STATISTICAL ANALYSIS PLAN

OVERVIEW AND SUMMARY OF DESIGN

The study is designed as an open-label phase 2 demonstration project, to evaluate the acceptability, safety, feasibility, as well as the adherence to daily oral PrEP (FTC/TDF) for HIV-1 prevention, among 150 male and female adolescent participants at two sites (with approximately 75 participants at each site), ranging in ages between 15 and 19 years old, over a 12 month period. The study will also investigate the uptake and utility of feedback to the participants real time about drug levels obtained from dried blood spots/plasma during adherence counselling.

STUDY ENDPOINTS

Primary Acceptability/Use Endpoints:

Consistent with the primary study objective to evaluate the acceptability and use of daily oral PrEP (FTC/TDF) in HIV-negative adolescents, ages 15-19 years, the following primary safety endpoints will be assessed:

- The proportion of participants who report willingness to use the study regimen, take up PrEP and remain on PrEP as part of a comprehensive prevention package.
- An acceptability questionnaire will be administered at the final study visit to assess the delivery of PrEP and the content of the study visits, including the user-friendliness of the medication regimen (including an assessment of side effects), the delivery format, the dosing strategy, and the clinic visit schedule.
- Number of adolescents recruited, enrolled, and retained in the study.

Primary Safety Endpoint:

- Grades 2, 3, and 4 clinical and laboratory adverse events;

Secondary Endpoints:

Consistent with the secondary study objectives to evaluate the adherence and sexual behaviour of daily oral PrEP (FTC/TDF) in HIV-negative adolescents, ages 15-19 years, the following endpoint will be assessed:

Adherence Endpoint:

- Adherence to daily regimens of oral PrEP as evidenced by self-report, pill counts, and proportion of participants assigned to FTC/TDF who have detectable drug levels.
  - Proportion of doses that are taken as instructed.

Sexual behaviour

- Reported number of steady and casual sex partners, condom use (and change after introduction of study product), substance use prior to or during sex, as evidenced by participant responses to interviewer-administered questionnaires and focus group.
- Proportion of scheduled HIV testing appointments that are missed, in relation to individual characteristics (age, sex, number of partners, use of other prevention methods) and the characteristics of the product.

Consistent with the exploratory objective to assess the HIV incidence in study participants during the course of study participation, the following endpoint will be assessed:

- HIV infection, as measured by seroconversion of study participants during the approximate 12 months of follow-up

Exploratory Endpoints:

- Consistent with the exploratory study objective to explore the feasibility of biofeedback of Blood Spot results as an adherence enhancing strategy in 150

HIV-negative adolescents, ages 15-19 years, the following exploratory endpoints will be assessed:

- Proportion and number of participants who are tested using DBS and received results of DBS compared to participants who are tested using DBS but do not receive biofeedback of results.
- Number of counselling opportunities in which drug level feedback is requested vs total number of counseling events
- Number and nature of participants who request drug level feedback .
- Correlation between self reported adherence and drug levels at each visit for each participant
- Correlation between self reported adherence and drug levels at each visit in the study overall.
- Proportion and number of participants whose adherence ito study product improves when receiving the results of DBS compared to those who do not receive feedback of results as measured by endpoints above.

Consistent with the exploratory study objective to investigate the relationship between daily PrEP usage and sexual practices and behavior in adolescents, the following exploratory endpoints will be assessed

Sexual behaviour data obtained using questionnaires and a focus group.

STUDY HYPOTHESIS

The DTHF protocol and research teams hypothesize that oral once daily PrEP (TDF/FTC) will be acceptable and utilized safely by healthy adolescents, between the ages of 15-19 years.

SAMPLE SIZE CONSIDERATIONS

The sample size calculation is based on the ability to detect the measure the primary

objectives (acceptability and use) in the participant population. The table below shows the precision (as 95% confidence limits) provided by various sample sizes in detecting the occurrence of acceptability endpoints (with a frequency non- uptake and use varying from 5%-20%).

Acceptability will be measured by proportion of youth who take up PrEP and remain on PrEP as part of a comprehensive prevention package. Also they will complete an acceptability questionnaire.

	*Precision **		
Frequency	*n=100*	*n=135*	*n=170*
5%	1.6; 11.2	2.1; 10.4	2.4; 9.8
10%	4.9; 17.6	5.7; 16.8	6.4; 16.2
15%	8.6; 23.5	9.9; 22.8	10.7; 22.2
20%	12.7; 29.1	14.2; 28.6	15.3; 28.1

* Asymptotic (exact) binomial 95% confidence limits

Based on these calculations, we anticipate that 135 participants, rounded up to 150 participants to account for premature censoring (due to loss to follow-up and/or withdrawal of consent, estimated at maximum 10%) will be required to detect non-acceptabiltiy endpoints that occur in 10% of subjects with approximately 5% precision.

DATA AND SAFETY MONITORING AND REVIEW

In addition to the informal safety monitoring done by the Site Investigator and study clinician, DTHF will convene an independent Protocol Safety Review Team (PSRT). The roles of the PSRT are:

- To conduct interim reviews of primary safety data as well as study conduct and progress, including rates of participant accrual, retention, rates of adherence to study product and protocol, as well as HIV rates. These reviews will take place approximately **every month,** or as needed. The PSRT may recommend early termination of the study or modification when there is clear evidence of harm towards participants. The PSRT will make a recommendation to the SMC for immediate review of data.

- To respond to queries from site staff on eligibility determination and conditions for exclusion from initial drug dosing.

- To consider and rapidly respond to queries from study staff regarding study drug dosing discontinuation or resumption following occurrence of toxicities as outlined in Section 9 of the study protocol. Additionally, when a site reports any condition that results in study drug being held with resumption pending PSRT consultation, the PSRT will be convened via conference call to make a final determination (unless this is done via email in the interim).

- To respond to queries from the study team when information is needed to guide clinical data management and safety reporting. The Principal Investigator is designated as the point person for queries, and will consult with other PSRT members as needed.

In addition, the Safety Monitoring Committee (SMC) will conduct interim reviews of safety data as well as study conduct and progress, including rates of participant accrual. Retention, rates of adherence to study product and protocol, as well as HIV rates. These reviews will take place approximately every 3 months, or as needed. The SMC may recommend early termination of the study or modification when there is clear evidence of harm towards

participants. Should the SMC or the PSRT detect evidence of harm such that the recommendation for halting or modifying the study is made, then this will be communicated to the PI, The protocol team and the DAIDS prevention division. The IRB and MCC will also be informed of the discontinuation of the study and/or the modification and reasons for modification of the study. All further enrollment, and study activities other than safety follow up will be put on hold. Participants, guardians and CABS will be notified as soon as possible after next steps are identified.

DATA ANALYSIS PLAN

A customized data entry and management system will be implemented for the project in Microsoft Access. Data will be analysed using the statistical programme STATA (Stata Corporation, College Station, Texas, USA). Simple descriptive statistics including frequency distributions will be used to assist in data cleaning and summarization.

The analysis plan for each of the study objectives (per 10.2 above) is detailed below. (Note that throughout the analysis, all statistical tests will be 2-sided using an alpha of 0.05.)

Endpoints:

A. PRIMARY

1. Primary: Use

Number of adolescents who continue to use PrEP after the initial 3 month period and the total time on PrEP for each adolescent and for the cohort as a whole.

Analysis:

Proportions of participants (a) willing to use and (b) taking up PrEP will be described with point estimates and exact 95% confidence intervals. "Remaining on PrEP" will be defined as any self-reported use of PrEP at each study interval (regardless of adherence) over the 9-month follow-up period, as a binary outcome. The first 3-month "use" will be 100% for all 150 participants. We will use product-limit methods to examine 'survival' proportions for

participants remaining on PrEP, accounting for censoring due to loss to follow-up. We will (i) examine the reasons for discontinuation and (ii) examine the baseline and/or pre-LTF characteristics of those LTF, and compare to those retained, in order to understand the possible associations between LTF and adherence and/or sexual risk taking. While our preliminary analyses will presume this censoring is uninformative, we will include analyses that assume LTF are non-adherent. Our intention is to (i) examine the reasons for discontinuation and (ii) examine the baseline and/or pre-LTF characteristics of those LTF, and compare to those retained, in order to understand the possible associations between LTF and adherence and/or sexual risk taking. While our preliminary analyses will presume this censoring is uninformative, we will include analyses that assume LTF are non-adherent. In subsidiary analyses for this primary endpoint, we will also examine the factors associated with PrEP willingness, uptake and "remaining on PrEP" using standard methods (including chi-square and exact tests for the comparisons of proportions, and log-rank tests for the comparison of 'survival' curves in product-limit analysis).

2. Primary: Acceptability

An acceptability questionnaire will be administered 3 monthly to assess reasons for continuing, stopping or restarting PrEP. A further questionnaire will be administered at the final study visit to assess the delivery of PrEP and the content of the study visits, including the user-friendliness of the medication regimen (including an assessment of side effects), the delivery format, the dosing strategy, and the clinic visit schedule.

Analysis Plan:

This acceptability questionnaire will contain both binary items as well as continuous (interval) measures from Likert-type items. Analysis of data from this questionnaire will begin with standard descriptive statistics for each item, overall and by key participant subgroups (a priori, these will be based on participant age, gender, sexual activity during the study and PrEP adherence during the study). We will compare responses to acceptability items using standard methods, including chi-square and exact tests; Student's t- and rank-

sum tests, replaced by analysis of variance or u-tests for global comparisons across polytomous groupings.

3. Primary: Safety

Grades 2, 3, and 4 clinical and laboratory adverse events

Analysis:

This analysis will be based on descriptive statistics.

B. Secondary Endpoints:

1. Secondary: Adherence

Adherence to daily regimens of oral PrEP as evidenced by the use self-report, pill counts, and proportion of participants on FTC/TDF who have detectable drug levels.

Proportion of doses that are taken as instructed.

Proportion of blood samples with detectable drug levels

Proportion of adolescents with detectable drug levels who report using PrEP

Analysis Plan:

There are 3 different measures of PrEP adherence that will be used for the primary adherence endpoint: Participant self-report, pill count results, and detectable drug levels. While the distributions of each of these measures may vary, a priori we anticipate that analysis of each of these 3 measures will use two different binary metrics from data at each study visit: (i) perfect adherence (defined as adherence data that do not suggest any missed doses, e.g., pill count data consistent with daily dosing, or no missed doses detected on self-report); and, (ii) effective adherence (defined as adherence levels that are consistent with protection against HIV exposure based on available data from PrEP efficacy trials, e.g., >85% doses taken, from pill count, or self-report; or therapeutic drug levels on biological assays). Note that drug levels will be used in analyses of effective adherence only.

Analyses of adherence data will be based on the binary constructs of perfect and effective adherence as well as continuous measures of the proportion of PrEP doses taken; both the

binary and continuous constructs will be analysed at each study visit (repeated measures) and in the overall study period (through aggregation of measures from each study visit). The analysis for this primary endpoint is largely descriptive, with the estimates of adherence and their 95% confidence intervals reported for different adherence measures for the cohort overall as well as a priori subgroupings by age, gender and sexual activity during the study. Subsidiary analyses will develop models to examine the predictors of adherence within the cohort; given the sample size, such analyses will have limited statistical power and should be considered exploratory. For repeated measures of adherence from each study visit, we will use a generalized linear modelling approach that includes fixed effects (e.g., participant characteristics) and random effects (e.g., characteristics of study visits). Following the GLLAMM routine in Stata, this will include mixed-effects linear models (for continuous measures of adherence) with maximum-likelihood extensions to binary adherence endpoints. (Skrondal A, Rabe-Hesketh S. Generalized latent variable modeling: multilevel, longitudinal and structural equation models. Boca Raton: Chapman and Hall, 2004.)

This approach converges with that of generalized estimating equations based on the Huber-White sandwich variance estimator.

2a. Secondary: Sexual Activity

Reported number of steady and casual sex partners, condom use (and change after introduction of study product), substance use prior to or during sex, as evidenced by participant responses to interviewer-administered questionnaires

Analysis plan:

The key analysis question for this secondary endpoint is whether different measures of sexual risk change with the availability and/or uptake of PrEP (in the initial 3-month mandatory study follow-up period) compared to periods "off-Prep". The measures of sexual risk include continuous measures (number of sexual partners) as well as binary measures (casual sex partners yes/no, substance use prior to sex, consistent condom use during sex). All measures are visit-specific, with a maximum of 12 measures during the study, per

participant. In analysis, we will first use graphical approaches to display the proportions of participants reporting binary sexual risk behaviours (or means of continuous measures) at each study visit over the 12-month period. Following this, we will use repeated measures methods (as described above) for both binary and continuous dependent variables, with PrEP access treated as a binary variable distinguishing visits in on vs. off PrEP phases; in the first instance, these models will not include other covariates.

2b. Proportion of scheduled HIV testing appointments that are missed, in relation to individual characteristics (age, sex, number of partners, use of other prevention methods) and the characteristics of the product.

Analysis:

"Missed appointments" will be defined as participants being >7 days late for a visit scheduled on a 28-day cycle. This measure will be analysed as a binary variable following the repeated measures approaches described above.

3. Secondary: HIV Incidence

HIV infection, as measured by seroconversion of study participants during the approximate 12 months of follow-up

Analysis:

The incidence of HIV will be described using person-time (rates per 100 person years of observation) and product-limit (survival proportion) approaches. For both methods, the date of seroconversion will be the midpoint of the interval between the last negative and first positive HIV test. For participants lost to follow-up, the date of censoring will be one day after the last visit attended.

C. Exploratory Endpoints:

1. Exploratory: Dried Blood Spot/plasma biofeedback

Correlation of self reported adherence and presence of drug levels in participants who are tested using plasma drug levels or DBS and received results compared to participants who are tested using plasma/DBS but do not receive biofeedback of results.

Analysis:

These data will be described as proportions at each study visit; while we will calculate differences in proportions attending between participant who do and do not receive biofeedback results, no formal statistical tests will be used for this exploratory endpoint.

2.Exploratory: Sexual behaviour by coital diaries

Sexual behaviour data obtained using a questionnaires and focus group.

Analysis:

Data from questionnaires will be compared to self-reported sexual behaviour on a per participant and per visit basis. While this endpoint is exploratory, we anticipate using the distributions of the observed data to defined post-hoc boundaries of acceptable agreement, and then use this to estimate the proportion of participants with acceptable versus unacceptable agreement at each visit, and overall during the study.

Analysis of Focus Group data:

All focus groups and interviews will be transcribed and then analysed using Framework Analysis (Ritchie and Lewis; 2003). This approach was developed in the 1980's by researchers at the UK National Centre for Social Research and has been used extensively for applied or policy relevant qualitative research (Pope et al; 2000). It uses an analytical framework to organize data according to key themes, concepts and emergent categories in grids or matrices. This method facilitates both between case and within case analysis to look for patterns and connections. It provides a systematic, consistent and transparent approach to handling data that also enables analysis to go beyond description to interpretation where appropriate. In addition, it facilitates the linking of findings with those of quantitative research, which will be relevant in comparing, contrasting and enhancing findings collected quantitatively in this study, in particular on PrEP acceptability and adherence. In addition, this triangulation of methods (qualitative and quantitative) has been advocated as an approach to increase the validity of findings (Mays and Pope; 2000, Shenton; 2004).

Up to three researchers experienced in qualitative data analysis will form the data analysis team and all will be trained on the process of Framework analysis to ensure a consistent approach to analysis. The initial coding framework will be developed and agreed on by the team after coding of half the transcripts at each time point. However, as this is an iterative process, the framework may be amended as coding of the second transcript ensues. Any amendments will be discussed and agreed upon by the team. Each transcript will be coded by two researchers and assessed for consistency in coding. Inconsistencies between two coders will be brought to the data analysis team and discussed until agreement is reached.

DATA HANDLING AND RECORD KEEPING

DATA MANAGEMENT RESPONSIBILITIES

Data collection is the responsibility of the research site staff under the supervision of the Site Investigator.
Case Report Forms (CRFs) will be developed by a multidisciplinary team of study staff members.

SOURCE DOCUMENTS AND ACCESS TO SOURCE DATA

Each participant will be allocated a study reference or participant identification number. All data will only be identified by this unique reference number. Data will be stored in a way in which participant identity is protected in a locked cupboard in a locked room on clinic premises and will be entered into computer by reference number only on password-protected computers. Consent forms with participant names and reference numbers will be kept separately from other data in a locked cupboard on clinic premises. Data will be stored for three years after the completion of the research, per DAIDS policy. Throughout this time, data held on computer will only be accessible by members of the research team who will have access to the computer system by way of security passwords. If the results of this research are published or presented, only group information will be given, not names of people in the study.

The Site Investigator or designee will maintain, and store securely, complete, accurate and current study records throughout the study, in accordance with local regulations. All study-related records must be maintained on site for the entire period of study implementation. No study records may be moved to an off-site location or destroyed prior to receiving approval from DTHF and/or DAIDS.

QUALITY CONTROL AND QUALITY ASSURANCE

Each site will be required to develop a robust SOP for quality control and quality assurance procedures.