Latent constructs identified in older individuals who smoke cigarettes and are eligible for lung cancer screening: Factor analysis of baseline data from the PLUTO smoking cessation trial

Abbie Begnaud a,⁎, Steven S. Fu b, Bruce Lindgren c, Anne Melzer d, Alexander J. Rothman e, Kelsey Schertz f, David M. Vock f, Anne M. Joseph a

a University of Minnesota Department of Medicine, USA
b Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System, USA
c Masonic Cancer Center, Clinical and Translational Science Institute, USA
d Pulmonary and Critical Care Medicine, Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System, USA
e University of Minnesota, College of Liberal Arts, USA
f University of Minnesota School of Public Health, USA

Abstract

Introduction: Lung cancer screening (LCS) combined with smoking cessation intervention is currently recommended for older individuals with a history of heavy smoking. Tailoring tobacco treatment for this patient population of older, people who smoke (PWS) may improve cessation rates while efficiently using limited smoking cessation resources. Although some older people who smoke will need more intensive treatment to achieve sustained abstinence, others may be successful with less intensive treatment. A framework to identify them a priori would be helpful to distribute smoking cessation resources accordingly.

Methods: Baseline demographic, smoking, and health data are reported from a randomized clinical trial of longitudinal smoking cessation interventions delivered in the setting of LCS. Candidate variables were factor analyzed to identify latent factors, or constructs, to identify subgroups of older participants among the heterogeneous population of older people who smoke.

Results: We identified three factor-derived constructs: self-reported health status, heaviness of smoking, and nicotine dependence. Nicotine dependence was moderately correlated with both of the other two factors.

Conclusions: This factor analysis of baseline participant characteristics identified a set of latent constructs – based on a few practical clinical variables – that can be used to classify the heterogeneous population of older people who smoke to identify. We propose this framework to identify subgroups of people who smoke who might successfully quit with less intense treatment at the time of lung cancer screening.

1. Introduction

Lung cancer kills more people every year than breast, colon and prostate cancer combined [1]. Lung cancer screening (LCS) with annual low-dose computed tomography (LDCT) was initially shown to reduce prostate cancer combined [1]. Lung cancer screening (LCS) combined with smoking cessation intervention is currently recommended for older individuals with a history of heavy smoking. Tailoring tobacco treatment for this patient population of older, people who smoke (PWS) may improve cessation rates while efficiently using limited smoking cessation resources. Although some older people who smoke will need more intensive treatment to achieve sustained abstinence, others may be successful with less intensive treatment. A framework to identify them a priori would be helpful to distribute smoking cessation resources accordingly.

Methods: Baseline demographic, smoking, and health data are reported from a randomized clinical trial of longitudinal smoking cessation interventions delivered in the setting of LCS. Candidate variables were factor analyzed to identify latent factors, or constructs, to identify subgroups of older participants among the heterogeneous population of older people who smoke.

Results: We identified three factor-derived constructs: self-reported health status, heaviness of smoking, and nicotine dependence. Nicotine dependence was moderately correlated with both of the other two factors.

Conclusions: This factor analysis of baseline participant characteristics identified a set of latent constructs – based on a few practical clinical variables – that can be used to classify the heterogeneous population of older people who smoke to identify. We propose this framework to identify subgroups of people who smoke who might successfully quit with less intense treatment at the time of lung cancer screening.

Abbreviations: cds, cigarette dependency scale; cpd, cigarettes per day.
⁎ Corresponding author. Delaware St SE, MMC 276, Minneapolis, MN, 55455, USA.
E-mail address: abegnaud@umn.edu (A. Begnaud).

https://doi.org/10.1016/j.conctc.2022.100977
Received 4 December 2021; Received in revised form 19 July 2022; Accepted 7 August 2022
Available online 18 August 2022
2451-8654/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
National Cancer Institute Smoking Cessation at Lung Examination (SCALE) trials test various forms of smoking cessation assistance delivered in the context of LCS, however results are not yet available.

The prevalence of smoking decreases with age [12], yet older people who smoke (PWS) cigarettes may be less likely to express interest in quitting or report a recent quit attempt [13]. There is a small literature [14,15] on ‘older smokers’ and ‘hard-core smokers’ that address these observations. Older PWS may be nihilistic about the benefits of quitting [16] and hold erroneous beliefs about the benefits of quit support [17, 18]. However, a recent more representative online survey of older PWS showed most are aware of health risks of smoking and are interested in quitting [19]. Similarly, a large study of PWS over 50 years of age showed sustained abstinence rates at one year of 55% with extended treatment, especially among younger participants in this age range and those with lower measures of nicotine dependence at initiation, similar to other treatment cohorts [20]. A meta-analysis of smoking cessation interventions in individuals over 50 years old showed treatment efficacy with pharmacologic, non-pharmacologic and multimodal therapies [21]. Tailored cessation messages to older PWS can also be effective [22]. More intensive treatment (e.g., multiple modalities in combination, such as nicotine replacement along with counseling or prescription medications) increases chance of successful cessation [23,24]. However, older PWS are not a homogenous group and unknown factors may impact what resources are needed to successfully quit smoking. Supporting older PWS may require a nuanced approach but with adequate support, many can successfully quit smoking. Therefore, it would be helpful to better understand specific phenotypes of older PWS.

In 2013, the NIH announced a request for proposals to fund clinical trials studying smoking cessation in the setting of LCS. The Program for Lung Screening and Tobacco Treatment study (PLUTO, ClinicalTrials.gov Identifier: NCT02597491), part of the National Cancer Institute Smoking Cessation at Lung Examination (SCALE) collaboration [25], has completed participant enrollment. We report here baseline demographic, smoking, and health data of the PLUTO sample. We assessed a comprehensive set of data including personal smoking history, prior quit attempts, risk perceptions, measures of nicotine dependence and withdrawal, smoking urges, pain, alcohol use, social support, measures of depression and anxiety, and medical co-morbidities. The objective of this analysis is to identify variables that describe latent factors to characterize older PWS cigarettes and are eligible for LCS.

2. Methods

RCT Study Design. The study design was previously described in detail [26]. Briefly, PLUTO is a sequential, multiple assignment, randomized trial (SMART) conducted at three sites with established LCS programs. Eligible participants were currently smoking and met criteria for LCS – men and women aged 55 to 80 with a cumulative smoking history of 30 pack-years or more. All participants received tobacco longitudinal care (TLC) for one year, defined as phone coaching and over-the-counter (OTC) nicotine replacement therapy. Treatment response dictated randomization to continued TLC vs. TLC + Medication Therapy Management (MTM) delivered by pharmacists who could provide prescription smoking cessation medications (e.g., varenicline or bupropion) (for nonresponders) or randomization to continued TLC vs. TLC at lower frequency (for responders). The study was approved by the institutional review board at each site. Written informed consent was obtained from each participant.

Data Collection. Data collection at baseline included validated instruments to assess smoking behavior (Population Assessment of Tobacco and Health Study questionnaires), cigarette dependence (CD5-27), self-efficacy for smoking abstinence [28], nicotine withdrawal [29], smoking urges (QSU-B [30]), and readiness to quit [31]. In addition, we calculated pack-years (calculated cumulative smoke exposure variable that determines eligibility for lung cancer screening, determined by multiplying peak smoking intensity by years smoked), and assessed personality (Regulatory Focus Scale [32]), anxiety (GAD-7 [33]), depression (PHQ-9 [34]), alcohol use (AUDIT-C [35]), and medical comorbidities (Charlson Comorbidity Index [36]). Participants completed selected questions describing respiratory symptoms and health status, from the Short Form 12 Survey (SF-12) [37] and the Chronic Obstructive Pulmonary Disease (COPD) Assessment Test [38], validated measures of the impact of health on an individual’s daily life.

Variable selection for analysis. PLUTO investigators, who collectively have expertise in smoking cessation, lung cancer screening and health behavior theory, considered baseline data variables for inclusion in the model based on probable relationship to smoking cessation support needs and feasibility to measure during a typical clinical encounter for LCS.

Factor Analysis An exploratory factor analysis was performed to determine the number and nature of the underlying factors that contribute to the covariation in the data. The selected candidate variables underwent a factor analysis using squared multiple correlations as prior communality estimates. The principal factor method extracted factors, followed by a promax (oblique) rotation, which allowed the factors to be correlated. The number of factors retained was based on the scree plot in addition to achieving at least 3 variables per factor with higher factor loadings (>0.35). Subsequent to factor analysis, standardized Cronbach’s alphas were calculated for these high loading variables on each factor to verify reliability of these combinations of variables. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Inc., Cary NC).

3. Results

Demographic characteristics of the 643 participants from the PLUTO trial are shown in Table 1. Complete information on all 20 candidate variables was available for 627 participants.

The factor analysis resulted in three distinct factors (Table 2). Eleven variables were found to have a factor loading >0.35 on at least one factor, and the distribution of these responses is shown in Table 3. Factor

| Table 1 Participant characteristics. |
|------------------------------------|
| N = 643                            |

| Age (years), mean (SD) | 64.4 (5.8) |
| Gender, n (%)          |            |
| Female                | 229 (35.6) |
| Male                  | 414 (64.4) |
| Hispanic ethnicity, n (%) (n = 641) | 6 (0.9) |
| Race (n = 638)         |            |
| American Indian/Alaska Native | 8 (1.3) |
| Asian/Native Hawaiian/Pacific Islander | 2 (0.2) |
| Black                 | 32 (5.0) |
| White                 | 571 (89.5) |
| More than one race     | 25 (3.9) |
| Education, n (%) (n = 642) |            |
| 8th grade or less      | 5 (0.8) |
| Some high school       | 41 (6.4) |
| High school graduate   | 176 (27.4) |
| Vocational training    | 69 (10.7) |
| Associates degree/Some college | 311 (22.9) |
| Bachelor’s degree      | 93 (14.5) |
| Graduate/Professional school | 47 (7.3) |
| Marital status, n (%) (n = 641) |            |
| Never Married          | 66 (10.3) |
| Married/Domestic partner | 291 (45.4) |
| Widowed                | 58 (9.1) |
| Separated              | 16 (2.5) |
| Divorced               | 210 (32.8) |
| Household income, n (%) (n = 589) |            |
| <$3,000                | 82 (13.9) |
| $3,000 - $24,999       | 139 (23.6) |
| $25,000 - $64,999      | 174 (29.5) |
| $65,000 - $100,000     | 126 (21.4) |
| >$100,000              | 68 (11.5) |
4. Discussion

We report the baseline characteristics of participants in the PLUTO

Table 2
Results of factor analysis: Factor loadings (standardized regression coefficients)∗.

| Variable                          | Health Status Self-Report (Factor 1) | Smoking Heaviness (Factor 2) | Nicotine Dependence and Self-Efficacy for Quitting (Factor 3) |
|-----------------------------------|--------------------------------------|-----------------------------|-----------------------------------------------------------|
| Time to 1st cigarette of the day  | 0.00                                 | 0.20                        | 0.36                                                      | 0.225 |
| Cigarettes smoked at peak level   | 0.08                                 | 0.55                        | –0.05                                                    | 0.299 |
| Average                           | 0.01                                 | 0.85                        | 0.00                                                     | 0.722 |
| Pack-years                        | 0.62                                 | 0.08                        | –0.10                                                    | 0.359 |
| Health Status Q1 (General Health) | 0.65                                 | 0.12                        | –0.12                                                    | 0.402 |
| Functional Status score           | –0.02                                | 0.02                        | 0.50                                                     | 0.251 |
| CDS Q2: Difficulty of quitting    | –0.08                                | 0.04                        | 0.58                                                     | 0.330 |
| CDS Q3: urge to smoke             | –0.03                                | 0.75                        | 0.13                                                     | 0.655 |
| Average CPD                       | 0.76                                 | –0.05                       | 0.12                                                     | 0.646 |
| PHQ-8 score                       | 0.66                                 | –0.08                       | 0.12                                                     | 0.496 |
| GAD score                         | 0.08                                 | –0.04                       | 0.58                                                     | 0.357 |

∗Factor loadings in Table 2 are standardized regression coefficients on a scale of –1 to 1 and indicate the linear “weight” in the corresponding factor associated with each variable. The variables with factor loadings greater than 0.35 are noted in bold and served to aid in interpretation of the factors.

Table 3
Distribution of responses to variables loading on Factors 1, 2 and 3.

| Variable                          | n (%) |
|-----------------------------------|-------|
| Cigarettes smoked at peak level   |       |
| 5–10                              | 18 (2.8) |
| 11–20                             | 258 (40.1) |
| 21–30                             | 138 (21.5) |
| 31–40                             | 149 (23.2) |
| 41+                               | 80 (12.4) |
| Average pack-years (n = 643)      |       |
| 9.5–19.99                         | 95 (15.0) |
| 20.0–24.99                        | 112 (17.7) |
| 25.0–29.99                        | 160 (25.2) |
| 30.0–34.99                        | 110 (17.4) |
| 35.0–39.99                        | 56 (8.8) |
| 40.0–44.99                        | 101 (15.9) |
| Cigarettes per day                |       |
| Mean (SD)                         | 17.6 (8.5) |
| Median [Min/Max]                  | 20.0 [1/50] |

Factor 2: Smoking Heaviness

1. Health Status Self-Report, loaded four variables related to patient report of health status: general health from the SF-12; activity limitation from the COPD assessment test; depression from the PHQ-8, and anxiety from the GAD. Factor 2, Smoking Heaviness, loaded three variables related to smoking intensity: average current cigarettes per day, average pack-years, and peak cigarettes per day. Factor 3, Dependence on Smoking, loaded with four variables related to nicotine dependence (one Cigarette Dependency Scale item: “Usually, how soon after waking up do you smoke your first cigarette?”), two Likert scale measures “For you, quitting smoking for good would be very easy to impossible” and “After a few hours without smoking, I feel an irresistible urge to smoke” (totally agree to totally disagree), and self-efficacy for abstinence. The remaining candidate variables were eliminated on the basis of low factor loading (standardized regression coefficients): the Charlson comorbidity index, previous quit attempts, prior use of quitting aids (programs, NRT, prescriptions), readiness to quit, risk perception, presence of other household smokers, age of smoking initiation, and alcohol consumption.

The Dependence on Smoking Factor was moderately, positively correlated with Health Status Self-Report (r = 0.35) and Smoking Heaviness (r = 0.39), but the latter two factors were not correlated with each other.

The internal consistency of each factor, as measured by the standardized Cronbach’s alpha, is shown in Table 4 and includes the correlation of each variable with the standardized total score of the factor.

4. Discussion

We report the baseline characteristics of participants in the PLUTO
trial, older people who smoke and are eligible for lung cancer screening. We use exploratory factor analysis to identify clusters of baseline variables and their relationships. The factor analysis suggests three constructs that may have implications for how people respond to cessation treatment: self-report of health status, heaviness of smoking, and dependence upon smoking. The factors are moderately correlated with one another, with the exception of health status and heaviness of smoking. The correlation of each variable with its specified factor is reasonably strong. We can postulate what types of treatment might benefit individuals who fall into each of the phenotypes that can be constructed from the three factors. For example, an older PWS who reports relatively poorer health (Factor 1) might benefit from counseling about how quitting smoking can positively impact their health status. On the other hand, an individual who has more heavy smoking (Factor 2), might benefit from aggressive nicotine replacement therapy and varenicline. Another type of older PWS who reports a greater degree of dependence upon smoking (Factor 3) might benefit from counseling and motivational interviewing to improve confidence in quitting along with aggressive pharmacological management.

A number of variables we considered did not load sufficiently onto any factor, including the Charlson Comorbidity Index (CCI), prior quit attempts, previous use of quitting aids, readiness to quit, risk perception, presence of other household smokers, age of smoking initiation, and alcohol consumption. Perhaps these factors are outweighed by the LCS experience (i.e., possibly or actually being diagnosed with lung cancer). On the other hand, it might be that these characteristics are simply unrelated to the three factors that we identified.

While some studies suggest that older PWS cigarettes are less likely to quit, they can achieve sustained abstinence with adequate treatment. PLUTO enrolled older individuals with a heavy smoking history who continued to smoke cigarettes and were eligible for LCS.

Limitations of this study include that smoking cessation outcome data are pending, and we enrolled a primarily non-Hispanic White—identifying sample of patients. While the racial demographics of Minnesota and this study are not representative of the United States as a whole, they are similar to many other states, especially rural states in the northern United States where smoking rates are moderately high. Our sample has higher educational attainment and household income than the general population of PWS in the US which is a reflection of Minnesota being the second highest ranked state for educational attainment. Similarly, our sample has relatively higher household income distribution than the general population of PWS; Minnesota median household income is also above the national median. In addition, participants in this study expressed willingness to make a quit attempt within 12 weeks of study enrollment, thus excluding those who were completely uninterested in stopping cigarette smoking. Finally, since study enrollment was completed, LCS eligibility guidelines have changed, with a recently published recommendation to expand eligibility to slightly younger individuals with slightly lower total smoking exposure (minimum 20 pack-years instead of 30) [4]; the results may not be generalizable to the expanded eligible population.

The long-term goal of this analysis is to identify a priori latent constructs to describe subgroups of older PWS. We set out to identify baseline variables collected during the PLUTO trial that group together, to describe latent constructs which might predict response to cessation intervention components for older PWS. We welcome replication of this analysis in samples representative of other geographic regions and the general population of LCS-eligible PWS in the US. Because routine clinical collection of all the data obtained in our study is not practical in the context of LCS, this approach has the potential to identify a few select variables that are practical to collect routinely. If we could identify individuals who need different forms of intensive therapy, it can be directed to those individuals rather than those who would be successful with only minimal intervention. Tools to identify patients, at the time of LCS, who smoke and will need more intensive treatment can prioritize the distribution of finite quitting support resources while maximizing the benefit of screening through smoking cessation.

Offering smoking cessation at the time of LCS is important to maximizing the benefit of screening. A shared decision-making visit is required at the time of LCS order and is an ideal opportunity to provide smoking cessation treatment. Personalized smoking cessation treatment at the time of LCS may have potential to efficiently use limited cessation resources (e.g., intensive counseling and prescription medications) while maximizing quit rates and improving all-cause mortality of individuals undergoing LCS.

5. Implications

This study seeks to describe phenotypes of older people who smoke from a smoking cessation clinical trial using factor analysis of baseline data. Ability to classify such individuals could be used to determine who would benefit from more intensive smoking cessation treatment and who would successfully quit with less intensive treatment. This is particularly impactful in lung cancer screening programs to prioritize scarce cessation assistance resources.

Funding

This study was funded by the National Cancer Institute (R01CA196873 and 2T32CA163184), National Institutes of Health. This material is the result of work supported with resources and the use of facilities at the Minneapolis VA HSR&D Center for Chronic Disease Outcomes Research. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, the National Institutes of Health, or the United States government.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We appreciate the efforts of Ms. Mariah Branson coordinating study team activities, regulatory requirements, and participant interactions.
