Measuring the potential role of frailty in apparent declining efficacy of HIV interventions

Felicia P. Hardnett, Charles E. Rose

Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA

Objective: In recent HIV intervention trials, intervention efficacies appear to decline over time. Researchers have attributed this to “waning,” or a loss of intervention efficacy. Another possible reason is heterogeneity in infection risk or “frailty.” We propose an approach to assessing the impact of frailty and waning on measures of intervention efficacy and statistical power in randomized-controlled trials.

Methods: Using multiplicative risk reduction, we developed a mathematical formulation for computing disease incidence and the incidence rate ratio (IRR) as a function of frailty and waning. We designed study scenarios, which held study-related factors constant, varied waning and frailty parameters and measured the change in disease incidence, IRR, and statistical power.

Results: We found that frailty alone can impact disease incidence over time. However, frailty has minimal impact on the IRR. The factor that has the greatest influence on the IRR is intervention efficacy and the degree to which it is projected to wane. We also found that even moderate waning can cause an unacceptable decrease in statistical power while the impact of frailty on statistical power is minimal.

Discussion: We conclude that frailty has minimal impact on trial results relative to intervention efficacy. Study resources would, therefore, be better spent on efforts to keep the intervention efficacy constant throughout the trial (e.g., enhancing the vaccine schedule or promoting treatment adherence).

Keywords: HIV, HIV intervention trial, HIV vaccine, HIV prevention, Frailty

Introduction

Background

An HIV intervention is defined as a “medical, clinical, or public health approach designed to moderate biological and physiological factors to prevent HIV infection, reduce susceptibility to HIV, and/or decrease HIV infectiousness.” Recent advancements in HIV intervention research suggest that a number of effective biomedical interventions to reduce HIV transmission will soon become available for widespread use. An HIV vaccine trial (RV144) reported an overall estimated decrease of 31% (95% CI: 1.1, 52.1; \( P = 0.04 \)) in the rate of HIV infection among those receiving the vaccine. A microbicide trial conducted in South Africa (CAPRISA) showed that Tenofovir gel resulted in an estimated 39% (95% CI: 6.0, 60.0; \( P = 0.017 \)) overall reduction in the rate of HIV acquisition.

Although these results seem promising, additional findings have been somewhat disappointing and difficult to explain. In both the RV144 and CAPRISA trials, the efficacy of the intervention appeared to decline over time (i.e., the estimated rate ratios approached 1). In the RV144 trial, the estimated average hazard ratio increased steadily from 0.40 at 12 months to 0.69 at 42 months. The CAPRISA trial showed a similar decline in intervention efficacy. The rate ratio comparing HIV incidence in the microbicide arm to the placebo arm increased from 0.53 at 6 months to 0.61 at 30-month follow-up. Researchers have attributed this apparent decline in efficacy to different intervention-related factors such as waning vaccine efficacy in the case of an HIV vaccine trial or a decrease in the proper use of the intervention, or adherence, in the case of the microbicide trial.

A recently published article by O’Hagan et al. suggested that, in addition to biological (e.g., waning) and behavioral (e.g., adherence) mechanisms by which an otherwise effective intervention might
appear to lose efficacy, there are largely unrecognized changes in the underlying study population that may partially explain what is observed in these studies. The article highlights the CAPRISA and RV144 trials and suggests frailty as a partial explanation for the observed decline in intervention efficacy.5

**Definition of frailty**

In the context of an HIV intervention trial, frailty is defined as heterogeneity in HIV infection risk among subjects in a study population. O’Hagan et al. pointed out that, after enrollment has closed and a trial progresses, high risk subjects are likely to become infected and, therefore, removed from the population at risk early on. As a result, disease incidence will decrease over time because those who remain in the study have a lower probability of infection independent of any treatment effect. This process is also referred to as “survivor bias,” “survivor cohort effect,” “crossing of hazards,” or “depletion of susceptibles”5

**Potential impact of frailty on the results of randomized controlled trials**

A randomized controlled trial (RCT) is often used to test the efficacy of a medical intervention on measurable outcomes (e.g., disease or death). Study participants are randomly assigned either to the arm(s) receiving the medical intervention (also called a “treatment” or “intervention” arm(s)) or the arm(s) not receiving the intervention (also called a “placebo” or “control” arm(s)). When there is heterogeneity in infection risk (i.e., frailty), the disease incidence observed early in the study is driven by those who are at highest risk of infection. However, O’Hagan et al. suggest that, during the course of a RCT of an effective intervention, high risk individuals in the intervention arm will likely remain uninfected longer when compared to high risk individuals in the placebo arm because of the protection conferred by the intervention. This difference in the rate of infection between the intervention and placebo arm will eventually resolve once the higher risk individuals are removed from both treatment arms. As a result, the IRR, which may have been significant at the beginning of the trial, will eventually progress toward the null value of 1.0 as the remaining population at risk in both treatment arms consists almost solely of lower risk participants. As support of this assertion, O’Hagan et al. suggest that waning immune response and decreased adherence do not explain the decreasing incidence observed in the placebo arms in both trials. Frailty, however, may partially explain decreased incidence in both treatment arms as well as the neutralizing of the IRR.

**Motivation for current study**

Although O’Hagan’s article raises an interesting issue in interpreting published HIV randomized trials, it does not attempt to measure the potential impact of frailty on the studies that were reviewed. A recent response to O’Hagan’s article posits decreased adherence (as measured by the number of returned gel applicators...
and a preliminary analysis of vaginal tenofovir concentrations) as the more likely cause of the apparent declining efficacy in the CAPRISA 004 trial by citing comments in the original article. However, a more mathematically rigorous assessment of the simultaneous impact of frailty and waning/adherence would provide a worthwhile contribution to this discussion. Additionally, measuring the potential impact of frailty under multiple study scenarios may help to inform the design and analysis of future RCTs.

Goals of the current study
We present a series of hypothetical study scenarios using both fixed and varied parameters that describe a study and its population. We then assess the impact of frailty in three separate contexts: (1) within a single population in the absence of an intervention; (2) within a RCT where the intervention efficacy remains constant over time; and (3) within a RCT where the intervention efficacy is projected to wane at varying degrees over time. We also examine the impact of both frailty and waning/adherence on statistical power. Hereafter, the term “waning” will be used to describe both the medical and behavioral factors, such as decreased immune response or decreased treatment adherence, which might negatively impact the efficacy of an intervention over time.

Mathematical Formulation
Consider a RCT of an effective intervention where each participant is randomly assigned to either the intervention or placebo arm. HIV-uninfected participants who enter the study are tested for HIV seroconversion at several, equal-length follow-up time intervals, and are followed until they test HIV-positive or until the study is terminated.

We first define several quantifiable study parameters to allow us to examine the effects of frailty and waning of intervention efficacy. These study parameters can be classified as either population-related parameters or intervention-related parameters that describe specific study characteristics pertinent to most HIV RCTs. Population-related parameters include: HIV risk groups (subsets of the study population with similar HIV risk) and frailty (the degree of heterogeneity in infection risk experienced across the different subgroups). Intervention-related parameters include: intervention efficacy (the percent reduction in infection risk conferred by the intervention) and waning (the reduction in intervention efficacy that occurs over time). We constructed the mathematical formulation needed to compute our outcome measures using these parameters as inputs (see Supplementary Material 1, which contains all definitions, notation and mathematical formulas). Our formulations utilize multiplicative risk reduction, a standard epidemiological framework, which incorporates selected factors that influence disease risk. These factors can be additive or multiplicative. In the multiplicative case, these factors are represented by multipliers that either increase (if $w > 1$) or reduce (if $w < 1$) the risk of disease. While there have been several approaches to examining the impact of frailty and waning in recent scientific literature, to our knowledge, this is the first known use of multiplicative risk reduction to model these factors.

Scenario design and formatting of results
To assess the impact of both frailty and waning, we constructed a series of hypothetical RCTs characterized by the population- and intervention-related study parameters described previously. We then calculated the IRR at each time interval. We held several study features constant while varying factors pertaining to waning and frailty and then measured the percent change in our final point estimates.

Our scenarios:
1. Assume that the study population consists of five risk groups ranging from very high to very low risk.
2. Assume that 50% of the subjects are at moderate risk of infection, a small proportion of subjects (5% each) are at very high or very low risk, and 20% of subjects fall into each of the “high” or “low” risk groups (Figure 2).
3. Assume that disease risk is balanced between treatment arms at baseline (i.e., the randomization scheme resulted in treatment arms that are identical with regard to their distribution of disease risk).
4. Assume that loss-to-follow-up is negligible.
5. Assume intervention efficacy is 50%. That is, the intervention reduces disease probability by half.
6. Assume that intervention efficacy is constant across all risk groups within the treatment arm on the multiplicative scale. In other words, the treatment is equally effective at preventing HIV transmission for all risky acts across all risk groups.
7. Assume that intervention efficacy waning is constant across all risk groups and time intervals on the multiplicative scale.

![Figure 2](image-url) The baseline risk distribution of the study population assumed for all scenarios.
We present study scenarios using five frailty levels and six waning levels for a total of 30 different study scenarios. We set the probability of disease for the highest risk group in each scenario at 60% and used a range of values for the frailty parameter, $\eta$, (0, 0.2, 0.3, 0.5, and 0.8) in order to model different degrees of heterogeneity in infection risk. Figure 3 illustrates how the different frailty levels compare with one another through the first year of the study. In the absence of frailty, infection risk is equal across all risk groups. However, as frailty increases from “Slight” to “Extreme” (i.e., as $\eta$ increases), the relative degree of infection risk changes dramatically. We established a follow-up time of 10 years and, for simplicity of calculations and interpretations, we set the length of each follow-up time interval at 1 year. We calculated the time-specific IRR at each follow-up time for each scenario.

### Analytic Method

We used Microsoft Excel 2010 to construct all 30 study scenarios, compute intermediate probabilities, and perform final incidence and IRR calculations at baseline ($t = 0$) and 10 additional follow-up time points. We also incorporated additional inputs that allowed us to modify factors such as intervention efficacy, length of follow-up, and population baseline risk distribution to assess the limitations imposed by our scenario assumptions (see Supplementary Material 2, which contains a written description of the spreadsheet calculations along with screenshots of actual scenarios).

The goal of the first stage of the analysis was to measure the change in the population risk distribution over time under the five different levels of frailty and compute the corresponding change in disease incidence. This is independent of the intervention efficacy so all intervention-related parameters are ignored. In so doing, we are able to assess the impact of risk heterogeneity alone on disease incidence over time. Results were obtained for 1, 3, 5, and 10-year follow-up times.

The second analytic stage introduces the intervention and its effect on individual disease probability (or incidence) over time. At this stage, the goal is to assess the impact of frailty on the IRR. We used the first-stage results as the control or placebo data and created a second arm, the intervention arm, which was identical to the placebo arm with the exception of having received an intervention with an efficacy of 50%. The intervention waned at varying rates. We then computed the time-specific IRR at each follow-up time point. Within each level of waning, we computed the IRR under each level of frailty and graphically displayed the results.

In the third and final stage of our analysis, we examined the impact of both frailty and waning on statistical power. Comparing survival curves between our intervention (I) and control (C) groups using the Cox Proportional Hazards model and assuming an intervention trial with 5-year follow-up, we wished to test the hypothesis $H_0$: $\text{IRR} = 1$ versus $H_1$: $\text{IRR} \neq 1$. We set the significance level at $\alpha = 0.05$ and compared statistical power across varying sample sizes (see Supplementary Material 3 for the mathematical formulation for our power calculations).

---

**Figure 3** The five levels of frailty (heterogeneity in infection risk) modeled in the study scenarios ranging from no frailty to extreme frailty.
Results

Change in population risk distribution over time

The change in the population risk distribution over time for each study scenario is presented both numerically and graphically in Figure 4. The graph for each scenario shows the risk distribution at baseline for each scenario as well as for follow-up years 1, 3, 5, and 10. The rows of the adjacent tables show the percentage of the at-risk population that falls into each risk group at each follow-up time. To assess

![Figure 4](image-url)

Figure 4  The risk distributions of a simulated study population in the absence of an intervention, the corresponding disease incidence and percent change in incidence from year 1 at 1, 3, 5, and 10-year follow-up under three levels of frailty.
the impact of frailty on incidence, we present the overall disease incidence as well as the corresponding percent change in incidence from Year 1.

The initial impact of frailty is evidenced by the Year 1 incidences under each scenario. Under the “No Frailty” scenario ($\eta = 0.0$), all risk groups have a disease probability that is equal to the 60% assumed for the highest risk group. As a result, Year 1 incidence is 60% and this remains constant throughout the course of the study.

Under the remaining scenarios, as frailty increases from “Slight” ($\eta = 0.2$) to “Extreme” ($\eta = 0.8$), Year 1 incidence decreases dramatically from $\sim 39\%$ to only $\sim 7\%$ because the probability of disease in the four lower risk groups decreases sharply with each scenario. Because the Year 1 incidence varies widely, a direct comparison of the percent change in incidence across scenarios is potentially misleading. However, tracking the percent change in incidence over time within each scenario suggests that frailty, in the absence of an intervention, does impact disease incidence. As discussed in more detail below, under each scenario, as the proportion in the highest risk groups is reduced (i.e., high risk individuals are removed from the population at risk) the lower risk groups comprise an increasing proportion of the study population resulting in a steady decline in disease incidence within each scenario.

Under “Slight” frailty (where the disease probabilities of neighboring risk groups differ by a factor of only 20%), the proportion of the population remaining in the two highest risk groups at 3-year follow-up declined from 25% (at baseline) to approximately 13%. While the proportion of the study population in these groups declined by half, incidence declined by only $\sim 5\%$. At the 10-year point, the proportion that remained uninfected in the two highest risk groups dropped to $\sim 2.5\%$ resulting in a total incidence decline of only $\sim 17\%$.

A similar trend can be observed in the other study scenarios leading up to “Extreme” frailty where neighboring risk groups differ in their disease probability by a factor of 80%. In this scenario, the proportion of the study population that remained in the two highest risk groups at the three-year mark dropped from 25 to $\sim 16\%$ resulting in an incidence decline of $\sim 39\%$ (from $\sim 7\%$ to $\sim 4\%$). At the 10-year mark, the proportion at highest risk fell to $\sim 8\%$ with a corresponding $\sim 62\%$ decline in incidence (from $\sim 7\%$ to $\sim 3\%$). Again, these results are in the absence of any treatment effect.

**Impact of frailty and waning on RCT results**

When intervention efficacy remains constant (i.e., no waning), dramatic increases in frailty had a relatively small impact on IRR (Figure 5). However, even a slight increase in waning from 0 to 10% resulted in a steep increase in the IRR. This increase becomes even more pronounced as waning increases and, at times, for example, when waning was assumed to be 50%, results in an IRR that exceeds 1 (i.e., the incidence in the intervention arm briefly exceeds that of the placebo arm) in the case of extreme frailty. Extending the follow-up time far beyond 10 years reveals that the IRR in each scenario progresses asymptotically toward 1 (Not shown).

**Impact of frailty on statistical power**

Table 1 shows the power calculations for all 30 study scenarios at various sample sizes. These results suggest that at 0% waning, statistical power is retained even at extreme levels of frailty but quickly reaches unacceptable levels at slight levels of waning for studies with small sample sizes. As expected, statistical power improves as sample size increases. However statistical power consistently remains well below 80% at moderate and extreme waning levels.

**Discussion**

Our analysis confirms that frailty alone can result in decreasing disease incidence over time. This finding is consistent with the apparent decline in incidence observed in the placebo arms of both the RV144 and the CAPRISA trials. However, the second stage of this analysis, which includes an intervention arm, demonstrates that frailty alone has limited impact on the final outcome measure of a RCT. Although frailty appears to influence the IRR, the results of this analysis suggest that the impact of frailty is minimal when compared to waning (or decreased adherence) and decreases with time as the infection probabilities within the remaining population at risk become more homogeneous. The study-related factor that appears to have the greatest influence on the IRR is the intervention efficacy itself and the degree to which the efficacy of the intervention is projected to wane.

The ability of an intervention to decrease the probability of disease can wane for various reasons including both biological and behavioral factors. In extreme cases, the rapid progression of the IRR toward 1 could result in the early termination of a study due to the apparent lack of efficacy of the intervention (or perceived harm in the case where waning is substantial and the IRR exceeds 1). This can be at least partially remedied by enhancing the vaccine schedule or implementing behavioral strategies that encourage treatment adherence.
Figure 5 Estimated incidence rate ratio (IRR) under varying degrees of frailty and waning. Frailty (heterogeneity in infection risk) is incorporated by sequentially multiplying the infection probability of a higher risk group by 1 (η, a fixed adjustment factor used to incorporate frailty). A higher value of eta corresponds to greater frailty.

Table 1 Power calculations for each study scenario at varying sample sizes*.

| Number of participants per group | Frailty | 0%    | 10%   | 20%   | 30%   | 40%   | 50%   |
|---------------------------------|---------|-------|-------|-------|-------|-------|-------|
| 100                             | 0       | 0.99  | 0.76  | 0.34  | 0.14  | 0.07  | 0.04  |
|                                 | 0.2     | 0.99  | 0.67  | 0.26  | 0.10  | 0.04  | 0.03  |
|                                 | 0.3     | 0.98  | 0.61  | 0.21  | 0.07  | 0.03  | 0.03  |
|                                 | 0.5     | 0.96  | 0.49  | 0.14  | 0.05  | 0.03  | 0.04  |
|                                 | 0.8     | 0.83  | 0.24  | 0.06  | 0.03  | 0.06  | 0.07  |
| 500                             | 0       | 1.00  | 1.00  | 0.94  | 0.50  | 0.18  | 0.07  |
|                                 | 0.2     | 1.00  | 1.00  | 0.84  | 0.31  | 0.08  | 0.03  |
|                                 | 0.3     | 1.00  | 1.00  | 0.74  | 0.21  | 0.05  | 0.03  |
|                                 | 0.5     | 1.00  | 0.99  | 0.52  | 0.10  | 0.03  | 0.07  |
|                                 | 0.8     | 1.00  | 0.80  | 0.13  | 0.04  | 0.12  | 0.18  |
| 1000                            | 0       | 1.00  | 1.00  | 1.00  | 0.79  | 0.31  | 0.11  |
|                                 | 0.2     | 1.00  | 1.00  | 0.99  | 0.54  | 0.12  | 0.03  |
|                                 | 0.3     | 1.00  | 1.00  | 0.96  | 0.37  | 0.07  | 0.04  |
|                                 | 0.5     | 1.00  | 1.00  | 0.81  | 0.16  | 0.03  | 0.09  |
|                                 | 0.8     | 1.00  | 0.98  | 0.22  | 0.05  | 0.21  | 0.32  |

*5-year follow-up, H0: RR = 1 versus H1: RR≠1, α = 0.05.
The impact of treatment adherence in particular on the measure of intervention efficacy in the CAPRISA trial has received recent attention in the literature. Dai et al. proposed a method for measuring the efficacy of an HIV intervention that could potentially mitigate the impact of treatment non-adherence on study results. This approach groups participants into treatment arms based on their level of treatment compliance as measured by drug concentration. By estimating intervention efficacy based on actual exposure rather than the treatment assigned at randomization (i.e., the intention-to-treat approach), the estimated efficacy of the intervention in the CAPRISA trial increased dramatically. This finding confirms the impact of actual intervention exposure on trial outcomes.

A sensitivity analysis was performed to assess the impact of our selection of the following fixed study parameters: baseline intervention efficacy and baseline population risk distribution. When we modified the baseline intervention efficacy, the rate at which subjects became infected changed but the overall study conclusion did not change. Shifting the baseline risk distribution so that more of the study population fell into the higher risk groups caused the study population to become depleted more quickly but this had no impact on our overall conclusions. Similarly, shifting the baseline risk distribution so that more of the study population fell into the lower risk group slowed the depletion of the population but did not change our conclusions. In each scenario, waning had the greater impact on study results. Additionally, in the context of this analysis, we considered extreme frailty values to give the frailty hypothesis its best chance of being supported even though, from a practical standpoint, these scenarios are unlikely to occur in most RCTs of HIV interventions which commonly target a more homogeneous, high-risk population.

Proposed approaches to adjust for the impact of frailty include adjustments to study design and data collection so as to monitor risk factors for each participant and censoring of their follow-up time during periods when they are not at risk. However, such modifications might prove to be inaccurate because it is not possible to measure all pertinent risk factors. Attempts to comprehensively monitor risk factors across an entire study population may also prove to be too resource-intensive in many trial settings. The current analysis of the impact of frailty on incidence, IRR and statistical power suggests that these approaches will only address a minor source of bias relative to the bias introduced by waning intervention efficacy. However, depending on the study and the available resources, it might be more feasible to make adjustments to the study design and statistical analysis to mitigate the impact of frailty rather than attempting to modify human behavior (i.e., promoting treatment adherence) or developing an intervention that is long lasting. It is worth noting that the underlying framework of this analysis (multiplicative risk reduction) and the Excel application developed to answer our study question can also be applied to assess the impact of other aspects of RCT study design. For example, our methodology can be extended to model the usefulness of stratified randomization by assuming different intervention efficacies for different risk groups. The impact of differential loss to follow-up, and other common study conduct issues, also can readily be modeled making this approach broadly applicable to investigations of issues related to the optimal allocation of study resources.

Disclaimer Statements
Contributors FH served as lead investigator. She designed the study and developed the application used to perform the analysis. She also drafted the manuscript. CR served as second author. He provided valuable input in the design of the study. He also reviewed and commented on the manuscript as it was being developed.

Conflicts of interest The authors do not have any commercial or other associations that might pose a conflict of interest (e.g., pharmaceutical stock ownership, consultancy).

Ethics approval Ethical approval was not required.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

References
1 Public health (ARTAS, CTR & CRCS), behavioral. overview of select interventions & strategies: biomedical (medication adherence). Available from: https://effectiveinterventions.cdc.gov/docs/default-source/general-docs/14-1021-hip-overview-factsheet.pdf?sfvrsn=0 (Accessed 2015 September 21).
2 Rerks-Ngarm S, Pitisutthithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med. 2009;361:2209–2220.
3 Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, et al. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science. 2010;329:1168–1174.
4 Ferrer RA, Morrow KM, Fisher WA, Fisher JD. Toward an information-motivation-behavioral skills model of microbicidal adherence in clinical trials. *AIDS Care*. 2010;22(8):997–1005.

5 O’Hagan JJ, Hernan MA, Walensky RP, Lipstitch M. Apparent declining efficacy in randomized trials: examples of the Thai RV144 HIV vaccine and South African CAPRISA 004 microbicidal trials. *AIDS*. 2012;26:123–126.

6 Grobler A, Karim SA, *et al*. Declining adherence as a more likely explanation than frailty of the apparent decline in efficacy in the CAPRISA 004 trial: response to O’Hagan *et al*. *AIDS*. 2012;26:2261–2265.

7 Dai JY, Gilbert PB, Hughes JP, Brown ER. Estimating the efficacy of preexposure prophylaxis for HIV prevention among participants with a threshold level of drug concentration. *Am J Epidemiol*. 2013;177(3):256–263.