STUDY PROTOCOL: GENETIC EPIDEMIOLOGY OF CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPDGene® 10.0)

I. BACKGROUND AND SIGNIFICANCE

Chronic obstructive pulmonary disease (COPD) was the third leading cause of mortality in the United States (1) prior to the SARS-CoV-2 pandemic and an important factor in the increase in deaths from COVID-19 and suicides. COPD is a heterogeneous disease, with varying contributions of emphysema and airway disease in each COPD subject. We have previously identified CT subtypes of disease based on visual scoring and found variation in progression of disease based on spirometry and CT metrics (emphysema, airway disease).

COPDGene was initiated in 2007 enrolling a cohort of more than 10,000 current and former smokers to study subtypes of disease and genetic factors. Subsequent phases (Phase 2 and Phase 3) of the study have returned the initial participants for additional evaluations at approximately five-year intervals.

Genetic studies of complex diseases have the potential to provide insight into the pathophysiologic mechanisms of COPD susceptibility and heterogeneity. We have performed comprehensive DNA sequencing studies and analysis of multiple omic parameters to identify genes and proteins that influence COPD risk and progression in two major racial/ethnic groups (non-Hispanic Whites and African Americans). Our primary hypotheses are:

(1) Precise characterization of COPD subjects using computed tomography – as well as clinical and physiological measures assessed longitudinally – will provide insight that will enable COPD to be decomposed into clinically significant subtypes.

(2) DNA sequencing studies and omics will identify determinants for COPD susceptibility that will provide insight into clinically relevant COPD subtypes.

(3) Distinct genetic determinants combined with social, physiologic and environmental factors influence the development of emphysema and airway disease.

COPDGene Phase 4

COPDGene has been funded for an additional five years under a contract issued by the US National Heart, Lung, and Blood Institute (NHLBI). The Multiple Principal Investigators (MPIs) are James Crapo, MD, National Jewish Health (NJH), and Edwin K. Silverman, MD, PhD, Brigham and Women's Hospital (BWH). The cross-sectional subject data collected at the baseline COPDGene study visit and longitudinal data from both 5 and 10-year follow up visits will be combined with the 15-year follow up to assess disease progression and incidence of COPD in smokers and as important additional endpoints for genetic and Omics association studies. In Phases 1 and 2, we

enrolled 10,718 subjects. In Phase 2, we completed 5,835 return visits and obtained information on an additional 574 subjects by other follow up methods. In Phase 3, we contacted 5,783 participants to date; this includes 3,457 in-person visits and 2,326 other types of contacts. For Phase 4, we will invite all COPDGene subjects to return for an in-person evaluation, 15 years after their initial visit, to assess disease incidence and/or progression. We anticipate approximately 3,500 subjects returning for the 15-year Phase 4 follow up study.

A summary of the visit procedures previously conducted in Phase 1, Phase 2 and Phase 3 are attached to this protocol as Appendix A (Phase 1), Appendix B (Phase 2), and Appendix C (Phase 3).

We will combine the information obtained in the previous 3 Phases of COPDGene with additional information collected during Phase 4 to meet our goals. This protocol covers the COPDGene Clinical Centers that perform study visits and collect information from study participants and Centers and Cores that process and analyze the data to meet the study aims.

1) **Clinical Centers (CC).** These 19 CC collect information from COPDGene participants during Phase 4 study visits and through other contacts, including longitudinal follow up and have conducted similar information during COPDGene Phases 1, 2 and 3. They provide information about mortality including medical records and informant interviews.
2) **Data Coordinating Center (DCC).** Located at National Jewish Health (NJH), the DCC accepts data from the CC and has stewardship of data from the previous 3 COPDGene Phases and Phase 4 in order to
   a) provide datasets to investigators with IRB-approved protocols for data analysis and publication,
   b) provide final datasets to NHLBI at the end of the contract period,
   c) perform searches for participants lost to contact,
   d) obtain information about those who have expired, as through National Death Index searches, and
   d) transfers data to NIH repositories as required by the NIH Data Sharing Policy.
3) **Administrative Core.** Located at NJH, the Admin Core oversees the implementation of Visit 4, including organizing, training, monitoring and evaluating the Clinical Centers, standardizing approaches to subject recruitment and study protocol, providing quality assurance of initial clinical data derived and convening Patient Advisory Boards
4) **Imaging Cores**. The Imaging Core at NJH, receives, stores, confirms de-identification, performs quality assessment and analysis of CT scans. The BWH Imaging Core performs additional analyses of CT scans.
5) **Biostatistics Core.** The Biostatistics Core at NJH provides technical oversight to the Data Coordinating Center and data analysis of primary study aims.
6) **Bioinformatics Core and Genetic Analysis Core**. Located at BWH, the Genetic Analysis Core analyzes genetic data, and the Bioinformatics Core performs quality control and analysis of Omics data and disease subtyping.
7) **Biomarkers Core**. Located at NJH, the Core analyzes Biomarker data.
8) **Subtyping Core**. Located at BWH, the Core analyzes COPDGene data to develop subtypes of the cohort.
9) **Epidemiology Core**. Located at the University of Colorado School of Public Health, the Center analyzes COPDGene epidemiological data.

10) **Pulmonary Function Core**. Located in Utah, this Core analyzes the quality of de-identified pulmonary function tests – spirometry and diffusing capacity.

II. SPECIFIC AIMS

This protocol involves 2 Task Areas for Phase 4 of COPDGene.

Task Area A: Collection of COPDGene study data and biospecimens and oversight of the COPDGene study.
Task Area B: Stewardship of Biospecimen and Data Repositories. Deliveries are the final datasets and biospecimen collections to be provided at NHLBI discretion at the end of the contract period.

The protocol addresses 2 specific aims for Phase 4:

Specific Aim 1: To Identify Imaging and Clinical Determinants of COPD Development and Progression Using Longitudinal Epidemiological Analysis
Data from all participants who return or have data for a 5-year, 10-year or 15-year visit and those who expire between the enrollment and the end of Phase 4 will be analyzed to confirm prognostication of clinical disease progression and mortality. We estimate that over 9,685 subjects will be analyzed in this aim - 6,756 subjects who 5-year visit data, 5,783 who have data from Phase 3, the estimated >2,920 who expire before the end of Phase 4, and the 3,500 participants we expect to return for a 15-year visit will be used to confirm prognostication of clinical disease progression.
In Phase 4 we will

1) develop a COPD Diagnostic System based on airway and emphysema phenotypes, and
2) improve longitudinal quantitative imaging with imaging noise adjustment, develop clinical calculators of COPD progression by modeling spirometric (forced expired volume in one second, FEV1) and imaging (emphysema) disease progression relevant for elderly individuals, deployed in a web-based interface.

Specific Aim 2: To Identify Genetic, Transcriptomic, Epigenetic, Metabolomic, and Proteomic Determinants of COPD Development and Progression Using Longitudinal Omics Analysis
All 10,718 subjects enrolled in COPDGene will be included to identify genetic, transcriptomic, and proteomic determinants of susceptibility and progression of specific COPD subtypes. All these subjects have had an initial visit but may or may not have participated in all the 5, 10 and 15-year follow-up visits. Through the longitudinal follow-up (LFU) program conducted twice a year since the initial visit, we have obtained follow-up data from 82.3% of the original 10,718 cohort by a combination of telephony, web survey and coordinator calls. We will continue LFU on all of the original 10,718 subjects who are alive and have not withdrawn from the study. All subjects will continue to be followed to assess vital status in Phase 4. We will perform death searches at least yearly to identify subjects that have expired. Death certificates, medical records and informant interviews will be obtained to determine cause of death, in an identical manner as in Phases 2 and 3.
In Phase 4, we will

1) identify omics (proteomic, transcriptomic, metabolomic, epigenetic, and genetic) biomarker panels that are associated with spirometric and imaging progression, and
2) identify sex-specific biomarkers associated with spirometric and imaging progression.

III. SUBJECT SELECTION - Phase 4

Phase 4 is the 15-year return visit for all available patients in the original COPDGene cohort.

All subjects to be studied in Phase 4 have been enrolled in the Genetic Epidemiology of COPD Study (COPDGene) in previous Phases 1 and 2. As part of their original COPDGene informed consent and subsequent consents at the 5- and 10-year follow-up visits, subjects have provided consent to be re-contacted for periodic follow-up and for other studies.

This project will invite all 10,718 subjects who are alive and still in the study to return for a Phase 4 study visit (15-year follow-up). At the time of writing this protocol, there are 2,909 deceased subjects. Phase 4 study visits will be scheduled 15 years ($\pm$ 6 months) after the Phase 1 enrollment visit or 10 years ($\pm$ 6 months) after Phase 2 for subjects enrolled in Phase 2. There must be at least 3 years in between Phase 3 and Phase 4 visits. Visits later than the scheduled return visit *are* permitted.

A key component of the project is to establish contact with all previously enrolled subjects who remain in the study and to confirm which subjects are deceased. Clinical Centers will utilize locally stored contact participant information including secondary contacts to establish contact. For Phase 4, subjects will be asked to allow personal contact information and Social Security numbers to be transmitted to and stored at the DCC. The Administrative Core may assist the Clinical Centers in locating and contacting subjects and to implement and maintain Participant Advisory Boards in order to gain their insight into study implementation and return of results, as applicable. The Administrative Core includes National Research Coordinators designated to assist CC research coordinators with all aspects of the study including on-site and remote training, quality assurance, contacting subjects, collecting death records and participating in collection of Phase 4 questionnaires. Personal information and Social Security information will be used to identify deceased subjects and cause of death from the National Death Index and other sources.

Eligible subjects will be contacted by mail, email and/or phone and invited to participate in Phase 4 at the Clinical Centers. The study protocol visit will be discussed in detail during an initial phone screening encounter; at that time, subjects will be administered an informed consent and will be asked pre-screening questions to determine if they are eligible to return for a study visit. Verbal informed consent will be obtained before collecting screening questionnaires. Before the in-person study visit procedures begin, written informed consent will be obtained by a research study staff member. A physician investigator from the Clinical Center will be available to answer any questions during the informed consent process.

Subject Inclusion/Exclusion Criteria – Phase 4 Return Visit

Inclusion Criteria

All previously enrolled participants in COPDGene are eligible for a Phase 4 Study Visit.

Temporary Exclusion Criteria

Subjects that meet the following will be temporarily excluded but may be rescheduled at a later time:

1. Pregnancy or suspected pregnancy. These subjects will be re-scheduled for a visit at least three months after childbirth or after they are no longer suspected of pregnancy. Given the minimum age of the cohort, 60 years old, at the time they would be eligible for a 15-year visit, it is expected that there will be no subjects who have this exclusion criterion.
2. Use of antibiotics for a respiratory infection within the past month. These subjects will be rescheduled for a visit at least 4 weeks after completion of the antibiotics.
3. Use of new prescription or increased dose of systemic corticosteroids for an acute respiratory disorder within the past month. These subjects will be rescheduled for a visit at least 4 weeks after completion of the steroids.
4. Chest or abdominal surgery in the past three months. These are contraindications to spirometry. These subjects will be rescheduled for a visit at least three months after surgery.
5. Heart attack in the last three months. This is a relative contraindication to spirometry. These subjects will be rescheduled for a visit at least three months after the heart attack.
6. Detached retina or eye surgery in the past three months. This is a contraindication for spirometry. These subjects will be rescheduled for a visit at least three months after surgery.
7. Hospitalization for a heart problem in the past month. These subjects will be rescheduled for a visit at least 4 weeks after hospital discharge.

Subjects with lung transplants will be excluded from a follow-up study visit; however, they will be invited to participate in long-term follow up to assess mortality.

Rationale: All previously enrolled subjects from COPDGene are eligible for the study. Previous exclusions for lung cancer and other conditions are not a reason to exclude subjects from follow-up evaluation because these are outcomes to be assessed.

We will temporarily exclude patients with recent use of antibiotics and/or systemic steroids for a COPD exacerbation since such subjects may have an ongoing inflammatory pulmonary process related to an infection that will impair our ability to characterize and phenotype patients in their usual stable state. We will temporarily exclude subjects who have a contraindication to spirometry (exclusion criteria 4 – 7 above).

Subjects with temporary exclusion criteria will be re-screened at a later date when they can be enrolled in Phase 4.

IV. SUBJECT SCREENING - Phase 4

All subjects previously enrolled in Phase 1 have provided consent with their initial study enrollment to be contacted again for further research studies. Subjects will be sent an invitation

letter advising them that funding has been provided for the Phase 4 visit and informing them that they will be contacted to schedule an appointment. Subjects will be asked to provide their contact information to ensure they have been previously enrolled in COPDGene. A time will be scheduled for their next study visit and the subject will be asked to bring the following to the Study Visit:

1) all current oral and inhaled medications,
2) names of any injectable study medications received on a regular basis,
3) social security number,
4) name, address, and phone number of their personal physician,
5) name, address, and phone of two other individuals likely to know the subject's whereabouts, at least one of whom is a relative not living with the subject, and
6) driver's license or other identification to ensure they are the individual previously enrolled.

Subjects will be given options for the Phase 4 Study Visit as described below. Screening questions will be asked upon scheduling a Study Visit to confirm that the subject is eligible for a visit.

Final eligibility will be confirmed prior to conducting the study visit.

V. PHASE 4 STUDY VISIT

Participants will sign a new informed consent for the Phase 4 Study Visit. If subjects cannot participate in the Phase 4 visit, limited information about their health status will be collected from the subject or a caregiver, as appropriate, at each of the visit types outlined below.

For the Phase 4, we will provide options to participants for the Study Visit. The participant will be offered the opportunity to have 2 types of Study Visits:

1) Single/comprehensive Study Visit. Questionnaire administration and other procedures will be done during an In-Person Study visit (see below), or
2) Remote (Pre Study Visit) and In-person Study visits. Questionnaires will be administered remotely via videoconference (preferred) or phone in 1 or 2 sessions (second session available as needed to limit participant fatigue in our aging cohort) followed by an in-person study visit. The remote visit for questionnaire administration is called the Pre Study Visit.

The option of allowing participants to select the visit easiest for them will enhance participation and limit participant burden.

Single/comprehensive study visit. Our experience in the 15 years of the COPDGene study has confirmed the ability of the CC to schedule and efficiently perform all questionnaires and procedures on a single visit whenever possible to limit subject burden related to additional visits. The vast majority of subjects (93%) have had one visit in Phases 1 and 2. Some subjects have had a second visit, dependent on the ability of the CC to schedule the chest CT scan on the same day that the spirometry and questionnaires are obtained. In the rare circumstances where spirometry has failed quality assurance, subjects have returned for repeat spirometry. At some CC, the CT scanner is at a site remote from the location used by the study coordinators for performance of spirometry and study questionnaires. In such cases, a second visit for performance of the chest CT scan has made the most efficient use of subject time. In rare cases, subjects will have up to three

visits over three months to the CC in order to complete the study protocol. The following situations may extend the number of visits beyond one.

1. If the chest CT scan cannot be scheduled simultaneously with other assessments, the CT scan will be performed on another date. This situation can occur if the CT scanner is at a remote location from questionnaire completion, spirometry and other procedures. In other cases, the CT scanner may be off-line due to malfunction or repair or may be unavailable due to an unexpectedly heavy burden of clinical cases.
2. If the subject returns for a Visit and has had a COPD exacerbation and/or respiratory infection within the previous 30 days, the visit will not be conducted. As in Phase 1, 2 and 3, we will limit this possibility by contacting subjects by phone prior to the scheduled visit to assure the required time has elapsed following an exacerbation. If subjects arrive at the visit within 30 days of an exacerbation, they will be rescheduled for a visit to occur at least 30 days after an exacerbation is completed (completed is defined by completion of antibiotic and/or systemic corticosteroid medications prescribed for an acute respiratory cause, such as an infection or COPD exacerbation). This stipulation is identical to the inclusion criteria for the initial COPDGene enrollment visit and 5 and 10-year follow-ups. In order to compare the information collected at previous visits, we require that subjects are in a similar stable condition for all visits.
3. If the subject has to return for repeat spirometry or questionnaire completion, a second visit may be necessary. Our Quality Assurance Program evaluates questionnaire completion. In the case of inadequate questionnaire completion, we will ask the CC to obtain requisite additional information and complete questionnaires over the phone. Spirometry tests undergo a rigorous quality control process. In cases where optimal spirometry tracings have not been obtained, we require an additional visit for repeat spirometry. In Phase 3, only 8 subjects thus far have had to return to repeat spirometry because of poor quality on the initial test.

Remote Pre Study Visit and In-Person Visit. During the pandemic, CC were required to cease in-person study visits to avoid SARS-CoV-2 spread to participants and research staff, particularly since spirometry is considered a potential aerosol-generating procedure. In order to continue data collection, IRB approval was obtained to collect questionnaires remotely via videoconferencing or phone. A total of 2,180 COPDGene participants had remote questionnaire administration. In-person visits, either for a) participants who had remote questionnaire visits to collect data that required the participant to be on-site, or b) as complete in-person visits for questionnaires and on-site data collection, were done when the pandemic lessened, as allowed by individual CC institutional policy. Procedures requiring on-site visits include spirometry, diffusing capacity, chest CT scan, sit-to-stand test, handgrip strength, 6-minute walk, mini-Cog, height, weight, arm span and oxygen saturation, and blood collection is optimally done on-site to allow rapid biospecimen processing. Participants were willing to perform questionnaires remotely and CC felt this process worked well. In most cases, questionnaires could be administered in one session. At the request of the participant, such as for participant fatigue, 2 remote sessions were conducted to administer all questionnaires.

Whenever possible, all evaluations for Phase 4 will be completed within a one-month time frame. These targets may be altered by the SARS-CoV-2 pandemic or other unforeseen circumstances.

Visit 2 Options for Subjects Who Have Difficulty with a Visit to Clinical Center

These options have also been in place for the previous COPDGene Phases.

Transfer to a closer COPDGene Clinical Center. In the event that a subject has moved away from their local Clinical Center and to an area closer to another participating COPDGene Clinical Center, the subject may elect to complete Visit 2 at this closer Center rather than traveling to their original Center. With the subject's verbal approval by phone, their new contact information will be provided from the Clinical Center where the subject had the first visit to the geographically proximate Clinical Center. The subject will sign a new Informed Consent Form and HIPAA with the coordinator at the new study location prior to their Phase 3 visit.

Travel assistance. In the event that a subject has financial hardship or has moved a considerable distance from their local Clinical Center, the COPDGene Administrative Core may consider providing transportation assistance to the CC for the subject for a return visit. The payment for this subject's return visit may include the actual reasonable cost of transportation, lodging, and meal expenses. Subject must provide receipts to Clinical Center for reimbursement. Use of this approach and funding for these expenses will be allocated on a case-by-case basis from the Administrative Core to the Clinical Center. Centers must contact the Administrative Core prior to the study visit to access the travel support.

COPDGene subjects who are unable to come to a COPDGene Clinical Center to complete a Phase 4 visit due to illness, disease progression, or inability to be transported to the Clinical Center, may be given the option of having a research coordinator complete a partial study visit at the subject's place of residence. The availability of a home visit will be at the discretion of the CC, given resources and local policies. The home visit will include spirometry before and after albuterol administration, blood drawing, and completion of all study questionnaires.

Complete Phone Visit. Phase 4 subjects who are unable to come to a Clinical Center and are not willing or able to undergo a home visit will be asked if they would complete study questionnaires over the phone (Complete Phone Visit). If the subject agrees, this Phone Visit will be completed.

Limited Follow-Up. If the subject is not willing or unable to complete study questionnaires over the phone they will be ask if they would complete a Limited Follow-Up Questionnaire and some or all of the regular study questionnaires. The Limited Follow-Up Questionnaire will be used to gather important longitudinal information on the subject related to respiratory symptoms, exacerbations, and comorbidities. A research coordinator will conduct this questionnaire with the subject by phone. These subjects may still participate in the bi-annual Longitudinal Follow-Up program by phone or Internet.

Subjects who do not agree to a Phase 4 visit or to participate in Longitudinal Follow-Up will be asked if they would be willing to complete a Limited Follow-Up Questionnaire by phone. If the subject refuses to complete this questionnaire, their wishes will be respected.

Home Visit. Previously enrolled COPDGene subjects who are unable to come to a COPDGene Clinical Center to complete the Phase 4 visit due to illness, disease progression, or inability to be transported to the Clinical Center, will be given the option of having a research coordinator

complete a partial study visit at the subject's place of residence., if allowed at the Clinical Center. It is recommended that 2 coordinators attend Home Visits to assure subject and coordinator safety.

If a COPDGene home visit is performed, it will include Home Visit Informed Consent and HIPAA administration, pre- and post-bronchodilator spirometry, blood draw, hand-grip strength, sit-to-stand test, and completion of all study questionnaires and procedures, *except* the six-minute walk, diffusing capacity, and CT scan; participants may return to the Clinical Center for these tests at a later time, if they are able to do so.

VI. PHASE 4 STUDY PROCEDURES

Subjects will be asked to provide affirmative information to ensure they have been previously enrolled in COPDGene and are correctly associated with their study ID number:

1) driver's license or other identification to ensure they are the individual previously enrolled; a copy of the identification should be kept in the subject's study folder locally at the Clinical Center,
2) verify full name, date of birth, current address (and previous address at the time of the last COPDGene Study Visit), phone number, email, and
3) social security number.

Clinical Center Staff must confirm identification using full name, date of birth and address at each contact before any study procedures are performed or data is collected. The identification materials must be compared/verified to stored identification for the participant. We will update participant identifiable information (address phone, email, secondary contacts and Social Security number) previously collected from COPDGene participants. This information is used to reimburse participants for their time, contact subjects for scheduling in-person visits, semiannual longitudinal follow-up (COPD exacerbations, newly diagnosed medical conditions), obtain information about mortality (National Death Index search, death certificates, informant interviews) and find participants who have moved or changed phone numbers. This information will be stored at the DCC so that the Administrative Core National Research Coordinators can assist CC in locating subjects who may have relocated, conduct longitudinal follow up calls, and search for death records. In addition, the personal identifiers of subjects who have participated in any of the previous phases of COPDGene but have not had a Phase 4 visit will also be transferred to the DCC to assist in locating those participants. We will also collect hospital and medical records to determine the cause of death. At the enrollment visit, we assigned each participant a unique study alpha-numeric identifier (study ID); that identifier was also used in Phases 2 and 3. We will continue to use the study ID and not use personal identifiable information on all data (e.g., questionnaires, procedures, CT scans) collected and COPDGene study records.

Subjects will also be asked to provide name, address, and phone number of their personal physician so that clinically significant study visit results can be shared with their personal physician. Subjects will also be asked to provide name, address, email and phone of two other individuals likely to know the subject's whereabouts, as Secondary Contacts, at least one of whom is a relative not living with the subject so that the participant can be located for longitudinal follow up phone calls or in the case the participant expires. Subjects will be asked to confirm or update the name of an individual who is their next of kin for requesting death certificate and medical records after death.

A letter/memo will be provided to the participant that describes their participation in the COPDGene study and their willingness to provide medical records and information in the event of their death. This memo should be shared with their physician, Secondary Contacts and next of kin.

Study Visit Procedures

At the Phase 4 visit, study staff will review the informed consent form with the subject and obtain written informed consent. The subjects will have the opportunity to review the informed consent in private, have questions answered and sign the informed consent. If necessary, subjects may review the informed consent with family members or their physicians. A separate HIPAA or combined consent and HIPAA will be reviewed and signed at the same time. Although subjects have signed a HIPAA in Phase 1, Phase 2 and Phase 3, a current authorization will be obtained to meet the requirements of hospitals and health care providers in order to determine cause of death. Phase 4 data collection includes transfer of subject identifiers to the Data Coordinating Center (DCC) and Administrative Core so that missing patients can be located and National Death Index searches can be completed to determine cause of death.

During the Phase 4 research study visit, the following procedures will be performed.

1. Informed Consent for participation in this proposed study. Subjects will be asked to sign a new Informed Consent for Phase 4 of COPDGene. The new consent will be based on the previously used Informed Consent and obtain subject permission to collect the identical information evaluated at the previous three visits and additional information.

2. Demographic Information - This information has already been collected in all previous COPDGene study visits and will be updated during the Phase 4 visit. All subjects will have their demographic information updated - name, address, e-mail address (if available), home telephone, cell phone, and date of birth. Subjects will also be asked their sex at birth, as required by the NIH, and gender identity. Contact information will be updated for two other people - family members not living with the subject or close contacts who may be knowledgeable about the subject in the event that the subject cannot be contacted for subsequent longitudinal follow-up or study visits. Subject and contact identifying information will be stored at the local Clinical Center in a password-protected database and locked filing system. As approved by the sIRB and local Institutional Review Board for Human Subjects, contact information and social security number for all subjects is transmitted to the Administrative Center for central searching of: 1) death records in the National Death Index, and 2) updated information to assist in finding and contacting subjects for return visits and longitudinal follow-up.

3. Questionnaires
    a. Eligibility Assessment - to assure the subject has not had a recent COPD exacerbation in order to perform the visit when the participant is free of acute respiratory illness.
    b. Safety Assessment - prior to administration of albuterol, including heart rate and blood pressure, and history of adverse events with albuterol. Questions about cardiac disease are used to assess patient safety for the six-minute walk test and to assess safety for administration of bronchodilator (albuterol) for spirometry testing.

c. Medical History – to assess the presence of other medical conditions identified since the last study visit.
d. Medication History – to record current respiratory and other prescription medications
e. Respiratory Symptoms - assessed with the previously used COPDGene versions including modified American Thoracic Society-Division of Lung Diseases Respiratory Questionnaire (3), presence and degree of dyspnea by the Medical Research Council Dyspnea Scale (4, 5), frequency and severity of COPD exacerbations, presence or absence of chronic bronchitis (chronic cough and chronic sputum production), smoking history in the last 5 years, and presence of other lung diseases.
f. Health Status - assessed with the questionnaires currently being used in COPDGene. The respiratory disease-specific St. George's Respiratory Questionnaire provides a total score and sub-scores for activities, symptoms and impacts (6). The general health status SF-36 allows comparison with other chronic respiratory conditions and comorbidities and includes a total score and physical and mental sub scores (7). The COPD Assessment Test (CAT) includes items related to symptoms and COPD-specific health status (8).
g. Hospital Anxiety and Depression Scale (HADS) – to assess symptoms of anxiety and depression as they are associated with exacerbations, health status and mortality (9).
h. Women's Only Questionnaire, using the questionnaire used at the 10-year visit, to assess hormonal use, menstruation onset and pregnancies, since these factors may be associated with COPD development or progression.
i. Residential and Occupational Questionnaire - to assess influence of environmental and occupational factors in the last 5 years on development and progression of COPD. We will use the identical questionnaire of residence and occupation used at the 10-year visit.
j. COVID-19 questionnaire – to assess COVID-19 infection and severity that may affect COPD development and progression, we will harmonize our questionnaires with those used in the Collaborative Cohort of Cohorts for COVID-19 Research (C4R) in which COPDGene is participating along with other NIH cohorts.
k. Frailty questionnaires – to assess physical function and frailty, the DASI (Duke Activity Status Index) questionnaire and PROactive physical activity questionnaire and questions about falls will be completed (10-12).
l. Mini-Cog - to assess cognition (13).
m. Fall Questions – Falls are common in older people and lead to hip fractures, immobility, traumatic brain injuries and mortality. Falls may thus impact COPD outcomes. We will have selected questions recommended by the CDC to assess falls (https://www.cdc.gov/steadi/pdf/STEADI-Algorithm-508.pdf).
n. Social Determinants of Health – Social determinants of health by the CDC/WHO definition include social and community context which we have not assessed in the past Phases of COPDGene. In Phase 4, we will add a single question to assess social support to limit participant burden.
o. Periodontitis questions – In some studies, periodontitis is associated with COPD exacerbations and COPD exacerbation severity. We will use one question to assess if participants are edentulous. We considered more extensive questionnaires but chose not to use one to limit participant burden and the gold standard for periodontitis is an oral exam by a dentist, which is costly and not readily available.

4. Physical Examination
    a. Height and weight - for calculation of Body Mass Index and for use in calculating predicted normal values for lung function.
    b. Arm span – for assessment of predicted lung function in subjects who lose height due to aging.
    c. Waist circumference – to assess adiposity.
    d. Pulse oximetry at rest – to assess oxygenation.
    e. Heart rate and blood pressure - to assess safety for spirometry and the walk test. Blood pressure will be measured 3 times with the subject seated at rest.
    f. Handgrip strength.
    g. Sit-to-stand, number of repetitions in 30 seconds – as a measure of lower extremity strength and endurance. This test will be performed as in Phase 3, and during this test we will also measure the time to rise 5 times as a frailty measure.

5. Blood sample – Blood will be drawn from all subjects for DNA (genetic and epigenetic studies), serum and plasma (for measurement of proteins, metabolites, and other analytes potentially related to COPD and other diseases), RNA (transcriptomic studies), and a CBC (complete blood count). A total volume of up to 50 ml of blood will be drawn for this study. A complete blood count will be performed at the Clinical Centers. The rest of the blood samples will be stored at the COPDGene Biological Repository at Brigham and Women's Hospital. We will ask the subject what time of day they last ate and/or drink anything other than water, including candy/chewing gum, and collect this information.

6. Six-Minute Walk Distance – the maximum distance walked in six minutes will be assessed following American Thoracic Society (ATS) guidelines for the six-minute walk test (16). Heart rate and pulse oximetry are recorded immediately at the end of the test and 1 minute after the test. During the six-minute walk, the time to walk 15 feet will also be measured to assess slowness of gait that has been used an index of frailty.

7. Spirometry will be performed in the same manner and with similar updated equipment as in COPDGene Phases 1, 2 and 3. Pre and post-bronchodilator (short-acting beta-agonist, albuterol), spirometry will include measurement of forced vital capacity (FVC), forced expiratory volume in one second (FEV_1), forced expiratory volume in six seconds (FEV_6), ratio of FEV_1/FVC, and ratio of FEV_6/FVC as well as full expired flow volume curve. Spirometric predicted values are determined based on the prediction equations in non-Hispanic white and African-American subjects developed by Hankinson from NHANES (17), for consistency as in COPDGene Phases 1, 2 and 3. We will also compare lung function to the Global Lung Function Initiative (GLI) normal values (18). Testing will be performed in all subjects using the EasyOne® spirometry system (NDD Medical Technologies, Andover, MA) following American Thoracic Society/Respiratory Society (ATS/ERS) testing standards used in all previous Phases of COPDGene (19). For post-bronchodilator testing, two puffs (180 mcg) of albuterol will be administered through an Aerochamber® spacer to assure optimal delivery of the medication. Twenty minutes after albuterol administration, spirometry will be done to ascertain post-bronchodilator FEV_1 and FEV_1/FVC to be used for determining the presence of airflow limitation and for COPD staging according to GOLD Spirometry Grade. Spirometry will be performed in the sitting position with nose clips. Infection control guidelines will be followed with a new disposable mouthpiece used for each subject, and with expiratory filter to prevent virus spreading during

the SARS-CoV-2 pandemic. At least three acceptable efforts will be performed to obtain results that meet reproducibility criteria, and the best effort will be used for analysis. Quality control will be performed as in Phases 1, 2 and 3 by the Pulmonary Function Quality Assurance Center headed by Dr. Robert Jensen, Professor, University of Utah, with secondary review of failed spirometry tests by Drs. Make and Silverman.

8. Diffusing Capacity for Carbon Monoxide – using the identical equipment used at the 5- and 10-year visits (EasyOne Pro®, NDD Medical Technologies, Andover, MA). Testing will be performed in accordance with established American Thoracic Society/European Thoracic Society (ATS/ERS) standards (20). Quality control will be performed by the Pulmonary Function Quality Assurance Center headed by Dr. Robert Jensen, with secondary review of failed tests by Drs. Make and Silverman.

9. Chest CT Scan Acquisition – Chest CT scans will be acquired at full inspiration (total lung capacity) and at the end of a normal exhalation (functional residual capacity), with standardized breathing instructions. In Phase 4 of COPDGene, low dose CT scans will be obtained in an identical manner to CT scans collected in Phase 3. The reduced dose was previously chosen to reduce potential subject adverse effects, translate COPDGene findings to clinical practice, and allow comparison of COPDGene scans with scans performed for clinical reasons including lung cancer screening CT scans. Use of low dose CT scans is critical to be able to apply the knowledge gained from COPDGene to the large population of cigarette smokers who will undergo reduced dose scanning for lung cancer screening (estimated at up to 10 million eligible smokers) (21). In devising the inspiratory CT scan protocol used in Phase 3 and to be used in the current proposal, we received valuable input from all of the major CT vendors, several expert physicists, and the Quantitative Imaging Biomarkers Alliance (QIBA) of the Radiology Society of North America (RSNA). This protocol has been endorsed by QIBA as a harmonized protocol for lung density assessment. Based on our measurements on 2,284 subjects in Phase 3, the protocol for the CT scan will result in a mean effective CT dose of 3.5 mSv. Scans will be anonymized locally at the CC and transferred electronically to the Quantitative Imaging Laboratory at National Jewish Health, where they will undergo quality evaluation for technical adequacy, absence of motion or other artifact, and assessment of significant visible changes from the prior studies. De-identified CT scans will be securely transmitted to the BWH Imaging Center for further analysis. A local Clinical Center radiologist will perform a clinical interpretation of the scan. The clinical interpretation will be provided to the CC research coordinator and PI. Clinically significant findings, based on the radiologist report and review of the scan and report by the CC PI, will be communicated to the patient and their physician.

10. COVID-19 infection and vaccination status - We will determine if participants have had a COVID-19 infection to evaluate the effect of a SARS-CoV-2 infection and its severity on the natural history of lung disease in our cohort. We have experience in assessing COVID-19 status in the COPDGene cohort through our participation in the NHLBI Collaborative Cohort of Cohorts for COVID-19 Research (C4R) and many of our investigators are also participating in the NIH Researching COVID to Enhance Recovery (RECOVER) study. Both of these studies use questionnaires and blood tests to determine the presence of previous COVID-19 infections. We propose primarily identifying past COVID-19 infections by either 1) a positive SARS-CoV-2 nucleocapsid IgG antibody from dried blood spot testing during COPDGene Phase 3 for the C4R study or from serum samples obtained at the time of the

Phase 4 visit, or 2) a history of a positive SARS-CoV-2 PCR test. At the Phase 4 visit, we will send a serum sample to Dr. Russ Tracy's laboratory for COVID-19 serology testing. This analysis will include testing for SARS-CoV-2 nucleocapsid antibody to identify participants who have had COVID-19 in order to determine the effect of COVID-19 on COPD development and progression. We will not rely solely on antibody testing as antibodies wane over time. We are currently assessing the history of a COVID-19 infection in C4R and RECOVER by use of questionnaires and are collecting medical records for those who have been hospitalized. We propose using our current questionnaire during the Phase 4 visit as a secondary definition of a COVID-19 infection. Patients may have had symptoms of a COVID-19 infection but not had a confirmatory blood test. Early during the pandemic, diagnostic testing was not widely available, patients were reluctant to get tested, and with more mild acute respiratory symptoms, patients may not have reported these to medical personnel. We will collect information about symptoms that may have been related to an acute COVID-19 infection when a confirmatory diagnostic test was not obtained. The severity of a COVID-19 infection may play a role in the effect of COVID-19 on COPD. For example, patients receiving mechanical ventilation for an infection may have a greater impact on COPD progression than those with mild infections that do not require oxygen or hospital admission. We are currently evaluating COVID-19 infection through our Longitudinal Follow Up.

11. Return of results. Results of CBC, chest CT scan and spirometry will be sent to the subject after the visit so they may share them with their primary health care provider. If results from the CBC or chest CT scan are felt to be clinically significant by the CC PI, they will also be sent to the subject's primary health care provider. We do not currently plan to identify and return clinically significant results from whole genome sequencing, although that may be considered through a future ancillary study. In Phase 3, information about potential hepatitis C infection was provided to the participants through an IRB-approved process. Return of information about potential HIV infection is now being proposed.

Additional Ongoing Study for COPDGene Participants as part of the NHLBI C4R Project:

Collaborative Cohort of Cohorts for COVID-19 Research (C4R) Protocol
Wave 3:
Date: 23 January 2023

1. Overview

C4R was originally funded to collect data on the COVID-19 pandemic 2020-2022. COPDGene participated in collecting Wave 1 and Wave 2 questionnaires from COPDGene participants and dried blood spots. Information from the C4R questionnaires and dried blood spots is used by COPDGene and include in our database. In light of the unprecedented, generalized, ongoing public health impact of the COVID-19 pandemic in the US, C4R has been extended by NIH to continue ascertainment and characterization of COVID- and pandemic-related health impacts in C4R over 2023-2025 via a third questionnaire (Wave 3 questionnaire), and ongoing events ascertainment. The Wave 3 questionnaire is being submitted to the COPDGene sIRB.

2. Study Purpose and Rationale

The effects of COVID-19 on US health, economy, and society, are widespread, deep, and will continue well beyond the initial wave of infections.[1] The Collaborative Cohort of Cohorts for COVID-19 Research (C4R) aims to advance our knowledge on the impact of the COVID-19 pandemic. C4R will leverage longitudinal data and deep phenotyping in fourteen US prospective cohort studies based at 36 academic medical centers to ascertain all cases of COVID-19 infection, to evaluate factors relevant to COVID risk and resilience, and to track the prognosis of both those with and those without infection. It will do this by questionnaire administration, and ascertainment of hospitalization and death records.

3. Study Aims and Hypotheses

The primary goals of the study are as follows.

1. **Ascertainment of SARS-CoV-2 infection and COVID-19 illness**. In order to generate a community-based estimate of the incidence of COVID-19 events and an inception cohort of COVID-19 cases relatively free of referral and disease spectrum biases, the C4R cohorts will:
    a. Accomplish COVID-19 case ascertainment through a standardized questionnaire deployed to study participants through phone, mail, or digital (online, email, smartphone) contact. Participants will have questionnaire-based COVID-19 case ascertainment assessed at least three times. Assessment of fatigue, depression, and anxiety will be performed at 1 or more time points.
    b. Leverage established individual cohort infrastructure for medical record ascertainment of COVID-related events, including electronic medical records (EMR) as available.
    c. Detect SARS-CoV-2 seroconversions through 0.25-0.50 cc serum aliquots from previously collected blood during Phase 3 of COPDGene, which will be analyzed by the C4R Biorepository and Central Laboratory (BCL).

VI. OTHER PHASE 4 STUDY PROCEDURES

Mail Consent

Prior to the Phase 4 return visit, subjects who were not seen in Phase 2 or 3 may be contacted by the Clinical Center via mail to request additional consent and authorization. This additional consent and authorization will request that the subject allow transfer of identifiers from the Clinical Center to the Data Coordinating Center at National Jewish Health. This Mail Authorization provides permission for the Clinical Center to transmit the subject's personal information (social security number, address, secondary contact address) to the DCC. A Personal Contact Information Update form will be sent with the mail consent, asking subjects to update contact information, including secondary contacts and primary care physician. Additionally, the subjects will be invited to sign an "Authorization to Release Protected Health Information" medical record release form.

This information will be used for central subject tracking, longitudinal follow up, vital status searching, and to obtain cost and utilization information on subjects from Medicare. The mail consent also allows the Clinical Center to seek perimortem medical records and informant interviews about the circumstances around the subject's death.

Phase 4 Options for Subjects Who Have Difficulty with a Visit to Clinical Center
In the event that a previously enrolled subject has moved away from their local Clinical Center and to an area closer to another participating COPDGene Clinical Center, the subject may elect to complete Phase 4 at this closer Center rather than traveling to their original Center. With the subject's verbal approval by phone, their new contact information will be provided from the Clinical Center where the subject had the first visit to the geographically proximate Clinical Center. The subject will sign a new Informed Consent Form and HIPAA with the coordinator at the new study location prior to their Phase 4 visit.

In the event that a subject has financial hardship or has moved a considerable distance from their local Clinical Center, the COPDGene Administrative Core may consider providing financial transportation assistance to this subject for a return visit. The reimbursement for this subject's return visit may include the actual reasonable cost of transportation, lodging, and meal expenses. Subject must provide receipts to Clinical Center for reimbursement. Use of this approach and funding for these expenses will be allocated on a case-by-case basis from the Administrative Core to the Clinical Center.

Subjects who are unable to come to a Clinical Center and are not willing or able to undergo a home visit will be asked if they would be willing to complete study questionnaires over the phone. If the subject agrees, this will be completed. The subject should be compensated $50 for their time and effort in completing the questionnaires over the phone. If the subject is not willing or unable to complete study questionnaires over the phone they will be asked if they would complete a Limited Follow-Up Questionnaire. The Limited Follow-Up Questionnaire will be used to gather important longitudinal information on the subject related to respiratory symptoms, exacerbations, and comorbidities. A research coordinator will conduct this questionnaire with the subject by phone. These subjects may still participate in the semi-annual Longitudinal Follow-Up program by phone or Internet.

Closed Clinical Center
Due to the longitudinal nature of the study, it is important to follow health and vital status on all of the COPDGene subjects. In the event that a clinical center is unable to perform follow-up contacts or visits, the COPDGene Administrative Core will work as an agent for the Clinical Center or provide a contract clinical research organization to manage the follow-up subject visits and contacts per protocol. Some subjects have been consented in Phase 2 or 3 to allow for a member of COPDGene Central Administration to contact them for follow-up and this has been tracked by the DCC. The original clinical center must work with the Administrative Core to facilitate efficient transfer of information and study management under the supervision of a local principal investigator and with the approval of the local IRB.

Longitudinal Follow-Up
Subjects will continue to be contacted as they were in Phases 1, 2 and 3, up to four times per year by telephone, mail, or email for up to 10 years after their Phase 4 consent. Questions will be asked about current health status, exacerbations, cancer, new illnesses or medical conditions and current smoking status. If lung cancer is diagnosed, a clinical coordinator may contact the subject via phone or email to ask them to provide medical records to aid in the completion of a Cancer Data Collection Questionnaire, to collect additional data on the lung cancer diagnosis. Subjects may be contacted on no more than three additional occasions per year to inform them of other research

studies and to update them about results of the COPDGene study. The longitudinal follow-up contact mechanism is primarily based on automated contacts to subjects via a computer server controlled by the local clinical center in which the clinical center securely uploads subject contact data and social security numbers to a secure server using secure sockets (SSL) technology and 128 bit or greater encryption with an HTTPS protocol. Subjects establish a preference for contact by email and web data entry or automated phone calls with data entry by telephone keypad. Subjects can also request only coordinator-conducted surveys. Subject contact information and identifying information is deleted automatically from the server after the contact is made or at the end of three weeks. The longitudinal follow-up telephone process only includes contacting the subjects. Data collected from all subjects are de-identified and made available to the Data Coordinating Center at National Jewish for storage and analysis. Subjects who fail to respond to automated contacts are contacted by the local clinical coordinator who then asks the questions of the subject and inputs the information obtained into the web-based data collection form. Subjects who become lost to follow-up from the longitudinal follow-up process will be traced by their secondary contacts and searched for in the Social Security Death Master File or National Death Index. Deaths identified from the longitudinal follow-up system are communicated to the Data Coordinating Center and the local clinical center. The local clinical center initiates collecting further information about the death as described below.

Mortality Assessment

Death is an endpoint of interest that will be analyzed as part of this protocol. The opportunity to assess vital status will occur in one of several ways. It is possible that during the process of longitudinal follow-up, the clinical center will be made aware of a subject's death. However, some subjects will be lost to follow-up with their vital status unknown. For these subjects, staff at the local clinical center or administrative personnel at National Jewish Health (if subject identifiers have been transferred) will conduct a search to determine vital status using subject's stored social security numbers. This search may include a National Death Index query, Social Security Death Index query and general internet search including obituary postings. Once vital status has been confirmed as dead, the clinical center will obtain additional information regarding the subject's death so that cause of death may be determined. This additional information will include: 1) Informant Interview conducted with next of kin or other close contact of the study participant or the subject's physician, 2) Death certificate, 3) Hospital Discharge Summary if it is determined that the death occurred in hospital or within 3 months of a hospitalization, and 4) Request for recent treating physician records if death did not occur in the hospital.

The above records will be de-identified except for date of death and dates of hospital admission and transferred to the Data Coordinating Center. A Mortality Adjudication Committee will review these records centrally. The data will be reviewed by the Mortality Adjudication Committee and a final cause of death will be assigned.

VII. BIOSTATISTICAL ANALYSIS

In Phase 1, genome-wide association analysis was performed using genome-wide SNP genotypes obtained in the entire COPDGene study population. In Phase 2, exome chip genotyping, and whole genome sequencing data were obtained. In Phase 3, additional RNA-seq, DNA methylation, metabolomic, and proteomic data were obtained.

For longitudinal epidemiological analysis, the primary outcomes for COPD development and progression are changes in FEV_1, CT emphysema [Adjusted Lung Density (PERC15 adjusted for lung volume)], and all-cause mortality. Secondary outcome measures include respiratory mortality and respiratory exacerbation frequency, as well as changes in CT airway wall thickness, CT gas trapping, six-minute walk distance (6MWD), and health-related quality of life.

Longitudinal analyses will be performed with these primary and secondary outcomes using measures from Phase 1, 2, 3 and 4 visits. Linear mixed models will be used to fit outcomes that are approximately normally distributed (or that can be log transformed to such); generalized linear mixed models will be used for binary and count outcomes. Repeated measures will be accounted for in these models using either subject-level random effects, a non-simple error covariance structure, or both. Random effects will also be included for study center and CT scanner model, when applicable. Time-to-event outcomes will be modeled using parametric (e.g., Weibull), semi-parametric (e.g., Cox regression) or nonparametric (e.g., Kaplan-Meier) survival models. Key predictor variables to be assessed for in these longitudinal or time-to-event models include age, gender, current smoking status (at each visit), pack-years, baseline lung function, baseline CT emphysema and airway measures, respiratory symptoms (e.g., chronic bronchitis, dyspnea), respiratory medications, and co-morbid illnesses (including lung cancer and coronary artery disease). Right-skewed variables will be log transformed. We will compare the demographic and clinical characteristics of subjects who have follow-up data to those subjects missing due to death and those subjects missing due to loss to follow-up. Longitudinal analysis will be performed using time-invariant groups and allow them to have their own trajectory. Covariates in these analyses include change in current smoking status and change in pulmonary medications. Disease progression and mortality will be simultaneously assessed to limit potential bias introduced by differential survival. This analytical approach considers the joint distribution of the vector of repeated measures (i.e., longitudinal progression of PFT abnormalities and CT measures) simultaneously with time to event (i.e., mortality) using the method developed by Hogan and Laird (10).

Power to detect differential 15-year changes between COPD subtypes was assessed for FEV1 and CT adjusted lung density (ALD). Extrapolating from previous COPDGene phases, we anticipate 3500 Phase 4 subjects will include 665 GOLD 2, 350 GOLD 3, and 210 GOLD 4 subjects. Using emphysema-predominant (EPD) and airway-predominant (APD) COPD defined by quantitative emphysema measures in GOLD 2-4 subjects, we anticipate 450 GOLD 2-4 subjects will provide $\geq$ 80% power to detect differences between EPD and APD in FEV1 decline and change in ALD. Calculations were based on expected 15-year drops in FEV1 of 8% predicted for APD and 12% predicted for EPD (SD=15%), and expected drops in ALD of 10 g/L for APD and 5 g/L for EPD (SD=18 g/L). Calculations used a=0.05 and were based on two-sample t-tests using a common variance. Thus, we should have adequate power to detect differences in disease progression between EPD and APD subjects.

The primary phenotypes for genetic, transcriptomic, and proteomic analysis will be each COPD subtype vs. smoking controls; the quantitative emphysema and airway disease axes; and the change in FEV_1 and adjusted lung density between the 5-year, 10-year and 15-year visits. In addition, a number of secondary analyses will be performed, including: a) presence/absence of COPD (comparing subjects with GOLD 2-4 COPD to smoking controls); b) cross-sectional measures of CT emphysema, CT gas trapping, CT airway wall area, FEV_1, 6MWD, oxygen saturation,

emphysema distribution, and chronic bronchitis; and c) longitudinal measures of exacerbation frequency and change in CT gas trapping, CT airway wall area, and 6MWD. Cross-sectional and longitudinal COPD-related phenotypes will be assessed within each COPD subtype and within the entire study population. Longitudinal imaging and lung function phenotypes will be assessed as both change over time (between the 5-, 10- and 15-year visits and across all four visits) and average outcomes over all four visits.

Phase 2 and 3 blood RNA-Seq is being used to generate gene, exon, and isoform-level estimates of relative expression abundance. RNA-Seq data is processed by the CDNM bioinformatics pipeline that addresses sequencing errors and other technical artifacts. Alignment to the human reference genome is being performed with STAR (13), followed by multiple quantification and analysis options, including DESeq2 (14), DEXSeq (15), and voom/limma (16). While isoform-level resolution is most reflective of underlying biology, it is technically challenging to accurately infer isoform levels from short-read fragments (17), and for some applications gene or exon-level quantifications may be preferable to isoform quantification (14, 15). Gene-level quantification of differential expression analysis is being performed using voom/limma, which accounts for the distributional aspects of RNA-Seq data through a weighted regression approach. Identification of differential exon usage can be performed using a novel implementation of DEXSeq adapted for large studies. These tools are all implemented in R/Bioconductor. Isoform quantification and differential expression analysis will be performed using pseudoalignment-based methods such as kallisto (18) and sleuth for isoform quantification and differential expression, respectively. Associations of gene expression to COPD subtypes can be adjusted for covariates, including complete blood count differentials. Surrogate variable analysis can be used to correct for potential batch effects and to remove unwanted sources of variability (19).

For proteomics analysis, the relationships between the categorical and quantitative outcomes and the predictor variables (~5000 individual plasma protein biomarkers in the SOMAScan panel) are assessed using standard regression models. Covariates in these analyses include age, gender, pack-years of smoking, and current smoking. To control the false discovery rate, the Benjamini-Hochberg method is used to adjust the p-values. These biomarker measurements are being integrated with pQTL genetic variants to improve accuracy of risk prediction.

We estimated power to detect Omics associations to FEV1 decline, based on expected 15-year FEV1 decline of 8% predicted (SD=15%). In a subtype of 1500 subjects, we will have 86% power to detect a biomarker accounting for 2% of FEV1 decline variation at alpha=0.00001 (adjusting for 5000 proteins); 500 subjects will provide 87% power to detect a biomarker accounting for 6% of FEV1 decline. Thus, we will have adequate power to detect significant Omics effects within COPD subtypes.

To determine power to detect genetic associations, we used simulation studies (in C) to estimate the statistical power of SKAT for grouped rare variants with variable directions of effect on COPD

Table 3: Power to Detect Rare Variant Associations to COPD Subtypes within the COPDGene NHW Population

	% of Variants in Region That Are Disease Susceptibility Loci			
Genetic Effect Size	**20**	**40**	**60**	**80**
Odds ratio of 2.25	24.3	43.1	68.2	82.7
Odds ratio of 2.5	32.1	58.8	84.9	95.4

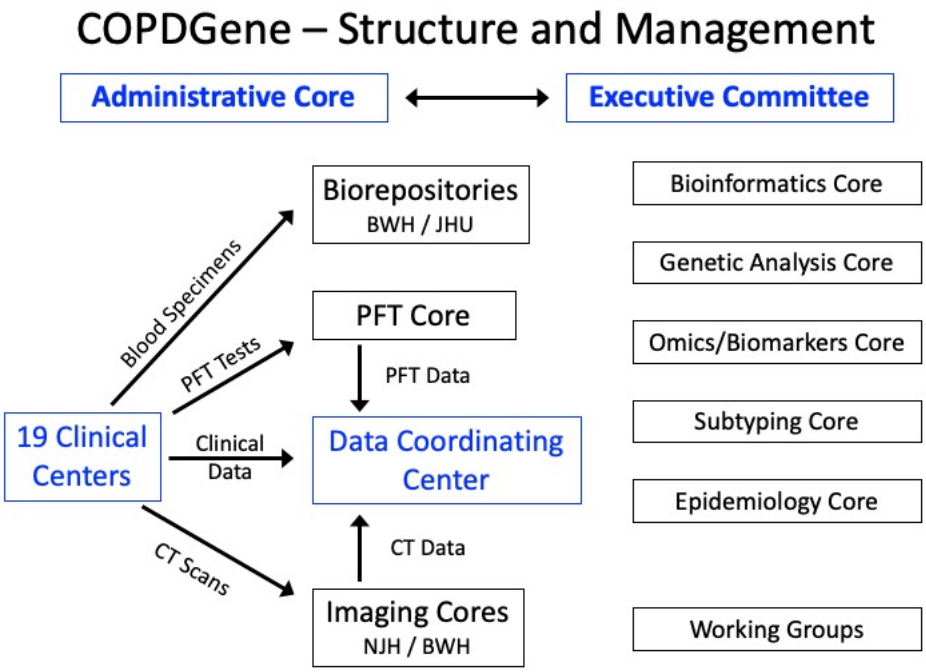
subtypes. We assumed COPD subtypes would include 25% of the affected population, with 20% of all causal variants having opposite/protective effects. We assumed overall odds ratios for the region of 2.25 and 2.5 for all variants combined. We used a 10kb region size, based on a simulated haplotype spanning 200kb, and a genome-wide significance level (assuming a Bonferroni genome-wide correction for 10kb non-overlapping windows) of $2x10^{-8}$. The number of tested loci will be reduced by phasing the entire genome and removing redundant variants. We will further increase statistical power by performing permutation testing at the genome-level. We examined only alleles with minor allele frequency <5%. Power estimates shown in **Table 3** for non-Hispanic White (NHW) subjects are based on 1000 replicates.

VIII. STRUCTURE OF CENTRAL ADMINISTRATION AND DATA MANAGEMENT

The primary purpose of the COPDGene Central Data Management is to oversee and manage the study, provide a mechanism to maintain the longitudinal contact with study subjects, assure protection of participant privacy and ensure responsible handling of data. All data is de-identified

Figure 1. The structure of data handling and transfer for the COPDGene Study, Phase 3 between clinical centers and the study "Cores". Types of data (identifiable, with indirect identifiers and genetic data) are specified with the planned transfers.

and a unique code is provided to each participant's data. The key to the code is kept securely at the DCC and will not be released to investigators. In Phase 4 of the study, we will be collecting



subject identifiers (name, address, phone numbers and social security numbers) and storing these in our Data Coordinating Center (DCC) located at National Jewish Health. In Figure 1 we illustrate the basic structure of data flow within the study and note data transfers of protected health information from our subjects that include direct and indirect identifiers.

The COPDGene study has a central Administrative Core that coordinates the study activities and works closely with the Data Coordinating Center to write and communicate the study protocol, consent template, manual of operations and to oversee training of the clinical center staff involved in the study. The Administrative Core also manages investigators meetings, provides scheduling and coordination for committee meetings and writing groups, oversees the longitudinal follow-up program, manages acquisition of Medicare cost information, performs Social Security Death Master File or National Death Index searching and maintains the study files for death adjudication. The Administrative Core has also employed National Research Coordinators (NRCs) who conduct COPDGene work at the 19 clinical centers in accordance with each center's specific institutional and IRB approvals and who also will assess site facilities and performance.

The Imaging Core receives subject CT scan image files and clinical center phantom data files along with dates of CT scan. The Imaging Core assesses the quality of the CT images and performs analysis of quantitative characteristics on the image files as well as overseeing the more complex scoring of images for visual analysis. Data and image files are shared between National Jewish Health and Brigham and Women's Hospital under the Imaging Core. Final output from the image analysis work is transferred and stored in the DCC.

The Pulmonary Function Testing (PFT) Core receives subject data with a date of study visit from the 19 clinical centers via a file depot located at National Jewish provided by the Data Coordinating Center. Dr. Robert Jensen is the head of the PFT Core. He reviews the Spirometry and DLCO data for quality assessment and returns this data with a quality grade to the Data Coordinating Center.

The Sample Storage Cores (Biorepositories) for COPDGene are located at Johns Hopkins University (JHU) and Brigham and Women's (BWH) Hospital. In phase 1 of the study biospecimens and DNA were stored at JHU, and for phases 2, 3, and 4 of the study, biospecimens are or will be stored at BWH. Backup storage of DNA from phase 1 is already in place at BWH. The date of sample collection is part of the information stored at each site. All samples are otherwise stored with a code but no other identifiers. The coded identifier for both sites is the COPDGene subject ID number. The BWH Biorepository prepares DNA, plasma, and serum samples for analysis at TOPMed-selected or other agreed-upon laboratories and also provides biospecimens to approved ancillary studies (which are responsible for the costs of sample preparation and shipping).

The Bioinformatics Center and the Genetic Analysis Center are located at Brigham and Women's Hospital. These Centers receive genetic and other Omics data for quality control and analysis. Data are coded and have no other identifiers associated with them.

The Epidemiology Core is located at the University of Colorado Denver on the Anschutz campus in Aurora. The Epidemiology Core participates in data analysis of phenotypes, genetic data, and imaging data. These data will include subject identifiers - dates and geocodes.

The Subtyping Core is made up of investigators from the whole study who meet by teleconference and collaborate on data analysis. Data shared to this Core will include geocodes and location to the level of zip codes.

The Biomarkers Core is located at National Jewish Health under the oversight of Dr. Russell Bowler but works collaboratively with the Sample Storage (Biorepository) Core sharing biospecimens to be used for analysis of biomarkers. Data will include dates of sample collection.

The Mortality Adjudication Center is part of the Administrative Core. They will receive redacted death certificates, redacted subject medical records and informant interviews. When the subject has consented to have personal identifiers transmitted to the Data Coordinating Center, staff for the Mortality Adjudication Center may contact next of kin for informant interviews and request death certificates. In Phase 3 after the identification of SARS-COV-2 virus and COVID-19 illness, the study is modified to include identification of COVID-19 illnesses and death. Identification of COVID-19 illness will include collection of medical records.

IX. PROCEDURES RELATED TO COPDGENE CLINICAL CENTERS

During the Phase 4 research study visit, the following procedures will be performed at the local Clinical Centers but are related to Central Study functions. Local Clinical Center functions are described to provide a context for the Central Study functions.

1. At the Clinical Centers: Informed consent will be obtained prior to any other study procedure. Clinical centers will use the COPDGene provided study template consent. An updated HIPAA consent to obtain medical records to adjudicate the cause of death will be obtained. Subjects will be asked to provide a signed release of medical records to be used in the event of their death, to obtain a next of kin interview/ knowledgeable friend interview or hospital and physician records related to the events and illnesses associated with their death.
2. At the Clinical Centers: Subjects will be asked to provide updated contact information (address and two phone numbers) including two secondary contacts with addresses and phone numbers. The subjects will be asked to give consent to have their participation in the COPDGene study disclosed to these secondary contacts in the event that we are unable to contact them, and also disclosed to their next of kin, personal representative or family members in the event of their death. Subjects will also be asked to permit transmittal of their personal information including social security number to the COPDGene Data Coordinating Center to use for central subject tracking, longitudinal follow-up, death searching through the Social Security Death Master File or National Death Index and obtaining cost and utilization information from the Centers for Medicare and Medicaid Services (CMS) databases.
3. At the Clinical Centers: Contact information (but not social security number) will be collected from two other individuals likely to know the subject's whereabouts, at least one of whom is a relative not living with the subject (these are called secondary contacts). The purpose of this information is to locate subjects who have moved to a different home and have a different address and/or phone number and ascertain vital status of the subject.

X. PROCEDURES RELATED TO COPDGENE CENTRAL STUDY COMPONENTS

1. Central Study Function: Phase 4 blood samples will be stored at the COPDGene Biological Repository at Brigham and Women's Hospital.
2. Central Study Function: CT Images will be transmitted to the NJH COPDGene Imaging Core and the BWH Imaging Center for analysis.
3. Central Study Function: Protected Health Information and Subject Identifiers from COPDGene subjects will be transmitted to the Data Coordinating Center via REDCap from the 19 Clinical Centers.
4. National Research Coordinators (NRCs) will be utilized to assess site facilities and train staff in COPDGene at 19 Clinical Centers and also potentially conduct the COPDGene study visits. National Research Coordinators have the capacity to function as a traveling clinical research coordinator to the Clinical Centers and work with local research teams to implement and coordinate research for the COPDGene study. Responsibilities include but are not limited to the following.
    - Conduct the COPDGene study according to the protocol. This includes: conducting COPDGene visits, conducting death adjudication tasks, calling subjects to schedule visits and ask follow-up surveys, locate lost subjects and ascertain vital status. Work may also be done remotely for the clinical centers.
    - Review and monitor Clinical Center functions for compliance with protocol and completion of necessary records. This task often includes working with the site to develop more efficient processes related to conduct of the COPDGene study. Monitoring may require on-site review of CC research coordinator performance during actual study visits.
    - Train Clinical Center site staff. This will involve group in-person and remote training sessions and may require on-site training at the CC.
    - Become familiar with CC policies and procedures for each CC that he/she travels to. Become familiar with each site-specific informed consent form.

Prior to visiting a Clinical Center, the NRC will work with the Principal Investigator, site manager, Human Resources and employee health departments, and others as applicable to ensure that they complete all of the paperwork and vetting needed to allow them to visit the site and perform COPDGene study work. NRCs will have access to personal information collected for participants in the study.

XI. PROCEDURES PERFORMED IN PART BY BOTH CLINICAL CENTERS AND THE COPDGENE ADMINISTRATIVE CORE

Longitudinal Follow-Up

Subjects will continue to be contacted up to four times per year by telephone, mail, or email for up to 10 years after their Phase 4 visit. Questions will be asked about current health status, exacerbations, cancer, new illnesses or medical conditions and current smoking status. The longitudinal follow-up contact mechanism is primarily based on automated contacts to subjects via a computer server controlled by the local clinical center in which the clinical center securely uploads subject contact data. Social security number and other subject personal information

requested in the Phase 2, Phase 3 and/or Phase 4 consents are uploaded to a secure server using secure sockets (SSL) technology and 256 bit or greater encryption with an HTTPS protocol.

For longitudinal follow up, subjects establish a preference for contact by email/ web data entry, automated phone calls with data entry by telephone keypad, or coordinator assisted survey. Subject contact information and identifying information is deleted automatically from the server after the contact is made or at the end of three weeks. Data collected from the subjects is de-identified and made available to the Data Coordinating Center at National Jewish to be stored. Subjects who fail to respond to automated contacts are contacted by the local clinical coordinator who then completes the clinical questions for the subject into the web-based data collection form. Subjects who become lost to follow-up from the longitudinal follow-up process are traced by their secondary contacts by the clinical center coordinators and searched for on the internet, in the Social Security Death Index or National Death Index using the same automated server system querying the social security master death file. Deaths identified from the social security death search are communicated to the Data Coordinating Center and the local clinical center. The local clinical center initiates collecting further information about the death as described below.

Central Study Function: In the event that the local clinical center is unable to complete the longitudinal follow process for their subjects the COPDGene central Administrative Core will take over that function using the stored subject personal identifiers to upload subject information to the computer server and will complete follow-up coordinator phone calls to subjects.

Communications to Study Subjects

In addition to the longitudinal follow-up program of phone and email contacts, subjects may be contacted on up to three additional occasions per year to inform them of other research studies and to update them about results of the COPDGene study. In general, these contacts will be made through staff at the local clinical center.

Central Study Function: Selected communications may be made by mail through the COPDGene Administrative Core for efficiency or cost reasons after there has been IRB approval of the content of the communication.

Central Study Function: Because the COPDGene Cohort is a national resource for understanding the impact of smoking on human disease and particularly on the lung, it is essential to retain the subjects effectively for the duration of possible research funding. Local clinical centers may not retain staff during periods when no active study visits are occurring, but maintaining contact and longitudinal data collection is necessary. The central Administrative Core of the study can assume subject follow-up contacts when local clinical centers are unable to complete this work. For Clinical Centers that are unable to complete study mailings or phone contacts, those functions can be assumed by the COPDGene Administrative Core.

Mail Consent Procedure:

Administrative Center Function: The Administrative Center will provide a consent addendum/HIPAA authorization template and a cover letter to the local Clinical Centers to be mailed to Phase 1 subjects who have not returned for a Phase 2, Phase 3 or Phase 4 visit. In the mailing, two copies of the informed consent addendum will be included. The cover letter instructs the subject to review the informed consent addendum and sign both copies if they are willing to

participate. They should return one signed copy in the enclosed pre-addressed envelope. In addition, a memo will be included that the subjects can give to their next of kin or personal representative explaining their participation in the study and planned follow-up with them in the event of death. The mailing will also include a form to provide updated contact information for the subject, their secondary contacts and their current treating physician.

The purpose of the informed consent addendum is to invite the subject to consent to an improved process to confirm vital status and obtain information about cause of death in the subjects who die before they are able to participate in the Phase 2, Phase 3 or Phase 4 visit. The consent addendum will allow the subject to designate next of kin, friends or a personal representative who can be contacted in the event of their death and allow that individual to disclose information about events surrounding their death. It will also allow the subject to authorize release of records from treating physicians and hospitals that they are admitted to. It will explicitly allow research data collection and will also include a provision to transfer subject name, address, social security number and phone numbers to the Data Coordinating Center and Administrative Core to be used for central data collection, including death searching the social security death master file, collecting records related to death and obtaining Medicare cost and usage information. The subject identifiers will be handled and stored in the same fashion as described for the Phase 2 subject visit data transfer.

If subjects cannot be reached or do not agree to sign the addendum, there will be no change in their status within the study. They will be contacted as usual for their Phase 4 visit and will continue having periodic contacts by email or telephone for updates on their health through the LFU system. Local clinical centers will continue to track subjects as described in the Phase 1 consent without a specific HIPAA authorization for medical records after death. Information regarding whether the subjects have signed the consent addendum will be maintained in the subject tracking system that is available to the local Clinical Center coordinators and the central Administrative Core.

Mortality Assessment

Local Clinical Center and Central Study Function: Death is an endpoint of interest that will be analyzed as part of this protocol. The opportunity to assess vital status will occur in one of several ways. It is possible that during the process of longitudinal follow up, the clinical center will be made aware of a subject's death. However, some subjects will be lost to follow up and vital status unknown. For these subjects all clinical centers will conduct a search to determine vital status including use of the local CC electronic medical record, obituaries and the internet. Periodic searches will also be done by the Administrative Center of the Social Security Death Master File, or National Death Index. Once the individual's vital status has been confirmed as deceased, the Clinical Center will obtain additional information regarding the subject's death so that cause of death may be determined. This additional information will include: 1) Informant Interview conducted with next of kin or other close contact of the study participant or the subject's physician, 2) Death Certificate, 3) hospital Discharge Summary if it is determined that the death occurred in hospital or within 28 days of a hospitalization, and 4) request for recent treating physician records if death did not occur in the hospital.

Central Study Function: The above records will be de-identified and transferred to the Data Coordinating Center. These records will be centrally reviewed by a Mortality Adjudication Committee. Cause of death will be assigned to one of the following causes: respiratory, cardiovascular, cancer, or other. The data will be reviewed by at least two members of the

Adjudication Committee and a final cause of death assigned according to the Principles of Adjudication. If the two Committee members disagree on cause of death, a third member of the Committee will review death records and cause of death will be assigned based on majority consensus.

Death adjudication will be coordinated by the central Administrative Core and all materials (death certificates, hospital records and physician records) for death adjudication will be maintained centrally. Death searching through the social security death master file or National Death Index will be automated through the longitudinal data collection server when permitted by local IRBs and manual searching will be performed by the central Administrative Core for subjects who have centralized personal information. Information on cost of care will be obtained from Medicare Databases by the central Administrative Core using the social security numbers stored centrally.

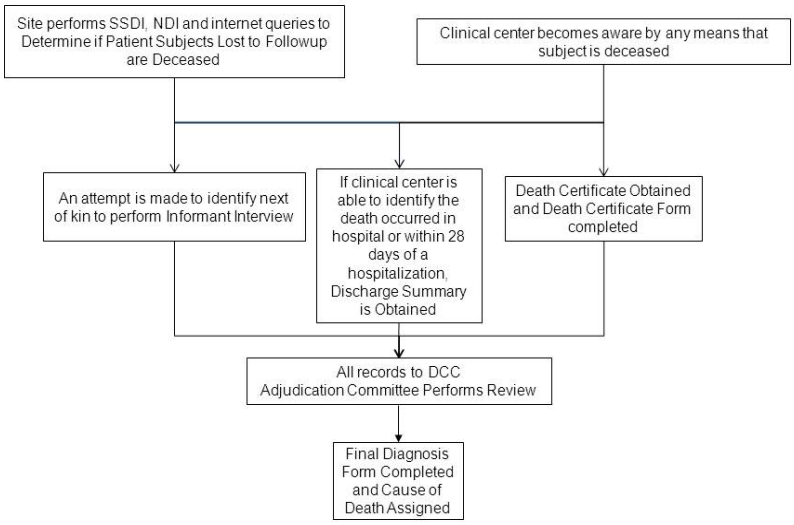


Figure 2 Procedures and Flow Diagram for Death Tracking and Death Adjudication

Study Procedures Performed by Imaging Core

The Imaging Core in its two locations (National Jewish Health and Brigham and Women's Hospital) receives CT images from the clinical centers. The images are accompanied by the date of study which is essential to the analysis to compare scans from Visit 1, Visit 2, Visit 3 and Visit 4. These images are analyzed for quantitative variables of emphysema, gas trapping and others. The images are also visually scored by investigators. Images can be fully de-identified when research procedures do not require knowing the date of study to be shared as a de-identified dataset. When images are shared outside of the Imaging Core with the date of study intact, the files will be treated as a limited dataset.

Study Procedures Performed by PFT Core

The PFT Core in two locations (National Jewish Health and Dr. Robert Jensen as a contract consultant) will receive protected health data from the spirometry and DLCO testing and the date of the study. After completing the quality review of the Spirometry and DLCO by Dr. Jensen, the data will be returned to a file depot in the DCC. It will be added to the main database. For Spirometry and DLCO results that fail quality assurance, a secondary review will be made by the PFT Committee (currently Drs. Make and Silverman).

XII. DATA COLLECTION AND DATA HANDLING

Each Clinical Center involved in the study will follow the COPDGene Study protocol which has been approved by the sIRB to obtain data for the study. Any information collected from any subject involved with COPDGene will be identified by a subject ID that consists of letters, numbers, and a three letter center code.

Phases 1, 2 and 3 Data
The data collection for Phases 1-3 has been done through the COPDGene Study under local IRB approval for the project. Subjects provided new informed consents to participate in Phases 2 and 3. This data is currently stored in the Data Coordinating Center as *existing* de-identified data that is described below.

Questionnaires. Respiratory Symptoms, Medical History, Medications, SF-36, St George's Respiratory Questionnaire, HADS, Longitudinal follow-up questionnaire, Residential and Environmental, CAT, For Women Only, Cognitive screening, CAPTURE (COPD Population Screener), PROACTIVE (physical activity), Duke Activity Status Index (DASI), Residential and Occupational history, COPD Assessment Test (CAT), e-cigarette use, Socioeconomic, and Longitudinal follow-up questionnaires. All questionnaire data to date has been stored at the DCC. The majority of the above questionnaires were collected at one or more of the Phase 1, 2 and 3 visits.

Breathing Test (Spirometry): Spirometry was collected on all subjects. Both a pre-bronchodilator and a post-bronchodilator measurement are collected from each subject.

DLCO: Electronic files transferred directly to DCC without personal identifiers but with date of study.

Physical Assessment: Subject blood pressure, height and weight were obtained.

CBC Results: Recorded on data form uploaded to DCC.

Six-Minute Walk Test: Subjects were asked to walk for 6 minutes on a level surface following a standard protocol, and a measurement of total distance walked was obtained.

High Resolution Chest CT Scan: Chest CT scans were obtained for all subjects following the COPDGene CT scan protocol. These scans were read by the clinical center radiologist and then de-identified and transferred to the COPDGene Imaging Core. Output from the analysis of the CT scans is stored in the DCC as de-identified data.

Personal identifiers: Name, address, date of birth, social security number, names and addresses of secondary contacts, date of death when applicable.

Exacerbations and Incident Medical Conditions: Subjects are contacted twice a year using a combination of automated phone calls, email and coordinator phone calls. Data output from these contacts are stored in the DCC as de-identified data.

Death Adjudication: Data has been collected regarding cause of death to include death certificate, next of kin interview, physician interview, office records and hospital records and searching the social security death master file or National Death Index. Records may include identifying information such as name, address and social security number when subjects have provided consent for identifying information to be transmitted to the central study. Records are redacted of

identifying information by the local clinical center prior to transmission if permission has not been granted.

Phase 4 Data

The data collection for Phase 4 will be done with sIRB approval (BRANY; https://www.brany.com) for the project and local IRB approval as required by each clinical center. Subjects will provide new informed consent to participate in Phase 4. Data will be transferred within the study cores as shown in Figure 1. REDCap will be utilized as the electronic data capture system (EDC) for Phase 4 of COPDGene. The DCC will receive and store identifiable data for Phase 4 only after receiving IRB approval for this activity.

Data Collection will include:

Study Questionnaires: Respiratory Symptoms, Medical History, Medications, SF-36, St George's Respiratory Questionnaire, HADS, Cognitive screening, PROACTIVE (physical activity), Residential and Occupational, CAT, e-cigarette use, Socioeconomic, Fall history, Social Determinants of Health, Periodontitis, and the Longitudinal follow-up questionnaire.

Physical Assessment: arm span, height, weight, pulse oximetry, heart rate and standardized blood pressure.

CBC Results: Recorded on data form uploaded to DCC.

Spirometry: Results with date of test.

DLCO: Electronic files transferred directly to DCC without personal identifiers but with date of study.

Six Minute Walk Test and Sit to Stand Test: results recorded on data form and uploaded to DCC.

CT Scan of the Chest: image files that are de-identified of subject identifiers except for date of study.

Personal identifiers: Name, address, date of birth, social security number, names and addresses of secondary contacts, date of death when applicable.

Death Adjudication: Data will be collected regarding date of death and cause of death to include death certificate, next of kin interview, physician interview and office records and hospital records. Records may include identifying information such as name, address and social security number when subjects have provided consent for identifying information to be transmitted to the central study. Records will be redacted of identifying information by the local clinical center prior to transmission if permission has not been granted.

C4R Project

Cohort data collection: Each cohort ,including COPDGene, will be directly responsible for accomplishing its own data collection in accordance with the master protocol. The protocol is as follows:

Consent: All participants have previously been consented for ongoing telephone contact, abstraction of medical records of hospitalizations, and data sharing (with some restrictions) by

the individual cohorts. All participants have previously consented to obtaining medical records for review; if exceptions are noted, records will not be obtained for those individuals.

COVID-19 Questionnaires: Each cohort will deploy standardized COVID-19 questionnaires. The Wave 3 questionnaire was developed to be standardized with both the original C4R questionnaire and also with the questionnaire instruments being used by the RECOVER initiative. The domains assessed by the Wave 3 questionnaire include infection and vaccination status, recovery from COVID-19, and symptoms associated with long COVID or post-acute sequelae of COVID-19 (PASC).

COVID-19 medical record ascertainment: Each cohort will use its own established infrastructure for ascertainment of medical records and death certificates, including use of the National Death Index (NDI), CMS, ICD codes,[120] and linkage to records from local departments of health.

Training and Certification:
Questionnaires: Cohort personnel will be trained to administer C4R questionnaires per protocol using existing cohort infrastructure, the usual COPDGene training.
Events*:* Collection of hospitalization and death records will follow the C4R Events Protocol. As much as possible, it will piggyback upon ongoing cohort events protocol, with minor modifications to include COVID-19 hospitalizations and deaths.

XII. DATA STORAGE AND DISTRIBUTION

Research Informatics Services (RIS) at National Jewish Health (NJH) is the Data Coordinating Center (DCC) for COPDGene. The COPDGene DCC will maintain the confidentiality of all protected health information collected under this protocol using physical security, database security, and web applications security.

Physical Security
Computer Room. National Jewish Health Information Services maintains servers for Research Informatics Services in their computer room, a secure environment whose access is restricted to essential Information Services personnel. This server room has dedicated power, cooling, lighting, and an environmental monitoring system.

Backup. Information Services runs backups of the DCC servers daily. The transaction logs for SQL databases are run multiple times daily. Data related to patients is encrypted by the backup process to ensure security.

Database Security
The main COPDGene database is stored in a Microsoft SQL Server 2014 database on a server dedicated to Microsoft SQL Server 2014 (SQL server). Each SQL database defines the users who may access the database. The COPDGene DCC incorporates Windows Authentication to define the members of *user groups* who can access a specific SQL database. In addition, we define permissible actions for each *user group* using Microsoft SQL's *grant* and *deny* commands.

The main COPDGene database uses a deidentified subject identification code to associate a particular subject with data submitted on a data collection form or data collected using a medical device such as a spirometer or a CT scanner. A subject's date of birth is collected on a demographics form and is used solely to calculate age; date of birth is not included in the main COPDGene® dataset that is provided to investigators. When data from a spirometer or a CT scanner are first submitted to the DCC, however, the spirometer data and the CT data do include the date of service.

In the second and third phases of COPDGene and again in the fourth phase, a new Subject Personal Information form will collect elements of Subject Identifiers, including name, address, and social security number. The data from this form will be used by the Administrative Core of COPDGene® for death adjudication, but all the subject identifiers are encrypted and stored in a SQL database that is separate from the main COPDGene database; moreover, the SQL database with subject identifiers will reside on its own SQL server. Microsoft SQL has its own encrypted backup system, and all subject identifier data will be backed up in an encrypted form. Direct access to this COPDGene subject identifier database will be restricted to essential DCC personnel; it will have a distinct owner, and it will not use the default SA owner; last, the *guest* user has been removed from this server. These measures reduce even further the chances that someone could hack their way into this subject identifier database.

REDCap will be the research electronic data capture system for Phase 4 data entry. It is a secure, web-based application designed to support data capture for research studies. REDCap is maintained by the REDCap Consortium (REDCap Consortium) which is comprised of over 2000 institutional partners, including NJH, and is administrated locally by NJH Research Informatics Services.

Data Security and Privacy

- Complete suite of features that support HIPAA compliance
- Access to database requires user authentication with password
- Data access based on individual's role on project
- Logging and audit trails on all data interactions
- All data stored on a secure server, and backups are encrypted
- Twice daily automated backups of REDCap servers

REDCap can remove identifiers from a dataset prior to exporting for analysis to create either a limited data set or a safe harbor data set. A safe harbor data set is created by the removal of the 18 pieces of information considered identifiers for the purposes of HIPAA compliance.

Two members of the Administrative Core (Associate Director, Dr. Elizabeth Regan and the COPDGene Central Study Project Manager), James Crooks, PhD, and 2 DCC members (Carla Wilson and Tricia Uchida) will have authority to access the subject identifier databaseand supervise internal analysts using the information to support external investigator project. Carla Wilson is the database developer for the SQL database, and on request, could provide Dr. Regan with an encrypted copy. This copy would reside on a single desktop computer in Goodman K706, encrypted using TrueCrypt (http://www.truecrypt.org/), free open-source disk encryption software

that is currently being used within NJH for Honest Broker protocols associated with the NJH Research Database. At this time, it is not anticipated that it will be necessary to produce an encrypted copy of this subject identifier database.

The only elements of the subject identifier database that we expect to associate with protected health information (PHI) are addresses and occupations in order to examine the relationships of environmental and occupational history to the development and progression of COPD. The NJH DCC and Biostatistics Core will produce de-identified datasets for these purposes.

Web Applications Security
The DCC maintains a password-protected, limited-access COPDGene website that is located at a secure https URL. The passwords are stored in a SQL table and protected by a sophisticated salt-and-hash technique. By definition, https websites are secured using Secure Sockets Layer (SSL) or Transport Layer Security (TLS) technology standards. SSL is used to encrypt data transmissions between the web server and a user's computer. In addition, the DCC maintains a security certificate that authenticates the website resides within NJH.

Among its other functions, the COPDGene website provides the ability for Clinical Center investigators to have access to existing datasets. Access to these features of the COPDGene website is controlled by defining distinct user groups that have different capabilities. The DCC will review on a monthly basis login activity from the Clinical Centers that participate in COPDGene. Users must change their passwords every 90 days. Accounts require legitimate institutional email addresses, which are used to verify the user's relationship at the institution. If the relationship has ended, the user will not be able to access the COPDGene study website. If a user has not logged in to the COPDGene within the previous 180 days, that user's access to the web site will be suspended until their status with COPDGene can be confirmed.

The DCC has enabled REDCap's Data Resolution Workflow to query fields that are out of range or trigger other data quality flags. DCC personnel initiate, review and close all queries. All data edits are logged in the REDCap logging system.

XIV. DISTRIBUTION OF DATA AND BIOSPECIMENS

The COPDGene Central Study acting through established policies in the Administrative Core and the Ancillary Studies and Executive committees will oversee and coordinate the data sharing approvals for collaborating investigators through the Data Coordinating Center (DCC). The final data sharing agreements between National Jewish Health and other institutions will be established managed by the National Jewish Compliance office. Agreements will include both the investigator and their institution.

Tracking of those distribution agreements will be managed by the Administrative Core and DCC staff. This includes verifying that Collaborating Investigators and the Recipients sign the appropriate agreements, as well as keeping those documents on file and reviewing data requests for appropriateness and merit. They are jointly responsible for ensuring data and images from COPDGene are stored securely and managed in accordance with existing regulations for protection of privacy and human subject research.

There are three major types of data distribution that are anticipated from COPDGene Study: 1) data sharing for "public" access to COPDGene data via dbGaP or other NIH controlled databases, 2) data distribution to investigators within the COPDGene study for analysis related to the study specific aims, and 3) data distribution to investigators within and outside of the COPDGene study who have ancillary studies approved by an IRB at the investigator's site.

1. Data Sharing to dbGaP or other controlled access database

The COPDGene Data Coordinating Center (DCC) has acted to oversee transfer of existing de-identified subject data to the NIH controlled dbGaP from Phase 1 of the study. dbGaP provides storage, oversight, and a process for distribution of phenotypic and genetic data as required by the NIH Data Sharing agreement. The DCC will oversee future data transfers of de-identified data to dbGaP or another controlled database as requested by the NIH. dbGaP is responsible for review and oversight of all data released from their database. Any request for data from researchers who are not associated with the COPDGene study and do not have approved ancillary studies will be directed to dbGaP.

2. Data Sharing with Internal COPDGene Investigators

Internal COPDGene Investigators are defined by the Executive Committee to include Clinical Center Directors and co-investigators, members of all study cores and data analysis working groups and other interested investigators who have been invited to join in the study.

<u>De-identified Datasets.</u>

The DCC will periodically post de-identified datasets at the instruction of the Executive Committee on the COPDGene password protected website, for the use of internal investigators. Internal investigators consist of co-investigators at the clinical centers and at the study cores. These members of the study have been provided logins/passwords to access the internal COPDGene website. Internal investigators are advised that the datasets downloaded from the COPDGene website are provided for their individual use and may not be shared within their local working groups that may include other research collaborators working with them and statisticians who are performing analytic work, without including those individuals on their protocols and having them sign data use agreements. Release of the de-identified dataset outside authorized recipients is not permitted without explicit permission of the COPDGene Executive Committee. COPDGene internal investigators who analyze the data are expected to submit a manuscript proposal to the Ancillary Studies and Publications committee for review and approval and when the manuscript is ready for submission to a journal they will submit the final manuscript draft for review and approval.

<u>Biospecimens.</u>

Biospecimen requests may involve additional procedures before being released to internal or external investigators. The COPDGene Study has an obligation to ensure that these are released in compliance with the subject consents and privacy regulations.

<u>CT Images.</u>

CT images are stored with the date of image acquisition in order to permit proper sequential analysis. Thus, these image files represent "limited datasets" under HIPAA regulations. Investigators who request access to the CT images will be provided with de-identified images.

Process for Access to Data, Images and Biospecimens.
For requests from investigators associated with the COPDGene study, written proposals for acquisition of specimens, images, and clinical data are submitted as a data request in REDCap to the DCC. These are reviewed, when necessary, by the Administrative Core and may be reviewed by the COPDGene Ancillary Studies and Publications Committee and the Executive Committee as needed. In the case of materials, the proposal will be discussed by the COPDGene administrative staff with the investigator to resolve any questions or missing information. Data requests by internal investigators of stored data can be released under a standard operating procedure by the DCC staff unless questions or problems are identified. Requests for biospecimens will be reviewed by the Ancillary Studies and Publication Committee. Once the manuscript proposal is approved by the Ancillary Studies and Publication Committee, it is sent to the COPDGene Executive Committee who will then review each proposal for merit and feasibility. Upon approval, the COPDGene Executive Committee will instruct the Biorepository, Imaging Core and DCC to release the approved data, specimens or images as needed.

Because storage of biological samples from research subjects are subject to the storage and sharing constraints imposed in the subject consent form, biospecimens that are provided to investigators must either be returned to the COPDGene Biorepository, destroyed, or used completely; and the receiving investigator must agree to the conditions and comply/attest to the final disposition. Biospecimens cannot be transferred to another investigator or used for other analyses without approval of the COPDGene Executive Committee.

3. Data Sharing for Investigators Within and Outside of COPDGene for their Ancillary Studies

In general, ancillary studies and associated requests for data, access to research subject or biospecimens will be made to the ancillary study committee using the Ancillary Study Proposal Form on the study-controlled website; this is a link to a REDCap form that must be completed by the investigator (also available from the Administrative Core, by request).

De-identified Datasets.
De-identified datasets can be provided to investigators outside of COPDGene after submitting a proposal through the process laid out above, completing a COPDGene review of the proposed work via the Ancillary Studies and Publication committee and confirmation that there is an existing IRB review at the investigators home institution. A data use agreement with National Jewish Health is required before data is released. Investigators can also be referred to dbGaP for access to the de-identified data.

Biospecimens
For requests from investigators outside the COPDGene study, written proposals for acquisition of specimens are initially reviewed by the COPDGene Ancillary Studies Committee. Biospecimen requests from investigators outside the COPDGene study must be accompanied by an approved IRB protocol. A data use agreement must be signed before any data or specimens can be released. The project must be reviewed by the Ancillary Study Committee and the Executive Committee.

As described for internal investigators, non-COPDGene investigators requesting biospecimens and images will have their proposal reviewed and discussed by the COPDGene administrative staff. The staff will contact the investigator to resolve any questions or missing information. During this time, the statistical plans and power analysis will be reviewed. Once the details of the proposal are

finalized by the administrative staff, it will be sent to the COPDGene Ancillary Studies Committee who will then review each proposal for merit and feasibility. Upon approval, the COPDGene Executive Committee will instruct the Biorepository, QIL or DCC to release the approved data, specimens or images.

Because storage of biological samples from research subjects are subject to the storage and sharing constraints imposed in the subject consent form, biospecimens that are provided to investigators must either be returned to the COPDGene Biorepository, destroyed, or used completely; and the receiving investigator must agree to the conditions and comply/ attest to the final disposition. Biospecimens cannot be transferred to another investigator or used for other analyses without approval of the COPDGene Executive Committee.

Distribution of Biospecimens

To protect the confidentiality and privacy of participants, Recipients of COPDGene Data and/or Biospecimens must adhere to the requirements of the appropriate Investigator Certification. Failure to comply with this Agreement could result in denial of further access to COPDGene Data and Biospecimens.

1. Application for Materials

An investigator wishing to use COPDGene specimens along with images or other phenotypic data should submit an application or Ancillary Study Proposal to the COPDGene Ancillary Study Committee. The application should include the title, investigators, hypotheses to be tested, identification of variables, specimens or images, analysis plan and proposed timetable.

2. Responsibility of Investigators

The investigator will pay any costs incurred by the DCC and Biorepository for data and biospecimen distribution. The investigator must agree to not distribute materials, avoid waste of precious materials, and return of materials when appropriate. Recipient will agree to retain control over Data, Genetic Analysis Data, and Biological Materials, their progeny, and unmodified or modified derivatives thereof, and further agrees not to transfer Data, Genetic Analysis Data or Biological Materials, their progeny, or unmodified or modified derivatives thereof, to any other entity or individual. The recipient will agree to make reasonable efforts to avoid contamination or waste of the samples during material handling.

Investigators must fulfill the following requirements:

a) Certify the investigators willingness to adhere to COPDGene guidelines for confidentiality, use of specimens, biosafety and indemnification.

b) Before receiving COPDGene® data, specimens or images the investigator must indicate that the data, specimens or images will be used only as agreed upon in the collaboration and will document this at completion of the study.

c) Provide an annual progress report.

3. Credit and Authorship

All presentations and manuscripts shall acknowledge that the data were collected through COPDGene. If there is collaboration with COPDGene researchers, authorship would be expected. The "right" to authorship is determined by substantive intellectual contribution of an individual to

at least 2 of 3 areas: study question concept; data collection and analyses; and manuscript preparation. Appropriateness of authorship for a given study may be reviewed by the COPDGene Publications and Ancillary Studies Committee.

We will be storing genetic data in addition to subject identifiers. An inadvertent breach of confidentiality of this data may affect the subject's employment or ability to obtain health care. A Certificate of Confidentiality has been obtained for the COPDGene study to provide additional protection for study participants in relationship to the genetic data that we will be holding.

This study is designed to be a national resource for scientific investigations. As such, medical information, genetic and Omics information and imaging data will be provided to dbGaP in order to make the study data available other investigators, with appropriate safeguards. Other researchers interested in using such information for scientific investigations will be required to apply to the Executive Committee for permission to access the data for studies that have received local IRB approval and with requirements to maintain subject confidentiality.

This study is designed to be a national resource for scientific investigations. As such, medical information, genetic and Omics information and imaging data will be provided to dbGaP in order to make the study data available other investigators, with appropriate safeguards. Other researchers interested in using such information for scientific investigations will be required to apply to the Executive Committee for permission to access the data for studies that have received local IRB approval and with requirements to maintain subject confidentiality. Subject identifying information will not be transmitted to other investigators. A Certificate of Confidentiality has been obtained for the COPDGene study to provide additional protection for study participants.

XV. RISKS AND DISCOMFORTS

The only drug administered in this study is albuterol, as part of the pulmonary function testing. Two puffs of albuterol (180 mcg) are given via metered dose inhaler, with a spacer. After waiting 20 minutes, spirometry is performed. This is a one-time dose only. Albuterol is a commercially available, FDA-approved drug. The common side effects associated with albuterol are transient tachycardia, tremulousness, and nervousness.

Subjects may become short of breath during the six-minute walk test as it is designed to assess subjects' maximal exercise capacity. There is a potential risk of syncope, light-headedness and fainting during these tests; these are uncommon. There is a potential for the subject to develop chest pain or cardiac arrhythmias; these are uncommon. There is a remote risk of heart attack or death during this test. All subjects in this study have previously had a walk test performed successfully without adverse events as part of the COPDGene enrollment visit, and walk tests were also performed during the 5-year and 10-year visits.

Subjects may become short of breath or have generalized muscle fatigue during or after the sit to stand test, as it is designed to assess subjects' muscle strength. There is a potential risk of syncope, light-headedness and fainting during these tests; these are uncommon. There is a potential for the subject to develop chest pain or cardiac arrhythmias; these are uncommon. There is a remote risk of heart attack or death during this test.

From spirometry and DLco, subjects may become short of breath. Syncope, light-headedness, and fainting may develop during this test; these are uncommon. Occasionally during or after pulmonary function testing, subjects may temporarily become short of breath. Subjects who are using oxygen may become short of breath when the oxygen is removed. However, this should rarely occur since oxygen will only be discontinued for 10 minutes and the oxygen saturation will be continuously monitored during this time.

There is a possible risk in questionnaire administration from inadvertent disclosure of medical history information. There is also a risk of loss of confidentiality. These potential risks are guarded against by maintaining completed questionnaires in a locked filing system in a locked room at the Clinical Centers, password-protected computers, and using secure transmission of information to the Data Coordinating Center (DCC). Pulmonary function and questionnaire data, identified by study ID number only, will be transferred by secure internet connection (with 132-bit encryption) to the Data Coordinating Center at National Jewish Health in Denver. Identifying information will be sent from the Clinical Centers to the DCC, but this identifying information will be stored separately from clinical and genetic information collected in COPDGene. Study personnel at Clinical Centers and the DCC will be required to meet local requirements for training in protection of confidentiality.

Potential risks of blood drawing are hematoma at the skin site and minimal pain of venous puncture.

Radiation exposure in the chest CT scan could theoretically increase the risk of cancer. Exposure of pregnant women to CT scan radiation could be harmful to the developing fetus. To minimize the likelihood of exposing pregnant women to CT scans, females of child-bearing potential will be asked if they are pregnant or have the possibility of being pregnant before the chest CT scan. CT scans will not be performed for women who state they are pregnant or that they may be pregnant; other pre-menopausal women will undergo urine pregnancy testing before a chest CT scan is performed. Chest CT scans could identify pulmonary nodules that may require follow-up outside of this study. Such pulmonary nodules could be curable lung cancers (a benefit) or scars/prior granulomas that could require additional radiation exposure or even surgery (a risk).

The average amount of radiation exposure, during the chest CT scan is approximately 3.5 mSv. The radiation dose differs with body size; thinner subjects will have less than this amount of radiation and heavier subjects will have more than this amount of radiation. The average amount of background doses of radiation that the general population is exposed to in the United States is 3 mSv per year. Thus, the average amount of radiation subjects will receive from each study visit is equivalent to slightly more than one year of normal background radiation.

There is a risk that the depression questionnaire (Hospital Anxiety and Depression Scale, HADS) will identify subjects with potential clinical depression. The HADS cannot diagnose clinical depression; the diagnosis of clinical depression can only be made by a health care professional. To assure these subjects have further evaluation and therapy if appropriate, subjects will be told if their score (16 or greater) (14) indicates possible depression and will be informed to seek further care from their personal physician. The HADS score will be immediately available to the Research Coordinator and if abnormal the Research Coordinator will inform the subject in person before the study visit is ended. In addition, this information will be mailed to the subject's primary care

provider who will be asked to perform a clinical evaluation to determine the possible need for further evaluation or treatment for depression. The DCC will collect information from the study coordinators to verify that subjects and their physicians have been informed about such scores.

There are no costs associated with participating in this study. Participants will be compensated for their time and expenses in this study as follows. Participants will be compensated $200, for completing questionnaires and procedures for the Phase 4 study visit. Participants who have a remote visit for questionnaire administration will receive $25 and receive an additional $175 when they have an in-person visit. Subjects who cannot attend an in-person visit will be offered compensation of $50, for time and effort required of completing all questionnaires over the phone.

This study is designed to be a national resource for scientific investigations. As such, medical information, genetic and Omic information will be provided to monitored data repositories such as dbGaP as one example and other Biorepositories in order to make the study data available to other investigators, with appropriate safeguards. Other researchers interested in using such information for scientific investigations will be required to apply to the COPDGene Executive Committee for permission to access the data for studies that have received local IRB approval and with requirements to maintain subject confidentiality. Subject-identifying information will not be transmitted to other investigators.

XVI. POTENTIAL BENEFITS

There are no expected benefits to the study participants. Improved understanding of COPD has occurred; COPDGene has generated over 500 scientific publications. As noted above, CT scans could identify pulmonary nodules, early lung cancer or other abnormalities that may require follow-up outside of this study.

XVII. DATA SAFETY AND MONITORING PLAN

Adverse Events and Protocol Deviations

COPDGene is an observational longitudinal investigation without a therapeutic intervention. It is expected that there will continue to be deaths among subjects enrolled in this study that are not related to study procedures. It is expected that there will be hospitalizations from a variety of causes not related to study procedures, including but not limited to newly discovered disorders, acute disorders requiring surgery, pre-existing conditions, and exacerbations of underlying COPD. Subjects may expire due to pre-existing or new diseases including cancer, cardiovascular conditions, and COPD. These are anticipated events that are not related to this investigation. These events will not be reported to IRBs as adverse events.

Adverse events for the purposes of this study will only be those that are directly related to study procedures done during the study visit(s). There are no expected Serious Adverse Events in this study related to study procedures. At each Clinical Center, subjects will be observed for the development of tremulousness and nervousness following bronchodilator medication. Unexpected Adverse Events related to study procedures will be reported to the IRB of the Clinical Center (and to the single IRB in Phase 4) and to the COPDGene Executive Committee. An Observational

Safety and Monitoring Board (OSMB) has been appointed by the National Heart, Lung, and Blood Institute and will continue to oversee this study. All protocol deviations and adverse events will be reviewed by the OSMB, and the OSMB will vote on modifying the study if needed.

As noted above, we anticipate that some subjects may expire during the next phase of this study due to a combination of pre-existing disease and the onset of new conditions. These events are not anticipated to be related to the study visit; however, they provide important information about the natural history of COPD and other medical conditions. We will monitor and collect information about deaths in the cohort but will not report them to IRBs as study-related events unless they occur during a study visit or within twenty-four hours of the study visit and are judged to be related to a study procedure.

Risks and Benefits of Study Procedures

The only drug administered in this study is albuterol, as part of the pulmonary function testing. Two puffs of albuterol (180 mcg) are given via metered dose inhaler, with a spacer. After waiting 20 minutes, spirometry is performed. This is a one-time dose only. Albuterol is a commercially available, FDA-approved drug. The common side effects associated with albuterol are transient tachycardia, tremulousness, and nervousness.

Subjects may become short of breath during the six-minute walk test as it is designed to assess subjects' maximal exercise capacity. There is a potential risk of syncope, light-headedness and fainting during these tests; these are uncommon. There is a potential for the subject to develop chest pain or cardiac arrhythmias; these are uncommon. There is a remote risk of heart attack or death during this test although these are very uncommon. All subjects in this study have previously had a walk test performed successfully without adverse events as part of the COPDGene enrollment visit, and walk tests were also performed during the 5-year and 10-year visits.

Subjects may become short of breath or have generalized muscle fatigue during or after the sit-to-stand-test, as it is designed to assess subjects' muscle strength. There is a potential risk of syncope, light-headedness and fainting during these tests; these are uncommon. There is a potential for the subject to develop chest pain or cardiac arrhythmias; these are uncommon. There is a remote risk of heart attack or death during this test.

From spirometry and DLco testing, subjects may become short of breath. Syncope, light-headedness, and fainting may develop during these tests; these are uncommon. Occasionally during or after pulmonary function testing, subjects may become temporarily short of breath. Subjects who are using oxygen may become short of breath when the oxygen is removed. However, this should rarely occur since oxygen will only be discontinued for 10 minutes and the oxygen saturation will be continuously monitored during this time.

There is a possible risk in questionnaire administration from inadvertent disclosure of medical history information. There is also a risk of loss of confidentiality. These potential risks are guarded against by maintaining completed questionnaires in a locked filing system in a locked room at the Clinical Centers, password-protected computers, and using secure transmission of information to the Data Coordinating Center (DCC). Pulmonary function and questionnaire data, identified by study ID number only, will be transferred by secure internet connection (with 132-bit encryption) to the Data Coordinating Center at National Jewish Health in Denver. Identifying information will be sent from the Clinical Centers to the DCC, but this identifying information will be stored separately from clinical and genetic information collected in COPDGene. Study personnel at

Clinical Centers and the DCC will be required to meet local requirements for training in protection of confidentiality.

Potential risks of blood drawing are hematoma at the skin site and minimal pain of venous puncture.

Radiation exposure in the chest CT scan could theoretically increase the risk of cancer. Exposure of pregnant women to CT scan radiation could be harmful to the developing fetus. To minimize the likelihood of exposing pregnant women to CT scans, females of child-bearing potential will be asked if they are pregnant or have the possibility of being pregnant before the chest CT scan. CT scans will not be performed for women who state they are pregnant or that they may be pregnant; other pre-menopausal women will undergo urine pregnancy testing before a chest CT scan is performed. Chest CT scans could identify pulmonary nodules that may require follow-up outside of this study. Such pulmonary nodules could be curable lung cancers (a benefit) or scars/prior granulomas that could require additional radiation exposure or even surgery (a risk).

The average amount of radiation exposure during the chest CT scan is approximately 3.5 mSv at each study visit. The radiation dose differs with body size; thinner subjects will have less than this amount of radiation, while heavier subjects will have greater than this amount. The average amount of background doses of radiation that the general population is exposed to in the United States is 3 mSv per year. Thus, the maximum amount of radiation subjects will receive from each study visit is equivalent to slightly more than one year of normal background radiation.

There is a risk that the depression questionnaire (Hospital Anxiety and Depression Scale, HADS) will identify subjects with potential clinical depression. The HADS cannot diagnose clinical depression; the diagnosis of clinical depression can only be made by a health care professional. Subjects and (where possible) their primary care physicians will be informed if a subject's HADS score (16 or greater) indicates possible depression.

This study is designed to be a national resource for scientific investigations. As such, medical information, genetics and Omics information will be provided to monitored data repositories such as dbGaP as one example and other Biorepositories in order to make the study data available to other investigators, with appropriate safeguards. Researchers interested in using such information for scientific investigations may also apply to the COPDGene Executive Committee for permission to access de-identified data for studies that have received local IRB approval and with requirements to maintain subject confidentiality. Subject-identifying information will not be transmitted to other investigators.

There are no expected benefits to the study participants. Improved understanding of COPD has occurred; COPDGene has generated over 500 scientific publications. As noted above, CT scans could identify pulmonary nodules, early lung cancer or other abnormalities that may require follow-up outside of this study.

Monitoring Procedures

Quality assurance of spirometry and diffusing capacity data will be insured by the Pulmonary Function Core in Utah, which will review de-identified spirometry data (the date of the test will be on each test) from each study participant. The PFT Committee will perform a secondary review of questionable results. Quality assurance of chest CT scans will be analyzed by the Imaging Core

at NJH. Questionnaires and other data will be quality controlled by the Data Coordinating Center at NJH. The DCC sends bi-weekly reports to each Clinical Center, the Administrative Core and NRCs describing missing items and study visit completion. If quality issues are identified in any of these areas, the Administrative Core will organize meetings with the relevant Clinical Centers for retraining as needed. Genetics and Omics data will undergo quality control review by the Genetic Analysis Center and Bioinformatics Center.

The COPDGene Administrative Core will continuously monitor study progress at the Clinical Centers. We strive to continue our excellent track record of high-quality study data and completeness. The Administrative Core's National Research Coordinators (NRCs) will visit Clinical Centers periodically to review site facilities, monitor study procedures and record-keeping, assist with continuing education of local coordinators, review local coordinator performance and observe study visits, and, as needed, assist in conducting study visits. Clinical Centers will be responsible for providing study records and logs for review to the National Research Coordinators. NRCs will report findings of their Clinical Center site visits regarding the completeness of record keeping and expertise in clinical protocol performance to the leaders of the Administrative Core. The NRCs will be in regular contact with Clinical Center staff to assess progress and answer questions. Clinical centers will facilitate these monitoring visits and provide appropriate approvals and access to the facilities to observe and participate in study visits.

If the COPDGene Executive Committee, OSMB, and/or NHLBI request changes or amendments to the protocol or consent form, IRB amendments will be submitted to the single IRB. Changes in the consent form require approval by the OSMB before they can be instituted. IRB actions in response to these amendments will be reviewed at weekly COPDGene Executive Committee meetings and with the NHLBI Project Officer.

Conflicts of Interest

Clinical Center Directors will be required to report potential Conflicts of Interest to the COPDGene Executive Committee annually. COPDGene Executive Committee members will be required to report potential Conflicts of Interest annually to the OSMB.

XVIII. DATA STORAGE AND DISTRIBUTION

The COPDGene study structure includes Cores, Centers and Clinical Centers.

Clinical Centers transmit data to the Data Coordinating Center (DCC) at National Jewish Health, biospecimens to the Biorepository at Brigham and Women's Hospital, and CT scans to the Imaging Center at NJH.

The Administrative Core oversees and coordinates all study activities. The DCC maintains confidentiality of all protected health information collected from study participants and stores subject data. The Administrative Center releases data and Biospecimens to internal and external study investigators. Internal Investigators consist of co-investigators at the 19 clinical centers and at the study centers. The Ancillary Studies and Publications Committee oversee distribution of data and Biospecimens to external investigators, outside of the COPDGene study.

The Imaging Core receives subject CT scan image files. Data and image files including study dates are shared between National Jewish Health and Brigham and Women's Hospital, under the Imaging Core.

The Pulmonary Function Testing (PFT) Core receives subject pulmonary function test data from the Clinical Centers via a file depot located at National Jewish Health provided by the DCC. The PFT Center is located in two locations, Dr. Robert Jensen, a contract consultant who assess test quality, and at National Jewish Health.

The Biorepositories, which collect, store, and distribute blood samples from the 19 Clinical Centers, are located at Johns Hopkins University and Brigham and Women's Hospital. The Phase 4 samples will be stored at the biorepository at Brigham and Women's Hospital. All samples are stored with a code but no other identifiers. Dates of blood collection are recorded in the DCC and are received but not stored at the Biorepository. Blood samples from the Biorepository may be shared with internal study investigators as well as with external investigators, as approved by the Ancillary Studies Committee.
If collected in the future, FFPE tissue blocks will also be stored in an established COPDGene Biorepository at Brigham and Women's Hospital, or at National Jewish Health.

The Genetic Analysis Core and Bioinformatics Core arelocated at Brigham and Women's Hospital. These Cores receive genetic, transcriptomic, epigenetic, metabolomic, and proteomic data for quality control and data analysis. Samples and data are coded and have no other identifiers associated with them.

The Biomarkers Core is located at National Jewish Health and works collaboratively with the Biorepository sharing biospecimens to be used for analysis of biomarkers. Data will include dates of sample collection.

The Subtyping Core is made up of investigators from the 19 clinical centers, who meet by teleconference and collaborate on data analysis. Data shared to this Center will include geocodes and location to the level of zip codes.

The Mortality Adjudication Center is located at National Jewish Health and is closely linked to the Administrative Center. They will receive death certificates, subject medical records and when the subject has consented to have personal identifiers transmitted to the DCC staff for the Mortality Adjudication Center will contact next of kin for informant interviews and request death certificates.

The Epidemiology Core is located at the University of Colorado, Denver. The Epidemiology Center participates in data analysis of phenotypes, genetic data, and imaging data. These data may include subject identifiers, including dates and geocodes (represents a geographic location).

The Principal Investigators for the study, COPDGene Associate Director, COPDGene Director of Clinical Centers, Director of the Imaging Core, and Heads of the DCC, Biomarker Core, Epidemiology Core, Genetic Analysis Core, Bioinformatics Core, Subtyping Core, and PFT Core, in consultation with the COPDGene Executive Committee, manage the central functions of the study. This team and the COPDGene Investigators seek to encourage appropriate collaborative

relationships with outside investigators to advance scientific knowledge and maximize the value of the study.

XIX. REFERENCES

1. Hoyert DL, Hsiang-Ching K, Smith BL. Deaths: Preliminary Data for 2003. National Vital Statistics Reports. 2005; 53(15):1-48.
2. Burrows B, Knudson RJ, Cline MG, Lebowitz MD. Quantitative relationships between cigarette smoking and ventilatory function. Am Rev Respir Dis. 1977; 115:195-205.
3. Silverman EK, Chapman HA, Drazen JM, Weiss ST, Rosner B, Campbell EJ, O'Donnell WJ, Reilly JJ, Ginns L, Mentzer S, Wain J, Speizer FE. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease: Risk to relatives for airflow obstruction and chronic bronchitis. American Journal of Respiratory and Critical Care Medicine. 1998; 157:1770-8.
4. Silverman EK, Palmer LJ, Mosley JD, Barth M, Senter JM, Brown A, Drazen JM, Kwiatkowski DJ, Chapman HA, Campbell EJ, Province MA, Rao DC, Reilly JJ, Ginns LC, Speizer FE, Weiss ST. Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. American Journal of Human Genetics. 2002; 70(5):1229-39. PMID: 11914989.
5. Hersh CP, DeMeo D, Silverman EK. COPD. In: Silverman EK, Shapiro SD, Lomas DA, Weiss ST, editors. Respiratory Genetics. London: Hodder Arnold; 2005.
6. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005; 6(2):95-108. PMID: 15716906.
7. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2017. 2016.
8. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P, Gustafsson P, Hankinson J, Jensen R, McKay R, Miller MR, Navajas D, Pedersen OF, Pellegrino R, Wanger J. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. European Respir J 2005; 26(4):720-35. PMID: 16204605.
9. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983; 67(6):361-70. PubMed PMID: 6880820.
10. Hogan JW, Laird NM. Mixture models for the joint distribution of repeated measures and event times. Stat Med. 1997; 16(1-3):239-57. PMID: 9004395.
11. Anthonisen NR, Connett JE, Murray RP. Smoking and lung function of Lung Health Study participants after 11 years. Am J Respir Crit Care Med. 2002; 166(5):675-9. PMID: 12204864.
12. Ohara T, Hirai T, Sato S, Terada K, Kinose D, Haruna A, Marumo S, Nishioka M, Ogawa E, Nakano Y, Hoshino Y, Ito Y, Matsumoto H, Niimi A, Mio T, Chin K, Muro S, Mishima M. Longitudinal study of airway dimensions in chronic obstructive pulmonary disease using computed tomography. Respirology. 2008; 13(3):372-8. PMID: 18399859.
13. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29(1):15-21. PMID: 23104886; PMCID: 3530905.
14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15(12):550. PMID: 25516281; PMCID: 4302049.
15. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012; 22(10):2008-17. PMID: 22722343; PMCID: 3460195.
16. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014; 15(2):R29. PMID: 24485249; PMCID: 4053721.
17. Steijger T, Abril JF, Engstrom PG, Kokocinski F, Hubbard TJ, Guigo R, Harrow J, Bertone P, Consortium R. Assessment of transcript reconstruction methods for RNA-seq. Nat Methods. 2013; 10(12):1177-84. PMID: 24185837; PMCID: 3851240.
18. Bray N, Pimentel H, Melsted P, Pachter L. Near-optimal RNA-Seq quantification. arXiv. 2015.
19. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet. 2007; 3(9):1724-35. PMID: 17907809; PMCID: 1994707.
20. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011; 89(1):82-93. PMID: 21737059.

21. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence Kernel Association Tests for the Combined Effect of Rare and Common Variants. Am J Human Genetics. 2013; 92:841-53. PMID: 23684009; PMCID: 3675243.
22. Prokopenko D, Hecker J, Silverman E, Pagano M, Nothen MM, Dina C, Lange C, Fier HL. Utilizing the Jaccard index to reveal population stratification in sequencing data: A simulation study and an application to the 1000 Genomes Project. Bioinformatics. 2016; 32(9):1366-72. PMID: 26722118.
23. Schlauch D, Fier H, Lange C. Identification of genetic outliers due to sub-structure and cryptic relationships. Bioinformatics. 2017 Jul 1;33(13):1972-1979. PMID: 28334167, PMC5870703.
24. Morrison AC, Voorman A, Johnson AD, Liu X, Yu J, Li A, Muzny D, Yu F, Rice K, Zhu C, Bis J, Heiss G, O'Donnell CJ, Psaty BM, Cupples LA, Gibbs R, Boerwinkle E. Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. Nature Genetics. 2013; 45(8):899-901. PMID: 23770607; PMCID: 4030301.
25. Fier H, Prokopenko D, Hecker J, Cho MH, Silverman EK, Weiss ST, Tanzi RE, Lange C. On the association analysis of genome-sequencing data: A spatial clustering approach for partitioning the entire genome into nonoverlapping windows. Genetic Epidemiology. 2017; 41:332-340.
26. Disease GIfCOL. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (Updated 2011) 2011 December 2, 2012. Available from: www.goldcopd.com.
27. Vestbo J, Hurd SS, Agusti AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, Stockley RA, Sin DD, Rodriguez-Roisin R. Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease, GOLD Executive Summary. American journal of respiratory and critical care medicine. 2012. Epub 2012/08/11. doi: 10.1164/rccm.201204-0596PP. PubMed PMID: 22878278.
28. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. American Journal of Respiratory and Critical Care Medicine. 1999; 159(1):179-87.
29. Ferris BG. Epidemiology Standardization Project. American Review of Respiratory Disease. 1978;118 (suppl.):1-120.
30. Society AT. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med. 2002;166(1):111-7. PMID: 12091180.
31. McCarten JR, Anderson P, Kuskowski MA, McPherson SE, Boorson Soo. Screening for Cognitive Impairement in an Elderly Veteran Population: Acceptibility and Results Using Different Versions of the Mini-Cog. J Am Geriatr Soc 2011; 59:309–313.
32. Martinez FJ, Raczek AE, Seifer FD, Conoscenti CS, Curtice TG, D'Eletto T, et al. Development and Initial Validation of a Self-Scored COPD Population Screener Questionnaire (COPD-PS). COPD: Journal of Chronic Obstructive Pulmonary Disease, 2008;. 5:2, 85-95.
33. Gimeno-Santos et al. The PROactive instruments to measure physical activity in patients with chronic obstructive pulmonary disease. ERS, 11 Jun 2015
34. Baldwin et al. Refining Low Physical Activity Measurement Improves Frailty Assessment in Advanced Lung Disease and Survivors of Critical Illness. Ann Am Thorac Soc Vol 14, No 8, pp 1270–1279, Aug 2017
35. Kim ES et al. Optimism and Cause-Specific Mortality: A Prospective Cohort Study. Am J Epidemiol, 185:21-29, https://doi.org/10.1093/aje/kww182.
36. Martinez CH, Diaz AA, Meldrum CA, McDonald MN, Murray S, Kinney GL, Hokanson JE, Curtis JL, Bowler RP, Han MK, Washko GR, Regan EA; COPDGene Investigators. Handgrip strength in chronic obstructive pulmonary disease. Associations with acute exacerbations and body composition. Ann Am Thorac Soc 2017 Nov; 14(11):1638-1645. PMID: 29090990 PMCid: https://www.ncbi.nlm.nih.gov/pubmed/29090990
37 Kaluza J, Larsson SC, Orsini N, Linden A, Wolk A .Fruit and vegetable consumption and risk of COPD: a prospective cohort study of men.Thorax 2017; 72 500-509 Published Online First: 22 Feb 2017. doi: 10.1136/thoraxjnl-2015-207851
38. Willett WC, Reynolds RD, Cottrell-Hoehner S, Sampson L, Browne ML. Validation of a semi-quantitative food frequency questionnaire: comparison with a 1-year diet record. *J Am Diet Assoc*. 1987;87(1):43-47
39. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985; 122(1):51-65
40. Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int J Epidemiol*. 1989; 18(4):858-867

The Phase 1, Phase 2 and Phase 3 COPDGene protocol items are included as appendices.

Appendix A: Phase 1(Enrollment Visit) – Completed

I. SUBJECT SELECTION: Phase I (Completed)

We are recruiting 12,000 smoking and non-smoking subjects.

Subject Inclusion/Exclusion Criteria
A total of 12,000 non-Hispanic White and African-American subjects will be recruited. These subjects will be classified into four groups:
(1) Up to 4500 subjects with COPD GOLD Stages 2 through 4,
(2) Up to 4500 control subjects – current or former smokers without airflow limitation,
(3) Up to 1500 subjects with minimal airflow limitation (COPD GOLD Stage 1) and including smokers unclassified by GOLD criteria (26, 27),
(4) Up to 1500 non-smoking control subjects, 500 African-American subjects and 1000 non-Hispanic White subjects, with no smoking history and no airflow limitation.
Subjects will be collected at 21 clinical centers in the United States.

General Inclusion and Exclusion Criteria for All Study Subjects
The following criteria will be required on ALL smoking and non-smoking study subjects in the first study visit:

Inclusion Criteria
- Age 45-80 years (Age 45-85 for non-smoking subjects)
- Non-Hispanic Whites and African Americans

Exclusion Criteria
- Other concomitant respiratory disorder (such as, but not limited to, diffuse bronchiectasis, cystic fibrosis, or interstitial lung disease)
- Lung surgery with removal of a lobe or more (including lung volume reduction and lung transplantation)
- Lung cancer, known or suspected
- Bronchoscopic lung volume reduction
- Pregnancy or suspected pregnancy
- Uncontrolled cancer, as defined as ongoing radiation therapy, ongoing chemotherapy, narcotics for pain control, or known metastatic disease
- History of radiation therapy to the chest (other than radiation for breast cancer)
- Use of antibiotics and/or systemic steroids (new prescription or increased dose) for a COPD exacerbation or any lung infection within the last month
- Inability to use albuterol
- First or second degree relative (parent, brother, sister, daughter, son, aunt, uncle, nephew, niece, half-sibling, grandparent, grandchild) of a subject enrolled in COPDGene®
- Subjects who indicate they are in more than one racial category

Metal objects that may interfere with chest CT quantification including presence of a cardiac pacemaker, defibrillator, metal prosthetic heart valve, metal projectile or metal weapon fragment (bullet, shrapnel, shotgun shot) or metal shoulder prosthesis

Subjects unable to perform spirometry due to:

- chest or abdominal surgery in the past three months
- a heart attack in the last three months
- detached retina or eye surgery in the past three months
- hospitalization for any other heart problem in the past month

Participation in the ECLIPSE study

Inability to provide telephone contact number(s) and two additional contacts

No place of permanent residence of three months or more

Each of the four subject groups contains specific criteria that define the distinct group. The criteria are relevant to the clinical and epidemiological categorization of the four study groups.

COPD Subjects:

Additional Inclusion Criteria

Smoking history of ≥ 10 pack-years

Diagnosis of COPD (post-bronchodilator $FEV_1/FVC < 0.70$) Stages 1, 2, 3 and 4 by GOLD criteria (26)

Additional Exclusion Criteria

Smoking history of < 10 pack years

The diagnosis of COPD includes airflow limitation on spirometry and history of risk factors (most commonly cigarette smoking) known to cause COPD (26). We will use NHANES predicted spirometry values obtained in the United States (28). There are no uniform criteria for the amount of cigarette smoking required for the diagnosis of COPD. We have chosen a threshold for cigarette consumption (10 pack-years) to ensure that all subjects have a substantial environmental stress to differentiate those who have an abnormal pulmonary response to cigarette smoke and those who do not have such a response. Review of medical records of patients with COPD frequently demonstrates a physician-listed diagnosis of asthma. These subjects will be included in the COPD groups if they have evidence of airflow limitation that is not fully reversible using the GOLD (Global Initiative on Obstructive Lung Disease) criteria (26). The interactions of asthma with COPD are common and complex. Elimination of asthma patients with fixed airflow limitation would inappropriately bias the findings of this study.

We will employ the FVC as the primary measure of lung volume since this allows comparisons to previous large-scale epidemiologic investigations. Normal values have been published and this maneuver is widely used. However, we will also collect the FEV_6, a more recently advocated index that may be easier for some patients to accomplish and reduce adverse effects of spirometry.

Subjects with known or suspected lung cancer will be excluded; if there are a substantial number of subjects with lung cancer in this study, then we may detect genes associated with lung cancer rather than with COPD. Subjects with a prior history of lung cancer, even if successfully resected and cured will be excluded. Subjects who may in the near future have a resection of a lesion suspected to be lung cancer will be excluded, but they may be included at a later time if the lesion proves not to be cancer and the resection removes less than one lobe of the lung.

Subjects with uncontrolled cancer of any type will be excluded because the cancer or cancer treatment may alter lung function and thus misclassify the respiratory status of the subject, and this cohort is being designed for long-term follow-up.

We will exclude patients with recent use of antibiotics and/or systemic steroids for a COPD exacerbation since such subjects may have an ongoing inflammatory pulmonary process related to an infection that will impair our ability to characterize and phenotype patients in their usual stable state. These patients can be re-screened 30 days after cessation of antibiotic or corticosteroid use.

GOLD Unclassified Subjects:

Additional Inclusion Criteria

Smoking history of ≥ 10 pack-years

Spirometry (Post-bronchodilator $FEV_1/FVC \geq 0.70$, $FEV_1 < 80\%$ predicted)

Additional Exclusion Criteria

Smoking history of < 10 pack years

Subjects meeting the required smoking history parameter but do not fall into either COPD or smoking control categories will be included as GOLD Unclassified subjects. These subjects will be grouped based on the occurrence of a normal FEV_1/FVC accompanied by a presence of reduced air flow (post-bronchodilator $FEV_1/FVC \geq 0.70$, $FEV_1 < 80\%$). To date, we have found approximately 10-12% of smoking subjects fall into this category. The inclusion of the Unclassified smoking group will provide additional and possibly novel information regarding the categorization, clinical presentation, and progression of COPD in smoking populations. Since this group is already present within the current subject pool, these subjects will be clarified as a distinct subject group.

Smokers without COPD

Additional Inclusion Criteria

History of cigarette smoking ≥ 10 pack-years

Post-bronchodilator $FEV_1/FVC \geq 0.70$ and $FEV_1 \geq 80\%$ predicted.(28)

Additional Exclusion Criteria

Smoking history of < 10 pack years

Some potential subjects may have exclusionary criteria that only temporarily limit their enrollment in COPDGene®, such as recent use of antibiotics or corticosteroids. Such subjects with temporary exclusion may be re-screened and enrolled at a later time when these features are no longer exclusionary.

Subjects who have been given a diagnosis of COPD by a health care professional, but who have normal spirometry, will be enrolled as smoking control subjects. The diagnosis of COPD may have been incorrectly given to a patient without confirmation by spirometry. For this study, the diagnosis of COPD or lack thereof (based on the objective presence or absence of airflow limitation) will be determined by the post-bronchodilator spirometry performed as part of this study.

Sources of subjects will vary from center to center, but will likely include inpatients and outpatients at the centers, spouses and friends of subjects with COPD, patients in primary care practices, local

patient support and educational groups, and local and national COPD voluntary organizations (such as the COPD Foundation and American Lung Association).

Subjects should not primarily be recruited from sources that include a high prevalence of asthmatics such as asthma clinics or asthma patient groups. While subjects with asthma are not excluded in either control or COPD subjects in order to assure similar inclusion/exclusion criteria in both populations, this study is not designed as a study of the genetics of asthma. Thus subjects who have asthma as their primary respiratory disease should not be targeted for recruitment.

Although the primary focus of this project is COPD and COPD-related phenotypes, subjects will also be informed that this cohort may be used to study the genetic and environmental determinants of other smoking-related illnesses such as lung cancer and coronary artery disease and, with their permission on the consent form, other disorders that are not smoking-related.

Non-smoking Controls

Additional Inclusion Criteria

No smoking history as defined by less than 100 cigarettes smoked in a lifetime
No airflow limitation (Post-bronchodilator $FEV_1/FVC \geq 0.70$, $FEV_1 \geq 80\%$ predicted)

Additional Exclusion Criteria

Smoking history of more than 100 cigarettes smoked in a lifetime
Smoking history of more than 52 cigars smoked in a lifetime
Smoking history of more than 12 oz. of pipe tobacco smoked in a lifetime
Any history of physician-diagnosed respiratory disease

One hundred and eight subjects with no smoking history and no airflow limitation were included as a reference population for comparison with those affected by smoking exposure. These subjects will offer information on the processes involved in normal lung aging. Non-smoking control subjects are critical for supporting genetic and pathological findings within smoking and diseased subject groups by acting as a baseline for normal pulmonary physiology within a genetically mixed population.

Subject Phase 1 Actual Enrollment:

Phase 1, which was performed from November 2007 until July 2012, enrolled 10,371 subjects. Of these subjects 3438 were African American and 6933 were Non-Hispanic White. Enrolled subjects classified by severity of lung disease were as follows:

Smoker Controls – 4391
GOLD 2 - 4 Subjects – 3692
GOLD Unclassified – 1258
GOLD 1 Subjects – 795
Non-smokers controls – 108
Lack final GOLD classification due to failed spirometry – 63
Significant ILD or bronchiectasis noted on CT scan, excluded from analysis – 64

All of the enrolled subjects had at least a 10 pack-year smoking history (except the non-smoker group) and were between 45 and 80 years of age. Enrollment of up to a total of 1500 non-smoker controls is still anticipated.

II. SUBJECT ENROLLMENT

Phase 1
This project will recruit a total of 12,000 subjects over five years from twenty-one clinical study centers. We anticipate recruitment will be completed early in the fifth year of the proposed study, allowing time for the genetic and epidemiologic analyses to occur by the end of the fifth year. Each Clinical Center will be expected to, on average, recruit approximately 160 subjects in each of the first four years meeting the program's defined recruitment goals for recruitment among both racial/ethnic classes, an approximately equal division of men and women, and with COPD subjects equally distributed among GOLD grades 1, 2, 3 and 4. The Executive Committee with the help of the Steering Committee will set and monitor goals for subject recruitment from each Clinical Center. The Administrative Core will have the right to modify recruitment goals among the centers to meet the overall goals of the project with respect to numbers of subjects recruited, minority subjects, or the balance of subjects in GOLD stages. The Administrative Core may incorporate additional qualified Clinical Centers if needed to meet recruitment goals.

Potentially eligible subjects who contact study staff or are referred by a health care professional may undergo screening either in person or by telephone to determine if they are likely to meet appropriate inclusion and exclusion criteria and to schedule the study visit to the Clinical Center. The study protocol will be discussed in detail during this screening encounter. A log will be maintained at each Clinical Center indicating the number of subjects who fail this initial screening. This screening log will be maintained at the Clinical Center and will not be transmitted to the Data Coordinating Center to assure confidentiality and protection of human subjects.

Before the study visit begins, written informed consent will be obtained by one of the study staff members. A physician investigator from the Clinical Center will be available to answer any questions regarding the informed consent process. Investigators may obtain consent from their own patients.

III. PHASE I STUDY PROCEDURES (Completed)

There will be one to two study visits for the majority of subjects depending on whether a separate visit is needed to schedule a chest CT scan. In some cases, subjects may have a third visit to re-collect information that does not meet quality control criteria. For example, if spirometry does not meet quality control criteria as judged by the Pulmonary Function Core, the subject may be called to schedule an additional visit to repeat the test.

During their research study visit, the following procedures will be performed:

1. Informed consent will be obtained prior to any other study procedure.
2. Prior to any other evaluations, an Eligibility Questionnaire will be administered to determine if potential subjects meet inclusion/exclusion criteria. This questionnaire is located on the protected study web site and will provide a check on whether subjects meet inclusion and exclusion criteria.
3. Contact Information will be collected from the subject. Name, home address, phone number, cell phone, email address, date of birth and social security number will be collected from the

subject. The purpose of this information is to maintain contact with the subjects up to four times a year for the purposes of longitudinal follow-up. Similar contact information (but not social security number) will be collected for two other individuals likely to know the subject's whereabouts, at least one of whom is a relative not living with the subject. The purpose of this information is to locate subjects who have moved to a different home and have a different address and/or phone number and ascertain vital status of the subject.

4. Safety Assessment. A Safety Assessment questionnaire will be administered to subjects. They will be asked if they have a history of adverse events with the use of albuterol prior to administration of this inhaled bronchodilator. This safety assessment will also be used to ascertain potential cardiac disorders prior to albuterol administration and the walk test. If the questionnaire uncovers known or potential adverse effects of albuterol or potential cardiac disorders, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a six-minute walk test.

5. Physical Assessment will be performed on all subjects: A limited assessment will be performed including height, weight, blood pressure, heart rate, and room air oxygen saturation by pulse oximetry with the subject seated at rest. In subjects using supplemental oxygen, the oxygen will be withheld and subjects will breathe room air for 10 minutes prior to recording oxygen saturation. If resting saturation falls below 82% or the subject becomes dyspneic prior to 10 minutes without oxygen, oxygen therapy will be restarted as a safety precaution. If heart rate is less than 50 or greater than 110 or if the blood pressure is greater than 170/100, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a walk test.

6. Pre and post-bronchodilator spirometry will be performed on all subjects. The maneuvers are performed in a standardized manner. To obtain three acceptable measures, the technician may ask the subject to perform up to eight attempts. Spirometry is performed before and then repeated twenty to thirty minutes following the administration of two puffs (180 mcg) of albuterol from a metered dose inhaler with a spacer. Spirometry will be performed with the subject in the seated position with a nose clip in place. Inspiratory capacity will also be measured post-bronchodilator.

7. Standardized questionnaires will be completed by all subjects that meet entry criteria to assess respiratory history and symptoms, smoking history, family history, and other medical history. These questionnaires include a modified version of the American Thoracic Society/Division of Lung Diseases Respiratory Epidemiology Questionnaire (29) to assess respiratory symptoms and the St. George's Respiratory Questionnaire to assess health-related quality of life. Medications and oxygen use will also be recorded. Questionnaires can be administered by the method judged to be most convenient at each Clinical Center. Questionnaires may be interviewer-administered or self-completed on either a paper copy or directly on a computer pdf file. Completion of all questionnaires is expected to take between 45 and 90 minutes.

8. Blood is drawn from all subjects for DNA (genetic association studies) and serum and plasma (for measurement of other proteins potentially related to COPD and other diseases). A total volume of approximately 40 ml of blood will be drawn for this study. Blood samples will be stored at the Johns Hopkins COPDGene® Biological Repository.

9. Six-Minute Walk Test will be performed on all subjects to determine exercise capacity (30). This will be used to calculate the BODE score. Subjects will be asked questions to assure they do not have significant or occult heart disease prior to the test to assure subject safety.
10. Chest CT Scan will be performed to assess for emphysema and airway disease in all subjects. An inspiratory chest CT scan will be performed with a radiation dose of 200 mAs to provide thorough assessment of small airway wall thickness and emphysema. An expiratory chest CT scan will be performed of lower dose (50 mA) to assess for air trapping. If a clinical chest CT scan with an appropriate CT protocol and data storage has been performed within 6 months, that clinical CT scan may be used for this study with prior approval of the Imaging Committee that the scan algorithm used meets study criteria and can be analyzed appropriately. If a CT scan may not be scheduled on the same day as the study visit, the scan must be performed within 6 months after the visit. All COPD, Unclassified, and control subjects will be required to have a chest CT scan. If the chest CT scan is done on the same day as the post-bronchodilator spirometry, the scan should be scheduled as soon as possible after albuterol. The time and date of the last bronchodilator prior to the CT scan will be captured.
11. A urine pregnancy test will be performed on all premenopausal women prior to the chest CT scan.
12. Medical Record Review: In subjects with COPD, medical records will be reviewed when available with subject permission to obtain information within the last year of pulmonary function tests (including lung volumes and diffusing capacity) and oxygenation (arterial blood gases).
13. Longitudinal Follow-Up: Subjects will be contacted up to four times a year by telephone, mail, email, and/or newsletter to assess survival status and respiratory illnesses for up to 10 years.
14. Linking to other studies: Subjects with COPD will be asked if they are currently participating or have participated previously in other COPD NIH and non-NIH clinical trials such as the NHLBI COPD Clinical Research Network MACRO or LEUKO studies, NHLBI Long-Term Oxygen Therapy Trial (LOTT), National Lung Cancer Screening Trial (NLST), GlaxoSmithKline International COPD Genetics Network, Boston Early-Onset COPD Study, NHLBI Lung Tissue Research Consortium (LTRC), NHLBI Lung Health Study, and NHLBI SPIROMICS Study. Subjects will be asked if their clinical and genetic data can be linked to the results of these other studies and future investigations in order to investigate the genetic associations between genotype data from this study with their outcomes in other trials.

After the genotyping of candidate genes, including alpha 1-antitrypsin (AAT), is performed, subjects who elected to learn about abnormal AAT test results, when IRB approved, will receive them. These results may not be available for several years after the blood samples are obtained. For newly diagnosed PI ZZ subjects, the Principal Investigator of the Clinical Center will telephone those subjects first, then send a follow-up letter. Subjects with other abnormal PI types will be informed by mail. Participants will be informed that the AAT test results are based on research laboratory test results and should be repeated in a clinical laboratory. If subjects provide permission, these results will also be communicated to their physician.

The questionnaire and pulmonary function test results will be stored in a locked filing cabinet at each Clinical Center. Questionnaire and pulmonary function data, identified by study ID number only, will be transmitted to the Data Coordinating Center in Denver by secure internet connection with 132-bit encryption. Blood samples and CT scans will be transmitted to the appropriate Biorepository and Imaging Cores by overnight delivery service.

Appendix B – Phase 2 (Year 5 and Year 1 for New Nonsmokers)- Completed

I. COPDGene Phase 2 Return Visit

In addition to the cross-sectional subject data collected in the baseline COPDGene study visit, disease progression and incidence of COPD in smokers are important additional endpoints for genetic association studies. We have completed enrollment of 10,300 subjects as the first phase of this project and now propose to invite these subjects to return for a second evaluation five years after their initial visit to assess disease incidence and/or progression in COPDGene subjects.

The primary goals of COPDGene Phase 2 are: 1) to identify new genetic loci that influence the development and/or progression of chronic obstructive pulmonary disease (COPD) and COPD-related phenotypes, and 2) to reclassify COPD into subtypes that can ultimately be used to develop effective subtype-specific therapies and prevention. The reclassification of COPD will be done using imaging, clinical and physiologic characteristics, longitudinal progression, long-term outcomes, and genetics.

In addition to identifying COPD genetic determinants, this program will characterize the natural history of COPD and identify well-characterized COPD subtypes. Improved understanding of COPD subtypes and genes controlling susceptibility to COPD could lead to novel patho-physiological insights, refined diagnostic criteria, and new treatment approaches. Moreover, the availability of comprehensive genetic data and longitudinal data on a large biracial group of smokers will be an invaluable national resource for other investigators.

New Non-Smoker Controls

To reach our Phase 1 goal of up to 12,000 subjects, we plan to enroll up to an additional 1,500 subjects in Phase 2. We will target enrollment of additional non-smokers without lung disease in Phase 2 of COPDGene. By providing insight into the normal effects of aging in healthy subjects, these subjects will allow us to correctly interpret the chest CT scans, lung function, health status and other features of our current cohort. For example, aging is associated with loss of lung elasticity, a characteristic feature of emphysema. Without knowledge of the effects of age, gender and race on the lungs of healthy subjects, we cannot accurately determine whether CT scans of subjects with a smoking history are normal or abnormal, or assess disease severity.

II. SUBJECT SELECTION – Phase 2

Return Visit

This project will invite all 10,371 subjects who successfully completed Visit 1 using the same 21 clinical study centers for a second study visit. Visit 2 will be scheduled five years (+ / - 3 months) after Visit 1. However, late visits are permitted if the subject becomes available for a Phase 2 visit more than three months after their five-year anniversary.

A key component of the project is to establish contact with all of the previously enrolled subjects or confirm the subjects are deceased. Clinical centers will utilize stored contact information including secondary contacts to contact previously enrolled subjects.

Eligible subjects will be contacted by mail, email and/or phone and invited to participate in Visit 2 at the Clinical Centers. The study protocol will be discussed in detail during this screening encounter. Before the study visit begins, written informed consent will be obtained by one of the study staff members. A physician investigator from the Clinical Center will be available to answer any questions regarding the informed consent process.

New Non-Smoker Control Subjects

Additional subjects without a smoking history will be enrolled in Phase 2 to meet the target of up to 1,500 subjects noted above. Subjects will be recruited from family members and friends of COPDGene enrolled subjects; this primary targeted recruitment will help assure age, race and gender distribution of control subjects is similar to the enrolled COPDGene subjects. If necessary, control subjects will also be enrolled from the community and physician offices using informational brochures and advertisements.

Subject Inclusion/Exclusion Criteria for Non-smoker Controls

Inclusion Criteria

- Age 45-85 years (for non-smokers)
- Non-Hispanic Whites and African Americans
- No smoking history as defined by less than 100 cigarettes smoked in a lifetime, less than 52 cigars smoked in a lifetime, and less than 12 oz. of pipe tobacco smoked in a lifetime

Exclusion Criteria

- Respiratory disorders (including, but not limited to, COPD, asthma, bronchiectasis, cystic fibrosis, or interstitial lung disease)
- Any lung surgery
- Lung cancer
- History of Pneumothorax
- Pleural procedure (including, but not limited to, chest tube placement, pleural surgery)
- History of any pulmonary procedures: diagnostic bronchoscopy, lung biopsy, except research bronchoscopy
- Uncontrolled cancer, as defined as ongoing radiation therapy, ongoing chemotherapy, narcotics for pain control, or known metastatic disease
- History of radiation therapy to the chest (other than radiation for breast cancer)
- Subjects receiving treatment for active TB
- Severe cirrhosis
- Renal disease requiring dialysis
- Pulmonary Hypertension
- Congestive heart failure
- Inability to use albuterol
- First or second- degree relative (parent, brother, sister, daughter, son, aunt, uncle, nephew, niece, half-sibling, grandparent, grandchild) of a subject enrolled in COPDGene
- Subjects who indicate they are in more than one racial category
- Subjects who do not authorize the storage of personal identifiers (social security number, name, address, phone number) in the COPDGene Data Coordinating Center
- Metal objects that may interfere with chest CT quantification including presence of a cardiac pacemaker, defibrillator, metal prosthetic heart valve, metal projectile or metal weapon fragment (bullet, shrapnel, shotgun shot) or metal shoulder prosthesis

Participation in the following studies: COPDGene Phase 1, ECLIPSE, Boston Early-Onset COPD Study, Boston COPD Exacerbations Study, Denver Genetics Study, GlaxoSmithKline International COPD Genetics Network, NCI: National Lung Screening Trial, NIH: COPD Clinical Research Network: Macrolide and Leukotriene Trials, NIH: Long-Term Oxygen Therapy Trial, NIH: Lung Health Study, NIH: Lung Tissue Research Consortium, NIH: SPIROMICS, Pittsburgh SCCOR

Inability to provide personal telephone contact number(s) and complete information for two additional contacts (next of kin, relative or close friend, not living with subject)

No place of permanent residence for the prior three months

Alpha-1 Antitrypsin deficiency

Regular use of an inhaled medication

Regular marijuana use (more than once per week)

HIV Positive

History of transplant of any organ

Active drug or alcohol dependence

History of employment in underground mining

Body mass index greater than 35

Temporary Exclusion Criteria

Subjects with the following may not have a study visit performed, but may be re-screened at a later time.

1. Pregnancy or suspected pregnancy. (These subjects will be re-scheduled for a visit at least three months after childbirth or after they are no longer suspected of pregnancy.)
2. Use of antibiotics for a respiratory infection within the past month. (These subjects will be rescheduled for a visit at least 4 weeks after completion of the antibiotics.)
3. Use of systemic corticosteroids (new prescription or increased dose) for a respiratory disorder within the past month. (These subjects will be rescheduled for a visit at least 4 weeks after completion of the steroids.)
4. Chest or abdominal surgery in the past three months. (These subjects will be rescheduled for a visit at least three months after surgery.)
5. Heart attack in the last three months. (These subjects will be rescheduled for a visit at least three months after the heart attack.)
6. Detached retina or eye surgery in the past three months. (These subjects will be re-scheduled for a visit at least three months after surgery.)
7. Hospitalization for any other heart problem in the past month. (These subjects will be rescheduled for a visit at least 4 weeks after hospital discharge.)

Subject Inclusion/Exclusion Criteria – Phase 2 Return Visit

Inclusion Criteria

Previously enrolled in COPDGene and meeting initial eligibility criteria at Visit 1.

Control subjects without a history of cigarette smoking (see criteria above).

Temporary Exclusion Criteria

Subjects with the following will be temporarily excluded but may be rescheduled at a later time:

1. Pregnancy or suspected pregnancy. (These subjects will be re-scheduled for a visit at least three months after childbirth or after they are no longer suspected of pregnancy.)
2. Use of antibiotics for a respiratory infection within the past month. (These subjects will be rescheduled for a visit at least 4 weeks after completion of the antibiotics.)
3. Use of systemic corticosteroids (new prescription or increased dose) for a respiratory disorder within the past month. (These subjects will be rescheduled for a visit at least 4 weeks after completion of the steroids.)
4. Chest or abdominal surgery in the past three months. (These subjects will be rescheduled for a visit at least three months after surgery.)
5. Heart attack in the last three months. (These subjects will be rescheduled for a visit at least three months after the heart attack.)
6. Detached retina or eye surgery in the past three months. (These subjects will be rescheduled for a visit at least three months after surgery.)
7. Hospitalization for any other heart problem in the past month. (These subjects will be rescheduled for a visit at least 4 weeks after hospital discharge.)

Rationale: All previously enrolled subjects from COPDGene are eligible for the study. Previous exclusions for lung cancer and other conditions are not a reason to exclude subjects from follow-up evaluation because these are outcomes of the disease to be assessed.

We will exclude patients with recent use of antibiotics and/or systemic steroids for a COPD exacerbation since such subjects may have an ongoing inflammatory pulmonary process related to an infection that will impair our ability to characterize and phenotype patients in their usual stable state.

We will exclude subjects who have contraindications to spirometry (exclusion criteria 4 – 7 above).

Subjects with exclusion criteria will be re-screened at a later date when they can be enrolled in Visit 2. The Administrative Core should be contacted regarding subjects who cannot complete all components of the study visit due to temporary exclusions or other medical conditions. The Administrative Core in consultation with the local clinical center director will determine which portions of a study visit can be completed. Those modifications will be considered if the subject is unlikely to be able to return at a later date for the study visit or unlikely to resolve medical issues that limit participation.

Phase 2

All Phase 1 subjects have provided consent with their initial study enrollment to be contacted again for further research studies. Subjects will be sent a letter advising them that funding has been provided for the second visit and informing them that they will be contacted by letter, email and/or phone to schedule an appointment for Visit 2. A telephone call will be made to each subject to schedule a date for his/her next study visit and ask them to bring the following to the study visit:

1) all their current oral and inhaled medications, 2) names of any injectable study medications received on a regular basis, 3) social security number, 4) name, address and phone number of their personal physician, 5) name, address and phone of two other individuals likely to know the subject's whereabouts, at least one of whom is a relative not living with the subject, and 6) driver's license or other identification to assure they are the individual enrolled in COPDGene Phase 1.

At the Phase 2 study visit the subject will review the informed consent form for the second study visit and its study procedures. The subjects will have the opportunity to review the study documents in private, have questions answered and sign the informed consent. If necessary, subjects may review the informed consent with family members or their physicians. A HIPAA authorization or combined consent and HIPAA authorization will be reviewed and signed at the same time. Although subjects have signed a HIPAA authorization in Phase 1, a current authorization will be obtained to meet the requirements of hospitals and health care providers. Phase 2 data collection includes transfer of subject identifiers to the Data Coordinating Center (DCC) as described in the Central Data, Imaging, and Human Subject Contacts Protocol.

Mail Consent

Prior to the return visit, subjects may be contacted by the Clinical Center via mail to request additional consent and authorization. This additional consent and authorization will request that the subject allow transfer of identifiers from the Clinical Center to the Data Coordinating Center at National Jewish Health. This Mail Authorization provides permission for the Clinical Center to transmit the subject's personal information (social security number, address, secondary contact address) to the DCC. A Personal Contact Information Update form will be sent with the mail consent, asking subjects to update contact information, including secondary contacts and primary care physician. Additionally the subjects will be invited to sign an "Authorization to Release Protected Health Information" medical record release form.

This information will be used for central subject tracking, longitudinal follow up, vital status searching, and to obtain cost and utilization information on subjects from Medicare. The mail consent also allows the Clinical Center to seek perimortem medical records and informant interviews about the circumstances around the subject's death.

Visit 2 Options for Subjects Who Have Difficulty with a Visit to Clinical Center

In the event that a Phase 1 subject has moved away from their local Clinical Center and to an area closer to another participating COPDGene Clinical Center, the subject may elect to complete Visit 2 at this closer Center rather than traveling to their original Center. With the subject's verbal approval by phone, their new contact information will be provided from the Clinical Center where the subject had the first visit to the geographically proximate Clinical Center. The subject will sign a new Informed Consent Form and HIPAA with the coordinator at the new study location prior to their Phase 2 visit.

In the event that Phase 1 subject has financial hardship or has moved a considerable distance from their local Clinical Center, the COPDGene Administrative Core may consider providing transportation assistance to this subject for a return visit. The payment for this subject's return visit may include the actual reasonable cost of transportation, lodging, and meal expenses. Subject must provide receipts to Clinical Center for reimbursement. Use of this approach and funding for these

expenses will be allocated on a case-by-case basis from the Administrative Core to the Clinical Center.

COPDGene subjects who are unable to come to a COPDGene Clinical Center to complete a Phase 2 visit due to illness, disease progression, or inability to be transported to the Clinical Center, will be given the option of having a research coordinator complete a partial study visit at the subject's place of residence. This home visit will include spirometry before and after albuterol administration, blood drawing, and completion of all study questionnaires.

Phase 1 subjects who are unable to come to a Clinical Center and are not willing or able to undergo a home visit will be asked if they would mind completing study questionnaires over the phone. If the subject agrees, this will be completed. If the subject is not willing or unable to complete study questionnaires over the phone they will be ask if they would complete a Limited Follow-Up Questionnaire and some or all of the regular study questionnaires. The Limited Follow-Up Questionnaire will be used to gather important longitudinal information on the subject related to respiratory symptoms, exacerbations, and comorbidities. A research coordinator will conduct this questionnaire with the subject by phone. These subjects may still participate in the bi-annual Longitudinal Follow-Up program by phone or Internet.

Phase 1 subjects who do not agree to a Phase 2 visit or to participate in Longitudinal Follow-Up will be asked if they would be willing to complete a Limited Follow-Up Questionnaire by phone. If the subject refuses to complete this questionnaire, their wishes will be respected.

Closed Clinical Center
Due to the longitudinal nature of the study, it is important to follow health and vital status on all of the COPDGene subjects. In the event that a clinical center is unable to perform follow up contacts or visits, the COPDGene Administrative Core will work as an agent for the Clinical Center or provide a contract clinical research organization to manage the follow-up subject visits and contacts per protocol. Subjects will be consented in Phase 2 to allow for a member of COPDGene Central Administration to contact them for follow up. The original clinical center must work with the Administrative Core to facilitate efficient transfer of information and study management under the supervision of a local principal investigator and with the approval of the local IRB.

III. PHASE 2 STUDY PROCEDURES

Similar to the first visit, for the Phase 2 study visit there will be one to two visits for the majority of subjects depending on whether a separate visit is needed to schedule a chest CT scan. In some cases, subjects may have a third visit to re-collect information that does not meet quality control criteria. For example, if spirometry does not meet quality control criteria as judged by the Pulmonary Function Core, the subject may be contacted by phone to schedule an additional visit to repeat the test.

During the Phase 2 research study visit, the following procedures will be performed:

1. Informed consent will be obtained prior to any other study procedure. An updated HIPAA consent to obtain medical records to adjudicate the cause of death will be obtained. Subjects will be asked to provide a signed release of medical records to be used in the event of their

death, to obtain an interview with a next of kin, secondary contact, or personal representative and with the treating physician. Hospital and physician records, related to the events and illnesses associated with a participant's death, will also be obtained through the signed release of medical records. Subjects will be provided a memo to give to their physician(s) for this purpose.

2. Subjects will be asked to provide or update contact information of two secondary contacts with addresses and phone numbers, one of who is a next of kin or personal representative, and the other a relative or close friend, not living with the subject. The purpose of this information is to locate subjects who have moved to a different home and have a different address and/or phone number and ascertain vital status of the subject. The subjects will be asked to give consent to have their participation in the COPDGene study disclosed to these secondary contacts in the event that we are unable to contact them, and also disclosed to their next of kin, personal representative or family members in the event of their death. Subjects will be provided a memo to give to their secondary contacts for this purpose.

3. Prior to any other evaluations, an Enrollment/Safety Questionnaire will be administered to determine if potential subjects have a condition that would require deferring their study visit. They will be asked if they have a history of adverse events with the use of albuterol prior to administration of this inhaled bronchodilator. This safety assessment will also be used to ascertain potential cardiac disorders prior to albuterol administration and the walk test. If the questionnaire uncovers known or potential adverse effects of albuterol or potential cardiac disorders, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a six-minute walk test. This questionnaire is located on the protected study web site.

4. Updated contact information will be collected from the subjects. Name, home address, home phone number, cell phone, email address, date of birth and social security number will be collected or verified from the subject. The purpose of this information is to maintain contact with the subjects twice a year and inform the subjects of other research studies and potential future COPDGene follow-up visits pending future funding. In order to perform central collection of mortality and cause of death, this information will be securely transmitted to the Data Coordinating Center at National Jewish Health, where it will be stored separately from the clinical, physiological, imaging, and genetic data. The transmittal of subjects' personal information to the DCC will be used for central subject tracking, longitudinal follow-up, death searching through the Social Security Death Master File (SSDMF) and to obtain cost and utilization information from the Centers for Medicare and Medicaid Services (CMS).

5. Physical Assessment will be performed on all subjects: A limited assessment will be performed including height, weight, arm span, waist circumference, blood pressure, heart rate, and room air oxygen saturation by pulse oximetry with the subject seated at rest. Blood pressure will be measured three times with the subjects seated and resting for at least five minutes. In subjects using supplemental oxygen, the oxygen will be withheld and subjects will breathe room air for 10 minutes prior to recording oxygen saturation. If resting oxygen saturation falls below 82% or the subject becomes dyspneic prior to 10 minutes without oxygen, oxygen therapy will be restarted as a safety precaution. If heart rate is less than 50 or greater than 110 or if the blood pressure is greater than 170/110, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a walk test.

6. Pre and post-bronchodilator spirometry will be performed on all subjects. The maneuvers are performed in a standardized manner. To obtain 3 acceptable measures, the technician may ask the subject to perform up to 8 attempts. Spirometry is performed before and then repeated approximately twenty minutes following the administration of two puffs (180 mcg) of albuterol from a metered dose inhaler with a spacer. Spirometry will be performed with the subject in a seated position with a nose clip in place.

7. Diffusing capacity will be measured on all subjects after post-bronchodilator spirometry. The test will be performed according to American Thoracic Society/European Respiratory Society standards (8). In subjects using supplemental oxygen, oxygen will be removed and subjects will breathe room air for 10 minutes prior to the test. If resting saturation falls below 82% or the subject becomes dyspneic prior to 10 minutes without oxygen, oxygen therapy will be restarted as a safety precaution and the test will be performed without the ten-minute oxygen washout. To obtain 2 reproducible measures, the test may be repeated up to five times.

8. Standardized questionnaires will be completed by all subjects to assess respiratory symptoms, smoking history, and other medical history. These questionnaires include the St. George's Respiratory Questionnaire to assess health-related quality of life, the Medical Outcomes SF-36 to assess generic quality of life, and the COPD Assessment Test (CAT) to assess COPD severity. Questionnaires will be administered to obtain information about respiratory history, medical history, anxiety and depression (Hospital Anxiety and Depression Scale) (9), menstrual, pregnancy and female hormone medication history in women (Women's Questionnaire), current and previous residential history (Residential/ Occupational Questionnaire) to ascertain air pollution and environmental exposures, and occupational history. Subjects will be asked about access to regular medical care, access to health insurance and attitudes about medical care for their lung disease (Socioeconomic Questionnaire) to determine the impact of these factors on lung disease and exacerbations. If a participant is found to have cancer, determined by the Medical History Questionnaire, a Cancer Data Collection Questionnaire will be administered during the study visit to collect additional data on the cancer diagnosis. The subject may be invited to sign a tissue release addendum in order to provide formalin fixed paraffin embedded (FFPE) tissue blocks if there was a cancer diagnosis or lung tissue removed. An e-Cigarette questionnaire will be administered. Subjects will answer questions to assess amount of use and attitudes.

9. Each patient will be asked to bring in all of their current medications in their original bottles. A list of current medications will be made at the local clinical center and transmitted to the Data Coordinating Center as a pdf document coded with the subject's study ID number. Oxygen use will also be recorded. A clinical coordinator will administer the Questionnaires. The results will be entered directly into the computer or recorded on paper forms to be input later by the coordinator. Completion of all questionnaires is expected to take between 90 and 120 minutes.

10. Blood will be drawn from all subjects for DNA (genetic and epigenetic studies), serum and plasma (for measurement of other proteins potentially related to COPD and other diseases) and a CBC (complete blood count). A total volume of approximately 50 ml of blood will be drawn for this study. A complete blood count will be performed at the Clinical Centers. The rest of the blood samples will be stored at the COPDGene Biological Repository at Brigham and Women's Hospital.

11. Six-Minute Walk Test will be performed in a standardized manner on all subjects to determine exercise capacity (11). Subjects will be asked questions to assure they do not have significant or occult heart disease prior to the test to assure subject safety. Subjects who have significant cardiovascular disease may require evaluation by the local physician investigator prior to six-minute walk test. Immediately and one minute after the end of the six-minute walk test, a subject's heart rate and oxygen saturation will be measured using a pulse oximeter.

12. Chest CT Scans will be performed to assess emphysema and airway disease in all subjects. An inspiratory chest CT scan will be performed with a radiation dose of 200 mAs in order to provide thorough assessment of small airway wall thickness and emphysema. An expiratory CT acquisition will be performed using a fixed x-ray tube current of 50 mAs to assess for air trapping. As of June 2015, in 3753 subjects scanned for COPDGene Phase 2 using a 200 mAs inspiratory scan and a 50 mAs expiratory scan, the estimated median total effective dose was 8.2 mSv. An additional inspiratory chest CT scan will be performed using a low dose protocol aligned with current clinical practice that utilizes dose modulation. Dose modulation automatically adjusts x-ray exposure for each patient to achieve a consistent level of image quality. The reference mAs for this scan will be set such that an average sized subject (~75 kg) with a normal build would receive an x-ray exposure (CTDIvol) of 3.0 mGy (corresponding roughly to an effective dose of 1.5mSv). The scans will be processed using filtered back projection and iterative reconstruction where feasible. X-ray exposure will automatically be adjusted higher or lower for larger or smaller subjects. The maximum exposure for the dose modulation acquisition will be set at 3.0 mSv. Comparison (quantitatively and qualitatively) of the two inspiratory CT scans will be performed to allow generalization of the study results to the low dose scans being performed clinically nationwide and enable application to clinical practice. Pending IRB approval, all Clinical Centers will perform this additional low dose inspiratory scan (absent local technical or logistical issues).

13. If the chest CT scans are done on the same day as the post-bronchodilator spirometry, the scan should be scheduled as soon as possible after albuterol. The time and date of the last bronchodilator prior to the CT scan will be captured. CT scans will be electronically transmitted to the Imaging Core at National Jewish Health by secure FTP site.

14. A urine pregnancy test will be performed on all premenopausal women prior to the chest CT scan to confirm that they are not pregnant.

15. Medical Record Review: medical records may be reviewed when available and with a signed release of records form documenting the subject's permission to obtain information about lung diseases, cancer, other medical conditions, and to determine cause of death, in the event of death.

Home Visit

If a COPDGene home visit is performed, it will include Home Visit Informed Consent and HIPAA administration, pre and post-bronchodilator spirometry, blood drawing, completion of all study questionnaires and procedures except six minute walk, diffusing capacity, and CT scan.

Visit 1 Study Procedures for Non-Smoker Controls

Non-smoker controls enrolled during the Phase 2 study will be assessed for eligibility using the inclusion and exclusion criteria listed previously for the new non-smoker controls. For the study visit there will be one to two study visits for the majority of subjects depending on whether a separate visit is needed in order to schedule a chest CT scan. In some cases, subjects may have a third visit to re-collect information that does not meet quality control criteria. For example, if spirometry does not meet quality control criteria as judged by the Pulmonary Function Core, the subject may be contacted by phone to schedule an additional visit to repeat the test.

During the Phase 2 research study visit, the following procedures will be performed:

1. Informed consent will be obtained prior to any other study procedure. An updated HIPAA consent to obtain medical records to adjudicate the cause of death will be obtained. Subjects will be asked to provide a signed release of medical records to be used in the event of their death, to obtain an interview with a next of kin, secondary contact, or personal representative and with the treating physician. Hospital and physician records, related to the events and illnesses associated with a participant's death, will also be obtained through the signed release of medical records. Subjects will be provided a memo to give to their physician(s) for this purpose.
2. Subjects will be asked to provide contact information of two secondary contacts with addresses and phone numbers, one of whom is a next of kin or personal representative, and the other a relative or close friend, not living with the subject. The purpose of this information is to locate subjects who have moved to a different address and/or phone number and ascertain vital status of the subject. The subjects will be asked to give consent to have their participation in the COPDGene study disclosed to these secondary contacts in the event that we are unable to contact them, and also disclosed to their next of kin, personal representative or family members in the event of their death. Subjects will be provided a memo to give to their secondary contacts for this purpose.
3. Prior to any other evaluations, an Enrollment/Safety Questionnaire will be administered to determine if potential subjects have a condition that would require deferring their study visit. They will be asked if they have a history of adverse events with the use of albuterol prior to administration of this inhaled bronchodilator. This safety assessment will also be used to ascertain potential cardiac disorders prior to albuterol administration and the walk test. If the questionnaire uncovers known or potential adverse effects of albuterol or potential cardiac disorders, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a six-minute walk test. This questionnaire is located on the protected study web site.
4. Updated contact information will be collected from the subjects. Name, home address, home phone number, cell phone, email address, date of birth and social security number will be collected or verified from the subject. The purpose of this information is to maintain contact with the subjects twice a year and inform the subjects of other research studies and potential future COPDGene follow-up visits pending future funding. In order to perform central collection of mortality and cause of death, this information will be securely transmitted to the Data Coordinating Center at National Jewish Health, where it will be stored separately from the clinical, physiological, imaging, and genetic data. The transmittal of subject's personal information to the DCC will be used for central subject tracking, longitudinal follow-up, death

searching through the Social Security Death Master File (SSDMF) and to obtain cost and utilization information from the Centers for Medicare and Medicaid Services (CMS).

5. Physical Assessment will be performed on all subjects: A limited assessment will be performed including height, weight, arm span, waist circumference, blood pressure, heart rate, and room air oxygen saturation by pulse oximetry with the subject seated at rest. Blood pressure will be measured three times with the subjects seated and resting for at least five minutes. If heart rate is less than 50 or greater than 110 or if the blood pressure is greater than 170/110, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a walk test.

6. Pre- and post-bronchodilator spirometry will be performed on all subjects. The maneuvers are performed in a standardized manner. To obtain 3 acceptable measures, the technician may ask the subject to perform up to 8 attempts. Spirometry is performed before and then repeated approximately twenty minutes following the administration of two puffs (180 mcg) of albuterol from a metered dose inhaler with a spacer. Spirometry will be performed with the subject in a seated position with a nose clip in place.

7. Diffusing capacity will be measured on all subjects after post-bronchodilator spirometry. The test will be performed according to American Thoracic Society/European Respiratory Society standards (8). To obtain 2 reproducible measures, the test may be repeated up to five times.

8. Standardized questionnaires will be completed by all non-smoker control subjects to assess respiratory symptoms, occupational exposures, and other medical history. These questionnaires include a modified version of the American Thoracic Society/Division of Lung Diseases Respiratory Epidemiology Questionnaire to assess respiratory symptoms and St. George's Respiratory Questionnaire to assess health-related quality of life, the Medical Outcomes SF-36 to assess generic quality of life, and the COPD Assessment Test (CAT) to assess COPD severity. Questionnaires will be administered to obtain information about respiratory history, medical history, anxiety and depression (Hospital Anxiety and Depression Scale) (13), menstrual, pregnancy and female hormone medication history in women (Women's Questionnaire), current and previous residential history (modified Residential/Occupational Questionnaire) to ascertain air pollution and environmental exposures, and occupational history. Subjects will be asked about access to regular medical care, access to health insurance and attitudes about medical care for their lung disease (Socioeconomic Questionnaire) to determine the impact of these factors on lung disease and exacerbations. If a participant is found to have cancer, determined by the Medical History Questionnaire, a Cancer Data Collection Questionnaire will be administered during the study visit to collect additional data on the cancer diagnosis. The subject may be invited to sign a tissue release addendum in order to provide formalin fixed paraffin embedded (FFPE) tissue blocks if there was a cancer diagnosis or lung tissue removed. An e-Cigarette questionnaire will be administered. Subjects will answer questions to assess amount of use and attitudes. Each patient will be asked to bring in all of their current medications in their original bottles. A list of current medications will be made at the local clinical center and transmitted to the Data Coordinating Center as a pdf document coded with the subject's study ID number. A clinical coordinator will administer the Questionnaires. The results will be entered directly into the computer or recorded on paper forms to be input later by the coordinator. Completion of all questionnaires is expected to take between 90 and 120 minutes.

9. Blood will be drawn from all subjects for DNA (genetic and epigenetic studies), serum and plasma (for measurement of other proteins potentially related to COPD and other diseases), and a CBC (complete blood count). A total volume of approximately 50 ml of blood will be drawn for this study. A complete blood count will be performed at the Clinical Centers. The rest of the blood samples will be stored at the COPDGene Biological Repository at Brigham and Women's Hospital.
10. Six-Minute Walk Test will be performed in a standardized manner on all subjects to determine exercise capacity (11). Subjects will be asked questions to assure they do not have significant or occult heart disease prior to the test to assure subject safety. Subjects who have significant cardiovascular disease may require evaluation by the local physician investigator prior to six-minute walk test. Immediately and one minute after the end of the six-minute walk test, a subject's heart rate and oxygen saturation will be measured using a pulse oximeter.
11. Chest CT Scans will be performed to assess emphysema and airway disease in all subjects. An inspiratory chest CT scan will be performed with a radiation dose of 200 mAs in order to provide thorough assessment of small airway wall thickness and emphysema. An expiratory CT acquisition will be performed using a fixed x-ray tube current of 50 mAs to assess for air trapping. As of June 2015, in 3753 subjects scanned for COPDGene Phase 2 using a 200 mAs inspiratory scan and a 50 mAs expiratory scan, the estimated median total effective dose was 8.2 mSv. An additional inspiratory chest CT scan will be performed using a low dose protocol aligned with current clinical practice that utilizes dose modulation. Dose modulation automatically adjusts x-ray exposure for each patient to achieve a consistent level of image quality. The reference mAs for this scan will be set such that an average sized subject (~75 kg) with a normal build would receive an x-ray exposure (CTDIvol) of 3.0 mGy (corresponding roughly to an effective dose of 1.5mSv). The scans will be processed using filtered back projection and iterative reconstruction where feasible. X-ray exposure will automatically be adjusted higher or lower for larger or smaller subjects. The maximum exposure for the dose modulation acquisition will be set at 3.0 mSv. Comparison (quantitatively and qualitatively) of the two inspiratory CT scans will be performed to allow generalization of the study results to the low dose scans being performed clinically nationwide and enable application to clinical practice. Pending IRB approval, all Clinical Centers will perform this additional low dose inspiratory scan (absent local technical or logistical issues).
12. If the chest CT scans are done on the same day as the post-bronchodilator spirometry, the scan should be scheduled as soon as possible after albuterol. The time and date of the last bronchodilator prior to the CT scan will be captured. CT scans will be electronically transmitted to the Imaging Core at National Jewish Health by secure FTP site.
13. A urine pregnancy test will be performed on all premenopausal women prior to the chest CT scan to confirm that they are not pregnant.
14. Medical Record Review: medical records may be reviewed when available and with a signed release of records form documenting the subject's permission to obtain information about lung diseases, cancer, other medical conditions, and to determine cause of death, in the event of death.

Appendix C – Phase 3 (Year 10 for Smokers and Year 5 for Nonsmokers Enrolled during Phase 2) – Completed

I. COPDGene Phase 3 Return Visit

Phase 3 obtained 10-year data on 5,783 subjects and visits were conducted on 4,720 subjects, to date.

The primary goals of COPDGene Phase 3 were:

1) To Confirm Prognostication of Clinical Disease Progression in Specific COPD Subtypes
To confirm whether clinical/imaging-based COPD subtypes will show differences in disease progression and mortality, we will perform a ten-year follow-up evaluation with spirometry, DLco, chest CT, and questionnaires from ~4,500 COPDGene subjects. *Hypothesis*: Specific COPD subtypes will show increased progression of emphysema (PSE, High Risk Airway Axis, High Risk Emphysema Axis, and high fSAD); greater loss of FEV_1 (PSE, High Risk Emphysema Axis); and increased mortality (High Risk Airway Axis, Diffuse Emphysema).

2) To Identify Genetic, Transcriptomic, and Proteomic Determinants of Susceptibility and Progression of Specific COPD Subtypes
We will utilize WGS, blood transcriptomics, and plasma proteomics to identify distinct genetic, transcriptomic, and proteomic factors influencing the development and progression of COPD. We will include expression quantitative trait locus (eQTL) and protein QTL (pQTL) analyses to identify networks of interacting biological determinants of COPD disease progression. *Hypothesis*: Protein and mRNA biomarkers, including sRAGE, will identify COPD subjects with progressive emphysema, lung function decline, and increased mortality.

3) To Test High-Risk Subgroups for Rapid COPD Progression
Although included in the original protocol, this aim was later deleted.

Among GOLD 1-2 COPD subjects, we will select 200 subjects from each of three groups identified in Phase 2 as high-risk for rapid FEV_1 decline and/or increased emphysema. Controls will be 200 low-risk subjects for disease progression. Enrolled at their ten-year return visit, each group will return for a Year 11 visit. A SOMAScan protein biomarker panel and peripheral blood RNA-Seq (from Year 10 samples) will be used to perform molecular phenotyping of rapid disease progression. The three high-risk groups are:

a) Extensive PSE. *Hypothesis*: PSE subjects will demonstrate more rapid COPD progression measured by changes in lung function and emphysema.

b) High COPD Disease Progression Risk Score. *Hypothesis*: Random Forests analysis of clinical, physiological, and imaging data identifies subjects who will show rapid progression measured by loss of FEV_1.

c) High Airway Disease Axis Score. *Hypothesis*: High-risk Airway Disease Axis COPD subjects will have increased risk for COPD progression measured by changes in emphysema.

II. SUBJECT SELECTION – Phase 3

Return Visit
A key component of the project is to establish contact with all of the previously enrolled subjects or confirm the subjects are deceased. Clinical Centers will utilize stored contact information including secondary contacts to contact previously enrolled subjects.

Eligible subjects were contacted by mail, email and/or phone and invited to participate in Visit 3 at the Clinical Centers. The study protocol was e discussed in detail during the screening encounter. Before the study visit began, written informed consent was be obtained by one of the study staff members.

Rationale: All previously enrolled subjects from COPDGene are eligible for the study. Previous exclusions for lung cancer and other conditions are not a reason to exclude subjects from follow-up evaluation because these are outcomes of the disease to be assessed.

We will exclude patients with recent use of antibiotics and/or systemic steroids for a COPD exacerbation since such subjects may have an ongoing inflammatory pulmonary process related to an infection that will impair our ability to characterize and phenotype patients in their usual stable state.

We will exclude subjects who have contraindications to spirometry (exclusion criteria 4 – 7 above).

Subjects with exclusion criteria were re-screened at a later date when they could be enrolled in Visit 3. The Administrative Core was to be contacted regarding subjects who could not complete all components of the study visit due to temporary exclusions or other medical conditions. The Administrative Core in consultation with the local clinical center director determined which portions of a study visit can be completed. Those modifications will be considered if the subject is unlikely to be able to return at a later date for the study visit or unlikely to resolve medical issues that limit participation.

Phase 3
All Phase 1 and Phase 2 subjects have provided consent with their initial study enrollment to be contacted again for further research studies. Subjects will be sent a letter advising them that funding has been provided for the second visit and informing them that they will be contacted by letter, email and/or phone to schedule an appointment for Visit 3. A telephone call will be made to each subject to schedule a date for his/her next study visit and ask them to bring the following to the study visit: 1) all their current oral and inhaled medications, 2) names of any injectable study medications received on a regular basis, 3) social security number, 4) name, address and phone number of their personal physician, 5) name, address and phone of two other individuals likely to know the subject's whereabouts, at least one of whom is a relative not living with the subject, and 6) driver's license or other identification to assure they are the individual enrolled in COPDGene Phase 1.

At the Phase 3 study visit the subject will review the informed consent form for the visit and its study procedures. The subjects will have the opportunity to review the study documents in private, have questions answered and sign the informed consent. If necessary, subjects may review the informed consent with family members or their physicians. A HIPAA authorization or combined

consent and HIPAA authorization will be reviewed and signed at the same time. Although subjects have signed a HIPAA authorization in Phase 1, a current authorization will be obtained to meet the requirements of hospitals and health care providers. Phase 3 data collection includes transfer of subject identifiers to the Data Coordinating Center (DCC) as described in the Central Data, Imaging, and Human Subject Contacts Protocol.

Mail Consent

Prior to the return visit, subjects may be contacted by the Clinical Center via mail to request additional consent and authorization. This additional consent and authorization will request that the subject allow transfer of identifiers from the Clinical Center to the Data Coordinating Center at National Jewish Health. This Mail Authorization provides permission for the Clinical Center to transmit the subject's personal information (social security number, address, secondary contact address) to the DCC. A Personal Contact Information Update form will be sent with the mail consent, asking subjects to update contact information, including secondary contacts and primary care physician. Additionally, the subjects will be invited to sign an "Authorization to Release Protected Health Information" medical record release form.

This information will be used for central subject tracking, longitudinal follow up, vital status searching, and to obtain cost and utilization information on subjects from Medicare. The mail consent also allows the Clinical Center to seek perimortem medical records and informant interviews about the circumstances around the subject's death.

Visit 3 Options for Subjects Who Have Difficulty with a Visit to Clinical Center

In the event that a subject has moved away from their local Clinical Center and to an area closer to another participating COPDGene Clinical Center, the subject may elect to complete Visit 3 at this closer Center rather than traveling to their original Center. With the subject's verbal approval by phone, their new contact information will be provided from the Clinical Center where the subject had the first visit to the geographically proximate Clinical Center. The subject will sign a new Informed Consent Form and HIPAA with the coordinator at the new study location prior to their Phase 3 visit.

In the event that a subject has financial hardship or has moved a considerable distance from their local Clinical Center, the COPDGene Administrative Core may consider providing transportation assistance to this subject for a return visit. The payment for this subject's return visit may include the actual reasonable cost of transportation, lodging, and meal expenses. Subject must provide receipts to Clinical Center for reimbursement. Use of this approach and funding for these expenses will be allocated on a case-by-case basis from the Administrative Core to the Clinical Center.

COPDGene subjects who are unable to come to a COPDGene Clinical Center to complete a Phase 3 visit due to illness, disease progression, or inability to be transported to the Clinical Center, will be given the option of having a research coordinator complete a partial study visit at the subject's place of residence. This home visit will include spirometry before and after albuterol administration, blood drawing, and completion of all study questionnaires.

Phase 3 subjects who are unable to come to a Clinical Center and are not willing or able to undergo a home visit will be asked if they would mind completing study questionnaires over the phone. If the subject agrees, this will be completed. If the subject is not willing or unable to complete study

questionnaires over the phone they will be ask if they would complete a Limited Follow-Up Questionnaire and some or all of the regular study questionnaires. The Limited Follow-Up Questionnaire will be used to gather important longitudinal information on the subject related to respiratory symptoms, exacerbations, and comorbidities. A research coordinator will conduct this questionnaire with the subject by phone. These subjects may still participate in the bi-annual Longitudinal Follow-Up program by phone or Internet.

Subjects who do not agree to a Phase 3 visit or to participate in Longitudinal Follow-Up will be asked if they would be willing to complete a Limited Follow-Up Questionnaire by phone. If the subject refuses to complete this questionnaire, their wishes will be respected.

Closed Clinical Center
Due to the longitudinal nature of the study, it is important to follow health and vital status on all of the COPDGene subjects. In the event that a clinical center is unable to perform follow up contacts or visits, the COPDGene Administrative Core will work as an agent for the Clinical Center or provide a contract clinical research organization to manage the follow-up subject visits and contacts per protocol. Subjects will be consented in Phase 3 to allow for a member of COPDGene Central Administration to contact them for follow up. The original clinical center must work with the Administrative Core to facilitate efficient transfer of information and study management under the supervision of a local principal investigator and with the approval of the local IRB.

III. PHASE 3 STUDY PROCEDURES

Similar to the first visit, for the Phase 3 study visit there will be one to two visits for the majority of subjects depending on whether a separate visit is needed to schedule a chest CT scan. In some cases, subjects may have a third visit to re-collect information that does not meet quality control criteria. For example, if spirometry does not meet quality control criteria as judged by the Pulmonary Function Core, the subject may be contacted by phone to schedule an additional visit to repeat the test.

During the Phase 3 research study visit, the following procedures will be performed:

1. Informed consent will be obtained prior to any other study procedure. An updated HIPAA consent to obtain medical records to adjudicate the cause of death will be obtained. Subjects will be asked to provide a signed release of medical records to be used in the event of their death, to obtain an interview with a next of kin, secondary contact, or personal representative and with the treating physician. Hospital and physician records, related to the events and illnesses associated with a participant's death, will also be obtained through the signed release of medical records. Subjects will be provided a memo to give to their physician(s) for this purpose.
2. Subjects will be asked to provide or update contact information of two secondary contacts with addresses and phone numbers, one of who is a next of kin or personal representative, and the other a relative or close friend, not living with the subject. The purpose of this information is to locate subjects who have moved to a different home and have a different address and/or phone number and ascertain vital status of the subject. The subjects will be asked to give consent to have their participation in the COPDGene study disclosed to these secondary

contacts in the event that we are unable to contact them, and also disclosed to their next of kin, personal representative or family members in the event of their death. Subjects will be provided a memo to give to their secondary contacts for this purpose.

3. Prior to any other evaluations, an Enrollment/Safety Questionnaire will be administered to determine if potential subjects have a condition that would require deferring their study visit. They will be asked if they have a history of adverse events with the use of albuterol prior to administration of this inhaled bronchodilator. This safety assessment will also be used to ascertain potential cardiac disorders prior to albuterol administration and the walk test. If the questionnaire uncovers known or potential adverse effects of albuterol or potential cardiac disorders, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a six-minute walk test. This questionnaire is located on the protected study web site.
4. Updated contact information will be collected from the subjects. Name, home address, home phone number, cell phone, email address, date of birth and social security number will be collected or verified from the subject. The purpose of this information is to maintain contact with the subjects twice a year, and inform the subjects of other research studies and potential future COPDGene follow-up visits pending future funding. In order to perform central collection of mortality and cause of death, this information will be securely transmitted to the Data Coordinating Center at National Jewish Health, where it will be stored separately from the clinical, physiological, imaging, and genetic data. The transmittal of subjects' personal information to the DCC will be used for central subject tracking, longitudinal follow-up, death searching through the Social Security Death Master File (SSDMF) and to obtain cost and utilization information from the Centers for Medicare and Medicaid Services (CMS).
5. Physical Assessment will be performed on all subjects: A limited assessment will be performed including height, weight, arm span, waist circumference, blood pressure, heart rate, and room air oxygen saturation by pulse oximetry with the subject seated at rest. Blood pressure will be measured three times with the subjects seated and resting for at least five minutes. In subjects using supplemental oxygen, the oxygen will be withheld and subjects will breathe room air for 10 minutes prior to recording oxygen saturation. If resting oxygen saturation falls below 82% or the subject becomes dyspneic prior to 10 minutes without oxygen, oxygen therapy will be restarted as a safety precaution. If heart rate is less than 50 or greater than 110 or if the blood pressure is greater than 170/110, a study physician will evaluate the subject to determine if it is safe to administer albuterol and perform a walk test.
6. Pre and post-bronchodilator spirometry will be performed on all subjects. The maneuvers are performed in a standardized manner. To obtain 3 acceptable measures, the technician may ask the subject to perform up to 8 attempts. Spirometry is performed before and then repeated approximately twenty minutes following the administration of two puffs (180 mcg) of albuterol from a metered dose inhaler with a spacer. Spirometry will be performed with the subject in a seated position with a nose clip in place.
7. Diffusing capacity will be measured on all subjects after post-bronchodilator spirometry. The test will be performed according to American Thoracic Society/European Respiratory Society standards (8). In subjects using supplemental oxygen, oxygen will be removed and subjects will breathe room air for 10 minutes prior to the test. If resting saturation falls below 82% or the subject becomes dyspneic prior to 10 minutes without oxygen, oxygen therapy will be

restarted as a safety precaution and the test will be performed without the ten-minute oxygen washout. To obtain 2 reproducible measures, the test may be repeated up to five times.

8. Standardized questionnaires will be completed by all subjects to assess respiratory symptoms, smoking history, and other medical history. These questionnaires include the St. George's Respiratory Questionnaire to assess health-related quality of life, the Medical Outcomes SF-36 to assess generic quality of life, and the COPD Assessment Test (CAT) to assess COPD severity. Questionnaires will be administered to obtain information about respiratory history, medical history, anxiety and depression (Hospital Anxiety and Depression Scale) (9), menstrual, pregnancy and female hormone medication history in women (Women's Questionnaire), current and previous residential history (Residential/Occupational Questionnaire) to ascertain air pollution and environmental exposures, and occupational history. Subjects will be asked about access to regular medical care, access to health insurance and attitudes about medical care for their lung disease (Socioeconomic Questionnaire) to determine the impact of these factors on lung disease and exacerbations. If a participant is found to have cancer, determined by the Medical History Questionnaire, a Cancer Data Collection Questionnaire will be administered during the study visit to collect additional data on the cancer diagnosis. The subject may be invited to sign a tissue release addendum in order to provide formalin fixed paraffin embedded (FFPE) tissue blocks if there was a cancer diagnosis or lung tissue removed. An e-Cigarette questionnaire will be administered. Subjects will answer questions to assess amount of use and attitudes.
9. Each patient will be asked to bring in all of their current medications in their original bottles. A list of current medications will be made at the local clinical center and transmitted to the Data Coordinating Center as a pdf document coded with the subject's study ID number. Oxygen use will also be recorded. A clinical coordinator will administer the Questionnaires. The results will be entered directly into the computer or recorded on paper forms to be input later by the coordinator. Completion of all questionnaires is expected to take between 90 and 120 minutes.
10. Blood will be drawn from all subjects for DNA (genetic and epigenetic studies), serum and plasma (for measurement of other proteins potentially related to COPD and other diseases) and a CBC (complete blood count). A total volume of approximately 50 ml of blood will be drawn for this study. A complete blood count will be performed at the Clinical Centers. The rest of the blood samples will be stored at the COPDGene Biological Repository at Brigham and Women's Hospital.
11. Six-Minute Walk Test will be performed in a standardized manner on all subjects to determine exercise capacity (11). Subjects will be asked questions to assure they do not have significant or occult heart disease prior to the test to assure subject safety. Subjects who have significant cardiovascular disease may require evaluation by the local physician investigator prior to six-minute walk test. Immediately and one minute after the end of the six-minute walk test, a subject's heart rate and oxygen saturation will be measured using a pulse oximeter.
12. Chest CT Scans will be performed to assess emphysema and airway disease in all subjects. An inspiratory chest CT scan will be performed with a radiation dose of 200 mAs in order to provide thorough assessment of small airway wall thickness and emphysema. An expiratory CT acquisition will be performed using a fixed x-ray tube current of 50 mAs to assess for air trapping. As of June 2015, in 3753 subjects scanned for COPDGene Phase 2 using a 200 mAs

inspiratory scan and a 50 mAs expiratory scan, the estimated median total effective dose was 8.2 mSv. An additional inspiratory chest CT scan will be performed using a low dose protocol aligned with current clinical practice that utilizes dose modulation. Dose modulation automatically adjusts x-ray exposure for each patient to achieve a consistent level of image quality. The reference mAs for this scan will be set such that an average sized subject (~75 kg) with a normal build would receive an x-ray exposure (CTDIvol) of 3.0 mGy (corresponding roughly to an effective dose of 1.5mSv). The scans will be processed using filtered back projection and iterative reconstruction where feasible. X-ray exposure will automatically be adjusted higher or lower for larger or smaller subjects. The maximum exposure for the dose modulation acquisition will be set at 3.0 mSv. Comparison (quantitatively and qualitatively) of the two inspiratory CT scans will be performed to allow generalization of the study results to the low dose scans being performed clinically nationwide and enable application to clinical practice. Pending IRB approval, all Clinical Centers will perform this additional low dose inspiratory scan (absent local technical or logistical issues).

13. If the chest CT scans are done on the same day as the post-bronchodilator spirometry, the scan should be scheduled as soon as possible after albuterol. The time and date of the last bronchodilator prior to the CT scan will be captured. CT scans will be electronically transmitted to the Imaging Core at National Jewish Health by secure FTP site.
14. A urine pregnancy test will be performed on all premenopausal women prior to the chest CT scan to confirm that they are not pregnant.
15. Medical Record Review: medical records may be reviewed when available and with a signed release of records form documenting the subject's permission to obtain information about lung diseases, cancer, other medical conditions, and to determine cause of death, in the event of death.