# INTRODUCTION

# METHODS

## Study Design and Setting

## Participants

2066 subjects

## Variables

Latent trajectories were determined using percent change from baseline in average cigarettes per day (CPD) at weeks 2, 10, 18, and 26. CPD was determined using responses to the question “how many cigarettes do you smoke/day on average?” When responses to that question were unavailable, values from “how many cigarettes do you smoke/week on average” divided by 7 days per week were used instead. If a participant reported they had stopped smoking, CPD was set to 0. 7 subjects were missing baseline CPD values and were not included in analyses. Of these, 1884 have at least one post-baseline average CPD value.

The baseline variables used to predict latent class were age started smoking, longest period without smoking, number of times tried to quit smoking, Fagerstrom nicotine dependency score (FTND), intention to quit, length of time since last quit attempt, experience of anxiety in the last 24 hours, experience of depression in the last 24 hours, SF-36, carbon monoxide parts per million (CO ppm), relief from smoking, study site, and parent trial treatment group (26 variables in all). Of the 1884 subjects in consideration, 101 (5.4%) were missing 7 or more baseline measures, and were removed from analyses. The remaining 1783 (94.6%) had fewer than 5 missing measures.

Two SF-36 sub score (General and Physical Functioning) values were 20.75% and 21.82% missing respectively and were discarded. Of the remaining missing values (< 7%), many were correlated with age, sex, and study site. To control for these differences, ages were binned into 4 equally sized groups, and all remaining baseline variables were imputed within age-sex-site combinations, using the mean and mode of those combinations where appropriate.

## Data sources

## Bias

## Study size

## Quantitative Variables

## Statistical Methods

### Analysis 1: Trajectories in cigarettes per day over time

A latent class mixture model (1–3) was used to determine longitudinal trends in cigarettes per day from baseline assessed at trial weeks 2, 10, 18, and 26. The primary dependent variable was percent change average CPD (e.g., a participant smoking 90% of their baseline level at Week 2 received a value of -10% for Week 2).

Models for classes greater than 1 were initialized using maximum likelihood estimates of the 1-class model. Grid search methods were performed for subsequent models, selecting the minimal Bayesian Information Criterion at each step. The optimal model was selected using the minimal BIC at the “knee of the curve” such that additional trajectories did not provide a substantially better model.

Posterior classification was then used to assign subjects to each latent class using Bayes’ theorem and the maximum probability of belonging to a particular latent class given the information (i.e., percent change in CPD at leach week) collected in the longitudinal model.

### Analysis 2: Predicting longitudinal trajectories in CPD using baseline variables

Elastic net logistic regression (5,6) was used to build predictive models for each latent class using all baseline characteristics as features (i.e., independent variables), and class membership as the target (i.e., dependent variable). Class membership was treated as a binary outcome.

Initially, the data were split into 80% training and 20% testing sets. The training data was then used to tune parameters for elastic net logistic regression models. Using the training data, mixture and penalty parameters were tuned using 10-fold cross validation, maximizing receiver operator characteristic area-under-the-curve (ROC AUC). Parameters from the model whose ROC AUC results were within one standard error of the optimal cross-validated results were selected (4). This procedure was repeated using each latent class as the target.

Using the model parameters selected for each class respectively, logistic regression models were fit and evaluated using internal 5-fold cross validation, where 80% of the training data was used for model training and 20% for internal validation, assessing the model’s performance out-of-sample. Again, classification accuracy was measured using ROC AUC. Finally, to confirm each model’s generalizability, an ROC AUC score was derived from training a model on all the training data and predicting the initial testing set. Null ROCs were computed for each predicted class, and overall statistical significance was assessed using to ROC AUC’s equivalence with the Mann-Whitney *U*-statistic (7).

### Analysis 3: Which trajectories in CPD predict smoking cessation?

A linear elastic net regression model was used to predict CO values at 1-year follow-up using (1) all baseline characteristics from Analysis 2, and (2) Latent Class (from Analysis 1) in addition to those predictors. The dependent variable was CO in parts per million. Predicted CO values <= 11 ppm were considered indicative of quitting smoking.

Each regression model was fit using the procedures outlined in Analysis 2, with model performance measured using the coefficient of determination, *R*2.

# RESULTS

## Participants

Smoking trajectories were modeled using a total of 5,578 observations (Weeks 2, 10, 19, and 26 had 1764, 1148, 1384, 1282 observations respectively). Distributions of average CPD changes from baseline are show in Figure 1.

A 3-class heterogeneous linear mixed model was selected (Maximum log-likelihood = 11.81, BIC = 43.75) (see Figure 3). Longitudinal fixed effects (week-by-class) were statistically significant (Wald statistics: Class 1 = -15.71, Class 2 = -4.20, Class 3 = 14.35, p’s < .001). Class-membership fixed effects showed Class 3 to be significantly different from Class 1 (Wald = 12.32, p = .0000), but not from Class 2 (Wald = 0.384, p = .7013). Average weekly CPD values for each class are shown in Figure 2.

Each subject was assigned to the class according to the maximum probability of belonging to that class. Characteristics associated with participants in each class are given in Table 1.

## Descriptive data

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Class 1** | **Class 2** | **Class 3** |
| N (%) | 216 (11.78) | 1003 (54.69) | 615 (33.53) |
| Study Site (%)  Australia  Denmark  Germany  Switzerland  USA |  |  |  |
| Sex (% male) |  |  |  |
| Age (mean (SD)) |  |  |  |
| Age started smoking (mean (SD)) |  |  |  |
| Baseline CPD (mean (SD)) |  |  |  |

## Predicting class membership using baseline characteristics

Elastic net logistic regression models were used to predict each latent class in turn.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean CV AUC | SE | Test Set AUC | p |
| Class 1 | 0.62 | 0.02 | 0.35 | 0.4170 |
| **Class 2** | **0.49** | **0.01** | **0.49** | **0.0000** |
| Class 3 | 0.56 | 0.01 | 0.42 | 0.0932 |
| **Contrast 1:2** | **0.56** | **0.04** | **0.40** | **0.0067** |
| **Contrast 1:3** | **0.67** | **0.02** | **0.30** | **0.0378** |
| **Contrast 2:3** | **0.54** | **0.01** | **0.45** | **0.0015** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| term | class\_1 | class\_2 | class\_3 | c12 | c13 | c23 |
| n\_quit\_attempts\_X2.TO.5.TIMES | 0.09083554 | 0 | -0.06661 | 0.087023 | 0.08724459 | 0.0417606 |
| n\_quit\_attempts\_X6.TO.10.TIMES | -0.1122922 | 0 | 0.04559821 | -0.1504042 | -0.0959373 | -0.0191311 |
| n\_quit\_attempts\_MORE.THAN.10.TIMES | 0.04208191 | 0 | 0.02821023 | 0.08980203 | -0.0186987 | -0.0268347 |
| n\_quit\_attempts\_NEVER | 2.71598988 | 0 | -0.8670753 | 2.96056622 | 1.15982438 | 0.39839577 |
| n\_quit\_attempts\_ONCE | -0.2127105 | -0.0015463 | 0.10797761 | -0.1522122 | -0.1282475 | -0.0692777 |
| depression\_EXTREMELY.SO | -1.617541 | 0 | 0.23744196 | -1.663674 | -0.8864929 | 0.00924579 |
| depression\_MODERATELY.SO | 1.38807398 | -0.001474 | -0.1005733 | 1.55245564 | 0.64935375 | 0.00523217 |
| depression\_NOT.AT.ALL | -0.0953516 | 0 | -0.0559574 | -0.0917189 | -0.0181885 | 0.02593849 |
| depression\_SOMEWHAT | -0.2179384 | 0 | -0.0214194 | -0.294082 | -0.075528 | 0.03894837 |
| depression\_VERY.MUCH.SO | -0.4194918 | -0.2409415 | 0.57037932 | -0.0939446 | -0.5395077 | -0.3518432 |
| int\_to\_quit\_NOT.AT.ALL | 0.74201472 | 0.01658298 | -0.2939604 | 0.70651982 | 0.50450422 | 0.14457113 |
| int\_to\_quit\_A.LITTLE | 0.07270863 | 0 | -0.0224927 | 0.10372176 | 0.04185542 | 0.00738965 |
| int\_to\_quit\_SOMEWHAT. | -0.1487701 | 0.06654428 | -0.072784 | -0.1930798 | -0.0386506 | 0.08709357 |
| int\_to\_quit\_A.LOT | -0.2875103 | -0.0879574 | 0.22435215 | -0.2098718 | -0.2685421 | -0.1594118 |
| site\_aus | -0.1140837 | 0 | -0.0818969 | -0.1659756 | 0.02400379 | 0.07714574 |
| site\_den | -0.1277005 | -0.1997882 | 0.30491884 | 0.0427225 | -0.1213226 | -0.199151 |
| site\_ger | -0.5479721 | 0 | 0.12324822 | -0.4989116 | -0.3455618 | -0.0144641 |
| site\_swi | 0.02604069 | 0 | 0.01943407 | -0.0201764 | 0.03114743 | -0.0215428 |
| site\_usa | 0.64017114 | 0 | -0.2937911 | 0.59641342 | 0.34006222 | 0.12192323 |
| ts\_last\_quit\_attempt\_X.12.MONTHS | 0.15779036 | 0 | -0.0611664 | 0.13780583 | 0.0265736 | -0.0032441 |
| ts\_last\_quit\_attempt\_X.12.24.MONTHS | -0.6098917 | 0.02565792 | 0.07338725 | -0.6336356 | -0.2240519 | 0.05739329 |
| ts\_last\_quit\_attempt\_X.6.12.MONTHS | 0.07736792 | -0.0474372 | 0.09372056 | 0.11686155 | 0.03559135 | -0.0658307 |
| ts\_last\_quit\_attempt\_X0.6.MONTHS | -0.0746874 | 0.030375 | -0.0673504 | -0.1096685 | 0.00824995 | 0.06586073 |
| trt\_recode\_active | -0.4211759 | -0.0283077 | 0.21506195 | -0.3672416 | -0.3743668 | -0.1188308 |
| trt\_recode\_placebo | 0.41625153 | 0.0275141 | -0.2118382 | 0.36236456 | 0.37421129 | 0.11872062 |
| anxiety\_EXTREMELY.SO | 0.34184907 | -0.1087973 | 0.31020896 | 0.86676771 | -0.2074545 | -0.2665516 |
| anxiety\_MODERATELY.SO | -0.2514134 | 0 | 0.14288694 | -0.1412051 | -0.1966089 | -0.0617203 |
| anxiety\_NOT.AT.ALL | -0.2039951 | 0 | 0.04517294 | -0.2142738 | -0.1951859 | -0.0276418 |
| anxiety\_SOMEWHAT | 0.27697972 | 0.0069893 | -0.1382494 | 0.2307771 | 0.27203916 | 0.08240073 |
| anxiety\_VERY.MUCH.SO | 0.37810407 | 0 | -0.0428168 | 0.30993928 | 0.38974544 | 0.01135126 |
| longest\_period\_wo\_smoking\_X1 | 0.36446974 | 0 | -0.0518387 | 0.33622153 | 0.2733531 | 0.0014645 |
| longest\_period\_wo\_smoking\_X2 | -0.0352316 | 0.0674906 | -0.0971957 | -0.1143424 | 0.04186514 | 0.08093823 |
| longest\_period\_wo\_smoking\_X3 | -0.1003065 | 0 | 0.06465493 | -0.0621927 | -0.0579676 | -0.0435583 |
| longest\_period\_wo\_smoking\_X4 | -0.177939 | 0 | 0.06848898 | -0.1195673 | -0.1927244 | -0.0338864 |
| rsq\_calming | -0.0704077 | 0.01709599 | -0.1092987 | -0.1219537 | 0.0310347 | 0.06216504 |
| rsq\_pepping\_up\_eff | 0.1273199 | 0 | -0.001169 | 0.14752421 | 0.05678373 | 0.00314693 |
| rsq\_last\_cig\_exp\_NEUTRAL | 0.06754261 | 0 | -0.0775738 | 0.06797527 | 0.0648932 | 0.05418231 |
| rsq\_last\_cig\_exp\_SOMEWHAT.PLEASANT | -0.1767582 | 0 | 0.04787494 | -0.1748697 | -0.1188128 | -0.0211402 |
| rsq\_last\_cig\_exp\_SOMEWHAT.UNPLEASANT | -0.144889 | -0.1148284 | 0.22930518 | -0.0188374 | -0.260153 | -0.1857501 |
| rsq\_last\_cig\_exp\_VERY.PLEASANT | 0.20002365 | 0 | -0.0365918 | 0.16977112 | 0.15558709 | 0.0177161 |
| rsq\_last\_cig\_exp\_VERY.UNPLEASANT | 0.24591935 | 0.0561327 | -0.3104261 | 0.02894335 | 0.30247759 | 0.20663975 |
| age | -0.2169849 | 0 | 0.08005624 | -0.2349869 | -0.139403 | -0.0409181 |
| ftnd | 0.20616428 | 0.00146812 | -0.1427872 | 0.14937109 | 0.16479348 | 0.05786812 |
| sex\_FEMALE | 0.15480213 | 0 | -0.0702501 | 0.13084236 | 0.14074115 | 0.04272232 |
| sex\_MALE | -0.1537316 | 0 | 0.06997214 | -0.1299709 | -0.1406992 | -0.0426997 |
| sf\_physheal | -0.0330821 | 0 | -0.023658 | -0.0667443 | -0.0019238 | 0.00846156 |
| sf\_emoprob | -0.0817343 | 0.03949433 | -0.0558954 | -0.1043903 | -0.0104177 | 0.04373608 |
| sf\_socfunc | -0.0607972 | 0 | 0.06163312 | -0.0341771 | -0.0386115 | -0.0219443 |
| sf\_pain | 0.14916814 | -0.0651846 | 0.06360217 | 0.21685324 | 0.03888745 | -0.0598524 |
| sf\_emowell | 0.04972194 | 0 | -0.0163039 | 0.04918337 | -0.0090438 | 0.00629864 |
| age\_started\_smoking | -0.0945436 | -0.0282678 | 0.11003129 | -0.0248336 | -0.1505748 | -0.065473 |
| cpd | 0.0332871 | -0.0256053 | 0.10542684 | 0.08365263 | 0.04269098 | -0.0466483 |
| co | 0.03824008 | 0 | -0.0385751 | 0.01555389 | 0.05412722 | 0.02219046 |

|  |  |  |  |
| --- | --- | --- | --- |
| Baseline Characteristic | class\_1 (Beta) | class\_2 (Beta) | class\_3 (Beta) |
| n\_quit\_attempts\_X2.TO.5.TIMES | 0.09083554 | 0 | -0.06661 |
| n\_quit\_attempts\_X6.TO.10.TIMES | -0.1122922 | 0 | 0.04559821 |
| n\_quit\_attempts\_MORE.THAN.10.TIMES | 0.04208191 | 0 | 0.02821023 |
| **n\_quit\_attempts\_NEVER** | **2.71598988** | **0** | **-0.8670753** |
| n\_quit\_attempts\_ONCE | -0.2127105 | -0.0015463 | 0.10797761 |
| **depression\_EXTREMELY.SO** | **-1.617541** | **0** | **0.23744196** |
| **depression\_MODERATELY.SO** | **1.38807398** | **-0.001474** | **-0.1005733** |
| depression\_NOT.AT.ALL | -0.0953516 | 0 | -0.0559574 |
| depression\_SOMEWHAT | -0.2179384 | 0 | -0.0214194 |
| **depression\_VERY.MUCH.SO** | **-0.4194918** | **-0.2409415** | **0.57037932** |
| int\_to\_quit\_NOT.AT.ALL | 0.74201472 | 0.01658298 | -0.2939604 |
| int\_to\_quit\_A.LITTLE | 0.07270863 | 0 | -0.0224927 |
| int\_to\_quit\_SOMEWHAT. | -0.1487701 | 0.06654428 | -0.072784 |
| int\_to\_quit\_A.LOT | -0.2875103 | -0.0879574 | 0.22435215 |
| site\_aus | -0.1140837 | 0 | -0.0818969 |
| site\_den | -0.1277005 | -0.1997882 | 0.30491884 |
| **site\_ger** | **-0.5479721** | **0** | **0.12324822** |
| site\_swi | 0.02604069 | 0 | 0.01943407 |
| **site\_usa** | **0.64017114** | **0** | **-0.2937911** |
| ts\_last\_quit\_attempt\_X.12.MONTHS | 0.15779036 | 0 | -0.0611664 |
| ts\_last\_quit\_attempt\_X.12.24.MONTHS | -0.6098917 | 0.02565792 | 0.07338725 |
| ts\_last\_quit\_attempt\_X.6.12.MONTHS | 0.07736792 | -0.0474372 | 0.09372056 |
| ts\_last\_quit\_attempt\_X0.6.MONTHS | -0.0746874 | 0.030375 | -0.0673504 |
| **trt\_recode\_active** | **-0.4211759** | **-0.0283077** | **0.21506195** |
| **trt\_recode\_placebo** | **0.41625153** | **0.0275141** | **-0.2118382** |
| **anxiety\_EXTREMELY.SO** | **0.34184907** | **-0.1087973** | **0.31020896** |
| anxiety\_MODERATELY.SO | -0.2514134 | 0 | 0.14288694 |
| anxiety\_NOT.AT.ALL | -0.2039951 | 0 | 0.04517294 |
| **anxiety\_SOMEWHAT** | **0.27697972** | **0.0069893** | **-0.1382494** |
| **anxiety\_VERY.MUCH.SO** | **0.37810407** | **0** | **-0.0428168** |
| longest\_period\_wo\_smoking\_X1 | 0.36446974 | 0 | -0.0518387 |
| longest\_period\_wo\_smoking\_X2 | -0.0352316 | 0.0674906 | -0.0971957 |
| longest\_period\_wo\_smoking\_X3 | -0.1003065 | 0 | 0.06465493 |
| longest\_period\_wo\_smoking\_X4 | -0.177939 | 0 | 0.06848898 |
| rsq\_calming | -0.0704077 | 0.01709599 | -0.1092987 |
| rsq\_pepping\_up\_eff | 0.1273199 | 0 | -0.001169 |
| rsq\_last\_cig\_exp\_NEUTRAL | 0.06754261 | 0 | -0.0775738 |
| rsq\_last\_cig\_exp\_SOMEWHAT.PLEASANT | -0.1767582 | 0 | 0.04787494 |
| rsq\_last\_cig\_exp\_SOMEWHAT.UNPLEASANT | -0.144889 | -0.1148284 | 0.22930518 |
| rsq\_last\_cig\_exp\_VERY.PLEASANT | 0.20002365 | 0 | -0.0365918 |
| rsq\_last\_cig\_exp\_VERY.UNPLEASANT | 0.24591935 | 0.0561327 | -0.3104261 |
| age | -0.2169849 | 0 | 0.08005624 |
| ftnd | 0.20616428 | 0.00146812 | -0.1427872 |
| sex\_FEMALE | 0.15480213 | 0 | -0.0702501 |
| sex\_MALE | -0.1537316 | 0 | 0.06997214 |
| sf\_physheal | -0.0330821 | 0 | -0.023658 |
| sf\_emoprob | -0.0817343 | 0.03949433 | -0.0558954 |
| sf\_socfunc | -0.0607972 | 0 | 0.06163312 |
| sf\_pain | 0.14916814 | -0.0651846 | 0.06360217 |
| sf\_emowell | 0.04972194 | 0 | -0.0163039 |
| age\_started\_smoking | -0.0945436 | -0.0282678 | 0.11003129 |
| cpd | 0.0332871 | -0.0256053 | 0.10542684 |
| co | 0.03824008 | 0 | -0.0385751 |

## Predicting smoking cessation using latent class

Of the 1783 participants assigned to latent classes, 911 had CO values at the 1-year follow-up point (6 months following the conclusion of the trial). Using a quit threshold of CO <= 11ppm, 261 (28.6%) had quit smoking at this point. More specifically, 66.9% (n = 95) of subjects in Class 1 had quit, 24.3% (n = 120) in Class 2, and 16.7% (n = 46) in Class 3.

An elastic net linear regression model was used to predict 1-year CO values using baseline characteristics alone, achieving an average cross-validated (mean [SE]) R2 = 0.22 [0.03], with R2 = 0.11 on the testing set. Adding latent class as a predictor improved average cross-validated prediction performance by 8% (CV R2 = 0.30 [0.03], test R2 = 0.18). Each showed greater performance than predicting 1-year CO values from predictors while limiting subjects to those within each class respectively.

Using the initial 20% testing data set to make predictions of 1-year follow-up CO values, subjects whose predicted values were > 11ppm were labelled smokers, and those with values equal to or below 11ppm non-smokers. This was also done with the actual 1-year CO values to assess the linear model’s ability to predict those who quit smoking. 138/186 (74.2%) of participants were accurately identified according to whether they quit smoking (see Figure 4).

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Mean CV *R*2 | CV *R*2 SE | Test *R*2 |
| Predictors + Class | 0.30 | 0.03 | 0.18 |
| Predictors Alone | 0.22 | 0.03 | 0.11 |
| Predictors w/in Class 1 Alone | 0.11 | 0.03 | 0.01 |
| Predictors w/in Class 2 Alone | 0.27 | 0.03 | 0.17 |
| Predictors w/in Class 3 Alone | 0.23 | 0.04 | 0.18 |

|  |  |
| --- | --- |
| Baseline Characteristic | Beta |
| **class\_X1** | **-8.5995268** |
| class\_X2 | 0 |
| **class\_X3** | **2.25842963** |
| **anxiety\_EXTREMELY.SO** | **5.9909219** |
| anxiety\_MODERATELY.SO | -0.4402322 |
| anxiety\_NOT.AT.ALL | 0 |
| anxiety\_SOMEWHAT | 0 |
| **anxiety\_VERY.MUCH.SO** | **-1.2045498** |
| **depression\_EXTREMELY.SO** | **-4.4486877** |
| depression\_MODERATELY.SO | 0 |
| depression\_NOT.AT.ALL | 0.35616685 |
| depression\_SOMEWHAT | 0 |
| depression\_VERY.MUCH.SO | 0.24587718 |
| **co** | **3.4610883** |
| **cpd** | **2.06709886** |
| site\_aus | 0 |
| site\_den | 0 |
| site\_ger | -1.8661252 |
| site\_swi | 1.17586879 |
| site\_usa | 0 |
| longest\_period\_wo\_smoking\_X1 | 0.0166343 |
| longest\_period\_wo\_smoking\_X2 | 0.46570704 |
| **longest\_period\_wo\_smoking\_X3** | **-1.638884** |
| longest\_period\_wo\_smoking\_X4 | 0 |
| n\_quit\_attempts\_X2.TO.5.TIMES | -0.1680086 |
| n\_quit\_attempts\_X6.TO.10.TIMES | 0 |
| **n\_quit\_attempts\_MORE.THAN.10.TIMES** | **1.11133142** |
| n\_quit\_attempts\_NEVER | 0 |
| n\_quit\_attempts\_ONCE | 0 |
| int\_to\_quit\_NOT.AT.ALL | -0.0807399 |
| int\_to\_quit\_A.LITTLE | 0 |
| int\_to\_quit\_SOMEWHAT. | -0.9687829 |
| int\_to\_quit\_A.LOT | 0 |
| ts\_last\_quit\_attempt\_X.12.MONTHS | -0.8546351 |
| ts\_last\_quit\_attempt\_X.12.24.MONTHS | -0.0307488 |
| ts\_last\_quit\_attempt\_X.6.12.MONTHS | 0 |
| ts\_last\_quit\_attempt\_X0.6.MONTHS | 0.88703307 |
| ftnd | 0.50221398 |
| trt\_recode\_active | -0.4625528 |
| trt\_recode\_placebo | 0.45371203 |
| sf\_physheal | 0 |
| sf\_emoprob | 0 |
| sf\_socfunc | -0.1040706 |
| sf\_pain | -0.2298105 |
| sf\_emowell | 0 |
| rsq\_calming | 0 |
| rsq\_pepping\_up\_eff | -0.1408797 |
| rsq\_last\_cig\_exp\_NEUTRAL | 0.05297322 |
| rsq\_last\_cig\_exp\_SOMEWHAT.PLEASANT | 0 |
| rsq\_last\_cig\_exp\_SOMEWHAT.UNPLEASANT | -0.1104583 |
| rsq\_last\_cig\_exp\_VERY.PLEASANT | 0 |
| rsq\_last\_cig\_exp\_VERY.UNPLEASANT | 0 |
| age | -0.0868007 |
| age\_started\_smoking | -0.0691899 |
| sex\_FEMALE | 0 |
| sex\_MALE | 0 |

# DISCUSSION

In summary, there were three distinct trajectories in smoking behavior throughout the trial, describing smokers who initially reduced and nearly eliminated their smoking (Class 1), reduced by approximately half and remained at that level throughout the trial (Class 2), and those who reduced initially but reverted to cigarette use near their baseline levels (Class 3). Predictive modelling showed relatively strong positive associations between Class 1 and never having tried to quit smoking before, as well as some levels of anxiety and depression. In particular, smokers who had never tried to quit before were substantially more likely to be in Class 1 compared to Class 2, rather than Class 1 compared to Class 3, suggesting quit attempts may have a role in differentiating smoking trajectories.

Paradoxically, while “moderate” depression in the past 24 hours of baseline was predictive of membership to Class 1, those reporting “extreme” depression were unlikely to receive that classification, with an even greater magnitude.

Finally, a regression model predicting CO values 6 months after the trial’s conclusion showed stronger performance with access to latent class as opposed to baseline characteristics alone, suggesting smoking trajectories may have implications for predicting smoking cessation. Those in Class 1 and who reported moderately high anxiety as well as “extreme” depression tended to have lower 1-year CO values, while those in Class 3 and who had higher baseline CO and CPD values, reported having tried to quit smoking 10 or more times, or who reported “extreme” anxiety tended to have higher CO values.

# FIGURES

Figure Distributions of CPD values at each time point.

Timeline

Description automatically generated

Figure Three-class smoking trajectories.

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

Figure Latent class mixture model goodness of fit.

Chart, scatter chart

Description automatically generated

Figure Predicting 1-year follow-up CO values using baseline predictors and latent class. 11ppm CO threshold is indicated with dashed lines.

1. Proust-Lima C, Philipps V, Liquet B. Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The *R* Package **lcmm**. J Stat Softw [Internet]. 2017 [cited 2022 Nov 14];78(2). Available from: http://www.jstatsoft.org/v78/i02/

2. Hagenaars JA, McCutcheon AL, editors. Applied latent class analysis. Cambridge ; New York: Cambridge University Press; 2002. 454 p.

3. Proust-Lima C, Philipps V, Diakite A, Liquet B. lcmm: Extended Mixed Models Using Latent Classes and Latent Processes [Internet]. 2022. Available from: https://cran.r-project.org/package=lcmm

4. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification And Regression Trees [Internet]. 1st ed. Routledge; 2017 [cited 2022 Nov 23]. Available from: https://www.taylorfrancis.com/books/9781351460491

5. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1):1–22.

6. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent. J Stat Softw. 2011;39(5):1–13.

7. Mason SJ, Graham NE. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. Q J R Meteorol Soc. 2002 Jul 15;128(584):2145–66.