Remote sensing of HIV care programmes using centrally collected laboratory results; Can we monitor ART programme effectiveness?

Carl Morrow(PhD)[1*], Alison Munro[2], Marisa Wilke (M.Cur)[3], Ruth Stark (PhD)[3], Robin Wood (DTM&H, MMed, FCP (SA))[1]

1: Desmond Tutu HIV Centre, Institute of Infectious Disease and Molecular Medicine, Department of Medicine, Faculty of Health Sciences, University of Cape Town, South Africa.

2: Southern Africa Catholic Bishops' Conference (SACBC) AIDS Office, PO Box 941, Pretoria 0001, South Africa

3: Catholic Relief Services/South Africa, 41 Wierda Road West, Wierda Valley, Johannesburg 2196, South Africa

*Corresponding Author
Carl.Morrow@hiv-research.org.za
Desmond Tutu HIV Centre
Institute of Infectious Disease and Molecular Medicine
Department of Medicine
Faculty of Health Sciences
University of Cape Town
Anzio Road, Observatory 7705
Tel: +27 (21) 650 6957
Fax: +27 (21) 650 6963

Other author contacts:
Alison Munro (AMunro@sacbc.org.za)
Marisa Wilke (marisa.wilke@crs.org)
Ruth Stark (Ruth.Stark@crs.org)
Robin Wood (Robin.Wood@hiv-research.org.za)

Conflicts of Interest
The Authors have no conflicts

Abstract (247 words)

Aims

This paper describes a population level monitoring system of patients on antiretroviral therapy using centrally collected laboratory data. We demonstrate an analogous process of remote sensing using a large set of laboratory results and illustrate the tremendous density of information stored. We moved from an individual to a community view of ART rollout which is similar to the earth and biological sciences where remote sensing is used once the spatial scale of the investigation is too large to be done at ground level.

Methods

The study was a retrospective cohort study of patients from 2004 to June 2011. A total of 188 759 individual laboratory results representing 26 445 patients were analysed for average CD4 and viral load by year.

Results

The data showed an increasing state of health of the population and allowed for hypothesis generation when the trends did not follow expected paths.

Conclusion

In this analysis we moved away from individual centred data to population level data in order to assess ART programme performance. We showed that routine patient monitoring data had great utility in assessment of population health.

These methods are useful in monitoring and evaluation and effectiveness studies as they are easy to collect, reliable (not needing much human matching or interventions) and scalable from a single clinic to an entire population. The larger the sample size the more reliable the results as confounders such as incorrectly identified transfers out, lost to follow up patients and transfers in would be removed.

Introduction

Since the introduction of treatment for the AIDS epidemic through the large scale roll out of Antiretroviral therapy (ART), the population of people receiving therapy has increased dramatically[1]. As a result of this, it is becoming increasingly difficult and expensive to collect and collate data from individual patients in order to compile reports of the patient entry and outcomes which provide insights into the success of the roll out programmes.[2,3] The need for continued improvement in monitoring and evaluation (M&E) systems was again highlighted in the National Strategic Plan for HIV and AIDS, STIs and TB, 2012-2016 (NSP).[4] Furthermore, the critical role of integrating M&E with electronic systems was emphasised by the World Health Organisation (WHO).[5]

It is accepted that the drugs used in ART are highly effective at suppressing HIV in humans as demonstrated by randomised controlled clinical trials.[6] The controlled environment that exists within trials is very different from the clinical setting where drug delivery occurs in an overburdened health system to sick, frequently impoverished individuals where adherence to treatment regimen cannot be assumed. This problem has been identified for a very long time where the impact of ART on a population will be primarily determined by programmatic issues of treatment availability, accessibility and delivery.[7]

These practical obstacles call for the need to develop an innovative approach to monitoring ART programmes that is simple, efficient and provides results that can be reliably interpreted. In the earth and biological sciences extensive use is made of remote sensing when the spatial scale of the investigation is too large to be done at ground level on individual plants or geographical units.[8] Essentially an image is taken of the landscape from a distance and the information contained in the image is analysed in order to interpret patterns that are present. In these examples, technology is used to assess health of forests and other landscapes, monitor nutrient levels in standing crop plants on farms and identify early outbreaks of diseases.[8] This information is then used to target specific interventions.

Currently routine laboratory tests are seen as a necessary expense in individual patient care.[9] These records are stored in the individual's clinical notes but also, importantly, in centralised databases maintained by the clinical laboratory services. The aim of this paper is to demonstrate an analogous process of remote sensing by making use of a large, well collected data set to illustrate the tremendous density of information stored in these data that can be used to gain valuable insights into programme performance and population responses to ARV treatment.

Methods

The study was a retrospective cohort study of patients within a South African NGO run clinic system for the period 2004 to June 2011. Nineteen data sets representing 17 clinics from 5 of the 9 provinces in South Africa were supplied from the Southern African Catholic Bishops' Conference/Catholic Relief Services (SACBC/CRS) ART programme. Each data set consisted of three.csv files containing data of laboratory tests, a main, annonomised, clinical and demographic data file and regimen information. The individual files were imported into a Microsoft® Access database. Duplicate records, records with date errors and mismatched dates (e.g. patient outcome before enrolment) were removed which resulted in 47 107 valid patient records, 34 907 regimen records and 410 352 blood tests comprising 141 019 blood sample records. 28 186 patients were recorded as having started ART. Approval to analyse these data was granted by the institutional ethics review board (RecRef

169/2007). HIV positive men, woman and children who were on ART and had one or more CD4 or viral load blood result were included in the analysis. The last Viral Load and CD4 prior to ART start (baseline) and all CD4 and Viral Load values while on ART (before the recorded outcome date) were extracted from the database. This yielded a database of 188 759 individual laboratory results representing 26 445 patients. No cut-off was applied to the period of time between baseline sample date and ART start date. Within this analysis there were 960 patients whose first CD4 or Viral Load or both was more than 6 months prior to treatment start.

CD4 analysis

Using the patent identification number, the first CD4 result for each patient was ascertained and then the time to the following samples was calculated. These values were then grouped into 6 month time periods to yield an average CD4 recovery for the study group. Each patient provides one point in each time period (depending on their time on treatment).

The mean CD4 for each patient in each year was determined. The means were then grouped into 100 unit CD4 strata. Each year was represented as a line with the CD4 category along the x-axis and the number of patients on the y-axis. Alternatively the strata were divided up into the biologically relevant categories (0 to 200, 200 to 350, 350 to 500 and greater than 500 cells.mm^{-3}). Each calendar year was plotted as a bar with the proportion of the population in each CD4 category making up the components. It should be noted that if a patient had been on treatment for more than one year then their results would appear in each of the years that they are on treatment.

Viral Load analysis

The viral load trajectory for each patient over time was determined by identifying the first viral load (baseline) and then grouping the viral loads into 6 month time periods and then presenting the proportion of the group with viral load in the different categories (suppressed, with a viral load of less than 50, 50 to 400 and greater than 400 copies.mm^{-3})

The mean $\log_{10}$ viral load for each patient for each year was calculated and the proportion of the population in each of the one log unit categories was plotted against the calendar year. It should be noted that if a patient had been on treatment for more than one year then their results would appear in each of the years that they are on treatment.

Results

When presenting data on the recovery of individual patients it is common to see curves showing the increase in CD4 value over time of patient treatment (Fig. 1a). This figure shows the efficacy of the drug in people who are adherent to their treatment. The error bars demonstrate that there is diversity in the recovery with treatment failures and non adherence influencing the median CD4 value. It should also be noted that there are fewer and fewer patients represented as one moves to the right of the figure

Summarising the mean CD4 value for each patient in each year yields Figs. 2 and 3. In these figures one is able to assess the health of the total clinic population. In Fig. 2 the size of the treated population can be seen (area under the curve) and the health can be assessed by seeing where the peak of the curve lies and the bulge seen in later years of people with CD4 counts above 400 cells.ml^{-1}. Taking the mean CD4

value for each patient for each year and categorising it in a CD4 stratum yields Fig. 3. This very succinctly shows the way in which the population health is improving thus the effectiveness of the roll out of the treatment.

This analysis did not take the age of the patients into account and so it could be argued that the increases seen are due to an influx of children into the clinics. Table I shows that the proportion of children (under 10 years of age) is small and has been decreasing over the years and the child data have not influenced the curves towards higher CD4 values.

The traditional view of the improvement of individual patient Viral Loads over time of treatment is shown in Fig. 1b. We would expect all ART naïve patients to have a viral load greater that 50 copies.mm^{-3}. The conclusion here is that over the full monitoring period around 25% of patients are transferring into the clinics and are virally suppressed and currently on ART treatment. This is significant from a monitoring and evaluation perspective as it is difficult to capture this proportion in a clinic setting.

Moving away from the patient centred to community centred view we can see that the population viral suppression rate has increased dramatically over the 7 year study period (Fig. 4) with over 60% of the population having an average viral load of below 400 in 2010 (the last complete year in the analysis).

Fig. 5 shows the mean $\log_{10}$ viral load for each patient for each year represented as a proportion for that year. A value of 1.75 is virally suppressed. One can see here again the way in which the population health has improved over time with less and less virus being present.

Discussion

This paper clearly demonstrates the way laboratory data can be used to view a care system from a distance - remote sensing of the population health. Once the data are compiled, a relatively simple analysis gives a very powerful view of the population level response to ART which is of great importance to the assessment of a public health ART roll out. This is very different to the traditionally held view that the patient folder is the primary source of information that needs to summarised through reporting tools to allow for the assessment of a treatment programme.

There are many advantages of using these laboratory data in this evaluation role. The data are collected at a more centralised location (one pathology laboratory servicing a region) and can be accessed as a single data download as opposed to trying to collate monthly or quarterly records into a continuous history. The results are reliable as they are directly used in clinical practice and so internal quality controls are already in place to ensure consistency and accuracy of the testing.

Minimal change is needed to the existing health infrastructure in order to implement this analysis. It can be started immediately and existing historical data allows for comparisons of current clinical systems to those of the past without having to wait for the accumulation of new indicators.

There is a minimal cost for this analysis as all of the data are routinely collected and so all that is needed is the time required for the analysis of the results. With some development, a reporting function can be built into the existing databases to generate these results automatically.

In this paper we have performed a very 'clean' analysis to show an ideal, well run clinic situation. Unselected, real world, data is more chaotic but can still be analysed in this way. Not taking the patient's ART start date into account resulted in a larger data set (222 901 records) but gave equivalent aggregate CD4 and viral load

curves. These are not shown because the CD4 and viral load recovery trajectories of the complete data are not comparable to the traditional recovery curves (Figs. 1a and 1b). The pre-treatment laboratory data shows an initial decline in CD4 after programme entry followed by a sharp rise when ART is initiated. Mirroring this is a steady increase in viral load followed by sharp decline. These data cannot be reliably used due to the selection bias in this group with the CD4 threshold and other clinical data governing when the patients initiate ART. Many clinics make use of electronic pharmacy systems[10] that could be linked to the laboratory data to give an accurate start date that can then be used in the analysis.

We have performed this analysis on other data sets of different data quality (RW & CM unpublished data) and it has been found that the curves seen in Fig. 2 can move 'backwards' with the peak in later sampling years at a lower CD4 value and the proportion of patients in higher CD4 categories becoming fewer. If the population size is still increasing, then this could be attributed to the clinic recruiting patients who are sicker, or the existing patients are failing treatment. If the population is decreasing then the conclusion could be that the clinic is losing healthy patients and retaining or recruiting sick ones. It can be seen that these analyses allow for hypothesis generation that can then be investigated within the clinic or district. In this situation, Fig. 3 would show 'flat' strata with little increase in the proportion of patients in the higher CD4 categories over time. It is not possible to elucidate the individual factors that may be contributing to a poor performance curve (dysfunctional clinic, poor adherence, cultural issues, associated issues such as alcohol or drug abuse etc) but it does allow for a broad scale analysis that quickly identifies areas of concern where investigations can be directed.

The minimum requirements for this analysis are that the data need to be in an electronic format, to include a date of sample, the sample type (CD4 or VL), the value, and a patient identifier to allow for the per person analysis. The location (area) of where sample was taken would need to be added if a comparison of clinics or other spatial analysis is needed.

Patient identifiers entered into the laboratory database are used to link records together. Although useful, they are not critical to these analyses. If there are no identifiers there would be double counting of patients that have more than one test in a year. The area under the curve becomes number of tests rather that number of people. In our experience it is difficult to link different laboratory systems together largely due to the challenges of linking patient identifiers. This means that large, centrally stored data is the most useful which does challenge the perspective of the current trend of devolving CD4 testing to the clinics using point of care machines to close the gap between testing and reporting which prevents the loss of patients.[11] A way of linking these point of care machines into a network would need to be created or these data will be lost.

The success of an ART programme relies on its effectiveness. As has been shown in other papers[12,13] there is a need to assess the viability, successful implementation and uptake of treatment interventions in a public health setting after the efficacy of a treatment has been demonstrated in the artificial environment of a clinical trial or pilot study. Using the methods described here we are now able to directly asses the effect the treatment regime is having on the population at large. It should be kept in mind that only the 'on treatment' population is being monitored[14] but through the increased roll out of the ART programme the size of the pool of uncontrolled HIV that has not entered the system will decrease.

In this analysis we have shown that population CD4 values are increasing which will result in decreases in co-morbid diseases such as tuberculosis.[15]   Mean population CD4 values in Africa have been summarised[16] and Fig. 2a shows that in spite of the treatment of the population, the distribution is still very much lower that that of an uninfected population.

Population viral load is decreasing which can be used as a surrogate measure of the transmission rates within the community.[14]   It is for this reason that we do not need to remove the newly recruited patients as they are still as source of infection until they have become virally suppressed.   This shows the value in taking a baseline viral load measurement at treatment start as this would allow health managers to assess the pool of virus in the community or health district that they are managing.   In our analysis we calculated average $\log_{10}$ viral load for each patient in the year and then presented this as a distribution of values while the community viral load (CVL) used the last viral load in each year.[14]

Monitoring the progress of a treated population is difficult.[2]   From clinic based records it is relatively easy to collect data on programme entry but becomes very difficult to ascertain programme losses and programme cycling (multiple transfers in and out of clinics).   With the widespread availability of treatment, an outcome of LTFU is no longer equivalent to death.[17]   Transfers among clinics are becoming more a more common as patients optimise their care strategy[18] and this is very difficult to monitor at a district level as the individual clinic's lost to follow up assessment may be unrecorded 'transferred out' patients moving to a nearby clinic and new patients entering a clinic's programme may be unreported 'transfer ins'.[14]   By looking at the programme entry viral load an estimate of the transferring in patients can be made and it is interesting that the 25.4 % unrecognised transfers rate seen previously[18] is similar to the results we are seeing from the laboratory records..   Furthermore, this analysis is designed to take a regional view of a programme and so if the scale is large enough (health district level) individual transfers would not influence the population level CD4 and viral load assessments.

Reporting indicators looking at a patient's functional status is often filled with ambiguity and mis classification.[2]   Monitoring the CD4 status shows a population level of health and, by association, predicted functional status.   There are moves to redefine indicators to be better at assessing coverage (programme and population) and programme quality (for example PEPFAR[19]).   This analysis provides data to assess both of these. In this paper we are developing new indicators that give a population view of programme effectiveness.

Conclusion

In this analysis we moved away from individual centred data to population level data in order to assess ART programme performance.   We showed that routine patient monitoring data had great utility in assessment of population health.   The numbers of patients included in our analysis are approximately equivalent to those needing ART (40 000) within a typical health sub district of four to five hundred thousand residents.[20]

These methods are useful in monitoring and evaluation and effectiveness studies as they are easy to collect, reliable (not needing much human matching or interventions) and scalable from a single clinic to an entire population.   The larger the sample size the more reliable the results will be as confounders such as incorrectly identified transfers out, lost to follow up patients and transfers in would be removed.

References

1: WHO, UNICEF, UNAIDS. Global HIV/AIDS response: epidemic update and health sector progress towards universal access: progress report 2011. Geneva, 2011:97 http://www.who.int/hiv/pub/progress_report2011/en/index.html (accessed 13 January 2012)

2: Hoskins S, Weller I, Jahn A, et al. An appraisal of indicators used to monitor the treated population in antiretroviral programmes in low-income countries. AIDS. 2010 Nov 13;24(17):2603-7.

3: Harries AD, Gomani P, Teck R et al. Monitoring the response to antiretroviral therapy in resource-poor settings: the Malawi model. Trans R Soc Trop Med Hyg. 2004 Dec;98(12):695-701.

4: National Strategic Plan for HIV and AIDS, STIs and TB, 2012-2016: Draft Zero for Consultation. August 2011. http://www.sanac.org.za/files/uploaded/519_NSP%20Draft%20Zero%20110808%20 pdf%20%20final.pdf (accessed 9 January 2012)

5: WHO. Global health sector strategy on HIV/AIDS 2011-2015. Geneva 2011. http://whqlibdoc.who.int/publications/2011/9789241501651_eng.pdf (accessed 13 January 2012)

6: Mathis S, Khanlari B, Pulido F. Effectiveness of protease inhibitor monotherapy versus combination antiretroviral maintenance therapy: a meta-analysis. PLoS One. 2011;6(7):e22003. Epub 2011 Jul 19. (http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0022003)

7: Lawn SD, Myer L, Wood R. Efficacy of antiretroviral therapy in resource-poor settings: are outcomes comparable to those in the developed world? Clin Infect Dis. 2005 Dec 1;41(11):1683-4 (http://cid.oxfordjournals.org/content/41/11/1683.long)

8: Navalgund Ranganath R.; Jayaraman V.; Roy P. S. Remote sensing applications: An overview. Current Science 2007 December 93(12):1747-1766

9: Stover J, Bollinger L. The Costs of AntiRetroviral Therapy in Low and Middle Income Countries Present Commitments and Future Needs. Futures Institute, 41-A New London Tpke, Glastonbury, CT 06033 USA (http://www.futureartcosts.org/The_Costs_of_Anti-Retroviral_Therapy.pdf, accessed 18 November 2011)

10: Wood R, Kaplan R, Bekker LG, Brown S, Rivett U. The utility of pharmacy dispensing data for ART programme evaluation and early identification of patients lost to follow-up. S Afr J of HIV Med 2008, (Autumn issue):44-48. (http://www.ajol.info/index.php/sajhivm/article/viewFile/34866/6496)

11: Peter T, Badrichani A, Wu E. et al. Challenges in implementing CD4 testing in resource-limited settings. Cytometry B Clin Cytom. 2008;74 Suppl 1:S123-30.

(http://onlinelibrary.wiley.com/doi/10.1002/cyto.b.20416/abstract;jsessionid=A7A6D
D2B54B4C2F76D30DC0C12561168.d01t04)

12: Coetzee D, Hilderbrand K, Boulle A, Draper B, Abdullah F, Goemaere E.
Effectiveness of the first district-wide programme for the prevention of mother-to-
child transmission of HIV in South Africa. Bull World Health Organ. 2005
Jul;83(7):489-94. (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2626297/)

13: Tchendjou P, Same-Ekobo C, Nga A. Effectiveness of multidrug antiretroviral
regimens to prevent mother-to-child transmission of HIV-1 in routine public health
services in Cameroon. PLoS One. 2010 Apr 29;5(4):e10411.
(http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0010411)

14: Das M, Chu PL, Santos GM. Decreases in community viral load are accompanied
by reductions in new HIV infections in San Francisco. PLoS One. 2010 Jun
10;5(6):e11068.
(http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0011068)

15: Lawn SD, Myer L, Edwards D, Bekker LG, Wood R. Short-term and long-term
risk of tuberculosis associated with CD4 cell recovery during antiretroviral therapy in
South Africa. AIDS. 2009 Aug 24;23(13):1717-25.

16: Thakar MR, Abraham PR, Arora S. Establishment of reference CD4+ T cell
values for adult Indian population. AIDS Res Ther. 2011 Oct 3;8:35.
(http://www.aidsrestherapy.com/content/8/1/35)

17: Dalal RP, Macphail C, Mqhayi M et al. Characteristics and outcomes of adult
patients lost to follow-up at an antiretroviral treatment clinic in johannesburg, South
Africa. J Acquir Immune Defic Syndr. 2008 Jan 1;47(1):101-7.

18: Government of Malawi Ministry of Health Quarterly HIV Programme Report on
HIV Testing and Counseling, Prevention of Mother to Child Transmission and
Antiretroviral Therapy: April - June 2011.
http://www.hivunitmohmw.org/uploads/Main/Quarterly%20HIV%20Programme%20
Report%202011%20Q2.pdf (accessed 18 November 2011)

19: The President's Emergency Plan for AIDS Relief (PEPFAR). Next generation
indicators reference guide. http://www.pepfar.gov/documents/organization/81097.pdf
(accessed 18 November 2011)

20: Statistics South Africa website
http://www.statssa.gov.za/publications/statsdownload.asp?PPN=P0302&SCH=4696
(Accessed 1 March 2011)

Table I

Table I: Number of laboratory samples attributed to children under 10 years of age showing the decreasing proportion of samples over time.

Year	Total samples	Age <= 10 years in year of sample	Age > 10 years in year of sample	Percentage of <= 10 year olds
2004	1645	164	1481	9.97
2005	8101	517	7584	6.38
2006	14472	946	13526	6.54
2007	21498	1384	20114	6.44
2008	30306	1641	28665	5.41
2009	40606	1992	38614	4.91
2010	52070	2424	49646	4.66
Overall	168698	9068	159630	5.68

Figure 1

a
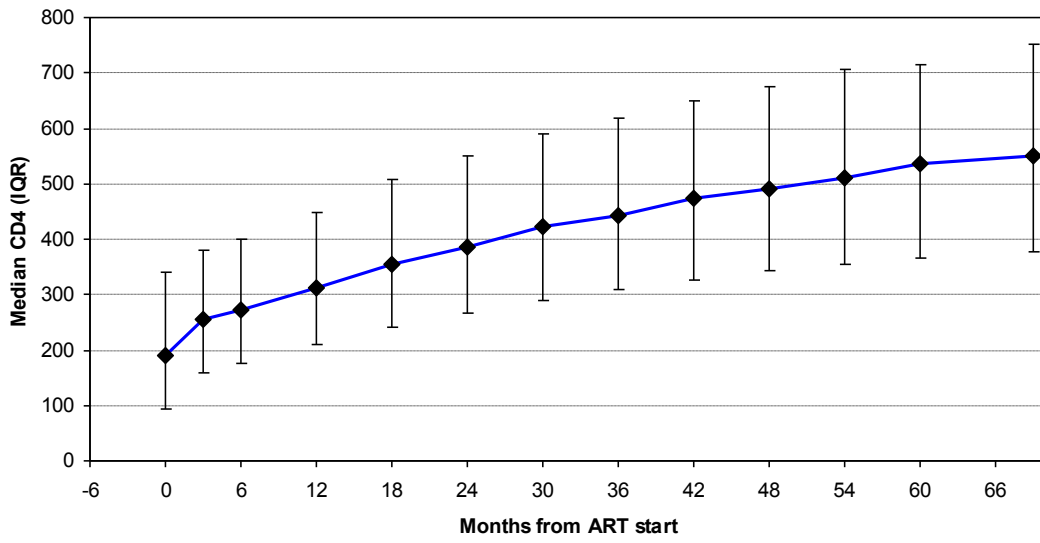
b
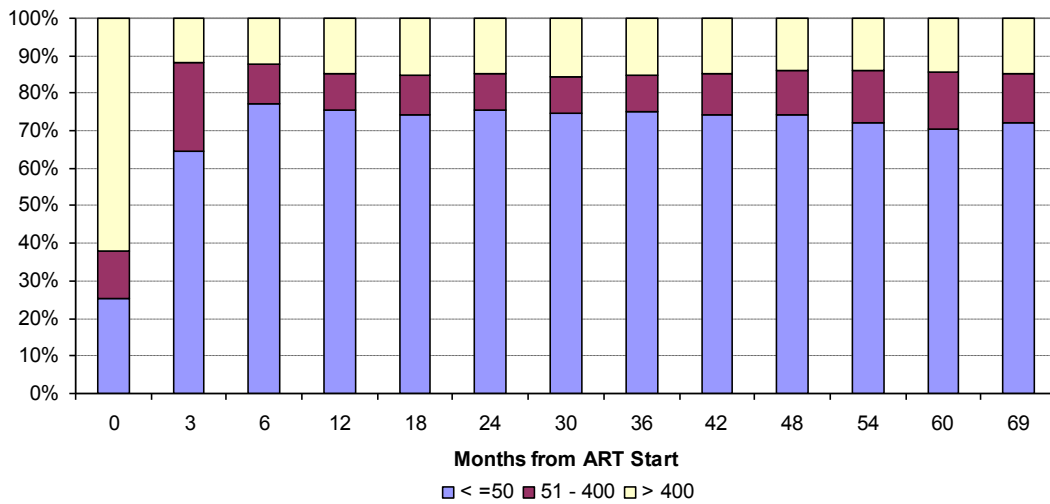
Fig. 1. Patient recovery (a) median CD4 and IQR (cells.mm^{-3}) and (b) proportions of viral load suppression over time showing the efficacy of the antiretroviral therapy in this clinic system.
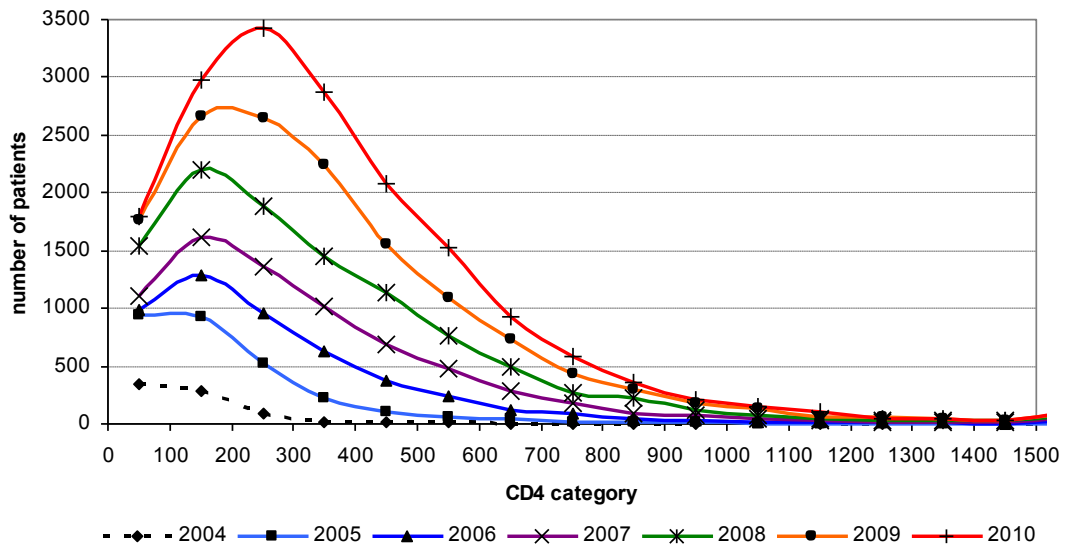
Figure 2

Fig. 2. Mean annual CD4 category (cells.mm^{-3}) showing increasing size of the patient population over time along with the improving health of the population demonstrated by peak moving to the left along with the increasing size of the proportion of the curve to the left of the peak.
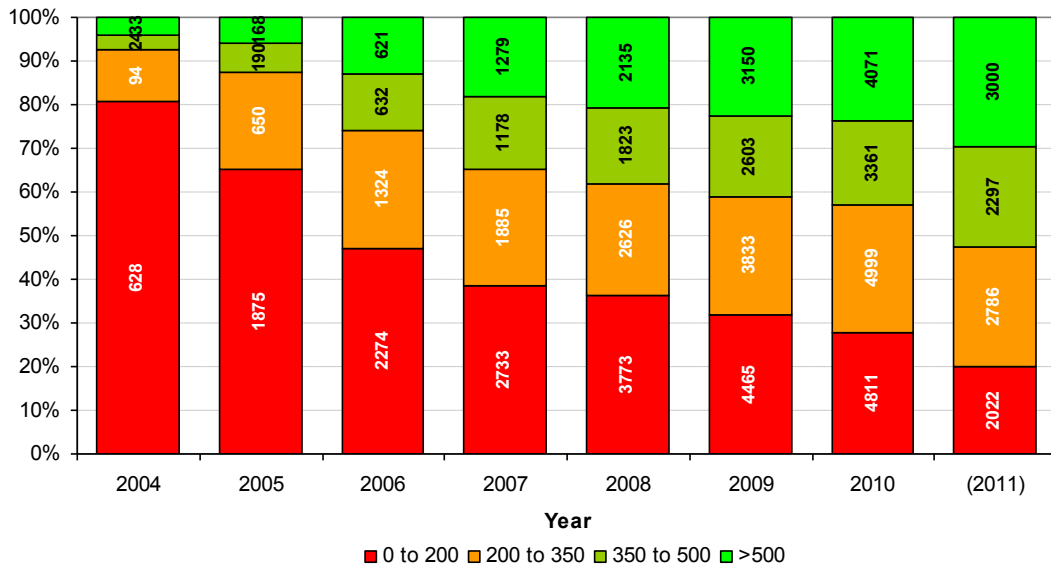
Figure 3

Fig. 3. Mean annual CD4 (cells.mm^{-3}) divided in strata the improving health of the population over time seen by the diminishing size of the lower categories. The values in the boxes are the number of patients in the category.
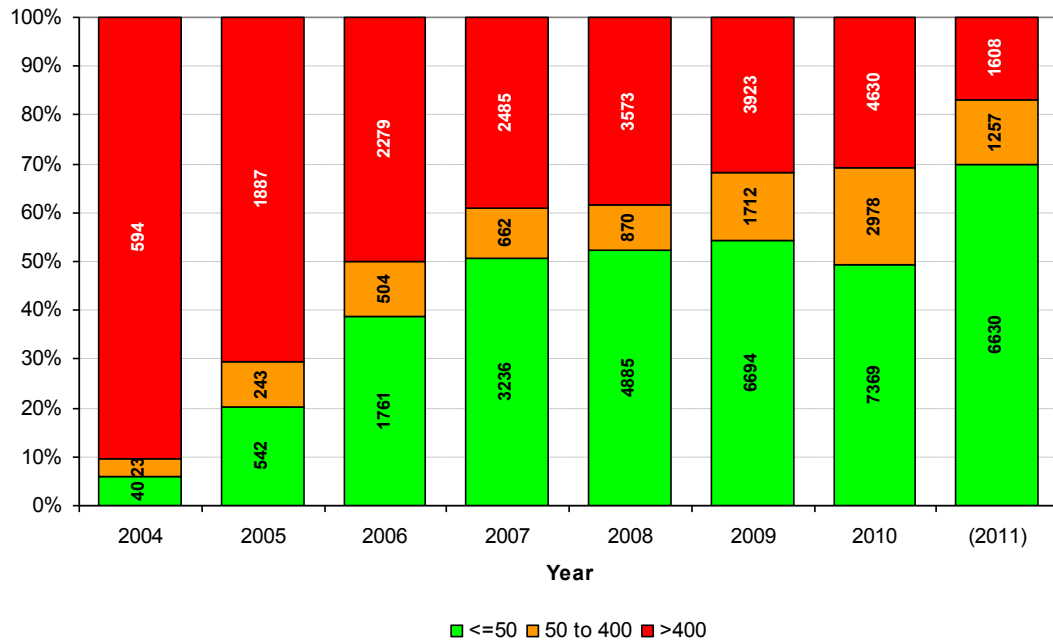
Figure 4

Fig. 4. Mean annual viral load (copies.mm^{-3}) showing increasing population viral suppression. The values in the boxes are the number of patients in the category.
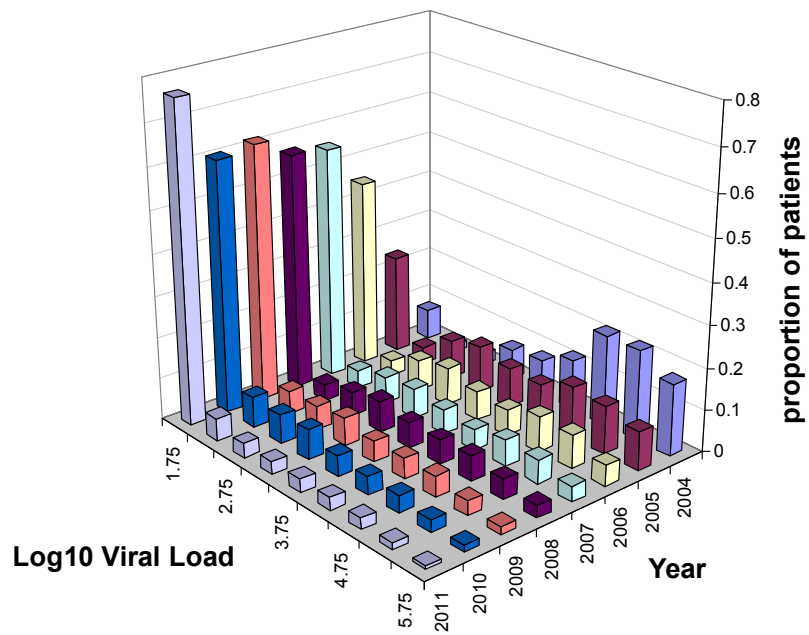
Figure 5

Fig. 5. Proportion of patients at different $\log_{10}$ viral loads over time. A value of 1.75 is virally suppressed. Here it can be seen how over time the proportion of virally suppressed patients has increased while the proportion of people with high viral loads has steadily decreased over time.