TITLE: ANALYTICAL ADVANCES IN THE *EX VIVO* CHALLENGE EFFICACY ASSAY.

AUTHORS: Nicola Richardson-Harman[1], Robert Parody[1,2], Peter Anton[3], Ian McGowan[4,5], Gustavo Doncel[6], Andrea Ries Thurman[6], Carolina Herrera[7], Kattayoun Kordy[8], Julie Fox[9], Karen Tanner[3], Glenn Swartz[10] and Charlene S. Dezzutti[4,5]

AFFILIATIONS:

[1]Alpha StatConsult, LLC, Damascus, MD, USA.

[2]Rochester Institute of Technology, School of Mathematical Sciences, Rochester, NY, USA.[3]Center for HIV Prevention Research, UCLA AIDS Institute, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA.

[4]University of Pittsburgh, School of Medicine, Pittsburgh, PA, USA.

[5]Magee Womens Research Institute, Pittsburgh, PA, USA.

[6]CONRAD, Eastern Virginia Medical School, Norfolk and Arlington, VA, USA.

[7]Division of Infectious Diseases, Faculty of Medicine, Imperial College, London, UK.

[8]Division of Pediatric Gastroenterology, Hepatology, Nutrition, University of Southern California, Los Angeles, CA, USA.

[9]Guys and St. Thomas' NHS Foundation Trust, London, UK.

[10]Advanced Bioscience Laboratories, Gaithersburg, MD, USA.

RUNNING TITLE: Advances in *ex vivo* challenge assay.

CORRESPONDING AUTHOR: Dr. Nicola Richardson-Harman, President, Alpha StatConsult, LLC. 25822 Bowman Acres Lane, Damascus, MD 20872. Email: nicola@alphastatconsult.com

KEYWORDS: HIV Prevention, Bioinformatics, Drug discovery, HIV.

ABSTRACT

(Currently 247 words, 250 max)

The *ex vivo* challenge assay is being increasingly used as an efficacy endpoint during early human clinical trials of HIV prevention treatments. There is no standard methodology for the *ex vivo* challenge assay although the use of different data collection methods and analytical parameters may impact results and reduce the comparability of findings between trials. In this analysis we describe the impact of data imputation methods, kit type, testing schedule and tissue type on variability, statistical power and *ex vivo* HIV growth kinetics. Data were p24 antigen (pg/mL) measurements collected from clinical trials of candidate microbicides where rectal (n=502), cervical (n=88) and vaginal (n=110) tissues were challenged with HIV-1_{BaL} *ex vivo*. Imputation of missing data using a non-linear mixed effect model was found to provide an improved fit compared to imputation using half the limit of detection. The rectal virus growth period was found to be earlier and of a relatively shorter duration than the growth period for cervical and vaginal tissue types. On average, only four rectal tissue challenge assays in each treatment and control group would be needed to find a one log difference in p24 to be significant (alpha = 0.05) but a larger sample size was predicted to be needed for either cervical (n=21) or vaginal (n=10) tissue comparisons. Overall, the results indicated that improvements could be made in the design and analysis of the *ex vivo* challenge assay to provide a more standardized and powerful assay to compare efficacy of microbicide products.

INTRODUCTION

The human immunodeficiency virus (HIV) infects predominantly through the mucosal tissue following sexual intercourse; therefore, the female genital tract (vaginal and cervical) as well as the rectal mucosae have been extensively studied. Consequently, HIV prevention products aim to prevent the sexual transmission of HIV-1. HIV-1 infection of human genital and rectal tract tissues biopsied from individuals following an *in vivo* HIV-1 prevention regimen has been used as a measure of drug efficacy and is referred to as the '*ex vivo* challenge assay.' In the *ex vivo* challenge assay, fresh tissue samples are infected with HIV-1 then, after a washout, HIV-1 growth is monitored in the tissue supernatant for up to 21 days post infection. Virus levels in tissue supernatant are tested every 1-4 days and low or no HIV-1 growth indicates treatment efficacy. There is currently no standard methodology for the *ex vivo* challenge assay although this assay is increasingly being used as an exploratory endpoint in Phase 1 and 2 clinical trials of candidate microbicides [1-6]. HIV-1 infection can be measured with HIV-1 RNA and DNA by real-time PCR [7-9] and fixing the tissue for immunohistochemistry to detect p24 expressing cells [10]. The p24 antigen release assay has been the most commonly used viral endpoint to measure *ex vivo* HIV-1 growth in rectal [11], cervical and vaginal tissue [10]. The purpose of this paper is to present an inter-laboratory retrospective analysis of *ex vivo* challenge p24 antigen release data to help improve and standardize *ex vivo* challenge assay methodology to allow for comparisons of treatment efficacy across laboratories, studies and tissue types.

Analytical measurements, such as the p24 antigen assay, have a limit of detection (LOD) that is the lowest concentration that can be determined to be statistically different from a blank measurement. For some assays a lower limit of quantification (LLOQ) is used as the lowest concentration measured with precision and accuracy. Values below the LOD or LLOQ are referred to as 'left censored', that is when a measurement below a lower limit is made, the concentration is unknown and reported as either

missing or zero. It has been shown that the logic underlying not reporting left censored values is flawed as distortions in values above the LOD can be worse than those below the LOD [12]. In addition, prediction of values below the LLOQ can provide more information than the mere statement that the value is less than the LLOQ [13]. Indeed, in the *ex vivo* challenge assay, left-censored values of low or no virus growth indicate successful virus suppression. As virus growth measurements are log-normally distributed [14] virus growth data are often log transformed to use standard, parametric statistical methods and so a data imputation method is needed to retain such instances of HIV suppression. Statistical methods that account for left-censored data have been developed particularly for measurement of environmental contaminants [15] and virological research [16]. Four imputation methods, including simple substitution and more complex model based approaches [17], were compared here with the goal of finding an optimum method for both ease of application and model fit for the *ex vivo* challenge assay.

In the *ex vivo* challenge assay, the frequency and sequence of testing days are not standardized but chosen by the laboratory to capture the time period of likely virus growth. Choice of the frequency and duration of testing days can be based on scientific and logistical factors especially if the assay is run within the context of a multi-site clinical trial. Ideally, fresh tissue is started in culture shortly after (e.g. 1-2 hours) biopsy collection [18] placing the burden of supernatant collection on a local laboratory team. Minimizing the number of testing timepoints and/or duration of the assay would increase the feasibility of running this assay. In this large retrospective analysis, virus growth kinetics were compared across laboratories and tissue types to determine the active virus growth periods in rectal, vaginal and cervical tissues during the *ex vivo* challenge assay.

The statistical power of an experiment is the likelihood that a study will detect a statistically significant effect when there is a true difference. Statistical power is driven by the size of the difference to be detected and the variability in the data. The statistical power of a treatment vs.

placebo/control comparison using the *ex vivo* challenge assay is affected by the virus endpoint, number of testing days collected and variability inherent in the tissue and assay methods used. Variability across different p24 endpoints has been compared. A cross-sectional index calculated from a growth curve reflective of the virus growth achieved in an assay (SOFT), a cumulative p24 endpoint (i.e. sum of all p24 measurements across the duration of the assay) and p24 on specific days were found to provide less measurement variability than the AUC (i.e. area under the virus growth curve) and slope of the virus growth curve [2, 14]. The cumulative p24 endpoint is now widely used as a readily calculable measurement that captures the total virus growth achieved in an assay [19-23]. The number of testing days used will have a direct impact on the value of the cumulative p24 endpoint as more frequent testing over longer assay durations will likely increase cumulative p24 measurements. In addition, variability in the cumulative p24 endpoint may not be equivalent across different tissue types or laboratory methods. The expected difference in cumulative p24 between a treatment and a control condition and the variability in these measurements will have a direct impact on statistical power. Statistical power was compared here for rectal, cervical and vaginal data to determine the number of tissue samples that would be needed, per treatment group, to find a one $\log_{10}$ difference in virus growth to be statistically significant.

Imputation methods, virus growth kinetics and statistical power were compared using a multi-study data set compiled from clinical trials of HIV microbicides where p24 measurements were collected during the *ex vivo* challenge assay [3-6, 10, 18, 24]. Only those tissue samples collected from non-treated subjects at baseline or following a placebo treatment were used to provide a large, homogeneous data set of non-drug treated *ex vivo* human tissues infected with HIV-1_{BaL} at approximately 10^4 $TCID_{50}$ (50% Tissue Culture Infective Dose; a measure of infectious virus titer).

METHODS

Data Sources and Inclusion Criteria

Data were p24 antigen (pg/mL) measurements collected from early phase clinical trials of candidate microbicides. The p24 antigen assay was used to measure HIV-1_{BaL} concentration post infection of biopsy tissues (rectal, vaginal and cervical) during the *ex vivo* challenge assay. The first *ex vivo* challenge Phase 1 experiment compared two infectious doses of HIV-1 BaL, 10^2 and 10^4 $TCID_{50}$: 100% of biopsies were infected at baseline with the higher titer whereas only about 60% were infected with the lower titer [3]. All data in this study were from tissue samples infected with 10^4 $TCID_{50.}$ Data were included in this study if the p24 measurements were from: (i) a rectal, vaginal, or cervical human tissue biopsy following *ex vivo* infection with approximately 10^4 $TCID_{50}$ HIV-1_{BaL}; (ii) a Phase 1or 2 clinical trial testing a candidate oral PrEP or microbicide product with *in vivo* treatment followed by *ex vivo* challenge assay; (iii) fresh (not frozen) tissue samples; and (iv) from a subject in a baseline, placebo or no treatment condition.

Ex vivo Challenge Assay

The assay methodologies used for the p24 data sets have been reported in detail elsewhere [3, 6, 10] and are summarized here in Tables 1a&b. Generally, the *ex vivo* challenge assays used tissue biopsies collected from human subjects by endoscopy (rectal biopsies) or direct biopsy (cervical or vaginal biopsies). Tissue samples were placed in a medium and transported to a laboratory and, within 1-3 hours following incubation or on ice overnight, samples were infected with HIV-1_{BaL} at approximately 10^4 $TCID_{50}$ and followed in culture for up to 21 days. During the culture period, supernatants for p24 quantification were collected every 1-4 days. IL-2 (50 U/mL; Roche Life Sciences, Indianapolis, IN) was added to the culture medium for data sets V1, V2, C1, C2 and C4. Concentrations of p24 in the supernatant were quantified using a range of ELISA kits (see Table 1a) where the LOD or LLOQ was provided by the laboratory or was the lowest non-zero p24 measurement for that study. Successive p24 concentrations (pg/mL), at each time point, were added to calculate *cumulative* p24 (pg/mL) [1]. The

cumulative p24 endpoint was not intended to be a measure of total p24 in the assay but, instead, a measure of the accumulation of successive p24 concentrations [23, 25-27].

Statistical Analysis

A total of 17 data sets were integrated using a "transform and recode" process [28] where all p24 measurements were transformed into pg/mL units, non-treated samples were defined as either placebo or baseline depending on the study design, assays from the same study but performed at different sites were coded as different data sets and data were selected only for those *ex vivo* tissue samples infected with approximately 10^4 $TCID_{50}$ HIV-1_{BaL}.

Three developmental (A-C) and one commonly used imputation method ('Z'; [17, 29]) were used to impute missing, zero or p24 measurements below the LOD or LLOQ. The commonly used method ('Z') imputes missing values with either ½ LLOQ or ½ LOD. The rationale for this imputation is that data below a lower cut-off (i.e., LLOQ or LOD) will have a normal distribution where the mean of the data falls halfway between the cut-off and zero. All developmental models (A-C) included a random effect to account for within subject repeatability. The first developmental imputation method ('A') used a non-linear mixed effect model [30-32] so that missing or zero results were imputed with predicted values from the model. Non-linear mixed effect models have been used to impute left censored values for longitudinal HIV measures of HIV-infected patients where RNA levels often drop below the LOD with highly active anti-retroviral treatments [29]. The second developmental imputation method ('B') took advantage of the virus growth kinetics during active infection where virus replication classically follows a non-linear s-shaped curve with gradual increase in virus before a vigorous growth phase followed by low or no additional growth [14]. In method B imputation was performed using a non-linear quadratic fit across days of culture and when enough detectable measurements (>3) where available for curve fitting, followed by method A for those tissue cultures that did not provide enough data to model virus growth curve. The third developmental method ('C')

combined methods A and B in an iterative fashion, randomly choosing single data values for imputation with method A after all possible data had been imputed with method B, thus reducing bias introduced into the non-linear mixed effect model by the order that values were imputed. To assess each method for variability, a summary measure called the sums of squares for imputation (SSI) was used. This is the sum of the squared differences between the measurements (imputed and detectible) and the predicted values from the model fit where a low SSI indicates a better fit.

A non-linear growth curve (Equation 1) was fit to each data set and the second derivative of each curve [33] was to identify the lower and upper inflection points as the beginning and end of the virus growth period. The virus growth periods were compared across tissue types and data sets.

$$\text{Log}_{10}\ p24 = a + (b-a)/ (1 + 10^{(c - Day)}) \qquad \text{Equation 1}$$

The number of tissue samples needed, per treatment group, to find a one $\log_{10}$ difference in cumulative p24 to be significant at alpha = 0.05, with 80% power by *t*-test was determined for each data set. Using this approach, a statistically significant one-log reduction in cumulative p24 was found in the *ex vivo* challenge assay following *in vivo* use of UC781 2.5% gel, a candidate microbicide [3] in the first Phase 1 pharmacodynamic study of *ex vivo* efficacy and Tenofovir 1% gel [6], in a subsequent Phase 1 pharmacodynamic study.

All statistical analyses were performed with SAS® v9.3 (Cary, North Carolina). The SAS® program written to perform and compare imputation methods A, B, C and Z is included in supplementary material.

RESULTS

HIV-1_{BaL} p24 antigen measurements were collected from a total of 700 tissue samples (Table 1a) where 1-3 biopsy samples, per donor, were entered into the analyses. Data sets were coded for each laboratory and study according to tissue type: rectal (R1-R8), cervical (C1-C4), and vaginal (V1-V5).

Comparison of p24 Imputation Methods

There was a pattern to the missing data where a greater proportion was found for the rectal data sets and at the earlier p24 testing time points (Table 1a). Imputation methods A, B, C, and Z were compared for data sets where at least 10 biopsies for *ex vivo* challenge assays were used, data were left-censored or missing and the assays included at least four p24 testing days, as the models were non-calculable when these requirements were not met. There were seven rectal data sets with missing or left censored data (R1-R5, R7 and R8; Table 1a), and of these data sets, R5 and R8 only collected p24 measurements on three testing days and R3 data set was comprised of only nine tissue samples (Table 1a). Following imputations, data were compared across methods A, B, C, and Z for sets R1, R2, R4, and R7 (Table 2). High SSI values indicated a poorer fit and more variability for imputation method Z compared to methods A-C for (bolded values in Table 2). A significant difference between the data sets was found when methods A, B and C were compared to method Z ($p<.0001$; Table 2). These results indicate that imputation using methods A-C provided both improved model fit and a different result outcome when compared to imputation method Z. Method A was the preferred method as imputations using this method could be readily calculated in commonly available software. Missing data were imputed with method A for the remainder of the analyses and the imputed data (open squares; Figure 1) were mostly within mean ±1 SD of the non-imputed data (filled squares, error bars; Figure 1).

Virus Growth in the *Ex vivo* Challenge Assay

Virus growth, as measured by the p24 assay, was modelled with a non-linear virus growth curve for each data set (Figure 1). There was considerable similarity in rectal virus growth curves across

laboratories and studies where active virus growth reached a p24 level of around 3 $\log_{10}$ (Figure 1a) and variability (vertical bars) was relatively low across the eight rectal data sets evaluated. Cervical virus growth reached a p24 level between 3-4 $\log_{10}$ (Figure 1b) and variability was markedly greater for the cervical compared to the rectal data sets (i.e. longer error bars). Vaginal tissue virus growth was variable both within and between data sets where virus growth reached between 1-4 $\log_{10}$ p24 (Figure 1c).

The active virus growth period was defined as the time span between the lower and upper inflection points of each curve ('Virus Growth'; Figure 2) where the first and last inflection points are indicated as the virus growth duration on the x axis. The virus growth period for rectal tissues was earlier (Days 6-8) and of a relatively shorter duration (3 days total) as compared to the cervical (Days 9-12) and vaginal (Days 9-15; Figure 2) tissues.

Statistical Power of the *Ex vivo* Challenge Assay

The *ex vivo* cervical, rectal and vaginal tissue protocols used various time points for up to 21 days post infection, several ELISA kits (Table 1a) and other within-laboratory potential sources of variability that were not measured but could nonetheless affect statistical power. For example, cumulative p24 would be expected to be higher when more testing time points were used: $\log_{10}$ mean cumulative p24 ranged from 4.13-4.57 pg/ml among cervical tissue experiments using the same six time points (C1, C2 and C4; Table 3) but dropped to 3.06 $\log_{10}$ mean cumulative p24 when only four time points were used (C3; Table 3). There was a trend for variability (i.e., $\log_{10}$ SDs) to be lower for the rectal data sets (0.19-0.54) compared to the cervical (0.85-1.30) and vaginal data sets (0.48-1.21; Table 3). The p24 kits used for each data set are listed in Table 1a and there was some evidence for potential differences in p24 results due to kit type: the V1 data set using the Perkin Elmer Alliance kit had a $\log_{10}$ mean cumulative p24 of 3.91 (3.7-4.1 95% CI) but, for the same testing days, the V2 data using the AlphaLISA kit had a $\log_{10}$ mean cumulative p24 of 4.31 pg/ml (4-4.6, 95% CI; Table 3). As

cumulative p24 could be affected by the number of days and interval of days of virus collection, the duration of the assay, tissue type, p24 kit and possibly other assay-related factors a power analysis was run per data set ('N per Group*'; Table 3). Statistical power was determined for comparisons of cumulative p24 between a non-drug treated condition (placebo and baseline) and an anticipated one $\log_{10}$ change in p24 following an *in vivo* drug treatment. The number of *ex vivo* challenge assays that would need to be performed to find a one $\log_{10}$ difference in p24 to be statistically significant was averaged across rectal, vaginal, and cervical tissue data sets (average N per Group in Table 3). On average, only four rectal tissue challenge assays in each treatment and control group would be needed to find a one $\log_{10}$ difference in p24 following a drug treatment when using rectal tissue in the *ex vivo* challenge assay. This relatively low number of tissue samples needed for 80% power was due to the low variability in the rectal tissue virus growth data (0.19-0.54 Log_{10} SD; Table 3). A larger sample of vaginal *ex vivo* tissues (n=10; Table 3) would be needed to find this same one $\log_{10}$ difference in cumulative p24 to be different between a treatment and control condition. The cervical tissue data were the most variable with $\log_{10}$ SDs ranging between 0.85-1.3 (Table 3) and this was reflected in the relatively larger number of cervical *ex vivo* tissue samples (n=21) that would be needed to find this same one $\log_{10}$ difference to be statistically significant.

DISCUSSION

The results of this retrospective analysis provide evidence to support a number of practical guidelines for conducting *ex vivo* challenge assays related to choice of: (i) imputation methods, (ii) testing days, and (iii) number of tissue samples to be used.

All three novel imputation methods tested here provided an improved model fit compared to the ubiquitous ½ LOD type methods. There was a pattern to the missing data; occurring only in the rectal

data sets and mostly at the early time points where low or no virus growth had yet occurred. A model based approach has been recommended for such 'missing not at random' data [34], especially when 50% or more data are missing [35]. Replacement of non-detected measurements with ½ LOD is considered acceptable when less than 15% data are missing [35]. Clearly, practical concerns will play an important role when choosing a data imputation method as ease of computation is important for any method that is to be routinely used by a scientific team. The non-linear mixed effect model was considered the simplest of the novel imputation methods tested, where imputations could be made with commonly available software (please see free imputation software at http://www.alphastatconsult.com//). Although the imputation methods were tested with p24 assay data sets, the non-linear mixed effect model here could equally be used to impute missing data collected from other HIV-1 strains or biomedical analytical procedures, such as RNA, cytokine and chemokine quantification.

Periods of active virus growth were found to vary across tissue types with the rectal tissue providing an early short growth period, cervical tissue providing a later growth period and vaginal tissue providing a later and longer growth period. The transformation zone is the area between the ectocervix and endocervix where the epithelium changes from stratified squamous to columnar. While there are resident dendritic cells, macrophage, natural killer cells and lymphocytes throughout the female genital tract, the transformation zone typically has the highest concentration of immune cells [36]. In contrast, the gastrointestinal mucosa contains the majority of the body's CD4+ lymphocyte population and likely represents the largest reservoir of HIV and site of viral replication [37]. Rectal subepithelial stromal tissues are densely populated with organized lymphoid tissue, dendritic cells, macrophages and T cells all susceptible to HIV infection whereas female genital tract tissue is less well defined with a higher density of immune cells and cervical columnar epithelial cells that produce mucus and anti-microbial proteins [38]. Differences have been found between rectal and genital tissue types in efflux transporter mRNA where OAT1 protein was detected in 100% of rectal tissues but not female genital tissues [39]. Such

anatomical and functional differences between rectal and female genital sites could account for the differences found here in HIV kinetics where the rectal tissue displayed rapid viral replication with less variable kinetics compared to cervical and vaginal tissue. Changes in the woman's menstrual cycle, contraception method and other cervicovaginal factors could impact infectivity in the *ex vivo* challenge assay [40, 41]. The finding of differences in *ex vivo* HIV growth support the standardization of p24 testing days per tissue type to allow direct comparisons in cumulative p24 to be made across studies and testing sites. Defining the active virus growth period and maximum levels of growth obtained using such a large body of data for each tissue type will allow research teams to choose and standardize the p24 testing days to provide a strong and reliable p24 signal in the baseline or placebo conditions to compare to expected HIV suppression in the active drug treatment arms.

Cervical and vaginal tissue assays present more challenges to the researcher due to greater variability in virus replication (cumulative p24) and kinetics. This variability in kinetics for the cervical and most markedly the vaginal tissue may hamper efforts to shorten the duration of this assay with active growth found out to 15 days although sampling time points beyond 15 days would not be predicted to capture any significant additional growth. This study demonstrates that the variability in the tissue *ex vivo* virus growth can have a direct impact on the power of placebo-controlled, microbicide trials using this exploratory efficacy endpoint. The *ex vivo* challenge assay has been an exploratory endpoint in Phase 1 and 2 clinical trials powered for the primary endpoints of pharmacokinetics, safety and acceptability resulting in the inclusion of predominantly males as rectal microbicides products were tested [3, 4, 6]. As the female participants in these studies usually provided both genital and rectal tissue, there can be over twice as many rectal tissue *ex vivo* data compared to cervical and vaginal tissue, per study. In addition, there are practical limitations to the number of mucosal biopsies that can be collected. Using flexible sigmoidoscopy it is possible to collect 20-30 mucosal biopsies and consequently there are no limitations on the numbers of biopsies that can be used for the *ex vivo* challenge assay. In contrast,

typically only 2-4 cervical or vaginal biopsies are collected. This problem is exacerbated by competing needs for biopsies that include measurement of drug concentration in mucosal biopsies. The finding that more vaginal and cervical data are needed to provide the equivalent statistical power as the rectal data is the exact opposite of the balance of *ex vivo* data that has been generated in recent studies. Rectal data sets are, therefore, more likely to be powered to find significant differences between treatment and control conditions and, conversely, a lack of significant treatment effects in cervical and vaginal tissue may be due to insufficient statistical power.

A difficulty that was not addressed in this paper was the effect of the various treatment regimens on the kinetics of virus growth as only non-treated (baseline or placebo) data were used in the analysis. This decision was based on the need to more fully understand how the untreated virus grows in the different tissue models so that suppression of growth, as would be expected from an efficacious microbicide, could be identified in the non-treated condition. The findings of this analysis are specific to the HIV-1BaL virus type used in this retrospective analysis of Phase 1 and 2 clinical trials of HIV prevention treatments. Although other virus isolates, for example Transmitter/Founder isolates and HIV-1 variants have been found to be equally infective in *ex vivo* cervical tissue [23], they have not been used for the *ex vivo* challenge assay as of yet. In addition, as assay parameters were not independently varied but a result of the choices made by each scientific team, factors other than those included here (e.g. testing days, p24 kits, tissue type, treatment) may have affected the p24 results in ways that were not discovered during this retrospective analysis.

A key feature of this paper is to demonstrate the utility of a retrospective analysis of data from multiple trials to improve assay methodology and statistical power in treatment comparisons. Overall, results indicated that improvements could be made in the design and analysis of the *ex vivo* challenge

assay to provide a more standardized and powerful tool to compare efficacy of oral PrEP and microbicide products designed to suppress HIV-1_{BaL} infection.

ACKNOWLEDGEMENTS

The authors thank Julie Elliott, Louise A. Ouattara, Kevin Uranker, Natalia Olejniczak and Jarret Engstrom for performing the *ex vivo* challenge assays and providing technical details of the assays for this manuscript.

This analysis was supported by a subcontract with Advanced BioScience Laboratories, Inc., Rockville, MD through a NIH/NIAID/DAIDS contract: "Comprehensive Resources for HIV Microbicides and Biomedical Prevention" (#HHSN272201000001C).

The authors would like to acknowledge the support of the following grants: NIH U19 CHARM-01 (U19AI082637), Bill & Melinda Gates Foundation (OPP1045325), NIH U19 MDP (#U19AI06061), UCLA's NIH CFAR Mucosal Immunology and Seed Grant Cores (#AI28697) and CTRC (#UL1TR000124), NIH FAME (U19 AI082637), an intra-agency agreement between the NIH/NIAID/DAIDS, USAID/OHA (NIAID Y1-AI-1756-01) and CONRAD (GPO-A-00-08-00005-00), and ViiV contract WS2184024.

The Microbicide Trials Network is funded by the National Institute of Allergy and Infectious Diseases (UM1AI068633, UM1AI068615, UM1AI106707), with co-funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Institute of Mental Health, all components of the U.S. National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Finally, the authors would like to thank all of the clinical trial participants and the clinical team members, without whom this work would not have been possible.

Please address reprint requests to: Dr. Nicola Richardson-Harman, President, Alpha StatConsult, LLC. 25822 Bowman Acres Lane, Damascus, MD 20872. Email: nicola@alphastatconsult.com

REFERENCES

1. Richardson-Harman N, Hendrix CW, Bumpus NN, Mauck C, Cranston RD, Yang K*, et al.* Correlation between compartmental tenofovir concentrations and an ex vivo rectal biopsy model of tissue infectibility in the RMP-02/MTN-006 phase 1 study. *PloS One* 2014,**9**:e111507.
2. Richardson-Harman N, Mauck C, McGowan I, Anton P. Dose-response relationship between tissue concentrations of UC781 and explant infectibility with HIV type 1 in the RMP-01 rectal safety study. *AIDS Res Hum Retroviruses* 2012,**28**:1422-1433.
3. Anton PA, Saunders T, Elliott J, Khanukhova E, Dennis R, Adler A*, et al.* First phase 1 double-blind, placebo-controlled, randomized rectal microbicide trial using UC781 gel with a novel index of ex vivo efficacy. *PLoS One* 2011,**6**:e23243.
4. McGowan I, Cranston RD, Duffill K, Siegel A, Engstrom JC, Nikiforov A*, et al.* A Phase 1 Randomized, Open Label, Rectal Safety, Acceptability, Pharmacokinetic, and Pharmacodynamic Study of Three Formulations of Tenofovir 1% Gel (the CHARM-01 Study). *PloS One* 2015,**10**:e0125363.
5. Chen BA, Panther L, Marzinke MA, Hendrix CW, Hoesley CJ, van der Straten A*, et al.* Phase 1 Safety, Pharmacokinetics, and Pharmacodynamics of Dapivirine and Maraviroc Vaginal Rings: A Double-Blind Randomized Trial. *J Acquir Immune Defic Syndr* 2015,**70**:242-249.
6. Anton PA, Cranston RD, Kashuba A, Hendrix CW, Bumpus NN, Richardson-Harman N*, et al.* RMP-02/MTN-006: A phase 1 rectal safety, acceptability, pharmacokinetic, and pharmacodynamic study of tenofovir 1% gel compared with oral tenofovir disoproxil fumarate. *AIDS Res Hum Retroviruses* 2012,**28**:1412-1421.
7. Nicol MR, Emerson CW, Prince HM, Nelson JA, Fedoriw Y, Sykes C*, et al.* Models for predicting effective HIV chemoprevention in women. *J Acquir Immune Defic Syndr* 2015,**68**:369-376.
8. Fletcher PS, Elliott J, Grivel JC, Margolis L, Anton P, McGowan I*, et al.* Ex vivo culture of human colorectal tissue for the evaluation of candidate microbicides. *AIDS* 2006,**20**:1237-1245.
9. Janocko L, Althouse AD, Brand RM, Cranston RD, McGowan I. The Molecular Characterization of Intestinal Explant HIV Infection Using Polymerase Chain Reaction-Based Techniques. *AIDS Res Hum Retroviruses* 2015,**31**:981-991.
10. Dezzutti CS, Uranker K, Bunge KE, Richardson-Harman N, Macio I, Hillier SL. HIV-1 infection of female genital tract tissue for use in prevention studies. *J Acquir Immune Defic Syndr* 2013,**63**:548-554.
11. Fischetti L, Barry SM, Hope TJ, Shattock RJ. HIV-1 infection of human penile explant tissue and protection by candidate microbicides. *AIDS* 2009,**23**:319-328.

12. Guo Y, Harel O, Little RJ. How well quantified is the limit of quantification? *Epidemiology* 2010,**21 Suppl 4**:S10-16.
13. Succop PA, Clark S, Chen M, Galke W. Imputation of data values that are less than a detection limit. *J Occup Environ Hyg* 2004,**1**:436-441.
14. Richardson-Harman N, Lackman-Smith C, Fletcher PS, Anton PA, Bremer JW, Dezzutti CS*, et al.* Multisite comparison of anti-human immunodeficiency virus microbicide activity in explant assays using a novel endpoint analysis. *J Clin Microbiol* 2009,**47**:3530-3539.
15. Slymen DJ, de Peyster A, Donohoe RR. Hypothesis Testing with Values below Detection Limit in Environmental Studies. *Environ Sci Technol* 1994,**28**:898-902.
16. Moulton LH, Halsey NA. A mixture model with detection limits for regression analyses of antibody response to vaccine. *Biometrics* 1995,**51**:1570-1578.
17. Krishnamoorthy K, Mallick A, Mathew T. Model-based imputation approach for data analysis in the presence of non-detects. *Ann Occup Hyg* 2009,**53**:249-263.
18. Dezzutti CS. Impact of cryopreservation on pharmacodynamic correlations following use of vaginal rings containing dapivirine and/or maraviroc. *Journal of Acquired Immune Deficiency Syndromes* Submitted.
19. Chinnadurai R, Rajan D, Munch J, Kirchhoff F. Human immunodeficiency virus type 1 variants resistant to first- and second-version fusion inhibitors and cytopathic in ex vivo human lymphoid tissue. *J Virol* 2007,**81**:6563-6572.
20. Cobos Jimenez V, Booiman T, de Taeye SW, van Dort KA, Rits MA, Hamann J*, et al.* Differential expression of HIV-1 interfering factors in monocyte-derived macrophages stimulated with polarizing cytokines or interferons. *Sci Rep* 2012,**2**:763.
21. Geuenich S, Kaderali L, Allespach I, Sertel S, Keppler OT. Biological signature characteristics of primary isolates from human immunodeficiency virus type 1 group O in ex vivo human tonsil histocultures. *J Virol* 2009,**83**:10494-10503.
22. Heigele A, Camerini D, van't Wout AB, Kirchhoff F. Viremic long-term nonprogressive HIV-1 infection is not associated with abnormalities in known Nef functions. *Retrovirology* 2014,**11**:13.
23. Merbah M, Arakelyan A, Edmonds T, Ochsenbauer C, Kappes JC, Shattock RJ*, et al.* HIV-1 expressing the envelopes of transmitted/founder or control/reference viruses have similar infection patterns of CD4 T-cells in human cervical tissue ex vivo. *PloS One* 2012,**7**:e50839.
24. Thurman AR, Chandra, N., Yousefieh, N., Zalenskaya, Z., Kimble, T., Asin, S., Rollenhagen, C., Anderson, S., Herold, B., Mesquita, P.M.M., Richardson-Harman, N., Cunningham, T., Schwartz, J., Doncel, G. Comparison of Follicular and Luteal Phase Mucosal Markers of HIV Susceptibility in Healthy Women. *AIDS Research and Human Retroviruses* in press.
25. Eckstein DA, Penn ML, Korin YD, Scripture-Adams DD, Zack JA, Kreisberg JF*, et al.* HIV-1 actively replicates in naive CD4(+) T cells residing within human lymphoid tissues. *Immunity* 2001,**15**:671-682.
26. Kmiec D, Iyer SS, Sturzel CM, Sauter D, Hahn BH, Kirchhoff F. Vpu-Mediated Counteraction of Tetherin Is a Major Determinant of HIV-1 Interferon Resistance. *MBio* 2016,**7**.
27. Yang H, Yorke E, Hancock G, Clutton G, Sande N, Angus B*, et al.* Improved quantification of HIV-1-infected CD4+ T cells using an optimised method of intracellular HIV-1 gag p24 antigen detection. *J Immunol Methods* 2013,**391**:174-178.
28. Carrig MM, Manrique-Vallier D, Ranby KW, Reiter JP, Hoyle RH. A Nonparametric, Multiple Imputation-Based Method for the Retrospective Integration of Data Sets. *Multivariate Behav Res* 2015,**50**:383-397.
29. Jacqmin-Gadda H, Thiebaut R, Chene G, Commenges D. Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics* 2000,**1**:355-368.

30. Vonesh EFaC, V. M. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker; 1997.
31. Davidian MaG, D. M. *Nonlinear Models for Repeated Measurement Data*. New York: Chapman & Hal; 1995.
32. Wu H, Wu L. Identification of significant host factors for HIV dynamics modelled by non-linear mixed-effects models. *Stat Med* 2002,**21**:753-771.
33. Thomas TJ. Finding Usable Portion of Sigmoid Curve. In. Edited by Office USP; 2008. pp. 27.
34. Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol* 2009,**60**:549-576.
35. Agency EP. Guidance for Data Quality Assessment. In. Edited by Information OoE. Washington DC: Environmental Protection Agency; 2000. pp. 153-165.
36. Pudney J, Quayle AJ, Anderson DJ. Immunological microenvironments in the human vagina and cervix: mediators of cellular immunity are concentrated in the cervical transformation zone. *Biol Reprod* 2005,**73**:1253-1263.
37. Mowat AM, Viney JL. The anatomical basis of intestinal immunity. *Immunol Rev* 1997,**156**:145-166.
38. McGowan I, Dezzutti C. Rectal microbicide development. *Curr Top Microbiol Immunol* 2014,**383**:117-136.
39. Nicol MR, Fedoriw Y, Mathews M, Prince HM, Patterson KB, Geller E*, et al.* Expression of six drug transporters in vaginal, cervical, and colorectal tissues: Implications for drug disposition in HIV prevention. *J Clin Pharmacol* 2014,**54**:574-583.
40. Thurman AR, Chandra N, Yousefieh N, Zalenskaya I, Kimble T, Asin S*, et al.* Comparison of Follicular and Luteal Phase Mucosal Markers of HIV Susceptibility in Healthy Women. *AIDS Res Hum Retroviruses* 2016,**32**:547-560.
41. Thurman AR. Comparison of Mucosal Markers of HIV Susceptibility in Healthy Premenopausal versus Postmenopausal Women. *Journal of Acquired Immune Deficiency Syndromes* 2016 (submitted).

Figure Legends

Figure 1. Tissue virus growth in *ex vivo* challenge assays for rectal (a), cervical (b) and vaginal (c) data sets. Log_{10} p24 means and SD are indicated for non-imputed (filled squares, ± 1 SD) and imputed (open circles) data. Missing data were imputed using a non-linear mixed effect model (Method 'A'). The solid black line indicates a non-linear growth curve model fit to each data set.

Figure 2. Non-linear growth curve models for rectal, cervical and vaginal data sets. The first and last inflection points of the non-linear curves, per tissue type, are indicated on the x axis as the beginning and end of the virus growth period.

Table Legends

Table 1. Ex Vivo Challenge Assay Methods for Rectal (R1–R8), Cervical (C1–C4), and Vaginal (V1–V5) Data Sets: Treatments, Kits, and p24 Testing Days.

Table 2. Ex Vivo Challenge Assay Methods for Rectal (R1–R8), Cervical (C1–C4), and Vaginal (V1–V5) Data Sets: Supernatant Volume, HIV-1 BaL Source, Tissue and Forceps Details.

Table 3. Comparison of imputation methods for missing or non-detected p24 measurements.

Table 4. Summary statistics and power analysis results for p24 (pg/mL) rectal, cervical, and vaginal data sets.

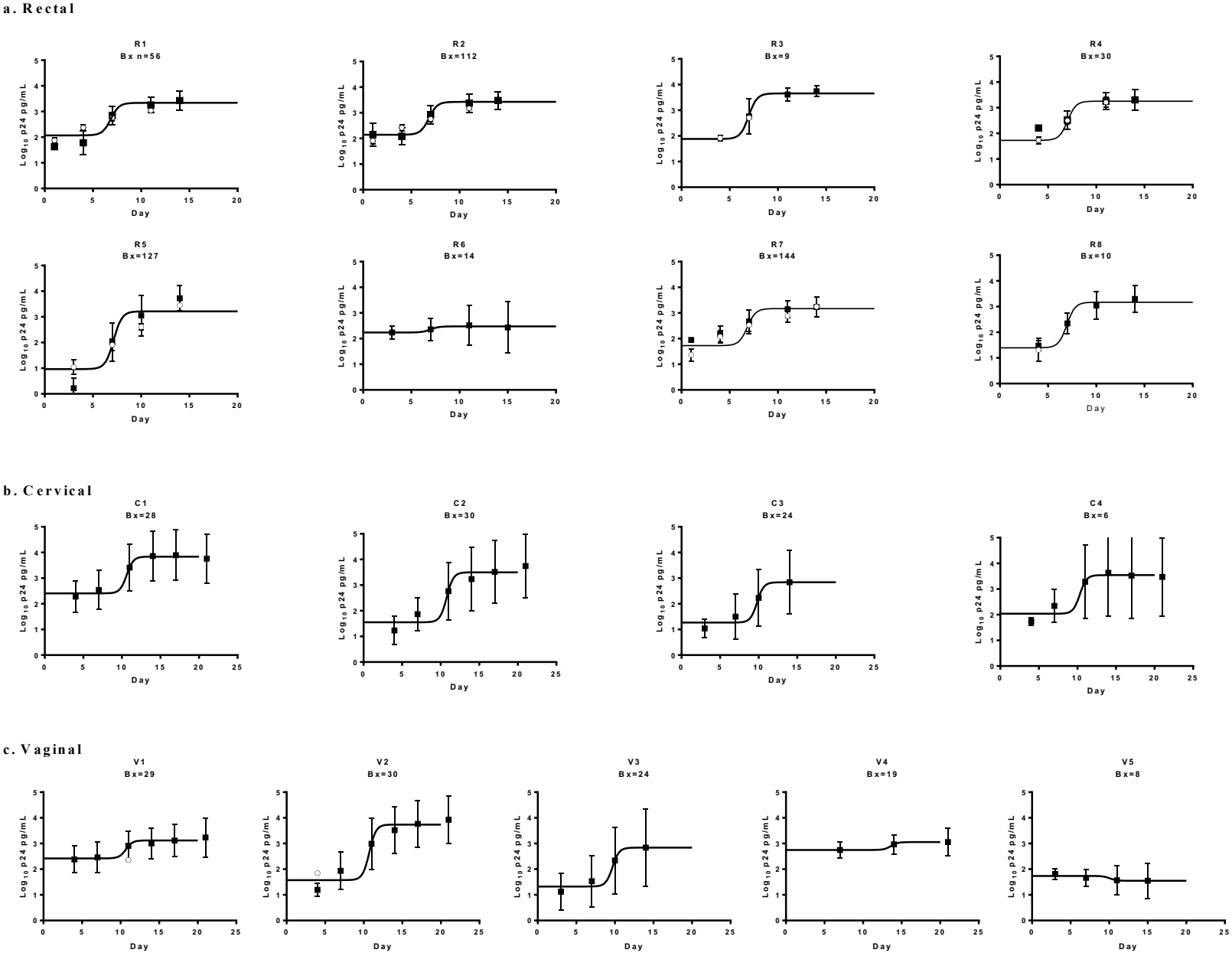
a. Rectal
R1
Bx n=56
R2
Bx=112
R3
Bx=9
R4
Bx=30
R5
Bx=127
R6
Bx=14
R7
Bx=144
R8
Bx=10
b. Cervical
C1
Bx=28
C2
Bx=30
C3
Bx=24
C4
Bx=6
c. Vaginal
V1
Bx=29
V2
Bx=30
V3
Bx=24
V4
Bx=19
V5
Bx=8
Log10 p24 pg/mL
Day

Figure 1

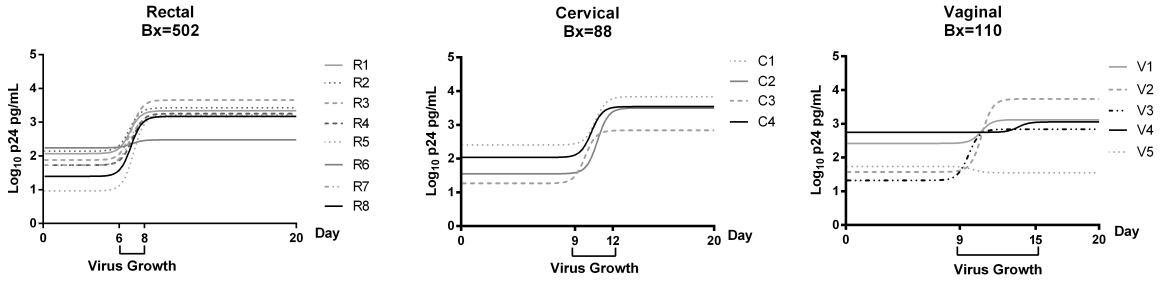
Rectal
Bx=502
Log10 p24 pg/mL
Day
Virus Growth
R1
R2
R3
R4
R5
R6
R7
R8
Cervical
Bx=88
C1
C2
C3
C4
Vaginal
Bx=110
V1
V2
V3
V4
V5

Figure 2.

Table 1

Data Set	Number *ex vivo* Samples[a] (Donors)	Treatment(s)	p24 Kit	p24 Testing Days (% missing)
R1	56 (24)	baseline, placebo	National Cancer Institute[b]	1 (93%), 4 (86%), 7 (34%), 11 (4%), 14 (0%)
R2	112 (48)	baseline, placebo	National Cancer Institute	1 (54%), 4 (54%), 7 (6%), 11 (1%), 14 (0%)
R3	9 (3)	baseline	National Cancer Institute	4 (100%), 7 (44%), 11 (0%), 14 (0%)
R4	30 (10)	baseline	National Cancer Institute	4 (90%), 7 (40%), 11 (10%), 14 (0%)
R5	127 (32)	baseline	Alliance[c]	3 (89%), 7 (52%), 14 (24%)
R6	14 (4)	baseline	Zeptometrix[d]	3 (0%), 7 (0%), 11 (0%), 15 (0%)
R7	144 (24)	placebo	National Cancer Institute	1 (99%), 4 (97%), 7 (51%), 11 (13%), 14 (6%)
R8	10 (5)	baseline	National Cancer Institute	4 (40%), 7 (0%), 10 (0%), 14 (0%)
C1	28 (28)	baseline	Alliance	4 (0%), 7 (0%), 11 (0%), 14 (0%), 17 (0%), 21 (0%)
C2	30 (30)	placebo	AlphaLISA[b]	4 (0%), 7 (0%), 11 (0%), 14 (0%), 17 (0%), 21 (0%)
C3	24 (24)	baseline	Alliance	3 (0%), 7 (0%), 10 (0%), 14 (0%)
C4	6 (6)	placebo	Alliance	4 (0%), 7 (0%), 11 (0%), 14 (0%), 17 (0%), 21 (0%)
V1	29 (28)	baseline	Alliance	4 (0%), 7 (0%), 11 (0%), 14 (0%), 17 (0%), 21 (0%)
V2	30 (30)	placebo	AlphaLISA	4 (0%), 7 (0%), 11 (0%), 14 (0%), 17 (0%), 21 (0%)
V3	24 (24)	baseline	Alliance	3 (0%), 7 (0%), 10 (0%), 14 (0%)
V4	19 (19)	baseline	Alliance	7 (0%), 14 (0%), 21 (0%)
V5	8 (2)	baseline	Zeptometrix	3 (0%), 7 (0%), 11 (0%), 15 (0%)

[a]700 total tissue samples

[b]Frederick National Laboratory for Cancer Research, Frederick, MD .

[c]Perkin Elmer, Waltham , Massachusetts.

[d]Zeptometrix, Buffalo, NY.

All cultures were non-polarized. Tissue set-up within 1 hour with the exception of V4, which was on ice overnight.

Table 2

Data Set	Supernatant Vol (ul)	HIV-1 BaL Source	Mean wt/size (pre/post)	Forceps Brand, Manufacturer	Forceps Size
R1[a]	400	NIH, Catalogue#510	20-30 mg (pre)	Radial Jaw 4, Boston Scientific	3.8 mm
R2[a]	400	NIH, Catalogue#510	20-30 mg (pre)	Radial Jaw 4, Boston Scientific	3.8 mm
R3[a]	500	Advanced	15-20 mg (pre)	Radial Jaw 4, Boston Scientific	2.8 mm
R4[a]	500	Advanced	15-20 mg (pre)	Radial Jaw 4, Boston Scientific	2.8 mm
R5[a]	500	Advanced	15-20 mg (pre)	Radial Jaw 4, Boston Scientific	2.8 mm
R6[a]	200	NIH	3x3x1 mm (pre)	Sarratt, Stericom	4 mm
R7[a]	400	NIH, Catalogue#510	20-30 mg (pre)	Radial Jaw 4, Boston Scientific	3.8 mm
R8[b]	500	NIH, Catalogue#510	20-30 mg (pre)	Radial Jaw 4, Boston Scientific	3.8 mm
C1[c]	700	Advanced	9-30 mg (post)	Tischler, BD	2.3x4.2 mm
C2[c]	700	Advanced	9-30 mg (post)	Tischler, BD	2.3x4.2 mm
C3[a]	500	Advanced	15-20 mg (pre)	Radial Jaw 4, Boston Scientific	2.8 mm
C4[c]	700	Advanced	4-13 mg (post)	Tischler, BD	2.3x4.2 mm
V1[c]	700	Advanced	11-26 mg (post)	Tischler, BD	2.3x4.2mm
V2[c]	700	Advanced	11-26 mg (post)	Tischler, BD	2.3x4.2 mm
V3[a]	500	Advanced	15-20 mg (pre)	Radial Jaw 4, Boston Scientific	2.8 mm
V4[d]	100	NIH	30 mg (pre)	Tischler, BD	3x5 mm
V5[a]	100	NIH	3x3x1 mm (pre)	Sarratt, Stericom	4 mm

HIV-1 BaL Sources: NIH=NIH AIDS Research & Reference Reagent Program, Bethesda, MD (http://www.aidsreagent.org/)
Advanced=Advanced Biotechnologies, Inc. (https://abionline.com/)
$TCID_{50}$ titration methods were:
[a] http://www.ncbi.nlm.nih.gov/pubmed/21116800
[b]Titrated on activated PBMC. Calculated by Reed-Muench Formula.
[c]Reed LJ, Muench H. A Simple method of estimating fifty per cent endpoints. Am J Hygiene. 1938;27(3):493-7.
[d]Kaerber G: Beitrag zur Kollektiven Behandlung Pharmakologischer Reihenversuche. Arch Exp Path Pharma 1931, 162:480-7.

Data Set	Parameter	Data Imputation Method[d]			
		A	B	C	Z
R1[a]	Difference to Z	<0.0001	<0.0001	<0.0001	na
	SSI[c]	76	91	89	**516**
R2[a]	Difference to Z	<0.0001	<0.0001	<0.0001	na
	SSI	325	434	434	**1223**
R4[a]	Difference to Z	<0.0001	<0.0001	<0.0001	na
	SSI	41	41	50	**173**
R7[b]	Difference to Z	<0.0001	<0.0001	<0.0001	na
	SSI	208	259	405	**1863**

[a] Differences for A vs. B, A vs. C and B vs. C non-significant.

[b] Differences for A vs. B and A vs. C significant at $p<0.05$, B vs. C non-significant

[c]SSI = sum of squares for imputation. This was the sum of the squared differences between the measurements (imputed and detectible) and the predicted values from the model fit where a low SSI indicates a better fit.

[d]Imputation Methods A = missing values were imputed with those predicted from a non-linear mixed effect model, B =missing values were imputed with a non-linear quadratic fit of virus growth across days of culture for assays with > 3 testing timepoints and a non-linear mixed effect model for assays with ≤ 3 testing time points. C = missing values were imputed with a combination of methods A and B in an iterative fashion and Z = missing values were imputed with ½ LOD.

Tissue	Data Set	p24 Sampling Days	Biopsy (n)	Cumulative p24	SD	95% CI	N per Group*
Rectal	R1	1, 4, 7, 11 & 14	56	3.77	0.27	3.7-3.8	3
	R2	1, 4, 7, 11 & 14	112	3.85	0.29	3.8-3.9	3
	R3	4, 7, 11 & 14	9	4.04	0.19	3.9-4.2	3
	R4	4, 7, 11 & 14	30	3.67	0.28	3.6-3.8	3
	R5	3, 7 & 14	127	3.77	0.47	3.7-3.9	5
	R6	3, 7, 11 & 15	14	3.17	0.54	2.9-3.5	6
	R7	1, 4, 7, 11 & 14	144	3.57	0.32	3.5-3.6	3
	R8	4, 7, 10 & 14	10	3.54	0.50	3.2-3.9	6
Average				***3.67***			***4***
Cervical	C1	4, 7, 11, 14, 17 & 21	28	4.57	0.85	4.2-4.9	13
	C2	4, 7, 11, 14, 17 & 21	30	4.13	1.15	3.7-4.6	22
	C3	3, 7, 10 & 14	24	3.06	1.08	2.6-3.5	20
	C4	4, 7, 11, 14, 17 & 21	6	4.32	1.30	2.9-5.7	28
Average				***4.02***			***21***
Vaginal	V1	4, 7, 11, 14, 17 & 21	29	3.91	0.48	3.7-4.1	5
	V2	4, 7, 11, 14, 17 & 21	30	4.31	0.88	4-4.6	14
	V3	3, 7, 10 & 14	24	3.18	1.21	2.7-3.7	24
	V4	7, 14 & 21	19	3.48	0.40	3.3-3.7	4
	V5	3, 7, 11 & 15	8	2.32	0.43	2-2.7	5
Average				***3.44***			***10***

*The number of *ex vivo* samples needed to provide 80% statistical power to detect a one $\log_{10}$ difference in cumulative p24 between treatment groups to be statistically significant at alpha = 0.05.

SAS CODE FOR COMPARISON OF IMPUTATION METHODS.

ANALYTICAL ADVANCES IN THE *EX VIVO* CHALLENGE EFFICACY ASSAY.

```
%macro expp24(exp);

*Comment 1 - Read the data in and stack the data based on day;

data &exp;
set "data location";
IDKey = _n_;
longp24 = day_1;  day = 1;  subject = IDKey; type=tissue; output;
longp24 = day_3;  day = 3;  subject = IDKey; type=tissue; output;
longp24 = day_4;  day = 4;  subject = IDKey; type=tissue; output;
longp24 = day_7;  day = 7;  subject = IDKey; type=tissue; output;
longp24 = day_10;  day = 10;  subject = IDKey; type=tissue; output;
longp24 = day_11; day = 11; subject = IDKey; type=tissue; output;
longp24 = day_14; day = 14; subject = IDKey; type=tissue; output;
longp24 = day_15; day = 15; subject = IDKey; type=tissue; output;
longp24 = day_17; day = 17; subject = IDKey; type=tissue; output;
longp24 = day_21; day = 21; subject = IDKey; type=tissue; output;
longp24 = day_0;  day = 1;  subject = IDKey; type=tissue; output;
run;

*End Comment 1;


*Comment 2 - Subset the data based on the experiment of interest;


data &exp;
set &exp;

exp_data = 0;
if exp = "&exp" then exp_data = 1;
if exp_data = 0 then delete;
Treat = translate(Treat,' ','_');
run;

*End Comment 2;

*Comment 3 - Segregate the data based on the key words (Placebo, control,
etc.) within the code;

data &exp;
set &exp;
Base_Data = 0;
run;

data &exp;
set &exp;
if Base_Data = 0 then delete;
```

```
run;

*End comment 3;

*Comment 4 - Determine the LOD from the data and replaces the unmeasurable
values with it;

proc means data = &exp;
var LOD;
output out=LOD_&exp mode=LOD_mode;
run;

data cutoff_&exp;
set LOD_&exp;
if  LOD_mode = "." then LOD_check = 0;
else LOD_check = LOD_mode;
run;



proc sql;
select LOD_check
into: LOD_value
from cutoff_&exp;
quit;


data &exp;
set &exp;
if longp24=0 then delete;
run;

proc means data=&exp;
var longp24;
output out=lod_&exp min=minlod;
run;


proc sql;
select minlod
into: LOD_min
from lod_&exp;
quit;

data &exp;
set &exp;
if LOD = "." then LOD = &LOD_min;
run;

proc means data = &exp;
var LOD;
output out=LOD_&exp mode=LOD_mode;
run;

data cutoff_&exp;
set LOD_&exp;
```

```
if  LOD_mode = . then LOD_check = &LOD_min;
else LOD_check = LOD_mode;
run;

*End Comment 4;



*Comment 5 - Use the LOD and prepare the data for modelling;

proc sql;
select LOD_check
into: LOD_value
from cutoff_&exp;
quit;




data modellong_&exp;
set &exp;
if longp24 >= 0 then impute = 1;
if longp24 = . then impute = 0;
if longp24 > &LOD_value then logp24 = log(longp24);
if longp24 <= &LOD_value then logp24 = .;
daysq = day**2;
if logp24 = . then event = 0;
if logp24 ne . then event = 1;
run;

data modellong1_&exp;
set modellong_&exp;
if impute = 0 then delete;
run;

*End Comment 5;

**Create the naive pooled model results for the imputation (Method A);


*Comment 6 - Get the starting values for the nonlinear mixed effects model;

proc sort data=modellong1_&exp;
by subject;
run;

ods output ParameterEstimates = beta1_&exp;
proc reg data = modellong1_&exp;
model logp24 = daysq day;
run;
quit;

proc transpose data = beta1_&exp out=beta2_&exp;
id variable;
var Estimate;
run;
```

```
proc sql;
select intercept
into: int
from beta2_&exp;
quit;

proc sql;
select day
into: linear
from beta2_&exp;
quit;

proc sql;
select daysq
into: sqr
from beta2_&exp;
quit;

*End Comment 6;


*Comment 7 - Use the starting values to run the nonlinear mixed effects model
and make predictions;

proc nlmixed data = modellong1_&exp;
parms b0=&int b11=&sqr b1=&linear s2=0.4 s2u=0.4;
pred = b0+b1*day+b11*daysq+ui;
model logp24 ~ normal(pred,s2);
random ui~normal(0,s2u) subject=subject;
predict pred out=nlout_&exp;
run;
quit;

*End Comment 7;

*Comment 8 - Use the results to replace the imputed values with the predicted
values;

data nlout1_&exp;
set nlout_&exp;
if event = 0 then newlogp24 = pred;
if event = 1 then newlogp24 = logp24;
**diffMethodA = (logp24 - pred)**2;
drop upper lower alpha probt df StdErrPred tValue pred;
run;


data MethodA_Results_&exp;
set nlout1_&exp;
respp24 = newlogp24;
Method = "A";
keep tissue site exp Treat day subject daysq Method respp24;
run;

*End Comment 8;
```

```
**End Method A;

**Create the results for the 2-step process of subject specific model then
method A with imputed data (Method B);

*Comment 9 - Use the model building data and split the subjects into groups
based on 3 or less measurable data points
            and more than 3 measurable data points;



proc means data=modellong1_&exp noprint;
by subject type notsorted;
var event;
output out=test_&exp sum=total;
run;


data test1_&exp;
set test_&exp;
if total ge 4 then quad = 1;
if total < 4 then quad = 0;
drop _type_ _freq_;
run;

*End Comment 9;

*Comment 10 - Creates a result to indicate if any of the subjects have
greater than 3 measureable observations;

proc means data = test1_&exp;
var quad;
output out = test1a_&exp sum=combsum;
run;

proc sql;
select combsum
into: sum1
from test1a_&exp;
quit;

*End Comment 10;


*Comment 11 - Uses the result above to indicate whether we should just use
Method A or not;

%if &sum1>0 %then
%do;

*End Comment 11;

*Comment 12 - Creates a dataset that only includes subjects with greater than
3 measureable observations;
```

```
proc sort data=test1_&exp;
by subject;
run;

data check_&exp;
merge modellong1_&exp test1_&exp;
by subject;
if total = 0 then delete;
run;

data check1_&exp;
set check_&exp;
if total < 4 then delete;
run;

*End Comment 12;



*Comment 13 - Creates the within subject quadratic models and saves the
results to a dataset;
proc reg data = check1_&exp;
model logp24 = daysq day;
by subject;
output out=check2_&exp p=yhat;
run;
quit;

libname lib "library location";
run;
data lib.subj_initialfits_&exp;
set check2_&exp;
run;

*End Comment 13;

*Comment 14 - Replaces the imputed observations with the predicted values
from the subject specific quadratic models;

data check3_&exp;
set check2_&exp;
if event = 0 then newlogp24 = yhat;
if event = 1 then newlogp24 = logp24;
**diffMethodB = (logp24 - yhat)**2;
run;

*End Comment 14;

*Comment 15 - Creates a new dataset merging the original dataset with the one
created from comment 14;

specific models for the imputed observations for the subjects with more than
3 measureable observations;
proc sort data=check3_&exp;
by subject;
run;
data check4_&exp;
```

```
merge check_&exp check3_&exp;
by subject;
run;

data check5_&exp;
set check4_&exp;
if total > 3 then delete;
if event = 1 then newlogp24 = logp24;
run;

data check6_&exp;
merge check4_&exp check5_&exp;
by subject;
run;

*End Comment 15;

*Comment 16 - Determine the starting values for the nonlinear mixed effects
model using the new dataset from comment 15;

ods output ParameterEstimates = beta5_&exp;
proc reg data = check6_&exp;
model newlogp24 = daysq day;
run;
quit;

proc transpose data = beta5_&exp out=beta6_&exp;
id variable;
var Estimate;
run;

proc sql;
select intercept
into: int2
from beta6_&exp;
quit;

proc sql;
select day
into: linear2
from beta6_&exp;
quit;

proc sql;
select daysq
into: sqr2
from beta6_&exp;
quit;

*End Comment 16;

*Comment 17 - Run the nonlinear mixed effect model and determine the
predcitions;

ods graphics on;
proc nlmixed data = check6_&exp;
parms b0=&int2 b11=&sqr2 b1=&linear2 s2=0.4 s2u=0.4;
```

```
pred = b0+b1*day+b11*daysq+ui;
model newlogp24 ~ normal(pred,s2);
random ui~normal(0,s2u) subject=subject;
predict pred out=testB_&exp;
run;
quit;

*End Comment 17;

*Comment 18 - Takes the predictions from comment 17 and replace the imputed
observations for those
 subjects other than the ones that got replaced in comment 14;

data check7_&exp;
set testB_&exp;
if total > 3 then delete;
if event = 0 then MethodB_newlogp24 = pred;
if event = 1 then MethodB_newlogp24 = logp24;
**diffMethodB = (newlogp24 - pred)**2;
drop upper lower alpha probt df StdErrPred tValue pred;
run;


data check8_&exp;
merge check7_&exp check3_&exp;
by subject;
if total > 3 then MethodB_newlogp24 = newlogp24;
run;


data MethodB_Results_&exp;
set check8_&exp;
respp24 = MethodB_newlogp24;
Method = "B";
keep tissue site exp Treat day subject daysq Method respp24;
run;
*End comment 18;

%end;

%else
%do;
*Comment 19 - If none of the subjects have more than 3 measureable
observations, then Method A results = Method B results;

data MethodB_Results_&exp;
set MethodA_Results_&exp;
Method = "B";
run;
%end;

*End Comment 19;


**End code for Method B;
```

```
**Code for Method Z;

*Comment 20 - Use given LOD or calculated alternative if not given to replace
imputed data;

data zimpute_&exp;
set Modellong_&exp;
zun = exp(logp24);
**if LOD = . then delete;
run;

proc means data=zimpute_&exp;
var zun;
output out=lod_&exp min=minlod;
run;

proc sql;
select minlod
into: LOD_min
from lod_&exp;
quit;


data zimpute1_&exp;
set Modellong1_&exp;
if LOD = . then LOD = (&LOD_min-1);
Half_LOD = LOD/2;
run;

data methodz_&exp;
set zimpute1_&exp;
zlogp24 = log(Half_LOD);
if event = 1 then zlogp24 = logp24;
run;



data methodZ_Results_&exp;
set methodz_&exp;
respp24 = zlogp24;
Method = "Z";
keep tissue site exp Treat day subject daysq Method respp24;
run;

*End Comment 20 ;

**End code for Method Z;

**Code to create SSI for method C;


%if &sum1>0 %then
%do;

*Comment 21 - Create a random sample of subjects that have more than 3
measureable observations;
```

```
proc means data=check5_&exp noprint;
by subject notsorted;
var event;
output out=sparse_&exp n=numsub;
run;


proc means data=sparse_&exp noprint;
var subject;
output out=sparse1_&exp n=num1;
run;

data sparse_&exp;
set sparse_&exp;
keep subject;
run;

proc sql;
select num1
 into :num_subject
from sparse1_&exp;
quit;

data random1_&exp;
  set sparse_&exp;
  x=ranuni(1234);
run;

/* Sort on the random variable X */

proc sort data=random1_&exp;
  by x;
run;

/* Keep the first n observations. Since the data points are randomly */
/* sorted, these observations constitute a simple random sample.     */

data sample1_&exp(drop=x);
  set random1_&exp;
run;


proc transpose data = Sample1_&exp out = widesub_&exp prefix=subject;
var subject;
run;

*End Comment 21;




%do i=1 %to &num_subject;


*Comment 22 - Take the subject number from i and create a dataset with that
data the resulting data from comment 18;
```

```
proc sql;
select subject&i
 into :rand_sub
from widesub_&exp
quit;



data check9_&exp;
set check5_&exp;
if subject ne &rand_sub then delete;
run;



data check10_&exp;
merge check3_&exp check9_&exp;
run;

proc sort data = check10_&exp;
by subject;
run;

*End comment 22;

*Comment 23 - Determine the starting values for the nonlinear mixed effect
model;

ods output ParameterEstimates = beta75_&exp;
proc reg data = check10_&exp;
model newlogp24 = daysq day;
run;
quit;

proc transpose data = beta75_&exp out=beta76_&exp;
id variable;
var Estimate;
run;

proc sql;
select intercept
into: int76
from beta76_&exp;
quit;

proc sql;
select day
into: linear76
from beta76_&exp;
quit;

proc sql;
select daysq
into: sqr76
from beta76_&exp;
quit;
```

```
*End Comment 23;


*Comment 24 - Run the nonlinear mixed effects model and create the
predictions;

ods graphics on;
proc nlmixed data = check10_&exp;
parms b0=&int76 b11=&sqr76 b1=&linear76 s2=0.4 s2u=0.4;
pred = b0+b1*day+b11*daysq+ui;
model newlogp24 ~ normal(pred,s2);
random ui~normal(0,s2u) subject=subject;
predict pred out=methodC_out_&exp;
run;
quit;

*End Comment 24;

*Comment 25 - Replace the imputed observations for subject i and add these
results to the results from comment 22;

data check11_&exp;
set methodC_out_&exp;
if total > 3 then delete;
if event = 0 then newlogp24 = pred;
**if event = 1 then newlogp24 = logp24;
**diffMethodC = (newlogp24 - pred)**2;
drop upper lower alpha probt df StdErrPred tValue;
run;


data methodC_out_&exp;
set methodC_out_&exp;
drop pred;
run;

data check11_&exp;
set check11_&exp;
drop pred;
run;




data check12_&exp;
merge check11_&exp check3_&exp;
by subject;
if subject = &rand_sub then total = 4;
**if total > 3 then diffMethodC = diffMethodB;
run;


data check3_&exp (replace=YES);
set check12_&exp;
run;
```

```
%end;

*End Comment 25;

*Comment 26 - Create a new data set replacing the imputed data with the
predictions;

data methodC1_Results_&exp;
set check12_&exp;
respp24 = newlogp24;
Method = "C";
keep tissue site exp Treat day subject daysq Method respp24;
run;

%end;



%else
%do;
data MethodC1_Results_&exp;
set MethodA_Results_&exp;
Method = "C";
run;
%end;

*End Comment 26;


*Comment 27 - Create a Summary of the SSI for each Method and combine;
proc sort data=methodA_Results_&exp;
by subject;
run;

proc sort data=methodB_Results_&exp;
by subject;
run;

proc sort data=methodZ_Results_&exp;
by subject;
run;

proc sort data = methodC1_Results_&exp;
by subject;
run;


data overall_&exp;
set methodA_Results_&exp methodB_Results_&exp methodC1_Results_&exp
methodZ_Results_&exp;

run;

*End Comment 27;

*Comment 28 - Prepare dataset to compare means;
```

```
data overall_dum_&exp;
set overall_&exp;
MethA=0;
MethB=0;
MethC=0;
if Method = "A" then MethA=1;
if Method = "B" then MethB=1;
if Method = "C" then MethC=1;
run;

*End Comment 28;

*Comment 29 - Determine Starting values and create contrasts to compare
methods;

ods output ParameterEstimates = beta50_&exp;
proc reg data = overall_dum_&exp;
model respp24 = daysq day MethA MethB MethC;
run;
quit;

proc transpose data = beta50_&exp out=beta60_&exp;
id variable;
var Estimate;
run;

proc sql;
select intercept
into: int20
from beta60_&exp;
quit;

proc sql;
select day
into: linear20
from beta60_&exp;
quit;

proc sql;
select daysq
into: sqr20
from beta60_&exp;
quit;

proc sql;
select MethA
into: DumA
from beta60_&exp;
quit;

proc sql;
select MethB
into: DumB
from beta60_&exp;
quit;
```

```
proc sql;
select MethC
into: DumC
from beta60_&exp;
quit;


ods pdf body = "output location";


ods graphics on;
proc nlmixed data = overall_dum_&exp;
parms b0=&int20 b11=&sqr20 b1=&linear20 b14=&DumA b24=&DumB b34=&DumC s2=0.4
s2u=0.4;
pred = b0+b1*day+b11*daysq+b14*MethA+b24*MethB+b34*MethC+ui;
model respp24 ~ normal(pred,s2);
random ui~normal(0,s2u) subject=subject;
estimate "A vs. B" b14-b24;
estimate "A vs. C" b14-b34;
estimate "B vs. C" b24-b34;
estimate "A vs. Z" b14;
estimate "B vs. Z" b24;
estimate "C vs. Z" b34;
predict pred out=testB1_&exp;
run;
quit;

*End Comment 29;

*Comment 30 - Create output for SSI and hypothesis tests for comparing
methods;
data SSI_&exp;
set testB1_&exp;
resid_SSI = (respp24 - Pred)**2;
run;

proc means data=SSI_&exp sum;
var resid_SSI;
by Method;
run;
quit;


proc sort data=Modellong_&exp;
by site tissue;
run;

proc means data=Modellong_&exp mean;
var logp24;
by site tissue;
run;
quit;

ods pdf close;

*End Comment 30;
```

```
%mend expp24;


option spool;
%expp24(data);

run;
```