Automated Telecommunication to Obtain Longitudinal Follow-up in a Multicenter Cross-sectional COPD Study

Jeffrey I. Stewart[1], **Sarah Moyle**[2], **Gerard J. Criner**[1], **Carla Wilson**[3], **Ron Tanner**[4], **Russell P. Bowler**[3], **James D. Crapo**[3], **Robert K. Zeldin**[5], **Barry J. Make**[3], **Elizabeth A. Regan**[3], and **the COPDGene Investigators**

[1] Pulmonary and Critical Care Medicine, Department of Medicine, Temple University School of Medicine, Philadelphia, PA, USA. [2] University of Bern, Bern, Switzerland. [3] National Jewish Health, Denver, Colorado, USA. [4] Inventive Labs Corporation, Greenwood Village, Colorado, USA. [5] Novartis Pharmaceuticals Corporation East Hanover, New Jersey, USA.

Abstract

Background—It can be challenging to maintain longitudinal follow-up of subjects in clinical studies. COPDGene is a multicenter, observational study designed to identify genetic factors associated with COPD and to characterize COPD-related phenotypes. To obtain follow-up data on patient's vital status and outcomes, the COPDGene Longitudinal Follow-up (LFU) Program was developed to supplement its parent study.

Methods/Results—We used a telecommunication system that employed automated telephone contact or web-based questions to obtain longitudinal follow-up data in our subjects. A branching questionnaire asked about exacerbations, new therapies, smoking status, development of co-morbid conditions, and general health status. Study coordinators contacted subjects who did not respond to one of the automated methods. We enrolled 10,383 subjects in the COPDGene study. As of August 29, 2011, 7,959 subjects completed 19,955 surveys. On the first survey, 68.8% of subjects who completed their survey did so by electronic means, while 31.3% required coordinator phone follow-up. On each subsequent survey the number of subjects who completed their survey by electronic means increased, while the number of subjects who required coordinator follow-up decreased. Despite many of the patients in the cohort being chronically ill and elderly, there was broad acceptance of the system with over half the cohort using electronic response methods.

Conclusions—The COPDGene LFU Study demonstrated that telecommunications was an effective way to obtain longitudinal follow-up of subjects in a large multicenter study. Web-based and automated phone contacts are accepted by research subjects and could serve as a model for LFU in future studies.

Correspondence to: Jeffrey I. Stewart, M. D., Temple Lung Center, 745 Parkinson Pavilion, 3401 North Broad Street, Philadelphia Pa, 19140, phone: 215-707-5864, fax: 215-707-6867, jeffrey.stewart@tuhs.temple.edu.

Introduction

Chronic obstructive pulmonary disease (COPD) is a major cause of morbidity and mortality in the United States (US), and is the third leading cause of death (1). Patients who have acute COPD exacerbations experience a more rapid decline in lung function as well as worse quality of life (2,3). Occurrences of co-morbid illnesses, such as cancer and cardiovascular disease, negatively impact patient outcomes (4). Strategies that decrease the frequency of COPD exacerbations, and recognize and treat the co-morbid diseases associated with COPD, will likely improve morbidity and mortality.

To better understand natural history and disease progression in COPD, it is essential to conduct longitudinal follow-up (LFU) in research subjects. When LFU is conducted solely by clinical coordinators, it is time consuming and expensive. Furthermore, when the goal is to obtain repeated surveys at various intervals, delays in subject contact can lead to missing data points or erroneous data secondary to inaccurate subject recall of prior events.

Asynchronous, bi-directional communication allows two parties to transmit information to each other in an intermittent fashion and include automated phone calls and web-based Internet surveys. Both of these techniques are being used in health care delivery, the management of chronic medical conditions, as well as in clinical research (5,6). When used as part of clinical research, these forms of communication are quicker and more efficient than personal calls conducted by study coordinators, and allow subjects to respond in a time frame and manner that is convenient for them. Questions can be asked in a consistent scripted manner that facilitates standardization and reduces transcription errors. However, the impersonal nature of the contact and the requirement for subjects to use advanced technology, especially elderly patients or those naïve to using new forms of electronic communication, may limit its effectiveness. Herein, we describe the use of such technology in the Genetic Epidemiology of COPD (COPD-Gene) Study, a large, multicenter, cross-sectional study of COPD that predominately enrolled subjects middle aged or older.

Methods

COPD Gene

The COPDGene Study is a multicenter, prospective, observational study that enrolled 10,383 subjects at 21 (academic) clinical centers across the United States. It was designed to identify genetic factors associated with COPD, and to characterize COPD-related phenotypes. Details regarding the study design have previously been published (7). Institutional review board (IRB) approval was obtained at each clinical center, and all subjects provided informed consent. This study recruited current and former cigarette smokers with and without COPD who were self-identified African-American and non-Hispanic white. Subjects were 45 to 80 years of age with at least 10 pack-years of smoking history. Subjects completed a modified American Thoracic Society (ATS) Respiratory Epidemiology questionnaire and a St. George's Respiratory Questionnaire. They also completed a 6-minute walk test and pre- and post- bronchodilator spirometry according to ATS guidelines. Volumetric chest computed tomography (CT) scans were completed on all subjects (7).

Study Design

A LFU strategy was developed to supplement the COPDGene baseline data collection and was adopted by 19 of the 21 clinical centers. It was conducted using a combination of three methods: telephone contact (telephony), web-based questions (web survey), and phone calls via the coordinator (manual follow-up). Written, informed consent was obtained from all subjects. Subjects consented to be contacted every 3–6 months. If a subject wanted to be contacted by a coordinator follow-up phone call only, this was noted.

The telephony system along with the web survey was developed in collaboration with Inventive Labs Corporation (Greenwood Village, CO). An automated calling system using pre-recorded outgoing messages that permitted both voice responses and keypad responses was selected. A branching questionnaire asked about exacerbations and treatments, new chronic therapies, current smoking status, other new co-morbid conditions and general health status (Fig. 1). The web-survey asked identical questions, in the same manner. If subjects failed to respond to the telephone or web survey within a week, a local study coordinator attempted to contact them. The study coordinators asked subjects the same questions by using a written script, and recorded the responses on a web site.

Operation

Each study center maintained a master file which included the following: subject identification number (SID), first and last name, primary phone number, secondary phone number (if available), e-mail address (if available), the date the first study visit was completed, which version of the informed consent they signed (allowing 3- versus 6-month follow-up), an optional e-mail password for accessing the web-survey, and an option for manual contact only. If a subject's information changed, they had withdrawn from the study, or were found to be deceased; the clinical coordinator updated the master file. The patient's identifiable information was only known to the individual centers, and not the data-coordinating center (DCC) of the COPDGene study.

Each month the clinical coordinator for each center uploaded the master file to the telecommunication system via the LFU website. The system recognized which subjects were due to be contacted that month based on the date of the initial study visit and the type of consent that was signed (3- versus 6-month follow-up). If the subject provided an e-mail address, the system sent an automated e-mail to the subject inviting them to go to the LFU website to complete their survey. The e-mail was customized to come from the individual study center, and instructed subjects on how to login via the secure website. If a subject did not complete the web-survey within seven days, they were contacted by the telephony system. Those subjects who did not provide an e-mail address were automatically routed to the telephony system (Fig. 2).

Each time a subject was called, they had the option to complete the survey, schedule a call at a more convenient time, or request a call from a clinical coordinator. The telephony system made 3 call attempts per phone number. Each call was automatically scheduled to occur at a different time of day to optimize the likelihood of contact. After 3 attempts to the primary phone number, 3 more attempts were automatically made to the secondary phone number (provided by the subject at the time of study enrollment). Subjects could also call the telephony server in which case the subject was recognized by the system by their primary or secondary phone number, or by the subject entering their telephone number after a machine prompt. If the subject did not complete the survey by the web-survey or telephony, their SID was automatically routed to the clinical coordinator task list on the LFU website. The clinical coordinator would then make 3 successive calls (on separate days) to the primary

number, and then 3 calls to the subject's secondary contact until contact was made. If no contact was made they were recorded as a failed contact for that cycle.

Data Transmission and Security

The system complied with all institutional standards for maintaining the security and confidentiality of patient data. All server access and databases were password protected, with periodic checks to guard against hacking. Transmission of subject information from the clinical centers was secured through password protection using unique logins specific to each study site. All identifying data was temporarily stored in a separate encrypted database, and the data was purged from the telephony server on completion of a survey, after attempts to contact the subject had been exhausted, or three weeks after the initial upload to the system.

Each completed survey was assigned a unique survey ID, which was mapped to the SID, and stored separately with no personal identifying information. Clinical coordinators had access to de-identified data reports of only their subjects, via secure login to the LFU website. The central LFU manager had access to all site deidentified data reports. The patient specific survey results coded by SID were exported to files, which the COPDGene DCC could download via a custom secure transfer program, to be integrated into the main COPDGene study database.

Implementation

An initial training program was developed to ensure data quality across the study centers. This included training in subject data collection and data entry. In addition to the initial in-person training, new coordinators were individually trained by the central LFU manager. Prior to roll out, all subjects received a letter from the clinical center explaining in detail the process of LFU contact. Roll out of the longitudinal follow-up program was done incrementally and began when subjects were about 1 year removed from their initial study visit. As study enrollment proceeded, initial LFU contact for subsequently enrolled subjects was made 6 months after enrollment.

LFU System Performance

Enrollment for COPDGene began in March 2008. A total of 10,383 subjects were enrolled in the study. The mean age for all subjects was 59.6 years (standard deviation 9.0). 3,657 (35.8%) subjects had GOLD stage 2 or greater (Table 1).

As of August 29, 2011, 7,959 subjects completed 19,955 surveys. Subjects who were enrolled in COPDGene during the first year (2008) were contacted for their initial follow-up survey about 1 year later, and every 6 months thereafter. These subjects completed up to 4-5 cycles of follow-up by August 2011. COPDGene completed enrollment in April 2011. Subjects who were enrolled later completed fewer cycles of follow-up. Therefore, the greatest amount of follow-up data exists for survey 1. Subjects who did not complete the first follow-up survey included: 109 subjects who were enrolled in one of the 2 centers that did not participate in the LFU study, 421 subjects who were contacted but declined to complete the survey, and 758 subjects who could not be contacted due to a disconnected phone number on file. 70 subjects were found to be deceased at time of contact.

On the first survey, 18.6% of subjects responded by website, 50.1% by telephony and 31.3% by coordinator follow-up. On each subsequent survey the number of subjects who completed their survey by electronic means (website or telephony) increased, while the number of subjects who required coordinator follow-up decreased. By the subject's fourth survey, 45.6% of subjects responded by website, 47.5% by telephony and only 10.0% required

coordinator follow-up. Our data from the fifth survey shows this trend continues, although the number of subjects who have completed this survey is small at this point (Fig. 3).

Automated calls were completed in less than 5 minutes for a majority of the subjects. For the first survey, contact success was higher among non-smoker controls, followed by those with COPD (GOLD 1–4), followed by smoker controls (94.4% vs. 82.6% vs. 75.8%). The contact success increased with age (<55 = 65.0%, 55–64 = 82.0%, 65 = 92%), and was higher among female subjects (84% vs. 75%). There was also a greater response rate among subjects with a higher degree of education (Fig. 4). There was a decrease in contact success among African-American subjects compared to non-Hispanic Whites (60.6% vs. 87.4%), and a larger proportion of African-American subjects had phone numbers in the system that were no longer valid (15.0% vs. 4.1%). On initial follow-up, African-Americans were less likely to respond by web-survey (Fig. 5).

Costs

Costs for obtaining ongoing information from study participants by this method included expenditures for coordinator time, central management of the system within COPDGene, development costs, and the per subject costs for hosting the system and phone usage. In 2011, the cost for obtaining follow up reports on 80% of the 10,383-subject cohort was just over $43,000, exclusive of development costs and administrative time to oversee the program. This breaks down to approximately $3.93 per contact using the automated system. This can be compared to an estimate of more than $200,000 for coordinator time to make all phone calls manually ($20.00 per subject). Therefore, use of the LFU system represents an 80% cost reduction compared to the more traditional method of follow-up. Coordinators were left with a smaller proportion of calls to be completed locally and with responsibility for local center uploading of the subject database.

Discussion

We used an asynchronous, bi-directional telecommunication system to obtain longitudinal follow-up in a large multicenter cohort. Using this system, follow-up data was obtained with reduced effort required by the study staff to conduct phone follow-ups. If necessary, the frequency of data collection or its content could be easily adjusted without the requirement for extensive new programming.

Our cohort consisted of many subjects who were chronically ill and elderly. Despite this, there was acceptance of the telecommunication system among the participants in the cohort, regardless of their age and severity of COPD. In fact, there was a higher response rate in COPD (compared to smoker controls) and in older subjects. Furthermore, once subjects became familiar with the system, they were more inclined to use one of the automated follow-up methods. This was evident by the increased number of subjects who used the web or telephony, and decreased use of manual calling at subsequent contacts. There was a decrease in contact success by all methods in African-American subjects, as compared to non-Hispanic Whites. A higher percentage of African-Americans were lost to follow-up because the phone numbers they provided were no longer valid.

Furthermore, African-Americans were less likely to respond by web-survey, which highlights the need to give subjects the option to respond to surveys by both web and telephony systems. The reduced need for coordinator phone calls in subsequent surveys was a result of the increased use of one of the automated methods of follow-up, as well as the elimination of attempts to contact subjects who were deceased, or were unable to be contacted during a prior survey because of a bad telephone number.

Limitations to the LFU system are related to the impersonal nature of the web survey and telephony system. If subjects chose to use one of the automated systems from the initiation of their involvement in the study, they would not speak to a study coordinator for the duration of the study unless further information from them is needed. To combat the impersonal nature of the system, all communication with the subject originated from the subject's own clinical site. The telecommunication system used the subject's name in all communications. The local clinical center managed the upload of LFU subject lists and was a resource for subject questions. The individuals also had the choice, at every contact, to change their preferred method of communication and provide their responses directly to a study coordinator. In spite of these limitations we found subjects willing to use the less personal communication system. This may be due to increased familiarity with using electronic communications that have extensively penetrated into the general public's daily life and the efficiency of the system. An additional limitation relating to study management was that revising survey questions required IRB re-approval.

With the growing use of the Internet to obtain and transmit data, especially protected health information, there was a major concern to maintain the subject's privacy. For that reason, strict protocols were followed to protect subject data and retain separation of research data from personal identifiers, such that the data stored in the DCC contained only the responses from the surveys, and not the uploaded contact information. Another potential limitation is the accuracy of data. If a subject who was completing a survey via one of the automated methods did not understand a question, accidently input the wrong answer, or the telephony system mischaracterized a subject's response, the data would be inaccurate. Furthermore, we have not yet assessed the reproducibility of data collected using telephony nor have we collected formal information about satisfaction with the survey process.

Payments were made to clinical centers to compensate them for calls completed by coordinators for those subjects who did not provide information by automated methods. The per-contact cost was reduced about 20-fold for subjects who reported their data electronically. While the initial costs for development of the telecommunications system were sizable, the per subject costs for hosting the system and phone calls were modest. Because a large proportion of the subjects reported via one of the electronic means rather than by coordinator, there were substantial cost savings over the duration of the project.

Conclusion

We developed a web and telephone based telecommunication system to obtain longitudinal follow-up at 3- or 6-month intervals in a large multicenter study of COPD subjects. Our cohort included many subjects who were chronically ill and elderly. Despite this, there was acceptance of both the web-based and telephony systems, which improved with subsequent survey contact. The development of this system could serve as a model for LFU in future studies.

Acknowledgments

References

1. Miniño, AM.; Xu, JQ.; Kochanek, KD. Deaths: Preliminary data for 2008. National Vital Statistics Reports. Vol. 59. National Center for Health Statistics; Hyattsville, MD: 2010.

2. Donaldson GC, Seemungal TA, Bhowmik A, Wedzicha JA. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. Thorax. 2002; 57(10):847–852. [PubMed: 12324669]

3. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 1998; 157(5 Pt 1):1418–1422. [PubMed: 9603117]

4. Sin DD, Anthonisen NR, Soriano JB, Agusti AG. Mortality in COPD: Role of co-morbidities. Euro Respir J. 2006; 28(6):1245–1257.

5. Abu-Hasaballah K, James A, Aseltine RH Jr. Lessons and pitfalls of interactive voice response in medical research. Contemp Clin Trials. 2007; 28(5):593–602. [PubMed: 17400520]

6. Biem JH, Turnell RW, D'Arcy C. Computer telephony: Automated calls for medical care. Clin Investig Med. 2003; 26(5):259–268. http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=11016331&site=ehost-live&scope=site. [PubMed: 14596488]

7. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. COPD. 2010; 7(1):32–43. [PubMed: 20214461]
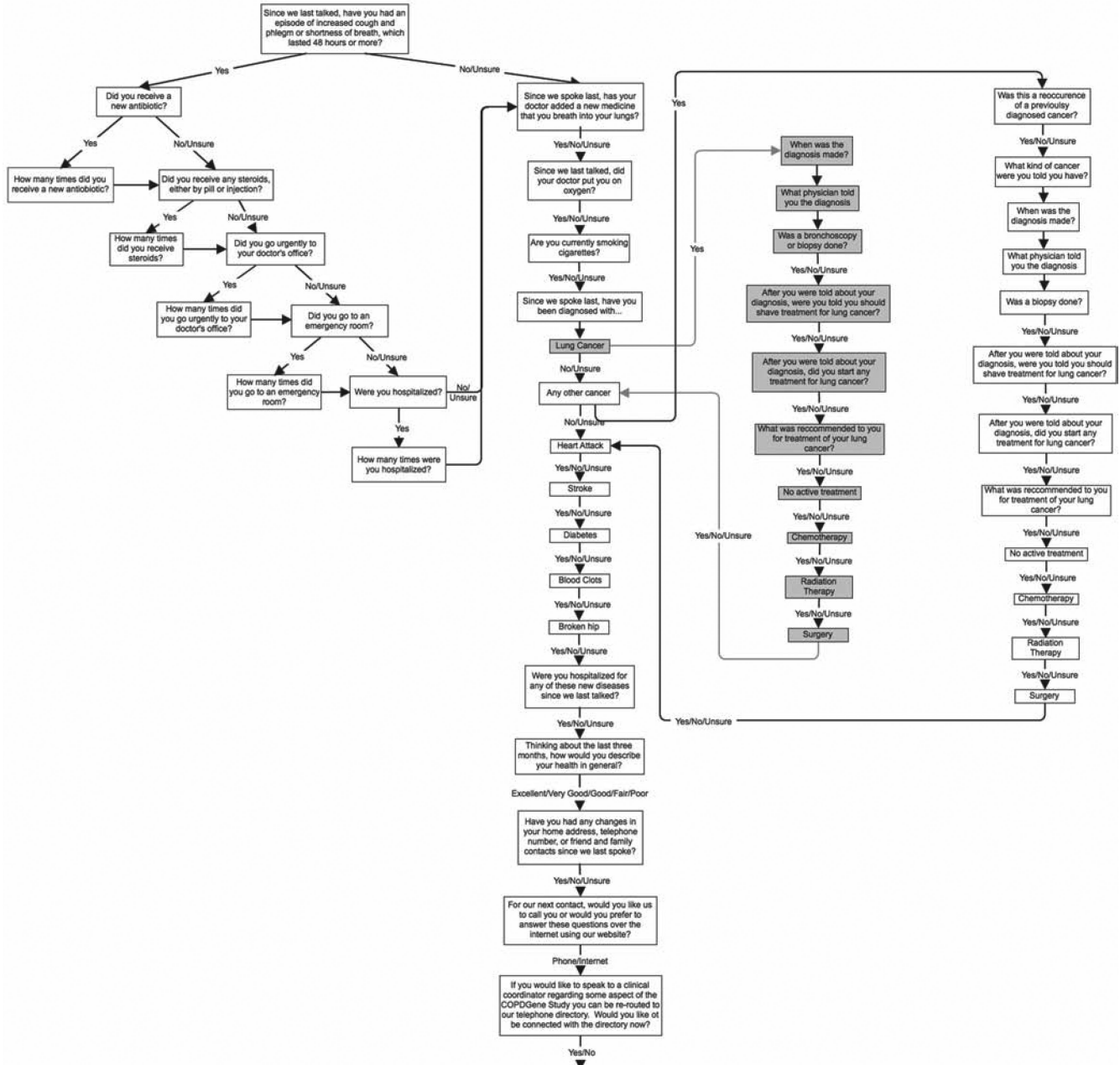
Figure 1.
Branching questionnaire with all questions asked to COPDGene study subjects via telephony system, web-survey or manually by clinical coordinators.
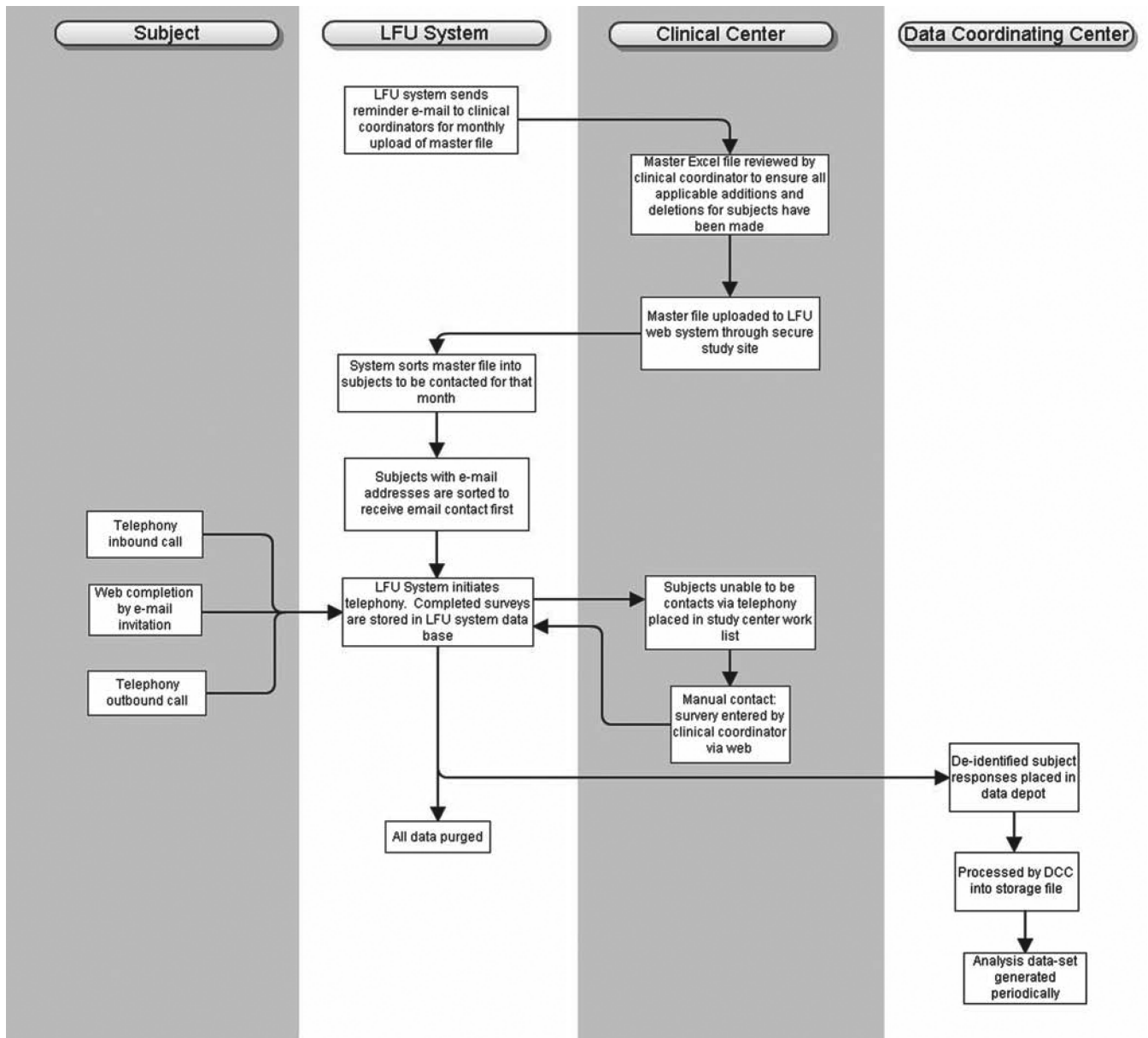
Figure 2.
Flow diagram highlighting transmission of data from between COPDGene subject, LFU
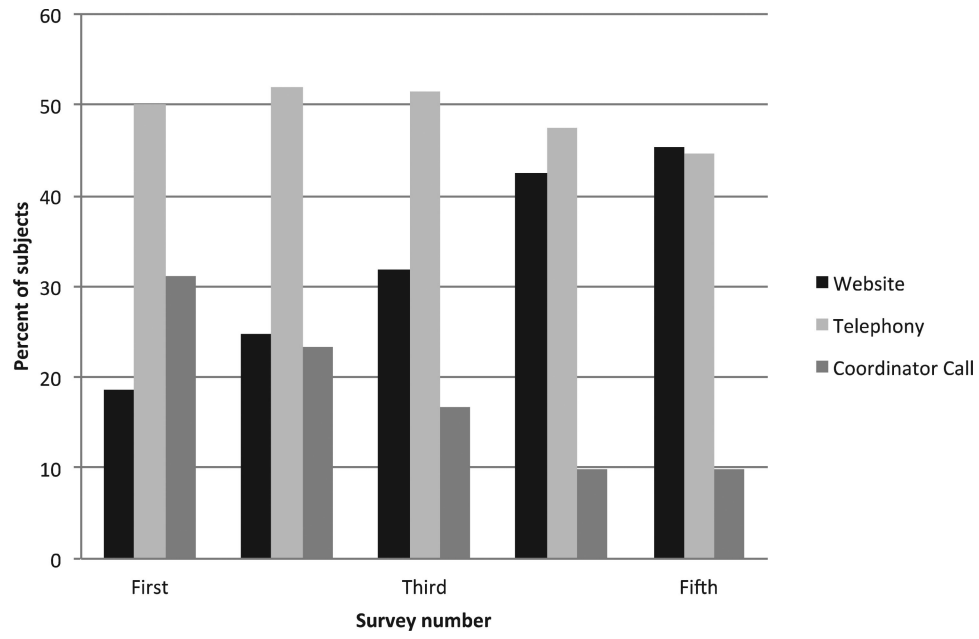system, clinical center, and data coordinating center.

Figure 3.
Percent of COPDGene subjects who responded by website, telephony, and coordinator call at each survey. Over time, patients were most likely to respond by web site.
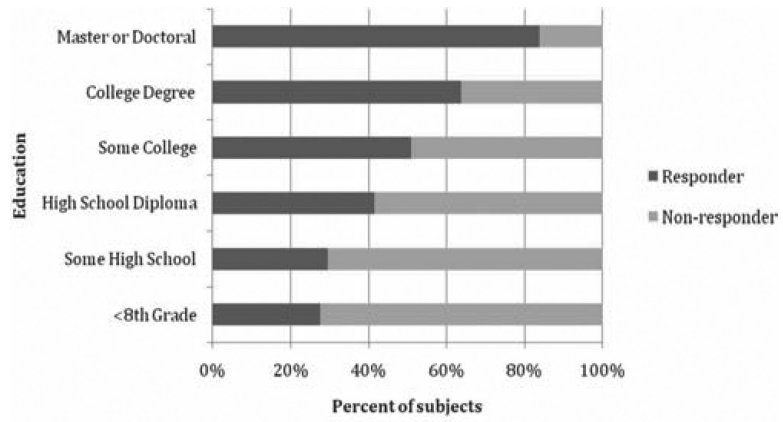
Figure 4.
Percent of subjects responding to survey 1 stratifi ed by education level. There was an increase in response rate among subjects with a higher education level.
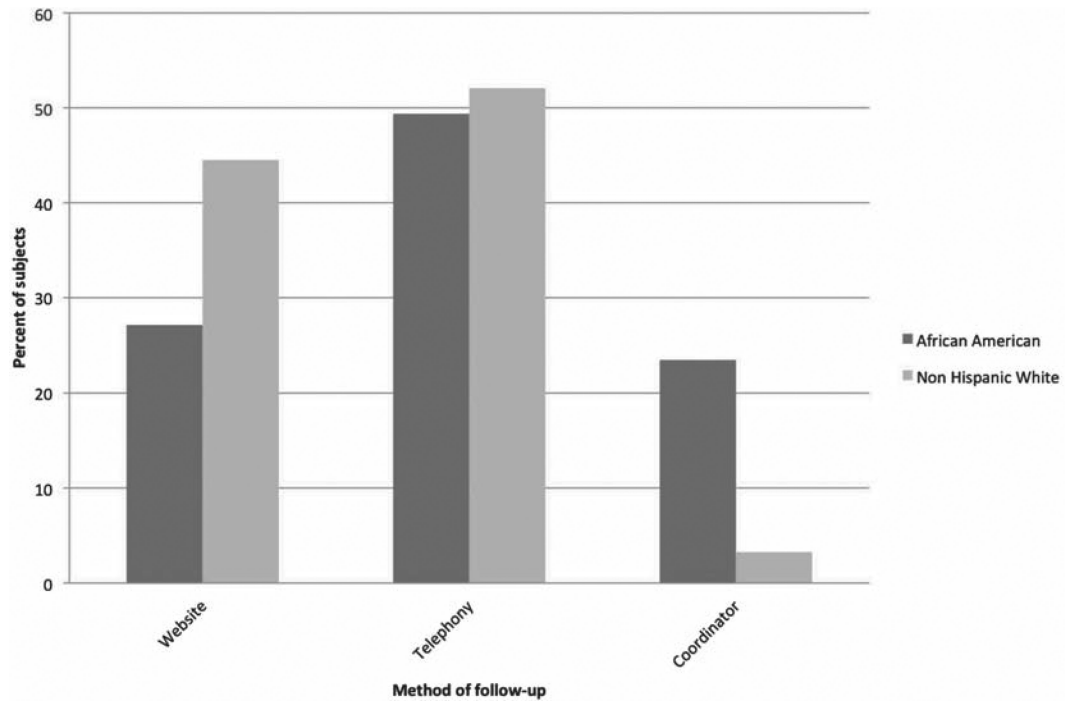
Figure 5.
Percent of subjects responding to survey 1 stratifi ed by method of follow-up and race.
African American's were less likely to respond by web-survey.

Table 1

Demographics of COPDGene Cohort by GOLD Stage

	Total Cohort	Non-Smoker Controls	Smoker Controls	GOLD Unclassified	GOLD 1	GOLD 2	GOLD 3	GOLD 4
N	10220[*]	108	4433	1204	818	1919	1153	585
Age, yr	59.6 ± 9.0	62.1 ± 9.2	56.6 ± 8.4	57.1 ± 8.3	61.7 ± 8.9	62.6 ± 8.9	64.4 ± 8.3	63.9 ± 7.6
Male, %	53.2	31.5	52.8	46.3	57.7	53.1	57.8	58.8
African American, %	33.1	7.4	41.1	43.9	23.1	25.1	20.8	19.5

Data are presented as mean ± SD or as a percentage of the group.

Non-smoker and smoker control subjects have a post-bronchodilator and FEV_1/FVC 0.7 and FEV_1 80% predicted.

GOLD Unclassified subjects have a post-bronchodilator FEV_1/FVC 0.7 and $FEV_1 < 80\%$ predicted.

[*] Missing from total - 59 subjects who did not have final group assignment due to failed spirometry.