PROTOCOL UPDATED 04/19/2013

LONGITUDINAL STUDIES OF HIV-ASSOCIATED LUNG INFECTIONS AND COMPLICATIONS

RFA HL07-008

(LUNG HIV)

PROTOCOL
VERSION 3

APRIL 2013

NOTICE

THE CONTENTS OF THIS PROTOCOL ARE CONFIDENTIAL AND ARE NOT TO BE CITED OR DISCUSSED EXCEPT WITH INDIVIDUALS TO WHOM IT HAS BEEN DISTRIBUTED ON BEHALF OF THE LUNG HIV STEERING COMMITTEE.

PREPARED BY:

LUNG HIV DATA COORDINATING CENTER
CLINICAL TRIALS & SURVEYS CORP.
10065 RED RUN BOULEVARD, SUITE 250
OWINGS MILLS, MD 21117
TELEPHONE: 410-435-0663

BRUCE W. THOMPSON, PHD
PRESIDENT AND CHAIRMAN OF THE BOARD

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV) Common Protocol

TABLE OF CONTENTS

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 1

TABLE OF CONTENTS

Page

CHAPTER 1: INTRODUCTION

1.1 SUMMARY

The National Heart, Lung and Blood Institute (NHLBI) Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV) project was established to examine a broad range of separate yet overlapping pulmonary topics. A collection of datasets and biological specimens will be created for use during this RFA as well as future investigations. The program is structured to facilitate both the development of these shared resources and the completion of the individual projects. Results of these efforts will be disseminated through publication in leading medical journals.

The concept of the Lung HIV study was developed by NHLBI to efficiently support multiple R01 efforts while simultaneously creating a shared database and specimen repository. The Lung HIV program will build on the knowledge and experience from existing studies and facilitate the start-up of new studies to further the understanding of the relationship between pulmonary disease and HIV infection.

The Lung HIV mission is to achieve a clear understanding of the clinical manifestations of HIV-associated pulmonary complications by fostering multidisciplinary research collaboration and establishing a high quality centralized specimen repository with an associated clinical dataset based on shared definitions.

1.2 PURPOSE OF THE STUDY PROTOCOL

The Protocol for the Lung HIV project details the rationale, specifies the objectives, and describes the design and organization of the study. It is very important to utilize standardized disease definitions, medical testing and procedures, and informatics parameters in order to accommodate the objectives of all eight projects as well as the collective Lung HIV study goals.

As the study progresses and the Protocol is modified, the Data Coordinating Center will prepare and distribute (via postings to the Web page) revised versions as necessary. Protocol modifications will be announced by the distribution of a numbered memo to all Clinical Centers, the Specimen Repository, NHLBI staff, and the DSMB. The memo will also be posted to the Lung HIV Web page for future reference.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 2

TABLE OF CONTENTS

Page

CHAPTER 2: OVERVIEW AND SPECIFICATIONS

2.1 OVERVIEW

The survival rate of HIV-infected persons continues to improve with the utilization of highly active antiretroviral therapy (HAART) and similar clinical strategies. Hence, it is becoming increasingly important to explore the prevalence and treatment of pulmonary diseases within this population.

2.2 SPECIFIC AIMS

The primary goal of the Lung HIV project is to facilitate the data and specimen collection efforts of eight individual HIV and pulmonary studies that operate under the direction of the National Heart, Lung and Blood Institute (NHLBI).

2.2.1 Collaborative Lung HIV Study

The collaborative efforts of these project teams will result in a collection of shared datasets and biological specimens that can be utilized by future investigators as well as those contributing directly to the collection. The Lung HIV study will build on existing studies and facilitate the start-up of new projects to further the understanding of the relationship between pulmonary disease and HIV infection. The result will be greater knowledge of the clinical manifestations of HIV-associated pulmonary complications through multidisciplinary research and collaboration.

2.2.2 Individual Lung HIV Projects

2.2.2.1 Johns Hopkins University – Medical (Site 1)

The “Longitudinal Study of HIV-Associated Lung Infections in Soweto” project is discussed in detail in Appendix A.

2.2.2.2 Johns Hopkins University – Bloomberg School of Public Health (Site 2)

The Study of HIV Infection in the Etiology of Lung Disease (SHIELD) project is discussed in detail in Appendix B.

2.2.2.3 New York University School of Medicine (Site 3)

The Longitudinal Studies of HIV-Associated Bacterial Pneumonia project is discussed in detail in Appendix C.

2.2.2.4 Ohio State University (Site 4)

The “Smoking Cessation and the Natural History of HIV-Associated Emphysema Prevalence” project is discussed in detail in Appendix D.

2.2.2.5 University of California, San Francisco (Site 5)

The “International HIV-Associated Opportunistic Pneumonias” (IHOP) study is discussed in detail in Appendix E.

2.2.2.6 University of Colorado (Site 6)

The “Longitudinal Studies of HIV-1 Nef in Pulmonary Hypertension” project is discussed in detail in Appendix F.

2.2.2.7 University of Pittsburgh (Site 7)

The “Prevalence and Pathogenesis of Pulmonary Disease in a Large Multi-Center HIV Cohort” project is discussed in detail in Appendix G.

2.2.2.8 University of Washington (Site 8)

The EXHALE (“EXamination of HIV-Associated Lung Emphysema”) study is discussed in detail in Appendix H

Longitudinal Studies of HIV-Associated

Lung Infections and Complications
(Lung HIV)

PROTOCOL
CHAPTER 3

TABLE OF CONTENTS

Page

CHAPTER 3 : CORE STUDY POPULATION

3.1 INTRODUCTION

The overall objective of the Lung HIV project is to build a collaborative collection of data and specimens that will facilitate the overall study and understanding of HIV and pulmonary conditions. Eight Clinical Centers, each with similar but separate cohorts and objectives, will contribute to this effort.

3.2 ELIGIBILITY CRITERIA

Due the variance between projects and cohorts across the eight Clinical Centers, the eligibility criteria will differ from site to site. However, all projects focus on two major categories of disease: Human Immunodeficiency Virus (HIV-1) and pulmonary conditions.

3.3 RECRUITMENT STRATEGIES

Strategies for recruiting patients vary widely due to the diverse nature of the individual Lung HIV projects. Each Clinical Center will be responsible for developing a cohesive group of researchers to identify potential participants at their individual location. All efforts should be made to recruit patients at the earliest time point in their clinical care possible to optimize the collection of adequate data.

The screening and consent processes will be the responsibility of the individual Clinical Centers. These efforts will be carried out while preserving the patient's privacy and confidentiality.

The processes of recruitment, screening, and enrollment will be different at each Clinical Center due to differences in staffing, infrastructure and interdepartmental relations.

3.4 RECRUITMENT GOALS

Target recruitment will vary across the Lung HIV Clinical Centers. A summary table is provided as Appendix I. Monthly reports will compare age, gender, and race between Lung HIV study subjects.

3.5 INFORMED CONSENT

Written informed consent of the participant will be required before any Lung HIV clinical visit takes place. The participant will be given the consent form to read after which the investigator or Study Coordinator will thoroughly explain the study and answer any questions.

Each Lung HIV Clinical Center will have an individualized consent form specific to the project. However, all Clinical Centers are required to include key elements within the consent document to reflect the collaborative study requirements.

Participants will have the study goals of the Lung HIV collaborative study as well as the individual Clinical Center project fully explained to them. The amount of time required for the interviews and procedures will also be explained. Subjects will be informed that refusal to participate in any part of the Lung HIV or individual projects will not change their current or future care at the Clinical Center.

Each Clinical Center will be able to customize its own consent form based on the approval of the Data Safety Monitoring Board (DSMB). Investigators at each Clinical Center will work with their Institutional Review Board (IRB) to develop an effective sampling and monitoring strategy to ensure that the approved procedures are being followed.

3.6 HIPAA COMPLIANCE

Each Lung HIV Clinical Center will be responsible for its compliance with the current HIPAA requirements. This includes familiarity with what data are considered personal identifiers and should not be forwarded to NHLBI or the DCC. The DCC will design all Lung HIV study forms

and databases to omit such variables. Personal identifiers will not be added to the Common Study Elements.

Each Clinical Center will fully explain their institution's HIPAA release form prior to obtaining the participant's signature. This form should include NHLBI and the DCC as institutions that may review the participant's data.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 4

TABLE OF CONTENTS

Page

CHAPTER 4 : CLINICAL CENTER PROCEDURES

4.1 OVERVIEW

Each Lung HIV Clinical Center will have individualized requirements regarding data and specimens to be collected. In addition to these project-specific items, the collective Lung HIV study has established a set of collaborative data to be collected from all study subjects, regardless of Clinical Center. Clinical questionnaires to determine the extent of symptoms, associated medical illnesses, smoking, limited environmental and occupational exposures, and quality of life will be administered. The primary goal of the Lung HIV project is to collect standardized data and specimens. The medical record will only be used to collect information that is relevant to the collection of the information listed in the study forms. A complete list of the Lung HIV Common Forms is given in Appendix J.

4.2 COLLABORATIVE INTERVIEWS AND QUESTIONNAIRES

4.2.1 Contact Form (Common Form 000-001)

This form is used to record information about the participant including his or her name, address, and other personal identifiers. This form is available in StudyCTMS for printing paper copies but these data are neither collected nor maintained in the Lung HIV data system.

4.2.2 Demographics Form (Common Form 000-002)

Demographic information is collected on all Lung HIV study participants regardless of Clinical Center or individual project.

4.2.3 HIV Pulmonary Questionnaire (Common Form 000-003)

A detailed questionnaire will collect data concerning the patient's medical history, current status, and socioeconomic factors. Queries concerning both HIV+ and pulmonary conditions will be

included. This form is collected at regular intervals for all Lung HIV study subjects irrespective of individual project or Clinical Center.

4.2.4 Pulmonary HIV Diagnosis Form (Common Form 000-004)

Data regarding infectious (bacterial pneumonia, TB, PCP, and other pneumonias) as well as non-infectious (asthma, COPD, lung cancer, PAH, sarcoidosis, and other) conditions are collected from each Lung HIV study subject.

4.2.5 Laboratory Abstraction Form (Common Form 000-005)

Key laboratory results are collected for study subjects at Lung HIV Clinical Centers where blood samples are taken in association with the individual project at that site.

4.2.6 Pulmonary Function Testing Form (Common Form 000-006)

Pulmonary Function data is collected for study subjects at Lung HIV Clinical Centers where PFTs are performed in connection with the individual project at that site.

4.2.7 CT Scan Report (Common Form 000-007)

CT scan data is collected for study subjects at Lung HIV Clinical Centers where CT scans are being done as part of the individual project at that site.

4.2.8 BAL Form (Common Form 000-008)

Bronchalveolar Lavage fluid and data are collected for study participants at Lung HIV Clinical Centers performing BAL for that individual project.

4.2.9 Missed Visit Form (Common Form 000-009)

This form is completed any time a study subject misses an expected visit.

4.2.10 Deactivation Form (Common Form 000-010)

This form is completed when a subject discontinues participation in the study regardless of the reason for doing so.

4.2.11 Death Form (Common Form 000-011)

This form is completed upon the death of a study subject.

4.2.12 Adverse Event Report (Common Form 000-012)

This form records events that are definable as Adverse Events or Serious Adverse Events according to the criteria established by the Lung HIV Investigators. Details are provided in the Lung HIV Manual of Operations.

4.3 LABORATORY SPECIMEN COLLECTION

In addition to any specimens that may be collected for the individual project at that site, each Lung HIV Clinical Center is requested to collect blood samples for the collaborative specimen repository. Details are provided in the Lung HIV Manual of Operations.

4.4 CHEST CT SCANS

The goal is to obtain CT image data on a certain number of Lung HIV participants. Details are provided in the Lung HIV Manual of Operations.

4.5 PULMONARY FUNCTION TESTS

The goal is to obtain PFT results from a certain number of Lung HIV participants. Details are provided in the Lung HIV Manual of Operations.

4.6 BLOOD SPECIMENS

Certain blood specimens will be collected from every Lung HIV study subject regardless of individual project or Clinical Center. Details are provided in the Lung HIV Manual of Operations.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 5

TABLE OF CONTENTS

Page

CHAPTER 5: STATISTICAL CONSIDERATIONS

5.1 SAMPLE SIZE CONSIDERATIONS

As part of the review for all Lung HIV Study proposals, power calculations will be made for the primary end point. Reports from the main Lung HIV database (StudyCTMS) will be used to ascertain if adequate numbers of patients are available to achieve the study objectives for a proposed study. The DCC may assist in these calculations or an external reviewer may submit their own as part of their proposal.

The approach to estimating study size and power depends on the type of study and the type of end point considered. There are two types of end points that are anticipated in the Lung HIV studies: 1) Categorical end points such as whether a participant has elevated creatinine levels in a tissue sample and 2) Continuous end points such as pulmonary function test results. Statistical power will be assessed for primary and secondary end points and prespecified secondary analyses for each approved study. The DCC staff recommend adjusting the alpha level for secondary end points to reduce the number of spurious associations that would be detected if an alpha level of 0.05 was used for the analysis of all secondary end points. In prior studies, the DCC has recommended reducing the alpha level for analyzing secondary hypotheses to 0.01 as an indicator of statistical evidence and 0.001 as an indicator of strong evidence between the risk factor and the event.

In Section 5.2, two commonly used study size formulas for analyses of continuous and categorical end points are presented. These calculations are based upon the comparison of two groups e.g., comparison of COPD patients to a control/comparison group of patients with other types of lung diseases (e.g., a case-control design), or comparison of patients with a specific lung disease who are in the early stages of the disease to patients with the same disease who are in the terminal stages of the disease. Both of these designs would allow for equal allocation of patients into the two groups, but this is not required. DCC staff will assist the investigators with issues concerning differential allocation versus balanced allocation in the design phase of each Lung HIV study. The DCC also has methods to estimate study size for comparing three or more groups and for estimating regression coefficients, but these types of designs are less frequently used and these formulas are not presented here. All study size estimates and power calculations will account for losses in power due to missing data. Sample size formulas will also account for the possibility that multivariate regression may be used in a study. In this circumstance, it is necessary to ensure that sufficient numbers of patients are present in a study

to allow regression to be reliably carried out. In general, if one adheres to a rule of having 10 observations for every regressor anticipated to be included into a regression equation, the design will have adequate numbers to perform the required analyses.

As stated above, the main difference in performing study size calculations and power calculations for observational studies is the unequal number of subjects that fall into the two comparison groups. We have designated these proportions by “a” and “1-a” in all of the study size formulas presented. The other features common to both of the sample size formulas are the critical values used to determine the alpha level and power of the test. We have designated these values as Z_α and Z_β respectively. “N” is the total sample size necessary for a study. The size of each comparison group can be obtained by multiplying “N” by “a” for one group and N by “1-a” for the other group. We have presented the formulas for study size calculations, but all of these formulas can be algebraically rearranged to give corresponding power calculations.

One of the most important aspects about the design of the analyses for the Lung HIV project is to ensure that the investigators are able to assess and test sufficient numbers of samples that represent the full spectrum of each patient’s condition. Sufficient numbers are required so that an etiological pathway can be constructed for each disease and so that different diseases can be compared to determine how the pathways of different lung diseases are the same or different. This will require stratification to ensure sufficient numbers of data points are present for each disease and to ensure that the samples for each disease are not over-weighted to specimens that have been collected long after the disease process has started.

5.2 POWER CALCULATIONS

In an effort to demonstrate that a proposed study will have sufficient power to achieve its primary goal, the DCC will present power analyses for the comparison of two groups with respect to a binary endpoint and a continuous endpoint. The power analyses will be adjusted to incorporate the impact of missing data and subgroup analyses in the proposed analysis plans. Below we have presented the major variables and the proposed analysis models that might be used for a variety of different types of end points that can be anticipated in the Lung HIV studies.

TABLE 5.1

Major Variables and Proposed Analysis Models

VARIABLE	TYPE OF END POINT	TYPES OF ANALYSES
Pulmonary Function Tests	Continuous measures	Chi-Square, ANOVA, and Regression Techniques
Biochemical assays on lung tissues	Continuous and categorical measures collected on lung tissue specimens.	Chi-Square, ANOVA, and Regression Techniques
CT scan results and clinical assessments on lung tissues	Ordinal measures collected at baseline	Chi-Square and Logistic Regression Techniques
Clinical evaluations	Dichotomous measures	Chi-Square and Logistic Regression techniques
Use of steroids and other treatments	Categorical measures	Chi-Square, ANOVA, and Regression Techniques

Independent Variables that will be used in these analyses include but are not limited to:

- Demographic Variables
- Clinical Variables (e.g., use of steroids, tissue morphological characteristics, CT scan results, type of lung disease, and stage of disease).
- Indicators of Genetic Alleles (from PCR on blood lymphocytes)

5.2.1 ANALYSIS OF MEANS

Under the assumptions that observations are independent and the variance is equal in two groups, the study size to detect differences between two means can be calculated. (8) The study size formula for a two-group comparison is:

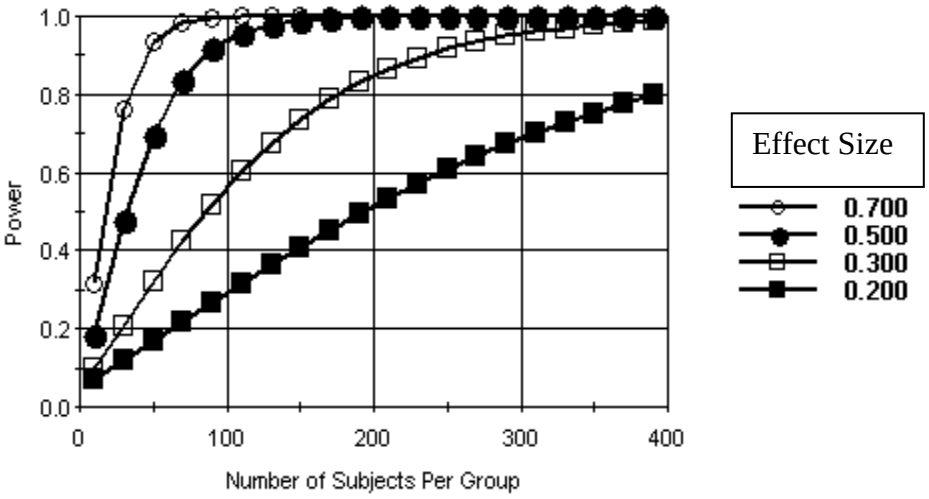
$$N = \frac{(Z_\alpha + Z_\beta)^2 \sigma^2}{a(1-a)(\mu_1 - \mu_2)^2}$$

In the above equation, $\mu_1 - \mu_2$ is the expected difference between two means and σ^2 is the common variance of the continuous variable.

Power is lowered as the proportion of the population (a) with an attribute becomes further away from 0.5, which in turn indicates that the proposed sample size will only be sufficient to detect larger effects sizes as the prevalence of the attribute decreases. In general, effect sizes that can be detected in this study are small to moderate (if the whole study size is used) and moderate to large (if a subgroup analysis is performed).

Below we present a graph showing the power of a two-group comparison of a continuous measure as a function of the effect size [$(\mu_1 - \mu_2)/\sigma$] and a common sample size in the two groups.

Figure 5.1
Comparison of Two Group Means
As a Function of Effect Size and the Sample Size
α = 0.05



As can be seen from the graph, if substantial effect sizes can be hypothesized, the number of required samples for a study will be small. For instance, if the effect size is one half

of a standard deviation, a sample size of at least 80 per group will be sufficient to detect this effect when testing at the α= 0.05 level. To have adequate power to detect smaller effects will require larger sample sizes. It is likely that most effect sizes being proposed in the Lung HIV project will fall into the range presented in the above table.

5.2.2 ANALYSIS OF PROPORTIONS

Study size calculations for analyses involving proportions are functions of the overall study size, the proportion of subjects in each of the comparison groups, and the difference between the expected proportion of events in the different groups (9). A study size calculation for a two-

$$N = \frac{(Z_\alpha \sqrt{\bar{p}\bar{q}/a(1-a)} + Z_\beta \sqrt{p_1 q_1 / a + p_2 q_2 /(1-a)})^2}{(p_1 - p_2)^2}$$

group comparison of proportions has the following formula:

In the above equation, p_1 is the probability that someone with the characteristic will have a particular event, p_2 is the probability that someone without the characteristic will have the event, $\bar{p}$ is the weighted average (weighted by a) of p_1 and p_2, and $\bar{q} = 1 - \bar{p}$. Studies of all 1600 patients can have low values of "a" (e.g., 0.1) and low probabilities of outcomes (e.g., 0.1) and still have adequate power to detect relative risks of 2 (93%). However, if a subgroup analysis of 300 patients is performed (such as might be the case if two groups were being compared that only differed with respect to the timing of disease progression), both "a" and the probability of and outcome would have to be higher (e.g., "a" = 0.3, probability of an outcome = 0.2, power = 0.94 to detect a relative risk of 2).

Below we have presented two figures showing the power to detect specified alternatives of p_1 and p_2 assuming that equal numbers of specimens are used in each group, and the proportion of specimens with an attribute (the control event rate) is low (0.1) or moderate (0.3). In both of these figures, we have assumed that the proportion of case specimens with the attribute (the case event rate) will be higher in the alternative. As can be seen from these figures, small sample sizes will be sufficient to detect large differences between event rates. If smaller differences between the two event rates are postulated in the alternative hypothesis, larger study sizes will be necessary.

Figure 5.2

Comparison of Two Proportions

Control Event Rate Low (0.1)

α = 0.05

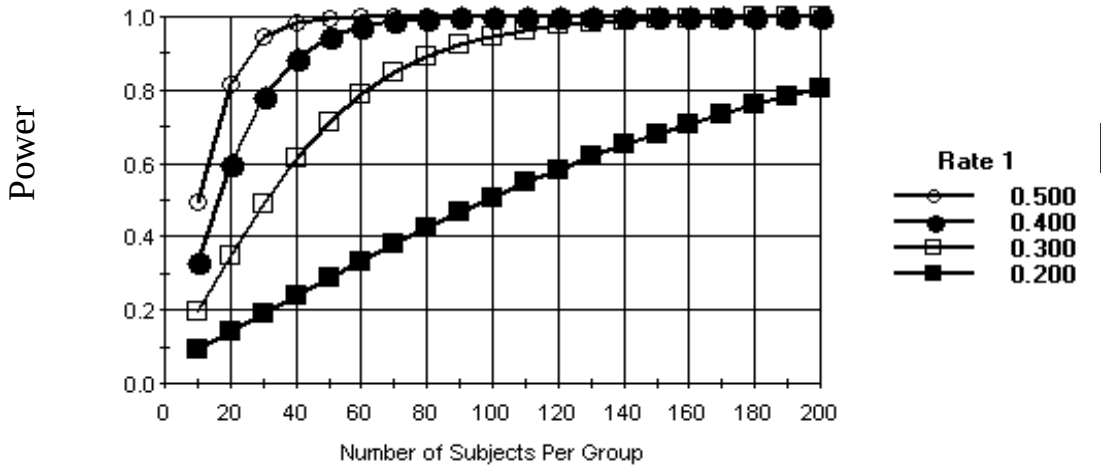
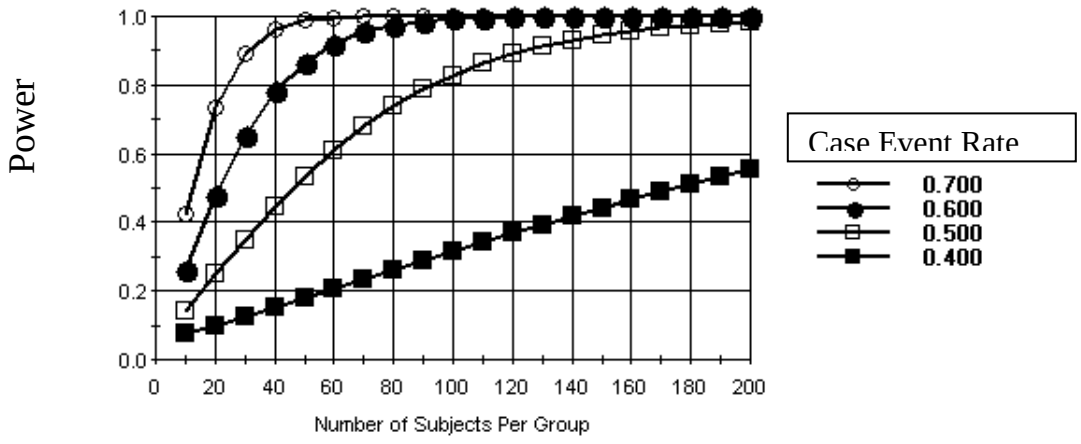


Figure 5.3

Comparison of Two Proportions

Control Event Rate Moderate (0.3)

α = 0.05



5.2.3 INFERENCES INVOLVING MULTIPLE COMPARISON GROUPS

Often investigators are not only interested in whether there are differences among multiple groups, but what pairwise differences are statistically significant. When performing many pairwise comparisons among several groups, one should adjust (or protect) the alpha level of the test to reduce the chance of finding spurious significant differences simply because more tests are performed. The DCC currently uses the methods of Hayter (10) to protect the overall alpha level for multiple comparisons. For three-group designs, the procedure is analogous to the Fisher Least Significant Difference (lsd) rule (10). That is, the test for testing whether all the means are equal is done first (alpha = 0.05), and only if the test is significant are the pairwise comparisons calculated. Once it is determined that the overall test is significant, the individual pairwise comparisons are carried out at the alpha level of 0.05 for each comparison.

5.2.4 REGRESSION MODELS

Multivariate regression models allow one to compare the statistical strength of associations among several risk factors in the presence of markers and co-factors. Power is usually increased when using regression models compared to simple univariate comparisons. For continuous endpoints, inclusion of important independent variables in the regression equation serves to reduce the error variance for all other comparisons. For logistic regression, there is also a bias in estimation of the odds ratio, but the direction of the bias can be positive or negative. Thus, regression models are important because they increase the efficiency of proposed comparisons. However, it is required to ensure that there are sufficient numbers of patients to allow regression analyses to take place. Lung HIV investigators will use the convention of having 10 observations for each planned regressor in a multivariate analysis to ensure that the sample size is adequate for this type of analysis.

5.2.5 MISSING DATA AND COMPLIANCE

The DCC will adjust study size requirements using the methods of Lachin (8). Under an assumption that data are missing at random one divides the complete data study size by the estimated proportion of individuals expected to have complete studies to arrive at the final study size estimate.

5.3 DATA ANALYSIS

5.3.1 Overview

Data analyses will be carried out in the Lung HIV study for two main purposes. One is to monitor CC performance and the second is to perform appropriate analyses of all study data with particular emphasis on evaluation of comparisons among participant groups. It is anticipated that recruitment and status reports will be generated on a monthly schedule and monitoring reports providing more detailed information about data quality and performance of the CCs and CLs will be generated on quarterly performance reports that will be presented to the SC, the NHLBI and the DSMB. Analyses of study specific objectives will be performed on an "as needed basis." This will include support, as directed by the NHLBI and the DSMB, for discrete and collaborative studies between the Lung HIV project and other investigators.

5.3.2 ANALYSIS FOR STUDIES

Primary analyses for each study will focus on estimating group differences for the designated primary end point and developing statistical models to determine associations and relationships between dependent variables and risk factors. Particular analysis methods will depend on the type of study being performed and the type of end point and covariates being collected.

Analysis of binary end points will be accomplished using contingency table analysis. Significance of results will be assessed with the Chi-square test uncorrected for continuity. Contingency table analysis will also be used for polychotomous dependent variables. Results for stratified analyses will be summarized using the Mantel Haenzel statistic and odds ratio (12). Results of matched pair designs (e.g. synthetic case control studies) will be analyzed using McNemar's test and the odds ratio derived from that test (13).

For continuous variables, comparisons of groups will be accomplished using Student's t-test or the Wilcoxon rank sum test depending on the distributional properties of the data. Stratified designs will be analyzed using regression methods with the strata represented as randomized blocks for the analysis.

5.3.3 REGRESSION ANALYSES AND ADJUSTMENT

It will be essential to adjust study results for potential confounding factors. Three general approaches are proposed: matching, stratification, and regression analyses. In observational studies the group divisions can lead to different profiles of known risk factors in the groups because the group divisions are not randomized. This creates problems when trying to determine which of the risk factors are responsible for a

particular outcome. Matching and stratification represent a non-parametric way to adjust for potential confounding variables and regression models represent a parametric method. Matching and stratification are used to “adjust out” confounding effects when one is not interested in estimating those effects and regression models are used when one is interested in comparing the effects due to a collection of risk factors.

Both of these techniques are important when performing outcome analyses. Not only do regression models control for potential confounding from markers and cofactors, they can also be used to determine if effect modifiers are present which accelerate or delay cellular and biochemical changes involved in the progression of a lung disease. These effect modifiers are usually included in regression models in the form of interactions.

The DCC uses SAS procedures to perform adjusted analyses. PROC FREQ is used to perform Mantel Haenzel analyses, and PROC GLM to perform regressions and analyses of variance. PROC LOGISTIC is used to perform unconditional logistic regression. The standard output from SAS procedures usually provide: point estimates for the regression coefficients, standard error estimates, and confidence intervals. The results of these analyses can be printed into computer files so that they can be directly inserted in progress reports using PROC REPORT.

In some instances, the procedures in SAS will not suffice since SAS procedures usually do not include methods to incorporate information about missing data, nor do they include complex models specifically designed to relate a biological process with the risk of disease progression. In these circumstances, the DCC uses PROC IML and PROC NLIN to program the required models. Utilization of SAS makes transporting the results to other institutions easy since most institutions have the capability to run SAS programs.

Since there are no plans to follow patients after they are enrolled into the study, it will be necessary to construct disease progression profiles from the individual data points that have been collected. Regression methods will be used to accomplish this aim. As part of the data collection efforts, the Lung HIV investigators will be asked to determine (or estimate) the amount of time that has transpired since the patient first developed the conditions that lead to the disease. It is expected that these estimates will be crude, but it may be possible to at least postulate that the disease is in its early stages, middle stages, or late stages.

If it is possible to actually estimate an “age” at which the disease process was initiated, the time since initiation of the disease is the difference between age of the patient at the time of the visit and the age at which the disease was initially developed. A time measure such as this can be used as a blocking factor in regression analyses, or as an actual regressor.

The identification of such a variable would allow the Lung HIV investigators to compare the cellular and biochemical progression of the different lung diseases in a very meaningful way. Under this scenario, it would be possible to develop disease type by time interactions to compare and contrast the different cellular and biochemical changes that take place between the different lung diseases being studied in the Lung HIV.

If it is only possible to create definitions for early, middle, and late stages, it will be more difficult to compare the disease processes of the different lung diseases since it will not be possible to determine if “early,” “middle,” and “late” mean the same thing in each disease. In this case, the “timing” variable would either not make sense or would be difficult to interpret when included in interaction analyses. However, the ordinal nature of the determinations would allow meaningful comparisons within a specific disease, and one could track the cellular and biochemical changes that occur within a lung disease type.

The DCC will develop regression and analysis of variance models that will provide meaningful interpretations of the Lung HIV data. It is anticipated that standard analysis methods would be used for continuous variables and for categorical variables, logistic regression would be used. All regression analyses will be performed using the SAS statistical analysis package. SAS has extensive routines to examine the goodness of fit of each analysis package. For linear regression, tests will be made to determine the goodness of fit by examining the residuals of the analysis using influence analysis and the Hosmer-Lemeshow test (14) will be used to test residuals.

5.3.4 MISSING DATA

Some missing data are anticipated. If data are missing at random, there will be a loss in efficiency of the proposed analyses, but bias will not be introduced into the study by not accounting for the missing data. We have assumed that we will have outcome information on 85% follow-up on Lung HIV patients. Power curves and tables have been reviewed that account for sample numbers that are lower than 1600 patients and review of these tables suggests that even if the final sample size is lower than 1000, there will be sufficient power to address the study objectives.

However, if data are not missing at random, there could be a bias in some of the estimation and inference routines used in this study. We will use the multiple imputation procedure developed by Rubin to correct for this type of bias. (15)

5.3.5 CONCLUSIONS

A preliminary power analysis has been performed for the Lung HIV project. It has been found that the proposed number of patients (N = approximately 1600) is sufficient to allow for small to moderate effects to be detected with adequate power (at least 80%) when testing at the alpha level of 0.05 (two-sided tests) and for larger effect sizes if subgroup analyses are to be performed.

The Lung HIV project offers a unique opportunity to incorporate variables depicting cellular morphology differences, lung morphology differences, lung tissue biochemical differences and genetic differences of different diseases to show how these variables impact the progression of the various lung diseases. The models presented here are general and robust enough to allow important research to be done in this area.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 6

TABLE OF CONTENTS

Page

CHAPTER 6 : DATA COLLECTION AND MANAGEMENT

6.1 DATA COLLECTION FORMS

The information recorded and collected on standard study forms will provide a large part of the database for the analyses and conclusions emanating from the Lung HIV project. Question by Question Responses (QxQs) will also be developed for every form, outlining specific instructions for that form as well as specific instructions for each question.

During the course of the Lung HIV project, it may be necessary to revise a form. This can occur because, despite prior testing, questions on a form may not be easily understood by the participants, the structure of the form does not "flow" well in routine use, or the fields are not of the correct size or type. The DCC will make every effort not to revise forms over the course of the Lung HIV project. There will be a yearly review to determine if minor revisions of a form are necessary. Major revisions that address faulty data collection will be implemented on an "as needed" basis; once forms are revised, DCC programming staff will modify the associated data entry screens. Form revisions will be dated and numbered. A numbered memo will announce the change and will be posted to the Lung HIV Web page for future reference.

The final SC approved versions of each Common and site-specific data collection form will be posted in the Lung HIV data system, StudyCTMS. If a form is revised, the old version will be replaced by the newest revision. Only the most current version of each form will be posted at any time. Study coordinators will be instructed to print out the necessary forms as a patient is recruited into the Lung HIV project.

Some of the information collected on Lung HIV patients will be collected by patient interviews. The Lung HIV investigators have designed questionnaires that address many potential factors affecting lung function and disease progression. The interview is designed not to exceed two hours in length.

Some questions have a "yes/no" component preceding a timeframe component. For example, if a participant is asked if they have been exposed to coal dust or powder, they would first answer "yes" or "no." If the answer is "no", the interviewer would proceed to the next exposure question. If "yes," the duration of exposure will be requested.

Due to the large number of questions anticipated, it will be imperative to have questionnaires that are clearly worded and straight forward to complete. As much as possible, answers will be given in predetermined, defined formats such as yes/no, multiple choice or

continuous numeric variables of predefined format. "Open-ended" or text answers will be avoided wherever possible. These types of answers are difficult to process, summarize and interpret for a large sample size.

6.2 INTERNET DATA MANAGEMENT SYSTEM

C-TASC data management staff have designed a Clinical Trials Management System (StudyCTMS) that will be used by the Clinical Center staff as well as by the members of the DCC. Clinical Center staff will have responsibility for data entry (and correction, if necessary) of all study forms and local laboratory data. The DCC has responsibility for storing all received study data.

StudyCTMS will facilitate Lung HIV data entry, quality assurance, and management of participant and laboratory data. The system will also provide a means for backup and off-site storage of the database and for the transfer of limited data sets to the NHLBI project office on a regular basis.

StudyCTMS provides the following basic functions:

1. Data entry and correction of forms with built-in error checking through screens that resemble the forms
2. Display and printing of keyed forms
3. Central edit of keyed forms resulting in locally printed queries
4. Direct entry of information into the C-TASC central database which is backed up daily at C-TASC (once every two weeks back up files are sent off site)
5. Online inventory display of participant materials and their status
6. Delinquent forms report
7. Delinquent specimen report

6.2.1 Data Entry

A primary function of StudyCTMS for the Lung HIV project is to facilitate the entry of forms into the DCC database. The system is designed to be used with a mouse, keyboard or both in combination to perform data entry and editing of study data. A Web browser serves as the communications connection to the DCC mainframe. A user-specific name and password is required for data entry to ensure that only certified staff can write to the database. Certified staff will choose a form to enter or to correct or display a form for a specific participant. Data entry and correction systems provide automated checking for participant identity, form appropriateness for the participant,

illegal codes, and range checks. In addition, remarks can be entered in free-form text. The data entry form screens are designed to look like the corresponding Lung HIV forms. Response codes are the codes in parentheses next to the form items or a write-in response and are entered as they appear on the form. The data are reviewed and then saved to the central database. The transmission of data from the Clinical Center to the server at the DCC will be encrypted.

During data entry all items are checked for valid codes. For example, if the preprinted codes for responses are 1 to 5, then only 1, 2, 3, 4, 5, not available or null is permitted. Items are checked for the correct format. For example, if a particular lab value is of the form xx.x then entering “8.23” will cause an error. At the end of each session of data entry (original keying or corrections to previously keyed data), StudyCTMS will automatically initiate an edit of the new or changed data from that session.

At the completion of data entry, a copy of the study form will be encrypted and displayed in the user’s browser to the data entry staff. Printing this produces a hard copy record of the information that has been put into the database.

6.2.2 Study CTMS

The DCC data management staff will be responsible for assembling and maintaining a communications network that will allow Clinical Center staff to distribute information, transfer laboratory data and download study-related materials in a quick, efficient, and secure manner. A key emphasis for this group will be to ensure that all communications are secure and easy to use for each Clinical Center staff member. Data transfers and much of the communication will be done using File Transmission Protocols (FTP) and encrypted e-mail. The DCC data management staff will be responsible for assembling, maintaining, and backing up a validated Internet data entry system for the Lung HIV project. This system will have a complete array of editing features, data management guides, and reports that will allow Clinical Center staff to enter and maintain data in a central location.

6.3 QUALITY CONTROL PROCEDURES

6.3.1 Studies and Reports

Recruitment will be closely monitored by the creation of regular recruitment reports that will be distributed to all Lung HIV staff and posted to the Web page. Tables

will summarize performance in: recruitment, submission of specimens, submission of CT scans, and submission of data collected at the Clinical Centers, edit status of the data, and quality of submitted materials. The DCC will provide information about the number of patients enrolled in the protocols, the number of minority participants, and the age distribution of the participants. This will assist the Lung HIV investigators in tracking these important distributions and in reporting these numbers to the NHLBI in their yearly reports. Below we have provided some specifics on the ways we will present performance statistics.

Performance is assessed by consideration of the following at quarterly intervals.

i. For enrolled patients, the number of study forms which are past due at the DCC, based on the date of enrollment
ii. Study forms not yet edited or failing edit
iii. Specimens and CT scans that are past due
iv. Procedures which are required by protocol but were not performed
v. Reports showing the number of specimens:
    - Received in good condition (frozen, fixed, etc.)
    - Meeting other required QC controls
vi. Reports on data showing:
    - Quality of scans
    - Image counts
    - Number of scans that are interpretable
    - Other required QC controls

As indicated above, performance reports include summary statistics for each Clinical Center. Large changes in these statistics from quarter to quarter within a Clinical Center may indicate changes in the way data are being collected. Period-to-period changes will be monitored using Shewart plots. Comparison of these statistics across Clinical Centers could suggest either differences in how data are collected or differences in the participant population, and may prompt further investigation.

6.3.2 Clinical Center Site Visits

This study will not require clinical siste visits to monitor the common protocol. This section will remain blank as recommended by the Lung HIV DSMB.

6.4 LABORATORY DATA MANAGEMENT

6.4.1 Specimen Tracking

The performance of each Clinical Center laboratory participating in a study will be summarized in regular reports. This report will include the number of specimens received and processed, the number of specimens not prepared or not labeled properly, studies of the monthly variation associated with reported results (Shewart plots), etc.

The DCC staff will compare the performance of each Lung HIV laboratory to its own past performance and to agreed upon study standards. Quality control charts (Shewart Plots) will be used to examine the means and frequencies of the assays and evaluations performed over time. An investigation will be undertaken to determine whether any shift represents a change in the population being studied or shift in the methods for performing specified assays.

6.4.2 On-Site Monitoring

This study will not require site visits to monitor the common protocol. This section will remain blank as recommended by the Lung HIV DSMB.

6.5 DCC QUALITY CONTROL

It is assumed that the DCC will undergo site visits by the NHLBI and the Lung HIV investigators on a periodic basis. The format for these visits will be left to the site visitors. DCC staff will be prepared to provide an overview of study operations, and demonstrate what has been done since the last visit.

There are certain activities DCC staff will carry out internally to ensure the quality of the data and analyses that are presented to the Lung HIV investigators.

a. Persons (such as the PI, Project Manager or other DCC staff) not involved in the development of the data editing programs fill out several study data forms, making deliberate errors. These forms are keyed and processed through the data editing system to see if all of the errors are detected by StudyCTMS.

b. A sample of original data forms are compared against the data on the DCC computer. This procedure is used not only to detect data entry errors, but also to detect problems with the editing software developed and implemented by the DCC.

c. For each continuous variable on the database, a point frequency distribution (i.e., a tabulation of the frequency of occurrence of every distinct value) is obtained. This

helps to identify many types of abnormalities in the continuous data such as: (a) digit preferences; (b) bimodality or other distinctive shapes of the distribution; (c) outliers (i.e., extreme values distinctly separate from the rest of the distribution); and (d) incorrect use of missing value codes. Once an observation has been identified as a true outlier, the first step is to go back to the original records and determine whether a recording or keying error was made. If such a value has been verified as correct through the distributed data system, an inquiry is made as to the reasons an outlier exists. The question of whether or not to include the value in the data analysis depends upon the nature of the analysis. There is no reason to exclude the value if the analysis is a count of the number of participants having a value exceeding a given cut-point. However, if measures of central tendency and variability are being computed, or if correlation or regression analyses are being carried out, non-parametric statistics may be preferable.

d. New analysis programs (including those that utilize standard statistical packages such as SAS) are tested by running these programs on a small subfile of 10 or 20 participants and independently reproducing the tabulations and statistical calculations from the original data. These procedures help to assure that the correct variables have been selected from the analysis file, variables and cut-points have been defined properly, and that transformations of the original variables on the analysis file have been formulated correctly.
e. When preparing data reports, different tables, which may have resulted from a variety of analysis programs, are checked for consistency of denominators.

Reliability measures are easily interpretable by CC staff if they are presented as graphical presentations or in tabular form. For continuous variables, the DCC will present scatter diagrams with the correlation presented on the graph. Categorical data can be presented using a “reader 1 versus reader 2” table with the intraclass correlation coefficient present below the table. This is a standard method of presentation and the specific graphics will not be presented here.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL
CHAPTER 7

TABLE OF CONTENTS

Page

CHAPTER 7: STUDY ADMINISTRATION

7.1 ORGANIZATION OVERVIEW

The Lung HIV project will be conducted by the collaborating investigators of a Data Coordinating Center (DCC), eight Clinical Centers (CC), a Specimen Repository, and the NHLBI Program Office. An organizational chart of the Lung HIV project is presented in Appendix K. The Data Study Monitoring Board (DSMB) will be responsible for reporting to the NHLBI on the overall progress of the study, the validity of Lung HIV evaluations, and on the safety profile of the approved study procedures.

The Steering Committee (SC) will be composed of the NHLBI Project Officer, the Principal Investigators of the Clinical Centers, and the DCC staff. This committee will be the focus for discussions and decisions on study design and performance. The DCC will abide by decisions of the SC and the NHLBI throughout the course of this study. This committee forms the basis of DCC distribution lists to ensure that all study staff receive the necessary materials to be trained in study procedures and information on the implementation of study procedures.

The DCC will compile an address directory that identifies the name, address, phone and fax numbers, and e-mail address of all Clinical Center, Specimen Repository, DCC, NHLBI staff, and all committee members. This address directory will be posted to the Lung HIV Web page. This directory will be updated as needed to reflect staffing changes in the Lung HIV project.

Once the study is underway, the DCC support staff will take primary responsibility for presenting information on the progress of the study with respect to collection, analyses, completeness and quality of required data. DCC support staff will be responsible for setting up meetings and conference calls of the Lung HIV project including the identification of the meeting site, meeting rooms, distribution of materials, and recording of minutes. All study materials and minutes not classified as confidential will be uploaded to the Lung HIV Web page.

7.1.1 National Heart, Lung & Blood Institute

This study is being funded by the National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (NIH). The NHLBI Project Officer, Hannah Peavy, is responsible for overseeing the study design, implementation, data quality and information dissemination. The Grant Management Office is responsible for monitoring the receipt of all deliverables as required by the NHLBI and dispersal of payment to the contractors.

7.1.2 Study Chair

The NHLBI has not appointed a study chair for the Lung HIV project.

7.1.3 Clinical Centers

The Principal Investigators of the Clinical Centers (CC) have agreed to abide by their individual protocol, comply with the requirements of the collaborative Lung HIV requirements, to have comparable staff, facilities and equipment and to ensure the proper conduct of the Lung HIV project including: recruitment and characterization of the patients as specified, accurate data collection and the transmission of information and specimens to the DCC and Specimen Repository. Clinical Center Principal Investigators are listed below:

Site 1 – Johns Hopkins University (Medical) - Richard Chaisson
Site 2 – Johns Hopkins University (Bloomberg School of Public Health) – Gregory Kirk
Site 3 – New York University - William Rom
Site 4 – Ohio State University – Phillip Diaz
Site 5 – University of California, San Francisco – Laurence Huang
Site 6 – University of Colorado – Sonia Flores
Site 7 – University of Pittsburgh – Alison Morris
Site 8 –University of Washington – Kristina Crothers

7.1.4 Data Coordinating Center

The Data Coordinating Center (DCC) is responsible for the logistics of study coordination, data management, study monitoring, quality control measures and descriptive data analysis. The DCC will be responsible for all documents and the Lung HIV Web pages. As study proposals are received, the DCC will be responsible for facilitating the review, querying the Lung HIV database for appropriate samples and supplying shipping lists to the Specimen Repository. DCC staff members are listed below:

Principal Investigator – Bruce Thompson
Senior Statistician: Zhaoyu Luo
Senior Coordinator – Arionna Stevenson

7.1.5 Specimen Repository

The Specimen Repository will be responsible for receiving and maintaining laboratory specimens associated with the Lung HIV project. This includes supplying appropriate shipping instructions, verifying receipt of all specimens, division and cataloging of specimens, and long-term storage of specimens. Once a study is approved, the Specimen Repository will be responsible for shipping selected specimens to the appropriate investigators as directed by the SC and DCC. The Lung HIV Repository is BBI SeraCare. The Repository contact person is Elizabeth Wagner (NHLBI). Details regarding the Specimen Repository can be found in the Lung HIV Common Manual of Operations.

7.2 COMMITTEES

7.2.1 Steering Committee

The Steering Committee (SC) consists of: The NHLBI Project Officer, the Principal Investigators of each of the eight Clinical Centers, and the staff of the DCC. The Steering Committee has the responsibility for developing study documents and managing the daily operations of the Lung HIV project.

7.2.2 Data Safety Monitoring Board

The Data Safety Monitoring Board (DSMB) will be charged with reviewing the Protocol and consent forms with respect to ethical and safety standards and advising the NHLBI. During the study, the DSMB will meet periodically to review safety issues and to monitor Clinical Center performance in execution of the protocols. Prior to each meeting, the DSMB will be provided with appropriate Lung HIV reports. At the beginning of each meeting, the NHLBI Project Officer and the DCC Principal Investigator will report on study progress and answer any questions.

7.3 WEB PAGES

The DCC will design and maintain a Web site for the Lung HIV project. All Web pages will be located on the DCC server and will be continuously updated with the most current information available. Information regarding Web pages can be found in the Lung HIV Manual of Operations.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL CHAPTER 8

TABLE OF CONTENTS

Page

CHAPTER 8: POLICY MATTERS

8.1 TRAINING AND CERTIFICATION

8.1.1 Study CTMS Data Entry

All Clinical Center staff members performing data entry for the Lung HIV Study will be required to undergo certification before being allowed to key or correct data in StudyCTMS. This will include the accurate keying and correct updating of sample forms into the data system.

The Lung HIV Principal Investigator and/or Clinic Coordinator from each site will provide the DCC with a list of individuals who will be testing for data entry certification. The DCC will distribute testing materials to all testers and work with them through the certification process.

8.1.2 CT Scanners

To ensure uniform quality standards, CT scanners at all Lung HIV Clinical Centers will be calibrated using phantoms purchased specifically for this project and purpose. Details can be found in the Lung HIV Manual of Operations.

8.2 PUBLICATION POLICY

The Lung HIV project will generate considerable new data about patients with HIV and lung disease. Topics to be developed into publications are generated from concept sheets that are reviewed by the SC with DCC input. If the SC approves the concept sheet, the NHLBI Project Office will authorize the release of specimens and/or data to the investigator. The investigators will be expected to perform their analyses in a rapid and efficient manner. Approved concept sheets and published manuscripts will be posted on the Lung HIV Public Web page.

Investigators at all Lung HIV Clinical Centers, the DCC, and the NHLBI Project Office have equal status with regard to developing concept sheets and collaborating in the

development and publication of research papers based on study material. Study coordinators and other staff at these centers are encouraged to submit studies.

8.3 REPORTS ON METHODOLOGY

Manuscripts concerning the overall design, protocol, procedures, or organizational structure of the study may be published prior to the end of the Lung HIV project. Such preliminary publications will be developed and reviewed according to the same guidelines used for other reports of findings.

Many public presentations about the Lung HIV that do not involve protocol data, or ancillary study data (e.g., grand rounds talks concerning the study's general design and objectives) do not require formal preliminary review and approval by the Steering Committee. However, if there is any doubt, investigators are asked to first consult with the Steering Committee indicating their intention to present the material, in order to avoid the premature release of study data or the inappropriate publication of confidential information.

8.4 RELEASE OF LUNG HIV DATA OR SPECIMENS TO NON-LUNG HIV INVESTIGATORS

Requests for study results, study data, or banked specimens may be submitted by investigators who are not participating in the Lung HIV project during the course of this investigation. These requests will arise primarily from colleagues and researchers who are interested in HIV and lung disease. Each request should be submitted in writing using the Lung HIV concept sheet format and provide the same information as required for ancillary studies submitted by Lung HIV Investigators. The SC reviews each request and the following principles are addressed in determining the disposition of each request.

1. Overlap with previously approved data bank studies.
2. The scientific importance of the request.
3. The efforts and costs of providing the information.

4. Willingness of individuals submitting a request to accept limitations, as specified by the NHLBI, on the uses that can be made of the data and data analysis.

This SC will be responsible for reviewing concept sheets from study Lung HIV Investigators as well as non-Lung HIV Investigators and making a recommendation to the NHLBI and staff at NHLBI will make the final decision. The decision for release of the specimens will be based on the availability of specimens, the scientific goals of the proposal and the order the requests are received after public notice has been issued to announce the availability of the specimens and the NHLBI Policy on Release of Specimens. This policy will remain in effect as long as there are specimens to release.

At least one month prior to the end of funding, the DCC staff will prepare data tapes and appropriate documentation for submission to the NHLBI Project Office. These tapes will not include personal identifiers. The release of these data tapes will be based on the NHLBI Policy on Release of Data from Large Scale NHLBI Sponsored Studies in existence at the time the study funding ceases.

8.5 CONFLICT OF INTEREST POLICY

8.5.1 General Principles

The Lung HIV investigators have agreed to a policy on conflict of interest which has few specific restrictions, but a broad indication for disclosure of potential conflicts of interest. The Lung HIV investigators wish to endorse the spirit and content of the 21st Bethesda Conference: Ethics in Cardiovascular Medicine (Frommer et al, 1990) dealing with these issues, and seek to make this policy consistent with the record of that conference.

8.5.2 Individuals to be governed by These Guidelines

Members of the Lung HIV group who will be governed by these guidelines include the Principal Investigator at each Clinical Center and key personnel in the Data Coordinating

Center. Co-Investigators and other staff who have major responsibility for enrollment, recruitment, follow-up or collection of data for the Lung HIV at Clinical Centers will also be governed by these guidelines. The Principal Investigator for each Lung HIV Clinical Center will submit a list of individuals who will be governed by these guidelines at the beginning of the study and revise, as necessary, annually. The Principal Investigator of each participating unit will review the guidelines with all appropriate staff prior to the start of patient recruitment and at least annually thereafter.

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

PROTOCOL

CHAPTER 9
REFERENCES

CHAPTER 9
Lung HIV PROTOCOL REFERENCES

1. Barnes PJ. Chronic obstructive pulmonary disease. *N Engl J Med*. 2000;343:269-80.

2. Croxton TL, Weinmann GG, Senior RM, Wise RA, Crapo JD and Buist AS. Clinical research in chronic obstructive pulmonary disease: needs and opportunities. *Am J Respir Crit Care Med*. 2003; 167:1142-49.

3. Suki B, Lutchen KR, and Ingenito EP. On the progressive nature of emphysema: roles of proteases, inflammation, and mechanical forces. *Am J Respir Crit Care Med.* 2003; 168:516-21.

4. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS Board of Directors, June 2001 and by the ERS Executive Committee, June 201. *Am J Respir Crit Care Med* 2002; 165:277-304

5. Crystal RG, Bitterman PB, Mossman B, Schwarz MI, Sheppard D, Almasy L, Chapman HA, Friedman SL, King TE, Jr., Leinwand LA, Liotta L, Martin GR, Schwartz DA, Schultz GS, Wagner CR and Musson RA. Future research directions in idiopathic pulmonary fibrosis: summary of a National Heart, Lung, and Blood Institute working group. *Am J Respir Crit Care Med*. 2002; 166:236-46.

6. Oga T, Nishimura K, Tsukino M, et al. Analysis of the factors related to mortality in chronic obstructive pulmonary disease: role of exercise capacity and health status. *Am J Respir Crit Care Med.* 2003; 167:544-9.

7. National Emphysema Treatment Trial Research Group. A randomized trial comparing lung-volume reduction surgery with medical therapy for severe emphysema. *N Engl J Med*. 2003; 348:2059-73.

8. King TE, Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2001; 164:1171-81.

9. Fleiss JL. Statistical methods for rates and proportions. Chapter 3, page 30. John Wiley & Sons, New York, New York, 1973.

10. Hayter AJ, The Maximum familywise error rate of Fisher's least significant difference test. *J Am Stati Assoc* 1986; 81:1000.

11. Lachin JM, Matts JP, and Wei LJ. Randomization in clinical trials: conclusions and recommendations. *Controlled Clin Trials* 1988; 9:365.

12. Mantel N. and Haenzel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 1959; 22:719-48.

13. McNemar Q. Note on sampling error of the differences between correlated proportions

or percentage. *Psychometrika* 1947; 12:153-7.

14. Hosmer DW, Jr. and Lemeshow, S (1989), *Applied Logistic Regression*, New York: John Wiley & Sons, Inc.

15. Rubin DB, Thomas N. Matching using estimated propensity scores: Relating theory to practice. *Biometrics* 1996; 52:249-64.

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix A

Lung Site 1
Johns Hopkins University

1.1 Longitudinal Study of HIV-Associated Lung Infections in Soweto
Richard Chaisson, PI

1.2

1.3

1.4

1.5

1.6

1.7 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix B

Lung Site 2
Johns Hopkins University - Bloomberg

1.8 Study of HIV Infection in the Etiology of Lung Disease (SHIELD)
Gregory Kirk, PI

1.9

1.10

1.11

1.12

1.13

1.14 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix C

Lung Site 3
New York University

1.15 Longitudinal Studies of HIV-Associated Bacterial Pneumonia
William Rom, PI

1.16

1.17

1.18

1.19

1.20

1.21 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix D

Lung Site 4
Ohio State University

1.22 Smoking Cessation and Natural History of HIV-Associated Emphysema
Philip Diaz, PI

1.23

1.24

1.25

1.26

1.27

1.28 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix E

Lung Site 5
University of California, San Francisco

1.29 International HIV-Associated Opportunistic Pneumonias (IHOP)
Laurence Huang, PI

1.30

1.31

1.32

1.33

1.34

1.35 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix F

Lung Site 6
University of Colorado

1.36 Longitudinal Studies of HIV-1 Nef in Pulmonary Hypertension
Sonia Flores, PI

1.37

1.38

1.39

1.40

1.41

1.42 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix G

Lung Site 7
University of Pittsburgh

1.43 Prevalence and Pathogenesis of Pulmonary Disease in a Large Multi-Center HIV Cohort
Alison Morris, PI

1.44

1.45

1.46

1.47

1.48

1.49 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated
Lung Infections and Complications
(Lung HIV)

Common Protocol

Appendix H

Lung Site 8
Yale University / VA Medical Center

1.50 Examination of HIV-Associated Lung Emphysema (EXHALE)
Kristina Crothers, PI

1.51

1.52

1.53

1.54

1.55

1.56 This Appendix contains site-specific Protocol(s) and Informed Consent(s).

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

Protocol Appendix I

Project Recruitment Goals

Site	Breakdown		Project Total
1 – Johns Hopkins (Medical)	South Africa	1,000	1,000
2 – Johns Hopkins (Bloomberg)	Parent Study COPD Lung Immune PAH	3,000 to 4,000 550 500 1,000	3,000 to 4,000
3 – New York University	New York South Africa	200 200	400
4 – Ohio State University	Ohio State University		365
5 – University of California, San Francisco	United Kingdom Uganda United States	480 2,000 520	3000
6 – University of Colorado	France United States	200 200	400
7 – University of Pittsburgh	Univ of Pittsburgh UCLA	300 300	600
8 –University of Washington	EXHALE	360	360
	Lung HIV Project Total		**9,125** to **10,125**

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)

Protocol Appendix J

Common Case Report Forms

Form 001 Contact Form

Form 002 Demographics Form

Form 003 Pulmonary HIV Questionnaire

Form 004 Pulmonary HIV Diagnosis Form

Form 005 Laboratory Abstraction Form (only at sites collecting specimens)

Form 006 Pulmonary Function Testing Form

Form 007 CT Scan Report (only at sites performing CT scans)

Form 008 BAL Form (only at sites performing BAL)

Form 009 Missed Visit Form

Form 010 Deactivation Form

Form 011 Death Form

Form 012 Serious Adverse Event Form

Longitudinal Studies of HIV-Associated Lung Infections and Complications (Lung HIV)
Appendix K

Lung HIV Organizational Structure