Reply to ICAP re: amfAR's Report "Blinded by Our Own Data - Recency Testing in PEPFAR"

10 June, 2022
amfAR Public Policy Office
Contact: Brian Honermann, Deputy Director, Public Policy (brian.honermann@amfar.org)

amfAR's *Blinded by Our Own Data - Recency Testing in PEPFAR* report was released on June 1st, 2022 raising concerns regarding the implementation, cost, utility, and ethics of PEPFAR's recency testing program. On June 7th, 2022, ICAP at Columbia University, one of the driving implementers of the recency testing program for PEPFAR, issued a response to the report. In the interests of transparency, we have placed both the report and ICAP's response on our website.[1]

ICAP's response disagrees with the "premise, details, and conclusions" of amfAR's report. We encourage all readers to read and review both amfAR's report as well as ICAP's response. Here, we reply to ICAP's points in the interests of advancing the discussion on the appropriate use of recency testing within PEPFAR and its comparative value relative to other programmatic needs.

1. **Regarding Sensitivity of Recency Tests**

ICAP claims we misunderstand the intended use of recency testing. This is incorrect. As we stated multiple times in the report, there is a distinction between utilizing recency testing for incidence estimation and utilizing it as the basis of a real-time (or near-time) HIV surveillance system.[2] We understand the intended use of recency testing, including the objectives stated in

[1] https://www.amfar.org/news/response-to-icaps-critique-of-blinded-by-our-own-data/

[2] amfAR Report, page 2, stating: "Recency testing has already been deployed [...] as the basis for estimating HIV incidence. PEPFAR, however, is proposing a new use of recency testing: to serve as the basis of a real-time HIV surveillance system to identify small geographic areas and sub-populations that are most at risk to acquire HIV and to reallocate resources to those areas in rapid fashion to stop any onward transmission."; page 9, stating: "LAg-based testing has been used substantially within research to estimate national and population HIV incidence in different contexts. But utilizing recency testing as part of a research study to estimate incidence is very different from using it as real-time (or near-term) surveillance and to inform programmatic targeting as PEPFAR is proposing."; page 13, stating: "Importantly, in incidence studies, false long-term results are less critical because they can be somewhat accounted for in the formulas and adjustments for producing the incidence estimates. The unknowns in these calculations explain why incidence estimates generally require large sample sizes and produce results with wide confidence intervals at geographic levels below the country level. However, for real-time surveillance and programmatic response—in which both financial and human resources are going to be redirected based on recency testing data as proposed by PEPFAR— false long-term results are absolutely critical to ensuring that resources are not misdirected away from sites with ongoing transmission because of the inherent NPV of LAg."; page 14, stating: "Again, this is a separate agenda and utilization profile than using LAg for the purposes of incidence estimation on large-scale studies and populations. In such studies, the algorithms applied are designed, and the sample sizes large enough, to accurately estimate population incidence with reasonable (if wide) confidence intervals. No algorithms or "fixes" are possible to clean up, correct, or adjust the small geographic or facility-level results with the accuracy necessary to inform geographic targeting.";

the ICAP response, namely to “detect areas and subpopulations with high transmission rates”, “provide data to assess and guide national HIV testing strategies”, “be informative to addressing a country’s epidemic and its programmatic response”, all of which align well with the objectives for the surveillance system identified in our report.[3] In short, we are in agreement with the objectives of recency testing as a surveillance system within the PEPFAR program.

However, ICAP’s response then diverts into a technical distinction on whether sensitivity and specificity are appropriate metrics to apply to a test utilized for surveillance purposes rather than individual diagnostic purposes, claiming that “[s]ensitivity and specificity are not applicable concepts to HIV recency tests because they are not diagnostic tests.”[4] However, this technical distinction that “sensitivity” is narrowly defined only to apply to individual diagnostic testing does not address the issue raised in the report about why the sensitivity matters to the intended use of the data generated from the application of the test.

ICAP’s claim is that the test is “only useful at a population level” and that the results of recency testing as implemented by PEPFAR “are intended to represent the true population of recents in programmatically relevant characteristics, i.e., geographical distribution, age, sex, and key or priority population status”.[5] However, the sensitivity of the test is critical to whether the data produced reliably describe these characteristics, particularly when the review of the data is of very small slices of the overall data set. It is agreed that at the broad population level at an annual or semiannual basis, some signal will likely make it through the noise in the data. However, PEPFAR (and ICAP) is not proposing broad population level review, but rather regularly reviewing smaller sections of the data as stated in ICAP’s response: “Data are examined at least weekly and interpreted and used by programs on a monthly to quarterly basis during work-planning and data review sessions.” Additionally, the sample dashboards and other studies emanating from the program[6] talk about facility level hotspotting of transmission areas.

Analyzing small samples of recency surveillance data is unlikely to accurately depict the true population proportion or geographic distribution of recent infections. Low test sensitivity

---

[3] amfAR Report, page 2, stating: “PEPFAR, however, is proposing a new use of recency testing: to serve as the basis of a real-time HIV surveillance system to identify small geographic areas and sub-populations that are most at risk to acquire HIV and to reallocate resources to those areas in rapid fashion to stop any onward transmission.”; page 6, citing to COP 2022 guidance outlining the purpose of recency testing to “identify and specifically support populations falling short of the benchmarks or populations where new transmission is occurring” and “signal recent HIV transmission among subgroups and geographic locations. Programmatically, these signals of potential hotspots of recent transmission can be investigated further to identify and address missed opportunities within routine HIV testing, treatment, and prevention services in order to prevent ongoing transmission;” Page 7, stating: “COP guidance between 2019 and 2022 has consistently talked about using the results of real-time recency testing surveillance to target testing programs and PrEP programming, determine eligibility for DREAMS program expansion, and for other services.” [internal citations removed]

[4] ICAP Response, page 2.

[5] ICAP Response, page 2.

[6] See, e.g. Telford C, Tessema Z, Msukwa M, et al., 2022, available at: https://doi.org/10.15585/mmwr.mm7109a

becomes a particular issue when comparing sub-groups of data at smaller geographical areas - as is required for hotspot finding. If 35-65% of the true recent infections never appear in the sample due to low test sensitivity, and these errors are not evenly or proportionally distributed across geographical locations, this substantially erodes confidence in our ability to identify where true recent infections are the highest. It is insufficient to claim that because the intended use of recency testing is surveillance rather than diagnosis, that the performance of the assay is not consequential. Indeed, poor test performance will negatively impact programmatic decision making and must be considered when assessing the utility of recency testing even at the population level.

Moreover, the results of recency tests are being attached to individual patient's medical records and/or identities in many countries and are being used for individualized programmatic follow-up - specifically by prioritizing index testing for clients who test recent. While not clinically relevant to the client's own treatment, the attachment to individual patient records implies the test provides individually important information for which the sensitivity and specificity of the test are relevant.

Finally, ICAP's statement that "[t]he largest impediment to this effort is the substantial number of re-testers noted in the HIV testing programs—those who present as undiagnosed and ART naïve but are in fact already aware of their HIV positive status and are on antiretroviral therapy ART)" is concerning. This ascribes blame for the poor quality of the data to the population being tested and is irrelevant. The population is the population, if the methodology doesn't work well on the population being tested, it is the methodology that must be changed, not the population. The reasons that people who are already aware of their status presenting as undiagnosed were discussed in the report and stem from substantial service delivery failures and other issues that recency testing will not help to resolve.[7]

2. Regarding Specificity of Recency Tests

We accept that most countries are implementing recency testing with viral load confirmation, however, 4 countries are not (Ethiopia, Zimbabwe, Tanzania, and Thailand). This may be changing as the programs progress, but for the time-being and for comparability to past results and trends, the data in these countries is subject to the false recency rate described in the report as between 15% and 90%. It is incorrect to state that those rates are "not applicable". They are also highly applicable when discussing the return of results to clients and whether the inaccuracy of the test can be properly conveyed to individuals when RTRI results are returned,

[7] amfAR Report, page 18 stating: "There are many reasons that people may re-test for HIV. Civil society organizations and others have been pointing out for many years that re-testing is partially a result of poor retention in treatment programs, poor quality filing systems, and inadequate systems for patients transferring their care to different facilities. In these circumstances, in order for the individual to start accessing treatment again, patients often are re-tested for HIV to confirm their result and re-start the treatment enrollment process. This is avoidable re-testing that should be remedied through programmatic improvements. However, many people may also re-test for other reasons as well—concerns that they may have been misdiagnosed, difficulty accepting their status, misunderstanding about whether HIV can be cured, or simple curiosity. This is likely to be unavoidable re-testing." [internal citations removed]

as in Thailand or when RTRI results are returned without waiting for to viral load confirmation to be completed.

ICAP then states that the 5% - 30% range provided for persistent false recent results even after viral load confirmation is "misleading as the upper bound of 30% is derived from very small numbers of recent cases identified in the general population in the Population HIV Impact Assessment (PHIA) surveys". This is incorrect on multiple fronts. First, the range of 5% - 30% is the range documented in the report. The 5% lower bound is also derived from a very small study in Kenya. Our report accurately presented the range at the top and bottom end.

Second, the 30% figure is not derived from the PHIAs, but from the Irish Surveillance system that was implemented using LAg and is documented in Table 1 in the report (reproduced below).[8]

Third, the issue of small numbers of recent cases is important to understand. It is agreed that the number and size of studies conducted in the field is not optimal. The fact that the studies are not conclusive raises questions about PEPFAR's intention of scaling "recency testing for surveillance at scale across all sites and all HTS service delivery points within each site, whether supported by PEPFAR or by other entities"[9] without further evidence of the validity and utility of the data being collected.

That said, the point on small sizes of the studies is also inaccurately portrayed in ICAP's response. As shown in Table 1, the PHIAs across all countries identified 2,450 individuals as "recent" based on LAg EIA alone. The Rice study ICAP references as representative of "larger facility-based studies" identified only 166 individuals as recent on LAg EIA alone across two locations in Kenya (Zimbabwe data are excluded as the RITA applied in Zimbabwe only included VL). Likewise, the Irish surveillance data identified 128 individuals as recent based on LAg-EIA alone. Each study applied a different RITA to confirm recent infection post viral load as shown in the table. The PHIAs found a 15.7% persistent false recent rate applying a RITA including ARV presence in the blood despite a VL > 1,000. The Rice study found persistent false recent rates of 9.1% and 4.2% based on RITAs inclusive of clinical records searches and ARV presence in the blood despite a VL > 1,000, respectively. The Irish Surveillance data used the most expansive RITA including clinical records searches, CD4 > 200, AIDS defining illness, and prior PEP use and found a 27.3% persistent false recent rate despite a VL > 1,000.[10] Importantly, because the RITAs applied are different, it's not clear that these are directly comparable rates.

---

[8] HSE Health Protection Surveillance Centre. Monitoring Recent HIV Infection in Ireland, 2018. Dublin: HSE HPSC; 2020. https://www.hpsc.ie/ a-z/hivandaids/hivdataandreports/2018reports/ HIV_2018_recentinfection.pdf. Accessed April 30, 2022.

[9] COP 2022 Guidance, page 562

[10] We note that prior PEP use is included in the RITA because of the ability for ARVs lead to false recent results. Cross-tabs are not available, but of the 80 overall false recents identified, 65 were excluded for prior ART use (clinical records searches), 62 for VL < 1000 copies, 4 for PEP use in prior 6 months, 3 for CD4<200, and 1 for AIDS defining illness. People may have been identified as false recents for multiple reasons. See, HSE Health Protection Surveillance Centre, page 5.

Table 1. Studies with Published False Recency Rates or Positive Predictive Values of LAg

COUNTRIES INCLUDED IN STUDY	RECENCY ASSAY	RITA-2 ALGORITHM	LAg TESTED	LAg RECENT POSITIVE	RITA-1 POSITIVE (VL ONLY)	RITA-2 POSITIVE	LAg FRR (RITA-1)	RITA-1 FRR	LAg FRR (RITA-2)	SOURCE
Cameroon, Cote d'Ivoire, Eswatini, Ethiopia, Kenya, Lesotho, Malawi, Namibia, Rwanda, Tanzania, Uganda, Zambia, Zimbabwe	HIV-1 LAg-Avidity EIA (plasma) Maxim HIV-1 LAg DBS EIA (DBS)	RITA-2: RITA-1+ARV spetronomy	23,887	2,450	357	301	85.4%	15.7%	87.71%	Voetsch, et. al. (*JAIDS*)
Kenya	Maxim HIV-1 LAg-Avidity EIA	RITA-2: RITA-1 + Clinical Records Search	426	106	11	10	89.6%	9.1%	90.57%	Rice, et. al. (*JIAS*)
		RITA-2: RITA-2 + ARV metabolite testing	530	60	48	46	20.0%	4.2%	23.33%	
Zimbabwe	Maxim HIV-1 LAg-Avidity EIA	NA	313	49	33		32.7%			
China	Beijing Kinghawk LAg-EIA (DBS)	RITA-2: RITA-1 + CD4 > 200	1,152	205	174	145	15.1%	16.7%	29.27%	Zhu, et. al. (*IJID*)
Ireland	Sedia HIV-1 LAg-Avidity EIA	RITA-2: RITA-1 + Clinical records, CD4 > 200, AIDS defining illness, PEP use	508	128	66	48	48.4%	27.3%	62.50%	HSE Health Protection Surveillance Centre
Malawi	Asanté HIV-1 Rapid Recency Assay	NA	9,162	556	304	NA	45.3%			Telford, et. al. (*MMWR*)
Rwanda	Asanté HIV-1 Rapid Recency Assay	NA	7,919*	753	479	NA	36.4%			RWIbasira, et. al. (*PLoS One*)

* Excludes 26% of patients who self-reported prior ART use.

DBS: dried blood spot; EIA: Enzyme Immunoassay; RITA: Recent infection testing algorithm; RITA-1: RITA testing algorithm of LAg-avidity + viral load <>1,000 copies; RITA-2: RITA testing algorithm of RITA-1 + additional criteria as shown; FRR: False recency rate; PPV: Positive Predictive Value.

In the end, we support further research into the persistent false recent results being experienced post-VL confirmation, but it is not clear that the rates cited by ICAP ("RITA [with only VL<1,000] recent cases are >90% correctly classified") is the consensus of the scientific literature at the moment.

3. Regarding Inter-observer Reliability of RTRI

ICAP objects to the citation of a Uganda study regarding only 72% concurrence between 2 laboratory implementations of RTRI based on 85 samples. Instead, ICAP cites its own internal data from certification and proficiency training that it has implemented across 6,130 individuals (and "nearly 3,000 testers that underwent proficiency testing") on a battery of at least 13 samples. We agree that the Uganda data alone is not dispositive, but there are questions regarding ICAPs data on proficiency testing.

First, it is unclear the environments in which these training and proficiency testing results were achieved. If they were done as part of training and proficiency testing in which the testers are aware that they are being monitored for protocol violations and where certification is at issue, it is not surprising that the results in such an environment may be different to how the test is

ultimately implemented in real world environments without in-person oversight of the process. Second, if testers were assessed outside their own clinic environments in such as grouped training sessions, this normalizes many variables that may affect the consistency of reading results (lighting conditions, time pressures, pressures of providing counseling for a person whilst implementing the test, availability of reliable equipment such as timers, and other real-world factors). Such conditions could have substantial impact on the inter-reliability of the results.

Suffice it to say that real-world implementation in busy, often poorly resourced, and under-staffed clinics and laboratories are different environments than training environments. The point made in the report is that monitoring of the inter-reliability of the results between thousands of different testers operating in different environments is a substantial challenge and ICAP's training and proficiency testing results do not resolve the issue, despite it speaking to the ability to train testers to achieve a proficiency standard. We would like to see analyses documenting inter-reliability of the results across sites, in practice.

4. **Regarding Biased Program Data**

ICAP again claims that we misunderstand the objective of recency testing. We do not. The objective of using "recent infection surveillance [...] to find high transmission areas and subpopulations and to inform HIV prevention and testing strategies" is well understood. (See footnote 1 above). The problem with routine program data identified in the report is that comparatively assessing facility level or small geographic regions to identify hotspots requires that the data between these facilities and small geographic regions is consistently gathered. They are not.

HIV testing implementation, recruitment strategies, program outreach, testing modalities, and other factors that affect HIV testing uptake in local areas are not consistent. As a thought experiment, consider comparing results across two regions supported by different partners: In region A, there are high quality social network testing services, assisted partner notification services, and good youth programs led by local youth champions and organizations to encourage HIV testing uptake by youth. In region 2 facilities are only implementing voluntary testing and counseling services and nurses are known to be stigmatizing to young women who come for testing. Region A identifies 3 times the rate of recent infections as region B on a quarterly basis. Is Region A a higher transmission area than region B?

While this simplified example is extreme, it is emblematic of the ways that routine programming implementation will skew the consistency and comparability of the data across regions and programs. Not all HIV testing is implemented equally, especially when the program will expand to non-PEPFAR supported sites where PEPFAR's insight and control of the programmatic implementation of HIV testing will be much more limited. ICAPs response does not deal with the very real heterogeneity of program implementation that will bias the recency data received into the surveillance system.

As to the examples provided of evidence of the utility of recency data, they go to the questions raised in the report about what is genuinely new that we are learning or likely to learn from recency testing. The three examples are:

- That out-patient HIV testing, while low overall yield (the proportion of people who test HIV positive) is numerically very large as a modality and is ultimately numerically important to the total number of recent infections detected. This is not surprising. General facility and voluntary testing and counseling modalities account for well over 50% of all HIV tests conducted with PEPFAR support.[11] It would be shocking if a significant proportion of overall recent results (even if less than 50%) did not emanate from out-patient testing. If it's necessary to confirm this finding, a sampling methodology would be more than adequate rather than implementing a full surveillance system.
- That community-based HIV testing has higher rates of "re-testers" than other modalities of HIV testing. Again, this is well known. Motivations for re-testing differ, but community settings bring testing into environments where people will re-test, including to not out themselves as positive to others in their group or just to confirm their status.[12] The point of community testing is that it does pick up people that otherwise do not engage with the health care system or other HIV testing modalities.
- ICAP points to a newspaper article from the *Times of Eswatini* discussing "recent infection data by residence vs. place of diagnosis and identified underserved communities in prevention and testing services". Initially, a newspaper article does not prove the math is accurate regarding recent infections. Second, identifying underserved communities based on residence vs place of diagnosis does not require recency data. There are many ways to identify service deserts.

Additionally, ICAP claims to have "found that identifying the proportion of long-term infections within newly diagnosed individuals is also important. Across the countries where recent infection surveillance was implemented, it was noted that ~90% of new diagnoses are long-term infections. [...] [T]hose long-term infected individuals who are presenting late in the course of HIV will continue to hinder the performance of the national programs to reach and sustain epidemic control."[13] This is an important statement and one of fundamental disagreement. The fact is, there is virtually no information provided by recency testing about whether these individuals are - in fact - presenting "late". Approximately 85% of people infected longer between 6 and 12 months will test long-term. Individuals infected between 12 and 24 months are unlikely to be experiencing symptoms. Lumping most/all individuals infected longer than 6 months and stating that they are late presenters is a standard of HIV testing that makes absolutely no sense and provides no nuance as to whether people are genuinely testing "late".

---

[11] See https://mer.amfar.org/location/World/HTS_TST

[12] Klerk J, Bartolani A, Meta J, Erio T, Rinke de Wit T, Moyer E, *'It is not fashionable to suffer nowadays': Community motivations to repeatedly participate in outreach HIV testing indicate UHC potential in Tanzania*, PlOS One (2021), https://doi.org/10.1371/journal.pone.0261408; Shamu S, Farirai T, Slabbert J, et al. *A community-based HIV counselling and testing programme found a decreasing proportion of new HIV testers in South Africa*. Afr J AIDS Res. 2020;19(1):34-39. doi:10.2989/16085906.2019.1676804;

[13] ICAP Response, page 5

Additionally, PEPFAR over the past 3-4 years has actively put in additional barriers as a way to discourage frequent or repeat testing, including risk assessment screenings to access HIV testing and putting heavy emphasis on HIV testing yield targets and holding implementing partners accountable to those targets. These policies actively undermine the likelihood that individuals will access testing routinely and frequently to fall within the period likely to be diagnosed as recent. The rationale for implementing these policies was the increasing costs associated with HIV testing and we don't comment here on whether these policies are appropriate, merely that they have an effect on whether people are likely to access testing services during the recent window period. The fact that >90% of people may be diagnosed long-term on RTRI or RITA does not provide actionable information on whether they are genuinely late presenters.

Finally, as will be discussed later, ICAP does not address the cost of the recency testing program relative to the value provided by the findings outlined above. Despite ICAP's examples and response, there remains little evidence that actions taken in response to recency data have provably reduced incidence comparative to other geographies. In fact, most interventions recommended based on recency testing (PrEP, index testing, etc) are services that are already meant to be provided to all eligible and consenting clients.

5. Regarding Real-time Information and Response

Again, ICAP claims the distinction between incidence estimation and surveillance insulates the program from the dynamics of how recency assays function. ICAP states, "[t]his test identifies recent infections that were likely acquired in the prior 6-12 months, and not, as described in the Report, in the prior two years. The two-year time frame has been used in studies looking at LAg avidity testing and other recency assays for incidence estimation, but NOT for RTRI use for recent infection surveillance."[14]

Incidence estimation and the "shadow" period (a measure of how far back in time incidence is being estimated)[15] is a function of the mean duration of recent infection (MDRI) and false recency rate. It's imputed because the dynamics of how a test performs is important to how one can and should use the data generated from that test. Researchers utilizing LAg EIA at an ODn of <1.5 and VL<1,000 and an MDRI of 134 days found the shadow period to be 690 days.[16] The difference between LAg-EIA based recency assessment and the Asante™ HIV-1 Rapid Recency® Assay (RTRI) is immaterial because the very premise of the RTRI is that it performs

[14] ICAP response, pages 5-6
[15] Konikoff J, Brookmeyer R, Longosz AF, Cousins MM, Celum C, et al. Performance of a limiting antigen avidity enzyme immunoassay for cross-sectional estimation of HIV incidence in the United States. PLoS ONE. 2013; 8(12):e82772. https://doi.org/10.1371/journal.pone.0082772
[16] Konikoff J, Brookmeyer R, Longosz AF, Cousins MM, Celum C, et al. Performance of a limiting antigen avidity enzyme immunoassay for cross-sectional estimation of HIV incidence in the United States. PLoS ONE. 2013; 8(12):e82772. https://doi.org/10.1371/journal.pone.0082772

comparably to EIA at an ODn<2.0, as the TRACE Initiative's own training materials discuss.[17] In fact, the RTRI has a longer MDRI of 180 days (though CEPHIA researchers have found that varies by whether the assay is read visually (105 days) or electronically (197 days)).[18] The same dynamics that cause incidence estimation to have to be backdated still apply to surveillance usage, it's just in the background of the data generated from the use of the test - whether EIA or RTRI.

If these challenges didn't exist for incidence estimation, EIA or RTRI could be used to calculate incidence in real-time, but the data generated from the tests do not allow for that. These data quality challenges are inherent to the biomarker and test and must be taken seriously when making decisions about how to use the data produced by this program, even when not specifically intended to calculate incidence.

Additionally, ICAP misunderstands the report's critique of the real-time utilization of the surveillance data generated from recency testing. They state, "[t]hese RITA results are then displayed on dashboards that refresh as data are uploaded. This timeline is certainly **near real time in the context of when HIV testing takes place**. Data are examined at least weekly and interpreted and used by programs on a monthly to quarterly basis during work-planning and data review sessions."[19] [emphasis original] The issue raised in the report is not that the data are not actually being processed in real-time or analyzed in real-time, but that by doing so, the broad population level assessments discussed above in the sensitivity discussion, are no longer broad.

Breaking down recency data into weekly, monthly, or quarterly analysis - especially when then used to identify comparative facility level or geographic transmission hotspots - necessarily reduces the broad level analysis into smaller and smaller comparison groups where the noise in the data resulting from all of the above challenges is amplified. If not handled with respect for the uncertainty in the data points being analyzed, it is likely to lead to misdiagnoses of the reality of on-going HIV transmission.

While PEPFAR and ICAP lean on claims that recency testing for surveillance is only one additional data source, neither provides any actual proposal of what other data sources are available on the same or similar time-frame to be integrated into the "transmission hotspotting" analysis. The Trace-recency materials likewise offer nothing in this regard nor does the sample dashboard available on their website.[20]

Conclusion

---

[17] TRACE INITIATIVE, *Overview of RTRI: Assay Principle and Test Performance*, available at: https://trace-recency.org/wp-content/uploads/tools/training-materials/modules/01.-Overview-of-RTRI-Principle-and-Performance_April2021.pptx (slide 26)

[18] CEPHIA, Asanté HIV-1 Rapid Recency® Assay Evaluation Report, available at: https://zenodo.org/record/3509834#.Ymr5jBzMJhE,

[19] ICAP Response, page 6

[20] https://trace-recency.org/

Finally, ICAP does not address the additional issues that were raised by the report including the cost of implementing the recency testing program, the additional burden it places on nurses and other HIV testers, the data systems investments, training requirements, and all other activities necessary to begin to make the recency testing actionable even were the data of high quality. These costs are not borne by PEPFAR alone and even were they, PEPFAR's resources have been flatfunded for over a decade and recency testing is a substantial new cost to bear that must be justified against alternative investment opportunities. To date, PEPFAR has provided no costing assessments aside from the most basic commodities cost of the Asante assay itself. As stated in the report - those costs are likely to be only a fraction of the total costs of the recency program. The training costs, the quality-assurance costs, the quality-improvement costs, the data systems, the health care worker time are all critical to understand whether the additional value of the recency testing program can justify the actual costs. ICAP is in a position to provide some of those costing data as an implementer of the program.

ICAP also does not address the human rights and ethical issues raised in the report, which can only be justified if the data coming out of the recency testing program are sufficiently actionable to make real and substantial inroads in incidence reduction. No such evidence has yet been presented.