Smoking reduction trajectories, and their association with smoking cessation: A secondary analysis of longitudinal RCT data

Anthony Barrows [[1]](#footnote-1), Gemma Taylor2, Elias Klemperer1, Hugh Garavan1, Nicholas Allgaier1

Running Head: SMOKING REDUCTION TRAJECTORIES

Word Count: 3,676 of 3500

Declaration of interests: None to declare.

Primary funding:

NIH/NIDA T32DA045593

# ABSTRACT

#### Background and aims

Cigarette smoking remains a leading cause of preventable death, with quitting the only available cure. Smoking reduction is often recommended to smokers who find it difficult to quit entirely, but little is known about the resulting smoking behavior. This study (1) examined smoking trajectories among participants told to reduce using latent class analysis, (2) used machine learning-based methods to predict these trajectories, and (3) determined whether smoking trajectories were associated with biochemically verified smoking cessation.

## Design

A secondary analysis of five randomized controlled nicotine replacement therapy (NRT) trials.

## Setting

Five university and medical centers in Australia, Denmark, Germany, Switzerland, and the USA.

## Participants

The full sample () had a baseline age of and was 45.1% male. On average, participants smoked cigarettes per day. Half (1032/2066) of the sample received active NRT (i.e., patch, gum, or inhaler).

## Measurements

Latent trajectories were determined using average self-reported cigarettes per day across four trial follow-up points. Predictive modeling used area under the curve (AUC) and the correlation coefficient .

## Findings

A three-class trajectory model was selected, characterized by one group (Class 1, ) which reduced from baseline smoking substantially, another (Class 2, ) which reduced by approximately half, and a third (Class 3, ) which remained near their baseline smoking levels. Predictive modeling showed relatively strong classifications to Class 1 (AUC = ) and Class 3 (AUC = ), but not to Class 2 (AUC = ). Older participants with lower baseline measures of anxiety and nicotine dependence were most likely to reduce their smoking. Latent class improved prediction of 1-year follow-up carbon monoxide (CO) ( using latent class alongside baseline characteristics versus using baseline characteristics alone). Relative to Classes 2 and 3, participants in Class 1 tended to have lower CO values at the 1-year follow-up point.

## Conclusions

Examining smoking behaviors among participants not looking to quit reveals heterogeneity, which may lend to identifying subgroups more receptive to cessation interventions.

# INTRODUCTION

Smoking is the number one cause of premature and preventable illness worldwide (1). Smoking accounts for approximately 5.7% of total health expenditure worldwide and 6.5% in high income countries like the UK and USA (2), and 1 in two smokers will die unless they quit (3).

Smoking reduction to quit is often recommended to smokers who have found it difficult to quit abruptly in the past. Though magnitude and duration vary, broadly, smoking reduction is defined as a decrease in the number of cigarettes smoked per day. In the UK and in some parts of Europe, smoking reduction is promoted as a second-line route to quitting smoking or as a harm reduction approach for smokers who cannot, or may not be ready, to stop smoking completely (4,5). In addition, the license for nicotine replacement therapy (NRT) explicitly states that it can be used for smoking reduction (6). In the USA smoking reduction is offered informally as part of smoking cessation treatment, and NRT is deemed appropriate for use as a reduction aid (7). There is limited evidence that reduction itself improves health (8), but clearer evidence that smoking reduction is as effective as abrupt quitting in achieving complete cessation (9), which has well established health benefits.

Smokers report that they perceive smoking reduction as an acceptable way to stop smoking (10–14). Smokers reducing their daily number of cigarettes could use a range of different approaches to do so, including a combination of behavioral counseling and NRT (9). Providing behavioral support to reduce to quit may help people to stop smoking than providing people with self‐help resources only. For example, setting a time period over which to reduce before quitting completely, progressively shortening the time periods in the day when smoking occurs (i.e., timed reduction (15)), or gradually eliminating the easiest to the most difficult cigarettes to give up in a given day (i.e., hierarchical reduction (16)). Other approaches include simply advising smokers to reduce their smoking as much as possible without giving any specific guidance on how to do so (17–21). However, there is no clear evidence to support any particular behavioral reduction method over others (9,14).

A Cochrane Review found that several trials have simply asked participants to reduce without giving them any specific instructions on how to do so (9). However, little is known about what the resulting, more unstructured, participant-led smoking reduction patterns look like. For example, after receiving basic instructions to reduce their smoking, do smokers cut down immediately, or later? Do they steadily reduce smoking until complete cessation? Do they reduce and then increase the amount smoked? Or are there other patterns of smoking reduction? Secondly, are certain patterns of smoking reduction are associated with better cessation outcomes, and if so, which ones? Third, are patterns of smoking reduction determined by participant characteristics or dependency profiles? This information could be important when developing and tailoring smoking reduction interventions and could aid clinicians who recommend and offer smoking reduction treatment.

In this secondary analysis of five randomized placebo-controlled trials of NRT for smoking reduction we:

1. Use latent class analysis (LCA) to determine if there are trajectories in cigarettes per day (CPD) overtime in people who are asked to reduce their smoking before quitting, without specific instructions on how to do so.
2. Determine which baseline participant characteristics predict latent class membership.
3. Determine whether latent class membership is associated with biochemically verified smoking cessation, while considering participant characteristics.

Although several studies have used LCA to explore heterogeneity in smoking behavior trajectories (e.g., 17–19), none to date have leveraged contemporary machine learning methods to determine characteristics associated with latent class membership. The present study uses regularized linear models to minimize noise associated with highly correlated predictors while emphasizing effects from relevant predictors of class membership and smoking cessation. These methods focus primarily on overall predictive performance rather than statistical significance.

# METHODS

This study was preregistered on [OSF](https://osf.io/qh378/), and all code used to produce results and figures is available through [GitHub](https://github.com/ajbarrows/mcneil-lca).

## Study Design

This secondary analysis examined individual-level patient data from five randomized placebo-controlled trials of NRT for smoking reduction (17–19). The trials were carried out to a consistent protocol. Participants were provided with active NRT (inhaler, patches, or gum) in the intervention arm and a placebo form of the same NRT in the control arm. In both arms participants received the same minimal level of behavioral support for smoking reduction and were encouraged to reduce their smoking as much as possible with cessation as the end goal.

The trials were conducted and funded by [McNeil AB](https://www.mcneilab.se/) . The current authors were not involved with the original trials.

## Study Details

The trials took place between 1997 and 2003 and were conducted in university and medical centers in Denmark, Switzerland, Australia, the USA and Germany. There were 2066 participants enrolled in 5 trials. All participants were at least 18 years of age, reported smoking 15 or more cigarettes per day, and were selected because they wanted to reduce but not stop smoking, and had smoked regularly for at least 3 years. Participants were excluded if they were pregnant, breastfeeding, under psychiatric care, deemed to be unfit by a general practitioner, or part of a cessation program.

## Data Collection

At baseline, trial investigators gathered data on participants’ demographic details, age started smoking, cigarettes per day (CPD), nicotine dependence (Fagerström Test of Nicotine Dependence; FTND (22)), intention to reduce, intention to quit, smoking history (e.g., number of previous quit attempts, longest period without smoking), and self-rated effects from smoking, the Short Form Health Survey-36 (SF-36/RAND-36, (23–25)). To preserve anonymity, some demographic data were unavailable for this secondary analysis.

Participants were followed up with consistent measures at baseline, and at 2-, 10-, 18-, 26- and 52- weeks from baseline. On each occasion, investigators collected data on CPD and recorded an expired air carbon monoxide (CO) reading. Participants answered the following questions via self-report: “how many cigarettes do you smoke/day on average?”, “how many cigarettes do you smoke/week on average”, “how many cigarettes do you smoke/month on average?”.

## Variables

Latent trajectories were determined using percent change from baseline in average cigarettes per day (CPD) at weeks 2, 10, 18, and 26. When CPD was unavailable, participants self-reported cigarettes per week were divided by 7 and used instead. If a participant reported they had stopped smoking, CPD was set to 0. Abstinence was determined using CO values from week 52.

The baseline variables used to predict latent class were age at trial intake, age started smoking, longest period without smoking, number of times tried to quit smoking, 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 subscales, carbon monoxide parts per million (CO ppm), relief from smoking, study site, and parent trial treatment group (22 variables in all).

## Bias

To address bias due to missing data, baseline predictors with greater than 25% missing observations were discarded. Additionally, subjects missing 7 or more baseline variables were excluded from analyses. Finally, missing observations for remaining baseline predictors were grouped by study site, sex, and age and imputed using average values within those groups.

## Statistical Methods

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

A latent class mixture model (26–28) 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 ranging from 1 to 6 classes were estimated, and models for classes greater than 1 were initialized using parameters from the 1-class model. Grid search methods were performed for subsequent models, seeking to minimize information criteria at each step. The model that balanced fit and parsimony was selected as optimal.

Posterior classification was then used to assign subjects to each latent class using maximum likelihood estimation given the information (i.e., percent change in CPD at each week) collected in the longitudinal model.

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

Elastic net logistic regression (29,30) 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, with membership to each class predicted using a separate regression model.

Initially, the data were split into 80% training and 20% testing sets. The training data were then used to tune parameters for elastic net logistic regression models. Model selection and evaluation was performed using a nested cross-validation framework. Training was conducted using 5-fold cross validation, with 80% of the training set used for model training and 20% for evaluation. Model hyperparameter selection was conducted using a further divide of this training set using 5-fold cross-validation within each outer fold. Prediction accuracy was measured using receiver operator characteristic area-under-the-curve (AUC). Parameters from the model whose AUC results were within one standard error of the optimal cross-validated results were selected (31). For each model, AUC represents the accuracy of prediction of the internal validation set. Hyperparameters from the most successful of these models were used to fit a model to the initial training set to obtain coefficients. To confirm model generalizability, these hyperparameters were also used to fit a final model to the initial testing set. Null ROCs were computed for each predicted class, and overall statistical significance was assessed using AUC’s equivalence with the Mann-Whitney U-statistic (32). This procedure was repeated using each latent class as the target.

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

Elastic net linear regression was used to predict smoking abstinence 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, and (3) baseline predictors with only subjects in each Latent Class, respectively. 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, .

## Quantitative Variables

Ordinal predictors (i.e., longest period without smoking, number of times tried to quit smoking, 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, and relief from smoking) were treated as numeric variables. These plus additional numeric predictors (i.e., SF-36 scores, CO, age at intake) and outcomes (i.e., CO-confirmed abstinence at 1-year follow-up) in the training set were normalized and scaled to have a mean of 0 and standard deviation of 1. Numeric predictors in the testing set were normalized blindly using the same transformations as the training set.

# RESULTS

## Participants

Seven subjects were missing baseline CPD values and were not included in analyses. Additionally, 101 subjects were missing 7 or more baseline measures and were not included in analyses. The 1783 remaining subjects had at least one post-baseline average CPD value.

Smoking trajectories were modeled using a total of 5,578 observations (Weeks 2, 10, 19, and 26 had 1764, 1148, 1384, 1282 observations respectively) from 1,783 participants (see flowchart in Figure 1).

## Descriptive Data

Participants were from 5 countries (Australia, Denmark, Germany, Switzerland, and USA), were 44.8% male (798/1783), with a mean (SD) age of 44.10 (10.72) years and smoked a mean (SD) of 27.32 (9.73) CPD (see Table 1).

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 (i.e., RSQ pepping-up, calming effects, and overall last cigarette experience: 7.31%, 2.81%, and 0.05% missing respectively; intention to quit: 2.81%; SF-36 subscales: < 1%), subjects’ 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.

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

Latent class model-fit information appears in Table 2, suggesting that models using 3-6 classes are optimal. The 3-class model was selected to balance minimal information criteria with parsimony; the 4, 5, and 6-class models did not reveal more information about smoking trajectories. Distributions of CPD changes for each follow-up point and a model-fit curve are available in the supplemental materials.

Class 1 (n = 186, 10.43%) is characterized by an average initial reduction in CPD of approximately 57% at week 2, followed by a trend toward 90% reduction by week 26. Participants in Class 2 (n = 803, 45.04%) initially reduced their smoking approximately 41% and remained consistent, for an average reduction of 46% from baseline at week 26. Class 3 participants (n = 794, 44.53%), in contrast, showed a modest (approximately 22%) reduction in smoking and a return to near-baseline levels (i.e., 6% reduction from baseline) by week 26. Average CPD trajectories for each of the selected 3 classes are shown in Figure 2.

Fixed effects maximum likelihood estimations show the initial smoking reduction (i.e., change from baseline smoking at week 2) for participants in Class 1 was significantly less than Class 3 (estimate (SE) = -1.31 (0.11), Wald = -12.32, ), but initial smoking reductions among Class 2 participants did not differ from those in Class 3 (estimate (SE) = 0.03 (0.08), Wald = 0.38, ). Within each trajectory, differences in CPD were observed for each week ().

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

Regularized logistic regression models predicting membership to Class 1 vs. all and Class 2 vs. all performed better than chance, while membership to Class 2 vs. all was difficult to discern. The strongest predictive performance is seen predicting Class 1 vs. Class 3 and Class 1 vs. Class 2, while Class 2 vs. Class 3 was within range of chance. During non-parametric statistical comparisons, all models (i.e., one-versus-all and class comparisons) performed significantly better than chance. To explore differences between participants who reduced their smoking a great deal but received placebo NRT vs. those who reduced their smoking little or not at all but received active NRT, a final regression model was used to differentiate between placebo NRT subjects in Class 1 and active NRT subjects in Classes 2 and 3 (see Table 3). See Figure 3 for ROC curves showing classification performance for the one-versus-all models.

To determine the relative contributions of each baseline characteristic to the overall model’s predictive capacity, regression coefficients for each predictor were recorded for each of the 5 outer validation folds. Averages of these coefficients are presented in Figure 4. Participants assigned to Class 1 – those who reduced smoking substantially – tended to be older, have started smoking later in life, and have lower measures of anxiety and nicotine dependence. They were more likely to want to quit smoking at study baseline and to have slightly higher SF-36 social and physical sub scores, but below average pain sub scores. Intuitively, those in Class 1 – the group which reduced its smoking the most – were more likely to have been assigned to the active NRT condition, while those in Class 3 were more likely to have received placebo treatment. Although it represents the largest group of smokers, the model predicting membership to Class 2 showed low overall performance. Consequently, no distinct pattern of features describing Class 2 emerges. Contrasting placebo-assigned subjects in Class 1 with active-assigned subjects in Classes 2 and 3, those with lower baseline anxiety and FTND scores were more likely to reduce smoking without NRT aid, while older, male smokers were less likely to reduce despite NRT aid.

Large site effects were observed. Relative to participants in the USA, Swiss, German, Danish, and Australian participants were more likely to follow Class 1’s trajectory and less likely to follow Class 3’s. Additional predictive models were used to determine characteristic patterns associated with individual sites, and no distinct patterns emerged.

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

Of the 1783 participants assigned to latent classes, 928 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, 268/928 (28.9%) had quit smoking at this point. More specifically, 67.1% (n = 96/143) of subjects in Class 1 had quit, 24.8% (n = 124/500) in Class 2, and 16.8% (n = 48/285) 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) = 0.245 , with on the testing set. Adding latent class as a predictor improved average cross-validated prediction performance by an average of 6.3% (CV , test ). Each model improved on performance observed using latent class alone (, test ). To simulate using CO prediction as a quit-smoking classifier, ROC curves were generated using exhaustive CO thresholds (see Figure 3). Predicted CO values at or below the threshold were considered accurate predictions. The model using latent class in addition to baseline characteristics outperformed the model using those characteristics alone.

As with the previous analysis, regression coefficients from each model’s validation folds were recorded to assess feature importance in predicting 1-year CO values (see Figure 5). Membership to Classes 2 and 3 were associated with higher follow-up CO values relative to participants from Class 1. Notably, latent class predictors were weighted more heavily than baseline CPD or CO.

# DISCUSSION

This study examined smoking patterns in a secondary analysis of five NRT trials and found three distinct repeated measures smoking trajectories. Specifically, some smokers initially reduced and nearly eliminated their smoking (Class 1), some reduced by approximately half and remained at that level throughout the trial (Class 2), and others reduced initially but reverted to cigarette use near their baseline levels (Class 3).

Predictive modeling revealed that smokers who reduced substantially tended to be older, have later smoking onset, and lower levels of anxiety and nicotine dependence. This trend was emphasized when comparing smokers who reduced despite receiving placebo NRT versus those who failed to reduce substantially yet received active NRT. Additionally, regression models predicting CO values 6 months after the trial’s conclusion showed stronger performance incorporating latent class as a predictor versus using baseline characteristics alone, suggesting smoking trajectories may have implications for predicting smoking cessation. Those in Class 1 have lower 1-year CO values, while those in Classes 2 and 3 tended to have higher follow-up readings.

Most importantly, despite random assignment to active or placebo NRT and consistent instructions to reduce smoking, cigarette use trajectories were not homogeneous. In particular, although those who received active NRT were more likely to reduce their smoking, many participants did not reduce. Lower levels of nicotine dependence and anxiety predicted reduction, even among those given placebo NRT.

As expected, those who failed to reduce their smoking during the trial or who only cut out a small number of cigarettes were less likely to have quit smoking following the trial. Although reduction and cessation patterns aligned with treatment assignment, baseline reports of anxiety, social functioning, and nicotine dependence were similarly important. Smoking trajectories were substantially more predictive of follow-up CO than assignment to NRT treatment group, which suggests that a focus on initial smoking reduction (i.e., as few as two weeks following instructions to reduce) is important when considering who is likely to benefit from a recommendation to reduce smoking if abruptly quitting is not feasible. Further, since participants who received placebo NRT but reduced smoking anyway tended to report lower levels of anxiety, treating anxiety may lead to smoking reduction.

Recent literature recommends lower cutoffs for biochemically-verified smoking cessation, with CO values as low as 6ppm (33), emphasizing the relevance of predictive modeling. However, particular smoking trajectories appear to improve prediction of CO values, and discerning between those trajectories occurred largely at two and four-week follow-up points. When asked to reduce smoking, those with substantial reductions after two weeks may follow paths toward substantial reduction or quitting, while those with minimal initial reductions may revert to their usual smoking levels. These results align with a previous LCA (34) which concluded that smokers who are not looking to quit are not homogeneous.

## Limitations

Several unmeasured variables might confound the associations in question, and we cannot evaluate the impact of residual confounding. For example, ethnicity, education, intention to reduce were either not measured, or not measured consistently across trials.

Additionally, the trials did not capture short term changes in CPD that occurred temporarily between measurements, which could have affected outcomes (35). Latent classes were developed based on self-reported CPD. It is possible that participants inaccurately reported their CPD, but it was not possible to biochemically confirm CPD using CO. Although a person may reduce their CPD, they may also inhale their remaining cigarettes more deeply, which could mean that the CO reading does not decrease in-line with the reduction in CPD. Furthermore, CO has a half-life of around 4 hours meaning that the magnitude of the measurement is affected by the recency of smoking (33). Therefore, even if a person has reduced their daily CPD overall this would not be reflected in the CO measurement if they had smoked their remaining cigarettes just before the measure was taken.

The McNeil trials recruited participants who wanted to reduce but not stop smoking and had smoked for at least 3 years. Therefore, our analyses may not generalize to participants consciously seeking to reduce smoking to quit, although many participants in these trials did quit smoking after smoking reduction intervention.

Examining latent trajectories in smoking behavior among a sample of smokers not looking to quit reveals heterogeneity in smoking patterns. Identifying these subgroups may allow for targeted intervention with the goal of reducing cigarette use particularly among populations predisposed to tobacco use.

# ACKNOWLEDGEMENTS

This research was supported by NIH/NIDA under award number T32DA043593. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

# REFERENCES

1. World Health Organization. WHO report on the global tobacco epidemic, 2011: warning about the dangers of tobacco. 2011 [cited 2023 Jan 20]; Available from: https://apps.who.int/iris/handle/10665/44616

2. Goodchild M, Nargis N, Tursan d’Espaignet E. Global economic cost of smoking-attributable diseases. Tob Control. 2018 Jan;27(1):58–64.

3. Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. The Lancet. 2013 Jan;381(9861):133–41.

4. NICE. National Institute for Clinical Excellence (NICE). Smoking: Harm reduction Public health guideline [PH45] [Internet] [Internet]. 2013. Available from: https://www.nice.org.uk/guidance/ph45

5. European Network for Smoking and Tobacco Prevention. Guidelines for treating tobacco dependence. 2018.

6. Medicine and Healthcare Products Regulatory Agency Committee. Report of the Committee on Safety of Medicines Working Group on Nicotine Replacement Therapy. London; 2005.

7. Fucito LM, Bars MP, Forray A, Rojewski AM, Shiffman S, Selby P, et al. Addressing the Evidence for FDA Nicotine Replacement Therapy Label Changes: A Policy Statement of the Association for the Treatment of Tobacco Use and Dependence and the Society for Research on Nicotine and Tobacco. Nicotine Tob Res. 2014 Jul 1;16(7):909–14.

8. Pisinger C, Godtfredsen NS. Is there a health benefit of reduced tobacco consumption? A systematic review. Nicotine Tob Res. 2007 Jun;9(6):631–46.

9. Lindson N, Klemperer E, Hong B, Ordóñez-Mena JM, Aveyard P. Smoking reduction interventions for smoking cessation. Cochrane Tobacco Addiction Group, editor. Cochrane Database Syst Rev [Internet]. 2019 Sep 30 [cited 2022 Dec 7];2019(9). Available from: http://doi.wiley.com/10.1002/14651858.CD013183.pub2

10. Beard E, Vangeli E, Michie S, West R. The Use of Nicotine Replacement Therapy for Smoking Reduction and Temporary Abstinence: An Interview Study. Nicotine Tob Res. 2012 Jul;14(7):849–56.

11. Shiffman S, Hughes J, Ferguson S, Pillitteri J, Gitchell J, Burton S. Smokers’ interest in using nicotine replacement to aid smoking reduction. Nicotine Tob Res. 2007 Nov;9(11):1177–82.

12. Schmidt L, Reidmohr A, Harwell TS, Helgerson SD. Prevalence and Reasons for Initiating Use of Electronic Cigarettes Among Adults in Montana, 2013. Prev Chronic Dis. 2014 Nov 20;11:140283.

13. Hughes J R. Smokers who choose to quit gradually versus abruptly. Addiction. 2007 Aug;102(8):1326–7.

14. Lindson N, Michie S, Aveyard P. Exploratory Analyses of the Popularity and Efficacy of Four Behavioral Methods of Gradual Smoking Cessation. Nicotine Tob Res. 2020 Dec 12;22(12):2257–61.

15. Cinciripini PM, Lapitsky L, Seay S, Wallfisch A, Kitchens K, Van Vunakis H. The effects of smoking schedules on cessation outcome: Can we improve on common methods of gradual and abrupt nicotine withdrawal? J Consult Clin Psychol. 1995;63(3):388–99.

16. Brockway BS, Kleinmann G, Edleson J, Gruenewald K. Non-aversive procedures and their effect on cigarette smoking. Addict Behav. 1977 Jan;2(2–3):121–8.

17. Haustein K. A double-blind, randomized, placebo-controlled multicentre trial of a nicotine chewing gum in smoking reduction. Study ID 980-CHC-9021-0013 Unpubl Data. 2001;

18. Batra A, Klingler K, Landfeldt B, Friederich H, Westin A, Danielsson T. Smoking reduction treatment with 4-mg nicotine gum: A double-blind, randomized, placebo-controlled study. Clin Pharmacol Ther. 2005 Dec;78(6):689–96.

19. Bolliger CT. Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety. BMJ. 2000 Aug 5;321(7257):329–33.

20. Rennard S, Glover E, Leischow S, Daughton D, Glover P, Muramoto M, et al. Efficacy of the nicotine inhaler in smoking reduction: A double-blind, randomized trial. Nicotine Tob Res. 2006 Aug 1;8(4):555–64.

21. Wennike P, Danielsson T, Landfeldt B, Westin Å, Tønnesen P. Smoking reduction promotes smoking cessation: results from a double blind, randomized, placebo-controlled trial of nicotine gum with 2-year follow-up: Smoking reduction with nicotine gum. Addiction. 2003 Oct;98(10):1395–402.

22. Fagerstrom KO, Heatherton TF, Kozlowski LT. Nicotine addiction and its assessment. Ear Nose Throat J. 1990 Nov;69(11):763–5.

23. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001 Jan;33(5):350–7.

24. Hays RD, Sherbourne CD, Mazel RM. The rand 36-item health survey 1.0. Health Econ. 1993 Oct;2(3):217–27.

25. Hays RD, Prince-Embury S, Chen H. RAND-36 health status inventory. Psychological Corporation San Antonio, TX; 1998.

26. 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/

27. 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

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

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

30. 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.

31. 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

32. 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.

33. Benowitz NL, Bernert JT, Foulds J, Hecht SS, Jacob P, Jarvis MJ, et al. Biochemical Verification of Tobacco Use and Abstinence: 2019 Update. Nicotine Tob Res. 2020 Jun 12;22(7):1086–97.

34. Borrelli B, Gaynor S, Tooley E, Armitage CJ, Wearden A, Bartlett YK. Identification of three different types of smokers who are not motivated to quit: Results from a latent class analysis. Health Psychol. 2018 Feb;37(2):179–87.

35. Klemperer EM, Hughes JR, Naud S. Distal Measurements Can Produce False Negative Results: A Prospective Secondary Analysis of a Natural History Study. Nicotine Tob Res. 2019 Nov 19;21(12):1727–30.

# Tables

Table Participant characteristics by assigned latent class.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Full Cohort | Analysis Sample | Class 1 | Class 2 | Class 3 |
| N (%) | 2066 | 1783 (100) | 186 (10.4) | 803 (45.0) | 794 (44.5) |
| Study Site (%) |  |  |  |  |  |
| Australia | 436 (21.1) | 360 (20.2) | 32 (17.2) | 159 (19.8) | 169 (21.3) |
| Denmark | 411 (19.9) | 340 (19.1) | 35 (18.8) | 175 (21.8) | 130 (16.4) |
| Germany | 385 (18.6) | 353 (19.8) | 60 (32.3) | 153 (19.1) | 140 (17.6) |
| Switzerland | 400 (19.4) | 301 (16.9) | 29 (15.6) | 139 (17.3) | 133 (16.8) |
| USA | 434 (21.0) | 429 (24.1) | 30 (16.1) | 177 (22.0) | 222 (28.0) |
| Study Trt. Group = active (%) | 1032 (50.0) | 900 (50.5) | 125 (67.2) | 413 (51.4) | 362 (45.6) |
| Sex = Male (%) | 931 (45.1) | 798 (44.8) | 96 (51.6) | 357 (44.5) | 345 (43.5) |
| Age (mean (SD)) | 43.98 (10.82) | 44.10(10.72) | 45.79 (11.41) | 44.26 (10.52) | 43.53 (10.73) |
| Baseline FTND (mean (SD)) | 6.12 (2.00) | 6.14 (2.00) | 5.60 (2.13) | 6.11 (2.01) | 6.30 (1.94) |
| Baseline CPD (mean (SD)) | 27.26 (9.74) | 27.32 (9.73) | 25.65 (10.37) | 27.42 (9.78) | 27.62 (9.50) |

Table Latent class model fit comparison (n = 1783). \*Selected model. BIC = Bayesian Information Criterion, AIC = Akaike Information Criterion.

| Number of Classes | BIC | AIC | Maximum Log-Likelihood |
| --- | --- | --- | --- |
| 1 | 2102.42 | 2085.97 | -1039.98 |
| 2 | 577.95 | 545.04 | -266.52 |
| 3\* | 43.75 | -5.62 | 11.81 |
| 4 | -36.02 | -101.85 | 62.92 |
| 5 | -87.61 | -169.90 | 99.95 |
| 6 | -116.79 | -215.54 | 125.77 |

# 

Table \*Average AUC across five-fold cross-validation SEM. \*\*AUC using unseen data. p values represent the probability that the test AUC value is greater than the average computed null AUC value for the test data.

|  | Mean CV AUC\* | Test AUC\*\* |  |
| --- | --- | --- | --- |
| Class 1 vs. All |  | 0.691 |  |
| Class 2 vs. All |  | 0.520 |  |
| Class 3 vs. All |  | 0.582 |  |
| Class 1 vs. Class 2 |  | 0.646 |  |
| Class 1 vs. Class 3 |  | 0.693 |  |
| Class 2 vs. Class 3 |  | 0.553 |  |
| C1 Placebo NRT vs. C1 & C2 Active |  | 0.617 |  |

# Figures

# 

Figure Participant record availability flowchart.

Chart, line chart

Description automatically generated

Figure Average smoking trajectories for the 3-class model (i.e., change from baseline smoking rate). Error bars represent SD.

Chart

Description automatically generated

Figure ROC curves for one-versus-all latent class membership predictions.

Table, calendar

Description automatically generated with medium confidence

Figure Average feature importance from elastic net logistic regression predicting class membership. Feature importance was determined using the average regression coefficient for each predictor across validation folds. Error bars represent SD.

Table

Description automatically generated

Figure Average feature importance from elastic net linear regression predicting CO 6 months following the trial’s conclusion. Latent class regression coefficients use Class 1 as a reference.

# Supplemental Materials

Timeline

Description automatically generated

Figure Distributions of changes in CPD as a percentage of baseline smoking rates (n = 1783).

Chart, line chart

Description automatically generated

Figure Latent class mixture model BIC curve (n = 1783).

1. Department of Psychiatry, University of Vermont, Burlington, VT, USA

   2 Department of Psychology, University of Bath, Bath, UK [↑](#footnote-ref-1)