UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance

Global surveillance of HIV and sexually transmitted infections is a joint effort of the World Health Organization (WHO) and the Joint United Nations Programme on HIV/AIDS (UNAIDS). The UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance, initiated in November 1996, is the main coordination and implementation mechanism for UNAIDS and WHO to compile the best information available and to improve the quality of data needed for informed decision-making and planning at the national, regional and global levels.
FIND & WHO WORKING GROUP ON HIV INCIDENCE ASSAYS MEETING REPORT

20–26 FEBRUARY 2016
BOSTON, MA, USA
# CONTENTS

## ABBREVIATIONS

<table>
<thead>
<tr>
<th>ABBREVIATIONS</th>
<th>iv</th>
</tr>
</thead>
</table>

## 1. BACKGROUND

<table>
<thead>
<tr>
<th>1. BACKGROUND</th>
<th>1</th>
</tr>
</thead>
</table>

## 2. METHODS OF WORK

<table>
<thead>
<tr>
<th>2. METHODS OF WORK</th>
<th>2</th>
</tr>
</thead>
</table>

## 3. ALIGNMENT ON CRITICAL ASSAY CHARACTERISTICS

<table>
<thead>
<tr>
<th>3. ALIGNMENT ON CRITICAL ASSAY CHARACTERISTICS</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 CEPHIA and the CDC</td>
<td>3</td>
</tr>
<tr>
<td>3.2 Johns Hopkins University</td>
<td>4</td>
</tr>
<tr>
<td>3.3 Time since infection specimen dating</td>
<td>4</td>
</tr>
<tr>
<td>3.3 MDRI: the differences and how to resolve them</td>
<td>5</td>
</tr>
<tr>
<td>3.5 Optimizing the LAg + viral load recent infection testing algorithm – implications for sample sizes</td>
<td>6</td>
</tr>
<tr>
<td>3.6 ARV medicine testing component in recent infection testing algorithms</td>
<td>8</td>
</tr>
<tr>
<td>3.7 Real-world experience with LAg</td>
<td>9</td>
</tr>
<tr>
<td>3.8 External Quality Assurance Program Oversight Laboratory for incidence assays</td>
<td>10</td>
</tr>
<tr>
<td>3.9 Update on the case-based surveillance meeting</td>
<td>10</td>
</tr>
<tr>
<td>3.10 New assay validation by CEPHIA: Geenius™, ARCHITECT®, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3</td>
<td>11</td>
</tr>
<tr>
<td>3.11 Update on the development of rapid assays and the potential applications and advantages</td>
<td>12</td>
</tr>
</tbody>
</table>

## 4. DISCUSSION AT THE END OF DAY 1

<table>
<thead>
<tr>
<th>4. DISCUSSION AT THE END OF DAY 1</th>
<th>14</th>
</tr>
</thead>
</table>

## 5. TARGET PRODUCT PROFILES FOR HIV INCIDENCE ASSAYS

<table>
<thead>
<tr>
<th>5. TARGET PRODUCT PROFILES FOR HIV INCIDENCE ASSAYS</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 New target product profiles for HIV incidence assays (presented by Peter Dailey, FIND)</td>
<td>15</td>
</tr>
<tr>
<td>5.2 Discussion of target product profiles</td>
<td>16</td>
</tr>
<tr>
<td>5.2.1 Test performance characteristics</td>
<td>16</td>
</tr>
<tr>
<td>5.2.2 Time the analyte is stable in specimen storage format</td>
<td>17</td>
</tr>
<tr>
<td>5.2.3 Specimen types</td>
<td>18</td>
</tr>
<tr>
<td>5.2.4 Recognition of HIV subtypes and circulating recombinant forms</td>
<td>18</td>
</tr>
<tr>
<td>5.2.5 Target test cost</td>
<td>18</td>
</tr>
<tr>
<td>5.2.6 Supplementary tests</td>
<td>18</td>
</tr>
<tr>
<td>5.2.7 Time from assay start until the results are available</td>
<td>18</td>
</tr>
</tbody>
</table>

## 6. UPDATED MARKET ASSESSMENT FOR HIV INCIDENCE ASSAYS AND BIOMARKERS IN THE PIPELINE

<table>
<thead>
<tr>
<th>6. UPDATED MARKET ASSESSMENT FOR HIV INCIDENCE ASSAYS AND BIOMARKERS IN THE PIPELINE</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Updated market assessment (presented by Charles Morrison and Rick Homan, FHI 360)</td>
<td>20</td>
</tr>
<tr>
<td>6.2 Updated market assessment: discussion</td>
<td>23</td>
</tr>
<tr>
<td>6.3 Update on biomarkers in the pipeline</td>
<td>23</td>
</tr>
</tbody>
</table>

## 7. CONCLUSIONS AND NEXT STEPS

| 7. CONCLUSIONS AND NEXT STEPS | 25 |

## REFERENCES

| REFERENCES | 26 |

## ANNEX 1. MEETING PROGRAMME

| ANNEX 1. MEETING PROGRAMME | 28 |

## ANNEX 2. LIST OF PARTICIPANTS

| ANNEX 2. LIST OF PARTICIPANTS | 30 |
ABBREVIATIONS

ART  antiretroviral therapy
ARV  antiretroviral
CDC  United States Centers for Disease Prevention and Control
CEPHIA  Consortium for the Evaluation of the Performance of HIV Incidence Assays
CI  confidence interval
DBS  dried blood spot
DREAMS  Determined, Resilient, Empowered, AIDS-free, Mentored, and Safe women
FIND  Foundation for Innovative New Diagnostics
FRR  false recent rate
Global Fund  Global Fund to Fight AIDS, Tuberculosis and Malaria
HPTN  HIV Prevention Trials Network
IDE  immunodominant epitope
LAG  limiting antigen
MDRI  mean duration of recent infection
PBS  phosphate-buffered saline
PCR  polymerase chain reaction
PEPFAR  United States President’s Emergency Plan for AIDS Relief
PrEP  pre-exposure prophylaxis
RT-PCR  reverse-transcription polymerase chain reaction
SACEMA  South African Centre for Epidemiological Modelling and Analysis
SHIMS  Swaziland HIV Incidence Measurement Survey
1. BACKGROUND

In 2008, WHO established a Technical HIV Incidence Assay Working Group to examine the issues and challenges involved in estimating HIV incidence using assays. This Group comprises epidemiologists, laboratory specialists and public health officials and has worked to standardize the terms in the fields of assay calibration and validation. Several meetings to advance the agenda have been held, and copies of the reports are available on the Working Group web page (www.who.int/diagnostics_laboratory/links/hiv_incidence_assay/en).

These meetings have been successful in bringing together a wide group of assay users, especially from countries affected by the HIV epidemic that may consider using HIV incidence assays in the future, together with experts in applying laboratory-based methods for estimating HIV incidence. The importance of HIV incidence as a key indicator of the success or failure of national HIV programmes has been highlighted, and health ministries clearly need to be aware of the complexities in producing estimates based on data generated by the currently available assays.

In collaboration with the United States Centers for Disease Control and Prevention (CDC), the Working Group produced a guidance document on how to estimate HIV incidence at the population level using HIV incidence assays in cross-sectional studies (1). In 2010, the Bill & Melinda Gates Foundation adjudicated a grant proposal to Public Health England and the Blood Systems Research Institute to continue laboratory work on existing HIV incidence assays. The Consortium for the Evaluation of the Performance of HIV Incidence Assays (CEPHIA) was established, which aims to validate existing and future HIV incidence assays, compare results from existing assays with direct incidence measurements and identify the key parameters to enable assay results to be interpreted correctly.

The Working Group last convened in October 2014 in Barcelona, Spain; this was funded by WHO, the CDC, CEPHIA and UNAIDS (2). Following this meeting, a technical update was published for the application of HIV incidence assays for surveillance (3), which was also incorporated into the new guidelines for monitoring the impact of the HIV epidemic using population-based surveys (4).

The purpose of the 2016 meeting of the Working Group was to convene a group of key opinion leaders in the field to discuss critical issues outstanding after the previous Working Group meeting. The main objectives of this meeting were:

- to obtain consensus on the mean duration of recent infection (MDRI) for the CDC-derived limiting antigen (LAg) assays (from Sedia Biosciences Corporation and Maxim Biomedical) and variability among HIV subtypes;
- to present the performance characteristics of five further incidence assays studied as part of the CEPHIA evaluation;
- to present the evaluation of recent infection testing algorithms including testing for viral load and for antiretroviral (ARV) medicines;
- to release the revised version of the sample size calculation tool by the developer, recommended by WHO (5);
- to determine methods for qualifying new kit lots made by commercial partners to ensure LAg quality management systems during manufacturing processes;
- to establish a proficiency testing programme for the LAg assay that may be expanded to other assays as the need for external quality assurance is demonstrated; and
- to obtain consensus on target product profiles for tests for recent HIV infection.

The 2016 Working Group meeting in Boston was funded by UNAIDS, the Foundation for Innovative New Diagnostics (FIND) and WHO.
2. METHODS OF WORK

HIV incidence and mortality remain important indicators for the Sustainable Development Goals and are used as target measures, especially by the Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund) and the United States President’s Emergency Plan for AIDS Relief (PEPFAR).

The current role of the Working Group is:

• to review and document the calibration, performance and validation of HIV incidence assays;
• to create a catalogue of studies reporting on assay validation and calibration;
• to define the assay development pathway and target product profiles; and
• to obtain consensus on statistical issues and terms.

Since publication, the guidance on how to apply incidence assays has since been accompanied by two technical updates (3).

The critical path to improve HIV incidence testing incorporates market, technical and regulatory issues and requires developing normative guidance, advocacy and support. Market issues include the need for a reliable commercial supplier of incidence assays; technical issues include the need for specimens to standardize the calibration and validation of assays; regulatory issues include identifying and executing a clear regulatory path to independent validation and approved use in the market. WHO/UNAIDS will need to develop updated guidance to be adopted by the HIV community, with dedicated individuals supporting new users.

Understanding of the characteristics and limitations of HIV incidence assays has improved, including the need for large sample sizes in low-incidence areas. However, more work is needed to be able to expand their use to produce subnational estimates and estimates for key populations. Challenges include the scaling up of antiretroviral therapy (ART) and whether this will require ARV testing in addition to viral load in a recent infection testing algorithm. The effect of early or discontinued ART or a history of pre-exposure prophylaxis (PrEP) use is still being evaluated, but recent studies demonstrate that these can affect recency assays, such as slowing antibody avidity maturation. In addition, a local estimate of the false recent ratio (FRR) is required, which will become increasingly difficult with widespread ART use. The FRR varies by HIV subtype, being especially high among people with subtype D and is unknown for subtypes CRF02_AG and CRF01_AE. Following these outstanding issues, and specifically in the context of the new “treat all” guidelines, recommendations on best practices in the context of tests for recent infection are needed.
3. ALIGNMENT ON CRITICAL ASSAY CHARACTERISTICS

For currently available incidence assays, the performance characteristics (MDRI and FRR) vary by HIV subtype. One purpose of this meeting was to present MDRI estimates based on available subtype data for defined recent infection testing algorithms and to explore how to address mixed-subtype populations.

### 3.1 CEPHIA and the CDC

The subtype-specific MDRI s were presented for the first seven HIV incidence assays evaluated by CEPHIA (ARCHITECT® HIV antigen/antibody Combo assay (Abbott Diagnostics), Sedia BED HIV-1 incidence enzyme immunoassay (Sedia Biosciences Corporation), CDC-modified GS HIV-1/HIV-2 PLUS O enzyme immunoassay (Bio-Rad Laboratories), Geenius™ HIV 1/2 Supplemental Assay (Bio-Rad Laboratories), the CDC-approved Sedia™ HIV-1 LAg-Avidity enzyme immunoassay (Sedia Biosciences Corporation), VITROS® Anti-HIV 1+2 assay (less sensitive method, Ortho Clinical Diagnostics) and VITROS® Anti-HIV 1+2 assay (avidity, Ortho Clinical Diagnostics)). These were evaluated for subtypes A, B, C and D. The MDRI s for subtypes B and C could be estimated to a high level of certainty. For subtypes A and D, there was wide variance around the MDRI s because fewer specimens of this type were available. The MDRI differed substantially by subtype; subtypes A and D seemed to be outliers for some of the evaluated assays.

Pooled data from CDC seroconverte panels and CEPHIA evaluation panels were used for the analysis of MDRI s. For the LAg assay, MDRI s for an optical density of 1.5 and 2.0 were presented. The MDRI for subtypes B and C were 136 (95% confidence interval (CI) 117–157) days and 174 (95% CI 152–198) days, respectively. At optical density 2.0, the MDRI for subtypes B and C were longer and closer with 205 (95% CI 177–233) days and 214 (95% CI 188–241) days, respectively.

Most recent infection testing algorithms used in the field include a supplementary viral load threshold to minimize the FRR among treated people, which affects the MDRI. The CDC seroconverte panels had no viral load data available. A viral load effect was estimated for various viral load thresholds from the CEPHIA data. This was subsequently applied to the estimates from the pooled analyses. Using an optical density of 1.5, a viral load threshold of 100 reduced the MDRI by 14 (95% CI 6.7–23.0) days. For viral load thresholds of 400, 1000 and 5000, the MDRI was reduced by 30 (95% CI 18.0–44.7), 43.4 (95% CI 29.0–59.6) and 88.1 (95% CI 68.3–109.1) days, respectively. At an optical density of 2.0, the reduction in the MDRI was slightly higher (1–2 days) at the chosen viral load thresholds.

The estimates for MDRI s are relative to the estimated date of detectable infection. The estimated dates of detectable infection are based on diagnostic testing histories and are complex to estimate. The data on MDRI s presented were relative to the estimated date of Western blot seroconversion. There is a move towards using the estimated date on which infection would become detectable on a hypothetical viral load assay with a sensitivity of 1 copy/ml. However, any standard may be chosen if the screening assay used is considered. Using a recent infection testing algorithm with an Optima qualitative RNA assay, the median date of detectable infection was about 24.3 (95% CI 18.8–31.0) days before Western blot seroconversion.

The application of these adjustments was demonstrated on data from South Africa, showing the difference in incidence estimates and required sample sizes. The recent infection testing algorithm included nucleic acid amplification testing for identification of HIV-positives and a recent infection testing algorithm included LAg (cut-off optical density 1.5) + viral load (cut-off at 1000 copies/ml). An FRR of 0.25% and an MDRI of 188 days (6) were used to generate an incidence estimate of 1.84% (95% CI 1.14–2.54%) (on a sample size of 10 000, among which 2000 were HIV positive and identified 80 recent samples). Using an MDRI of 154 days (to adjust for screening assays and LAg performance in subtype C alone) generated an incidence estimate of 2.25% (95% CI 1.48–3.02%). Measuring a 40% decline in incidence in this population requires a sample size of 28 101 with an MDRI of 188 days versus 33 010 for an MDRI of 154 days (with prevalence of 19% and incidence of 1.27% in the first survey and 18% and 0.78%, respectively, in the second survey, design effect of 1.5 on HIV prevalence and the prevalence of recency among positives and assuming that the probability of correctly inferring a reduction to this extent is 80%).

Christopher Pilcher et al. (University of California, San Francisco, USA) will soon publish estimates for the estimated date of detectable infections for various screening tests. It was recommended that the guidelines specify that MDRI s are specific to HIV subtype, recent
infection testing algorithm supplementary tests (such as viral load) and the initial screening assay.

3.2 Johns Hopkins University

Johns Hopkins University has about 7000 data points for people with known duration of HIV infection by subtype. The assays evaluated were the Sedia™ BED HIV-1 incidence enzyme immunoassay, Sedia™ HIV-1 LAg-Avidity enzyme immunoassay and the protocol for the Bio-Rad GS HIV-1/HIV-2 PLUS O enzyme immunoassay developed by Johns Hopkins University and the HIV Prevention Trials Network (JHU/HPTN) for subtypes A, B, C and D.

The sample panel included elite suppressors and people receiving ART. MDRIs and FRRs were presented for each assay by subtype without a supplementary viral load. For the Sedia™ BED HIV-1 incidence enzyme immunoassay (<0.8 optical density), the MDRI was high, ranging from ~200 to 450 days, as was the FRR, ranging from ~6% to 20% and thus above the target of 2% for all subtypes.

The Bio-Rad GS HIV-1/HIV-2 PLUS O (JHU/HPTN) (avidity index <40%) had a much lower MDRI for all subtypes (~120–220 days) and an FRR <2% for all subtypes but D (FRR ~14%). For Sedia™ HIV-1 LAg-Avidity (optical density <1.5), the MDRI was marginally higher than for Bio-Rad GS HIV-1/HIV-2 PLUS O (JHU/HPTN) for each subtype, but the FRR was <2% only for subtype A and ~3% for subtype C.

Sequence data were examined for subtypes A and D on individuals from Rakai infected for >2 years (ARV medicine naive) to explore any differences in the virus and consequent antibody response. No evidence was found that the high FRR in people with subtype D resulted from the same manufactured assay and a slightly different procedure, with CDC having marginally higher MDRI and FRR values. With a recent infection testing algorithm, Sedia™ BED HIV-1 LAg-Avidity assay demonstrated low viral loads (<400 copies/ml, adjusted odds ratio 3.7 (1.6–8.6) and CD4 <50 cells/mm3, adjusted odds ratio 5.4 (95% CI 1.9–15.7)) were factors associated with misclassification (7). Further, the individuals who were misclassified at an earlier point of infection were five-fold more likely to be misclassified later, after controlling for other factors. This was especially evident for people with subtype D.

The CDC and JHU/HPTN generated different results using the same manufactured assay and a slightly different procedure, with CDC having marginally higher MDRI and FRR values. With a recent infection testing algorithm, Sedia™ HIV-1 LAg-Avidity (optical density 2.9) + Bio-Rad GS HIV-1/HIV-2 PLUS O (JHU/HPTN) (avidity index < 85%) + viral load >400 copies/mL + CD4 >50 cells/mm3, it was possible to obtain an MDRI of ~170 days for subtypes A, B and C and an FRR <1% (8). Of note, subtypes A and D are prevalent in the same areas (eastern Africa) and had substantially different MDRIs and FRRs (Sedia™ HIV-1 LAg-Avidity + viral load >400 copies/mL: the MDRI for subtype A was 130 days and subtype D was 236 days; the FRRs were 0.5% and 4.1%, respectively).

The Sedia™ HIV-1 LAg-Avidity data (using optical density 1.5) were reanalysed for samples from people not receiving ART using the new software package to calculate the MDRI developed by the South African Centre for Epidemiological Modelling and Analysis (SACEMA). Among all people tested, the misclassification frequency was examined by subtype and cohort. The MDRI was based on third-generation screening assay seroconversion; for 20%, an acute time point was used 13 days after infection. For subtypes A, B and C, the MDRI ranged from ~100 to 170 days, with overlap in the confidence intervals. The FRR was ≤2% for subtypes A and B and ~3% for subtype C. MDRI and FRR values were significantly higher for subtype D.

Johns Hopkins University compared Sedia™ HIV-1 LAg-Avidity and Maxim HIV-1 LAg enzyme immunoassay (Maxim Biomedical) avidity data on 1601 specimens of subtype C from 176 seroconverters. The optical density values demonstrated good agreement between the two assays ($R^2 = 0.93$), with a small coefficient of variation (0.82%). The normalized optical density values showed slightly less agreement ($R^2 = 0.86$), with Sedia™ values slightly higher than Maxim (coefficient of variation 12.4%). For samples with optical density <2 for either assay, Sedia™ values were slightly higher than Maxim ($R^2 = 0.82$, coefficient of variation = 15.4%). The MDRI for Maxim was 20 days greater than that for Sedia™ (140 days) using only 1.5 optical density, 22 days greater when adding viral load >400 copies/mL in the algorithm (Sedia™ + viral load >400 copies/mL: 117 days) and 19 days greater if using viral load >1000 copies/mL (Sedia™ + viral load >1000 copies/mL: 104 days). This difference seemed related to the calibrator. A suggestion was made to re-examine the results fitting the data to a polynomial and to review the number of recents classified by each assay.

3.3 Time since infection specimen dating

Individual infection timing is useful for targeting rapid or emergency ART, interpreting clinical symptoms and research on pathogenesis, ART outcomes and preventing HIV infection. The accuracy of infection timing is important for reporting early-stage infections, targeting public health interventions and designing research studies. The purpose of this session was to introduce an improved method of infection dating proposed by CEPHIA as a standardized method that is less cumbersome than the gold standard Fiebig staging method (9).

The Fiebig method identifies five stages after infection based on the results of the p24 antigen immunoassay, third-generation antibody immunoassay, second-generation antibody immunoassay and Western blot. However, some of these tests are now obsolete, and insufficient information is often available in datasets to conduct Fiebig staging. Symptoms could be used (onset = infection + 14 days), although this may be problematic,
since acute symptoms are not well understood. Testing history data may be used (assuming Western blot seroconversion, midway between the last negative and the first positive test) or tests for recent infection.

CEPHIA aimed to develop a more generalizable method to estimate date of detectable infection by incorporating information routinely available, such as other tests performed on the same day (Fiebig method) or different days that could be applicable to patient samples obtained months after infection. Data on two methods were presented, first using viral load ramp-up (RNA viral load + antibody negative) exploiting quantitative viral load increase dynamics; second, using seroconversion history based on information from testing history and test conversion dynamics to obtain a point estimate. A total of 53 plasma donor panels (468 specimens) with consecutive measurements <7 days apart for which the first measurement was <100 copies/ml were used. These included 174 Western blot–negative specimens and 60 specimens in the original Fiebig publication. Only specimens from the Fiebig dataset were used for the viral load ramp-up model, which was validated on specimens not in the Fiebig study. The increase in viral load was estimated to be 0.33 logs per day (base 10). This was used to examine the performance of Fiebig estimates in 119 specimens not in the Fiebig data, comparing the observed time in days since the date of detectable infection. It showed that the predicted time since the date of detectable infection had a dramatically improved correlation with the observed date of detectable infection using the viral load ramp-up model compared with the Fiebig staging method.

The method to estimate time since infection using the midpoint of a first-positive and last-negative HIV test took into account the screening assay (such as 30 days before the midpoint using Western blot). In the field, several different tests are used and Western blot is not used for screening, and he conversion times between the most and least sensitive tests could differ by up to one month. Taking the different screening tests into account, a suggested formula for the estimated date of detectable infection was:

\[ \text{Estimated date of detectable infection} = \frac{t_1 - d_1 + t_2 - d_2}{2}, \]

where

- \( t_1 \) = the time of the most sensitive negative test;
- \( d_1 \) = the delay associated with the most sensitive negative test;
- \( t_2 \) = the time of the least sensitive positive test; and
- \( d_2 \) = the delay associated with the least sensitive positive test.

The conversion times for all available screening tests have been published (10,11).

This method and the Fiebig staging method compared with the observed number of days since infection showed similar estimated dates of detectable infection.

CEPHIA intends to review data on early seroconversion dynamics and explore whether low-end results on incidence assays could be used to extend Fiebig staging beyond the antibody-positive stage in the absence of a last-negative test result. Data for this are available on 242 specimens (107 people). Preliminary analysis indicated that the assays performed differently among individuals, especially during the earlier stages of infection.

### 3.3 MDRI: the differences and how to resolve them

The purpose of this session was to resolve discrepancies in the MDRI values that were calculated between JHU/HPTN and CEPHIA/CDC. In 2012, the International Centre for AIDS Care and Treatment Programs conducted a survey in South Africa that included HIV incidence testing using the LAg assay (12) and applied an MDRI of 130 days (6,12). Incidence estimates did not take into account varying subtypes or consider viral load or ARV testing. The FRR was assumed to be 0%. Johns Hopkins University estimated the MDRI to be 151 days for subtype C (the most prevalent subtype in South Africa). Reanalysis of CDC and CEPHIA pooled specimens estimated the MDRI for subtype C to be 174 days, although adding viral load >1000 copies/mL produced an MDRI of 130 days (Eduard Grebe et al., South African Centre for Epidemiological Modelling and Analysis, Stellenbosch University, 2016, unpublished). A summary of the MDRIs for the various subtypes (not adjusting for viral load) was shown comparing CDC/CEPHIA estimates with those of Johns Hopkins University using Sedia™; the estimates from CDC/CEPHIA were greater for all subtypes except subtype D (Table 1). The difference was about 40 days for subtypes A and B and 23 days for subtype C. For subtype D, Johns Hopkins University estimated the MDRI to be 20 days longer; however, the confidence intervals were wide.

SACEMA has made available an R-code programme to estimate the MDRI with confidence intervals. Possible reasons cited for the differences were:

- the nature of the cohorts, although both had excluded people receiving ART;
- the assay provider, although Sedia™ was used in most instances;
- the specimen type: plasma used for CEPHIA/CDC specimens and plasma or serum for JHU/HPTN;
- the differences in infection dating methods and reference points: Western blot versus third-generation antibody testing;
- the method of determining the subtype, although these were broadly the same;
• the method of viral load testing: unlikely to have had a great effect, since the people receiving ART were excluded; and
• the method for estimating MDRI: in this case, the same.

It was suggested to compare the characteristics of the people included in each of the samples to explore any innate differences.

A proposal was made to develop an MDRI guidance tool for non-experts, which may include HIV subtype, assay threshold, recent infection testing algorithm components, method for HIV screening, kit manufacturer and specimen type. The same was suggested for the FRR, incorporating relevant variables. Key challenges will remain mixed-subtype populations. The preliminary results from a study on 100 CRF02_AG ART-naive specimens from Cameroon using LAg (optical density <1.5) produced an FRR of 3.2% (95% CI 0.7–9.0%). HIV infection had been established using the national algorithm in Cameroon (Alere Determine™ HIV 1/2½ (Alere) and Murex® Combo antigen/antibody kit (DiaSorin)) and the HIV-1 subtype was determined with RT-PCR of gp41 and phylogenetic analysis.

### 3.5 Optimizing the LAg + viral load recent infection testing algorithm – implications for sample sizes

The purpose of this session was to evaluate the feasibility and sample size requirements for using HIV incidence assays to obtain incidence estimates for two different use scenarios: national-level estimates and impact assessments in key populations. Population-based HIV impact

<table>
<thead>
<tr>
<th>HIV subtype</th>
<th>Source</th>
<th>Number of specimens</th>
<th>Number of panels</th>
<th>MDRI without viral load</th>
<th>CI for MDRI without viral load</th>
<th>MDRI viral load 100 copies/mL</th>
<th>MDRI viral load 400 copies/mL</th>
<th>MDRI viral load 1000 copies/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>CEPHIA/CDC</td>
<td>185</td>
<td>89</td>
<td>212</td>
<td>157–277</td>
<td>199</td>
<td>182</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN</td>
<td>587</td>
<td>76</td>
<td>168</td>
<td>137–201</td>
<td>149</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CEPHIA &gt; JHU/HPTN; confidence intervals overlap</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRF01_AE</td>
<td>CEPHIA/CDC</td>
<td>644</td>
<td>97</td>
<td>124</td>
<td>106–149</td>
<td>111</td>
<td>94</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>CEPHIA/CDC</td>
<td>715</td>
<td>166</td>
<td>136</td>
<td>117–157</td>
<td>122</td>
<td>106</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN</td>
<td>698</td>
<td>270</td>
<td>96</td>
<td>62–137</td>
<td>66</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CEPHIA &gt; JHU/HPTN; confidence intervals overlap</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>CEPHIA/CDC</td>
<td>731</td>
<td>247</td>
<td>174</td>
<td>152–198</td>
<td>160</td>
<td>143</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN</td>
<td>669</td>
<td>221</td>
<td>151 (140*)</td>
<td>129–189</td>
<td>111</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CEPHIA &gt; JHU/HPTN; confidence intervals overlap</td>
<td>23 (34*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>CEPHIA/CDC</td>
<td>139</td>
<td>42</td>
<td>256</td>
<td>165–361</td>
<td>242</td>
<td>225</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN</td>
<td>222</td>
<td>28</td>
<td>276</td>
<td>192–360</td>
<td>270</td>
<td>236</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JHU/HPTN &gt; CEPHIA; confidence intervals overlap</td>
<td>-20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*New subtype C data (Sedia assay).
assessments implemented by the International Center for AIDS Care and Treatment Programs (ICAP) (13) measure HIV incidence as one of several indicators. National HIV incidence is to be determined using a recent infection testing algorithm consisting of LAg + viral load. DREAMS (Determined, Resilient, Empowered, AIDS-free, Mentored and Safe women) is a partnership funded by the Bill & Melinda Gates Foundation, PEPFAR and the Girl Effect aiming to reduce the number of adolescent girls newly infected with HIV in 10 countries in sub-Saharan Africa. Specifically, they aim to reduce the HIV incidence by 25% among women 15–24 years old by 2016 and by 40% by 2017 (14).

The factors affecting survey sample sizes for the precise estimation of incidence are local prevalence and incidence, MDRI, FRR as well as the desired precision and power and survey design effect. To calculate the effect on survey sample size requirements to obtain incidence estimates, a recent infection testing algorithm (LAG + viral load) was used. Prevalence and incidence estimates were taken from UNAIDS data from 2016 (15), except for Ethiopia, which was from 2014. For key populations, incidence and prevalence estimates were taken from published studies (16–18). Two scenarios were modelled:

- a single population survey with a desired precision of the incidence estimate: relative standard error 20% (as recommended by WHO/UNAIDS) or 30%; and
- two consecutive surveys to assess the impact of population-level prevention interventions: to enable detection of a statistically significant reduction of 25% or 40%.

The sample sizes were estimated using the spreadsheet tool developed by SACEMA (5) adapted to an R programme (19,20). A design effect of 1.3 was used for the proportion of people living with HIV and the proportion recent among positives. The MDRI was estimated using pooled longitudinal data on subtype C specimens from CEPHIA. A supplementary viral load >1000 copies/mL at LAg optical density 1.5 was estimated to reduce the MDRI by 43 (95% CI 29–60) days; the viral load effect using a threshold of 100 copies/mL at LAg optical density 2.0 reduced the MDRI by 16 (95% CI 8–25) days. Further adjustments (24 days) were made for the nucleic acid amplification

Table 2. Subtype C MDRI estimates for sample size calculations

<table>
<thead>
<tr>
<th>Assay and recent infection testing algorithm (cut-offs)</th>
<th>Adjusted MDRI (days)</th>
<th>Specimen requirements</th>
<th>HIV screening requirements</th>
<th>Feasibility</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAg (1.5) + viral load (1000 copies/mL)</td>
<td>130</td>
<td>Dried blood spot (DBS) or plasma</td>
<td>Antibody only</td>
<td>High</td>
<td>Current</td>
</tr>
<tr>
<td>LAg (1.5) + viral load (1000 copies/mL)</td>
<td>154</td>
<td>DBS or plasma</td>
<td>Pooled nucleic acid amplification testing + antibody</td>
<td>Moderate</td>
<td>Current</td>
</tr>
<tr>
<td>LAg (1.5) + viral load (100 copies/mL)</td>
<td>160</td>
<td>Plasma</td>
<td>Antibody</td>
<td>High</td>
<td>Current</td>
</tr>
<tr>
<td>LAg (1.5) + viral load (100 copies/mL)</td>
<td>184</td>
<td>plasma</td>
<td>Pooled nucleic acid amplification testing + antibody</td>
<td>Moderate</td>
<td>Current</td>
</tr>
<tr>
<td>LAg (2.0) + viral load (100 copies/mL)</td>
<td>222</td>
<td>plasma</td>
<td>Pooled nucleic acid amplification testing + antibody</td>
<td>Moderate or low</td>
<td>Current (?)</td>
</tr>
<tr>
<td>New assay</td>
<td>300</td>
<td>To be determined</td>
<td>To be determined</td>
<td>To be determined</td>
<td>Future</td>
</tr>
</tbody>
</table>

*50% of specimens yielded positive results on the Aptima® HIV-1 RNA qualitative assay (Hologic) 26 days before Western blot seroconversion. An adjustment of MDRI by 24 days was used to account for the loss of sensitivity associated with pooled nucleic acid amplification testing (11).
testing screening assay (11) (Michelle Owen, CDC, personal communication, 2016), and the FRR was assumed to be 0.25%. Table 2 presents MDRI estimates for the sample size calculations.

The estimated sample sizes were shown for incidence point estimates for 14 African countries using a relative standard error of 20% and MDRIs of 130, 154, 184, 222 and 300 days. People 15–49 years old accounted for about two thirds of the population, and it was therefore assumed that a sample size <20 000 would be feasible for population-based HIV impact assessment (total feasible sample size 30 000). Eight countries had feasible sample sizes (<20 000) using an MDRI of 130 days; this increased to nine countries at an MDRI of up to 184 days, 11 at 222 days and 13 at 300 days (all but Kenya). The calculations did not account for variation in subtype. Only one country had a feasible sample size (Lesotho: <20 000) for population-based HIV impact assessment to detect a 40% reduction in incidence over two time points, at a minimum MDRI of 154 days. This increased to two countries at 184 days, four at 222 days and six at 300 days. Feasible sample sizes to detect a 25% reduction could not be achieved in any of the countries at any of the studied MDRIs.

For key population surveys, a feasible sample size was considered to be <5000, which may be considered high depending on the population. Because the HIV incidence and prevalence are high in these groups (prevalence between 16% among men who have sex with men 18–21 years old in Thailand and 52% among South African women >18 years old in KwaZulu Natal), population surveys for incidence point estimates were feasible in most cases using an relative standard error of 20% at the studied MDRIs. For impact assessment scenarios, very few key populations were able to meet the <5000 sample size condition.

The limitations of this analysis were that the sample sizes of a survey may be driven by factors other than incidence estimation, the FRR and design effect values were assumed and may not have been appropriate in some settings and the sample sizes (and FRRs) may have been underestimated in countries with non–subtype C HIV variants.

Incidence assays with longer MDRIs and lower FRRs will enable HIV impact assessment and result in more cost-effective national and key population surveys. It was commented that countries are currently conducting serial cross-sectional incidence estimates, and interpreting a change in incidence resulting from the incorrect application and design of survey methods must be prevented.

### 3.6 ARV medicine testing component in recent infection testing algorithms

Since ARV medicine testing has been applied to some national surveillance surveys in the past (in South Africa), the use of ARV medicine testing in a recent infection testing algorithm for estimating incidence is currently being debated. The purpose of this session was to present the findings of the recent infection testing algorithms with and without ARV medicine testing to inform recommendations for the use of this testing data. The national HIV household survey in South Africa in 2016 will collect DBS specimens for HIV antibody screening (and PCR among infants); positive specimens will be tested for HIV incidence (among people >2 years old), ARV medicine, viral load and HIV drug resistance. The recent infection testing algorithm applied will use the LAg assay (optical density 1.5); recent specimens will be tested for ARV medicine and, those testing negative, additionally for viral load (viral load >1000 copies/mL). Testing for ARV medicine will use high-performance liquid chromatography coupled to tandem mass spectrometry for qualitative determination of the ARV medicines nevirapine, efavirenz, lopinavir, atazanavir and darunavir. It was assumed for this work that the presence of ARV medicine indicates long-term infection. (The Working Group participants debated this assumption.)

The DBS specimen sizes are 3 by 3.2 mm punches, and the limit of detection was 0.2 μg/mL. ARV medicines can be detected in blood for 24–48 hours and efavirenz for up to 4–6 days.

In the 2012 survey, 2758 people living with HIV were tested for recent infection, of which 195 were recent; 99 were negative for ARV medicine, among which 10 had a viral load <1000 copies/mL (assumed to have been elite suppressors or people on treatment holidays), leaving 89 subjects. Examining the recent specimens by viral load and ARV medicine status, 10 of 63 (16%) ARV-negative specimens had a viral load <1000 copies/mL, and 13 of 66 (20%) specimens with a viral load >1000 copies/mL tested positive for ARV medicine, demonstrating the false recent cases missed if only one of these tests is incorporated into the recent infection testing algorithm. Of the 13 LAg recent on ARV medicine with a viral load >1000 copies/mL, 11 were women and none was pregnant; five had viral load >20 000 copies/mL; they were unlikely to have been on treatment holidays but could have been late ART initiators who had viral failure or recently infected people who had initiated ART and not yet obtained suppressed viral loads. Since pregnancy had been assessed by self-report only, some may have been pregnant.

The FRR had been chosen to be 0% and MDRI 130 days (21). HIV incidence among people 15–49 years old was estimated to be 1.72% (95% CI 1.38–2.06%) and was compared against estimates from other models (12): synthetic cohort (average annual incidence rates between 2009 and 2012) 1.9% (95% CI 0.8–3.1%), UNAIDS EPP/Spectrum (v.5) 1.5% (95% CI 1.43–1.62%) and THEMBISA 1.47% (95% CI 1.23–1.72%). The incidence estimates were compared using different recent infection testing algorithms: LAg + ARV medicine (1.98%), LAg + viral load (2.24%) and only LAg (3.58%), illustrating that adding viral load and ARV medicine
3.7 Real-world experience with LAg

Despite the high sensitivity and specificity of enzyme-linked immunosorbent assays, testing large numbers of specimens, such as in household surveys, may produce many false positives. Misdiagnosed false positives may be diagnosed as recent infections, and a testing algorithm should be used that accounts for this (22–24). Higher rates of false positives were observed in eastern Africa (thought to have resulted from people who had many more antibodies having been exposed to more infections). This was stated to have implications for the UNAIDS 90–90–90 target if the initial prevalence estimates were too high. The CDC algorithm for recency testing includes serology confirmation for any avidity result <0.4.

Data from the Swaziland HIV Incidence Measurement Survey (SHIMS) were presented as an example of the application of LAg in a population with high HIV incidence and prevalence. Cross-sectional HIV incidence testing was conducted among 18 172 people; 5802 were positive and were tested for recency and viral load. Pooled nucleic acid amplification testing was performed among negative specimens to identify acute infection. HIV-negative people were followed up and retested after 6.5 months (n = 11 944); 145 were positive, enabling the observed incidence to be calculated and compared with the initial LAg estimate. Overall prevalence was high (31.8%, 95% CI 30.82–32.18%) and was higher among women than men (38.10% (95% CI 36.89–38.72%) versus 21.93% (95% CI 20.79–22.73%)). The directly observed incidence was 2.38% (95% CI 2.06–2.75%) overall, 1.65% (95% CI 1.28–2.11%) among men and 3.14% (95% CI 2.63–3.74%) among women. The estimated incidence from nucleic acid amplification testing (assumed window period 15 days) was 2.58%, 1.32% among men and 3.6% among women (confidence intervals not calculated because of uncertainty around MDRI). Applying LAg + viral load, with an FRR of 0%, at baseline, the HIV incidence was 2.59% overall (95% CI 1.73–3.46%), 1.42% among men (95% CI 0.62–2.23%) and 3.55% among women (95% CI 2.29–4.81%). Adjusting for an FRR of 0.3% (found in Zimbabwe), the incidence was slightly lower, with 2.22% (95% CI 1.67–2.77%) overall, 1.57% (95% CI 0.94–2.19%) among men and 2.75% (95% CI 1.95–3.56%) among women. Analysis of the LAg estimates by age and sex showed that the incidence among men was higher at 18–24 years (~2%), peaked at 25–34 years (2.5%) and subsequently declined at 35–49 years (~1%). Among women, the incidence was highest among those 18–24 years old (4.2%) and declined with age (2.0% for those 25–34 years old and ~1.5% among those 35–49 years old). Geographically, the data showed two hotspot clusters in the cities Mbabane and Manzini for both directly observed and LAg estimates.

A further example presented was the Zvitambo study, which enrolled postpartum women for cross-sectional testing and a longitudinal cohort in Zimbabwe from October 1997 to January 2000. Women were tested and enrolled within 96 hours of giving birth. Of HIV-negative women followed up and tested at regular intervals for 24 months (90% follow-up at 12 months of 9562 at baseline), 351 seroconverted (182 women had ≥2 specimens for MDRI calculations). Of the 4495 women positive at baseline, about 3000 were seen at 12 months and included in an FRR study. A study was conducted to examine whether LAg provided the same incidence estimate for a time frame matching the directly observed incidence. In this population, the window period was 112 days at optical density 1.5 (subtype C), and the FRR was 1.3–1.6%. The results showed that the MDRI was shorter among women who became positive after the first year of pregnancy compared with the second (101 versus 140 days). It was suggested that postpartum women may have an enhanced response with increased antibodies, which will have implications for cross-sectional studies. The estimated postpartum incidence using LAg + viral load, MDRI 101 days and FRR 2.6% was 3.1%; the observed incidence in the cohort was 3.4%. At baseline, the incidence was higher (5.1%). This was similar to observations in Uganda, where the postpartum incidence was twice that seen in other women.

The application of LAg in a population with low HIV incidence and prevalence was shown using data from the Namibia Blood Bank between February 2012 and June 2013. Among 14 352 donations, 62 were antibody positive. Among repeat donors, the prevalence was 0.31% (95% CI 0.21–0.44%) and directly observed incidence 304 (95% CI 216–430) per 100 000 person-years. Nucleic acid amplification testing identified 11 positive cases, generating an incidence of 280 per 100 000 person-years. Among all donors, using LAg, 16 of the 62 positive cases were recent infections, and the prevalence is estimated to have been 0.43% (95% CI 0.33–0.55%) and incidence 309 (95% CI 153–465) per 100 000 person years. Using LAg, the prevalence was higher among first-time versus repeat donors: 0.74% (95% CI 0.5–1.05%) versus 0.31% (95% CI 0.21–0.44%). The incidence was higher among repeat donors: 405 (95% CI 196–614) versus 62 (95% CI 0–199) per 100 000 person-years at baseline. This was also slightly higher than that found in the directly observed estimates, although the inter-donation interval was not taken into account, which may have been shorter than the MDRI. This, however, demonstrated that LAg was able to generate useful estimates even in a population with low HIV incidence and prevalence.

The population-based HIV impact assessment countries in 2015 were Malawi, Zambia and Zimbabwe. Surveys planned for 2016 were for Cameroon, Namibia, Uganda and the United Republic of Tanzania and, for 2017, Côte d’Ivoire, Ethiopia, Haiti and Kenya. In population-based
HIV impact assessment surveys, training is currently underway for HIV, syphilis, CD4, Bio-Rad Geenius™ HIV 1/2 Supplemental Assay, early infant diagnosis, viral load (plasma and DBS) and LAg testing. LAg testing will be performed at the end of the surveys in the survey country. Training also includes quality control, quality assurance and data management. To date, 520 staff members have been trained by 18 people over 21 trips.

3.8 External Quality Assurance Program Oversight Laboratory for incidence assays

The External Quality Assurance Program Oversight Laboratory group in collaboration with the Blood Systems Research Institute and the CDC are developing an external quality assurance programme for the LAg assay to assess the reliability of results. The purpose of this session was to provide an update of current and planned activities of the external quality assurance programme for LAg. Initial programme objectives were to develop and test distribution materials, to evaluate data from multiple sites to determine the proficiency assessment criteria and to establish samples and a sample testing scheme. The Blood Systems Research Institute is responsible for collecting and shipping specimens. The programme will oversee programme management, statistics, reports and remediation, and the CDC will provide assay expertise and participate in the programme as a site. The algorithm laboratories are to follow was assessed; initial screening was performed in singlet, and specimens with optical density <2.0 repeat tested in triplicate; optical density <1.5 was considered a recent infection. Confirmation testing for HIV antibodies was performed on specimens with optical density <0.4.

Standard operating procedures for the LAg assays were developed in November 2014, and pilot sites were recruited and questionnaires distributed the following month. In January 2015, the pilot site results were received, which were reviewed with the sites in April. In July, approval was granted for an additional send-out. In January 2016, panel 1 was distributed and the results automatically uploaded onto the programme website, enabling laboratories prompt access. Five laboratories took part across the United States of America and the United Kingdom. The results of panel 1 included information on the sample scheme, the assay and serology kit types used, site questionnaires, panel 1 included information on the sample scheme, the assay and serology kit types used, site questionnaires, and specimen volume.

was insufficient. The responses were equally divided between the Sedia™ and Maxim assays; the experience of operators was rated between 3 and 5, with 5 being most experienced. One laboratory reported an incorrect incubation temperature; however, on further investigation, this was established to have been a clerical error.

The testing results revealed that all laboratories identified the negative sample. All but one laboratory appropriately classified the low-reactive recent specimens (one stated that it was a recent infection). The calibrator results of initial reactivity and confirmation reactivity were compared for the Maxim and Sedia™ assays and showed that the mean optical density of the Maxim calibrator was lower than that of Sedia™.

The optical density values of the samples tested within each of the laboratories using the various assays (Sedia™ or Maxim) showed good reproducibility of results but lower normalized optical density on the Maxim assay compared with Sedia™, and a wider spread among Sedia™ assay results between the different laboratories. On the negative specimen, the Sedia™ assay produced modestly reactive results (~0.1) but below 0.4, prompting confirmatory testing. Further investigation discovered several different lots of Sedia™ reagents. The Maxim results all stemmed from the same lot between the three laboratories. These implied that the differences in the results were probably related to lot variation. If future grading criteria are to include proficiency assessment of optical density values, these would need to be kit-specific because of differences across assays.

To date, the programme has distributed panel 1 to 17 sites. Each site received the 10 samples to run on Sedia™ or Maxim assays. The sites are to report data using the programme’s web-based system. The programme will examine the range of responses detected by laboratories, the precision of measurements among replicates and the ability to correctly classify recent infections. The programme is expected to continue for the next 2–3 years and may be expanded to include other assays.

3.9 Update on the case-based surveillance meeting

An overview of the case-based surveillance held in December 2015 in Glion, Switzerland was presented, and the future directions for WHO and UNAIDS to support the use of recent infection testing algorithms in case-based surveillance systems were discussed.

Case-based surveillance is a system for the routine event notification of newly diagnosed HIV infections (historically AIDS cases) from a facility to a subnational or national public health entity. The data reported may include a unique patient identifier, demographic information, clinical staging and immune status (CD4 or viral load),
transmission risk information and recency of infection (such as the last known negative test or recency testing). These data contribute to estimating the number of people newly infected and assessing the population-level impact of the epidemic. Case-based surveillance systems are now beginning to feature in low-income as well as middle- and high-income countries, and the data from these systems increasingly generate the information required for HIV care cascades.

WHO’s role in supporting case-based surveillance includes recommending countries to collect and report HIV and AIDS diagnoses and expanding antenatal clinic and sexually transmitted infection sentinel surveillance in sub-Saharan Africa. In the early 2000s, early assays for diagnosing recent infections were integrated into case-based surveillance in some countries (1). The December meeting discussed issues on incorporating HIV incidence assays into recent infection testing algorithms and estimating HIV incidence from case-based surveillance to inform the upcoming WHO consolidated guidelines on HIV case-based surveillance and patient monitoring. The objectives of the session were to review methods for estimating HIV incidence using case-based surveillance data and provide recommendations for countries planning to implement recent infection testing algorithms in case-based surveillance. Country experiences were presented from Brazil, Canada, England, France, Netherlands and the United States of America as well as other examples of incidence estimation tools such as Spectrum and key modelling papers (25–28). Many high-income countries have robust systems for case-based surveillance, but experience presents challenges with the completeness of clinical data and the reliability of transmission risk information. France, the United Kingdom and the United States of America rely in part on recent infection testing algorithms to inform incidence estimates, whereas other countries have developed methods adapted to the specific attributes of their case-based surveillance system.

Low- and middle-income countries have renewed emphasis on building case-based surveillance systems with limited planning for incorporating recent infection testing algorithms. Brazil has introduced recent infection testing algorithms in case-based surveillance in two provinces, and the CDC is considering supporting work in central Asia. The benefits of incorporating the recent infection testing algorithms into case-based surveillance are that the specimens used to diagnose HIV also can be used to determine recency, can potentially identify geographical hotspots or outbreaks and can enable improved understanding of the priority needs for prevention. The challenges of estimating incidence using case-based surveillance are that the number of newly diagnosed cases (and recent infections) depends on HIV testing patterns, which may be motivated by symptoms or risky behaviour. In addition, the laboratory infrastructure required for recent infection testing algorithms are considerably greater than those needed for current HIV diagnostic assays. The policies for returning recency results will need to be clear, especially for point-of-care testing.

The outcomes of the session identified many opportunities for monitoring the HIV epidemic using case-based surveillance. Countries with robust systems with testing for recent infection have been able to use the results to triangulate other estimates of incidence. Views were mixed, however, on whether recency testing is necessary given the availability of other clinical data (such as CD4 at diagnosis and testing history information). Countries initiating case-based surveillance may give priority to approaches to estimate incidence, which do not require recency testing.

A meeting report will be published describing the incidence estimation approaches using case-based surveillance data, and a manuscript on how to incorporate recent infection testing algorithms into case-based surveillance is currently being prepared. The WHO consolidated guidelines on HIV case-based surveillance and patient monitoring will include an annex on the considerations and recommendations for incorporating recent infection testing algorithms into these systems and a matrix listing various methods for HIV incidence estimation.

3.10 New assay validation by CEPHIA: Geenius™, ARCHITECT®, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3

The CEPHIA has evaluated more than 10 HIV incidence assays. In this session, CEPHIA presented the latest data on their evaluation of the Geenius™ HIV 1/2 Supplemental Assay, ARCHITECT® HIV antigen/antibody Combo assay, Glasgow-modified Genscreen™ HIV-1/2 Version 2 (Bio-Rad Laboratories) and IDE-V3 incidence assays (Luminex assay data were being analysed). The CEPHIA has three sets of developmental panels that contain 50 specimens with known recent or long-term infections for developers to use and a qualification panel (containing 250 specimens) for initial review and reproducibility. The specimens are primarily subtypes B (50%) and C (27%), with specimens also for subtypes A (12%), D (6%) and other recombinants (2%). Independent evaluation is conducted on an evaluation panel containing 2500 specimens. This panel contains a greater and more diverse representation of false-recent specimens and diversity of subtypes (overall 1168 specimens from 474 men who have sex with men and 1256 specimens from 445 heterosexuals). Most of the specimens are from North America (52%) and Africa (45%) and few from South America (3%). It includes specimens from 655 acute infections with known duration of infection, 155 other people who are ARV medicine naïve, 36 elite controllers, 163 people treated with ARV medicine
and 128 people diagnosed with AIDS, with multiple specimens available for most. One in five specimens are from 270 people with a duration of infection <6 months, with others infected up to >5 years.

The Bio-Rad Geenius™ HIV 1/2 Supplemental Assay is a HIV discriminatory and confirmatory assay. It is a lateral flow-based test that measures antibody responses by band intensity to four HIV-1 antigens (gp41, gp160, p31 and p24) and two HIV-2 antigens (gp36 and gp140) (band intensity data are not normally available to the user; Bio-Rad will need to amend how these data are made available). A Geenius™ recency index was defined as the sum of readings for gp41, gp160 and p31 bands divided by the reading for the control band. A result below a chosen threshold, proposed to be 1.5 by developers, was interpreted as recent infection.

The ARCHITECT® HIV antigen/antibody Combo assay is a chemiluminescent microparticle immunoassay for detecting HIV p24 antigen and antibodies to HIV type 1 and/or type 2 in serum or plasma. HIV antibody–positive specimens were tested in duplicate and subjected to a pre-analysis 1:10 dilution, one aliquot diluted in 1 M guanidine and the other in PBS. The difference in reactivity between the two aliquots enabled an avidity index to be calculated; <80% indicated a recent infection.

The Glasgow-modified Genscreen™ HIV-1/2 Version 2 avidity assay is now routinely used in Scotland. Plasma samples were initially diluted 1:400 in sample diluent and added, in duplicate, to two wells of a 96-well plate. These were washed for 5 minutes at room temperature in either kit buffer or 7 M urea. An avidity index was calculated by taking the optical density of the urea-treated sample and dividing it by the optical density of the wash buffer–treated sample. Samples with an optical density <1.0 were retested at a 1:100 dilution. If the optical density was ≥4.0, then these were retested at a dilution of 1:1600. Specimens with an avidity index of <40% were considered recently infected.

The IDE-V3 assay is based on two conserved highly immunogenic epitopes found in the envelope glycoproteins of HIV-1. One is derived from the immunodominant epitope (IDE) of the transmembrane gp41 glycoprotein and the other from the V3 loop of the gp120 outer glycoprotein. The IDE antigen comprises two consensus oligopeptides of 30 amino acids from HIV-1 group M and from subtype D. The V3 component comprises a blend of oligopeptides derived from the HIV-1 subtypes A, B, C, D and CRF01_AE. A dilution of each specimen was tested against both the IDE and V3 antigens. To discriminate recent from long-standing infection, this assay employed a mathematical formula comparing the reactivity of the specimen with the two antigens. An advantage of this assay is that it can be performed on DBS, although it has short shelf life (IDV-3 plates: 1–2 months).

The MDRI and FRRs were presented for varying thresholds of the assays. At optical density 1.5 for Geenius™, the MDRI was 179 (95% CI 154–205) days and FRR 6.06% (95% CI 3.17–10.35%) (among ARV medicine–naïve specimens). Supplementary viral load >1000 copies/mL reduced the MDRI to 129 (95% CI 110–149) days and the FRR to 5.05% (95% CI 2.45–9.09%).

At an avidity index of 80%, the ARCHITECT® HIV antigen/antibody Combo assay had an MDRI of 128 (95% CI 105–153) days and an FRR of 1.52% (0.31–4.36%), which is below the 2% recommended in the target product profiles. Addition of viral load reduced the MDRI to 109 (95% CI 88–133) days; the FRR remained at 1.52% (95% CI 0.31–4.36%).

For Glasgow-modified Genscreen™ HIV-1/2 Version 2 at an avidity index of 40%, the MDRI was very low (88 (95% CI 77–101) days) and the FRR was low (0.96%, 0.20–2.77%). With viral load, the MDRI declined to 71 (95% CI 61–83) days and the FRR remained stable. The FRR remained stable up to an avidity index of 60% (MDRI 149 days). At an avidity index of 85%, an MDRI of 231 (95% CI 207–256) was possible with an FRR of 2.07% (95% CI 0.80–4.33%).

IDE-V3 at a cut-off value of 0.5 produced an MDRI of 216 (193–240) days and an FRR of 5.73% (95% CI 3.43–8.91%). Adding viral load made the MDRI 158 (95% CI 139–179) days and generated an FRR of 3.03% (95% CI 1.43–5.58%).

The target product profiles criteria considered were an MDRI of 4 months, an FRR of <2%, a stand-alone assay, a shelf life of at least nine months and minimal training required. The Glasgow-modified Genscreen™ HIV-1/2 Version 2 assay may meet the target product profiles with an increased MDRI. All assays used low volume of specimen and required minimal training.

### 3.11 Update on the development of rapid assays and the potential applications and advantages

The purpose of this session was to provide an update on the effort to develop assays to translate the limiting antigen-based assay concept developed by the CDC into a rapid test format to simultaneously diagnose HIV and identify recent infection. Currently, Sedia Biosciences, Immunetics and Maxim are developing rapid test platforms. The design envisaged was a strip with a control line and two lines, one with high and one with low antigen concentration. Depending on high or low avidity, the three lines show HIV infection and indicate recency (all three lines signify a positive, long-term infection; no binding antibody presents a weak line). A characterized panel of specimens used to examine the performance of these tests
showed that they worked well. Sedia Biosciences initiated commercial kit development, and the CDC is currently evaluating this. The Sedia Rapid Incidence strip is a dipstick test. It has an optional reader providing a quantitative reading for each line, which can be compared against the cut-off. Sedia Biosciences has evaluated the assay on a panel with LAg data ($n = 500$) and showed good correlation. The CDC is conducting a larger evaluation with cross-sectional panels. If the results are promising, seroconversion panels will be used for evaluation. Testing among long-term infections will determine the proportion of false recent. Consistency among the three lots and between operators will be reviewed along with quality control of specimens. The Immunetics rapid recency test is similar in principle but uses two strips including a control and multiple antigen lines in a cassette format.
The Working Group was asked whether a technical update for the application of incidence assays was required based on the work presented during the meeting. This may include the outcome of the evaluation of assays by CEPHIA and issues concerning ARV medicine testing, the increasing prevalence of PrEP and HIV subtype variation. Another 10–12 population-based surveys collecting CD4 and viral load data were scheduled for 2016–2017.

It was stated that the Working Group should provide normative guidance to avoid potential confusion. In particular, the concept of minimum sample sizes to evaluate the impact of programmes and accounting for FRRs would be an important message to convey. In addition, it would be helpful to provide clarity to users on appropriate MDRIs and how these are affected, thereby ensuring comparable surveys and incidence estimated over time.

An annex to the guidelines for measuring national HIV prevalence in population-based surveys (4) incorporates issues concerning sample sizes based on the work completed in the previous June meeting. However, it was suggested that it would be helpful for it to be summarized in a smaller directive document associated with incidence testing (such as 2–4 pages listing current issues put out with a press release), allowing users to apply these assays with more confidence. In addition, the Working Group was considered a forum to review how assays are applied in the field and the type of results generated. It was stated that the role of WHO/UNAIDS as a provider of normative guidance is crucial and the Working Group is the vehicle. It was proposed to update the terms of reference to clarify the role and areas of work for the next two years.
5. TARGET PRODUCT PROFILES FOR HIV INCIDENCE ASSAYS

5.1 New target product profiles for HIV incidence assays (presented by Peter Dailey, FIND)

The purpose of this session was to obtain feedback and consensus on target product profiles and to use cases of HIV incidence assays. A Delphi process was used for reviewing the draft target product profiles. Stakeholder interviews were conducted to develop preliminary use case scenarios, which were reviewed by the working group on target product profiles. This resulted in eight use cases consolidated into three target product profiles (Table 3), each including 20–30 characteristics. A survey was distributed among a broader group to obtain consensus on the optimum and minimum requirements for target product profiles, and the results were reviewed at the meeting.

Survey participants had received questionnaires for each target product profile on optimum and minimal requirements. The results were reviewed at the meeting.

Table 3. Main use cases for tests of recent infection for HIV and corresponding target product profiles

<table>
<thead>
<tr>
<th>Use</th>
<th>Description of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uses for tests for recent HIV infection related to estimating incidence</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Point estimates of incidence</strong></td>
<td></td>
</tr>
<tr>
<td>National surveillance</td>
<td>National estimate of incidence; may be part of a broader demographic study</td>
</tr>
<tr>
<td>Programme, prevention or trial planning</td>
<td>Incidence estimate in subpopulations for planning, setting priorities or other instances when an estimate of incidence is required. Often may be for only a city or region</td>
</tr>
<tr>
<td>Key or sentinel populations</td>
<td>Incidence estimates in special subpopulations using targeted sampling methods*</td>
</tr>
<tr>
<td>Case-based surveillance</td>
<td>National or regional incidence estimates via case-based reporting of newly identified people living with HIV</td>
</tr>
<tr>
<td><strong>Impact assessment</strong></td>
<td></td>
</tr>
<tr>
<td>National surveillance</td>
<td>Comparing a reduction in incidence before and after an intervention to assess the impact of interventions at the national level</td>
</tr>
<tr>
<td>Key or sentinel populations</td>
<td>Comparing a reduction in incidence before and after an intervention to assess the impact of interventions in key populations</td>
</tr>
<tr>
<td><strong>Uses for tests for recent HIV infection not related to estimating incidence</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Individual level</strong></td>
<td></td>
</tr>
<tr>
<td>Research purposes</td>
<td>Identification of individuals with “recent” infections for multiple potential applications, such as recruiting recently infected individuals into longitudinal cohort studies</td>
</tr>
<tr>
<td>Individual management</td>
<td>Identification of people with recent infections to guide clinical management and/or public health programmes, such as selecting therapy and/or giving priority to contact tracing</td>
</tr>
<tr>
<td><strong>Population level</strong></td>
<td></td>
</tr>
<tr>
<td>Targeted prevention planning</td>
<td>To provide population-level data on recent infections to enable risk factor analysis or identify hot-spots to inform targeted prevention planning (no incidence estimate is obtained)</td>
</tr>
</tbody>
</table>

*Respondent-driven sampling method (www.respondentdrivensampling.org).
characteristics using the Likert scale for scoring (1 = disagree, 5 = fully agree). The survey was conducted in two rounds: first among the working group on target product profiles and second among the larger stakeholder input group. Before the results were examined, consensus was defined as >50% of participants agreeing with the proposed criteria for each characteristic. In total, 21 responses were collected. Nearly all (95%) had >10 years of experience in HIV incidence. The responses were from a variety of areas, including sub-Saharan Africa, the United States of America, South-East Asia, Europe and South America. Half the respondents identified themselves as researchers as well as laboratory experts and programme implementers. Policy-makers, employees of NGOs and personnel in the in vitro diagnostics industry were also among the respondents. Familiarity with public health and research use cases for tests for recent infection was high apart from for individual management (use case 7). Consensus was achieved (>50% agreement) for all but one of the characteristics for target product profile A (95% of the characteristics of target product profile A exceeded 50% agreement), and all but two on target product profile B (90% of the characteristics of target product profile B exceeded 50% agreement). Three characteristics differed between target product profiles A and B: test performance (MDRI and FRR), throughput and the time the analyte must be stable in specimen storage format.

The performance of tests for recent infection is usually measured with metrics such as the MDRI and FRR. To ensure a minimal acceptable test performance, pairs of MDRI and FRR values were derived under the assumption of a hypothetical scenario, with HIV prevalence of 5% and HIV incidence of 0.3% in the tested population. This scenario corresponds to the lowest prevalence and incidence settings recommended by WHO for incidence surveillance via cross-sectional surveys (29).

The rationale used to derive a combination of MDRI and FRR that would define a target product profile was to consider the sample size needed for a survey to obtain a reliable incidence estimate. The sample size is considered to be a critical parameter, since it is a crucial driver of cost and feasibility. It was decided that the minimal target product profile should enable an estimate of incidence screening to be derived in a population ≤30 000 individuals, and an optimal target product profile should enable a population of ≤10 000 individuals to be screened (except for key populations, where the bounds were set at ≤10 000 and ≤5000 respectively).

Since MDRI and FRR are related, several different pairs of values may meet the criteria based on sample size. In this case, they are all considered equivalent. Additional factors influencing the survey sample size are the baseline HIV incidence and prevalence (when the intention is to measure a change in incidence before and after an intervention), the assumption on the design effect and the desired precision and power: these values were kept constant during all the evaluations.

5.2 Discussion of target product profiles

The characteristics of target product profiles with low agreement from the survey were presented to obtain feedback on any modifications to the characteristics that did not obtain consensus. Table 4 lists the characteristics of target product profiles and the resulting percentage consensus achieved for minimal and optimal characteristics from the survey. The characteristics highlighted in grey were discussed at the consensus meeting.

The group discussion and recommended revisions to the target product profiles are summarized below for each characteristic discussed.

5.2.1 Test performance characteristics

Overall, for target product profile A (use cases included national surveys, programme prevention or trial planning, key populations and targeted prevention planning) there was considerable discussion about the key population use case and whether decoupling key populations from target product profile A might be better. Further, the sample size used for the minimum criteria for key populations was recommended to be reduced. It was noted that sample techniques for key populations are a limiting factor and that feasible sample sizes vary widely depending on the key population of interest. Some suggested that sample sizes around 2500 would be feasible, and others suggested that sample sizes as low as 1000 were more likely. The sample sizes will be revised, and the working group on target product profiles will review splitting off key populations into a different target product profile. Options to include it in target product profile B or to break down the test performance characteristics based on use case were suggested as two possible approaches for the revised target product profiles.

For target product profile B (use cases for assessing the impact of population interventions and case-based surveillance), there was consensus that the proposed sample size was too large for case-based surveillance and that few countries currently reported such numbers. The sample sizes required for calculating incidence using case-based surveillance data depend on the expected number of diagnoses and recent cases, for which the CDC has guidelines. It was commented that, in the market assessment exercise (see Section 6), most tests for recent infection were used for case-based surveillance. In addition, many countries are currently reviewing the strengths and weaknesses of various incidence estimation models using case-based surveillance, and tests for recent infection have large resource implications compared with other methods. An assay that may only meet the requirements for a minimum sample size of 10 000 people
tested was not considered competitive in such settings. There was consensus that combining the target product profiles for impact assessment and case-based surveillance is inappropriate.

It was suggested to add “national” to use case 4 and to add a new use case assessing the impact of interventions in key and sentinel populations (suggested sample size 5000). Previously presented data highlighted challenges in assessing changes in key populations because of sample size restrictions. A proposal for the impact of interventions was that the optimal target product profiles could be specific to key populations and the minimal for national evaluations, omitting the need for another target product profile.

5.2.2 Time the analyte is stable in specimen storage format

For target product profile A, there was 76% agreement for minimal (one year) and 60% agreement for optimal (also one year). For target product profile B, there was 81% agreement for both minimal (one year) and optimal (three years) time periods. Feedback from the survey:

Table 4. Summary of consensus results from the survey on target product profiles

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Target product profiles A and B</th>
<th>Target product profile A only</th>
<th>Target product profile B Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Target user</td>
<td>100% 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Specimen storage conditions in the laboratory</td>
<td>100% 83%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Specimen preparation in the laboratory</td>
<td>94% 88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Data export to user</td>
<td>89% 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Specimen preparation at the point of collection</td>
<td>88% 82%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Other data export</td>
<td>88% 100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Product registration and regulatory path</td>
<td>88% 80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Infrastructure level</td>
<td>83% 72%*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Shipping conditions for test kits</td>
<td>82% 88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Reagent stability</td>
<td>82% 82%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Target countries</td>
<td>78% 72%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Stability of the specimen between collection and arrival at the laboratory</td>
<td>76% 94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Requirement for supplementary tests to achieve the desired FRR</td>
<td>67% 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Target cost per result</td>
<td>61% 83%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Time from assay start until the results are available</td>
<td>59% 53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Specimen type</td>
<td>56% 89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Recognition of HIV subtypes and circulating recombinant forms</td>
<td>47% 94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Test performance (MDRI and FRR)</td>
<td>59% 76%</td>
<td>44% 63%</td>
<td></td>
</tr>
<tr>
<td>19 Throughput</td>
<td>75% 75%</td>
<td>81% 87%</td>
<td></td>
</tr>
<tr>
<td>20 Time the analyte must be stable in the specimen storage format</td>
<td>71% 60%</td>
<td>81% 81%</td>
<td></td>
</tr>
</tbody>
</table>
Optimal storage would be more than one year. Since many countries have samples stored for >1 year, consider extending the optimal time cut-off if the samples are in good condition. Optimal: should be 5 years or longer. Prefer if up to 5 years.

DBS samples may have to be processed in less than 12 months if viral load testing is part of the recent infection testing algorithm.

It was suggested that differentiation by specimen type may be required and that other tests forming part of the algorithm may be restrictive. There was consensus to increase the optimal time period to 5–10 years, although many samples may have been stored for longer.

5.2.3 Specimen types

There was 56% agreement for minimal specimen types that any of the following would be acceptable: whole blood, urine, saliva, peripheral blood mononuclear cells and stool. There was 89% agreement on the optimal being an easy-to-collect specimen requiring minimal training (such as finger-prick blood and DBS). The rationale for including all the specimen types in the minimal criteria was that, if a large MDRI and small FRR were to be achieved, an attempt would be made to collect any sort of specimen. The group at large did not feel any changes were required to the target product profile characteristic but that it be clearly stated that any of the specimen types would be acceptable as a minimal criterion.

5.2.4 Recognition of HIV subtypes and circulating recombinant forms

There was 47% agreement on minimal performance requirements for subtype C. There was high agreement (94%) that the optimal should be for subtypes A, B, C, D and major circulating recombinant forms. It was commented that many countries with high usage of tests did not have subtype C prevalent. There was consensus that subtypes B and C should be included in the minimum requirement.

5.2.5 Target test cost

The target cost per test had 61% agreement for minimal up to US$ 5 and 83% agreement for <US$ 5 as optimal. Comments included:

- Minimal: US$ 5 per stand-alone test result, US$ 10 per recent infection testing algorithm; optimal: <US$ 2.5 per test result (either stand-alone or recent infection testing algorithm).
- I would think we can tolerate higher costs under minimal.
- The cost advantage may depend on the quality of the test.

I think US$ 5 is likely not achievable for any final result that requires a confirmation or multiple test algorithm.

It was suggested to refer to this as a target cost per assay rather than result, since this may be misunderstood as the result of the recent infection testing algorithm.

5.2.6 Supplementary tests

There was 67% agreement for minimal requirements to include supplementary tests to achieve the desired FRR (not more than three). There was 89% agreement that the optimal should be that no other tests are required. This was considered to depend on the donor and whether there would be interest in the results of supplementary tests such as viral load for other purposes. Feedback from the survey was:

- I think allowing three supplementary tests is too open-ended – maybe two is a good place to draw the line, knowing actually that there is little appetite beyond viral load, maybe a second serology. But is dual serology a case of “primary” and “supplementary”?
- Minimal: three additional tests seem excessive but, ultimately, this is the decision of the end-user and what they are willing to accept.
- I am not sure that it is realistic to have supplementary tests to perform, particularly in resource-limited countries
- No more than one supplementary assay, since there might not be sufficient specimen collected to go around.
- Three supplementary tests are too many.

5.2.7 Time from assay start until the results are available

The Working Group discussion regarding this characteristic mainly pointed out that, for the use cases described in target product profile A, since the test is not performed at the point of care, this characteristic is not of great importance. We recommend removing this from the list of critical target product profile characteristics and only including it for target product profile C (individual use).

The Working Group agreed that the next steps were to update the target product profiles with the provided feedback and compare these with existing assays. The Sedia™ HIV-1 LAg-Avidity assay met the minimal requirement for target product profile A, although the FRR may be achieved only in certain settings (such as high HIV incidence and high relative standard error) (Table 5). The assay did not meet the minimal requirements for target product profile B.

Only six survey participants responded to questions on target product profile C (research and individual management), although all were highly experienced
There was significantly less agreement on the elements of the 17 characteristics. Technologies to support tests for individual use remain in the early stages of development and would be highly influenced by the setting and context in which they were to be used. Target product profile C may be redefined only once these technologies can provide accurate information on individual time since infection and the intended use of such assays are clearly defined to inform the key product requirements.

Table 5. Sedia™ HIV-1 LAg-Avidity enzyme immunoassay comparison to the key requirements of target product profiles A and B

<table>
<thead>
<tr>
<th>Key requirements</th>
<th>Target product profile A</th>
<th>Target product profile B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Use cases (incidence point estimates at national, subnational and key population levels)</td>
<td>Use cases (impact assessment and case-based surveillance)</td>
</tr>
<tr>
<td></td>
<td>Meets minimal requirements?</td>
<td>Meets optimal requirements?</td>
</tr>
<tr>
<td>Target setting, level, intended user, operating environment</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Type of result</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Supplementary test (recent infection testing algorithm) requirements</td>
<td>Yes, because the requirements permit supplementary tests</td>
<td>No, supplementary tests not permitted</td>
</tr>
<tr>
<td>HIV subtypes</td>
<td>Yes, performance met for subtype C</td>
<td>No, performance not met for some subtypes</td>
</tr>
<tr>
<td>Reagents in format amendable to the anticipated batch sizes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Accessories and the other supplies required are present at this level</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Stability of reagents</td>
<td>Yes</td>
<td>No (freezer required)</td>
</tr>
<tr>
<td>MDRI + FRR</td>
<td>Yes, MDRI of ≥120 days and FRR of ≤1.5% anticipated for LAg in certain populations</td>
<td>No</td>
</tr>
<tr>
<td>Time until the results are available</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample type and volume</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

[1]
6. UPDATED MARKET ASSESSMENT FOR HIV INCIDENCE ASSAYS AND BIOMARKERS IN THE PIPELINE

6.1 Updated market assessment (presented by Charles Morrison and Rick Homan, FHI 360)

In 2009, FHI 360 undertook a landscape analysis and market assessment for HIV incidence assays (30). This was updated in 2015 (unpublished) because of the improving technology of HIV incidence assays and their expanded application. The objectives of the assessment were:

• to determine the current and planned use of HIV incidence assays during the next three years;
• to estimate the changes in demand and contributing factors during the next five years,
• to construct potential market scenarios for the development of HIV incidence assays and to estimate demand during the next 5–10 years;
• to determine the cost of developing and delivering an assay based on a newly identified HIV incidence assay biomarker;
• to estimate the economic value of an accurate HIV incidence assay to global public health in terms of public health dollars spent: cost savings related to improved surveillance, evaluation, decision-making and expenditure on HIV prevention.

The methods applied to explore the first three objectives were key stakeholder interviews (n = 7), systematic interviews with experts (n = 20) and previously interviewed stakeholders (n = 2). Usage estimates by country and use case were based on expert informants, peer-reviewed articles, published reports, proceedings of international working groups and normative bodies and the actual HIV incidence assays used or sold in 2014 (Table 3 provide use cases for HIV incidence assays). Interviews were conducted among a range of organizations, including guidance bodies (n = 3, including WHO and UNAIDS), funding bodies (n = 6, including the CDC, the United States Agency for International Development and the United States National Institute of Allergy and Infectious Diseases), manufacturers (n = 3, Sedia Biosciences Corporation, Maxim Biomedical and Immunetics), developers (n = 2, including the CDC), a regulator (n = 1, the United States Food and Drug Administration), surveillance (n = 3, Istituto Superiore di Sanità (Italy), Institut de Veille Sanitaire (France) and Public Health England), clinicians (n = 2, University of California, San Francisco and Wits Reproductive Health and HIV Institute) and other users (n = 7, Blood Systems Research Institute, University of California, San Francisco; Johns Hopkins University; Epicentre; International Centre for AIDS Care and Treatment Programs at Columbia University; Human Sciences Research Council; and ICF International).

The number of incidence assays used in 2014 was calculated to be 308 893. The highest use was in Asia, dominated by China, with 138 364 (~45% of total global use). This was followed by the Americas (n = 56 050, 18%), Africa (n = 33 950, 11%) and Europe (n = 25 029, 8%). Some sold in 2014 could not be attributed to a specific region (n = 43 000, 14%). Most incidence assays were used in case-based surveillance (n = 188 564, 61%), followed by population-based surveys (n = 25 300, 8%), research evaluations of incidence assays (n = 17 300, 6%), sentinel surveillance (n = 11 929, 5%), assessing the impact of population-level interventions (n = 6000, 2%) and identifying individuals with recent infection for research purposes (n = 1200, <1%). There were also 58 600 (19%) used or sold in 2014 that could not be attributed to a specific use case. No incidence assays were found in use for use cases of subpopulation-level incidence point estimates or individual use (research purposes or individual management). Countries with the greatest use for national and population-based surveillance were Swaziland (47%) followed by Zimbabwe (19%), Zambia (16%), Botswana (8%), Lesotho (6%) and others (4%). The countries with the greatest use for sentinel surveillance were China (23%) followed by France (7%), Brazil (6%), South Africa (6%), Malaysia (6%) and others (7%). The greatest use for case-based surveillance was in China (62%), followed by the United States of America (20%), England (3%), Germany (3%), Thailand (3%) and others (9%). In 2014, the most frequently used incidence assay globally remained the Sedia™ BED HIV-1 incidence enzyme immunoassay with 145 800 (47% of all used or sold), followed by the Sedia™ HIV-1 LAg-Avidity enzyme immunoassay (n = 105 814, 34%) and the Bio-Rad Avidity assay (n = 43 600, 14%). Other incidence assays used (n = 13 679, 4%) included ARCHITECT® HIV antigen/antibody Combo assay, Geenius™, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3.
The planned global use from 2015 to 2017 is for case-based surveillance (may be discontinued in China but possibly increased in Europe and Africa), population-based surveys (by the CDC (population-based HIV impact assessments), United States Agency for International Development, Médecins Sans Frontières, HIV prevention trials (HPTN)) and patient management (contact tracing). The average planned annual use was calculated to be 108 281. The planned annual use during this period was highest for the Americas (n = 58 583, 54%) followed by Europe (n = 24 567, 23%), Africa (n = 12 445, 11%) and Asia (n = 11 436, 11%). Annually, the most incidence assays planned for use during this period continued to be for case-based surveillance (n = 72 803, 67%), followed by population-based surveys (n = 11 028, 10%), sentinel surveillance (n = 8200, 8%), assessing the impact of population-level interventions (n = 7500, 7%), research evaluations of incidence assays (n = 7250, 7%) and identifying individuals with recent infection for research purposes (n = 1500 or 1%). Use in population-based surveillance was to be 83% for population-based HIV impact assessment and 17% for non-population-based HIV impact assessment countries. The planned population-based HIV impact assessment countries included: Cameroon, Côte d’Ivoire, Lesotho, Malawi, Namibia, Swaziland, Uganda, United Republic of Tanzania, Zambia and Zimbabwe. Non-population-based HIV impact assessment countries included: Nigeria, Mozambique (United States Agency for International Development) and South Africa (Human Sciences Research Council). During this period, the planned incidence assay use was expected to shift from the Sedia™ BED HIV-1 incidence enzyme immunooassay (average use n = 14 100 or 13% of planned use/year) to CDC-modified G5 HIV-1/HIV-2 PLUS O enzyme immunooassay (n = 44 600 or 41%) and the Sedia™ HIV-1 LAg-Avidity enzyme immunooassay (n = 39 281 or 36%) assays. Several other assays, including the ARCHITECT® HIV antigen/antibody Combo assay, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3 were expected to be used in smaller numbers in countries in which they already are in use.

The changes in the global demand for incidence assays during the next 5–10 years were estimated and the factors influencing these considered. The following potential market scenarios were explored:

- scenario 1: current situation – no changes or improvements in the performance of incidence assays;
- scenario 2: an assay or assay algorithm with improved performance becomes available: for example, improved performance for people receiving ART, varying subtypes, less expensive or longer MDRI; and
- scenario 3: a game-changing scenario: for example, incidence assays find expanded use in patient care and management or an incidence assay or recent infection testing algorithm is used as part of rapid screening or diagnostic testing (a prevalence-incidence test).

For scenario 1, interviewees were asked to project the number of incidence assays their institution would use in 5–10 years if there were no changes in the available incidence assays. Most responded that the quantity of incidence assays used would remain the same; several said that there would be an increase of indeterminate size; few felt that unless the technology improved, demand would decline because measuring low incidence is difficult. The estimated annual number of incidence assays used under scenario 1 was 108 281 (±25% for upper and lower estimate bounds), which was similar to the planned use between 2015 and 2017. The largest number of assays would be used in the Americas followed by Europe and Africa. Most assays would be used for case-based surveillance, especially if case-based surveillance were to become more prevalent in Africa. No use was projected for use cases of subpopulation-level incidence point estimates or individual use (research purposes or individual patient management), mirroring the planned use for 2015–2017.

For scenario 2, most thought that demand would increase – from either modestly or moderately to doubling in number if there were to be better precision or lower cost (such as from a longer MDRI). People using HIV impact assessment primarily for case-based surveillance were of the view that there would be no change in use if the number of new HIV diagnoses in their country remained constant. One interviewee responded that demand would not change as survey sample sizes were driven by viral load estimates.

It was assumed that, with an improved assay, China would redeploy its use for case-based surveillance and that the number of assays used would increase modestly or moderately, with 304 670 annually (±25%). Use would be highest in Asia followed by the Americas, Europe and Africa. Use would be highest for case-based surveillance, assuming that the increases in use are proportional to current use patterns.

For scenario 3, most felt that there could be a dramatic increase if a recent infection testing algorithm combined with a screening test yielded incidence information. It could be used in clinics as part of routine care for the entire screening market. Some felt there could be a 2-to 10-fold increase if there were to be use for contact tracing, case finding, treatment management and screening. Others thought there would be an indeterminate increase and that it would predominantly be used by governments, PEPFAR, programme implementers and for research purposes but not by clinicians. When an incidence assay or recent infection testing algorithm was part of HIV confirmatory testing and could be used for individual management, based on all identified positive cases in a year for countries with case-based surveillance and estimates for the number of people newly infected with HIV in other countries (2014 estimates), estimated use was 1 992 698, with a lower bound of 1 672 155 (65% of the total). In this instance,
most assays would be used in Africa, where the number of newly diagnosed cases was highest.

If a HIV incidence assay or recent infection testing algorithm were to be part of all rapid screening and diagnostic tests used globally, based on WHO estimates of the number of HIV diagnostic tests used annually for low- and middle-income countries, the annual use was estimated to be ~180 million for low- and middle-income countries from 2020 to 2022. Tens of millions of additional tests would be used in high-income countries (data not currently available). Assuming that not all countries would adopt this new type of screening or diagnostic test in 5–10 years, a lower bound of at least 117 million assays (65% of 180 million plus the tests in high-income countries) is projected under this scenario.

A significant factor affecting the future use of assays is the need for greater confidence in assay results and incidence estimates. There are currently no guidelines on which assays or recent infection testing algorithms should be used for different use cases, resulting in confusion. Some said that interpreting the results is difficult. Others speculated that the demand for assays or recent infection testing algorithms would continue, since government and international agencies valued estimates of HIV incidence for measuring progress. However, one respondent felt that the priority was shifting towards treatment and viral suppression and away from incidence measures. It was noted that the declining incidence in many countries could undermine the ability to use assays or recent infection testing algorithms, since the necessary sample sizes would become too large. Other qualities of incidence assays that interviewees felt would affect future demand were assays with improved performance, having a test with multiple uses, staging in different ways and the potential to obtain information on recency among individuals.

The estimated cost of developing and delivering an incidence assay based on a new biomarker was presented. Assumptions about the type of technology were made and several likely scenarios evaluated. A Microsoft Excel-based cost-estimation tool organized around the four phases of development (technical feasibility, development, manufacturing transfer and external evaluation) and the parties involved (manufacturer, academic partner and an external evaluator (CEPHIA)) was used to estimate the full-time equivalent efforts on a monthly basis and the equipment and evaluation costs as needed. The full-time equivalents were valued at US$ 225 000 per year and represented a fully loaded cost per full-time equivalent. Two scenarios were examined corresponding to different levels of complexity: multi-well enzyme immunoassay and highly complex multiplex similar to Luminex, leveraging existing test platforms. Depending on the complexity of the incidence assay, the development cost was estimated to range from US$ 2.4 million to US$ 3.2 million over a three-year period. Most of the time and costs were devoted to the development phase. Other key assumptions were that the developer was ISO 13485 certified and that products would be used for research properties only and were not intended for patient management. Therefore, no costs for registration and certification for CE-IVD (in vitro diagnostic medical devices in Europe) or the United States Food and Drug Administration were accounted for.

This investment was compared with the economic value of an accurate incidence assay to global public health in terms of cost savings related to improvements to surveillance efforts, evaluation, decision-making and expenditure on HIV prevention efforts. Estimation was based on use cases to obtain incidence point estimates at the national level as well as for key populations for the following 2–3 and 5–10 years. The total annual sample sizes of surveys were based on past use, such as the estimated total annual cost of recent integrated biobehavioural surveys in Botswana and Ghana. Data from Parkin et al. (31) were used to estimate the reduction in sample size with MDRI increasing from 130 days to 154, 240 or 280 days and the associated reduction in variable costs. The estimated annual cost to prevent one person from becoming infected with HIV through cost-effective prevention interventions was taken from the literature. The maximum number of people prevented from becoming infected with HIV if annual survey cost savings were redirected towards the most cost-effective prevention interventions was estimated. The value of these averted infections was assessed based on the annual discounted present value of HIV treatment. Finally, the “return on investment” was estimated as the annual value of HIV infections averted divided by the annual savings from the improved incidence assays. The annual sample sizes for national- and key population–level incidence point estimates for 2–3 years were 300 731 and 405 987 for 5–10 years, and the total annual cost by use and time frame was US$ 46 026 950 2–3 years and US$ 61 584 646 for 5–10 years. The fixed costs were estimated to be US$ 220 000 and a variable cost of US$ 180 per subject for use case 1; fixed costs were US$ 110 000 and variable costs US$ 70 per subject for use case 3. Over the 2–3 years, about two thirds of subjects would be under the national incidence point estimates, at a cost of about US$ 37 million per year and key population incidence point estimates would cost US$ 9.2 million per year. Figures for 5–10 years represented about a 35% increase over 2–3 years to US$ 61.5 million per year.

The potential annual cost savings from reduced sample size requirements due to increased MDRI over 2–3 years were estimated to range from US$ 7.6 million to US$ 24.4 million per year. This was estimated to increase to about US$ 10.2 million to US$ 32.8 million per year over 5–10 years. The “return on investment” from increased MDRI survey cost savings versus the development costs ranged from US$ 4.4 million to US$ 30.4 million. If the resources saved from reduced sample sizes were to be invested in cost-effective HIV prevention interventions and
the investor chose the most cost-effective interventions, the estimated cost to prevent one person from acquiring HIV was estimated to be about US$ 58. Therefore, the savings could reduce the number of people newly infected with HIV by between 104 130 and 333 745 per year over the 2–3 years. This was estimated to increase to 140 575 to 450 556 per year over 5–10 years. Converting the number of HIV infections averted into an estimated public health savings using the discounted annual present value of HIV treatment cost (US$ 612), the annual cost savings in this best-case scenario were between US$ 63.7 million and US$ 204.2 million per year over 2–3 years, rising to US$ 86.0 to US$ 275.7 million per year over 5–10 years. These were considered optimistic figures but were thought to demonstrate how improving MDRI can reduce the resources required for estimating HIV incidence and the extent to which the funds could be put towards cost-effective prevention interventions.

6.2 Updated market assessment: discussion

For the market assessment, it was suggested to provide more detail on the assumptions related to reducing the sample size resulting from increased MDRI, since incidence is often not the primary driver of sample size calculations. A further suggestion was to review the information on successful interventions; regional estimates based on different interventions may be more precise (although WHO provided the estimates to prevent one person from becoming infected with HIV for sub-Saharan Africa). One query was how the investor would recover development costs if it were a diagnostics company, although it was suggested that the public sector could be the investor. The figures were deemed reasonable for scenario 3, but justifying investment from manufacturers would be difficult for scenario 2.

It was stated that Immunetics and Sedia are working towards developing a combined rapid test technology that can provide prevalence and incidence tests. The path forward for regulatory approval for a diagnostic recency-based assay would be through the United States Food and Drug Administration if it were for individual use. With a diagnostic component for prevalence, premarket approval through the Center for Biologics Evaluation and Research of the United States Food and Drug Administration would be required. From a performance perspective, the Asante test was thought promising, with discussions ongoing with the United States Food and Drug Administration on how this can be advanced. It was thought that the game-changing scenario may already be occurring with diagnostic assays that have a very long dynamic range: for example, ARCHITECT® in its normal mode can be interpreted as a diagnostic test and recency test with no additional cost. This also applies to the Geenius™ assay if the software were to be modified.

A consideration is that the cost of a regular rapid test is as low as US$ 0.25, and the additional cost of a combined test could therefore be a barrier to the market. Other work is ongoing to use incidence assays as markers of viral suppression as an alternative use.

6.3 Update on biomarkers in the pipeline

Several biomarker discovery efforts are ongoing, including those sponsored by the Bill & Melinda Gates Foundation, the United States National Institutes of Health and others that are independently funded. These exploratory biomarker efforts have used the CEPHIA repository panels for the discovery and proof-of-concept stages of development. Within the Bill & Melinda Gates Foundation group, the five projects are:

- circulating cellular microRNAs in plasma as biomarkers of HIV-1 infection;
- multiplex antibody markers, cell-associated viral load;
- detection of recent HIV-1 infections based on naturally inspired synthetic oligomers;
- novel gastrointestinal biomarkers for HIV incidence; and
- urinary metabolite HIV recency biomarker profile for incidence measurement.

The original intent was to explore many specimen types (such as stool and urine), but all promising projects that have been identified during the course of the discovery exploration resulted in antibody-based approaches in plasma.

Data were presented on the multiplex antibody HIV-1 incidence assay developed at Duke University, which screens for novel antibody subtype or epitopes associated with recent infection. A total of 281 antibody-antigen measurements were taken (Env gp41, gp120, non-Env antigens, peptide epitopes (including identification and new epitopes) versus IgG, IgG3, IgG4, monomeric IgA and dimeric IgA in a high-throughput screening modality. The output measurements were magnitude, dilution and avidity. Combinations of 4–5 antibody–antigen measurements were identified. The proof of concept MDRI was ~253 days and the FRR ≤~3%.

The detection of recent HIV-1 infections based on naturally inspired synthetic oligomers screened random shape libraries (peptoids) for antibodies associated with recency. The results from specimens divided into groups by the estimated time from seroconversion showed potentially useful temporal patterns; of the four peptoids, some showed decreasing and others increasing optical density with the duration of infection. The MDRI calculated from the panel was 342 (95% CI 227–537) days. This now requires validation on a subsequent panel.
For the use of microRNAs in plasma as biomarkers of HIV-1 infection, in early development, the samples used to develop the model showed separation between recent and non-recent specimens. On a second panel, for five microRNAs, the panel was divided into recent and non-recent specimens; none of the microRNAs associated with recency in the first round were associated with recency in the second.

For novel gastrointestinal biomarkers for HIV incidence, the gut inflammation markers analysed with random forest models did not predict recent infection. Plasma-based antibodies or cytokine measurements in plasma were retained, and random forest prediction models seemed promising. The MDRIs were ~200 days, with reasonable FRRs.

The next steps are to use the biomarker validation panel to reproduce MDRIs and FRRs for directly comparing discovery groups. There is also the potential need for a challenge panel including subtype C. The magnitude of improvement versus current tools that make new assay development worthwhile is being assessed, examining cost savings through MDRI, simplification (such as no viral load or ARV medicine testing) and good performance for subtypes D and CRF02_AG.

Incidence assay development projects funded by the United States National Institutes of Health were:

- novel single-well avidity enzyme immunoassays to determine the recency of HIV-1 infections (Sedia Biosciences);
- a rapid test for recent HIV infection (Immunetics); and
- developing and commercializing an innovative rapid HIV-1 incidence assay (Sedia Biosciences).

The discovery projects were all based on viral diversity:

- high-resolution melting assay based on increasing HIV genetic diversity over time since infection (Johns Hopkins University);
- viral biomarker based on inter-sequence Hamming distance via deep sequencing (University of Southern California);
- measuring changes over time in the variability of the HIV envelope gene sequence (Harvard University);
- adapting a low-cost assay based on DNA hybridization kinetics (AmpliCot) to measure HIV gene complexity to distinguish recent from chronic infection (Blood Systems Research Institute); and
- using mutations detected by clinical drug resistance assays to distinguish recent from chronic infection (United States National Cancer Institute and United States National Institutes of Health).

All of these markers are associated with recency, although their performance has yet to be determined based on MDRI and FRR, except for the high-resolution melting assay approach from Johns Hopkins University. They used tools from SACEMA to determine an MDRI (141 days, FRR 0%); three accessed the CEPHIA repository.

Other organizations (n = 7) have made requests to access the CEPHIA repository for other recency assay approaches, including using Bio-Plex 2200 for use as incidence assay; antigen/antibody (from Bio-Rad Laboratories), a multi-clade HIV-1 proteomic chip (Antigen Discovery) and Lactobacillus in early and late infection (University of California, San Francisco).
7. CONCLUSIONS AND NEXT STEPS

This Working Group meeting presented key updates to experts and stakeholders in assays for estimating HIV incidence and identified critical gaps in which further research, technical updates and recommendations are needed. Day 1 focused on critical issues, including variability among HIV subtypes and considerations for incorporating ARV medicines into a LAg + viral load recent infection testing algorithm. In addition, sample size requirements and considerations for national and key population surveys were presented. The results from the five incidence assays evaluated by CEPHIA were presented, concluding a multi-year evaluation funded by the Bill & Melinda Gates Foundation. Updates to the sample size calculation tool, improved specimen dating methods, an overview of a new external quality assurance programme for LAg, comparison of LAg performance between manufacturers (Sedia Biosciences and Maxim Biomedical) and a brief update on rapid recency test development were addressed. On day 2 of the meeting, target product profiles for tests for recent infection were presented to obtain consensus from meeting attendees. Findings from the updated market assessment and a landscape of new biomarker and assay development projects in the pipeline were also presented.

At the conclusion of the meeting, it was noted that some key gaps still remain. The CEPHIA evaluation of 10 available HIV incidence assays has demonstrated that no single assay alone satisfies the target product profiles, suggesting that a recent infection testing algorithm is necessary until improved tests become available. Further, challenges associated with variation in assay performance because of HIV subtypes and the recommended methods of addressing mixed-subtype populations to derive incidence estimates with current assays remains unsolved. Further research is needed to explore approaches to address these challenges. An updated method for infection dating was also presented, but standardized uptake of this new method needs to be addressed. It was suggested that a follow-on Working Group meeting would be helpful as a forum to present cutting-edge research on the best practices and approaches to account for these challenges. The Working Group also suggested that a technical update may be helpful to inform users about the appropriate use of assays for incidence estimation and to provide clear recommendations for the best recent infection testing algorithms to use.

Future work is needed to incorporated attendee feedback on the target product profiles and market assessment. These documents will be updated and published on the Working Group and FIND websites. Previous Working Group meeting reports will also be available on the Working Group website. The next Working Group meeting will be planned in conjunction with the Conference on Retroviruses and Opportunistic Infections on 13–16 February 2017 in Seattle, WA, USA to address these remaining challenges.
REFERENCES


20. HIV incidence assays sample size calculator: one survey [website]. Stellenbosch: South African Centre for


### ANNEX 1. MEETING PROGRAMME

**Saturday, 20 February 2016**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30–09:00</td>
<td>Breakfast</td>
<td></td>
</tr>
<tr>
<td>09:00–09:20</td>
<td>Welcome remarks</td>
<td>Txema Calleja, Peter Dailey</td>
</tr>
<tr>
<td>09:20–09:30</td>
<td>Introductions, objectives and expected outcome, review agenda and outstanding issues</td>
<td>Txema Calleja, Kim Marsh</td>
</tr>
<tr>
<td>09:30–09:45</td>
<td>HIV subtypes – MDRI best estimates (CEPHIA and CDC)</td>
<td>Eduard Grebe, Alex Welte</td>
</tr>
<tr>
<td>09:45–10:15</td>
<td>HIV subtypes – MDRI best estimates (Johns Hopkins University)</td>
<td>Oliver Laeyendecker</td>
</tr>
<tr>
<td>10:15–10:45</td>
<td>Time since infection specimen dating</td>
<td>Christopher Pilcher</td>
</tr>
<tr>
<td>10:45–11:15</td>
<td>Discussion – what are the differences and how to resolve them</td>
<td>Neil Parkin</td>
</tr>
<tr>
<td><strong>11:15–11:35</strong></td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>11:35–11:55</td>
<td>Optimizing the LAg + viral load recent infection testing algorithm – implications for sample sizes</td>
<td>Neil Parkin, Stefano Ongarello</td>
</tr>
<tr>
<td>11:55–12:15</td>
<td>Discussion</td>
<td>All</td>
</tr>
<tr>
<td><strong>12:15–13:30</strong></td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>13:30–14:00</td>
<td>ARV medicine testing component in recent infection testing algorithms</td>
<td>Thomas Rehle</td>
</tr>
<tr>
<td>14:00–14:15</td>
<td>Real-world experience with LAg</td>
<td>Bharat Parekh</td>
</tr>
<tr>
<td>14:15–14:30</td>
<td>External quality assurance programme for incidence assays</td>
<td>Mike Busch</td>
</tr>
<tr>
<td>14:30–14:45</td>
<td>Update on case-based surveillance meeting</td>
<td>Kim Marsh</td>
</tr>
<tr>
<td>14:45–15:30</td>
<td>New CEPHIA assay validations (Geenius™, ARCHITECT®, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3)</td>
<td>Gary Murphy</td>
</tr>
<tr>
<td><strong>15:30–16:00</strong></td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>16:00–16:30</td>
<td>Update on rapid assay development and potential applications and advantages</td>
<td>Bharat Parekh</td>
</tr>
<tr>
<td>16:30–17:00</td>
<td>Discussion</td>
<td>All led by Txema Calleja</td>
</tr>
<tr>
<td>17:00–17:15</td>
<td>Preview of day 2</td>
<td>Peter Dailey</td>
</tr>
</tbody>
</table>

**Alignment on critical assay characteristics**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30–09:45</td>
<td>HIV subtypes – MDRI best estimates (CEPHIA and CDC)</td>
<td>Eduard Grebe, Alex Welte</td>
</tr>
<tr>
<td>09:45–10:15</td>
<td>HIV subtypes – MDRI best estimates (Johns Hopkins University)</td>
<td>Oliver Laeyendecker</td>
</tr>
<tr>
<td>10:15–10:45</td>
<td>Time since infection specimen dating</td>
<td>Christopher Pilcher</td>
</tr>
<tr>
<td>10:45–11:15</td>
<td>Discussion – what are the differences and how to resolve them</td>
<td>Neil Parkin</td>
</tr>
<tr>
<td><strong>11:15–11:35</strong></td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>11:35–11:55</td>
<td>Optimizing the LAg + viral load recent infection testing algorithm – implications for sample sizes</td>
<td>Neil Parkin, Stefano Ongarello</td>
</tr>
<tr>
<td>11:55–12:15</td>
<td>Discussion</td>
<td>All</td>
</tr>
<tr>
<td><strong>12:15–13:30</strong></td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>13:30–14:00</td>
<td>ARV medicine testing component in recent infection testing algorithms</td>
<td>Thomas Rehle</td>
</tr>
<tr>
<td>14:00–14:15</td>
<td>Real-world experience with LAg</td>
<td>Bharat Parekh</td>
</tr>
<tr>
<td>14:15–14:30</td>
<td>External quality assurance programme for incidence assays</td>
<td>Mike Busch</td>
</tr>
<tr>
<td>14:30–14:45</td>
<td>Update on case-based surveillance meeting</td>
<td>Kim Marsh</td>
</tr>
<tr>
<td>14:45–15:30</td>
<td>New CEPHIA assay validations (Geenius™, ARCHITECT®, Glasgow-modified Genscreen™ HIV-1/2 Version 2 and IDE-V3)</td>
<td>Gary Murphy</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Presenter</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>07:30–08:00</td>
<td>Breakfast available</td>
<td></td>
</tr>
<tr>
<td>08:00–08:15</td>
<td>Welcome remarks, introduction, objectives and expected outcomes</td>
<td>Peter Ghys, Neil Parkin</td>
</tr>
<tr>
<td>08:15–09:00</td>
<td>Review agenda and summary from day 1</td>
<td>Peter Dailey, Neil Parkin</td>
</tr>
<tr>
<td>09:00–10:30</td>
<td>New target product profiles for HIV incidence assays</td>
<td>Jen Osborn</td>
</tr>
<tr>
<td><strong>10:30–11:00</strong></td>
<td><strong>Break</strong></td>
<td></td>
</tr>
<tr>
<td>11:00–12:00</td>
<td>Target product profiles: discussion</td>
<td>All, led by Peter Dailey</td>
</tr>
<tr>
<td><strong>12:00–13:15</strong></td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td>13:15–14:15</td>
<td>Updated market assessment</td>
<td>Charles Morrison</td>
</tr>
<tr>
<td>14:15–14:45</td>
<td>Updated market assessment: discussion</td>
<td>Led by Tim Mastro</td>
</tr>
<tr>
<td>14:45–16:00</td>
<td>Update on biomarker pipeline</td>
<td>Neil Parkin</td>
</tr>
</tbody>
</table>
ANNEX 2. LIST OF PARTICIPANTS

Adamma Aghaizu
Public Health England
London, United Kingdom

David Burns
Division of AIDS, National Institute of Allergy and Infectious Diseases, United States National Institutes of Health
Bethesda, MD, USA

Michael Busch
Blood Systems Research Institute
San Francisco, CA, USA

Kelly Curtis
Division of HIV/AIDS Prevention, CDC
Atlanta, GA, USA

Peter Dailey
FIND (consultant)
Geneva, Switzerland

Daniela DeAngelis
Medical Research Council
Cambridge, United Kingdom

Shelley Facente
University of California
San Francisco, USA

Jesus (Txema) Garcia Calleja
World Health Organization
Geneva, Switzerland

Peter Ghys
UNAIDS
Geneva, Switzerland

Eduard Grebe
Stellenbosch University, SACEMA
Stellenbosch, South Africa

Irene Hall
Division of HIV/AIDS Prevention, CDC
Atlanta, GA, USA

Wolfgang Hladik
CDC, Division of Global HIV/AIDS
Atlanta, GA, USA

Rick Homan
FHI 360
Durham, NC, USA

Sheila Keating
Blood Systems Research Institute
San Francisco, CA, USA

Oliver Layendecker
United States National Institute of Allergy and Infectious Diseases, Johns Hopkins University
Baltimore, MD, USA

Laurie Linley
Division of HIV/AIDS Prevention, CDC
Atlanta, GA, USA

Kim Marsh
UNAIDS
Geneva, Switzerland

Tim Mastro
FHI 360
Durham, NC, USA

Charles Morrison
FHI 360
Durham, NC, USA

Gary Murphy
Public Health England
London, United Kingdom

Stefano Ongarello
FIND
Geneva, Switzerland

Bharat Parekh
Division of HIV/AIDS Prevention, CDC
Atlanta, GA, USA

Neil Parkin
FIND (consultant)
Geneva, Switzerland

Christopher Pilcher
University of California
San Francisco, USA

Tom Quinn
Johns Hopkins University
Baltimore, MD, USA

Thomas Rehle
Human Sciences Research Council
Pretoria, South Africa

Renee Ridzon
FIND (consultant)
Geneva, Switzerland
Connie Sexton  
Division of Global HIV/AIDS, CDC  
Atlanta, GA, USA

Usha Sharma  
Division of AIDS, National Institute of Allergy and Infectious Diseases, United States National Institutes of Health  
Bethesda, MD, USA

Alex Welte  
Stellenbosch University, SACEMA  
Stellenbosch, South Africa

Ernest Yufenyuy  
Division of Global HIV/AIDS, CDC  
Atlanta, GA, USA

Day 2 only

Gavin Cloherty  
Abbott Diagnostics  
Des Plaines, IL, USA

Pat Garrett  
Immunetics  
Boston, MA, USA

Andy Levin  
Immunetics  
Boston, MA, USA

Ron Mink  
Sedia Biosciences  
Portland, OR, USA

Denise Naniche  
ISI Global  
Barcelona, Spain

Kelly Seaton  
Duke University  
Durham, ND, USA

Debra Hanson  
Division of HIV/AIDS Prevention, CDC  
Atlanta, GA, USA

Joyce Neal  
Division of HIV/AIDS Prevention, CDC  
Atlanta, GA, USA

Jennifer Osborn  
FIND (consultant)  
Geneva, Switzerland

Michelle Owen  
Division of HIV/AIDS Prevention, CDC  
Atlanta, GA, USA

Christine Rosseau  
Bill & Melinda Gates Foundation  
Seattle, WA, USA

Ruiguang (Rick) Song  
Division of HIV/AIDS Prevention, CDC  
Atlanta, GA, USA

Mitchell Warren  
AVAC  
New York, NY, USA