Diagnostic accuracy of simulated HIV testing algorithms following the latest WHO recommendations

Cara S. Kosack, Greet Beelaert, Tumwesigye Benson, Aboubacar Savané, Anne Ng'ang’a, Bita Andre, Jean-Paul B. N. Zahinda, Leslie Shanks, Katrien Fransen, Anne-Laure Page

Background

- In resource-constrained settings, HIV testing is based on rapid diagnostic tests (RDT) combined in an algorithm.
- According to the latest WHO guidelines (2012 and 2015), the algorithm is composed of 3 RDTs used sequentially according to the HIV prevalence in the tested population (Figure 1).
- These algorithms have not yet been widely adopted and there is limited data on their accuracy in different settings.
- We have used the results of a multi-centre study of the accuracy of HIV testing in 6 sites in 5 sub-Saharan African countries to simulate the WHO-recommended algorithms and present an accuracy data of selected algorithms globally and on specimens from DRC, where individual RDTs showed the lowest specificity.

Methods

- **Design**: laboratory diagnostic evaluation study
- **Samples**: At least 220 HIV-positive and 220 HIV-negative (according to on-site algorithm) specimens were obtained from each of six African sites in five countries (Guinea, Uganda, Kenya, Cameroon, Democratic Republic of Congo) and sent to the AIDS reference laboratory at Institute of Tropical Medicine, Belgium.

Results

- **2785 specimens were included in the study (438-500 per site)**
- **Table 1** shows the adjusted performance of selected algorithms on all specimens, and **Table 2** illustrates the results on specimens from DRC

Discussion

- The algorithms selected here all contained Determine as first test, since this is currently the first test used in most settings, and HIV 1/2 Stat-Pak as second or third test since this test showed better specificity than all other tests in the evaluation of individual RDTs.
- The performance of the algorithms for high HIV settings varies greatly with the choice of the second test, with the algorithms using HIV 1/2 Stat-Pak as second test showing significantly better specificity and PPV than most other algorithms.
- All the algorithms for low HIV settings presented here (and all including HIV 1/2 Stat-Pak) showed good performance, with slightly different proportions of inconclusive results.
- The variety of results on specimens from DRC, which showed the highest proportion of false-negative results in the RDT evaluation, illustrates the need for a proper local evaluation of different tests and algorithms for selection of an appropriate algorithm.
- Further modelling of the predictive values with varying prevalence should be performed for a more generalizable assessment of the performance in high- and low-prevalence settings.