>CDC [26]</td>
<td>Low- and low–middle-income countries (not specified)</td>
</tr>
<tr>
<td></td>
<td>Crucitti et al. [27]</td>
<td>Benin, India, South Africa, Uganda and India</td>
</tr>
<tr>
<td></td>
<td>da Costa et al. [28]</td>
<td>Brazil</td>
</tr>
<tr>
<td></td>
<td>Eller et al. [29]</td>
<td>Uganda</td>
</tr>
<tr>
<td></td>
<td>Fogel et al. [23]</td>
<td>Multiple countries in Africa</td>
</tr>
<tr>
<td></td>
<td>Galiwango et al. [30]</td>
<td>Cameroon</td>
</tr>
<tr>
<td></td>
<td>Granade et al. [31]</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>Gray et al. [32]</td>
<td>Uganda</td>
</tr>
<tr>
<td></td>
<td>Hsiao et al. [33]</td>
<td>South Africa</td>
</tr>
<tr>
<td></td>
<td>Jentsch et al. [34]</td>
<td>South Africa, Tanzania, Uganda and Zambia</td>
</tr>
<tr>
<td></td>
<td>Kanal et al. [35]</td>
<td>Cambodia</td>
</tr>
<tr>
<td></td>
<td>Karugaba et al. [36]</td>
<td>Uganda</td>
</tr>
<tr>
<td></td>
<td>Khan et al. [37]</td>
<td>Swaziland</td>
</tr>
<tr>
<td></td>
<td>Klarkowski et al. [38]</td>
<td>Central Africa Republic, Congo, DRC, Ethiopia, Haiti, India, Cote d’Ivoire, Myanmar, Uganda and Zimbabwe</td>
</tr>
<tr>
<td></td>
<td>Klarkowski et al. [10]</td>
<td>DRC</td>
</tr>
<tr>
<td></td>
<td>Kufa et al. [39]</td>
<td>South Africa</td>
</tr>
<tr>
<td></td>
<td>Learmonth et al. [40]</td>
<td>Multi-country study (26 countries)</td>
</tr>
<tr>
<td></td>
<td>Manak et al. [16]</td>
<td>Nigeria</td>
</tr>
<tr>
<td></td>
<td>Maparo et al. [41]</td>
<td>Zimbabwe</td>
</tr>
<tr>
<td></td>
<td>Martin et al. [42]</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>Masina et al. [43]</td>
<td>Malawi</td>
</tr>
<tr>
<td></td>
<td>Mayaphi et al. [44]</td>
<td>South Africa</td>
</tr>
<tr>
<td></td>
<td>Mehra et al. [45]</td>
<td>India</td>
</tr>
<tr>
<td></td>
<td>Mine et al. [46]</td>
<td>Botswana</td>
</tr>
<tr>
<td></td>
<td>Nelson et al. [47]</td>
<td>Mozambique</td>
</tr>
<tr>
<td></td>
<td>Sacks et al. [48]</td>
<td>UK</td>
</tr>
<tr>
<td></td>
<td>Shanks et al. [11]</td>
<td>DRC, Burundi and Ethiopia</td>
</tr>
<tr>
<td></td>
<td>Shanks et al. [13]</td>
<td>Ethiopia</td>
</tr>
<tr>
<td></td>
<td>Shanks et al. [12]</td>
<td>Ethiopia</td>
</tr>
<tr>
<td></td>
<td>Simoncini et al. [49]</td>
<td>Niger</td>
</tr>
<tr>
<td></td>
<td>Stetler et al. [50]</td>
<td>Honduras</td>
</tr>
<tr>
<td></td>
<td>Tchounga et al. [51]</td>
<td>Burkina Faso, Cote d’Ivoire and Mali</td>
</tr>
<tr>
<td></td>
<td>Wolpow et al. [52]</td>
<td>South Africa</td>
</tr>
<tr>
<td></td>
<td>Viani et al. [53]</td>
<td>USA and Mexico</td>
</tr>
<tr>
<td></td>
<td>Young et al. [54]</td>
<td>Mozambique</td>
</tr>
<tr>
<td><strong>Focus on misdiagnosis of HIV-negative serostatus</strong></td>
<td>Bassett et al. [55]</td>
<td>South Africa</td>
</tr>
<tr>
<td></td>
<td>Kahemele et al. [56]</td>
<td>Tanzania</td>
</tr>
<tr>
<td></td>
<td>Matambo et al. [57]</td>
<td>Zimbabwe</td>
</tr>
<tr>
<td></td>
<td>Oiaru et al. [58]</td>
<td>Zimbabwe</td>
</tr>
<tr>
<td><strong>General quality issues from sites conducting HIV testing services</strong></td>
<td>Adebayo et al. [59]</td>
<td>Nigeria</td>
</tr>
<tr>
<td></td>
<td>Benzaken et al. [60]</td>
<td>Brazil</td>
</tr>
<tr>
<td></td>
<td>Bile et al. [61]</td>
<td>Botswana</td>
</tr>
<tr>
<td></td>
<td>Cham et al. [62]</td>
<td>30 countries in Africa</td>
</tr>
</tbody>
</table>
Sixteen studies reported clerical errors [8,11,26,28,29,31,34,35,45,50,63–65,67,73,75]. Errors included poor record-keeping [35], data reporting problems, labelling and transcription mistakes [73] and specimen mix-ups. Poor record-keeping, according to one study, resulted in nearly 30% of errors leading to incorrect status [67]. Clerical errors were not always clearly defined and may not have always led to misdiagnosis [28].

Fourteen studies reported challenges related to weak reactive test results, particularly difficulty with interpretation [8,10,24,27,30,32,36,38,40,42,44,48,62,74]. A study, which assessed the proficiency of laboratory technicians, found that specimens with very weak levels of HIV-1/2 antibodies were less accurately reported [40]. In Uganda, two studies found that the majority of false reactive results came from weak reactive RDTs [32,36]. 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 than true positive specimens.

Table 1. Reported HIV testing errors and factors potentially related to misdiagnosis

<table>
<thead>
<tr>
<th>Category</th>
<th>No. of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect/suboptimal testing strategy or algorithm (e.g. testing strategies not aligned to the World Health Organization recommendations, such as a tiebreaker or parallel testing strategies, use of a single RDT to make an HIV-positive diagnosis)</td>
<td>37</td>
</tr>
<tr>
<td>User error (e.g. errors performing RDT or interpreting results, misapplication of buffer, inaccurate reading time and other human errors)</td>
<td>25</td>
</tr>
<tr>
<td>Poor or inadequate management and supervision (e.g. work load stress, staff shortages, lack of training, poor adherence to testing strategy or testing algorithm, substandard operating procedures, testing in window period)</td>
<td>21</td>
</tr>
<tr>
<td>Other factors (e.g. acute infection, cross-reactivity, known HIV status/on ART)</td>
<td>18</td>
</tr>
<tr>
<td>Clerical/technical errors (e.g. mislabelling, poor record-keeping and clerical mistakes)</td>
<td>16</td>
</tr>
<tr>
<td>Weak reactive test results (e.g. faint or ghost lines appearing on test strip)</td>
<td>14</td>
</tr>
</tbody>
</table>

RDT: rapid diagnostic test; ART: antiretroviral therapy.

Table includes 63 reporting studies. One study (Bile et al. 2017) did not report a specific factor or error related to misdiagnosis. Some studies reported multiple factors related to poor quality testing and factors that could be related to potential misdiagnosis.
False positive diagnostic errors and misdiagnosis rates
Thirty studies reported on false positive diagnostic errors (43 reports; n = 16,777 total positive diagnoses). In general, error rates were small (median: 3.1%; IQR: 0.4%-5.2%) with the exception of a few studies where a tiebreaker test was used to resolve discrepant results [10–14,23,32,33,37,39,41,44,46,47,75] (Figure 2). Of these, six studies (eight reports) exclusively among people with HIV enrolled in care or ART reported that between 0.1% (95% CI: 0–0.3) and 6.6% (95% CI: 4.5–9.6) of people were misdiagnosed (median: 1.6%, IQR: 0.3–4.7%) [10–13,33,37,39,41,47] (Table 3).

In studies reporting false positive diagnoses, nearly all reported the use of a suboptimal testing strategy [11–13,33,39,49,75]. Sixteen studies reported the use of a tiebreaker testing strategy were related to false positive HIV diagnoses [5,12,13,16,24,25,27,29,30,32,34,36,37,68]. In one of these studies, 95% (123/129) of false positive statuses resulted specifically from using a tiebreaker test [32]. Additionally, one study reported which used a tiebreaker testing strategy in Burundi, Democratic Republic of Congo (DRC) and Ethiopia reported some may have been provided an HIV-positive diagnosis based on a single HIV RDT [11].

False negative diagnostic errors and misdiagnosis rates
Twenty-eight studies reported on false negative diagnoses (40 reports, total negatives = 55,626) (median: 0.4%, IQR 0–3.9%) (Figure 3) [10,12,13,37,47,55–57]. The studies reporting the highest proportions, for example, Olaru et al., which was designed to assess how ART impacts test performance [58], of false negative diagnoses were exclusively among people with HIV on ART who were retested using an HIV RDT-based algorithm.

Nearly all studies reporting false negative diagnoses also reported using a suboptimal testing strategy. Four studies in South Africa reporting false negative diagnoses reported that HIV testing was conducted with an algorithm using a first-line test with high specificity and poorer sensitivity [14,33,39,55]. According to one of these studies [14], between 2014 and 2016, the testing algorithm changed four times in an effort to address the high proportion of false negative diagnoses resulting from these algorithms.

Clerical and user errors [57], early/acute infection [16,44,45], presentation late in disease stage [56] and individuals with known HIV status on ART who sought retesting, or were retested using oral fluid-based RDTs [58], were also reported as factors contributing to false negative diagnostic errors [14,16,39,44,47]. In Zimbabwe, all the reported false negative diagnoses were among children on ART who were retested with an oral fluid-based HIV RDT [58]. In South Africa and Zambia, individuals on ART comprised 44% (26/59) and 14% (5/38) of false negative diagnoses, respectively [14]. In Mozambique, 88% (21/24) of all true HIV-positive clients with a false negative test result were confirmed to know their HIV status and 62% (13/21) were reportedly on ART [47]; reasons for retesting in study reportedly included users misunderstanding the question or hoping to receive health services and emotional or mental health issues.

Discussion
This review identified and described a number of diagnostic errors and poor HIV testing practices that may lead to misdiagnosis. Data on the magnitude of misdiagnosis was identified but limited, and no study could determine or quantify the exact cause(s) of misdiagnosis. Although no studies could determine and quantify the exact cause(s) of misdiagnosis, several identified the following factors to have strongly contributed: (1) suboptimal testing strategies, (2) poor management of supplies, (3) user errors including difficulty interpreting weak reactive lines and (4) retesting among people with known HIV status on ART.

No assay is perfect. False reactive and false non-reactive results are inevitable when using a single RDT and should be anticipated. However, the risk of misdiagnosis should be very low when a validated testing algorithm for high (≥5%) or low (<5%) prevalence settings is used [76]. In this review, we identified that many studies reporting diagnostic errors - both false positive and false negative - utilized suboptimal testing strategies which were not aligned to international guidance. Studies reviewed clearly showed the use of a tiebreaker strategy to rule-in HIV infection increases the likelihood of false positive statuses and possible misdiagnosis. This is concerning because a third of national testing strategies reviewed in 2015 recommended using a tiebreaker testing strategy [9].

In addition to adopting a proven testing strategy, national or regional validation is critical to determine which RDTs, and in which order, perform the best as a complete algorithm. As previously reported [38,77–83], tests and
and supervision. In several studies, this was not the case; nurses and untrained and uncertified providers were performing HIV testing [7,72]. Training, including pre-service, in-service and periodic refresher training, is important to maintain and improve the quality of services. Participation in EQA schemes is another way to monitor performance and improve testing services. Several studies also reported user and clerical errors resulted from inadequate support,·circumstances.

Table 2. Rates of false positive diagnostic error rates diagnosis (n = 30 studies, 43 reports).

<table>
<thead>
<tr>
<th>Study</th>
<th>Proportion false positive (95% CI)</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boeras et al. 2011</td>
<td>45.95 (31.04, 61.62)</td>
<td>C, F</td>
</tr>
<tr>
<td>Gray et al. 2007</td>
<td>43.73 (38.18, 49.43)</td>
<td>C, E, F</td>
</tr>
<tr>
<td>Fogel et al. 2015c</td>
<td>38.64 (35.72, 53.38)</td>
<td>A, C</td>
</tr>
<tr>
<td>Sacks et al. 2012</td>
<td>34.62 (19.41, 53.78)</td>
<td>E</td>
</tr>
<tr>
<td>Fogel et al. 2015d</td>
<td>33.33 (21.01, 48.45)</td>
<td>A, C</td>
</tr>
<tr>
<td>Fogel et al. 2015a</td>
<td>33.33 (9.69, 70.00)</td>
<td>A, C</td>
</tr>
<tr>
<td>Bock et al. 2017 SA</td>
<td>32.78 (27.16, 38.94)</td>
<td>B, C, F</td>
</tr>
<tr>
<td>Fogel et al. 2015b</td>
<td>22.22 (6.32, 54.74)</td>
<td>A, C</td>
</tr>
<tr>
<td>Kanai et al. 2005</td>
<td>18.18 (5.14, 47.70)</td>
<td>A</td>
</tr>
<tr>
<td>Mehra et al. 2014 a</td>
<td>12.20 (5.32, 25.54)</td>
<td>A, F</td>
</tr>
<tr>
<td>Mine et al. 2015b</td>
<td>11.11 (4.41, 25.31)</td>
<td>B, D</td>
</tr>
<tr>
<td>Mehra et al. 2014 b</td>
<td>10.26 (4.00, 23.58)</td>
<td>A, F</td>
</tr>
<tr>
<td>Jentsch et al. 2012</td>
<td>10.24 (7.95, 13.10)</td>
<td>A, B, C</td>
</tr>
<tr>
<td>Crucitti et al. 2011</td>
<td>8.70 (3.43, 20.32)</td>
<td>B, C, D, E, F</td>
</tr>
<tr>
<td>Shanks et al. 2015 (VL)</td>
<td>7.84 (3.09, 18.50)</td>
<td>C, F</td>
</tr>
<tr>
<td>Shanks et al. 2015</td>
<td>7.31 (4.55, 11.54)</td>
<td>C</td>
</tr>
<tr>
<td>Bock et al. 2017 ZAM</td>
<td>7.02 (5.84, 8.43)</td>
<td>B, C, F</td>
</tr>
<tr>
<td>Eller et al. 2007</td>
<td>6.58 (3.43, 20.32)</td>
<td>A, C</td>
</tr>
<tr>
<td>Kliarowski et al. 2009</td>
<td>6.58 (4.46, 9.60)</td>
<td>C, E, F</td>
</tr>
<tr>
<td>Shanks et al. 2015 (No VL)</td>
<td>6.54 (3.94, 10.68)</td>
<td>C</td>
</tr>
<tr>
<td>Kula et al. 2017</td>
<td>6.00 (2.00, 24.00)</td>
<td>C, F</td>
</tr>
<tr>
<td>Shanks et al. 2013 c</td>
<td>4.81 (3.61, 6.40)</td>
<td>B, D, C</td>
</tr>
<tr>
<td>Shanks et al. 2013 b</td>
<td>4.70 (2.29, 9.38)</td>
<td>B, D, C</td>
</tr>
<tr>
<td>Granade et al. 2004a</td>
<td>4.21 (1.65, 10.33)</td>
<td>A, B, D, F</td>
</tr>
<tr>
<td>Granade et al. 2004b</td>
<td>4.11 (1.41, 11.40)</td>
<td>A, B, D, F</td>
</tr>
<tr>
<td>Viani et al. 2013</td>
<td>4.00 (0.71, 19.64)</td>
<td>C</td>
</tr>
<tr>
<td>Shanks et al. 2013 a</td>
<td>2.56 (1.71, 8.88)</td>
<td>B, D, C</td>
</tr>
<tr>
<td>Mine et al. 2015a</td>
<td>2.43 (1.28, 4.56)</td>
<td>B, D</td>
</tr>
<tr>
<td>Aghokeng et al. 2009a</td>
<td>1.96 (0.54, 6.87)</td>
<td>C, F</td>
</tr>
<tr>
<td>Steifer et al. 1997</td>
<td>1.61 (0.44, 5.69)</td>
<td>C, F</td>
</tr>
<tr>
<td>Baveewo et al. 2012</td>
<td>1.38 (0.83, 2.31)</td>
<td>C, E</td>
</tr>
<tr>
<td>Manak et al. 2015 (HP)</td>
<td>1.16 (0.21, 6.30)</td>
<td>C, F</td>
</tr>
<tr>
<td>Baltazar et al. 2014a</td>
<td>0.78 (0.30, 1.96)</td>
<td>C, F</td>
</tr>
<tr>
<td>Khan et al. 2017</td>
<td>0.55 (0.33, 0.93)</td>
<td>B, C</td>
</tr>
<tr>
<td>Martin et al. 2011</td>
<td>0.54 (0.15, 1.95)</td>
<td>B, C, E</td>
</tr>
<tr>
<td>Hsiao et al. 2017</td>
<td>0.32 (0.11, 0.92)</td>
<td>B, C</td>
</tr>
<tr>
<td>Maparo et al. 2015</td>
<td>0.28 (0.11, 0.71)</td>
<td>D</td>
</tr>
<tr>
<td>Baltazar et al. 2014a</td>
<td>0.19 (0.03, 1.09)</td>
<td>C, F</td>
</tr>
<tr>
<td>Nelson et al. 2016</td>
<td>0.09 (0.03, 0.28)</td>
<td>F</td>
</tr>
<tr>
<td>Manak et al. 2015 (LP)</td>
<td>0.00 (0.00, 1.59)</td>
<td>C, F</td>
</tr>
<tr>
<td>Bassett et al. 2011</td>
<td>0.00 (0.00, 2.75)</td>
<td>C</td>
</tr>
<tr>
<td>Mayaphi et al. 2016</td>
<td>0.00 (0.00, 6.31)</td>
<td>E, F</td>
</tr>
</tbody>
</table>

To ensure correct results, all staff providing HIV testing must be trained, certified and provided ongoing support and supervision. In several studies, this was not the case, and algorithm performance vary across settings, often due to cross-reactivity caused by HIV subtypes, co-infections, comorbidities and possible environmental or population characteristics. Without validating a testing algorithm at a country or regional level, it would not be possible to fully understand the causes of poor performance. Furthermore, to ensure correct diagnoses, it is important to retest people diagnosed HIV positive before they enrol in care and ART. This is a cost-effective approach [84] which is increasingly critical as more people with HIV are being offered immediate treatment.
Table 3. Rates of false positive diagnosis rates among people diagnosed with HIV and/or enrolled in care or antiretroviral therapy (ART)

<table>
<thead>
<tr>
<th>Study/author</th>
<th>Sample size</th>
<th>Total no. of retested</th>
<th>No. of false positives</th>
<th>Percentage of false positive diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klarkowski et al. 2009</td>
<td>365</td>
<td>229</td>
<td>24</td>
<td>6.6</td>
</tr>
<tr>
<td>Shanks et al. 2013c</td>
<td>914</td>
<td>54</td>
<td>44</td>
<td>4.8</td>
</tr>
<tr>
<td>Shanks et al. 2013b</td>
<td>149</td>
<td>149</td>
<td>7</td>
<td>4.7</td>
</tr>
<tr>
<td>Shanks et al. 2013a</td>
<td>78</td>
<td>78</td>
<td>2</td>
<td>2.6</td>
</tr>
<tr>
<td>Khan et al. 2017</td>
<td>2533</td>
<td>88</td>
<td>14</td>
<td>0.55</td>
</tr>
<tr>
<td>Hsiao et al. 2017</td>
<td>952</td>
<td>37</td>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>Maparo et al. 2015</td>
<td>1447</td>
<td>1447</td>
<td>4</td>
<td>0.28</td>
</tr>
<tr>
<td>Nelson et al. 2016</td>
<td>3160</td>
<td>3146</td>
<td>3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Figure 3. False negative diagnostic error rates (n = 28 studies, 40 reports).

LP: low prevalence; HP: high prevalence; SA: South Africa, discrepant results; ZAM: Zambia, discrepant results; VL: visceral Leishmaniasis; Data reported includes reports of misdiagnosis of HIV-negative statuses. Misdiagnoses of HIV-negative statuses were reported 28 studies (40 reports), total negative = 55, 626. Kufa et al. 2017 reported proportion misdiagnosed but did not report full sample size information.

Note Olaru et al. was exclusively among people with HIV on ART, accounting for the high rate of false negative diagnoses.
User error interpreting weak reactive lines was a common challenge which contributed to false positive results. To address this, specialized training for health workers and site-level standard operating procedures including the use of a “second-reader” to validate the correct interpretation of test results may be needed, as well as work with manufacturers to improve RDTs and instructions on how to interpret faint lines and weak control lines. Several studies hypothesized that weak reactive lines may be caused by other user errors, for example, misapplication of buffer and reading test results too early and cross-reactivity. Further investigation into the cause of weak reactive and other faint lines, and how they can be prevented, is needed.

False negative test results among people with HIV and on ART were observed and contributed to a substantial proportion of misdiagnoses [14,16,39,44,47]. While it is unclear why people on ART would seek retesting, some reports suggest it may be due to wanting to “check” or “confirm” one’s HIV status and religious beliefs about being “cured” [85], as well as misunderstandings and emotional or mental health issues [47]. It is important for programmes and users to be aware of the potential risk of false negative results, as the presence of ART can lead to confusing test results and could result in individuals unnecessarily stopping treatment which could have dire individual and public health implications. As “treat all” policies are rolled out, it will be increasingly critical for programmes to address this issue and ensure clients and health workers are aware that testing individuals on ART is not recommended [76].

Strengths and limitations

This analysis is the first to bring together a diverse set of studies with the aim of identifying and describing suboptimal HIV testing practices and misdiagnosis. The results indicate the problem of misdiagnosis deserves attention. However, there are several limitations to this review.

As with all literature reviews, publication bias may be an issue and for this topic is inevitable and information on misdiagnosis is often unreported. This review was also limited to reports in English and may have missed reports in other languages. The majority of reports are from Africa and may not be representative of other geographies. Because the review was designed to identify reports of misdiagnosis, it is possible studies reporting errors and quality of HIV testing may have been missed.

Due to both the paucity and heterogeneity of data, it was not possible to conduct more quantitative analyses. Studies included were generally not designed to determine the exact cause or causes of misdiagnoses, a weakness cited across research on diagnostic errors [86].

This review focused on human errors and quality system failures. While we did identify some reports of cross-reactivity [10,22,24,25,27,32,38,56,60], reports did not provide conclusive information on what exactly caused cross-reactivity. Possible biological factors due to antibodies from inter-current infections, adverse environmental exposure to assay components, HIV subtype or shared false cross-reactivity in RDTs within an algorithm may be issues requiring further investigation.

Acute and early infection did not appear to be a significant cause of false negative diagnoses; however, few studies identified reported on acute infection. Retesting among HIV-positive individuals taking ART did emerge as a key factor contributing to a substantial proportion of false negative diagnostic errors and misdiagnoses. Further research is needed to understand how ART, as well as the use of antiretroviral drugs for prevention, for example, pre-exposure prophylaxis, may impact the performance of HIV RDTs, as well as how frequently people previously diagnosed with HIV and on ART retest.

Conclusions

Our review has identified a number of factors and practices that may contribute to diagnostic error and HIV misdiagnosis. Although no study could fully determine and quantify the exact cause(s) of misdiagnosis, our review elucidated four key factors: (1) suboptimal testing strategies, primarily the use of a tiebreaker testing strategy to rule in HIV infection, (2) user errors including interpretation of weak reactive lines, (3) inadequate management and supervision of testers and (4) retesting among people with HIV on ART. Most, if not all, are avoidable with appropriate guidelines, training and supervision. The consequences of misdiagnoses are serious at an individual and public health level. With the momentum to scale-up HIV diagnosis and linkage to ART, a parallel push to improve the quality of HIV testing services and prevent misdiagnosis is essential.

Authors’ affiliations

1Department of HIV, World Health Organization, Geneva, Switzerland; 2Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, UK; 3Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA; 4Department of Essential Medicines and Health Products, World Health Organization, Geneva, Switzerland; 5Department of Epidemiology and Population Health, American University of Beirut, Beirut, Lebanon; 6Department of International Health, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA; 7US Agency for International Development, Washington, DC, USA

Competing interests

The authors declare no competing interests. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.

Authors’ contributions

RF, CMCO, CI and AS conceived of and provided overall guidance to the study. VF and ST conducted the primary literature review and drafted the initial report. VF and CI conducted the update of the literature review and data extraction. CI primarily drafted the manuscript, performed quality assessment and data analyses and undertook supplementary literature reviews. NF and CI conducted the analyses. All authors have contributed to the conceptualization of the study, contributed to the development of drafts and read and approved the final version.

Funding

This publication was supported in part by a grant from the US National Institutes of Health (NIH; 5 R01 HD053268), Carla Makhlouf Obermeyer, PI,
disparity and the U.S. Agency for International Development (USAID), with support from the U.S. President’s Emergency Plan for AIDS Relief. The content of this supplement is solely the responsibility of the authors and does not necessarily represent the official views of the USA Government or the World Health Organization.

Disclaimer

The funders of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

References


29. CDC. HIV testing for surveillance: evidence from assessments of routine diagnostic testing in ANC. Presented at: WHO HIV testing and counselling consultation; 2014 Sep 10; Geneva, Switzerland.


54. Young PW, Mahomed M, Horth RZ, Shiraishi RW, Jani IV. Routine data collection from preventive to mother-to-child transmission (PMTCT) HIV testing not yet ready for HIV surveillance in Mozambique: a retrospective analysis of matched test results. BMC Infec Dis. 2013;13(1):.


64. Kalou M, Masha M, Mwa S, editors. 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. Presented at: First International Conference of the African Society for Laboratory Medicine; 12 S December 2012; Cape Town, South Africa.

65. Kithene F, Umuro M, Mwa S, editors. Implementation of 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; 12 S December 2012; Cape Town, South Africa.


