C.H.A.I.N. REPORT

Report 2003_3

Validation of Self-Reported Viral Load Levels

Peter Messeri
Robert Frey
Gunjeong Lee

Columbia University
Mailman School of Public Health
In collaboration with Medical and Health Research Association of New York, the NYC Department of Health and Mental Hygiene, and the NY Health & Human Services HIV Planning Council

DRAFT ONLY!!!

Originally submitted September 24, 2003
HRSA Contract H89 HA 0015-12

ACKNOWLEDGMENTS

A Technical Review Team (TRT) provides oversight for the CHAIN Project. In addition to Peter Messeri, PhD, David Abramson, and Angela Aidala, PhD, of Columbia University's Mailman School of Public Health, TRT members include Mary Ann Chiasson, DrPH, MHRA (chair); Susan Forlenza, MD MPH, NYCDOHMH; Robert Cordero, Office of AIDS Policy Coordination; JoAnn Hilger, NYCDOHMH; Julie Lehane, PhD, Westchester County DOH; Jennifer Nelson, MHRA; and Fatima Prioleau, PWA Advisory Group.

This research was supported by grant number H89 HA 0015-12 from the US Health Resources and Services Administration (HRSA), HIV/AIDS Bureau with the supported of the HIV Health and Human Services Planning Council, through the New York City Department of Health and Mental Hygiene and the Medical and Health Research Association of New York City, Inc. Its contents are solely the responsibility of the researchers and do not necessarily represent the official views of the U.S. Health Resources and Services Administration, the City of New York, or the Medical and Health Research Association of New York.

Introduction

This report is the first of a series reports that have as their major objective an understanding of the determinants of successful and failed HAART treatment with a focus on how patient access to the "continuum" of HIV services may best support long-term treatment success. Before we address the main purpose of this sequence of reports, we begin with a technically-oriented report that evaluates the quality of the self reports of viral load test results that CHAIN participants have provided since 1996, in the fourth round of interviews.

Background

The two most proximal measures of HIV-related health status are indicators of immune system functioning, CD4 lymphocyte cell counts, and HIV suppression and replication, viral loads measured as replicates / per milliliter of serum. Consequently CD4 counts and viral loads are important markers of both the natural history of the disease and the efficacy of antiretroviral treatment. Contemporaneous blood drawn (phlebotomy) as part of a research protocol or chart abstraction of the most recent lab test results are preferred methods for obtaining patient information on CD4 cell counts and viral loads (Kalichman et al. 2000). However, phlebotomy and chart abstraction data are prohibitively expensive for studies conducted outside of clinical settings such as the CHAIN study. Therefore the CHAIN study relies upon participant self reports for this information. We have been asking CHAIN participants' for their most recent CD4 test results since the baseline study was initiated in the fall of 1994. We began asking participants for the results of viral loads test with round 4 interviews that were initiated during 1996, soon after this test was introduced to assess the success of combination antiretroviral therapy.

Patient self reports of the results of recent medical diagnostic tests are subject to inaccuracies and biased responses (Kalichman et al. 2000). However HIV positive individuals are very attentive to their test results, which may produce better reporting of HIV-related test results than for other medical tests. Our experience administering the CHAIN interview supports this impression; although the supporting evidence is indirect. CHAIN participants have had little difficulties in reporting results of recent CD4 count tests. Moreover findings involving CD4 counts in numerous CHAIN reports exhibit aggregate trends and association with other health status measures, including mortality consistently associate lower CD4 counts with expedited poorer health outcomes. A few other studies have demonstrated good correspondence between self reports of CD4 counts and chart abstraction data (Kalichman et al.2000).

Because of the much more recent introduction of viral load testing into clinical practice and the large range of values, it is likely that viral loads are less reliably reported than CD4 counts; a situation confirmed in the only known published validation study (Kalichman et al. 2000) of self reports of these two HIV health status indicators.

The current report examines the quality of viral load reports in two ways. First we

examine whether there are differences in the health status and demographics of individuals who are and are not able to report the results of their viral load tests. There is some evidence that education level is associated with the knowledge of test results (Kalichman et al. 2000). Here we extent the range individual correlates to factors such as housing, mental health functioning and drug use. We also investigate the quality of reports by examining the association of viral loads with other HIV-related health status measures, self reports of CD4 counts and occurrence of Opportunistic Infections (O.I.'s) as well as its association with current use of HAART. We conclude this data with trends in our proposed measure of HAART treatment success and failure based upon the viral load self reported data.

METHODS

Data Sources

The purpose and general methodological strategy of the CHAIN project is well described in numerous other CHAIN reports. Study data are taken from the fourth through eighth rounds of the original CHAIN cohort. We also report comparable findings for the first 444 interviews with the new CHAIN cohort. Recruitment for the new cohort was initiated in 2002 and will continue through the fall of 2003.

Viral load questions

CHAIN participants were first asked for the results of viral load tests with round four interviews. The initial wording of the question was:

> "Have you ever had a test for viral load?
> Yes / No
> " IF YES, what was your recent viral load test?

The interviewer would right down a numeric value or "don't know" if the respondent could not give a numeric value.

Since a large percentage of individuals reported they did not know their viral load test results we modified the wording of the above question in round five to include a "good" and "bad" results categories as well as an undetectable category. If the respondent could not give a numerical response, the interviewers could probe for one of the above qualitative responses before coding a "don't know." The expanded response format sharply reduced the number of "don't know" test results from 39% in round 4 to 12% in round 5. Because the change in wording had a substantial impact on responses, the current analysis is restricted to data collected in rounds 5 and onward. Beginning in round 5, CHAIN participants were also asked the date of the most recent test result.

To assess the overall quality of reporting, responses to the viral load questions were

grouped into one of three categories: 1) reported results of viral load test, 2) tested but did not report results, 3) Never tested. For test results reports, exact and qualitative responses were combined into a three-point interval scale. Level 1 combined reports of undetectable viral loads, numerical values <=400 c/ml (copies per milliliter), or being told viral level was "good". Level 2 grouped numeric reports between 401 c/ml and 9999 c/ml. Level 3 group all numerical reports of 10,000 c/mL and larger with being told viral level was "bad." Level 2 was included to capture an intermediate stage of partial suppression with possible health benefits. Assignments of "good" and "bad" test results to level 1 and level 3 categories was based on intuitive appeal. A formal assessment of the congruence of this assignment is presented as an appendix to this report.

As a further refinement of our analysis we often compare all reported test results with those for tests reported to be completed within a half year of the interview. Unless otherwise indicated presentations of findings are restricted to test results done within six months of the intervals (186 days to be precise). Although we often conducted additional runs that included all test results.

Data on HAART and other study variables follow coding procedures used in previous CHAIN studies.

Findings

Completeness of response

The response to the viral load questions over five rounds of interviews with the original CHAIN cohort combined with initial interviews with the new cohort as summarized in Table 1 give reason for optimism that the quality of self reports are improving as patients become more familiar with this test procedure. A remarkable 96% of the cohort reported "ever having" a viral load test by round 6 interviews which were initiated in 1998, just two years after the test's introduction into general clinical practice.

There is also evidence from this table indicating that some CHAIN participants experiences difficulty in reporting their viral load counts. A small but persistent minority of CHAIN participants indicate that they didn't know the results of their most recent viral load test. Four in ten of CHAIN respondents didn't know their test results when we first introduced the question at round 4, although the number of don't knows dropped in half and has stabilized to under 20% of those reporting ever had a viral load test in all subsequent rounds of interviews. Part of this decline is attributable to including in round 5 and subsequent interviews the qualitative categories desribed above. Among participants who every had a viral load test, between 10 to 20 percent at each round of interviews reported that they were told that their viral loads were "good" or "bad" rather than reporting an exact count or reporting "undetectable" viral loads. The similarity in responses to viral questions between the new cohort and those for the later rounds of interviews with the original cohort suggests that there is generality to the pattern

rather than simply an idiosyncratic response from repeated interviews of the same group of individuals who become familiar with the structure and content of these questions

Characteristics of individuals who know their test results.

By the late rounds of interviewing virtually the entire sample reports getting a viral load test at some point. Yet a small but persistent number of individuals who are tested were not told or could not recall their test results. How best to handle this missing information requires that we examine whether the failure to recall is systematically related to patient attributes or health status. Table 2 presents percentages of those able to report viral load test results among all observations in which individuals report ever having a viral load test. We restrict analysis to interviews conducted for rounds 6 and subsequent interview, so that we are dealing with a period of time by which viral load had become a routine part of clinic management. Although there are significant differences in reporting test results related to gender ethnicity, drug use and education, none of the differences are of a substantial magnitude. Another problem in excluding cases which don't have viral load results is that those not reporting results may be healthier or sicker than those who report. Table 3 examines this for CD4 counts and reports of opportunistic infection. It appears from Table 3 that those not reporting viral loads have more compromised immune system functioning than those reporting results. This is most evidence in the difference in percentage in individuals with CD4 counts above 500. Among individuals who knew their test results, 36 percent reported CD4 counts above 500. Whereas among those who didn't recall their viral test results only 20% reported a CD4 count above 500. In contrast rates of OI do not differ with respect to viral load reporting status. Similar there is little difference in use of HAART and HAART adherence between those reporting and not reporting viral test results.

Dating of most recent test results

A second potential source of error concerns how close to an interview the viral test was performed. For most analyses it is desirable that reports of viral loads are relatively recent. This generally is the case. A total of 1,973 test result were reported when first round interviews for the new cohort are combined with observations for the original cohort for the fifth through 8th rounds of interviews. The great majority of reports 1,435 or 73% are for test results performed within a half year of the interview. Among the remaining tests results relatively small number, 122 report the duration since most recent test results of 188 days or more and 420 are missing dates. Virtually all the test results with missing dates are from rounds 5 and 6 interviewers. A more detailed examination of the dates of test results indicate that more than half the viral load data reported are for tests done within two months of the interview and 90 percent are within six months of the interviews. For analysis examining the association between viral loads and other health status measures, our analysis is restricted to observations with viral load test dates within a half a year of interview. Analysis were repeated for all observations, the findings reported below do not differ when all test results are included regardless of date.

Correspondence Between Viral Load Testing and Health Status

The next set of tables examine the concordance between reports of viral load and other measures of HIV-related health status. For this analysis we pool observations for interview rounds 5 through 8 for the original cohort with initial interviews with the new cohorts that report results of viral load tests done within six months of the interview. Gamma reported in these tables is a measures of association between two variables with ordered categories as opposed to continuous measures. Table 4 shows that viral loads has the expected association with both CD4 counts and O.I's. Lower viral loads is significantly associated with both increasing CD4 counts and lower rates of O.I.'s Observe that the distribution of CD4 counts for the partially suppressed viral load level, more closely approximates the undetectable level than the "high" load level. This middle category of viral load more nearly occupies a middle ground with respect to rates of O.I. We have also included distributions for CD4 counts and O.I.'s for observations in which a viral test was reported but results were not known. For this group of individuals, their distribution of CD4 count most closely approximates that for the high viral load category but has O.I. rates that virtually identical to that for the low viral load category.

Table 5 reports the distribution of viral load categories by HAART therapy status. When we percentage on HAART category there is a consistent pattern in which being on HAART is associated with lower viral laod status. Among observations when individuals are not on HAART, 50% report that their most recent viral loads was undetectable or below 400 c/m. The percent of individuals with undetectable/low viral loads increases to 60% among individuals who are partially adhereing to HAART regimens and the percentage with undetectable/low viral loads increases to 70% for individuals reporting complete adherence to their HAART regimen. A reverse progression is evident among the precentage of those with high viral loads. Finally not that individuals that partial suppressing is higher as well among individuals on HAART regardless of degree of adherence.

Tables 6 , 7 and 8 report on trends in the association from round 5 interviews onward between viral load levels and other HIV-related health status measures. Although not entirely consistent, the general pattern of association is similar across rounds of interviews and between the original and new CHAIN cohorts. At each round of interviews there are large differentials between those with high viral loads and lower loads in both elevated and suppressed CD4 counts. Observe that the sharpest break is between high viral loads and the intermediate viral load category. In table 7 a relative stable pattern of association across rounds of interviews is evident between viral load levels and percent reporting O.I.s.

Table 8 shows that using HAART and HAART adherence are usually but not always associated with increased percentages of individuals with low and undetectable viral loads and reduced percentages of people with undetectable viral loads. Complete adherence is usually but not always associated with improved viral load outcome when compared to partial adherent use of HAART.

Trends in treatment success

In anticipation of subsequent reports, Table 9 presents summary data on trends monitoring HAART treatment success.. The table also distinguishes between individuals who continue to have low viral loads in the absence of HAART and those that are benefitting from HAART. The trend data are restricted to the three most recent rounds of interviews with the original cohort. For the new cohort information is lacking on the history of HAART use.

In this table we offer a provisional definition of succeeding on HAART as an individual reporting undetectable viral load, viral load below 400 c/ml or being told that the level is "good" based on a viral test done within six months of the interview and is currently or in the past has been on HAART. Table 9 indicates that the number of individuals who are HAART naive has been cut by a third between the sixth and 8th round of interviews while those currently on HAART has remained at just over half the sample interviewed. The third column of table 9 further indicate that the percentage of all cohort members "succeeding" on HAART has increased from the sixth to the 7th and 8th rounds of interviews. Among those who have every been on HAART approximately two-thirds have low to undetectable viral loads. In concert with the sustained benefits the cohort continues to enjoy from HAART, not that the number of individuals who are able to maintain low and undetectable viral loads is becoming an increasingly smaller percentage of the cohort with each successive round of interviews. The final column prevents the reverse site of the picture. At each round of interview a stable 25% to 30% of the cohort is experiencing HAART failure. That is to say they have or were on HAART but have recent viral load levels > 400 c/ml.

Discussion

This report is a first in a series that is investigating the long-term success of HAART in reducing viral loads. Given concerns about the accuracy and reliability of self reports, this first report explored the quality of viral load data reported by CHAIN cohort members. The general completeness of viral reports and their consistent association with other HIV-related health status measures offer evidence that self reported viral load are acceptable indicators for statistical analysis of treatment success and failure. Our results suggest a couple of guidelines for using these data.

1. Viral load analysis is best restricted to observations collected in rounds six and later. From round six onward, virtually the entire cohort reports having had a viral load test.

2. Even in the latter interviews, useful viral load data is not universal. Roughly one quarter and perhaps less of those interviews either can't recall their viral load value or report a test result that is not very recent. Respondents with missing or unusable viral load do not appear to be very different that those reporting viral load except that those not reporting viral loads have lower CD4 counts. Furthermore the exclusion of viral loads from tests done more than 6 months prior to the interview have little or no impact on the associations between viral loads and other

HIV-related health status measures. Taken together these suggest that dropping cases that don't provide useful viral load data will not seriously distort generality of findings. Nonetheless imputation of missing viral load data may be considered that takes advantage of known correlations between viral loads and other information.

3. Our presentation indicates that people with partially suppressed viral loads (between 400 and 1000) is a useful middle category. It appears that partial suppression may confer some benefits, although this hypothesis requires further exploration. Supplemental analysis included in the appendix of this report, suggest that it si appropriate to assign the "good" and "bad" reports of viral load into the two extreme categories.

4. In this report we have defined treatment failure as evidence by viral loads above 400. This is consistent with clinical guidance in the 2003 edition of John's Hopkins Medical Management of HIV Infection. However this report suggest that there may value in exploring the impact of raising the level of "success" to under 10,000 c/ml.

In conclusion we believe that the large majority of CHAIN participants can reliably report viral loads given the relatively coarse categorization proposed for future study. It would still be desirable to obtain viral load data from some of the CHAIN participants to provide a stronger validation of self reports. CHAIN staff is in the process of arranging such an analysis.

References

Kalichman, Seth, Rompa David, Cage, Marjorie. 2000 Reliablity and validity of self-reported CD4 lymphocyte count and viral losd test results in people living with HIV/AIDS. International Journal of STD & AIDS. 11: 579-585

Table 1: Summary of Response to Viral Load Questions

Round of Interview	4	5	6	7	8	New Cohort
Sample N=	**420**	**651**	**508**	**444**	**388**	**444**
Ever Viral Load Test						
YES	65%	89%	96%	97%	97%	98%
No	33%	11%	4%	3%	3%	2%
Don't Know / Missing	2%	<0%	<0%	0%	0%	0%
Among Who Have Ever Had a Test (N=)	**(274)**	**(576)**	**(489)**	**(432)**	**(374)**	**(427)**
Undetectable	13%	30%	35%	39%	41%	36%
Know Exact Score: Good (<= 400)	5%	4%	3%	5%	3%	7%
Know Exact Score: Bad (> 400)	42%	39%	34%	21%	30%	38%
Don't know, but has been told "Good"	<0%	10%	8%	12%	7%	7%
Don't know, but has been told "Bad"	0%	4%	3%	5%	3%	3%
Don't know the score	39%	12%	17%	17%	16%	9%
Refuse to answer or Missing Value	0%	0%	<0%	<0%	0%	0%
Range of Viral Score						
Minimum Score	50	23	50	8	50	25
Maximum Score	3400000	1000000	688288	500020	600000	5000000

Table 2: Percent who reported test results among those tested by demographic characteristics (*analysis restricted to round 6 through 8 interviews and new cohort*)

	% Knows Test Results	P-value
Gender		
Male (N=952)	87%	.03
Females (N=766)	83%	
Race/Ethnicity		
Black (N=765)	81%	.03
White (N=171)	86%	
Latino(N=347)	86%	
Drug Use		.01
Never Used (N=305)	83%	
Past User (N=1,061)	87%	
Current user (N=358)	81%	
Housing Status		NS
Stable (N=1,400)	84%	
Doubled Up (N=89)	90%	
Unstable (N=235)	88%	
Mental health Functioning		NS
Normal (N=950)	84%	
Low (N=346)	82%	
Age		
Mean Age Know results	44	NS
Mean Age Don't know Results	43	
Education		
Mean grades completed know Results	11.7	.004
Mean grades comploeted don't know results	11.2	

Table 3: Health Status and Viral Load Knowledge of Viral Load Results (*analysis restricted to round 6 through 8 interviews and new cohort*)

	Knows Test Results (256)	**Tested but did no know results (1,468)**	**P**
CD4 Count > 500	36%	20%	<..000
CD4 Count <200	20%`	28%	
Reported O..I.	30%	27%	NS
HAART/adherent	33%	28%	NS
HAART/Not Adherent	19%	19%	
Not on HAART	47%	52%	.

Table 4: Distribution of CD4 Counts Occurrence of O.I..'s by Viral Load Levels
(Restricted to viral loads from tests results within a half year of the interview)

Viral Load Levels (N)	10000+* (311)	9999 - 400 (285)	<400** (839)	Unknown*** (256)
CD4 Count (Gamma = 0.30)				
0 - 100	12%	5%	3%	11%
101 - 200	27%	12%	12%	17%
201 - 300	18%	15%	16%	20%
301 - 500	26%	29%	26%	31%
500<	17%	39%	43%	20%
% Reporting Opportunistic Infection (Gamma = -0.29)	44%	35%	26%	27%

*Includes Individuals who said their viral load was "Bad"
** Includes individuals who reported undetectable viral loads or viral loads were good
***Tested but don't know for rounds 6 through 8
N are pooled Observations for rounds 5 through 8 and initial interviews with new cohort

Table 5: Viral Load Levels by HAART Therapy status

Viral Load Levels	HAART/Adherent (N=489)	HAART/Not Adherernt (N=278)	Not on HAART (N=668)
1000+*	16%	20%	26%
999-400	14%	19%	24%
<400**	70%	6o%	50%

Gamma=.26
*Includes Individuals who said their viral load was "Bad"
** Includes individuals who reported undetectable viral loads or viral loads were good
N are Pooled Observations for rounds 5 through 8 and initial interviews with cohort

Table 6: Association between CD4 Counts and Viral Load by Round of Interviews
(*restricted to individuals reporting date of test within half- year of interview*)

% Reporting CD4<200				
Viral Load Levels:	**10000+***	**9999 - 400**	**< 400****	**Gamma**
Round 5	46%	27%	29%	0.29
Round 6	34%	15%	16%	0.30
Round 7	31%	12%	18%	0.23
Round 8	33%	22%	12%	0.34
New Cohort	63%	39%	34%	0.31

% Reporting CD4>500			
Viral Load Levels:	**10000+***	**9999 - 400**	**< 400****
Round 5	18%	32%	35%
Round 6	9%	47%	43%
Round 7	22%	42%	44%
Round 8	22%	33%	49%
New Cohort	13%	38%	38%

*Includes Individuals who said their viral load was “Bad”
** Includes individuals who reported undetectable viral loads or viral loads were good

Table 7: Association between O.Is And Viral Load by Round of Interviews

% Reporting O.I.s				
Viral Load Levels:	**10000+***	**9999 - 400**	**<400****	**Gamma**
Round 5	58%	34%	26%	-.45
Round 6	40%	29%	19%	-.37
Round 7	42%	37%	25%	-.31
Round 8	41%	37%	26%	-.25
New Cohort	43%	38%	32%	-.16

*Includes Individuals who said their viral load was “Bad”
** Includes individuals who reported undetectable viral loads or viral loads were good

Table 8: Undetectable Viral Load and HAART Therapy Status by Round of Interviews

% Undetectable Viral Load (<400 c/ml)				
	No HAART	**Non-Adherent**	**Adherent**	**Gamma**
Round 5	43%	37%	60%	0.19
Round 6	41%	67%	66%	0.29
Round 7	60%	61%	82%	0.31
Round 8	52%	69%	71%	0.30
New Cohort	49%	60%	63%	0.22

% Viral Load > 10000			
	No HAART	**Non-Adherent**	**Adherent**
Round 5	33%	32%	22%
Round 6	28%	16%	22%
Round 7	23%	24%	12%
Round 8	26%	12%	14%
New Cohort	26%	19%	15%

Table 9: Trends in HAART Treatment Success (*restricted to observations with viral load test results in last half year***)**

Round of Interviews	HAART naive	Currently on HAART	Succeeding on HAART	Ever on HAART succeeding	Undetected viral load & HAART Naive	Not succeeding on HAART
6 (N=286)	27%	57%	43%	59%	13%	30%
7 (N=333)	20%	57%	57%	69%	11%	25%
8 (N=283)	18%	53%	53%	64%	9%	29%

Except for "every on HAART succeding" values are all percentaged on number participants interviewed and reporting a date of viral load test within a year of interview. For "Ever on HAART succeeding" Values are percentaged on HAART experienced respondents (Current and former users of HAART).

Succeeding on HAART =Viral loads <400 c/ml, undetectable or told "good" & either currently or in the past on HAART