# Viral Load is Critical in Limiting False-Recent Results from HIV Incidence Assays

**Reshma Kassanjee[1], Shelley Facente[2] , Sheila Keating[3], Elaine McKinney[4], Kara Marson[2], Christopher Pilcher[2], Michael Busch[3], Gary Murphy[4], Alex Welte[1]**

**For the Consortium for the Evaluation and Performance of HIV Incidence Assays (CEPHIA)**

Consortium for the Evaluation and Performance of HIV Incidence Assays

www.incidence-estimation.org/page/cephia

[1] The South African Centre Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa. [2] University of California - San Francisco, San Francisco, CA, USA. [3] Blood Systems Research Institute, San Francisco, CA, USA. [4] Public Health England, London, United Kingdom

## Background

### The Basic Idea behind incidence surveillance using recent infection tests

The cross-sectional use of (biomarker) tests for 'recent HIV infection' in principle offers affordable, low bias options for incidence estimation. The driving heuristic is that **a high prevalence of binary recent infection classification is indicative of a lately high incidence**.

### Critical Concepts

- This notion of recent infection is not defined by a gold standard such as a specific period post infection – rather, the tendency of the biomarker to manifest over various times post infection can be investigated, and the properties of the test summarised into a '**mean duration of recent infection**' (**MDRI**) and a '**false recent rate**' (**FRR** – the proportion of spurious 'recent' results among individuals infected for long periods of time).
- **A convenience cut-off time scale (T) is required** to make all this precise: Infections occurring more than a time T in the past are those for which 'recent' results are pronounced 'false'. This is crucially different to summarising test performance into a sensitivity and specificity.
- When these concepts and test properties are at least approximately in the useful regime, the **MDRI is a mainly context-independent property of the test**. On the other hand, **there appears to be some inevitable context dependence in the FRR -** significantly so for currently available *assays* (such as those evaluated in this study) in that viral suppression (due either to elite control or antiretroviral treatment) is predictive of long established infections being falsely classified as 'recent'.

### The Difficulty of developing a good recent infection test

Practical HIV surveillance applications require a not-too-short **MDRI – preferably at least 6 months** – to lead to usefully precise estimates of the prevalence of recent infection, and a sufficiently low **FRR - no more than about 2 percent**. When the FRR is low enough, and sufficiently precisely and accurately known, it can be consistently accounted for. The precise requirements vary in detail in complex ways from case to case, depending on the local epidemiology and programmatic context, feasible sample sizes, and surveillance needs, but **existing assays appear not to perform adequately – at least when used in isolation.**

***Table 1*** *False recent rates for infection test algorithms, with and without a viral load threshold criterion*

| | | | Scenario I Prevalence=20%; Incidence=2%; Treatment coverage=20% | | Scenario II Prevalence=40%; Incidence=1%; Treatment coverage=95% | | Scenario III Prevalence=5%; Incidence=2%; Treatment coverage=0% | |
|---|---|---|---|---|---|---|---|---|
| MDRI | Serological Assay | VL threshold | FRR | CoV of incidence | FRR | CoV of incidence | FRR | CoV of incidence |
| 200 | ArchAv | None | 7.21% | 64% | 2.98% | >100% | 9.70% | 23% |
| 200 | BED | None | 6.16% | 54% | 18.41% | >100% | 4.50% | 18% |
| 200 | BioRad2 | None | 1.37% | 20% | 2.87% | >100% | 2.60% | 16% |
| 200 | Geenius | None | 12.32% | >100% | 30.65% | >100% | 11.30% | 25% |
| 200 | LAg | None | 4.68% | 41% | 17.28% | >100% | 2.10% | 16% |
| 200 | LSVitDil | None | 10.56% | >100% | 29.68% | >100% | 10.40% | 24% |
| 200 | VitrosAv | None | 6.61% | 58% | 16.92% | >100% | 7.40% | 20% |
| 200 | ArchAv | 75 | 7.24% | 65% | 0.33% | 32% | 10.60% | 24% |
| 200 | BED | 75 | 1.68% | 22% | 0.13% | 28% | 4.20% | 17% |
| 200 | BioRad2 | 75 | 0.84% | 18% | 0.09% | 27% | 2.90% | 16% |
| 200 | Geenius | 75 | 4.95% | 43% | 0.36% | 33% | 11.50% | 26% |
| 200 | LAg | 75 | 0.46% | 17% | 0.06% | 26% | 1.80% | 16% |
| 200 | LSVitDil | 75 | 2.97% | 29% | 0.29% | 31% | 9.20% | 22% |
| 200 | VitrosAv | 75 | 2.51% | 26% | 0.23% | 30% | 7.50% | 20% |

*Note standard value of MDRI (200 days) achieved by choice of serological assay threshold. Results for MDRI of 300 (not shown) are barely distinguishable.*

## The Study

### Assay Evaluation

It is widely understood that using an explicit viral load assay in conjunction with a primary recent infection serological assay can in principle support a substantially improved practical definition of 'recent infection'. This study evaluates this idea by estimating the performance characteristics (MDRI and FRR) of 7 well known recent infection assays, considered alone, or in conjunction with a viral load test, subject to the interpretation that **low serological assay readings (below a chosen threshold), in the absence of viral suppression (defined by another tunable threshold) indicate 'recent' infection**. Assay results on a panel of 2500 well-characterised specimens drawn from the CEPHIA repository formed the basis of the evaluation. This is an additional analysis to a primary evaluation of individual assays (reference 3) The seven assays evaluated are: BED, Limiting Antigen (LAg), Less-Sensitive (LS) Vitros, Vitros Avidity, BioRad Avidity, Architect Avidity, and the Geenius assay.

### Making a Balanced Comparison Between Assays

Recency test properties **MDRI and FRR interact in a complex way** with incidence, Prevalence and sample size to yield the operational precision of the incidence estimate, which is the ultimately relevant metric. (**A common misconception is that FRR leads to bias, but a known FRR does not lead to bias, under correct analysis** – see reference 2) To make a simple direct and fair comparison, assay thresholds have been adjusted, away from standard developer recommendations, to provide a final MDRI of the total algorithm, of either 200 or 300 days. **Within each of these groups then, lower FRR immediately means better performance.**

**Demonstrative heuristic model scenarios have been provided**, which allow the calculation not only of the context dependent FRRs from the FRRs in underlying strata (such as treatment naïve patients (further stratified by time since infection), elite controllers, or those virally suppressed on treatment:

- **Scenario I**: Incidence = 2% p.a.; Prevalence = 20%; ART coverage = 20% of infected individuals: not unlike an **entrenched epidemic in sub-saharan Africa**, with an immature treatment programme.
- **Scenario II**: Incidence = 1% p.a.; Prevalence = 40%; ART coverage = 95% of infected individuals: a caricature of a **mature epidemic with declining incidence** against an ART-sustained high prevalence.
- **Scenario III**: Incidence = 2% p.a.; Prevalence = 5%; no ART: an extreme case of high incidence with relatively low prevalence, such as in a **high risk population of young women -** the kind of context where FRR has minimal impact on estimation.

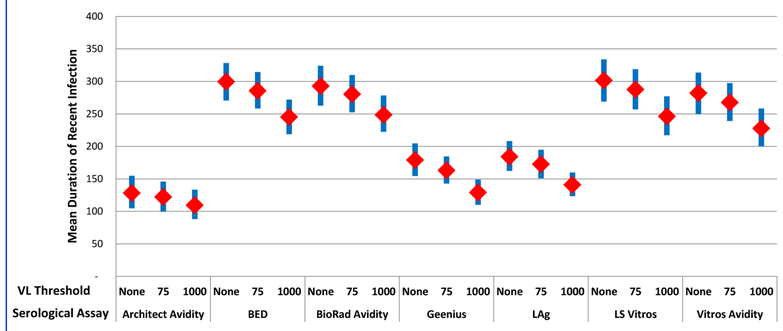
### Supplementary Viral Load testing is not a Free Lunch

**Figure 1** demonstrates a key point often overlooked in the focus on the use of viral load to reduce FRR. Even recently infected individuals may exhibit low viral loads. This should not be framed as 'false longstanding' results, but can be accounted for by consistent estimation of MDRI. Inevitably, additional **criteria which make it less likely for a specimen to be classified 'falsely' recent also make it less likely to be classified recent at any other times, thus reducing the MDRI**.

### Core Results

**Table 1** reports the FRRs and precision of incidence estimates (assuming non clustered sampling of 5000 individuals) for the three demonstrative scenarios.

***Figure 1*** *Mean Duration of recent infection under manufacturers threshold, using various Viral Load thresholds as supplementary algorithm criteria*




*Sensitivity of FRR is minimal, and data is relatively scarce, for viral load thresholds in the range 75 to 1000, but there is notable loss of MDRI, hence low thresholds are reported*

## Conclusions

### Central Lessons

- Currently available serological assays which form the core of a recent infection test are sufficiently confounded by viral suppression that **a direct viral load measurement should be regarded as an essential supplementary test** in any practical Recent Infection Test 'Algorithm', independent of context.
- Using supplementary viral load criteria, **currently available recent infection tests are viable for informative surveillance, in some contexts**.
- In principle, **there is a complex, context dependent trade-off between MDRI and FRR** as assay thresholds are varied, **but** we see little variation in the critical indicator (variance of incidence estimates) over a range of thresholds that vary MDRI from 200 to 300 days, which reassures us that **threshold optimisation is little more than choosing a reasonable value**.

### Ongoing work

- The existing repository has a **shortage of specimens from suboptimally treated individuals**, which would facilitate an investigation of the FRR in the realistically inevitable appearance of such subjects in practical applications. Such specimens are being procured.
- **Estimation of FRR in realistic settings requires some refinement** of the methods applied to the heuristic scenarios in the present analysis.

## References

1. CEPHIA project website: www.incidence-estimation.org/page/cephia.
2. Kassanjee R, McWalter TA, Bärnighausen T, Welte A. A new general biomarker-based incidence estimator. Epidemiology 2012; 23(5): 721–28.
3. Kassanjee R et al on behalf of CEPHIA. Independent assessment of candidate HIV incidence assays on specimens in the CEPHIA repository, AIDS. 28(16):2439-2449, October 23, 2014.

## Acknowledgements

The authors acknowledge, with thanks, the many people who contributed to the original evaluation (ref 3), notably the Bill and Melinda Gates Foundation, whose funding has made CEPHIA possible.

BILL & MELINDA GATES foundation

Public Health England

UCSF University of California San Francisco

Blood Systems Research Institute

SACEMA DST/NRF Centre of Excellence in Epidemiological Modelling and Analysis

UNIVERSITEIT STELLENBOSCH UNIVERSITY