

Tests of BDL Bias and Variance - v.1a

James R. Hunter, D.Sc.
Laboratório de Retrovirologia, UNIFESP

2025-12-30

Purpose

This file will test the determination of bias and variance when calculating viral loads and other variables that are subject to detection limits. For HIV-1 patients, the objective of antiretroviral treatment is to make their plasma RNA viral load zero.

However, there is a practical limit below which the assays cannot distinguish between the presence of virus and having no virus (the zero objective). This is called the “detection limit” and patients who achieve this level are said to have a viral load “BDL”, below detection limit. One of the most important consequences of exactitude in measuring true lack of the presence of virus is the phrase “undetectable = untransmittable” [G. C. Chang et al. [1]][2]. This frees, at least emotionally, the individual from concern about transmitting the virus to a partner.

There have been various strategies to deal with calculations of viral load in a panel of patients. One has been to eliminate the cases from the calculation, another has been to treat the viral level of those patients who are BDL as 0, assuming they are really without any virus. Still another strategy has been to assign a value equal to the detection limit. Others have used a value of half the detection limit, i.e., half of the difference between the detection limit and zero.

Name	Strategy	Value Assigned to “BDL”
censor	eliminate case	no value
zero	treat as zero	0
dl	treat as detection limit	detection limit
halfdl	treat as half of detection limit	detection limit/2

This study was inspired by a blog post by Nikolas Siccha, “Lower Limits of Detection or Quantification”. [3] He develops a Bayesian model to explore the pharmacodynamics of drugs that cause the measurement of interest to fall below detection limits. As it deals with the dynamics over time of a drug’s presence and efficacy in a patient, his question is more complex than the question I would like to address here.

! Important

What is the relationship between the true value of viral load below detection limit and the strategies used to estimate it and how does that effect analyses of the viral load of a sample or population?

Method

I will create a panel of 1,000 PLWH artificially with estimates of their viral load at a baseline and after 48 simulated weeks of antiretroviral treatment (ART) when some of them will have viral loads reduced to BDL, which I will define here as 50 copies per mL.¹ According to a 2009 clinical trial (STARTMRK) of Raltegravir, a 1st generation integrase inhibitor, 86.1% of the RTG arm of naive patients had viral loads BDL by week 48. [4] Likewise, a clinical trial (SINGLE) of a second generation integrase inhibitor, Dolutagravir, showed 88% of the DTG arm with viral loads below detection limit by week 48. [5] For my panel, I will use a proportion of 85% of patients achieving true viral loads between 0 and 50 by the 48 week mark. 85% is a bit more conservative than either of the real trials, but consistent with their results.

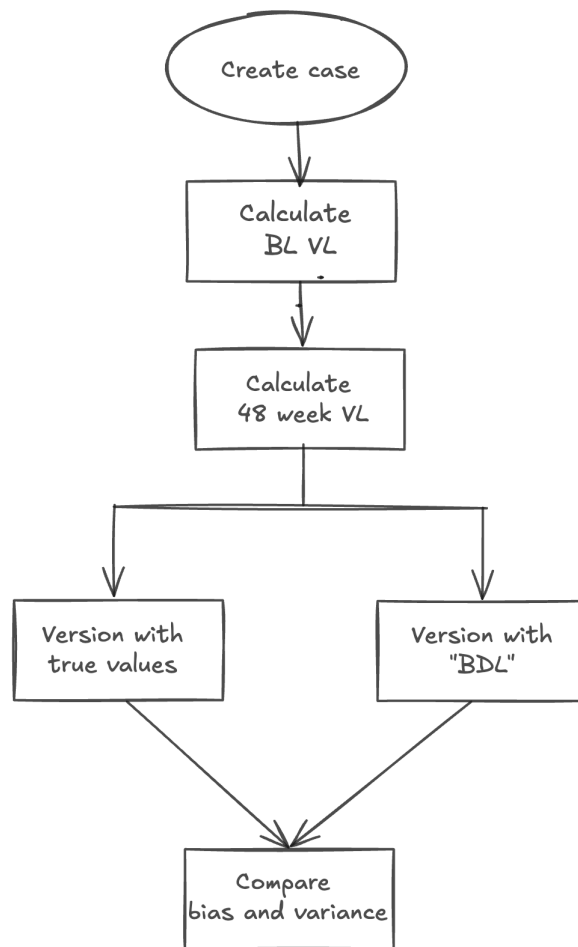
The baseline viral loads will be determined by calculating 1,000 normal variates from a truncated normal distribution with a mean of 900 copies per mL and a standard deviation of 1,000 copies. Since values below detection limit make no sense for this simulation, these values will be removed by using the truncated normal distribution.[6]

The 48 week data will be reduced dramatically (and unrealistically)² by 97.5% and the number of values below 50 copies per mL will be calculated. I will then create a second version substituting the true values below the detection limit with BDL. Using the four strategies outlined in Table 1, I will then compare the bias and variance of each of these strategies for substituting the true value of the viral load. Finally, I will propose a method for imputing values that is commonly applied in studies using the R statistical language.

The flow chart below shows a schematic of the workflow of this study.

¹While 50 copies was a common level for detection limit until 2024, since then it has been reduced to 30 or 20 copies depending on the measurement technology used. I am using 50 here as the purpose is to explore the bias and variance of samples rather than to test the detection limit itself.

²While unrealistic, this large reduction is useful to give us a number of values BDL that can be tested against the 4 strategies.



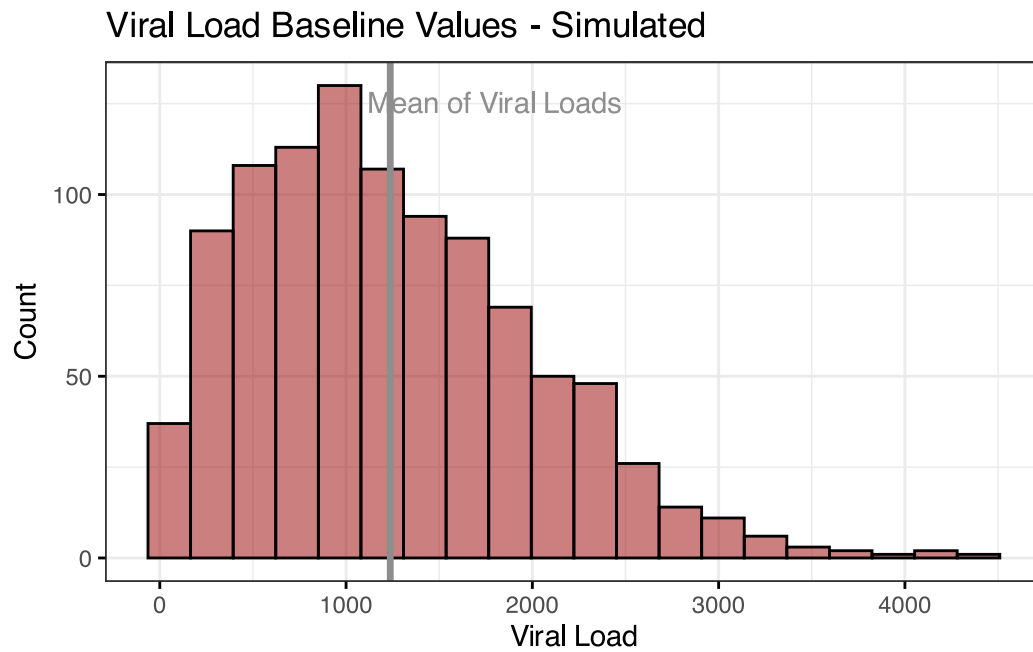
Baseline

Descriptive Statistics

bl

N: 1000

	bl
-----	-----
Mean	1237.22
Std.Dev	753.15
Min	51.00
Median	1116.50
Max	4395.00
IQR	1058.25
CV	0.61



Week 48 Viral Load

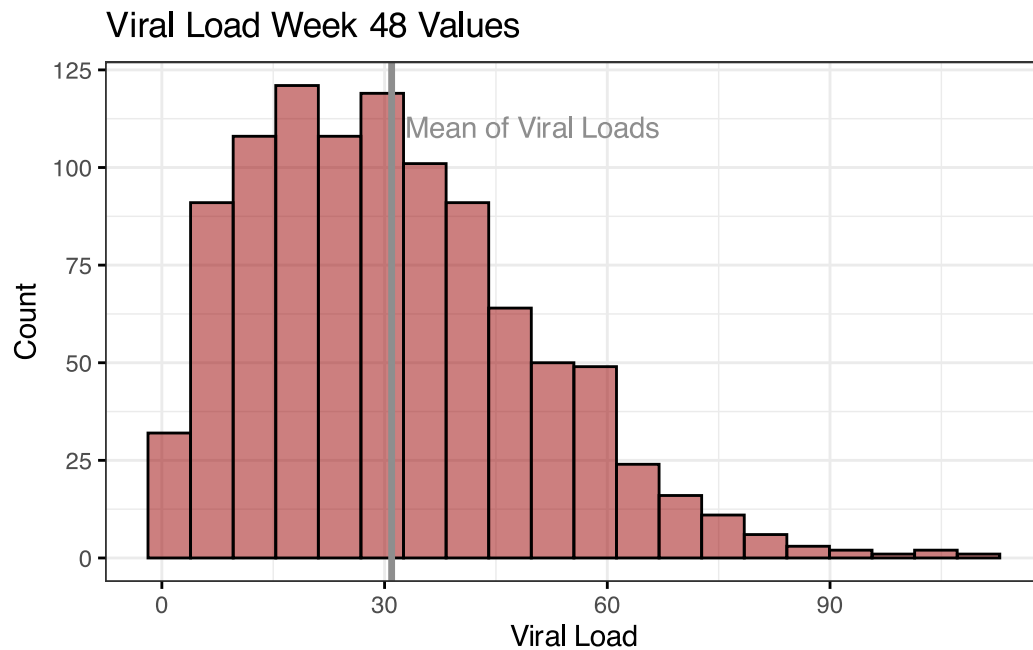
We will calculate the viral load at 48 weeks to be 97.5% less than the original viral load.

Descriptive Statistics

vldf\$w48

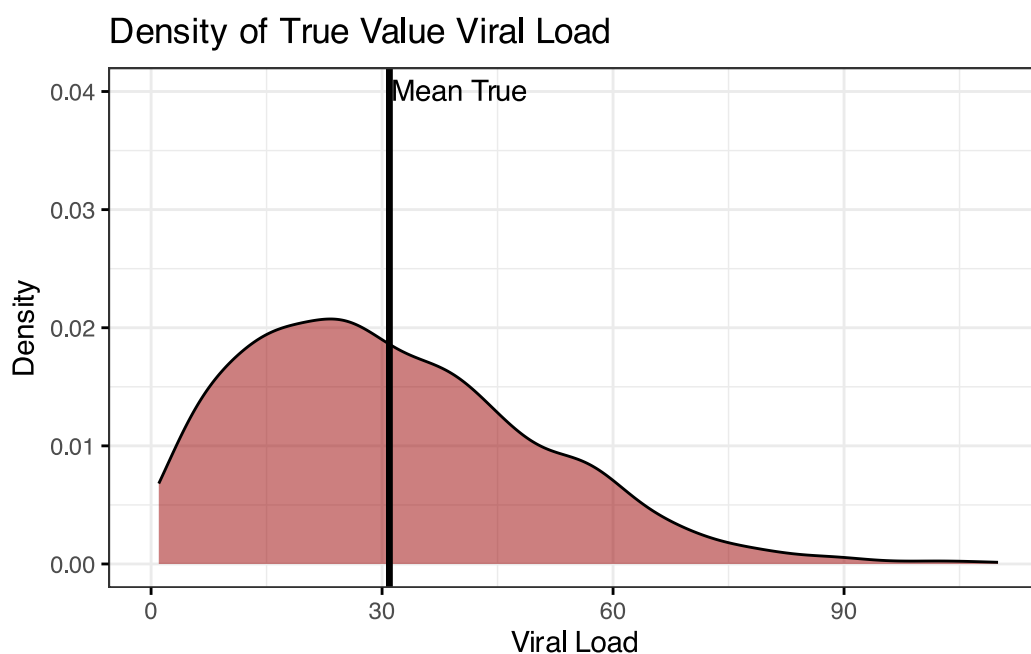
N: 1000

	w48
-----	-----
Mean	30.95
Std.Dev	18.84
Min	1.00
Median	28.00
Max	110.00
IQR	27.00
CV	0.61



To give you an idea of what we have now, here are the first 10 cases in our dataset. While we know the true values of the viral load, the assay itself will only give us a BDL signal instead of the value. This occurs 835 times out of the 1,000 cases. The question is what can we do about it and what does that do to the distribution of viral loads that we are trying to use to make an inference about HIV-1 treatment strategies.

	Week 48	
	True Value	BDL
Baseline		
2,271	57	57
335	8	bdl
1,263	32	bdl
1,533	38	bdl
1,304	33	bdl
794	20	bdl
2,412	60	60
805	20	bdl
2,918	73	73
837	21	bdl



Bias and Variance

The measures I am going to use to test the strategies listed at the beginning of this analysis are bias and variance. Bias measures how much the strategy distorts the distribution of true values. I will show this by comparing the mean, median, and standard deviation of the results of the strategies to those of the true values and through a density visualization of the distributions. The variance is rather more straightforward as we can find the statistical variance as determined by the classical formula for this quantity.

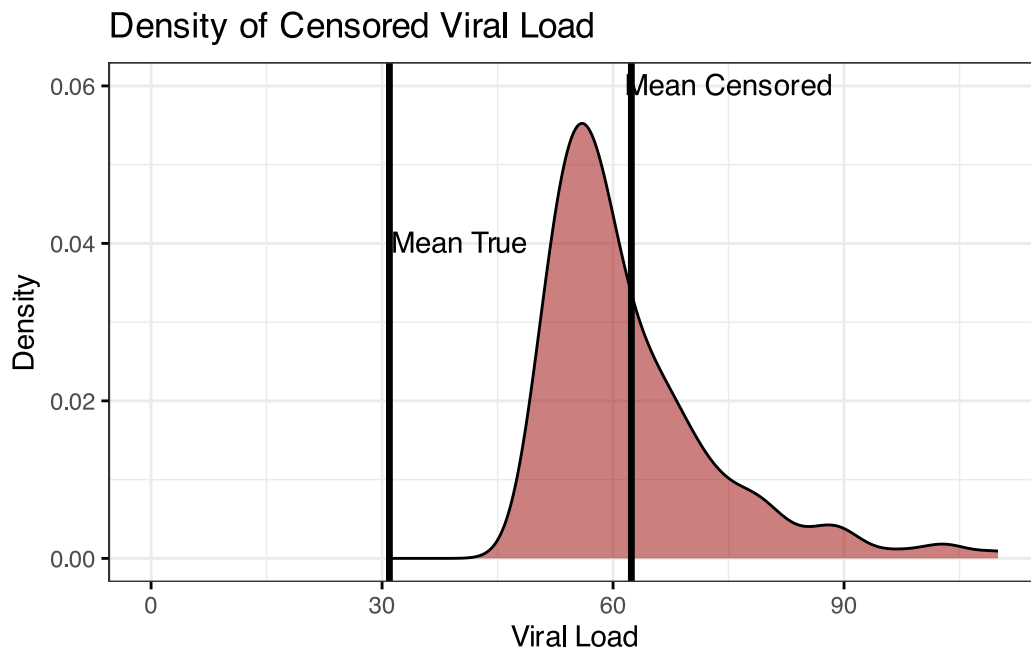
The Four Strategies

Censorship

In this strategy, which is very commonly applied when laboratories confront results of assays that do not indicate numerical values, simply removes those cases from the study.

When we do that here, we throw away 835 cases, meaning 83.5% of the cases, which likely contain many other nuggets of useful information. Before we even arrive at judging bias and variance, this reckless disposal of worthwhile information is the principal defect of the strategy. This dataset is not arbitrarily skewed toward having cases with BDL viral load after 48 weeks. As pointed out at the beginning, many studies show results in the range of 85%.

Week 48 Viral Load					
Strategy	Mean	Std. Dev.	Variance	Median	# of Cases
true	30.95	18.84	355.02	28	1000
censored	62.38	11.23	126.21	59	165



As you can see from the table and the density graph, the mean of the censored viral loads at 48 weeks is more than twice the true value as is the median value. That is because all values less than 50 were removed from the dataset because we did not know their true value; they were “BDL”. The censorship strategy leaves a highly biased dataset. The variance is greatly reduced, approximately a third the variance of the true values. Again, this is due to censoring—eliminating values within a certain range. Here the eliminated range is from the lowest value among the true values (1) up to 49.

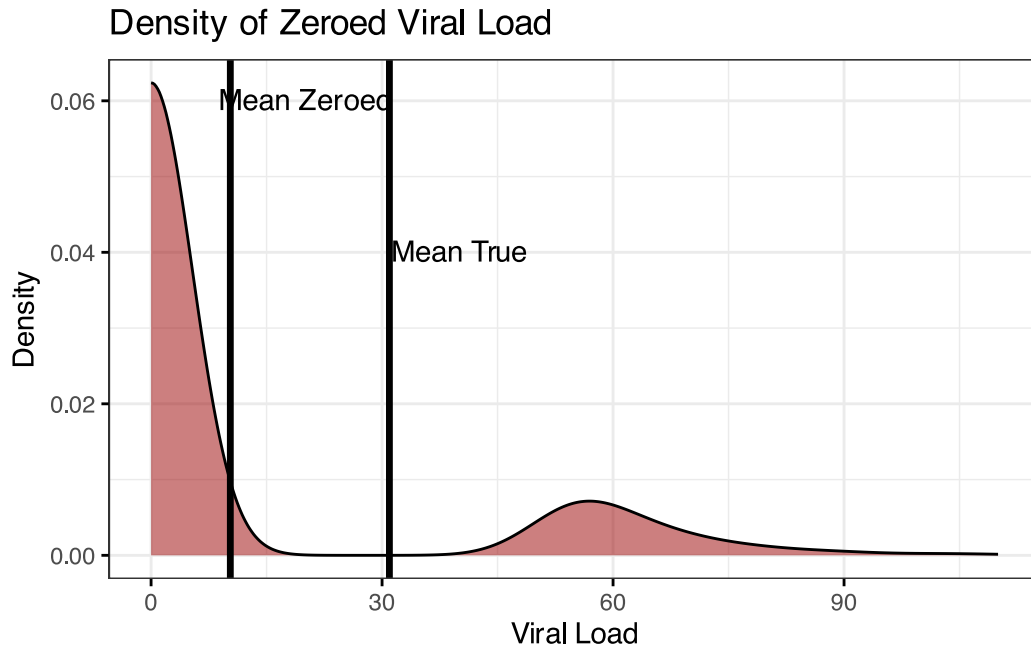
Treat as 0

This is a very attractive and tempting option for any research team. The objective of HIV treatment and cure projects is to eliminate the virus, drive it to zero. It would be very tempting to simply treat any viral load value that cannot be specified for being below the detection limit is to treat it as indistinguishable from zero. This has support in many cases by assuming that someone whose viral load is BDL is no longer communicable. The tendency is to assume that if the viral load cannot be counted, it is because it is not there. However, we know that is not the case. Below Detection Limit does not substantively

translate to a value of zero. So, this strategy as well comes with conceptual as well as calculation issues that argue against implementing it.

However, we need to test the implications of this strategy in terms of bias and variance as well. In this strategy, we will take all those values that are “bdl” in the dataset and substitute 0 for them.

Week 48 Viral Load					
Strategy	Mean	Std. Dev.	Variance	Median	# of Cases
true	30.95	18.84	355.02	28	1000
censored	62.38	11.23	126.21	59	165
Zero	10.29	23.61	557.30	0	1000



Here, the mean of the distribution now reflects the 835 cases that have viral loads that cannot be accurately measured. Although we know that if we had been able to measure these values, they would have produced a mean of 30.952 and a variance of 355.0227187. However, because of the huge numbers of false zeros that have been inserted, the mean has now fallen to `wk48_compare$mean_wk48[3]`. Because there are now the mass of zero values and a lesser number of high values, with few in the middle (the trough of the graph on the density graph), the variance has increased to 557.302038. Since the zero values represent such a large portion of the dataset, the median has fallen to 0.

Thus, while tempting, this strategy also distorts dramatically the true nature of the data and bases medical and research decision-making on a false premise.

Bibliography

- [1] G. C. Chang et al., “Brief Report: Undetectable HIV-1 Viral Load Among Virally Suppressed People Living With HIV: Implications for Undetectable Equals Untransmittable (U=U) in Resource-Limited Settings Using Dried Spot Testing,” *Journal of Acquired Immune Deficiency Syndromes* (1999), vol. 100, no. 4, pp. 289–293, Dec. 2025, doi: 10.1097/QAI.0000000000003745.
- [2] C. Combescure et al., “How reliable is an undetectable viral load?,” *HIV medicine*, vol. 10, no. 8, pp. 470–476, Sep. 2009, doi: 10.1111/j.1468-1293.2009.00714.x.
- [3] N. Siccha, “Lower limits of detection or quantification.” [Online]. Available: https://www.generable.com/post/lloqs?utm_source=substack&utm_medium=email
- [4] J. L. Lennox et al., “Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naïve patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial,” *Lancet (London, England)*, vol. 374, no. 9692, pp. 796–806, Sep. 2009, doi: 10.1016/S0140-6736(09)60918-1.
- [5] S. L. Walmsley et al., “Dolutegravir plus abacavir-lamivudine for the treatment of HIV-1 infection,” *The New England Journal of Medicine*, vol. 369, no. 19, pp. 1807–1818, Nov. 2013, doi: 10.1056/NEJMoa1215541.
- [6] O. Mersmann, H. Trautmann, D. Steuer, and B. Bornkamp, *Truncated Normal Distribution*. 2023. [Online]. Available: <https://github.com/olafmersmann/truncnorm>