Chronic Obstructive Pulmonary Diseases:
Journal of the COPD Foundation®

Original Research

Factors Associated with Attrition in a Longitudinal Cohort of Older Adults in the Community

Noah Katsuno, BA[1] Pei Z. Li, PhD[2] Jean Bourbeau, MD[2] Shawn Aaron, MD[3] Francois Maltais, MD[4] Paul Hernandez, MD[5] Kenneth R. Chapman, MD[6] Brandie Walker, MD[7] Darcy D. Marciniuk, MD[8] Denis D. ODonnell, MD[9] Don D. Sin, MD[1] James C. Hogg, MD[1] Michael Cheng, MBBS[1] Jeremy Road, MD[10] Wan C. Tan, MD[1] on behalf of the CanCOLD Collaborative Research Group*

Abstract

Introduction: Retaining participants in longitudinal studies increases their power. We undertook this study in a population-based longitudinal cohort of adults with COPD to determine the factors associated with increased cohort attrition.

Methods: In the longitudinal population-based Canadian Cohort of Obstructive Lung Disease (CanCOLD) study, 1561 adults >40 years old were randomly recruited from 9 urban sites. Participants completed in-person visits at 18-month intervals and also were followed up every 3 months over the phone or by email. The cohort retention for the study and the reasons for attrition were analyzed. Hazard ratios and robust standard errors were calculated using Cox regression methods to explore the associations between participants who remained in the study and those who did not.

Results: The median follow-up (years) of the study is 9.0 years. The overall mean retention was 77%. Reasons for attrition (23%) were: dropout by participant (39%), loss of contact (27%), investigator-initiated withdrawal (15%), deaths (9%), serious disease (9%), and relocation (2%). Factors independently associated with attrition were lower educational attainment, higher pack-year tobacco consumption, diagnosed cardiovascular disease, and a higher Hospital Anxiety and Depression Scale score: adjusted hazard ratios (95% confidence interval) were 1.43 (1.11, 1.85); 1.01 (1.00, 1.01); 1.44 (1.13, 1.83); 1.06 (1.02, 1.10) respectively.

Conclusion: Identification and awareness of risk factors for attrition could direct targeted retention strategies in longitudinal studies. Moreover, the identification of patient characteristics associated with study dropout could address any potential bias introduced by differential dropouts.

1. University of British Columbia, Centre for Heart Lung Innovation, St Paul's Hospital, Vancouver, British Columbia, Canada
2. Research Institute, McGill University Health Centre, McGill University, Montreal, Quebec Canada
3. Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada
4. Institut Universitaire de Cardiologie et de Pneumologie de Québec, Université Laval, Quebec City, Quebec, Canada
5. Department of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada
6. University of Toronto, Toronto, Ontario, Canada
7. Department of Medicine, University of Calgary, Calgary, Alberta, Canada
8. Respiratory Research Centre, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
9. Queen's University, Kingston, Ontario, Canada
10. Department of Medicine, University of British Columbia and Vancouver General Hospital, Vancouver, British Columbia, Canada

* *Listed in Acknowledgements section*

Abbreviations:

ATS=American Thoracic Society; **BMI**=body mass index; **CanCOLD**=Canadian Cohort of Obstructive Lung Disease (longitudinal cohort); **CAT**=COPD Assessment Test; **CI**=confidence interval; **COLD**=Canadian Chronic Obstructive Lung Disease study (cross-sectional baseline study); **CT**=computed tomography; **COPD**=chronic obstructive pulmonary disease; **CVD**=cardiovascular disease; **DLCO**=diffusing capacity of the lungs for carbon monoxide; **FEV_1**=forced expiratory volume in 1 second; **FEV_1 %pred**=FEV_1 percentage of predicted normal; **FVC**=forced vital capacity; **GOLD**=Global initiative for chronic Obstructive Lung Disease; **HADS**=Hospital Anxiety and Depression Scale; **LLN**=lower limits of normal; **mMRC**=modified Medical Research Council; **SF-36**=Short Form 36 Health Survey

Funding Support:

The Canadian Cohort Obstructive Lung Disease (CanCOLD) (NCT00920348) study is currently funded by the Canadian Respiratory Research Network and the industry partners AstraZeneca Canada Ltd, Boehringer Ingelheim Canada Ltd, GlaxoSmithKline Canada Ltd, and Novartis. Researchers at RI-McGill University Health Centre Montreal and iCAPTURE Centre Vancouver lead the project. Previous funding partners were the Canadian Institutes of Health Research (CIHR; CIHR/

Rx&D Collaborative Research Program Operating Grants- 93326), the Respiratory Health Network of the Fonds de la Recherche en santé du Québec, and industry partners: Almirall, Merck Nycomed, Pfizer Canada Ltd, and Theratechnologies. The funding sponsors had no role in the study design, in the collection, analysis, and interpretation of data, in the writing of this manuscript, or in the decision to submit this manuscript for publication. The funders had no role in the study design, data collection, and analysis, or in the preparation of the manuscript.

Citation:

Katsuno N, Li PZ, Bourbeau J, et al. Factors associated with attrition in a longitudinal cohort of older adults in the community. ***Chronic Obstr Pulm Dis***. 2023;10(2):178-189. doi: https://doi.org/10.15326/jcopdf.2022.0380

Publication Dates:

Date of Acceptance: March 29, 2023
Published Online Date: April 27, 2023

Address correspondence to:

Wan C. Tan, MD
University of British Columbia
Centre for Heart Lung Innovation
St Paul's Hospital, Rm166
081 Burrard Street, Vancouver, B.C., V6Z 1Y6, Canada
Phone:1-(604)682-2344 ext. 62749
Email: wan.tan@hli.ubc.ca

Keywords:

population-based longitudinal cohort; predictor of attrition; retention strategies

Note: This article has an online supplement

Introduction

In epidemiological research, the longitudinal study is a design of choice for assessing the association between specific exposures with an outcome.[1] Such a study design monitors change and may identify factors that influence the progression and development of a disease.[2] However, retaining participants in a longitudinal study remains a challenge.[2,3] Attrition is a concern as it could lead to selection bias and compromise the generalizability of a study.[2,4-6] To mitigate preventable attrition, reasons for attrition must be addressed to ensure that appropriate retention strategies are applied.[7]

Previous research has identified older age as an independent predictor of attrition.[8] A comprehensive assessment of reasons for attrition is particularly important when studying older adults because they are vulnerable to visual, auditory, and cognitive impairments and multiple chronic conditions that challenge their continued study participation.[9] Maximizing the retention of samples of older people is key to understanding change over time and identifying factors responsible for the change to obtain generalizable findings in older people.[9]

In this study, we examined the different reasons for participant loss in a longitudinal cohort of adults with and without chronic obstructive pulmonary disease (COPD) in the general population and determined the modifiable factors associated with an increased attrition rate.

Methods

Study Population

The Canadian Chronic Obstructive Lung Disease (COLD) prevalence study recruited a sample of non-institutionalized adults 40 years and older by random digit dialing using census data in 2005–2009 from 9 urban communities in Canada.[10] The Canadian Cohort of Obstructive Lung Disease (CanCOLD), a nested community-based case-control study, then enrolled COLD participants with COPD, in addition to representative random subsets of COLD non-smoking participants, and smoking participants without COPD, matched on age and gender in 2010–2014. (ClinicalTrials.gov Identifier: NCT00920348).[11] This study included 1561 CanCOLD study participants, followed up as a longitudinal cohort. Details on the study's sampling methodology can be found elsewhere.[11] All participants provided written informed consent before completing study assessments. The research ethics board of each participating institution approved the study protocol.

Follow-up Assessment

Participants completed in-person visits at intervals of at least 18 months. For all visits, participants completed an interview-administered questionnaire, performed pre- and post-bronchodilator spirometry using a portable spirometer (Easyone) according to American Thoracic Society (ATS) criteria,[10,12] and venous blood samples were obtained. At baseline and for every other visit, additional tests performed included lung volume measurements by whole-body plethysmography and diffusion capacity of the lungs for carbon monoxide, which were obtained according to standard techniques[13]; a 6-minute walk test (performed according to the ATS guidelines[14]; a symptom-limited incremental cardiopulmonary exercise test conducted on an electronically braked cycle ergometer according to recommended guidelines[15,16]; and low-dose expiratory and inspiratory computed tomography scans (CT) of the chest were acquired using a multi-slice CT scanner (≥16 detectors).[11]

Additionally, participants were followed up every 3 months over the phone or by email. At these virtual visits, a standardized exacerbation questionnaire was administered to determine the participants' state of health and report any new onset of respiratory symptoms.

Retention Strategies to Reduce Attrition

The CanCOLD protocol[11] has a retention plan that includes, but is not limited to, the following strategies:

1. Participants are contacted by email or by phone at 3-monthly intervals as described above throughout the study. Attempts at establishing contact by telephone are made at different times and days of the week. A reminder is sent to participants who do not call back or respond to the email within the week of contact. After 6 weeks of being unreachable, a minimum of 3 attempts are further made over 2 weeks to re-establish contact. A participant is deemed "noncontactable" only after 7 documented, unsuccessful attempts to reach the participant. In general, with these participants, we receive an email or a telephone message of "returned to sender" due to participant relocation, telephone number "not in service,"' or emails were "undeliverable."
2. Regular feedback is provided to participants following each visit. Participants receive a report in brief lay language explaining the tests and questionnaire responses. This report is independent of test results that can be provided to the participant if required, or if needed for follow-up (e.g., CT scan) as part of good clinical practice.
3. An annual report of the study's progress is created.
4. Enrolled participants receive a birthday card and a New Year's card, thanking them for their participation in the study. All these strategies are coordinated by the study central office and were implemented upon approval by the local sites' research ethics boards.

Definitions for Reasons of Attrition

The 6 categories of reasons for attrition are as follows:

1. "Dropout" = participant-initiated termination of participation.
2. "Withdrawal" = investigator-initiated withdrawal of participant from the study.
3. "Died" = deceased.
4. "Relocation" = permanently left the region.
5. "Medical Exemption"' = initiated by the family physician for physical or medical reasons of the participant that impede continued participation.
6. "Loss to Follow-up" = outdated contact information.

Table 1a shows a detailed description of the wide range of individual reasons which are re-grouped for simplicity and frequency computation under the 6 categories shown in Table 1b.

Statistical Analysis

Participants in each of the 9 study sites were described by baseline demographic characteristics, respiratory symptoms, self-reported comorbidities, and spirometry results. Retention rates for the whole cohort for each site and for each year from 2011 (cohort start) to 2019 (cohort end) were estimated by a formula: the number of participants kept in the cohort at the end of that year divided by the total participants at the beginning of that year. The average retention rate was also calculated. Unpaired *t*-tests were used to compare continuous variables and Chi-squared tests for categorical variables between the 2 groups: those who remained in the cohort and those who did not in the period of 2011 to 2019. In the regression models, we included variables that would potentially be associated with attrition including self-reported ethnicity, exacerbation rate, forced expiratory volume in 1 second percentage predicted (FEV_1 %pred), pack years of tobacco smoked, cardiovascular diseases, physician-diagnosed asthma, COPD spirometrically defined as FEV_1 to forced vital capacity (FVC) ratio <lower limits of normal (LLN),[17] and COPD exacerbations in the previous year, and mutually adjusted for all potential risk factors in the models. To assess the factors associated with attrition rate, univariate and multivariate proportional sub-distribution hazards models with competing risks[18] were conducted. In the analysis, the data were classified into statuses: censored (keep in the cohort), events of interest (dropout/withdrawal), and competing events (died). Hazard ratio (95% confidence interval [CI]) and standard error ratio were estimated using the SAS macro for proportional and non-proportional sub-distribution hazards regression[19] with a robust sandwich covariance matrix. Significance was considered at alpha <0.05. Statistical analyses were performed in SAS 9.4 (SAS Institute, Cary, North Carolina).

Results

Cohort Characteristics

Table 2 shows the baseline demographics and characteristics of the 1561 participants in the study, stratified by those who remained in the study and those who did not. The mean age is 66.7 years with 56.1% of the cohort (876/1561) consisting of men. A third of the population (467/1561; 30.7%) have COPD, most with mild (Global initiative of chronic Obstructive Lung Disease [GOLD] stage 1)[20] airflow obstruction. Table 2 shows that, compared with people who remain in the cohort, those who are not retained in the study tend to be older, heavier (higher body mass index), have fewer years of education (<12 years), have greater tobacco consumption, have cardiovascular comorbidities, report more dyspnea, have lower lung function, higher COPD Assessment Test (CAT) scores, lower health status score on the Short Form-36 (SF36) Health Survey- physical

Table 1a. Reasons for Non-Retention

Dropout *(Study discontinued by participant)*	Health-related Issues	Dizziness
		Increasing cough and phlegm
		Heart disease and multi-comorbidities
		Sciatica and mobility issues
	Personal Perceptive Issues	Depression, arrhythmia and cataract
		Post thoracotomy
		Study too long
		Study too time consuming
		Lost interest
		Transportation problem
		Family/ domestic problems
Withdrawal *(Study discontinued by investigator)*		Recurrent pneumonia
		Cerebrovascular event
		Advanced dementia
		Transportation issue
		Participation in an in interventional study which precludes continuation in CanCOLD ,an observational study.
Death Documentation		Deaths reported by relatives and friends
		Deaths published in obituaries
Relocation *(Study discontinued by participant)*		Moved to another city
		Moved to another province within Canada
		Moved to another country
Medical Exemption *(Study discontinued by family physician)*		Mobility issues
		Lung cancer
		Other life-threatening disease (unstable aneurysm)
Reasons for Loss to Follow-up *(Study discontinued by investigator)*		Phone not in service
		Unable to establish contact by phone, e-mail, or mail
		No alternate contact

CanCOLD=Canadian Cohort of Obstructive Lung Disease

Table 1b. The Reasons for Attrition by Site Shown As Numbers of Participants[a] and Frequencies[b] Censored in 2019

Site	Total n	6 Reasons for Attrition						Attrition Rate for Each Site
		Dropout n (%)	Withdrawal n (%)	Died n (%)	Relocation n (%)	Medical Exemption n (%)	Loss to Follow-up	
1	435	57 (13.1)	5 (1.2)	13 (3.0)	1 (0.2)	3 (0.7)	14 (3.2)	93 (21.4)
2	355	34 (9.6)	0 (0.0)	3 (0.9)	0 (0.0)	0 (0.0)	35 (9.9)	72 (20.3)
3	74	0 (0.0)	11 (14.9)	1 (1.4)	1 (1.4)	11 (14.9)	4 (5.4)	28 (37.8)
4	127	4 (3.2)	6 (4.7)	4 (3.2)	0 (0.0)	0 (0.0)	0 (0.0)	14 (11.0)
5	129	24 (18.4)	2 (1.6)	2 (1.6)	0 (0.0)	0 (0.0)	1 (0.8)	29 (22.5)
6	116	7 (6.0)	14 (12.1)	2 (1.7)	1 (0.9)	0 (0.0)	2 (1.7)	26 (22.4)
7	136	14 (10.3)	1 (0.7)	3 (2.2)	1 (0.7)	0 (0.0)	30 (22.1)	49 (36.0)
8	85	0 (0.0)	7 (8.2)	3 (3.5)	1 (1.2)	3 (3.5)	5 (5.9)	19 (22.4)
9	104	0 (0.0)	9 (8.7)	2 (1.9)	1 (1.0)	15 (14.4)	6 (5.8)	33 (31.7)
Total	**1561**	**140 (9.0)**	**55 (3.5)**	**33 (2.1)**	**6 (0.4)**	**32 (2.1)**	**97 (6.2)**	**363 (23.3)**

[a](n)
[b](%)
From 1-9 are the following sites retrospectively: Vancouver, Montreal, Toronto, Halifax, Calgary, Ottawa, Kingston, Quebec, and Saskatoon.
The table above shows the numbers and frequency of the reasons for attrition for the year evaluated (2019) across the 9 sites.

Dropout=participant-initiated withdrawal; Withdrawal= investigator-initiated withdrawal; Died=deceased; Relocation=out-of-region; Medical Exemption=reason initiated by family physician; Loss to Follow-up=outdated contact information, loss of contact

Table 2. The Baseline Characteristics of the Patients for All the Participants, Stratified by Those Who Remained in the Cohort and Those Who Left the Cohort

	n	Total n=1561	Retained n=1198	Not Retained n=363	*P*-value
Age, in years	1561	66.7±9.8	66.2±9.4	68.3±10.8	<0.001*
Sex, male, n (%)	1561	876 (56.1)	671 (56.0)	205 (56.5)	0.876
Race, White, n (%)	1561	1486 (95.2)	1142 (95.3)	344 (94.8)	0.662
BMI, kg/m^2	1558	27.7±5.3	27.5±5.2	28.3±5.4	0.006*
Education (≤12 years), n (%)	1548	336 (21.7)	234 (19.7)	102 (28.4)	<0.001*
Never Smokers, n (%)	1561	546 (35.0)	450 (37.6)	96 (26.4)	<0.001*
Former Smokers, n (%)	1561	765 (49.0)	564 (47.1)	201 (55.4)	0.006*
Current Smokers, n (%)	1561	250 (16.0)	184 (15.4)	66 (18.2)	0.199
Pack Years	1539	17.1±22.9	15.6±21.5	21.8±26.8	<0.001*
COPD-LLN, n (%)	1521	467 (30.7)	364 (30.7)	103 (30.7)	0.985
Marijuana Smokers, joint years	1439	2.6±13.8	2.4±11.6	3.5±19.0	0.168
Four Study Groups, n (%)					
Healthy	1521	336 (22.1)	276 (23.3)	60 (17.9)	0.037*
At Risk	1521	466 (30.6)	353 (29.8)	113 (33.7)	0.164
GOLD1	1521	402 (26.4)	325 (27.4)	77 (23.0)	0.105
GOLD2+	1521	317 (20.8)	232 (19.6)	85 (25.4)	0.021*
CVD, n (%)					
Including Hypertension	1561	787 (50.4)	573 (47.8)	214 (59.0)	<0.001*
Excluding Hypertension	1561	461 (29.5)	317 (26.5)	144 (39.7)	<0.001*
Asthma, n (%)	1561	361 (23.1)	286 (23.9)	75 (20.7)	0.204
Chronic Cough, n (%)	1561	248 (15.9)	182 (15.2)	66 (18.2)	0.172
Chronic Phlegm, n (%)	1561	208 (13.3)	149 (12.4)	59 (16.3)	0.061
Wheeze, n (%)	1561	435 (27.9)	329 (27.5)	106 (29.2)	0.517
mMRC Scales (1-5), n (%)					
1	1481	919 (62.1)	742 (64.8)	177 (52.7)	<0.001*
2	1481	470 (31.7)	350 (30.6)	120 (35.7)	0.075
3+	1481	92 (6.2)	53 (4.6)	39 (11.6)	<0.001*
FEV_1, L	1521	2.6±0.8	2.6±0.8	2.4±0.8	<0.001*
FVC,L	1521	3.7±1.1	3.7±1.1	3.5±1.0	<0.001*
FEV_1, % predicted-NHANES	1521	91.7±20.3	92.0±19.4	90.5±23.2	0.272
FVC, % predicted-NHANES	1521	98.9±17.2	99.3±16.7	97.4±18.9	0.09
FEV_1/FVC, %	1521	69.5±10.4	69.7±10.0	69.1±11.7	0.896
COPD Exacerbation in the Previous Year, n (%)	1560	70 (4.5)	53 (4.4)	17 (4.7)	0.885
CAT Score	1527	6.9±6.0	6.5±5.6	8.3±7.2	<0.001*
SF36 Mental Component Score	1557	50.3±9.3	50.5±9.1	49.7±10.1	0.305
SF36 Physical Component Score	1558	50.7±8.9	51.3±8.4	48.8±10.3	<0.001*
HADS-Anxiety Score	1547	3.9±3.2	3.9±3.1	4.1±3.4	0.501
HADS-Depression Score	1551	2.8±2.7	2.7±2.6	3.3±3.2	0.010*
Self-reported Major Depression	1561	81 (5.2)	60 (5.0)	21 (5.8)	0.559

Unpaired *t*-tests were used to compare continuous variables and Chi-squared test compared categorical variables between the two groups (retained and not retained) in the period of 2011 to 2019.

CVD assessments are hypertension, stroke, transient ischemic attack, coronary artery disease, myocardial infarction, unstable angina, stable angina, heart failure, arrhythmia (atrial fibrillation, etc), peripheral artery disease, valvular heart disease, other cardiovascular conditions.

Asthma included only physician-diagnosed asthma.

mMRC Scale indicates dyspnea levels: 1=not troubled by breathlessness except with strenuous exercise, 2=troubled by shortness of breath when hurrying on the level of walking up a slight hill., 3=walks slower than people of the same age on the level because of breathlessness, or has to stop for breath when walking at own pace on the level, 4=stops for breath after walking about 100 years (90m) or after a few minutes on level, 5=too breathless to leave the house or breathless when dressing or undressing.

Marijuana smoking: total lifetime exposure in joint years (number of joints per day x number of years smoked).

SF-36: to measure perceived general health and well-being with 2 overall summary scores—physical and mental component scores.

Retained=remained in the cohort; Not retained = did not remain in the cohort due to attrition;

BMI=body mass index; COPD=chronic obstructive pulmonary disease; LLN=lower limits of normal; GOLD=Global initiative for chronic Obstructive Lung Disease; CVD=cardiovascular disease; mMMR=modified Medical Research Council; FEV_1=forced expiratory volume in 1 second; FVC=forced vital capacity; CAT=COPD Assessment Test; SF-36=Short Form 36 Health Survey; HADS=Hospital Anxiety and Depression Scale

component score, and higher Hospital Anxiety and Depression Scale (HADS) depression scores.

Attrition

Table 3 shows the number of participants recruited and retained in the study and the estimated cumulative retention rates for each year of follow-up. The duration of follow-up ranged from 4 to 9 years for groups of participants recruited in different years. The mean annual retention rates (calculated by excluding deaths) for the whole cohort remained above 80% for 8 years of follow-up with a final 77% at year 9 of follow-up. Figure 1 compares the retention rates expressed as a percentage of all recruited participants when deaths were excluded or included in the calculation and shows a difference of 1%–2%, which remained relatively constant throughout the years of follow-up. Among the 9 sites, there were differences in retention, with average retention rates ranging from 90% to 40% (Table S1-S9 in the online supplement).

Reasons for Dropout From the Study

The specific reasons for dropout are listed in Table 1a. These specific reasons for dropout were recorded at the time of dropout and are descriptive and wide-ranging but were related to either poor health or personal reasons, which included loss of interest in the study, did not like the study, problems of transportation, or presence of a domestic conflict. For clarity, we present the reasons for dropouts under 6 categories as shown in Table 1b.

Table 1b shows the reasons and frequencies for attrition in the study by site censored in December 2019. Of the 363 cases who did not remain in the study, the reasons in descending order of frequency and expressed as a percentage of all dropouts were: participant-initiated withdrawal (dropout) accounting for 39% (140/363); loss to follow-up (loss of contact) 26.7% (97/363); investigator–initiated withdrawal (withdrawal)15.2% (55/36); mortality (died) 9.1% (33/363); medical exemption (family physician initiated withdrawal) 8.8% (32/363); and out-of-region withdrawal status (relocation), 1.7% (6/363).

Modifiable Risk Factors for Attrition: Comparisons of Cohort "Remainers" versus "Non-Remainers"

To identify factors associated with increased attrition the proportional sub-distribution hazards model was used to conduct a univariate and multivariate analysis with death as the competing event. (Table 4).

In the univariate unadjusted model, age, education less than or equal to 12 years, pack years of tobacco smoking, cardiovascular disease as a comorbidity, CAT score, SF36 physical component score, and HADS-depression score were significantly associated with increased attrition. After adjustments for all other variables in the model, factors independently associated with attrition were lower educational level, pack years of smoking, cardiovascular disease, and the HADS depression score: adjusted hazard ratios (95%CI) were 1.43 (1.11, 1.85), 1.01 (1.00, 1.01), 1.44 (1.13, 1.83), and 1.06 (1.02, 1.10), respectively.

Discussion

Long-term longitudinal studies are invaluable for evaluating the natural history of disease and the salient risk factors for disease progression or remission but their findings can be undermined by attrition of participants over time, a problem that is particularly common among older participants.[7,8,21] In CanCOLD, we achieved a high overall annual retention rate (>80%) over 9 years with the retention rate decreasing below 80% only in year 9 among individuals in the community who were predominately in their sixth, seventh, or eighth decade of life. These rates are comparable to other published longitudinal studies with high retention rates.[1]

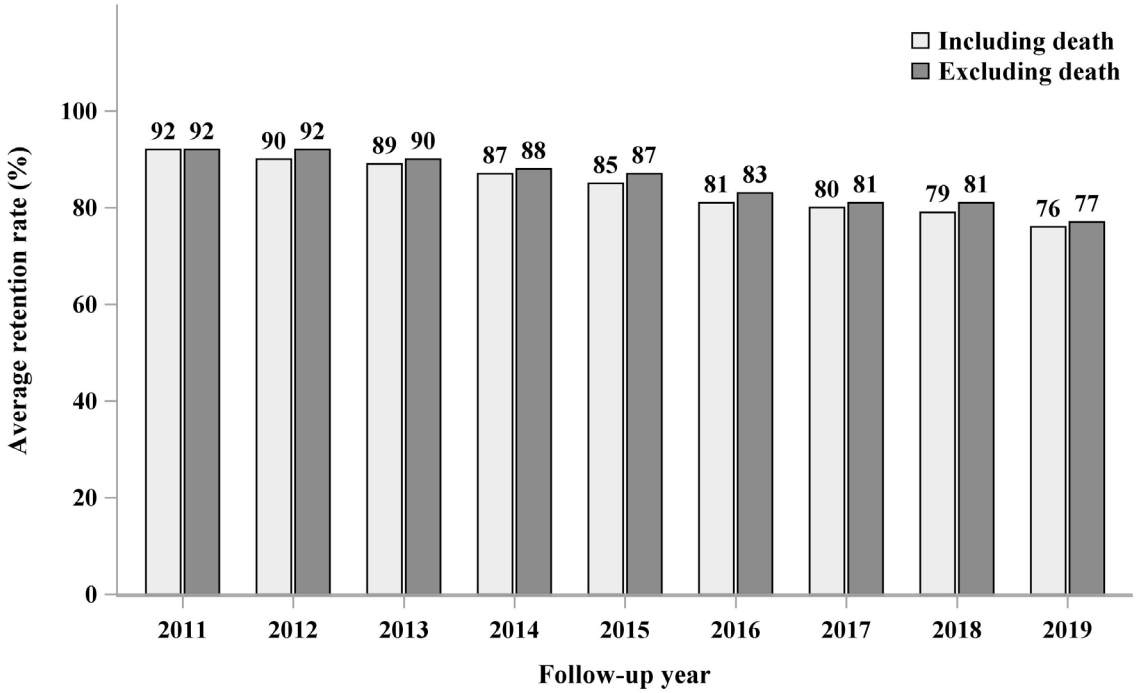
Table 3. The Number of Participants and the Retention Rates for Each Follow-up Year, 2011 to 2019[a]

Recruited in Year	Newly Recruited Participants	Retention Rate Over Follow-up Year, n (%)								
		2011	2012	2013	2014	2015	2016	2017	2018	2019
2010	137	126 (92%)	119 (87%)	115 (84%)	107 (78%)	106 (77%)	106 (77%)	106 (77%)	106 (77%)	106 (77%)
2011	242		233 (96%)	219 (90%)	209 (86%)	205 (85%)	205 (85%)	204 (84%)	203 (84%)	200 (83%)
2012	389			376 (97%)	356 (92%)	331 (85%)	317 (81%)	315 (81%)	315 (81%)	310 (80%)
2013	454				441 (97%)	421 (93%)	388 (85%)	375 (83%)	369 (81%)	357 (79%)
2014	246					232 (94%)	212 (86%)	203 (83%)	196 (80%)	182 (74%)
2015	60						48 (80%)	48 (80%)	48 (80%)	43 (72%)
Total	**1528**	**126 (92%)**	**352 (92%)**	**710 (90%)**	**1113 (88%)**	**1295 (87%)**	**1276 (83%)**	**1251 (81%)**	**1237 (81%)**	**1198 (77%)**

[a]Excluding those who have died.

The retention rate=number remaining in the cohort / (number recruited minus death).

Figure 1. Average Retention Rate With Time of Follow-up



Comparison of the retention rates expressed as a percentage of all recruited participants when deaths were excluded with the rates when deaths were included in the calculation. The difference between the 2 rates is 1%–2%, which remains relatively constant through the years of follow-up.

The present study also examined the common reasons for attrition in a longitudinal cohort of older adults in the population and identified several risk factors that increased attrition rates. We found that the main explanation for attrition was participant-initiated and that the underlying independent participant risk factors associated with increased dropouts were lower educational level, higher pack-year tobacco consumption, the presence of cardiovascular disease, and an increased HADS-depression score based on a questionnaire, all of which are potentially modifiable. Participant attrition is a major issue in longitudinal studies involving older people, which could result in selection bias if those who leave the study differ in characteristics from those who remain. Knowledge of the predictors of attrition could help identify individuals "at risk" for attrition and enable the channeling of efforts, strategies, and resources for maximizing retention to these subgroups to reduce the attrition rate over time.

The results of the study concur with previously observed findings of factors associated with attrition in longitudinal studies in which lower socioeconomic level (fewer years of education as a surrogate) and poorer health have been shown to be predictors of attrition.[3,21-23] Here we extend this body of knowledge by showing that cardiovascular disease as a comorbidity rather than poorer general health per se, and cigarette smoking as well as depression represented by a high HADS-depression score, are additional predictors of attrition in a longitudinal study. An unadjusted comparison between participants who remained in the study and those who left the study initially suggested that age was a risk factor for attrition. However, after full adjustment and stringent modeling, only 4 factors remained significant, namely fewer years of education, smoking, the presence of cardiovascular disease, and depression, as independent predictors of attrition. It is interesting that the multivariable hazards model suggested that participants with COPD, even those with a history of exacerbations in the previous year or self-reported physician–diagnosed asthma in our study, were not more likely than those without the disease to drop out of the study. A potential explanation is that people with health problems tended to be more diligent participants in studies, as they were interested in receiving medical attention.[8] Further, the participants with COPD in our study were mostly mild – few had severe or very severe disease— and this might also help explain why having a chronic disease was not an independent risk factor in predicting attrition.

The findings in this study differ from those in studies with a specific focus on aging[24,25] that examined preventable reasons for attrition and found patterns of social vulnerability such as older age, poor functioning, cognitive impairment, living alone, and not being married were associated with more dropouts.[8] In our study, age after adjusting for confounders, was not a predictor of withdrawal from the study. This difference could be explained by the fact that the CanCOLD study is not a specific study on aging, rather it is a cohort of a population sample composed of healthy smokers

Table 4. Factors Associated With Increased Attrition Rate, Results From the Proportional Subdistribution Hazards Model

	Univariate Analysis				Multivariate Analysis With Stepwise Approach			
	HR (95% CI)	Standard Error	Standard Error Ratio	*P*-value	Adj. HR (95% CI)	Standard Error	Standard Error Ratio	*P*-value
Age	1.02 (1.00, 1.03)	0.007	1.104	0.017*	-	-	-	-
Sex, male	1.08 (0.85, 1.36)	0.119	0.995	0.537	-	-	-	-
Race, White	1.20 (0.66, 2.18)	0.305	0.993	0.55	-	-	-	-
Education (≤12 years)	1.63 (1.27, 2.10)	0.128	0.991	<0.001*	1.43 (1.11, 1.85)	0.131	0.979	0.006*
Pack Years	1.008 (1.004, 1.012)	0.002	0.975	<0.001*	1.005 (1.001, 1.009)	0.003	0.976	0.032*
CVD	1.58 (1.25, 2.00)	0.121	0.994	<0.001*	1.44 (1.13, 1.83)	0.122	0.991	0.003*
Marijuana Smokers Joint Years	1.005 (1.000, 1.011)	0.003	0.871	0.055				
Asthma	0.89 (0.67, 1.17)	0.142	0.993	0.403	-	-	-	-
FEV_1, % predicted-NHANES	1.00 (0.99, 1.01)	0.003	1.073	0.986	-	-	-	-
COPD-LLN	0.95 (0.75, 1.22)	0.126	0.989	0.706	-	-	-	-
COPD Exacerbations in the Previous Year	0.90 (0.51, 1.59)	0.288	0.977	0.722	-	-	-	-
CAT Score	1.03 (1.02, 1.05)	0.009	1.023	<0.001*	-	-	-	-
SF36 Mental Component Score	0.99 (0.98, 1.00)	0.006	1.04	0.115	-	-	-	-
SF36 Physical Component Score	0.98 (0.97, 0.99)	0.007	1.047	0.003*	-	-	-	-
HADS-Anxiety Score	1.02 (0.98, 1.06)	0.018	1.023	0.257	-	-	-	-
HADS-Depression Score	1.08 (1.03, 1.12)	0.02	1.027	<0.001*	1.06 (1.02, 1.10)	0.02	1.01	0.006*
Self-reported Major Depression	1.04 (0.63, 1.72)	0.255	0.99	0.864	-	-	-	-

COPD-LLN was spirometrically-defined COPD where FEV_1/FVC < lower limits of normal.

CVD assessments excluded hypertension but included stroke, transient ischemic attack, coronary artery disease, myocardial infarction, unstable angina, stable angina, heart failure, arrhythmia (atrial fibrillation, etc), peripheral artery disease, valvular heart disease.

Asthma included only physician-diagnosed asthma.

- indicates not significant

In univariate analysis, age, education, pack years, and CVD are significant.

In multivariate analysis, only education, pack years, and CVD are significant.

P-values are based on the robust sandwich covariance matrix.

Standard error ratio columns contain the ratios of the robust estimate of the standard error relative to the corresponding model-based estimate standard error.

HR=(subdistribution) hazard ratios; Adj. HR=adjusted hazard ratios; CI=confidence interval; CVD=cardiovascular disease; FEV_1=forced expiratory volume in 1 second; COPD=chronic obstructive pulmonary disease; LLN=lower limit of normal; CAT=COPD Assessment Test; SF-36=Short Form 36 Health Survey; HADS=Hospital Anxiety and Depression Scale

and never smokers with and without COPD. Further, we did not detect any sex/gender bias in non-retention in the study, as has been noted previously in men.[8,25]

Site Variability in Reasons for Dropout

We noted variability in the categories of participant dropout across study sites, but the reason for this was not immediately obvious. It is conceivable that there could be minor variations in individual site investigators' interpretations of the different categories of attrition. However, all attempts were made to minimize this by ensuring that at the start of the study, the personnel at each site were trained collectively and had a consensus on categorizing the reasons as either health-related or personal for *dropout* and *withdrawal*. The personnel at all sites were also briefed on all definitions:

(1) Deaths were recorded when reported by relatives and by searching for obituaries on the internet when contact failed.

(2) There were 2 types of medical exemptions—mobility issues and life-threatening conditions and the request for exemption was made by the family physician of the participant.

(3) Specific criteria was used for loss to follow-up when the staff was unable to establish contact by phone (not in service), email, or mail and when there was no alternate contact.

In this study, geography was an unlikely reason for significant non-retention as at the time of recruitment the participants all lived within 10 miles of the study site. However, 6 participants did not continue in the study because they moved to another city, another province within Canada, or another country.

Strengths and Potential Limitations of This Study

There are several strengths of the study. First, a high retention rate of greater than 80% for 8 out of 9 years period of the longitudinal study of older individuals randomly recruited from the general population compares very favorably with those reported in longitudinal cohorts with the highest retention rates[1] and with studies of aging of shorter duration of follow-up.[7,8,21]

Second, we have included quality of life and mental health measures in our analyses. These assessment tools included the SF-36 (to measure perceived general health and well-being), the CAT, and depression and anxiety variables namely the HADS-anxiety score and the HADS-depression score, and self–reported major depression. A further novel analysis is the impact of marijuana smoking expressed as joint years on attrition.

Third, our analyses identified modifiable risk factors namely, low education level, the burden of tobacco exposure, the presence of cardiovascular disease, and depression measured by a high HADS-depression score, to be associated with non-retention and found no significance associated with non-modifiable factors such as age and sex/gender, as reported in the literature. Interestingly, marijuana smoking adjusted for cigarette smoking was not a risk factor for attrition in the study. To our knowledge, the impact of marijuana smoking on attrition in a longitudinal study has not been previously evaluated.

Fourth, tailored retention strategies and practices were instituted immediately after recruitment and maintained throughout the follow-up to reduce bias and enhance retention efforts over time. These include repeated mailings, telephone contacts, hybrid inquiries with the options of responding by either postal or internet services, holiday greeting cards, and financial reimbursement for expenses, and were supported by a positive, consistent relationship between participant and research staff, all of which have been previously advocated as effective motivational measures in longitudinal studies of randomized clinical trials, birth cohorts, and disease-specific cohorts.[2,3,21,26-29] We have shown that retention efforts are also effective for achieving the goal of good retention in a non-disease-specific observational cohort comprising population-derived older individuals.

Fifth, the CanCOLD cohort was recruited from a population-based sample and is not a disease-specific cohort but includes a spectrum of healthy people, never smokers, and smokers with and without disease of varying severity, thus providing the opportunity to examine health-related issues as reasons for study attrition of healthy and disease comparator subgroups in the general population. This adds to the information on the risk of attrition found in disease-specific cohorts.[1]

Lastly, we used a competing risk model (sub-distribution model by Fine and Gray)[18] to analyze our data, as it applies to the sub-hazard underlying the cumulative incidence function (sub-distribution hazard), not the cause-specific hazard. In the presence of competing events, a sub-distribution hazard model better predicts the cumulative incidence function than the classical cause-specific Cox proportional hazard model.

There were limitations in the study. Our study did not evaluate *a priori* potential risk factors associated with social vulnerability, in particular sociopsychological factors such as having a lower job grade or not being a homeowner, and neurocognitive details, all of which had been associated with a greater probability of withdrawal.[8,9]

In this longitudinal study we did not have formal measures of health literacy, a recognized patient risk factor for poor health. Instead, we found that low education was an independent predictor of participant dropout. Since low or limited health literacy has been associated with low education,[30-32] we speculate that health literacy could conceivably be an explanation for the association between dropout and educational level in this study. Measurement of health literacy in future longitudinal studies is needed to better define its role in study attrition.

In conclusion, good retention rates are achievable with the inclusion of *a priori* retention strategies in a longitudinal study of unselected population-derived adults and underscore the general effectiveness of retention strategies for longitudinal studies regardless of age groups and participant types. Finally, although patient factors such as education, cardiac disease, smoking, and depression are shown to be the risk factors in this study, a comprehensive and logical approach would be to tailor retention strategies for individuals with different life circumstances to promote continued participation in the study. Such strategies would include using remote procedures, reducing participant burden for individuals with multiple chronic diseases, and ensuring appropriate health literacy for individuals with low education.

Acknowledgements

Author contributions: All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work. WCT and JB had full access to all data in this study and had final responsibility for the decision to submit this manuscript for publication. All authors approved the final version of the manuscript.

Data sharing: The CanCOLD study makes its population-based longitudinal data available to Canadian and international investigators and trainees. Applications for data access can be made by visiting the website http://www.cancold.ca and its data portal.

The authors would like to thank the men and women who participated in the study and individuals in the CanCOLD Collaborative Research Group: Jonathon Samet (the Keck School of Medicine of University of Southern California, Los Angeles, California); William M Vollmer (The Kaiser Permanente Center for Health Research in Portland, Oregon.) Milo Puhan (John Hopkins School of Public Health, Baltimore, Maryland); Qutayba Hamid, Carolyn Baglole, Benjamin M. Smith, Palmina Mancino, Pei-Zhi Li, Zhi Song, Dennis Jensen (McGill University, Montreal, Quebec, Canada); Yvan Fortier (Sherbrooke University, Sherbrooke, Quebec, Canada); Patricia McClean, Jane Duke, Andrea S. Gershon, Teresa Toh (University of Toronto, Toronto, Ontario, Canada); Mohsen Sadatsafavi (University of British Columbia, Vancouver, British Columbia, Canada); Christine Lo, Sarah Cheng, Elena Un, Michael Cheng, Cynthia Fung, Faize Faroon, Olga Radivojevic, Sally Chung, Carl Zou, Rena Choi, Joe Comeau, Harvey Coxson, Miranda Kirby, Jonathon Leipsic, Cameron Hague (James Hogg Research Center, University of British Columbia, Vancouver, British Columbia, Canada); Curtis Dumonceaux, (University of Calgary, Calgary, Alberta, Canada); Scott Fulton, (Dalhousie University, Halifax, Nova Scotia, Canada); Kathy Vandemheen, (University of Ottawa, Ottawa, Ontario, Canada); Matthew McNeil, Kate Whelan (Queen's University, Kingston, Ontario, Canada); Cynthia Brouillard (University of Laval, Quebec City, Quebec, Canada); Ron Clemens, Janet Baran (University of Saskatchewan, Saskatoon, Saskatchewan, Canada).

Declaration of Interest

WCT reports grants from the Canadian Institute of Heath Research (CIHR)(Rx&D Collaborative Research Program Operating Grants- 93326) with industry partners Astra Zeneca Canada Ltd., Boehringer-Ingelheim Canada Ltd., GlaxoSmithKline Canada Ltd., Merck, Novartis Pharma Canada Inc., Nycomed Canada Inc., and Pfizer Canada Ltd., during the conduct of the study; and personal fees from GlaxoSmithKline Canada Ltd., and Astrazeneca Canada Ltd., outside the submitted work. JB reports grants from CIHR, and grants from the Canadian Respiratory Research Network during the course of the study; personal fees from the Canadian Thoracic Society and CHEST, grants from the Foundation of the McGill University Health Centre and Aerocrine, and grants and personal fees from AstraZeneca, Boehringer Ingelheim, Grifols, GlaxoSmithKline, Novartis, and Trudell outside the submitted work. DS reports grants from Merck, personal fees from Sanofi-Aventis and Regeneron, grants and personal fees from Boehringer Ingelheim and AstraZeneca, and personal fees from Novartis, outside the submitted work. FM reports grants from CanCOLD, during the conduct of the study, grants from GlaxoSmithKline, AstraZeneca, Sanofi, personal fees from GlaxoSmithKline, Boehringer Ingelheim, Grifols, Novartis, unrestricted grants paid to their institution from Novartis, Boehringer Ingelheim, Grifols, and financial participation in Oxynov, a company which is developing an oxygen delivery system, outside the submitted work. DDM reports grants from McGill University, during the conduct of the study; grants from AstraZeneca, Boehringer-Ingelheim, GlaxoSmithKline, Canadian Health for Healthcare Improvement, the Lung Association of Saskatchewan, CIHR, Grifols, Novartis, Sanofi, the Saskatchewan Health Research Foundation, and Schering-Plough. DDM is an employee of the University of Saskatchewan, outside the submitted work. PH reports personal fees from AstraZeneca, grants from Boehringer Ingelheim, Grifols, and Vertex, and fees for medical advisory boards from Actelion, GlaxoSmithKline, and Novartis, outside the submitted work. DDO, KRC, BW, SA, JCH, PZL, NK, MC, JR have nothing to declare.

References

1. Abshire M, Dinglas VD, Cajita MI, Eakin MN, Needham DM, Himmelfarb CD. Participant retention practices in longitudinal clinical research studies with high retention rates. *BMC Med Res Methodol.* 2017;17(1):30-40. https://doi.org/10.1186/s12874-017-0310-z

2. Booker CL, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. *BMC Public Health.* 2011;11:249-261. https://doi.org/10.1186/1471-2458-11-249

3. Launes J, Hokkanen L, Laasonen M, et al. Attrition in a 30-year follow-up of a perinatal birth risk cohort: factors change with age. *PeerJ.* 2014;2:e480-e499. https://doi.org/10.7717/peerj.480

4. Fewtrell MS, Kennedy K, Singhal A, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Arch Dis Child.* 2008;93(6):458-461. https://doi.org/10.1136/adc.2007.127316

5. Walker J, Fairley CK, Urban E, et al. Maximising retention in a longitudinal study of genital Chlamydia trachomatis among young women in Australia. *BMC Public Health.* 2011;11:156-164. https://doi.org/10.1186/1471-2458-11-156

6. Siddiqi AE, Sikorskii A, Given CW, Given B. Early participant attrition from clinical trials: role of trial design and logistics. *Clin Trials.* 2008;5:328-335. https://doi.org/10.1177/1740774508094406

7. Davies K, Kingston A, Robinson L, et al. Improving retention of very old participants in longitudinal research: experiences from the Newcastle 85+ study. *PLoS One.* 2014;9(10): e108370. https://doi.org/10.1371/journal.pone.0108370

8. Mein G, Johal S, Grant RL, Seale C, Ashcroft R, Tinker A. Predictors of two forms of attrition in a longitudinal health study involving ageing participants: an analysis based on the Whitehall II study. *BMC Med Res Methodol.* 2012;12:164-171. https://doi.org/10.1186/1471-2288-12-164

9. Heid AR, Cartwright FP, Wilson-Genderson M, Pruchno R. Understanding attrition and bolstering retention in a longitudinal panel of older adults: ORANJ BOWL. *Innov Aging.* 2021; 5(2):1-13. https://doi.org/10.1093/geroni/igab010

10. Tan WC, Bourbeau J, O'Donnell D, et al. Quality assurance of spirometry in a population-based study -predictors of good outcome in spirometry testing. *COPD.* 2014;11(2):143-151. https://doi.org/10.3109/15412555.2013.822857

11. Bourbeau J, Tan WC, Benedetti A, et al. Cancold Study G. Canadian Cohort Obstructive Lung Disease (CanCOLD): fulfilling the need for longitudinal observational studies in COPD. *COPD.* 2014;11(2):125-132. https://doi.org/10.3109/15412555.2012.665520

12. Standardization of spirometry--1987 update. Statement of the American Thoracic Society. Am Rev Respir Dis. 1987;136(5):1285-1298. https://doi.org/10.1164/ajrccm/136.5.1285

13. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma.. *Am Rev Respir Dis.* 1987;136(1):225-244. https://doi.org/10.1164/ajrccm/136.1.225

14. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med.* 2002; 166(1):111-117. https://doi.org/10.1164/ajrccm.166.1.at1102

15. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med.* 2003;167(2):211-277. https://doi.org/10.1164/rccm.167.2.211

16. O'Donnell DE, Revill SM, Webb KA. Dynamic hyperinflation and exercise intolerance in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2001;164(5):770-777. https://doi.org/10.1164/ajrccm.164.5.2012122

17. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159(1):179-187. https://doi.org/10.1164/ajrccm.159.1.9712108

18. Fine JP, Gray, R.J. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999;94(446):496-509. https://doi.org/10.1080/01621459.1999.10474144

19. Kohl M, Plischke M, Leffondre K, Heinze G. PSHREG: a SAS macro for proportional and nonproportional subdistribution hazards regression. *Comput Methods Programs Biomed.* 2015; 118(2):218-233. https://doi.org/10.1016/j.cmpb.2014.11.009

20. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, 2022 report. GOLD website. Published 2022. Accessed July 2022. https://goldcopd.org/2022-gold-reports/

21. Provencher V, Mortenson WB, Tanguay-Garneau L, Belanger K, Dagenais M. Challenges and strategies pertaining to recruitment and retention of frail elderly in research studies: a systematic review. *Arch Gerontol Geriatr.* 2014;59(1):18-24. https://doi.org/10.1016/j.archger.2014.03.006

22. Gul RB, Ali PA. Clinical trials: the challenge of recruitment and retention of participants. *J Clin Nurs.* 2010;19(1-2):227-233. https://doi.org/10.1111/j.1365-2702.2009.03041.x

23. Robinson L, Adair P, Coffey M, Harris R, Burnside G. Identifying the participant characteristics that predict recruitment and retention of participants to randomised controlled trials involving children: a systematic review. *Trials.* 2016;17(1):294-311. https://doi.org/10.1186/s13063-016-1415-0

24. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol.* 2005;58(1):13-19. https://doi.org/10.1016/j.jclinepi.2004.05.006

25. Jacobsen E, Ran X, Liu A, Chang CH, Ganguli M. Predictors of attrition in a longitudinal population-based study of aging. *Int Psychogeriatr.* 2021;33(8):767-778. https://doi.org/10.1017/S1041610220000447

26. Eakin MN, Eckmann T, Dinglas VD, et al. Association between participant contact attempts and reports of being bothered in a national, longitudinal cohort study of ARDS survivors. *Chest.* 2020;158(2):588-595. https://doi.org/10.1016/j.chest.2020.02.050

27. Heins SE, Wozniak AW, Colantuoni E, et al. Factors associated with missed assessments in a 2-year longitudinal study of acute respiratory distress syndrome survivors. *BMC Med Res Methodol.* 2018;18(1):55-66. https://doi.org/10.1186/s12874-018-0508-8

28. Parkinson B, Meacock R, Sutton M, et al. Designing and using incentives to support recruitment and retention in clinical trials: a scoping review and a checklist for design. *Trials.* 2019;20(1):624-638. https://doi.org/10.1186/s13063-019-3710-z

29. Bonk J. A road map for the recruitment and retention of older adult participants for longitudinal studies. *J Am Geriatr Soc.* 2010;58 Suppl 2:S303-S307. https://doi.org/10.1111/j.1532-5415.2010.02937.x

30. Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR. The prevalence of limited health literacy. *J Gen Intern Med.* 2005;20:175-184. https://doi.org/10.1111/j.1525-1497.2005.40245.x

31. Nutbeam D. The evolving concept of health literacy. *Soc Sci Med.* 2008;67(12):2072-2078. https://doi.org/10.1016/j.socscimed.2008.09.050

32. Leak C, Goggins K, Schildcrout JS, et al. Effect of health literacy on research follow-up. *J Health Commun.* 2015;20(0): 83-91. https://doi.org/10.1080/10810730.2015.1058442