



---

## UW Biostatistics Working Paper Series

---

2-18-2009

# Pooled Nucleic Acid Testing to Identify Antiretroviral Treatment Failure during HIV Infection

Susanne May

*University of California, San Diego*, [smay@ucsd.edu](mailto:smay@ucsd.edu)

Anthony Gamst

*University of California, San Diego*, [agamst@ucsd.edu](mailto:agamst@ucsd.edu)

Richard Haubrich

*University of California, San Diego*, [rhaubrich@ucsd.edu](mailto:rhaubrich@ucsd.edu)

Constance Benson

*University of California, San Diego*, [cbenson@ucsd.edu](mailto:cbenson@ucsd.edu)

Davey Smith

*UCSD, VA San Diego Hlthcare System*, [davey@ucsd.edu](mailto:davey@ucsd.edu)

---

## Suggested Citation

May, Susanne; Gamst, Anthony; Haubrich, Richard; Benson, Constance; and Smith, Davey, "Pooled Nucleic Acid Testing to Identify Antiretroviral Treatment Failure during HIV Infection" (February 2009). *UW Biostatistics Working Paper Series*. Working Paper 343. <http://biostats.bepress.com/uwbiostat/paper343>

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

Copyright © 2011 by the authors

**Title: Pooled Nucleic Acid Testing to Identify Antiretroviral Treatment Failure during HIV Infection**

Running Head: Comparison of Pooling Algorithms

**Authors:**

Susanne May<sup>1</sup>, <smay@ucsd.edu> , Anthony Gamst<sup>2</sup>, <agamst@ucsd.edu>, Richard Haubrich<sup>3</sup>, <rhaubrich@ucsd.edu>, Constance Benson<sup>3</sup>, <cbenson@ucsd.edu> , Davey M. Smith<sup>3,4</sup>, <davey@ucsd.edu>

**Affiliations:**

<sup>1</sup> Department of Biostatistics, University of Washington Seattle (This work was performed while at the Division of Biostatistics and Bioinformatics, Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, California,) USA

<sup>2</sup> Division of Biostatistics and Bioinformatics, Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, California, USA

<sup>3</sup> Department of Medicine, University of California San Diego, UCSD, La Jolla, California, USA

<sup>4</sup> Veterans Affairs San Diego Healthcare System, San Diego, California, USA

**Word Count:** 3531

**Abstract:** 195

**Address correspondence to:**

Davey M. Smith, M.D.,  
University of California San Diego,  
9500 Gilman Drive 0679,  
La Jolla, CA 92093-0679, USA  
Tel: (858) 552-4339;  
Fax: (858) 552-7445;  
E-mail: davey@ucsd.edu

Collection of Biostatistics  
Research Archive

## **Abstract**

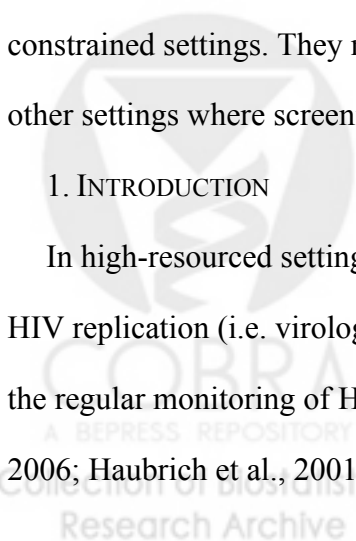
**Background:** Pooling strategies have been used to reduce the costs of polymerase chain reaction based screening for acute HIV infection in populations where the prevalence of acute infection is low ( $<1\%$ ). Only limited research has been done for conditions where the prevalence of screening positivity is higher ( $>1\%$ ).

**Methods and Results:** We present data on a variety of pooling strategies that incorporate the use of PCR-based quantitative measures to monitor for virologic failure among HIV-infected patients receiving antiretroviral therapy. For a prevalence of virologic failure between 1% and 25%, we demonstrate relative efficiency and accuracy of various strategies. These results could be used to choose the best strategy based on the requirements of individual laboratory and clinical settings, such as required turnaround time of results, and availability of resources.

**Conclusions:** Virologic monitoring during antiretroviral therapy is not currently being performed in many resource constrained settings largely because of costs. The presented pooling strategies may be used to make such monitoring feasible and to optimally limit the development and transmission of HIV drug resistance in resource constrained settings. They may also be used to design efficient pooling strategies for other settings where screening involves quantitative measures.

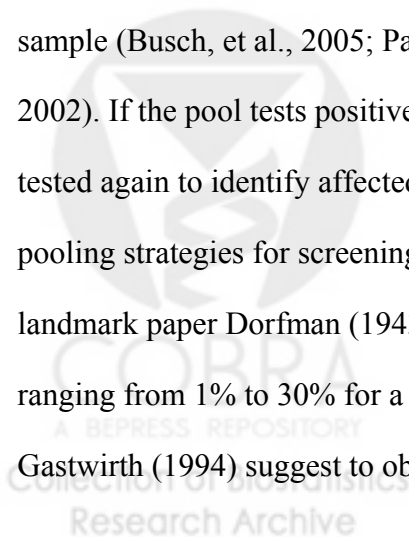
### 1. INTRODUCTION

In high-resourced settings, the failure of antiretroviral therapy (ART) to suppress HIV replication (i.e. virologic failure) while a patient is receiving ART is detected by the regular monitoring of HIV RNA levels in the blood (viral loads) (Hammer et al., 2006; Haubrich et al., 2001; Hughes et al., 1997). Viral loads are not performed



regularly in resource-constrained settings, where changes in CD4 cell count are used as a surrogate, although CD4 –based criteria have been shown to be insensitive in detecting virologic failure during ART (Bisson et al., 2008; Bisson et al., 2006; Fiscus et al., 2006; Moore et al., 2006; Petti et al., 2006). It is, therefore, imperative to develop and implement less expensive methods to monitor for viral replication during ART in resource-constrained settings to limit the development and transmission of drug resistant HIV (Calmy et al., 2007; Durant et al., 1999; Miller and Larder, 2001; Phillips et al., 2008; Smith and Schooley, 2008).

Various efforts demonstrated that screening for HIV RNA among people presenting for HIV testing or blood donation can be used efficiently to identify individuals who were acutely infected with HIV despite a negative HIV antibody test, since individuals testing during the window period between acute HIV infection and seroconversion will escape routine detection (Busch et al., 2005; Patterson et al., 2007; Pilcher et al., 2005; Pilcher et al., 2002). Because testing for HIV RNA in each blood sample would be expensive, a commonly used strategy is to pool blood samples from a group of individuals and perform one HIV RNA assay on a pooled sample (Busch, et al., 2005; Patterson, et al., 2007; Pilcher, et al., 2005; Pilcher, et al., 2002). If the pool tests positive, individual samples or pools of smaller size might be tested again to identify affected individuals. Performance and advantages of different pooling strategies for screening purposes have been extensively investigated. In a landmark paper Dorfman (1943) calculated the optimal pool size for a prevalence ranging from 1% to 30% for a single stage pooling algorithm. Hammick and Gastwirth (1994) suggest to obtain two samples per individual an form two

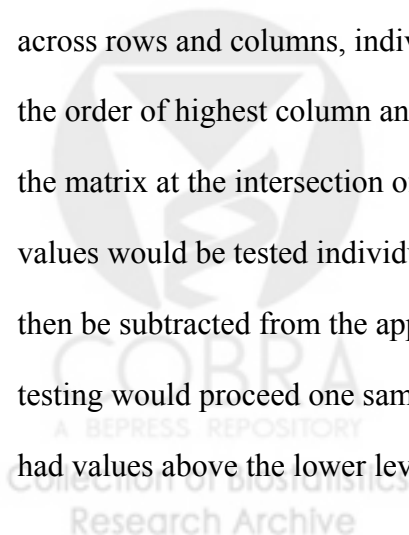


independent sets of groupings to estimate the prevalence of HIV while preserving confidentiality. An advantage of this procedure is that it can be efficient for a prevalence of up to about 30%. Brookmeyer (1999) developed an estimator of disease prevalence and its variance for multistage pooling studies. Westreich et al (2008) investigate the performance of three different pooling strategies (two hierarchical and one matrix based approach) to acute HIV infection. Kim et al. (2007) derive and compare operating characteristics of hierarchical as well as square array-based testing algorithms in the presence of testing error. For additional references see (Brookmeyer, 1999; Hammick and Gastwirth, 1994; Kim et al., 2007; Westreich et al., 2008)}. Nevertheless, all of the existing work is based on settings where the outcome is binary (disease versus no disease). Also, the testing for HIV RNA in blood samples pooled from patients receiving ART to identify instances when ART is failing to suppress viral replication is considerably different from nucleic acid testing on pooled blood samples to identify instances of acute HIV infection. Specifically, more efficient algorithms are needed because of the higher prevalence of instances of virologic failure, the inherent variability of the viral load assay, and the need to identify lower viral load levels. In this report, we demonstrate how screening for HIV RNA among pooled blood samples can be used to monitor for active HIV replication during ART in order to identify individuals who are experiencing virologic failure. These results could have profound impact on the ability to virologically monitor HIV-infected patients receiving ART in resource constrained settings and ultimately limit the development and transmission of drug resistant HIV strains.

## 2. METHODS

Simulations were used to generate sample values of viral loads that represent a population of patients on ART. The distribution of viral load values of the simulated population was based on a natural history cohort of HIV infected individuals. Briefly, for viral load values below 500 copies/mL, 85%, 5% and 10% of values were generated to be uniformly distributed below 50 copies/mL, between 50 and 100 copies/mL and between 100 and 500 copies/mL, respectively. For values above 500 copies/mL, simulated viral load values were based on two gamma distributions with modes at about 1000 and 6000 copies/mL (on  $\log_{10}$  scale the gamma location parameter was 2.7, scale parameter was 0.5 and shape parameter was 1.6 for the first and 3.18 for the second gamma distribution). The proportion of values generated for the first and second gamma distribution were 93% and 7% respectively. Both individual and pooled values were generated to include measurement error from viral load assay variability based on normally distributed measurement error with mean zero. Assay standard deviation was assumed to be zero, 0.12 and 0.20 on the  $\log_{10}$  scale. The value of zero represented a best case scenario, the value of 0.12 was based on published data (Brambilla et al., 1999) and 0.20 represented a more conservative alternative (Jagodzinski et al., 2000). Threshold values to define virologic failure during ART of 500, 1000 or 1500 copies/mL were used based on clinically meaningful thresholds at which drug resistance testing could be performed and decisions to amend therapy might be expected (Harrigan et al., 2005; Haupts et al., 2003; Hirsch et al., 2003).

Three algorithms were evaluated (Figures 1 and 2). The first was based on Dorfman's (Dorfman, 1943) two-stage "minipool" approach where a fixed number of samples would be combined into one pool. If the pooled samples yielded an assay value above the lower limit of interest for the pool (e.g.  $\geq 50$  copies/mL), then all individual samples would be tested in the second stage. Samples of pools with assay values below the lower limit of interest would be considered virologically suppressed and no further tests would be performed. For the second algorithm, minipools would be formed in the same manner as for Dorfman's approach; however, if the pooled samples yielded a viral load value above the lower limit of interest, then individual samples would be tested consecutively (in random order) and subtracted from the pooled sample estimate until the lower limit of interest for the pool was achieved. We call this approach "minipool + algorithm". The third approach applied a similar algorithm to a  $k$  by  $k$  matrix. For this approach, samples would be organized in a matrix structure and pools formed over  $k$  rows and  $k$  columns. Samples of pools would be considered virologically suppressed if either the row or column pools yielded values below the threshold of interest. After obtaining  $(k + k)$  values for pools across rows and columns, individual samples would be consecutively tested based on the order of highest column and row value, i.e. the individual sample that resides in the matrix at the intersection of the row and column pools with the highest viral load values would be tested individually. Viral load values from the individual test would then be subtracted from the appropriate column and row pool totals, and individual testing would proceed one sample at a time until no further row and column pools had values above the lower level of interest. If a single row (or column) showed a

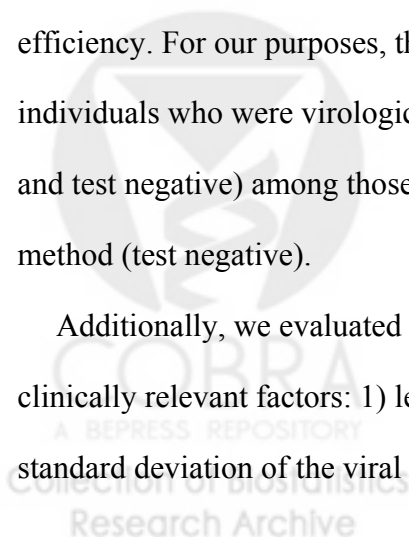


value above the lower limit of interest, but none of the columns (or rows) did, all samples from such a row (or column) would be considered virologically suppressed.

Based on various prevalences of and thresholds for defining virologic failure, we compared relative efficiency and negative predictive value for four approaches (individual samples, minipool, minipool + algorithm and matrix approach) and various pool sizes within each of the pooling approaches (minipool, minipool + algorithm and matrix approaches). Relative efficiency was defined as one minus the average number of assays performed divided by the number of samples. We use a different definition of efficiency than others (Westreich, et al., 2008) because of the natural interpretation that the best algorithm will have the highest values of efficiency relative to individual testing. With this definition, an algorithm with efficiency zero has no advantage over individual testing and an algorithm with efficiency of 0.25 uses 25% fewer samples than would be used with individual testing.

Negative predictive values were compared for the various approaches to identify individual samples that had viral load values below the same range of thresholds (500, 1000 and 1500 HIV RNA copies/mL) similar to the comparisons of relative efficiency. For our purposes, the negative predictive value was the proportion of individuals who were virologically suppressed below the threshold of interest (true and test negative) among those who were considered suppressed by each proposed method (test negative).

Additionally, we evaluated the various approaches in the context of the following clinically relevant factors: 1) level of viral load for defining virologic failure, 2) standard deviation of the viral load assay and 3) prevalence of virologic failure in the





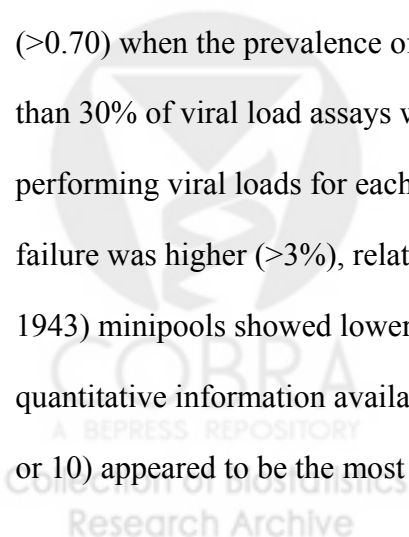
sampled population. Screening approaches for higher prevalence of virologic failure and assays with larger standard deviation were expected to be less efficient than for lower prevalence of virologic failure or assays with smaller standard deviation.

Pool sizes of 3, 4, 5, 6, 7, 8, 10 and 20 were considered. Since the lowest possible level of viral load detection depends on the dilution factor of the pool, the level of detection was considered in relation to the size of the pool. For example, with a lower detection limit for a viral load assay of 50 copies/mL and a pool size of 10, the on average lowest level of virologic failure that we can expect to detect for any individual sample based on the pooled analysis was 500 copies/mL. When pools of 20 individual samples were considered, the lowest detectable level of virologic failure was 1000 copies/mL. All simulations were run using the statistical software Stata (StataCorp, 2007) and were repeated one thousand times for all conditions.

### 3. RESULTS

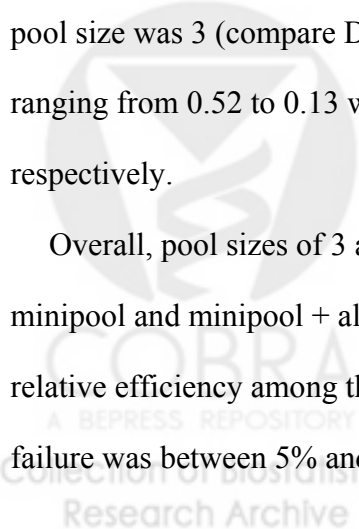
#### *3.1 Relative efficiency*

Of the three pooling approaches, most demonstrated relatively high efficiency ( $>0.70$ ) when the prevalence of virologic failure was low ( $<3\%$ ) (Figure 3), i.e. less than 30% of viral load assays would be used in a pooling approach relative to performing viral loads for each individual sample. When the prevalence of virologic failure was higher ( $>3\%$ ), relative efficiency varied markedly, Dorfman's (Dorfman, 1943) minipools showed lower relative efficiency than algorithms that used the quantitative information available. A matrix approach with intermediate pool size (8 or 10) appeared to be the most efficient approach when: 1) virologic failure was



defined as  $\geq 1500$  copies/mL, 2) the standard deviation of the assay was 0.12 and 3) the prevalence of virologic failure of the sampled population was between 5% and 20%. Specifically, the relative efficiency for the matrix approach with a pool size of 10 ranged from 0.68 to 0.33 and for a matrix approach with a pool size of 8 it ranged from 0.65 to 0.32 when the prevalence of virologic failure was 5% and 20%, respectively. Among the approaches that use the minipools + algorithm, the one with 5 samples per pool appeared most efficient under similar conditions with relative efficiency from 0.63 to 0.26 for a prevalence of virologic failure of 5% and 20%, respectively. Thus, the relative efficiency of the 5 minipool+algorithm was close (no more than 0.07 difference) to the most efficient matrix approaches. The relative efficiency over the various approaches changed only slightly when the standard deviation of the assay was 0.20. For example, when the prevalence of virologic failure was 5% in the sampled population, the relative efficiency for the matrix approach with a pool size of 10 was: 1) 0.66 when the standard deviation of the viral load assay was 0.20, 2) 0.68 when the standard deviation was 0.12, and 3) 0.73 when the standard deviation was zero. For Dorfman's minipool approach, the most efficient pool size was 3 (compare Dorfman (Dorfman, 1943), Table 1) with relative efficiency ranging from 0.52 to 0.13 when the prevalence of virologic failure was 5% and 20%, respectively.

Overall, pool sizes of 3 and 4 had the highest relative efficiency among the minipool and minipool + algorithm approaches and a pool size of 10 had the highest relative efficiency among the matrix approaches when the prevalence of virologic failure was between 5% and 20% (Figure 4). Results were not substantially different



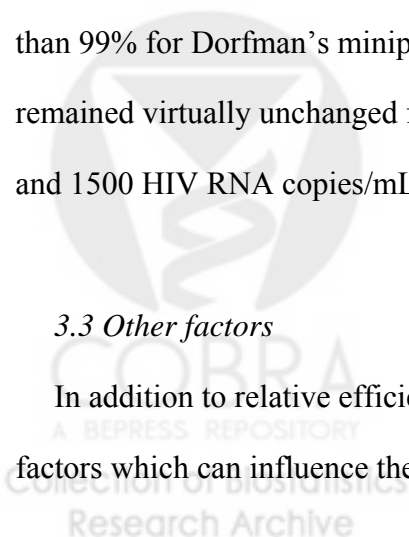
when virologic failure was defined as  $\geq 500$ ,  $\geq 1000$  copies/mL or  $\geq 1500$  copies/mL. For example, when using the matrix approach with a pool size of 10, assay SD of 0.12 and a prevalence of virologic failure of 20%, the relative efficiency was 0.68 when virologic failure was defined at  $\geq 1500$  copies/mL, 0.66 at  $\geq 1000$  copies/mL and 0.65 at 500 copies/mL.

### *3.2 Negative predictive values*

The results for negative predictive values were uniformly good. For a prevalence of virologic failure up to 10%, negative predictive values were at least 99% and 97% for the matrix approach with assay standard deviations of 0.12 and 0.20 copies/mL respectively. Similarly, the corresponding negative predictive values were 99% and 98% for all minipool + algorithm approaches. For a prevalence of virologic failure up to 20%, negative predictive values were always at least 92% for all matrix approaches and always above 96% for the minipool + algorithm approach when the standard deviation of the assay was 0.12. These results changed only minimally for an assay standard deviation of 0.20. Negative predictive values were always higher than 99% for Dorfman's minipool approach (without algorithm). These results remained virtually unchanged for different definitions of virologic failure (500, 1000 and 1500 HIV RNA copies/mL).

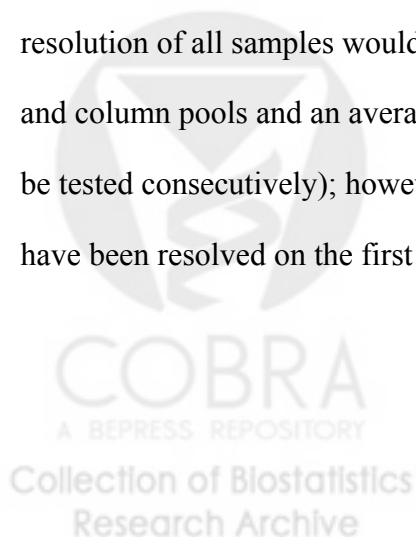
### *3.3 Other factors*

In addition to relative efficiency and negative predictive value, there are other factors which can influence the clinical utility of these approaches. One is the local



need of turnaround time of the results, which will depend on availability of the assays, personnel performing the assays and size of the clinical population receiving ART. The proposed pooling strategies will vary in the time from sample collection to the time individual viral load results are available to the clinician for management of their patients. For example, if a laboratory has the capability of performing 20 viral loads a day, then 20 individuals could be screened daily using individual viral load testing and it would take 5 days to screen 100 individuals. If minipools of size 5 are used, then 100 individuals could be screened in the first day, but some pools would require additional testing and, therefore, additional days would be required before results are available. The number of additional days would depend on the prevalence of virologic failure in the sampled population and the specific pooling approach.

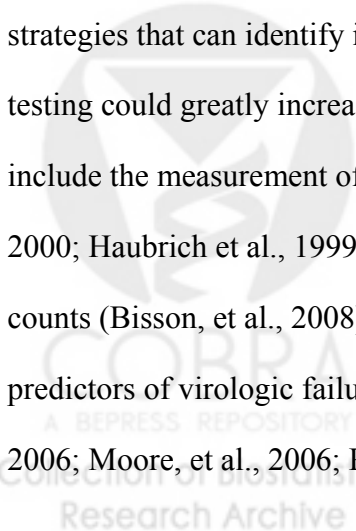
When 100 patients are screened using minipools of size 5 + algorithm and virologic failure is defined as  $\geq 1500$  HIV RNA copies/mL and the prevalence of virologic failure is 10%, there would be an average of 2 days until the viral loads of all samples have been resolved. Similarly, the initial screening of the same 100 patients using the matrix approach with a pool size of 10 would require one day, but the complete resolution of all samples would require on average 28 days (1 day for testing all row and column pools and an average of about 27 days for 27 individual assays that must be tested consecutively); however, on average 65% of the individual samples would have been resolved on the first day.



#### 4. DISCUSSION

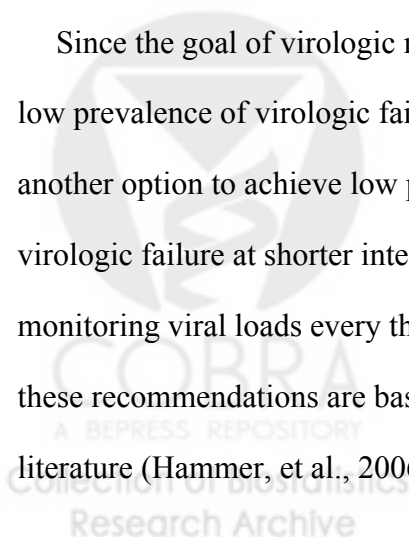
Because commercial viral load assays cost between US \$50.00–\$150.00 per test [Clinton Foundation. Diagnostics pricing. Available at: <http://www.clintonfoundation.org/cf-pgm-hs-ai-work3.htm>]; Accessed May 23, 2008. (Fiscus, et al., 2006)] the costs of virologic monitoring during ART may exceed those of ART itself, but the development and transmission of drug resistant HIV may ultimately compromise the effectiveness of ART in many populations (Harrigan, et al., 2005; Haupts, et al., 2003; Little et al., 2002; Miller and Larder, 2001; Vijayaraghavan et al., 2007). In resource-constrained settings less expensive methods to monitor viral replication during ART are necessary to make such monitoring feasible and ART sustainable. In this report, we demonstrate how nucleic acid testing on pooled blood samples can be used to reduce the overall number of viral load tests needed to screen patients receiving ART compared to individual testing.

Since the prevalence of virologic failure in the sampled population will impact the relative efficiency and accuracy of the presented pooling and testing methods, other strategies that can identify individuals with virologic failure before nucleic acid testing could greatly increase the usefulness of the proposed methods. Such strategies include the measurement of adherence to ART (Bisson, et al., 2008; Gifford et al., 2000; Haubrich et al., 1999; Paterson et al., 2000) or longitudinal trajectory of CD4 counts (Bisson, et al., 2008); however, these methods are by no means perfect predictors of virologic failure (Bisson, et al., 2008; Bisson, et al., 2006; Fiscus, et al., 2006; Moore, et al., 2006; Petti, et al., 2006); and will need to be evaluated in



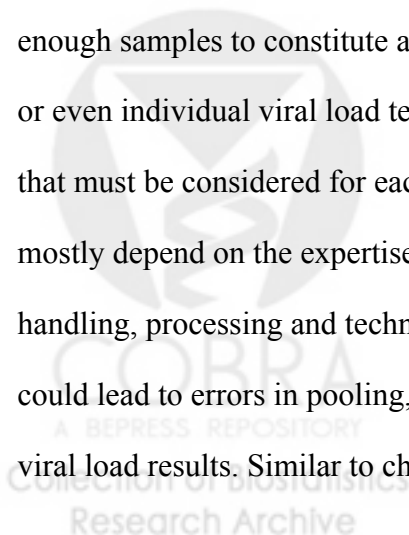
individual clinical settings. Additionally, the definition of virologic failure is an obvious factor in determining the prevalence of virologic failure in a population. Although some research considers the consequences of different definitions of virologic failure (Macias et al., 2005; Raboud et al., 1999; Raboud et al., 1998), it remains unclear which level of HIV RNA viral load constitutes the most important clinical cutpoint to define virologic failure. Differences regarding negative predictive values between different definitions of virologic failure (500, 1000 and 1500 HIV RNA copies/mL) appear to be very small ( $\leq 2\%$ ) and reasonably small ( $\leq 7\%$ ) for prevalence of virologic failure up to 10% and 20% respectively. Differences in relative efficiency for different definitions of virologic failure (500, 1000 and 1500 HIV RNA copies/mL) were also small. In essence, results regarding relative efficiency and negative predictive value did not depend on the definition of virologic failure. It might be possible to further improve the efficiency of the matrix approach (e.g. choosing and individual testing of the cell with largest column and row pool; subtracting individual test values from column and row estimates might not be optimal).

Since the goal of virologic monitoring during ART, is to achieve and maintain a low prevalence of virologic failure in the population of patients receiving ART, another option to achieve low prevalence would be to monitor the population for virologic failure at shorter intervals. Recent HIV treatment guidelines recommend monitoring viral loads every three to four months (Hammer, et al., 2006); however, these recommendations are based solely on expert's interpretation of the published literature (Hammer, et al., 2006), and more frequent monitoring has been proposed



based on prospective clinical trial data (Haubrich, et al., 2001). More likely, the frequency of virologic monitoring needed to optimize clinical outcomes will vary by clinical setting, which would include factors such as the potency, tolerability and durability of the available ART regimens, support available for patient adherence and prevalence of transmitted drug resistance. Taken together, the methods used to monitor for virologic failure during ART will need to be evaluated based on the needs of each clinical setting.

Similarly, each clinical program will most likely have unique requirements for assay characteristics (overall costs, turnaround time of results and monitoring accuracy). For example, in some areas the overall cost of performing a viral load may be the greatest limiting factor for monitoring for virologic failure; therefore, a clinical program in this setting might best be served by a method that has the highest relative efficiency for the assays, independent of the turnaround time of individual results or maximal accuracy. Alternatively, other programs might require more rapid turnaround times for obtaining viral load results, such as programs with smaller patient populations receiving ART. These settings would require more time to obtain enough samples to constitute a pool, and more frequent testing or smaller pool sizes or even individual viral load testing might be the most cost-effective. Another factor that must be considered for each clinical setting is quality assurance, which will mostly depend on the expertise of the personnel performing the assays (sample handling, processing and technical consistency with the viral load assay), which could lead to errors in pooling, resolution testing and inconsistency in calculating viral load results. Similar to choosing a method to screen for virologic failure, the



extent and nature of technical training required to perform various pooling approaches and maximize quality assurance will need to be evaluated in each local laboratory and clinical setting.

We present characteristics for a variety of algorithms and pool sizes that can incorporate available quantitative viral load data and data on how pooling methods can be used efficiently and accurately to monitor for virologic failure in patients receiving ART. These data could be used by individual laboratory and clinical settings to make choices about optimal local virologic monitoring strategies. They may also be used to design efficient pooling strategies for other settings where screening involves quantitative measures. Although promising, further investigation in resource-constrained settings is required to determine if these methods are feasible and cost-effective with respect to the factors that could not be included in the simulations like turnaround time of results, additional personnel costs in local settings, and individual and public health costs of a patient population with virologic failure during ART.





AUTHOR CONTRIBUTIONS, ACKNOWLEDGEMENTS AND FUNDING

SM and DMS designed the study and prepared the manuscript. SM implemented the simulations. AG, RH and CB contributed substantially to the design of the study and manuscript preparation. All authors read and approved the final manuscript. We would like to thank Drs. Susan Little, Robert Schooley, and Matthew C. Strain for insightful comments.

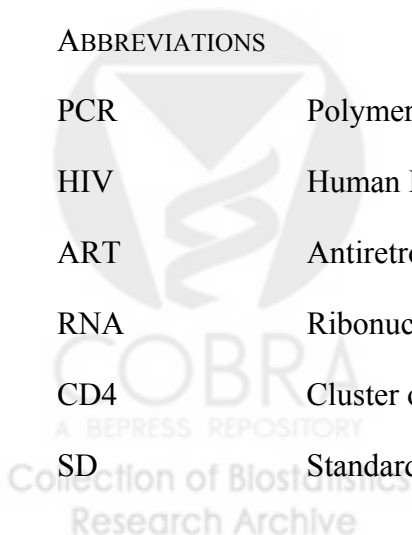
This work was supported by Grants AI27670, AI43638, the UCSD Center for AIDS Research AI 36214, AI29164, AI47745, AI57167, AI55276, K24-AI064086, (CHRP) CH05-SD-607-005, MH62512 from the National Institutes of Health, and the San Diego Veterans Affairs Healthcare System. The funding agencies had no role in the study design, implementation or interpretation of the results; manuscript preparation; or decision to submit the work for publication.

COMPETING INTERESTS

None of the authors have competing interests regarding the research presented.

ABBREVIATIONS

PCR	Polymerase Chain Reaction
HIV	Human Immunodeficiency Virus
ART	Antiretroviral Therapy
RNA	Ribonucleic Acid
CD4	Cluster of Differentiation 4
SD	Standard Deviation



## REFERENCES

- BISSON, G.P., GROSS, R., BELLAMY, S., CHITTAMS, J., HISLOP, M., REGENSBERG, L., FRANK, I., MAARTENS, G. AND NACHEGA, J.B. 2008. Pharmacy Refill Adherence Compared with CD4 Count Changes for Monitoring HIV-Infected Adults on Antiretroviral Therapy. *PLoS Med* **5**, e109.
- BISSON, G.P., GROSS, R., STROM, J.B., ROLLINS, C., BELLAMY, S., WEINSTEIN, R., FRIEDMAN, H., DICKINSON, D., FRANK, I., STROM, B.L., GAOLATHE, T. AND NDWAPI, N. 2006. Diagnostic accuracy of CD4 cell count increase for virologic response after initiating highly active antiretroviral therapy. *AIDS* **20**, 1613-1619.
- BRAMBILLA, D., REICHELDERFER, P.S., BREMER, J.W., SHAPIRO, D.E., HERSHOW, R.C., KATZENSTEIN, D.A., HAMMER, S.M., JACKSON, B., COLLIER, A.C., SPERLING, R.S., FOWLER, M.G. AND COOMBS, R.W. 1999. The contribution of assay variation and biological variation to the total variability of plasma HIV-1 RNA measurements. The Women Infant Transmission Study Clinics. Virology Quality Assurance Program. *AIDS* **13**, 2269-2279.
- BROOKMEYER, R. 1999. Analysis of multistage pooling studies of biological specimens for estimating disease incidence and prevalence. *Biometrics* **55**, 608-612.
- BUSCH, M.P., GLYNN, S.A., STRAMER, S.L., STRONG, D.M., CAGLIOTI, S., WRIGHT, D.J., PAPPALARDO, B. AND KLEINMAN, S.H. 2005. A new strategy for estimating risks of transfusion-transmitted viral infections based on rates of detection of recently infected donors. *Transfusion* **45**, 254-264.
- CALMY, A., FORD, N., HIRSCHL, B., REYNOLDS, S.J., LYNEN, L., GOEMAERE, E., GARCIA DE LA VEGA, F., PERRIN, L. AND RODRIGUEZ, W. 2007. HIV viral load monitoring in resource-limited regions: optional or necessary? *Clin Infect Dis* **44**, 128-134.
- DORFMAN, R. 1943. The detection of defective members of large populations. *Annals of Mathematical Statistics* **14**, 436-440.
- DURANT, J., CLEVENBERGH, P., HALFON, P., DELGIUDICE, P., PORSIN, S., SIMONET, P., MONTAGNE, N., BOUCHER, C.A.B., SCHAPIRO, J.M. AND DELLAMONICA, P. 1999. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet* **353**, 2195-2199.
- FISCUS, S.A., CHENG, B., CROWE, S.M., DEMETER, L., JENNINGS, C., MILLER, V., RESPESS, R. AND STEVENS, W. 2006. HIV-1 viral load assays for resource-limited settings. *PLoS Med* **3**, e417.

GIFFORD, A.L., BORMANN, J.E., SHIVELY, M.J., WRIGHT, B.C., RICHMAN, D.D. AND BOZZETTE, S.A. 2000. Predictors of self-reported adherence and plasma HIV concentrations in patients on multidrug antiretroviral regimens. *J Acquir Immune Defic Syndr* **23**, 386-395.

HAMMER, S.M., SAAG, M.S., SCHECHTER, M., MONTANER, J.S., SCHOOLEY, R.T., JACOBSEN, D.M., THOMPSON, M.A., CARPENTER, C.C., FISCHL, M.A., GAZZARD, B.G., GATELL, J.M., HIRSCH, M.S., KATZENSTEIN, D.A., RICHMAN, D.D., VELLA, S., YENI, P.G. AND VOLBERDING, P.A. 2006. Treatment for adult HIV infection: 2006 recommendations of the International AIDS Society-USA panel. *Jama* **296**, 827-843.

HAMMICK, P.A. AND GASTWIRTH, J.L. 1994. Group-testing for sensitive characteristics: extension to higher prevalence levels. *International Statistical Review* **62**, 319-331.

HARRIGAN, P.R., HOGG, R.S., DONG, W.W., YIP, B., WYNHOVEN, B., WOODWARD, J., BRUMME, C.J., BRUMME, Z.L., MO, T., ALEXANDER, C.S. AND MONTANER, J.S. 2005. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. *J Infect Dis* **191**, 339-347.

HAUBRICH, R.H., CURRIER, J.S., FORTHAL, D.N., BEALL, G., KEMPER, C.A., JOHNSON, D., DUBE, M.P., HWANG, J., LEEDOM, J.M., TILLES, J. AND MCCUTCHAN, J.A. 2001. A randomized study of the utility of human immunodeficiency virus RNA measurement for the management of antiretroviral therapy. *Clin Infect Dis* **33**, 1060-1068.

HAUBRICH, R.H., LITTLE, S.J., CURRIER, J.S., FORTHAL, D.N., KEMPER, C.A., BEALL, G.N., JOHNSON, D., DUBE, M.P., HWANG, J.Y. AND MCCUTCHAN, J.A. 1999. The value of patient-reported adherence to antiretroviral therapy in predicting virologic and immunologic response. California Collaborative Treatment Group. *AIDS* **13**, 1099-1107.

HAUPTS, S., LEDERGERBER, B., BONI, J., SCHUPBACH, J., KRONENBERG, A., OPRAVIL, M., FLEPP, M., SPECK, R.F., GRUBE, C., RENTSCH, K., WEBER, R. AND GUNTARD, H.F. 2003. Impact of genotypic resistance testing on selection of salvage regimen in clinical practice. *Antivir Ther* **8**, 443-454.

HIRSCH, M.S., BRUN-VEZINET, F., CLOTET, B., CONWAY, B., KURITZKES, D.R., D'AQUILA, R.T., DEMETER, L.M., HAMMER, S.M., JOHNSON, V.A., LOVEDAY, C., MELLORS, J.W., JACOBSEN, D.M. AND RICHMAN, D.D. 2003. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. *Clin Infect Dis* **37**, 113-128.

HUGHES, M.D., JOHNSON, V.A., HIRSCH, M.S., BREMER, J.W., ELBEIK, T., ERICE, A., KURITZKES, D.R., SCOTT, W.A., SPECTOR, S.A., BASGOZ, N., FISCHL, M.A. AND D'AQUILA, R.T. 1997. Monitoring plasma HIV-1 RNA levels in addition to CD4+ lymphocyte count improves assessment of antiretroviral therapeutic response. ACTG 241 Protocol Virology Substudy Team. *Ann Intern Med* **126**, 929-938.

JAGODZINSKI, L.L., WIGGINS, D.L., MCMANIS, J.L., EMERY, S., OVERBAUGH, J., ROBB, M., BODRUG, S. AND MICHAEL, N.L. 2000. Use of calibrated viral load standards for group M subtypes of human immunodeficiency virus type 1 to assess the performance of viral RNA quantitation tests. *J Clin Microbiol* **38**, 1247-1249.

KIM, H.Y., HUDGENS, M.G., DREYFUSS, J.M., WESTREICH, D.J. AND PILCHER, C.D. 2007. Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics* **63**, 1152-1163.

LITTLE, S.J., HOLTE, S., ROUTY, J.P., DAAR, E.S., MARKOWITZ, M., COLLIER, A.C., KROUP, R.A., MELLORS, J.W., CONNICK, E., CONWAY, B., KILBY, M., WANG, L., WHITCOMB, J.M., HELLMANN, N.S. AND RICHMAN, D.D. 2002. Antiretroviral-drug resistance among patients recently infected with HIV. *N Engl J Med* **347**, 385-394.

MACIAS, J., PALOMARES, J.C., MIRA, J.A., TORRES, M.J., GARCIA-GARCIA, J.A., RODRIQUEZ, J.M., VERGERA, S. AND PINEDA, J.A. 2005. Transient rebounds of HIV plasma viremia are associated with the emergence of drug resistance mutations in patients on highly active antiretroviral therapy. *J Infect* **51**, 195-200.

MILLER, V. AND LARDER, B.A. 2001. Mutational patterns in the HIV genome and cross-resistance following nucleoside and nucleotide analogue drug exposure. *Antivir Ther* **6 Suppl 3**, 25-44.

MOORE, D.M., MERMIN, J., AWOR, A., YIP, B., HOGG, R.S. AND MONTANER, J.S. 2006. Performance of immunologic responses in predicting viral load suppression: implications for monitoring patients in resource-limited settings. *J Acquir Immune Defic Syndr* **43**, 436-439.

PATERSON, D.L., SWINDELLS, S., MOHR, J., BRESTER, M., VERGIS, E.N., SQUIER, C., WAGENER, M.M. AND SINGH, N. 2000. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med* **133**, 21-30.

PATTERSON, K.B., LEONE, P.A., FISCUS, S.A., KURUC, J., MCCOY, S.I., WOLF, L., FOUST, E., WILLIAMS, D., ERON, J.J. AND PILCHER, C.D. 2007. Frequent detection of acute HIV infection in pregnant women. *AIDS* **21**, 2303-2308.

PETTI, C.A., POLAGE, C.R., QUINN, T.C., RONALD, A.R. AND SANDE, M.A. 2006. Laboratory medicine in Africa: a barrier to effective health care. *Clin Infect Dis* **42**, 377-382.

PHILLIPS, A.N., PILLAY, D., MINERS, A.H., BENNETT, D.E., GILKS, C.F. AND LUNDGREN, J.D. 2008. Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. *Lancet* **371**, 1443-1451.

PILCHER, C.D., FISCUS, S.A., NGUYEN, T.Q., FOUST, E., WOLF, L., WILLIAMS, D., ASHBY, R., O'DOWD, J.O., MCPHERSON, J.T., STALZER, B., HIGHTOW, L., MILLER, W.C., ERON, J.J., JR., COHEN, M.S. AND LEONE, P.A. 2005. Detection of acute infections during HIV testing in North Carolina. *N Engl J Med* **352**, 1873-1883.

PILCHER, C.D., MCPHERSON, J.T., LEONE, P.A., SMURZYNSKI, M., OWEN-O'DOWD, J., PEACE-BREWER, A.L., HARRIS, J., HICKS, C.B., ERON, J.J., JR. AND FISCUS, S.A. 2002. Real-time, universal screening for acute HIV infection in a routine HIV counseling and testing population. *Jama* **288**, 216-221.

RABOUD, J.M., RAE, S., HOGG, R.S., YIP, B., SHERLOCK, C.H., HARRIGAN, P.R., O'SHAUGHNESSY, M.V. AND MONTANER, J.S. 1999. Suppression of plasma virus load below the detection limit of a human immunodeficiency virus kit is associated with longer virologic response than suppression below the limit of quantitation. *J Infect Dis* **180**, 1347-1350.

RABOUD, J.M., SEMINARI, E., RAE, S.L., HARRIGAN, P.R., HOGG, R.S., CONWAY, B., SHERLOCK, C., SCHECHTER, M.T., O'SHAUGHNESSY, M.V. AND MONTANER, J.S. 1998. Comparison of costs of strategies for measuring levels of human immunodeficiency virus type 1 RNA in plasma by using Amplicor and Ultra Direct assays. *J Clin Microbiol* **36**, 3369-3371.

SMITH, D.M. AND SCHOOLEY, R.T. 2008. Running with scissors: using antiretroviral therapy without monitoring viral load. *Clin Infect Dis* **46**, 1598-1600.

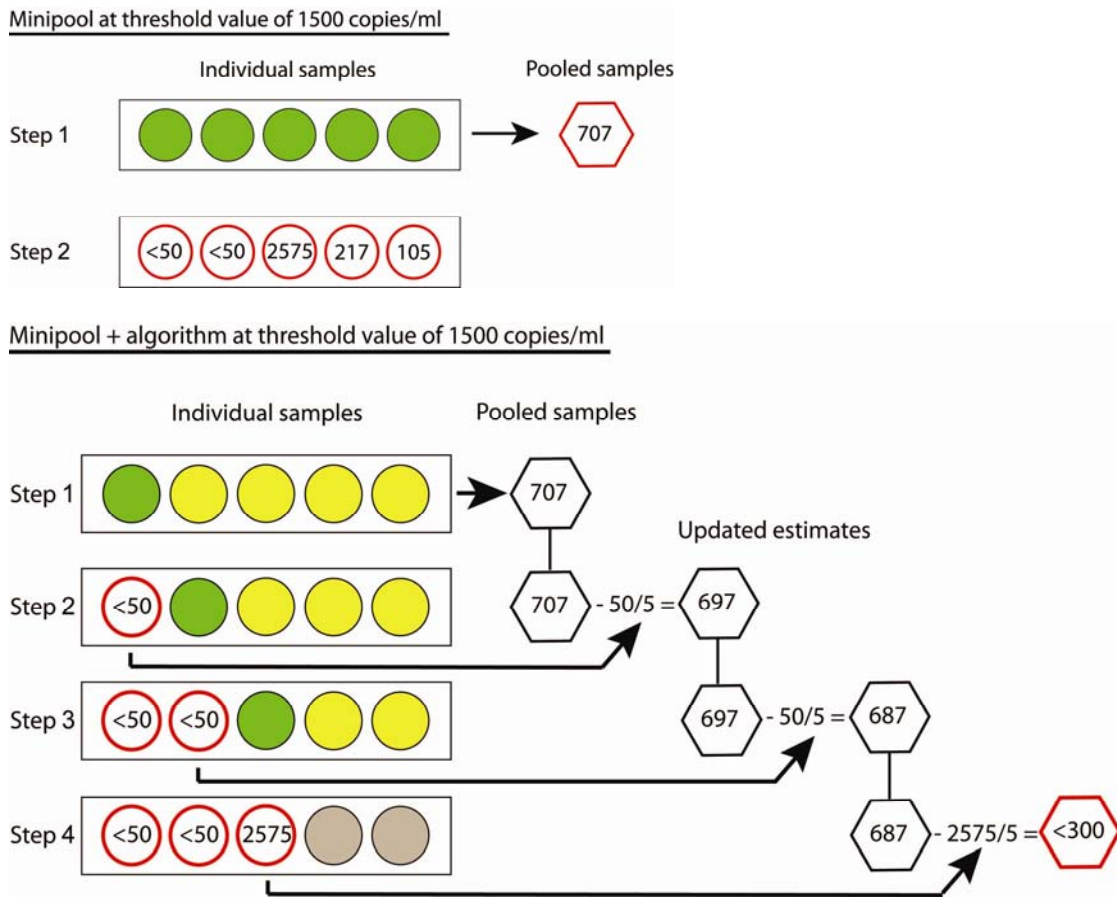
STACORP 2007. *Stata Statistical Software: Release 10*. StataCorp, College Station, TX.

VIJAYARAGHAVAN, A., EFRUSY, M.B., MAZONSON, P.D., EBRAHIM, O., SANNE, I.M. AND SANTAS, C.C. 2007. Cost-effectiveness of alternative strategies

for initiating and monitoring highly active antiretroviral therapy in the developing world. *J Acquir Immune Defic Syndr* **46**, 91-100.

WESTREICH, D.J., HUDGENS, M.G., FISCUS, S.A. AND PILCHER, C.D. 2008. Optimizing screening for acute HIV infection with pooled nucleic acid amplification tests. *J Clin Microbiol*.





**Figure 1:** Schematic overview of minipool and minipool + algorithm approaches.

Each circle represents an individual sample, each hexagon a pooled sample. The numbers within the non-shaded circles are viral load values measured in the individual samples. Green circles indicate individual samples that will have viral loads measured in the next step. Yellow circles indicate samples that might have viral loads measured depending on the outcome of the measures for the green circles.

In step 1 for the minipool approach (A), individual samples (green circles) are pooled (hexagon) and a viral load measured in the pooled sample. In this example, since the pooled sample is greater than the threshold of interest (here  $\leq 300$  copies/mL

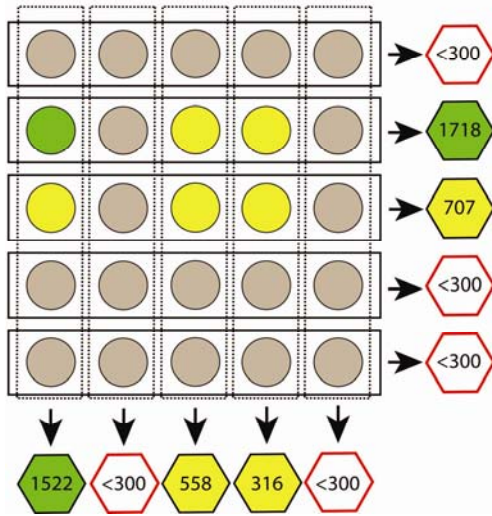
because there are 5 individual samples in a pool), then all individual samples are tested individually in step 2.

In step 1 for the minipool + algorithm approach (B), individual samples (green and yellow circles) are pooled (hexagon) and a viral load measured in the pooled sample. In steps 2-4, since the pooled sample is greater than the threshold of interest, individual samples of pools with values above the threshold are consecutively tested and individual estimates (divided by pool size) are subtracted from pool total (hexagon) until the threshold value is reached (red hexagon). Remaining samples (grey circles) that have not been tested individually are assumed to be below threshold (here  $\leq 1500$  copies/mL). Note, estimated values for individual samples will typically not add up to estimated values for pooled samples because of assay variability.

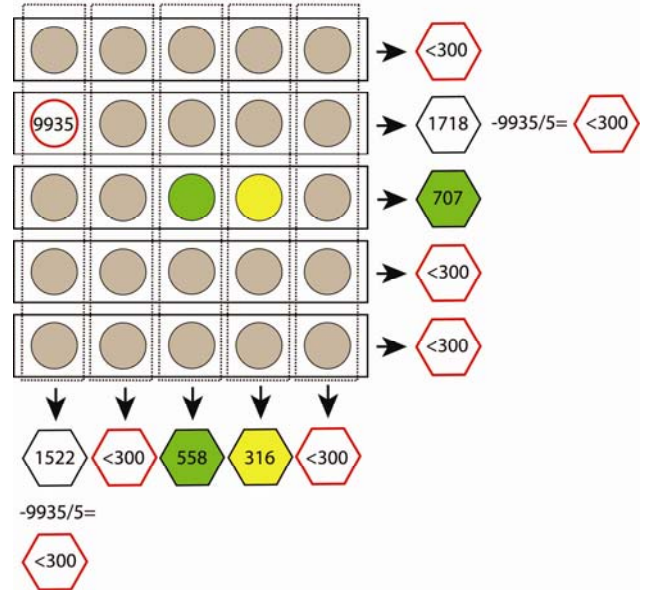




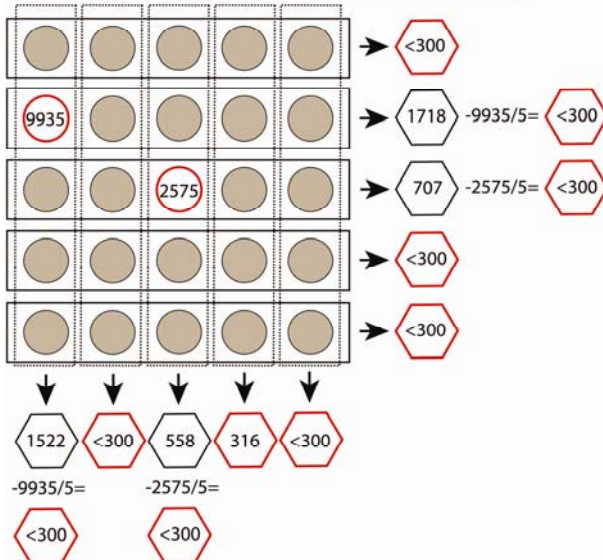
Matrix algorithm - Step 1 at threshold value of 1500 copies/ml



Matrix algorithm - Step 2 at threshold value of 1500 copies/ml

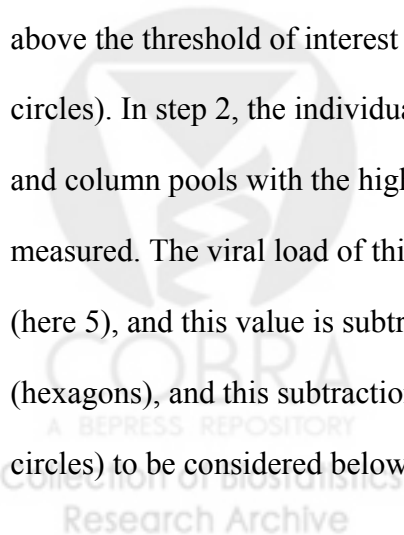


Matrix algorithm - Step 3 at threshold value of 1500 copies/ml



**Figure 2:** Schematic overview of matrix approach. Each circle represents an individual sample, each hexagon a pooled sample. The numbers within the non-shaded circles are viral load values measured on individual samples. Viral load values for grey circles are assumed to be below the threshold (here  $\leq 1500$  copies/mL). Green circles indicate samples that will be measured in the next step. Yellow circles indicate samples that might be measured depending on the outcome of the measures for the individual samples marked by the green circles. Red circles and hexagons indicate that no further testing/estimation will be performed for the sample or pool. Note, estimated values for individual samples will typically not add up to estimated values for pooled samples because of assay variability.

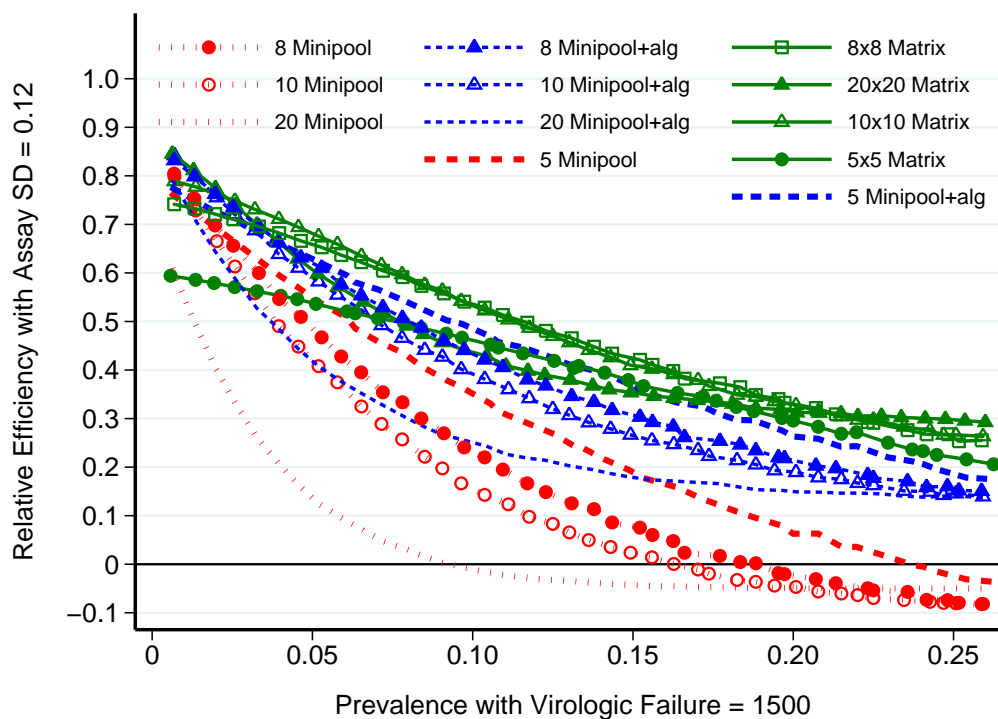
In step 1 of the matrix approach, individual samples (circles) are pooled across rows and columns (hexagons), and viral loads are then measured in each pool. All individual samples are considered to be below the threshold level (grey circle) if their respective row *or* column pool has a viral load below the threshold level of interest. Individual samples that belong to a row *and* a column pool that have viral loads above the threshold remain ambiguous as to whether or not they contain viral loads above the threshold of interest and might be tested individually (green and yellow circles). In step 2, the individual sample that belongs to the intersection of the row and column pools with the highest viral load values (green circle) has its viral load measured. The viral load of this individual sample is divided by the size of the pools (here 5), and this value is subtracted from the respective row and column pool totals (hexagons), and this subtraction can resolve previously ambiguous samples (yellow circles) to be considered below the viral load level of interest (grey circles). In step 3,



this continues until all individual samples belong to a row or column pool that has reached the threshold level.

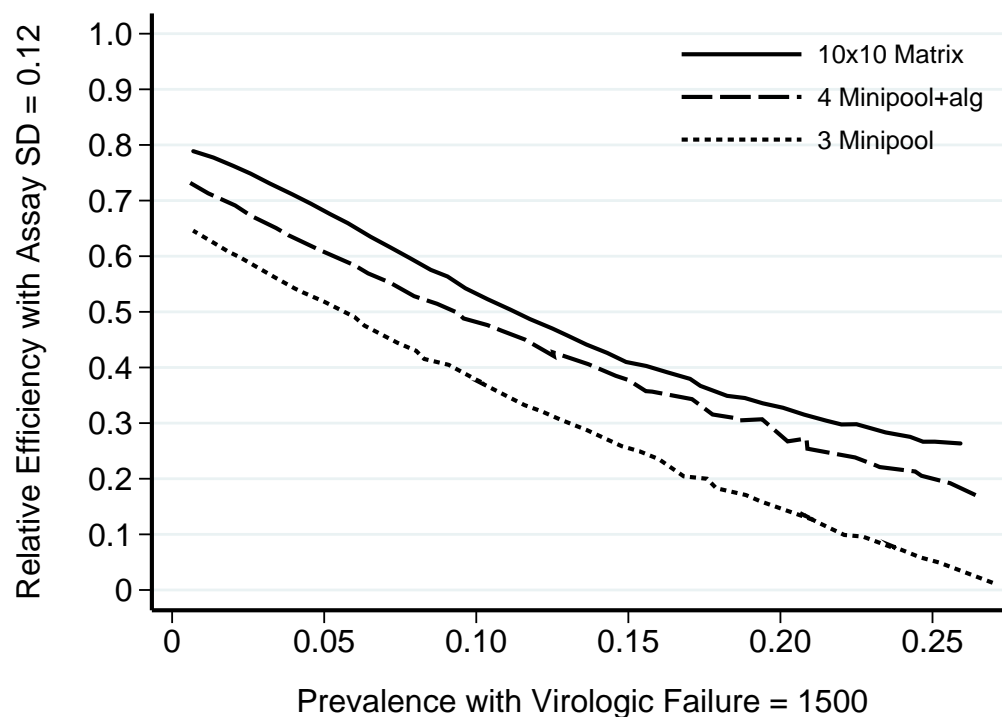
In this example the first individual sample that is tested has a viral load of 9935. This value divided by the pool size ( $9935/5 = 1987$ ) is subtracted from column ( $1522 - 1987 < 0$ ) and row estimate ( $1718 - 1987 < 0$ ). Samples or rows and columns with pool estimates below the threshold (here  $\leq 300$ ) are not tested individually and are assumed to be below the threshold (here  $\leq 1500$ ).





**Figure 3:** Relative efficiency of minipool, minipool + algorithm and matrix approaches for various pool sizes using a definition of virologic failure of HIV RNA  $\geq 1500$  copies/mL and assay standard deviation (SD) 0.12 copies/mL (on  $\log_{10}$  scale).





**Figure 4:** Relative efficiency for minipool, minipool + algorithm and matrix approaches with highest (among different pool sizes) relative efficiency compared to individual testing using a definition of virologic failure of viral load  $\geq 1500$  copies/mL and assay standard deviation (SD) 0.12 copies/mL (on  $\log_{10}$  scale).

