

## ANALYSIS AND USE OF HEALTH FACILITY DATA

# Exercise book

For HIV  
programme managers

WORKING DOCUMENT, MARCH 2019



# ANALYSIS AND USE OF HEALTH FACILITY DATA

## Exercise book for HIV programme managers

WORKING DOCUMENT, MARCH 2019

© World Health Organization 2019

All rights reserved. This is a working document and should not be quoted, reproduced, translated or adapted, in part or in whole, in any form or by any means.

# Contents

How this tool is organized .....	1
Exercise 1: Introduction to HIV dashboards .....	2
Exercise 2: Analysis and use of HIV disease burden data .....	6
Exercise 3: Assessing the HIV test-link-treat-retain-suppress cascade.....	10
Exercise 4: Using district ranking tables to assess performance across geographic areas .....	15
Exercise 5: Using maps to assess performance by geographic areas .....	18
Exercise 6: Using time trends to detect problems in implementation .....	21

# Acknowledgements

This exercise book has been developed by the World Health Organization, with the support of grants from Bloomberg Philanthropies Data for Health Initiative, Gavi, the Vaccine Alliance, The Global Fund to Fight AIDS, Tuberculosis and Malaria, and The Norwegian Agency for Development Cooperation.

# How this tool is organized

The exercises in this booklet are part of a package of materials used to develop skills in using routine facility data for by HIV programme managers at national and sub-national unit (SNU) level.

The package includes:

- a Facility Analysis Guidance Module<sup>1</sup> with a recommended set of standard dashboards used for programme management and improvement at national/SNU1 (e.g. regional), SNU2 (e.g. district), and F (facility) levels. The version of these HIV dashboards are referred to as 'HIV Burden 2.0' and 'HIV 2.0'<sup>2</sup>
- a training instance (or version) in DHIS2 demonstrating the use of the standard dashboards
- this Exercise Book reviewing how managers can interpret each figure included in the standard dashboards
- a supplement to the Exercise Book containing instructions for navigating and using the standard dashboards in DHIS2.

The Exercise Book is designed to demonstrate the use of each chart featured in the standard dashboards using data from a real world HIV programme to understand performance and improve services or conduct further investigation.<sup>3</sup> The module, dashboards and exercise book are anchored to existing normative guidance including the 2015 consolidated SI guidelines and 2017 person-centered monitoring guidelines. In addition, they are meant to complement a Cascade Analysis Manual which will be launched in 2018. These exercises are designed to be compatible with any HMIS used, however the supplemental instructions are tailored to use of the HIV app in DHIS2. The exercises feature charts from the dashboards designed for a national or SNU1 level programme manager, but in each exercise there are notes on how to modify the material to coincide with the SNU2 and facility level manager perspective.

## How to customize the training package with Exercises

The exercises in this document can be used modularly to fit a training suitable to the skills, data availability, training time, and technical capacity of different participant groups. A trainer may also prefer to use examples from participant's country's HMIS/DHIS2 instance and can replicate the equivalent charts using participant country data to replace the figures presented here.

## Orientation to the DHIS2 Training Dataset

The sample data used in these exercises come from a dataset constructed from real world data but reconfigured to represent a fictional country. This country's organizational hierarchy goes from national, regional (SNU1), district (SNU2), and facility level.<sup>4</sup> Although some of the names of facilities and geographic areas resemble those from actual places, because the data have been adjusted to create good training examples, the figures should not be assumed to represent the actual situation of places with the same names.

---

<sup>1</sup> Analysis and use of health facility data: Guidance for HIV programme managers, 2018 (in preparation)

<sup>2</sup> For easy reference between the guidance document, Exercise book, and Training instance the name of each figure in the dashboards starts with a geographic level and letter in addition to a short description of the content.

<sup>3</sup> With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities. It should also be noted that to maintain a more generic approach to referencing different sub-national levels, the training materials refer to SNU1 as 'regional' and SNU2 as 'district' even though these are not terms used in Kenya. And to make the training instance easier to use by multiple users simultaneously in a workshop setting, the number of regions and districts has been pared down.

<sup>4</sup> In this dataset, there is also a sub-district level, while this level is not used in the exercises for HIV dashboards, they are used more extensively in other programme areas such as immunization and RMNCAH.

## Exercise 1: Introduction to HIV dashboards

### Learning Objectives:

- Orient users to the general content of the recommended HIV dashboards (HIV 2.0 – N/SNU1 level, HIV 2.0 – SNU2 level, HIV 2.0 F level, HIV 2.0 Burden) and their relationship to each other;
- Demonstrate the differences between what national and SNU2 level managers need from a dashboard and the implications for how the charts are configured;
- Provide examples of how the layout of a chart can help managers make more effective use of HMIS data

The following table describes the recommended dashboards for HIV programme managers at different levels. The indicators used for these dashboards are aligned with the standard core indicators included in the Core 100 Health Indicators (from facility based data).<sup>5</sup> The main dashboard (HIV 2.0) has a version appropriate for managers at national/SNU1 (e.g. regional level), SNU2 (e.g. district) level, and those responsible at the facility level. These variants of the main dashboard use similar figures with modifications to make each more useful for managers at that level. The HIV Burden 2.0 dashboard displays the epidemiologically relevant data which provides important contextual data for tracking the performance of the response.<sup>6</sup>

### Types of HIV Dashboards

Dashboard Name	Primary User	Recommended Frequency of Use	# of Figures
HIV 2.0 N/SNU1	National and SNU1 (e.g. regional) programme managers	Quarter/Annual	10
HIV 2.0 SNU2	SNU2 (e.g. district) programme manager	Month to Month	8
HIV 2.0 F	Facility programme manager	Month to Month	6
HIV Burden 2.0	Programme managers at all levels; Surveillance officers	Annual	3

1. Figure 1 below shows the HIV 2.0 dashboard at national level. Note the use of various styles of visualisation used: Bar charts, tables, maps, and line graphs.

<sup>5</sup> Global Reference List of 100 Core Health Indicators, WHO 2015. Available at: <http://www.who.int/healthinfo/indicators/2015/en/>

<sup>6</sup> Exercise 2, goes into more detail in understanding and interpreting this dashboard.



Figure 1: The HIV 2.0 Dashboard – National/SNU1 level of the Data Use Training instance of DHIS2

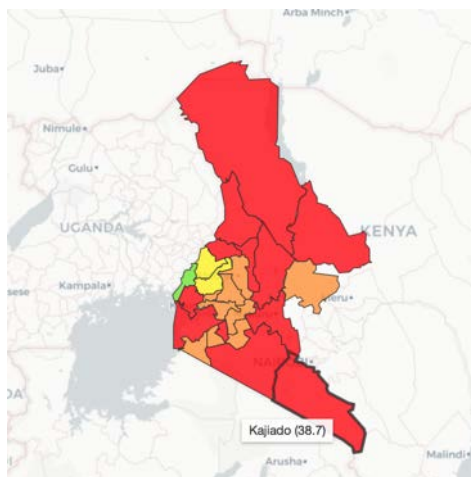


Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.

- a. List two ways that looking at core indicators through regional level maps can help **national** programme managers make decisions.

[Maps give national programme managers a summary view of performance by region that is intuitive to use. A map easily shows the performance level of different parts of the country in a specific service area, e.g. first, second, or third 90. Maps also allow managers to see if the same geographic patterns in performance across indicators or if performance is correlated to geographic characteristics (e.g. border areas, transport corridors, remote areas, etc.)]

- b. The same map can have different uses for managers at different levels. Look at the first map in the HIV 2.0 dashboard now generated for a user at a regional/SNU1 level, *ART Coverage Rate – last month (by district)*:

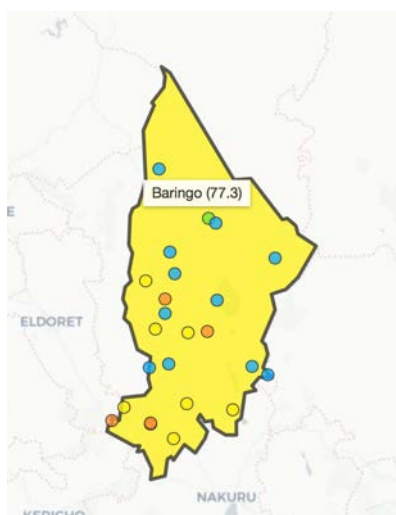


How can a **district** manager use this map to help her with programme management?

[District managers can compare their county's performance to other similar or nearby geographic areas to judge their performance in relative terms as well as against targets (shown through the choice of color-coding). If the district is performing poorly compared to other areas, seeing what other counties have been able to achieve established targets may help motivate under-performers to do better.]

- c. Think about the second map in the dashboard, ART Retention Rate – 12 months. How else could a map be changed to become more informative to a district level manager?

[The map can be switched to show sub-district level performance, e.g. ART facilities color-coded by retention rate, so district managers can pinpoint where in the district performance needs to be improved. In the example below, although Baringo district (in the Rift Valley Region) has a moderately high retention rate (77%), there are some facilities with lower levels of retention, shown by the orange dots. If ART facilities are categorized by type in the HMIS it is also possible to display different types of facilities, e.g. hospitals vs. health centers or distinguish sites by funder, etc. through the use of different shaped symbols. District level managers can then look for patterns to see if retention rates vary according to type of facility.]



*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

- d. The 2.0 dashboard includes trend lines showing the changes in core indicators from month to month.

- a. What kind of pattern does a programme manager look for in a time series chart to identify problems in programme implementation?

[A trend is a helpful way to detect sudden changes in reported performance. A stable system will show a stable trend. When a change in trend is detected, especially one that is marked by an unusual spike or drop, managers must investigate whether these sudden changes are a result of a data quality problem (e.g. transcription error, incomplete reporting, change in indicator definitions or reporting protocols); or from an actual change in service delivery. Dramatic changes due to a known programmatic event could be related to e.g. stock out, vacancy of key staff positions, or the launch of a major campaign to promote services. These types of programmatic events should be known about by managers and should be expected to be reflected in service trends. Managers should also be on the look-out for unexpected shifts in service delivery or more gradual increases in demand for services or declining levels of service utilization. The trend line should be compared to expected performance, i.e. targets to determine whether there are implementation problems. Trend lines that fluctuate a lot may also indicate an unstable programme.]

- b. Both of the figures with trends show several indicators (i.e. multiple lines in the same figure), what can a manager understand about programme performance by including both the number of newly diagnosed PLHIV and the number newly on ART in the same chart (See *Trend in newly diagnosed and newly on ART.*)

[Some indicators have a relationship or dependency with another indicator. By showing these indicators side by side, programme managers can see if performance is related. For example, in a “Treat All” ART programme, all PLHIV newly diagnosed should initiate ART as soon as possible. Assuming there are no data quality issues, a manager may expect to see the line showing newly on ART be at least as high as the line of newly diagnosed. For some period of time, the number newly on ART may exceed the number newly diagnosed if patients on ART wait lists are cleared and begin receiving treatment.]

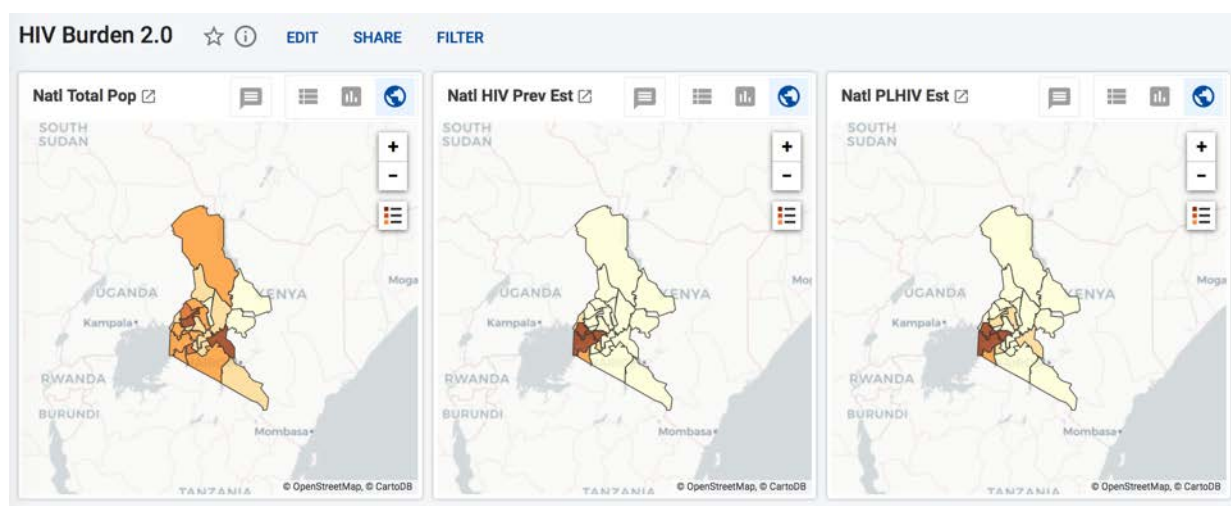
## Exercise 2: Analysis and use of HIV disease burden data

### Learning Objectives:

- Orient the user to the types of maps included in the HIV burden dashboard
- Describe how each map can be useful in contextualizing programme data in the HIV Main dashboards
- Understand the source of the data used in the HIV burden dashboard and how that can influence interpretation of the charts in both the HIV Main dashboard and the Burden dashboard
- Show how data from the HIV burden dashboards play a role in target setting for core indicators.

The HIV Burden dashboard presents three maps of districts showing a) the size of the general population, b) the estimated number of PLHIV, and c) the estimated HIV prevalence. The maps show the lowest geographic level for which there are estimates of HIV prevalence and PLHIV estimates. In this training DHIS2 instance, estimates are available down to SNU2/district level. The data presented in these dashboards are the inputs for the denominators of some core indicators and are the basis for calculating targets used for others. Denominator can be helpful when comparing performance between different geographic areas. Ideally, the data used for these three maps are uploaded to the HMIS where the routine facility based data are housed.

1. Review the maps included in the HIV Burden dashboard. Note the titles and the legends used for each map. (In the DHIS2 dashboard view, the longer names of the charts are not shown. The full names are: Natl Total Pop (Census, 2017); Natl HIV Prev Est (Spectrum, 2017), and Natl PLHIV Est (Spectrum, 2017)



*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

- a. What would be the data source for each of the three maps if this dashboard were generated with data from your country? When were each of these data sources last updated?

[General population data are usually based on the Census conducted by the National Statistics Office. In-between years when the Census is conducted, most countries employ a system of projections to account for population growth and movement. To estimate HIV prevalence and the number of PLHIV, countries often rely on modelled results such as those from SPECTRUM. Because the SPECTRUM package is updated annually, to incorporate additional parameters or modelling features that can improve the fit of the data to the model, previous years' estimates of PLHIV may change when the most recent estimates are generated. When the estimates data are imported into DHIS2 any previous years' estimates should be generated using the same version of SPECTRUM as used for the most recent estimates to enable appropriate comparisons (i.e. trends over time). Direct estimates of HIV prevalence from general population surveys may be another source, though these are likely used as a critical input for model-based estimates. Some countries use models other than SPECTRUM for their estimates, depending on the type of epidemic and type of input data available.]

- b. Describe the strengths and weaknesses of each data source (that would be used in your country) with respect to how reliable a source it is for the indicator each map represents, i.e. is the data source inclusive of the total population or are some populations less represented by these data sources? e.g. key populations, migrant populations, ethnic minority communities, institutionalized populations, etc.

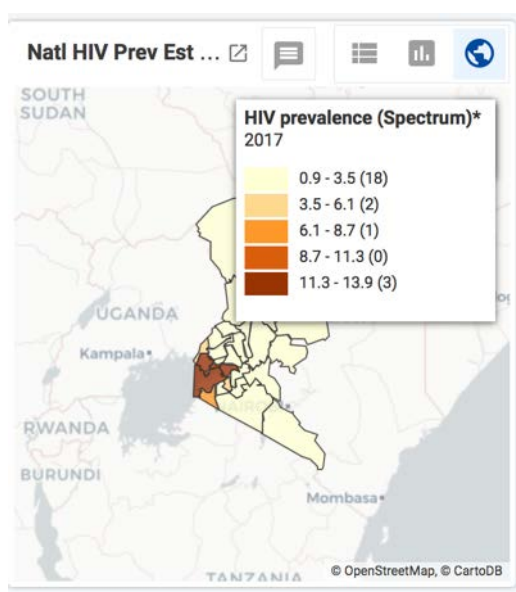
[Issues to consider include: how recent the data are (including the inputs used for modelled estimates); whether data are equally reliable and up to date in all parts of the country; whether the data source takes into account specific populations who may be disproportionately affected by HIV, e.g. key populations, prisoners, refugees, lower income individuals, etc. In some cases, data sources for important inputs may have been affected by one-time events, e.g. mass movement related to natural disasters, change in policy, conflict, economic crisis, etc. that should be noted. Data based on surveys or models come with uncertainty ranges that should be taken into account when interpreting programme performance that uses these data as denominators. Taking uncertainty ranges into account when interpreting programme performance makes sure that the differences observed between areas, groups or over time are statistically meaningful.]

2. What is the difference in interpretation of the map showing HIV prevalence and the map showing estimated PLHIV?

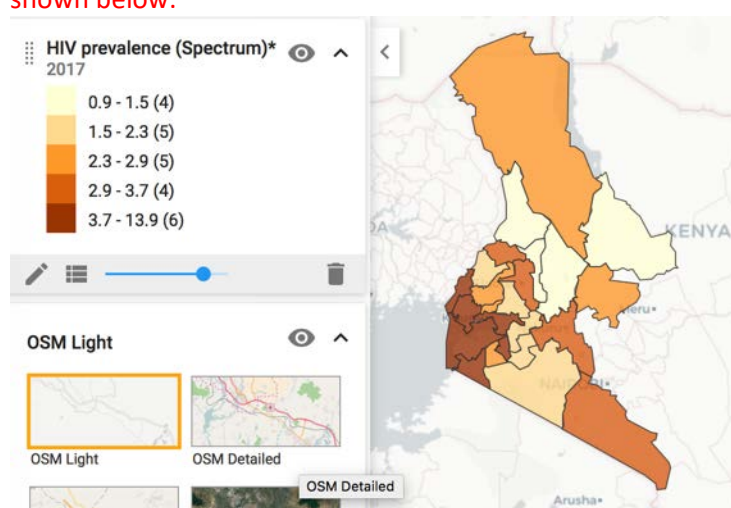
[The HIV prevalence is a percentage calculated as the number of PLHIV divided by the general population in the same geographic area. The estimated number of PLHIV is only the numerator, i.e. gives a number rather than a percentage. The prevalence of HIV can be high in a small population, but still have a small number of PLHIV compared to an area with a large general population but a moderate HIV prevalence. A high HIV prevalence area may be important because it indicates high rates of transmission have occurred or are occurring and prevention interventions are needed. A large absolute number of PLHIV must also be given attention by managers in terms of ensuring adequate services and resources are available to provide people with testing and treatment.]

3. These maps are designed to color code the districts into five 'equal intervals.' This means, the range of values across all districts is calculated and divided into five categories. For example, the HIV prevalence map has a range of 0.9 to 13.0% prevalence. Each color category represents a spread of 2.6 percentage points. The legend also shows the number of districts that fall into each category. In this example, the vast majority (18 out of 24) fall into the lowest prevalence category. How would the map be interpreted differently if the districts were divided into quintiles instead of using this 'equal intervals' method?





[When quintiles are used to color code maps, this means that there are an equal number of districts in each category, regardless of the range. The same map color coded by quintiles is shown below:



*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

When using this method a manager can see the five highest prevalence districts easily, however, the range in prevalence in the highest category is very wide, it ranges from 3.7 to 13.9% prevalence. It may not make sense to group a district with 3.7% prevalence with a district that has a prevalence over 10%. Similarly, in the lower prevalence categories, it may not be programmatically or epidemiologically meaningful to distinguish a district that has a prevalence of 2.3 and 3.7%. Both relative ranking and absolute values are important for managers to consider when prioritizing areas for more managerial attention or resources.]

4. What is the relationship between data displayed in the HIV burden dashboard and target setting for core indicators?

[Because the source of data on the HIV burden dashboard is updated at most annually, it is not a dashboard that will change weekly or monthly. The HIV burden dashboard is most useful for giving programme managers the context for reviewing programme performance. e.g. ensuring strong

programme performance is most important in areas where the HIV burden is greatest. At a more quantitative level, the PLHIV estimates give the size of the population requiring services, i.e. it sets the upper bound for target setting. Some core indicators reference estimates of PLHIV explicitly, for example, the ART coverage rate indicator uses the number of PLHIV from the HIV burden dashboard as a denominator, e.g. 75% ART coverage is calculated as 75% of the estimated number of PLHIV in the country. A similar calculation is made for targets at sub-national level, e.g. 75% ART coverage at the SNU2 level is calculated as 75% of the estimated number of PLHIV in the district. Many countries also use HIV prevalence estimates as a benchmark for determining the efficiency of HIV testing programmes. (Although testing yield targets must take into account history of testing and the proportion of PLHIV who are already diagnosed.)

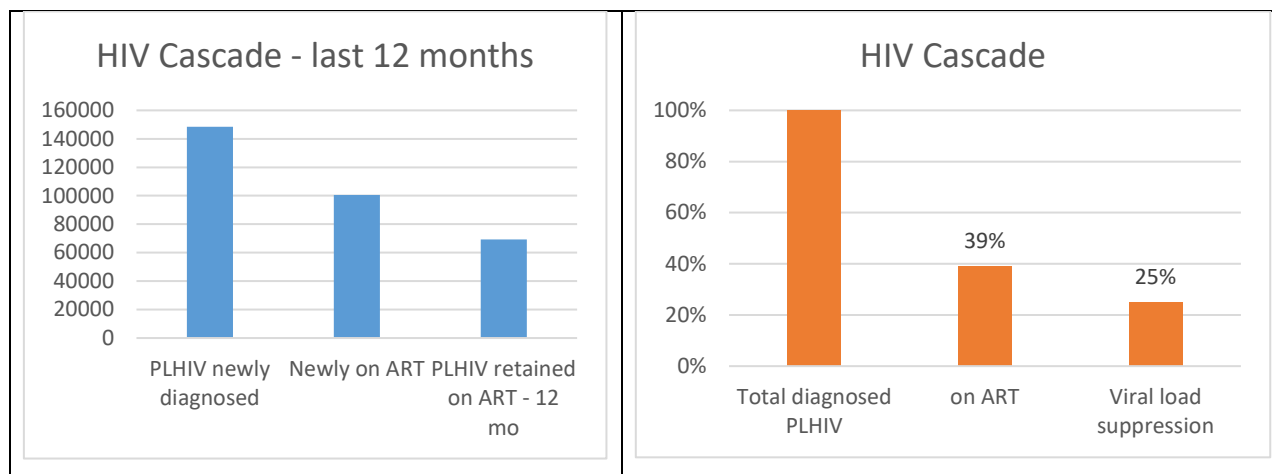
Because of the importance of estimates to target setting and measuring programme performance, calculations of targets and coverage based on the PLHIV estimates should take the uncertainty associated with these estimates into consideration. Uncertainty for PLHIV estimates can be wide depending on the availability and quality of data used as inputs for the models used to generate the estimates. The more data points over time from a wide range of geographic areas in the country, the better the fit of the model and the smaller the uncertainty of estimates at the national level. Estimates made at the regional or district level may have even greater degrees of uncertainty due to the relatively fewer data points likely to be available to fit the models used.]

## Exercise 3: Assessing the HIV test-link-treat-retain-suppress cascade

### Learning Objectives:

- Orient the user to the charts used to represent the HIV cascade in the HIV 2.0 dashboards using cross sectional data from facilities
- Become aware of typical patterns observed from cascades in the HIV 2.0 dashboard
- Demonstrate how showing to identify service challenges and inequity (e.g. by sex) through disaggregated cascades
- Use facility characteristics as a variable for disaggregation to describe the performance of different service delivery models
- Understand how the interpretation of VL suppression rates is influenced by VL testing coverage data.

1. The HIV 2.0 dashboard starts with an HIV cascade chart that features data readily available from routine HMIS systems: # of newly diagnosed, # PLHIV newly on ART, and # of PLHIV retained on ART – 12 months, over the last 12 months. The HIV cascade presents the fundamental analysis needed to assess progress in diagnosis (first 90), starting treatment (second 90) and achieving viral suppression (third 90).



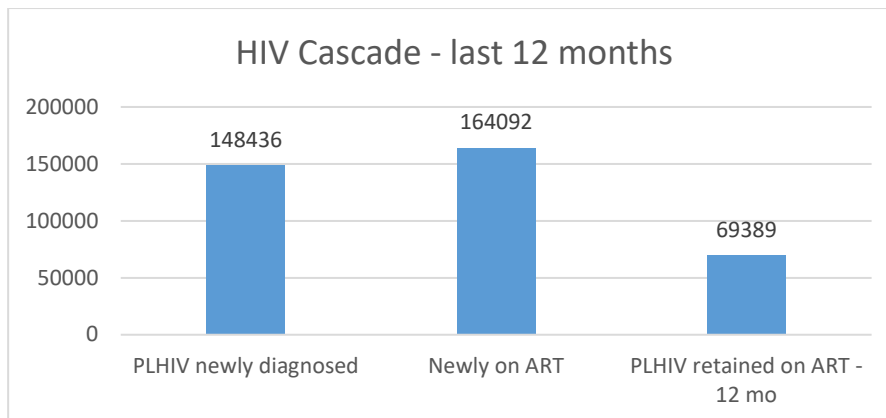
- a. The chart (above) on the left is the chart included in the HIV 2.0 dashboard using facility based data. The chart on the right shows a different way to present the cascade which may be familiar to programme managers. What are the differences in data sources and in interpretation of these two formats assuming they reflect data from the same country and time period:

[One fundamental difference is that the cascade on the left gives absolute numbers, where as the cascade on the right uses percentages. More specifically, the cascade on the left only gives the numbers of people receiving the service during a 12-month period, likely corresponding with a standard financial and implementation period. This means that the bars in the left-hand chart include different, partially overlapping groups of PLHIV. While the chart on the right represents the data for all bars as a percentage of all PLHIV EVER diagnosed and still alive. The cascade on the



right shows the cascade in an intuitive way, but these figures cannot be obtained from routine facility HMIS, especially if reported in an aggregate form. Often times they are derived from national population-based surveys or from a variety of different data sources.]

- b. In a typical health service cascade, each bar is a subset of the people captured in the previous bar of the chart. The decreasing height of each subsequent bar illustrates the drop-off in service utilization between steps in the cascade. However, in some countries the HIV cascade chart as rendered in the HIV 2.0 dashboard may show the # of newly on ART is larger than the # of newly diagnosed, as illustrated below. Describe some programmatic situations that could explain this unexpected result.



[Despite near-universal uptake of 2015 WHO normative guidance recommending universal test and start and same-day ART initiation, implementation may remain incomplete in some countries and there may be a wait list of PLHIV who have been diagnosed but did not meet clinical criteria for ART eligibility or where there were not sufficient resources to enroll them in ART. When policies change or resources become more available, countries may make a big push to increase ART enrollment, especially for those who have been on treatment waitlists. If this happens, but HIV testing and case finding is unchanged, it is possible for the bar showing the number new on ART may exceed those who are newly diagnosed during the same reporting period. Currently, this is a more common finding at sub-national levels, where testing and treatment service capacity may not be perfectly aligned requiring newly diagnosed PLHIV from one SNU to seek treatment in another SNU.]

In some SNUs this mismatch between where PLHIV are diagnosed and where they initiated ART may occur regularly due to patterns of migration or perceptions where testing or treatment services are provided with greater quality or confidentiality. Detecting and understanding the reasons for these recurrent patterns will be critical for managers to correctly interpret the cascades for their service sites.]

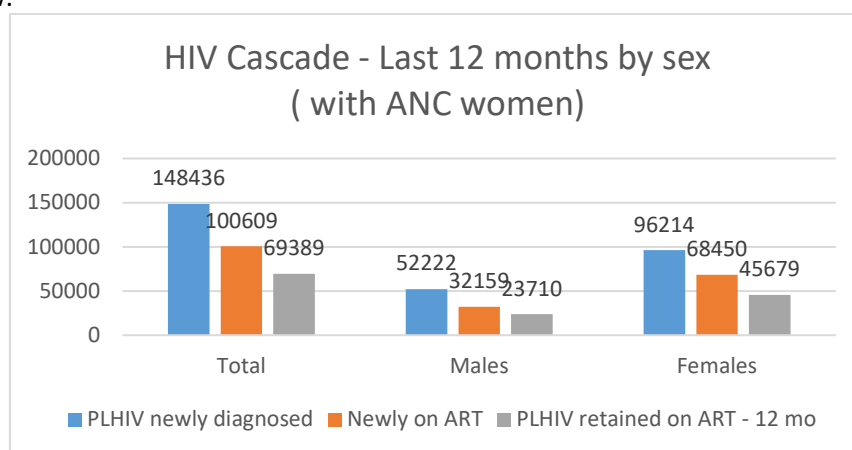
- c. To be consistent with the “three 90s” it would be preferable for the HIV cascade in the HIV 2.0 dashboard to feature VL suppression rather than just ART retention at 12 months. Describe why showing the “number of patients with viral load tests <1000 copies / mL in the last 12 months” would be problematic as the third bar in the chart labeled ‘HIV cascade – last 12 months,’ shown above.

[The indicator calculated from an aggregate reporting dataset would include any patient receiving a VL test during the time period, regardless of the reason for testing, i.e. routine or targeted based on signs or symptoms. Over a 12-month period, the expected number of patients with a VL tests should be close to the total number of patients in care or ART, because most countries’ treatment guidelines recommend viral load tests annually for PLHIV enrolled in care or treatment. In

contrast, the cascade figure shown above limits the number of patients for each bar to the people diagnosed in one year or retained on ART for a period of one year. A more comparable indicator would be number of patients with a VL test <1000 copies/mL at 12 months on ART. (By definition this group of ART patients would be those who initiated ART in the year prior to the reporting period included in the chart and not a match for those who got diagnosed during the reporting period.) However, in an aggregate reporting system, it is unusual to capture viral load test data so that it can be disaggregated by the duration of treatment.

In addition to the challenges of including a comparable patient population in each bar of the cascade, “the number of patients with viral load tests <1000 copies/mL” by itself is difficult to interpret without a denominator, i.e. the number of viral load tests conducted. In a cascade based on a cohort of patients for which each bar is a clear subset of the previous bar the interpretation is clear. However, the types of cascades that can be produced with cross-sectional aggregate data cannot make that assumption about the denominator.]

2. Where possible, the HIV 2.0 dashboard should show the HIV cascade disaggregated by sex, as shown below.



- a. What is it useful for programme managers to compare cascade data among males and females?

[Comparison of cascade performance between groups and geographic-administrative units is critical to identifying gaps in service access and quality and/or identifying factors associated with strong or weak performance. Cascade analysis is a basic assessment of programme performance that can be applied to any sub-group of patients (by sex, geographic area, facility type, etc.), illuminating gaps where rates of new infections and HIV-associated mortality may be higher. Sex disaggregation may be particularly important in countries where a large portion of women PLHIV are diagnosed or enrolled in ART via a PMTCT programme. These types of programmes including linkage to treatment, may operate differently than for people seeking HIV testing in a general population testing site. In this example, the crude ratio of linkage to treatment for males is 62% (32159/52222) compared to females at 71% (68450/96214) Assessing gender inequity in service delivery is an important criteria for rating the quality of services.]

- b. Note that in this country example, there are many more female PLHIV being diagnosed compared to men. Does that mean that there is a higher burden of HIV among women than among men? Why or why not?

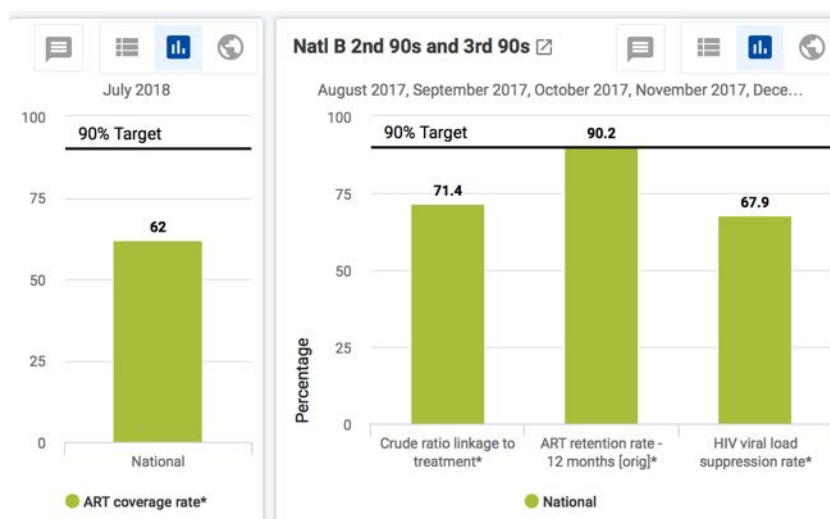
[The number of PLHIV diagnosed reflects both the burden of disease and the way testing programmes target and promote testing among different populations. In many countries, pregnant women are routinely screened for HIV. Women diagnosed in this manner comprise a

relatively large portion of new cases found in areas where HIV testing coverage is low. To determine whether the burden of HIV is higher among women compared to men, the test positivity, i.e. # of PLHIV diagnosed / # of people tested, should be compared among men and women. If test positivity among men is higher, it is worthwhile to characterize how men are targeted for testing and whether the group tested is a higher risk group and if more testing should be targeted among male populations.]

c. What facility type disaggregation variables are available in your country's HMIS?

[These might include: programme area (TB, ANC, etc.); level of facility (hospital, health center, dispensary, etc.); funder (government, Global Fund, PEPFAR, etc.) ownership (public/private/NGO, etc.)]

3. The second and third charts in the dashboard show calculations that are derived by cascade analysis, representing fall off from one service to another in terms of percentages. These percentages correspond to the "2<sup>nd</sup> 90s," i.e. ART coverage rate, the Ratio of newly on ART to newly diagnosed and the % ART retention for 12 months; and the "3<sup>rd</sup> 90," i.e. % VL suppression rate."



- a. What does the ART coverage rate indicator tell you about the treatment program that is different than the Crude Ratio Linkage to Treatment? Is it possible to have a good performance on Linkage to Treatment, but poor performance on ART coverage? And vice versa (is it possible to have good performance on ART coverage but poor performance on Linkage to Treatment)?

[ART coverage rates are calculated using the denominator of the total estimated number of PLHIV who may or may not know their status. The Crude Ratio uses the number of newly on ART as the numerator and the number newly diagnosed PLHIV as a denominator. The ART coverage rate and the Crude Ratio measure different aspects of programme quality. The ART coverage looks at how the programme performs overall against the epidemic need and 2<sup>nd</sup> 90 goals and targets, while the Crude Ratio of Linkage to Treatment looks at the more specific aspect of making sure PLHIV who are diagnosed are linked to treatment, and meet treatment initiation targets towards 2<sup>nd</sup> 90 goals.]

- b. This chart shows both the ART retention rate – 12 months and HIV viral load suppression rate – last 12 months. What are the advantages of using one measure over the other to determine the effectiveness of the ART programme?

[The primary outcome of a successful ART programme is determined by whether PLHIV achieve viral suppression through effective use of ART (based on retention, adherence, and prevention of HIV drug resistance). This is why the ideal measure of performance is the viral load suppression rate, rather than simply the ART retention rate – 12 months. The viral load suppression rate uses the number of VL tests conducted during the reporting period as its denominator. In countries where VL testing is not available to all PLHIV on ART, there may be some bias in the rate of suppression measured, i.e. the result is not generalizable to all ART patients' experience.

When VL suppression rates are not generalizable the ART retention rate – 12 months may give a more reliable measure of ART programme performance. For this reason, knowing the percentage of PLHIV on ART who have had a viral load test gives the programme manager a sense of the potential bias in the VL suppression measure. This is especially important at lower SNU level because of the greater disparity of where VL testing is available. This information is available in the time trend of retention and VL suppression shown in the last part of the dashboard. Interpreting these charts together will be addressed in Exercise 6. In this example, the ART retention rate appears much higher than the viral load suppression rate. This may reflect that adherence and retention among patients in the first 12 months of treatment may be better than patients who have been on ART for many years. But without looking at VL testing coverage in more detail it is not possible to determine which is more appropriate to use.]

4. If the '2<sup>nd</sup> 90s and 3<sup>rd</sup> 90s' chart were presented for a shorter period of time, e.g. last 3 months, how would this change the interpretation of each chart?

[In general, the last 3 months timeframe gives managers a more recent view of programme performance compared to last 12 months. This can be helpful if the programme has recently implemented changes in how ART retention and treatment adherence are promoted, e.g. via active monitoring and tracing of PLHIV lost to follow-up. One disadvantage of using a shorter time frame to review performance is that the number of patients included in the denominator is smaller and may show greater fluctuation. There may also be a problem if viral load testing availability is uneven through the year because the shorter time frame measure may not be representative of the whole 12-month period.]

## Exercise 4: Using district ranking tables to assess performance across geographic areas

### Learning Objectives:

- Understand the importance in reviewing all four indicators: testing volume, numbers diagnosed, test positivity, and crude ratio linkage to treatment to assess local performance of a testing programme.
- Show how tables can be used to identify high and low performing geographic areas with respect to testing and ART programmes.
- Distinguish between indicators that can be used to compare district programme performance directly and those where other contextual factors are needed.

There are two tables included in the HIV Main Dashboard: Testing Performance and ART Performance by SNU. Each table shows the value of multiple indicator side by side for each district. The tables can be sorted alphabetically and/or from high to low for each indicator.

1. The table showing Testing Performance by district includes four indicators: number tested, number positive tests, the test positivity, and the crude ratio linkage to treatment. (The examples shown in this exercise will assume the perspective of an SNU1/regional manager reviewing performance at the SNU2/district level. However, a similar analysis could be done for a national level manager reviewing performance of the regional level.)
  - a. What are the advantages of sorting the table, i.e. ranking districts, according to the number of tests performed vs. the number of new positive tests?

[Districts ranked by tests performed indicates where testing programmes are most active. Districts ranked by the largest number of new diagnoses tells managers where case finding has been successful and districts which have the largest number of PLHIV who need to be linked to treatment. (However the latter is better interpreted by looking at the # of new diagnoses as a percentage of # tested, i.e. test positivity). This kind of information can be helpful for forecasting the need for consumables and staffing.]

- b. How do the columns showing test positivity and crude ratio linkage to treatment help managers determine whether the HIV testing programme in a district is performing as expected?

[Test positivity gives programme managers a sense of the efficiency of the district testing programme. Some districts may test a large number but find very few cases. This may indicate that the right groups are not being targeted by the testing programme. Programmes that have a higher test positivity are more efficient, however, direct comparisons of test positivity across districts must take into consideration what proportion of PLHIV have been diagnosed already and the underlying HIV prevalence in the area. The crude ratio linkage to treatment in the a table showing testing programme performance reinforces the responsibility for HIV testing programmes

to ensure those newly diagnosed are successfully linked to treatment. A programme strong on case finding, but weak in linkage to treatment is not a successful testing programme. Ultimately testing programme performance assessment must be based on a composite profile of all four indicators, with appropriate socio-demographic disaggregations.]

The table(s) showing ART Performance - by district includes seven indicators: # newly on ART, Crude ratio linkage to treatment, # currently on ART, ART retention at 12 months, VL suppression rate, estimated PLHIV, ART coverage rate, and estimated PLHIV.<sup>7</sup>

July 2017, August 2017, September 2017, October 2017, November 2017, December 2017, January 2018, February 2018, March 2018, April 2018, May 2018, June 2018					
Organisation unit / Data	Newly on ART*	Crude ratio linkage to treatment*	Currently on ART (LV)	ART retention rate - 12 months [orig]*	HIV viral load suppression rate*
Homa Bay	13 597	64.4	93 474	90.2	69.8
Kisii	4 468	68.5	26 481	89.2	71.5
Kisumu	11 828	59.7	92 081	83.2	66.4
Migori	9 061	67.4	61 646	88.3	79.7
Nyamira	2 372	70.4	12 614	94.7	80.2
Siaya	13 554	94.1	73 655	85.5	62.3

*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

- Assuming the quality of the data is the same, which of these indicators can be used to directly compare ART performance of different districts?

[Denominators are generally used to standardize indicators so that comparisons between areas or groups can be made. Crude ratio linkage to treatment, the ART coverage rate, and ART retention rate – 12 months and VL suppression rate can be compared across districts. The appropriate benchmarks are set at absolute values rather than being relative to how other areas or groups performed. Newly on ART, currently on ART, and estimated PLHIV are absolute numbers and are not standardized in a way that can be used to judge programme performance. These indicators are included in the table to provide context, e.g. where is the ART programme expanding rapidly, which areas have the highest burden, etc.]

- In this table, note that the cells for the last 3 indicators appear in different colors: Red, orange, yellow, green, and blue. What is the relationship between the color and the numerical value of the data?

[The color coding indicates the level of performance against standard benchmarks. Cells in red indicate poor performance, orange and yellow show moderate to good performance with green or blue fonts show districts which are excelling. The standard categories used for these indicators are Red= <60, Orange = 60-75%, Yellow=75-90%, Green= >90%, Blue = >100%. As discussed in earlier exercises, for some indicators the coverage % exceeds 100% due to mismatches in the data sources for the numerator and the denominator.]

<sup>7</sup> In the DHIS2 interface, all the indicators shown in a table must be aggregated by the same time period. For this reason, ART coverage and Estimated PLHIV are shown in a separate table because these indicators are more useful when calculated on a shorter time period.

- c. What indicator would you sort the table by in order to compare and assess whether districts with large ART programmes are more likely to have higher retention rates than districts with smaller ART programmes? Using the sample data, test your method and describe your conclusions.

[Sort by Current number on ART to see if there are more red cells in the ART retention rate column at the top compared to the bottom of the list. The result will look like this:

July 2017, August 2017, September 2017, October 2017, November 2017, December 2017, January 2018, February 2018, March 2018, April 2018, May 2018, June 2018					
Organisation unit / Data	Currently on ART (LV) ‡	Newly on ART* ‡	Crude ratio linkage to treatment* ‡	ART retention rate - 12 months [orig]* ‡	HIV viral load suppression rate* ‡
Homa Bay	93 474	13 597	64.4	90.2	69.8
Kisumu	92 081	11 828	59.7	83.2	66.4
Siaya	73 655	13 554	94.1	85.5	62.3
Migori	61 646	9 061	67.4	88.3	79.7
Kisii	26 481	4 468	68.5	89.2	71.5
Nyamira	12 614	2 372	70.4	94.7	80.2

*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

This quick analysis shows that ART retention rates are generally good, but some of the smaller programmes have higher levels of ART retention at 12 months.

A similar analysis should be done for assessing whether geographic areas with large need (i.e. large estimated PLHIV) are performing well.]

- d. To understand whether a district is performing well on # newly on ART or # currently on ART, what additional information is necessary to have?

[Specific targets for newly on ART or currently on ART, or performance thresholds for 12-month retention, will help compare district performance to district expectations. This is important because not all districts receive the same level of resources. For these reasons incorporating district level targets in this type of ART performance table would be helpful. However, because targets are based on estimates that have some uncertainty, targets should similarly be presented as having some level of uncertainty themselves.]

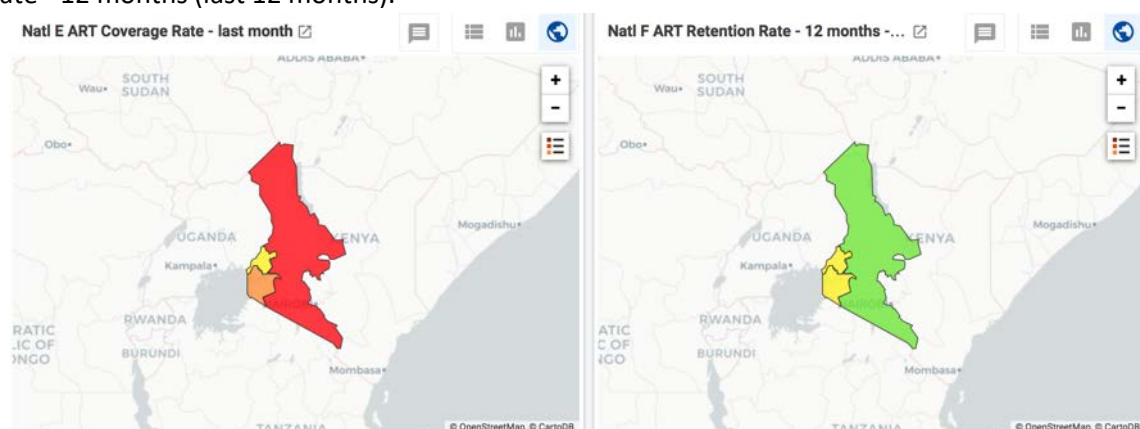


## Exercise 5: Using maps to assess performance by geographic areas

### Learning Objectives:

- Demonstrate how maps can be used to identify high and low performing geographic areas with respect to testing, treatment coverage, and ART retention
- Become familiar with the different sources of denominator data used in specific indicators and the implications for interpreting programme performance
- Explore how analyzing data over different time frames can change the interpretation with respect to programme performance

1. There are two maps included in the HIV main dashboard: ART coverage rate, and ART retention rate - 12 months (last 12 months).



*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

- a. Both of these indicators have denominators. Write out the denominator for each map's indicator, what is the data source of the denominator?

[The denominator for ART coverage rate is the estimated number of PLHIV (usually from SPECTRUM). The denominator for ART retention rate – 12 months is the number of PLHIV who started on ART 12-24 months prior to the date the chart is generated.]

- b. An advantage of using routine facility based data is that, assuming the data are complete, there is no statistical error or uncertainty associated with the measure. For each of the denominators used in the maps included in the dashboard, which is based on routine facility based data and which is based on survey samples or modeled estimates. What level of uncertainty is associated with each of these data sources (in your country)?

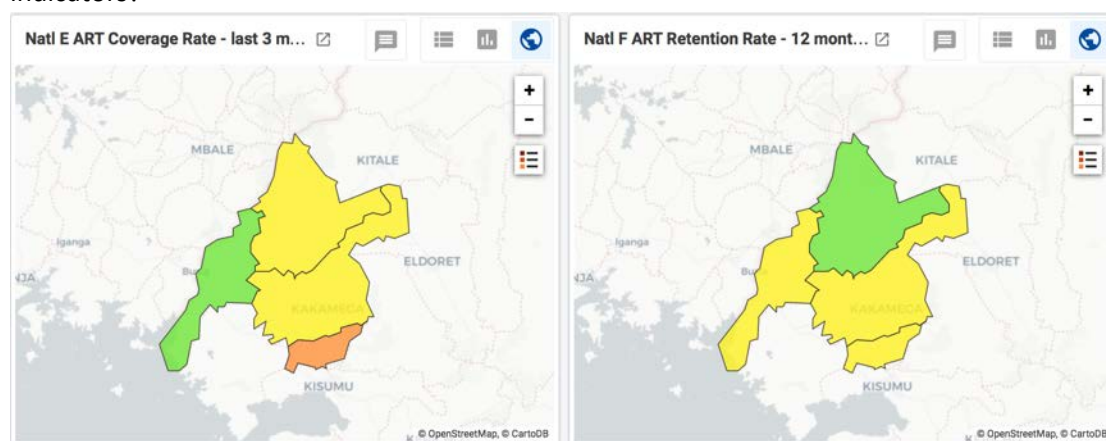
[In the case of the training instance, the estimated number of PLHIV is a model based estimate that has an uncertainty range associated with it. The number of PLHIV started on ART 12-24 months prior to when the chart was generated comes from routine monitoring data and has no uncertainty bound associated with it.]



- c. The purpose of using denominators is to “standardize” the indicator when comparing geographic areas. How would the interpretation of each map be different if only the numerator were used instead (i.e. # currently on ART, # retained on ART for 12 months)

[Without the denominator it would be difficult to compare one geographic area to another in terms of judging the expected level of performance. A geographic area with a large patient load may be doing poorly with respect to coverage rate or retention rate, but the absolute number currently on ART or retained for 12 months may appear larger than an area with a small PLHIV burden that has an excellent quality ART programme. Similarly, the denominator allows for determination of the absolute and percentage treatment gaps which can be vital to informing planning priorities.]

2. The map below, shows the dashboard maps for a SNU1 manager from the Western Region. Note that the same color-code category are applied for both maps, i.e. yellow = 75-90%. Which districts are the lowest performing for each map? Are the underperforming districts the same for both indicators?



*Note: With the permission of NASCOP, the DHIS2 training instance used in this package features data from the Kenya that have been modified to provide useful training examples from a real-world context. The figures generated from the training instance should not be interpreted as actual Kenya data or attributed to specific sub-national areas or facilities.*

[In this example, performance against both indicators is generally good. However, there does not appear to be a strong correlation with respect to a district having excellent, >90%, achievement in ART retention – 12 months and excellent ART coverage. There are several possible explanations. On the one hand, ART coverage is a cumulative indicator since the start of the treatment program, whereas 12-month retention is an indicator of fairly recent performance. Programme efforts to increase ART coverage may function separately from efforts to retain patients on ART. If a country relies heavily on a community based model to support adherence and retain ART patients on treatment, changes in resources or management issues for community based services may be separated from resources that flow through public sector health facilities that determine the number of ART patients that are diagnosed and enrolled in treatment. Also, as discussed in the earlier part of the exercise, the categorization of districts by ART coverage comes with some degree of uncertainty due to the range associated with the denominator, estimated number of PLHIV in the district. Coverage data must be considered in the context of levels of resource investment and programmatic effort, and monitored as modeled estimates are improved by new data sources.]

3. The ART retention rate – 12 months indicator looks at ART retention over the first 12 months of treatment. This indicator can be calculated over different periods of time, e.g. one month, 3 months, 12 months, etc. The ART retention rate – 12 months calculated for 3 months looks at how many ART patients enrolled over a 3 month period were still on ART 12 months after initiating treatment.

List the advantages and disadvantages for using a 12 month time frame as opposed to a 3 month or last month period for the ART retention – 12 months map.

[Advantages: The 12 months period may be more stable and less subject to seasonal fluctuation or a single time period when there was an interruption in service.

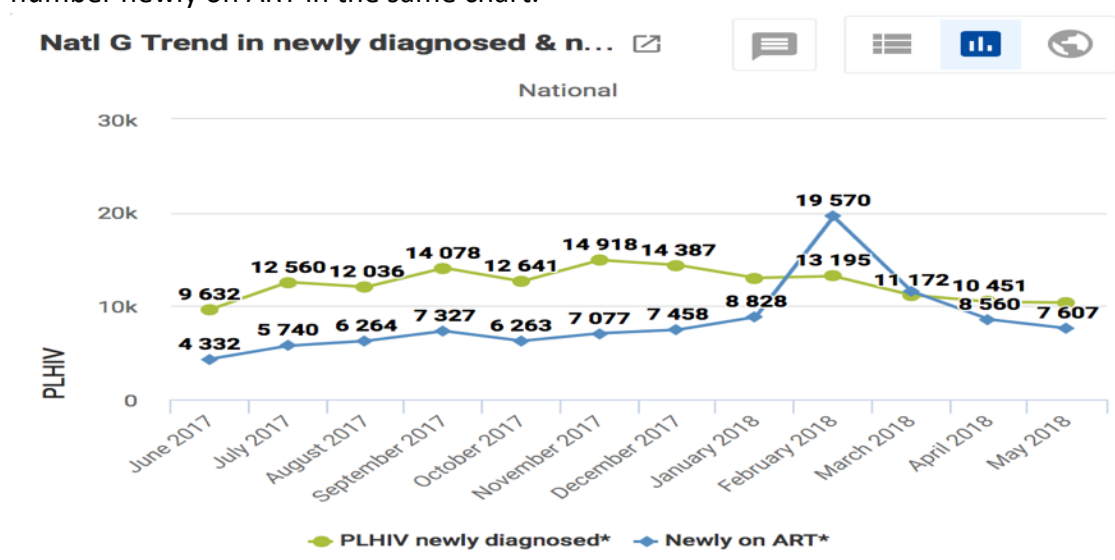
Disadvantages: If the programme has undergone recent changes, the longer timeframe may mask the effect and make it more difficult to correlate changes to improvement in retention.]

## Exercise 6: Using time trends to detect problems in implementation

### Learning Objectives:

- Demonstrate how time trends can be used to identify problems in programme performance
- Understand instances when including two trend lines in the same chart can help in the assessment of programme performance
- Show examples of how understanding programme guidelines or policy can be critical for interpreting time trends

1. The first time trend in the Main dashboard shows the number newly diagnosed and the number newly on ART in the same chart.



- a. What is the expected relationship between these two trends?

[In a “treat all” programme, any diagnosed PLHIV should be enrolled in ART soon after confirmed positive. For this reason, the trend line of newly diagnosed and newly on ART should be very similar.]

- b. How does the expected relationship change in a time period when a “treat all” policy has not fully been implemented?

[When “treat all” is not implemented fully, there may be a backlog of PLHIV who have been diagnosed but are on a wait list for treatment enrollment. Depending on how resources are allocated, clearing the backlog of previously diagnosed PLHIV may be uneven and not match the trend in newly diagnosed.]

- c. How would you interpret and further investigate the reason for the spike in those newly on ART during February 2018?

[The increase is large (~ 75% increase) so this could only be explained by a large-scale change in the programme during this month and it might also be expected for this change to be sustained for more than a single month. This type of programmatic change should be well known to managers. Another explanation to be investigated is a data quality issue, e.g. an extreme outlier/data error. This would require looking to see if the spike is observed across a large number of geographic areas, or just one or two facilities/areas. For a programmatic change, the spike would be observed in geographic areas affected/involved. A data error is more likely to be random and isolated to a specific site.]

- d. What additional information about the treatment guidelines or treatment programme is useful to have in order to interpret this figure to determine whether performance was as expected?

[Critical information includes: treatment eligibility criteria, the resource availability for ART, and the targets for ART enrollment in each area by implementation time period.]

- e. What additional programme information is necessary for interpretation if this time trend chart were generated for a specific facility?

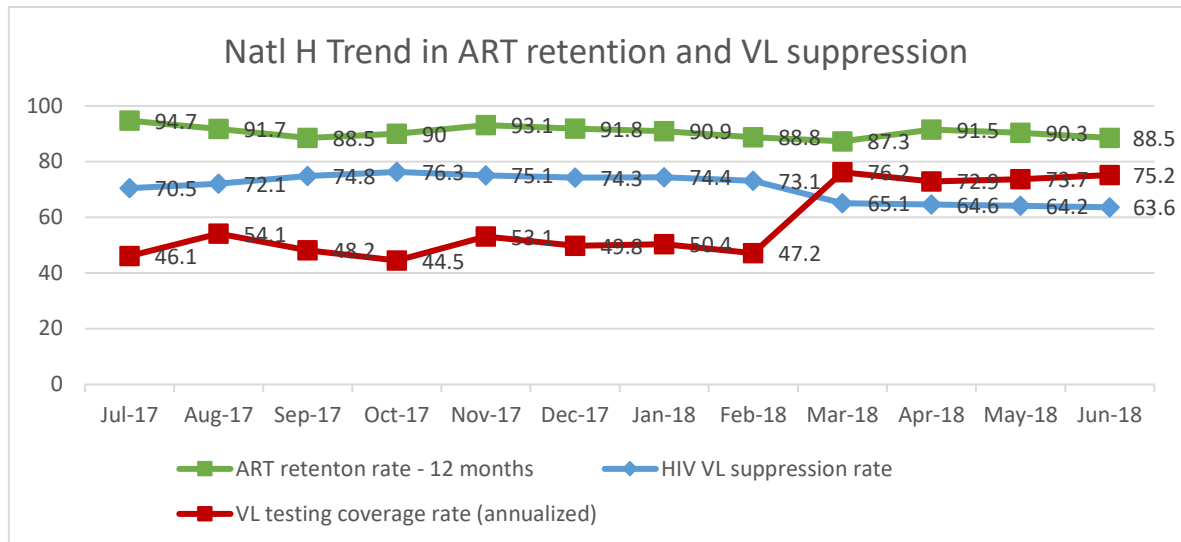
[The service availability and patient flow from testing to treatment at this facility need to be understood. For example, what proportion of new ART enrollees are diagnosed at the same facility vs. are referred from other testing sites. If the facility is funded by different donors are there differences in how linkage to treatment happens depending on different patient populations and how has this changed over different funding periods.]

2. The second time trend shows the ART retention rate at 12 months (green) and the viral load suppression rate (blue). A third (red) line shows the % of ART patients who received a viral load test (annualized).<sup>8</sup>

---

<sup>8</sup> The VL testing coverage must be annualized because in each month only a portion of ART patients are expected to get a VL test. If VL testing is recommended once a year for all patient on ART, roughly 1/12 of all ART patients should get a VL test each month. In this annualized monthly indicator is calculated as:

$(12 * \text{the \# of VL tests conducted in the month}) / \text{\# of patients currently on ART in that month}$



- a. What is the expected relationship between the trend in ART retention rate- 12 months and VL suppression rate?

[In general, a patient retained on ART for a prolonged period of time is likely to achieve viral suppression. It is unlikely for VL suppression rates to be high if ART retention at 12 months is low. However, even if ART retention at 12 months is high, there may be other factors which cause the VL suppression rate to be low, e.g. poor long term adherence or development drug resistance.]

- b. The annualized viral load testing coverage is useful for assessing whether the VL suppression rate is likely to be representative of all patients on ART, and for monitoring progress as viral load testing services are scale-up. In this example, interpret the trend in ART retention vs. VL suppression and describe how the annualized VL testing coverage data is used in your interpretation.

[In this country example, the ART retention at 12 months is very consistent with the VL suppression rate for the first seven months shown. Starting in March 2018, the VL suppression rate drops. At the same time, there appears to be a large shift in the coverage of VL testing. The rate of testing increases by 40-60%. This suggests some change in the profile of the patients who are receiving viral load testing and who are included in the VL suppression rate data. It may be that in the earlier period those patients getting VL tests were from select facilities with a higher quality of care. Then as VL testing coverage expanded into other facilities, the patient population became more generalizeable. The manager might conclude that the VL suppression rate measured before March 2018 was less representative of the patient experience in the country, rather than there being a sudden drop in quality of care occurring from March 2018 onward. To test this hypothesis, the trends could be generated separately for facilities with a longer history of routine VL testing, from facilities where VL testing has been recently become more broadly available. In addition to viral load tests, managers could also look for changes in ART regimens among patients in these facilities which would be associated with treatment failure.]

- c. What follow up would a programme manager do if the trend in ART retention at 12 months were high, but the trend in viral load suppression were low or declining? And vice versa (what if the trend in ART retention at 12 months were low or declining, but the trend in viral load suppression were high?

[If ART retention rates appears high but VL suppression rates appear low, managers must determine whether VL testing is representative. Or this may reflect differences in practices for managing patients in the first year of treatment and patients who have been on treatment for a long time, or background HIV drug resistance. When the opposite relationship is found, the same issues may apply, i.e. VL testing may not be representative or there are explanations why the patients included in the ART retention measure are very different than those who received VL testing. Either way, managers must characterize the patient populations and look at more comparable groups of patients.]



World Health Organization  
20, Avenue Appia  
1211 Geneva 27  
Switzerland