

BETTER MONITORING HIV TREATMENT ADHERENCE IN SOUTH AFRICA: USING EXISTING DATA IN NEW WAYS TO GUIDE IMPROVEMENTS

There has been considerable expansion in routine data generation over the last decade in health care (and beyond) as well as in methodologies and technologies allowing innovative analysis and use of such data. Linking vast numbers of records and subsequent analyses is one such 'Big Data' method that has become an increasingly common activity for governments and private sector organizations in helping them deliver services more efficiently and effectively (see Annex A).

This policy brief, part of a World Bank series, tells the story of an innovative Big Data analysis of laboratory viral load tests undertaken in South Africa, which, when combined with data about HIV treatment coverage and access, provide new strategic information on viral suppression among South Africa's antiretroviral treatment (ART) clients by geography and demography. The analysis emphasises types of ART clients, health facilities and districts that need enhanced adherence support and identifies success stories that provide an opportunity for learning. Such strategic information should inform targeted ART programme supervision, mentoring and improvements.

THE PROBLEM

Most ART clients are treated in public sector facilities, but there is incomplete information at national, district and health facility level about the proportion of clients who adhere to treatment ...

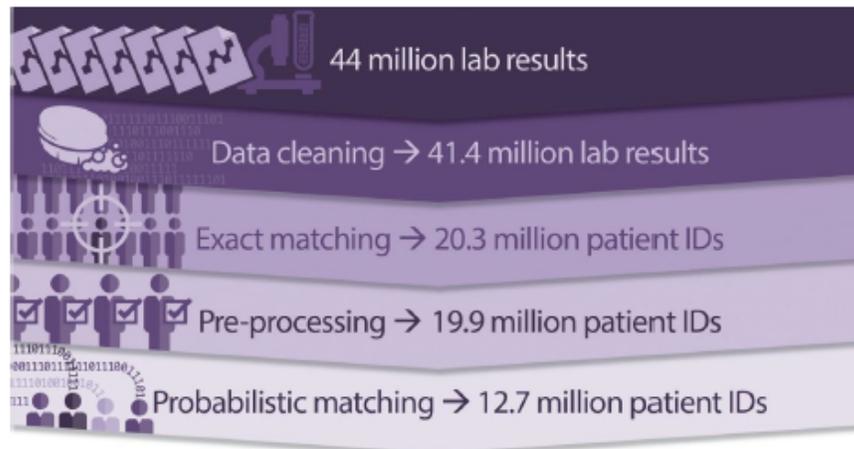
- ▶ South Africa has the largest ART programme in the world. Most ART clients are treated in public sector facilities, **but there is incomplete information at national, district and health facility level about the proportion of clients who adhere to treatment by monitoring the amount of HIV in their body i.e., measuring changes in their HIV viral load (VL).** Very low levels of HIV in ART clients (or viral load suppression, VLS) is the desired outcome of an HIV treatment programme. VLS brings benefits for the individual client like feeling better and living longer. It also brings benefits for their partners and for society at large through reduced HIV transmission, as virally-suppressed ART clients are less likely to transmit HIV to others.
- ▶ **Laboratory monitoring test results, including VL results from clients on HIV treatment, are stored in the National Health Laboratory Services' (NHLS) database but are not available in a format at health facilities or the district level to understand how clinics (and districts) are faring in terms of VLS of their clients on HIV treatment.** This national

- ▶ laboratory database has the potential to provide important strategic information on the HIV treatment programme’s reach and quality if VL test data were to be systematically linked to client data. Currently, the VL information used at the clinic level and for client management is incomplete due to fragmented health information systems between the laboratory (NHLS system) and the health facilities (Tier.net¹). The NHLS test results are not automatically merged with the client data stored in Tier.net due to a variety of reasons such as the lack of a unique identifier on most client records to allow linkage of client data, and the fact that the NHLS and Department of Health use separate lists of health facility names and identifications. In addition, there is insufficient capacity in some clinics to capture all laboratory data in the electronic patient record.
- ▶ To address the problem, several organisations worked together with the World Bank in a coordinating role, including the NHLS and its Corporate Data Warehouse, the Health Economics and Epidemiology Research Office, Boston University, University of Witwatersrand, University of Stellenbosch, University of Michigan, Clinton Health Access Initiative, Right to Care, HISP, BiTanium, and the South Africa National Department of Health ■

WHAT WAS DONE TO ADDRESS THE PROBLEM

.....
 The multi-stage procedure started off with 44 million VL test results and ended with 12.7 million estimated unique clients.

- ▶ We merged DHIS and NHLS facility lists to produce one consolidated master list of South African public sector health facilities providing ART.
- ▶ We then created a time-bound patient dataset (a ‘patient-linked cohort’) from the NHLS VL test data points using Big Data analytical methods including probabilistic record linkage. The multi-stage procedure (see figure) started off with 44 million VL test results and ended with 12.7 million estimated unique clients.

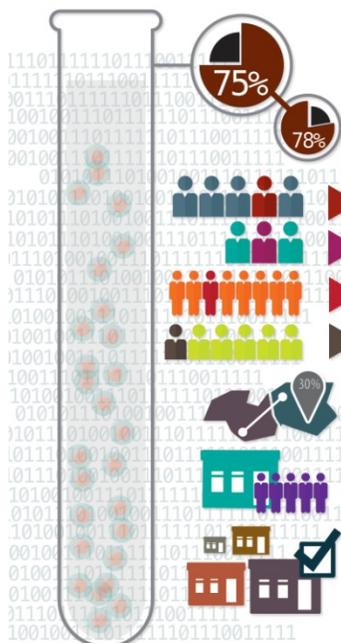


¹ Tier.net is the National Department of Health’s official patient HIV database which holds longitudinal patient demographic and clinical information about cohorts of patient started on ART.

- ▶ Next, using the newly developed “patient-linked cohort”, we estimated the proportion of clients receiving a VL test in a 12-month period at the facility level. Where clients had more than one VL test in a 12-month period, the most recent test was used.
- ▶ We grouped VL test results in four categories (<400, 400–1000, >1000, and >10,000 copies/mL), as per the VL-based client management guidance in the National ART guideline.
- ▶ Using the patient-linked cohort, we estimated the proportion of viral load tests done (VLD) and proportion of ART clients virally suppressed (VLS) by province, district, sub-district and health facility.
- ▶ We then assessed if there is any relationship between facility size (determined using the number of clients on ART at each facility) and viral suppression levels.
- ▶ Finally, we determined if poorer-performing facilities were spatially grouped (i.e. in one district). To do this, VL results from neighbouring clinics were compared to see if they were more similar than what would be expected if there was no spatial pattern or correlation ■

KEY FINDINGS AND POLICY RECOMMENDATIONS

SUMMARY OF RESULTS



75% had received at least one VL test in the past 12 months, Health Information System **only reports half**

Of these, **78%** clients virologically suppressed, but:

- ▶ **1 in 5** not suppressed
- ▶ **1 in 3** of the under 25s not suppressed
- ▶ **1 in 8** had VL >10,000 (high risk of transmission)
- ▶ **1 in 6** male patients >10,000

Best performing districts had 30% higher VLS than worst performing districts

Facilities and districts with higher ART patient numbers do better on VLS

200 clinics with VLS below 50%

3.6% of clinics reach VLS of 90% or more

1 SOUTH AFRICA IS NOT YET MEETING ITS TARGET OF 100% OF ART CLIENTS HAVING A VIRAL LOAD TEST ONCE A YEAR

The patient-linked cohort provided new insights: from April 2014 through March 2015, 3,775 public facilities reported 2,993,125 ART clients. During this period, 2,199,890 unique clients received 2,995,133 VL tests. This means that based on the patient-linked cohort, 75% of clients had received a VL test in the previous 12 months, short of the 100% of clients who should have had a VL test (this target is in each of the country's 52 District Implementation Plans).

KEY POLICY RECOMMENDATION:

To ensure that all ART clients have an annual VL result documented, South Africa needs to further create demand for, scale up, promote and electronically capture VL testing. Reaching this target will be a major step toward achieving high quality of care for ART clients as well as the early identification of non-adherence to treatment and potential drug resistance.

2 SOUTH AFRICAN ART CLIENTS HAVE A VIRAL SUPPRESSION RATE OF 78%. THERE IS AN INTENSIFIED EFFORT REQUIRED TO ACHIEVE THE 90% TARGET

Using the new patient-linked cohort, 78% of ART clients were virally suppressed (<400 copies/ μ l). This means that over 1 in 5 ART clients were not virally suppressed in South Africa.

58% of ART clients are known to be suppressed. This is the product of VL done and VL suppressed, and indicates the actual level of documented viral suppression in the client population.

KEY POLICY RECOMMENDATION:

VL results should be used systematically by health care workers and clients to jointly monitor the effectiveness of treatment. VL results should always be communicated to the client, as a "you are doing well" or "you are struggling, we need to monitor and change practices" message, with appropriate follow-up as per ART guidelines. Clients should be differentiated in terms of the VLS results. This will increase demand for VLS results from ART clients, frontline workers and facility managers.

3 SPATIAL ANALYSIS REVEALS VERY LARGE DIFFERENCES IN VIRAL SUPPRESSION LEVELS ACROSS SOUTH AFRICA

- ▶ Some provinces, districts, sub-districts and facilities did much better than others in terms of annual VL testing of their ART clients and achieving viral suppression (table). Three provinces had 25% or more clients not virally suppressed (i.e., VL test results above 1,000 copies/ μ l), and one province had 20% of clients still highly viremic (VL test results above 10,000 copies/ μ l). The best performing districts had 30% higher VLS levels than the worst

The best performing districts had 30% higher VLS levels than the worst performing districts.

performing districts. There were 200 clinics with VLS below 50% while one-in-30 facilities had 90% or more of ART clients virally suppressed.

Summary table of viral load results by level

	VL Test done in 12 month period, VLD		VL<400 cp/mL, VLS		VL>10000 cp/mL		Known to be suppressed (VLD x VLS)	
	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
National	75%	–	78%	–	12%	–	58%	–
Province	71%	82%	69%	82%	10%	20%	52%	65%
District	54%	99%	47%	86%	8%	35%	34%	73%
Facility	n/a	n/a	20%*	96%	1%	67%	n/a	n/a

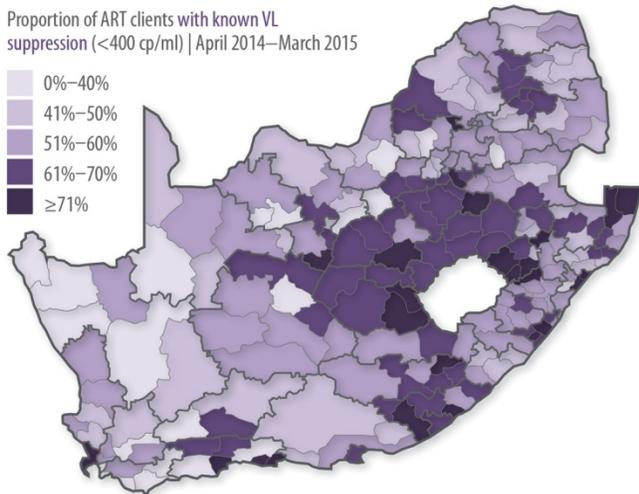
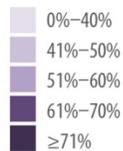
Note: *= the few facilities with lower percentages had sample sizes too small to take into account.

- ▶ The findings on viral load results at local level pinpoint the successes and shortfalls in the South African ART programme by identifying places where good practice can be studied and learnt from, as well as facilities and districts in need of additional supervision.

Viral suppression results: Identifying success

Can we learn from the-dark shaded sub-districts?

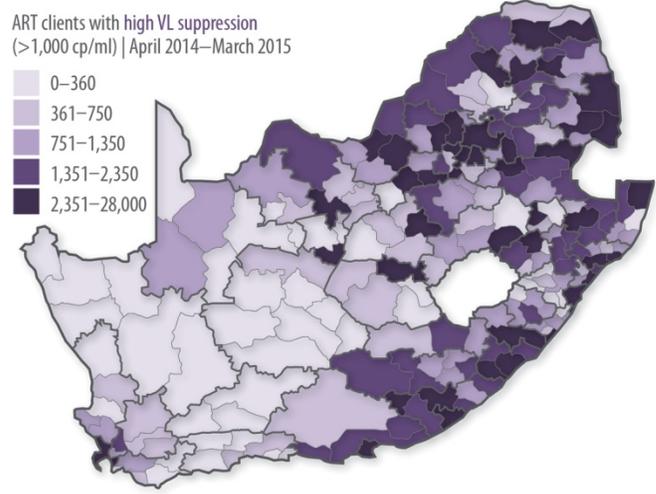
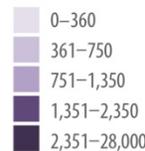
Proportion of ART clients with known VL suppression (<400 cp/ml) | April 2014–March 2015



Viral suppression results: Identifying failure

Low hanging fruit for better adherence support

ART clients with high VL suppression (>1,000 cp/ml) | April 2014–March 2015



- ▶ South Africa’s target of 90% VL suppression among ART clients is not yet being met by any entire district or province.
- ▶ The analysis also identified a pattern of VL suppression by size of client population. The districts and health facilities with larger client populations had better VLS levels than districts and facilities with low client numbers. This suggests that even if facilities have high client loads they can still achieve good VLS results, whereas facilities with small ART populations seem to have more difficulties in achieving VLS in their clients. Additional exploration of this trend may provide an opportunity for further learning.

KEY POLICY RECOMMENDATIONS:

VL data, disaggregated to the district and facility level where decisions about ART service delivery and which ART clients to support, are taken, should be made available. These results should be used to guide the allocation of resources for ART programme improvements, particularly better compliance with patient monitoring guidelines and targeted adherence support to those in need.

Health facilities achieving 90% VL suppression among their clients should be used as learning sites for quality improvement initiatives.

4 YOUNG PEOPLE BELOW 25 YEARS AND MEN ARE AT HIGHER RISK OF NOT ACHIEVING VIRAL SUPPRESSION AND HAVING VERY HIGH VIRAL LOADS

- ▶ 1 in 3 young people below 25 years were not virally suppressed. The rate of unsuppressed VLs was particularly high in the 0–4 year-olds (49%). This may in part be due to the lack of understanding of treatment and its monitoring by the parents or guardians of this population.
- ▶ Individuals with VL levels >10,000 copies/μl are highly infectious and therefore likely contribute to ongoing HIV transmission. One in six males and one in nine females on ART had a VL >10,000 copies/μl. 1 in 5 ART clients aged 15–24 years (a group with high sexual activity) had a VL >10,000 copies/μl. This requires the urgent targeting of treatment adherence and support to these clients.

KEY POLICY RECOMMENDATION:

There is an urgent need to strengthen treatment monitoring and adherence support in young people and men. If treatment as prevention is to work, the proportion of ART clients with a very high VL needs to be drastically reduced.

5 BETTER PERFORMING AND POORER PERFORMING SITES ARE SPATIALLY CORRELATED (MEANING THEY ARE MORE LIKELY TO BE IN THE SAME DISTRICT). THIS POINTS TO “ABOVE-FACILITY” FACTORS DETERMINING THE EFFECTIVENESS OF THE ART PROGRAMME

- ▶ Neighbouring facilities have more similar levels of virologic suppression than facilities that are further away suggesting that area factors also influence viral suppression.
- ▶ Possible explanations are shared patient populations with similar socio-economic profiles or shared provincial and district governance with similar policies, programmatic support and health systems factors.

1 in 3 young people below 25 years were not virally suppressed. The rate of unsuppressed VLs was particularly high in the 0–4 year-olds (49%).

KEY POLICY RECOMMENDATION:

In order to reach the VLS targets across South African districts, there needs to be better understanding of what drives high and low levels of viral suppression in districts, clinics and patients. This is an important research agenda and the NDOH/World Bank evaluation of the country’s new national treatment adherence guideline for chronic conditions will contribute to knowledge generation.

6 IMPROVED DATA AND LINKAGES ARE NEEDED. THE COUNTRY DOES BETTER THAN WHAT DOH REPORTS SUGGEST, AS A LARGE NUMBER OF VIRAL LOAD RESULTS ARE NOT CAPTURED IN THE DHIS

It is therefore essential to link and de-fragment health data information systems to better understand the performance and impact of the ART programme and ensure high quality clinical care.

- ▶ The 75% of clients who had received a VL test in the previous 12 months is short of the 100% of clients who should have had a VL test but much higher than what was reported in the DHIS (46%) in the same time period.
- ▶ One possible reason for this is the fact that the VL test results received by the facility are not entered into the DHIS (it needs to be done manually).
- ▶ If it is assumed that these VL test results that are not entered into the electronic patient system are not being used at the clinic level for programme monitoring or client management, then this wastage of resources for inconsequential VL testing would amount to more than 30 million US dollars annually.
- ▶ Improvement in linkages between data systems in South Africa are important to not only save costs, but to improve ART adherence monitoring and clinical patient management.

KEY POLICY RECOMMENDATION:

Important strategic data – like VL results—must not be lost across data systems. It is therefore essential to link and de-fragment health data information systems to better understand the performance and impact of the ART programme and ensure high quality clinical care ■

HOW THE DATA HAS BEEN USED AND NEXT STEPS

.....
 Hundreds of thousands of viral load tests not previously recorded and used in statistics could inform and support the efforts of the South African government to prioritise programme improvements by location and population for highest impact.

- ▶ The analysis represents an excellent baseline prior to South Africa’s fast-tracking of treatment scale-up through focused District Implementation Plans from 2016 onward. Recent initiatives to attain the third “90”² can be evaluated against this baseline.
- ▶ Hundreds of thousands of viral load tests not previously recorded and used in statistics could **inform and support the efforts of the South African government to prioritise programme improvements by location and population for highest impact.**
- ▶ Using the same patient-linked cohort, **CD4 immune reconstitution among ART clients** has also been studied in the largest ever national CD4 data analysis.³
- ▶ The patient-linked cohort has also been used to analyse the **HIV care cascades in different demographics and geographical areas, as well as to assess system-wide retention of ART clients in public sector care** in South Africa.
- ▶ Furthermore, health facility VL suppression results have been instrumental to **match intervention and control facilities in the NDOH/World Bank impact evaluation of South Africa’s first national adherence guidelines to improve adherence to chronic disease medications**, including ART. We expect that the various interventions articulated in the country’s national adherence guidelines, such as the adherence clubs, and the recently implemented “treat all”⁴ policy will further increase the number of those receiving ART.

High standards in VL detection will be essential for differentiated HIV care in South Africa, so that virally suppressed clients can be identified for decentralised medicine delivery options, while clients with elevated viral load can benefit from enhanced adherence counselling and alternative treatment regimens ■

² Refers to UNAIDS’ 90-90-90 targets: By 2020, 90% of all people living with HIV will know their HIV status, 90% of all people with diagnosed HIV infection will receive sustained ART, and 90% of all people receiving ART will have viral suppression.

³ NICD, 2016.
<http://documents.worldbank.org/curated/en/851301474884707261/Determinants-of-CD4-immune-recovery-among-individuals-on-antiretroviral-therapy-in-South-Africa-a-national-analysis>

⁴ All populations and age groups living with HIV are eligible for ART with no limitations on eligibility.

ANNEX A | BIG DATA APPLICATIONS FOR BETTER DECISION-MAKING

Big data is not just defined by volume but is as much about data complexity and heterogeneity. Advancements in big data analysis offer cost-effective opportunities to improve decision-making in critical development areas ...record linkage has become a more common activity for governments and private sector organisations as the extent of administrative and other big data has increased and as computing power has improved.

Several areas including science, economics, finance, business intelligence and health are exploring big data as a way to produce new information, make better decisions, and advance their systems and technologies. **Big data is not just defined by volume but is as much about data complexity and heterogeneity.** Some small datasets can be considered big data although they don't consume much physical space but are particularly complex in nature while some large datasets that require significant physical space may not be complex enough to be considered big data.ⁱ

Advancements in big data analysis offer **cost-effective opportunities to improve decision-making in critical development areas** including health care and disease prevention. Health care data frequently come from different information systems and disparate databases; to assess public health policies and monitor drug interventions these may need to be combined. Complex and varied data often require alternative approaches in processing to provide new insights and help enhance decision making. These methodologies go well beyond the traditional linkage of health records with other databases, which has been successfully implemented in numerous large studies. Similar approaches to record linkage have also been used to support public health surveillanceⁱⁱ, prevention research^{iii,iv}, and studies on the use and outcomes of health services^{v,vi}.

In situations where there is no unique identifier for record linkage, sophisticated methods need to be developed to match data across systems. Although each big data analysis has its specific objectives and methods, the work-flows are often similar. They may involve

- a) assessment of data quality
- b) pre-processing with data transformation, cleansing, anonymisation and blocking
- c) record linkage with deterministic and probabilistic algorithms for pairing, and
- d) validation^{vii}

Specifically in health, big data represent a challenge due to the need to retrieve, aggregate and process large data volumes from disparate databases, and sometimes deal with poor quality of data. Health informatics is also giving attention to techniques of privacy preserving record linkage.^{viii} The analysis of such matched data can uncover aspects of groups or individuals that are not obvious when a single database is analysed separately.^{ix} As a consequence, record linkage has become a more common activity for governments and private sector organisations as the extent of administrative and other big data has increased and as computing power has improved. In the absence of a unique identifier, probabilistic or "fuzzy" matching techniques can be employed to develop an identifier. The key task in fuzzy record linkage is to create a unique identifier that simultaneously minimises over-matching (falsely combining records that should remain separate) and under-matching (falsely separating records that should be combined).

- i What is big data? <http://www.villanovau.com/resources/bi/what-is-big-data/#.V2exSfkrLIU>, accessed 20 June 2016
- ii Gill L, Goldacre M, Simmons H et al. (1993). Computerised linking of medical records: Methodological guidelines. *J Epidem Comm Health* 47(4), 316-19.
- iii Guend P, Engholm G, Lyng E. (1990). Laryngeal cancer in Denmark: A nationwide longitudinal study based on register linkage data. *Br J Ind Med*, 47(7), 473-79.
- iv Van der Brandt PA, Schouten LJ, Goldbohm RA et al. (1990). Development of a record linkage protocol for use in the Dutch Cancer Registry for epidemiological research. *Int J Epidemiol* 19(3), 553-58.
- v Tyndall RM, Clarke JA, Shimmins J (1987). An automated procedure for determining patient numbers from episodes of care records. *Med Inform* 12, 137-46.
- vi Thomas JW, Holloway JJ (1991). Investigating early readmission as an indicator of quality of care studies. *Med Care* 29(4), 377-94.
- vii Pita R, Pinto C, Melo P et al. (2015). A Spark-based workflow for probabilistic record linkage of healthcare data. Workshop Proceedings of the EDBT/ICDT 2015 Joint Conference (March 27, 2015, Brussels, Belgium) on CEUR-WS.org (ISSN 1613-0073).
- viii Hassanien AE, Azar AT, Snasel V, Kacprzyk J, Abawajy JH (eds). *Big Data in Complex Systems: Challenges and Opportunities*
- ix Christen, P. (2006). A Comparison of Personal Name Matching: Techniques and Practical Issues. Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06), (September), 290–294. <http://doi.org/10.1109/ICDMW.2006.2>

For more information, please contact:

Dr Sergio Carmona (Sergio.Carmona@nhls.ac.za)