

BMJ Public Health

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

Journal:	<i>BMJ Public Health</i>
Manuscript ID	Draft
Article Type:	Original research
Date Submitted by the Author:	n/a
Complete List of Authors:	Barrows, Anthony; University of Vermont, Psychiatry Klemperer, Elias; University of Vermont, Psychiatry Garavan, Hugh; University of Vermont, Psychiatry Allgaier, Nicholas; University of Vermont, Psychiatry Lindson, Nicola; University of Oxford Taylor, Gemma; University of Bath
Keywords:	Anxiety , Public Health, Sociodemographic Factors, methods

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

Anthony Barrows, MS^{1*}, Elias Klemperer, PhD¹, Hugh Garavan, PhD¹, Nicholas Allgaier, PhD¹, Nicola Lindson, PhD², Gemma Taylor, PhD³

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

² Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

³ Department of Psychology, University of Bath, Bath, UK

* Please address correspondence to ajbarrow@uvm.edu (802) 656-4849

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Introduction

Tobacco smoking remains the leading cause of preventable death worldwide, with quitting the only cure. Smoking reduction can be recommended to people unmotivated to quit, but evidence on outcomes associated with reduction is mixed.

Methods

In a secondary analysis of five randomized, placebo-controlled trials of nicotine replacement therapy, we determined latent smoking trajectories using cigarettes per day (CPD) across 26 weeks. Participants were adults who smoked daily and were unmotivated to quit. We used predictive modeling and ROC AUC to assess smoking cessation at follow-up week 52.

Results

Participants (N=2066) smoked a mean 27.26±9.74 CPD at baseline. We selected a three-class trajectory model: Class 1 (n=186) achieved the greatest reduction in CPD (57-90%); Class 2 (n=803) reduced by ~50%; Class 3's (n=794) CPD remained stable.

Regularized regression showed older, male participants with lower baseline anxiety were most likely to reduce their smoking (cross-validated AUCs = 0.520 to 0.684, p's < .01).

Latent class improved prediction of smoking cessation at the 52-week follow-up over prediction using baseline characteristics by 14.4% (AUC = 0.776±0.010, p = 0.002). Those who reduced their CPD minimally were nearly 90% less likely to achieve cessation than those who reduced by over 50% (ORs: Class 2 = 0.111±0.013, Class 3 = 0.070±0.005).

Conclusions

Findings suggest people who are unmotivated to quit at baseline but reduce their smoking by more than half are most likely to achieve smoking cessation. Therefore, a lack of reduction success could indicate that greater support is needed to help people to quit.

Key Messages

The present work quantifies the level of smoking reduction (i.e., more than 50%) most associated with cessation, and identifies participant characteristics that predict greater magnitudes of reduction among people who were unmotivated to quit smoking at baseline. Our findings demonstrate the importance of smoking reduction during the first two weeks and may allow clinicians to identify both those most likely to achieve complete cessation and those in need of additional support.

Keywords

Smoking cessation; structural equation modeling; latent class analysis; Elastic Net regression; machine learning; secondary analysis

Word Count: 3830/5000

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 Approximately 50% of people who smoke will die unless they quit.³

Reducing smoking to quit is sometimes recommended to people who have found it difficult to quit smoking 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 people who cannot, or may not be ready, to stop smoking completely.^{4,5} In addition, nicotine replacement therapy (NRT) is licensed 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 health⁸, but clearer evidence that smoking reduction is as effective as abrupt quitting in achieving complete cessation⁹, which has well established health benefits.

People who smoke report that they perceive smoking reduction as an acceptable way to stop smoking.^{10–14} Those wishing to reduce 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 more than providing people with self-help resources only.⁹ 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 reduction¹⁵), or gradually eliminating the easiest to the most difficult cigarettes to give up in a given day (i.e., hierarchical reduction¹⁶). Other approaches include simply advising people who smoke 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}

When people are asked to reduce, but not provided with instructions on how to do so, it is unclear how they choose to do this themselves, e.g., whether they cut down immediately, or later, whether they steadily reduce smoking until complete cessation or, instead, whether they reduce and then increase the amount smoked. There may be other patterns of smoking reduction, and some patterns may be associated with better cessation outcomes. Furthermore, it is unknown whether patterns of smoking reduction are determined by participant characteristics or dependency profiles. This information could be important to those developing and tailoring smoking reduction interventions, and it 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) and machine learning (ML)-based regression to determine if there are trajectories in cigarettes per day (CPD) over time in people who are asked to reduce their smoking before quitting without being given 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 ML-based 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 in an effort to capture nuanced use patterns as compared to simply more-versus-less reductions in CPD.²²

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

METHODS

This study was preregistered on Open Science Framework ²³, and all analytical code is available through GitHub.²⁴

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 or gum) in the intervention arms and a placebo form of the same NRT in the control arms. In both arms, participants received the same minimal level of behavioral support for smoking reduction and were simply encouraged to reduce their smoking as much as possible, with cessation as the end goal.

The trials were conducted and funded by [McNeil AB](#) who manufacture NRT. The current authors were not involved with the original trials, and the funder was not involved with the planning, analysis, interpretation, or funding of this study.

Study Details

The trials took place between 1997 and 2003 and were conducted in university and medical centers in Denmark ²¹, Switzerland ¹⁹, Australia ¹⁷, the USA ²⁰, and Germany ¹⁸. There were 2,066 participants enrolled across all five trials. At baseline, all participants were at ≥18 years old, reported smoking ≥15 CPD, were selected because they wanted to reduce but not stop smoking, and had smoked regularly for ≥3 years. Participants were excluded if they were pregnant, breastfeeding, under psychiatric care, deemed to be unfit by a general practitioner, or currently enrolled in a smoking cessation program.

Patient and Public Involvement

Patients who wanted to reduce their smoking were recruited for participation in the included trials, and all participants were given behavioral support to do so. We expect to make the results of this study publicly available.

Data Collection

At baseline, trial investigators gathered data on participants’ demographic details, age they started smoking, nicotine dependence, intention to reduce, intention to quit, smoking history (number of previous quit attempts, longest period without smoking, time since last quit

attempt), self-rated effects from smoking, and self-reported physical and emotional health. To preserve anonymity, some demographic data were unavailable for this secondary analysis.

Participants also provided a breath carbon monoxide (CO) sample and answered the following questions at baseline and at 2-, 10-, 18-, 26- and 52- weeks from baseline: “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?”

Latent trajectories were determined using percent change from baseline in average CPD at weeks 2, 10, 18, and 26. When CPD was unavailable, participants’ self-reported values for cigarettes per week were divided by 7. If a participant reported they had stopped smoking, CPD was set to 0. Abstinence at 52 weeks was determined using CO values <6 parts per million (ppm), consistent with recent guidance.²⁵ See supplement for cessation prediction with the threshold used at the time of the trial.

The baseline variables used to predict latent class were age at trial intake, age started smoking, longest period without smoking, number of previous quit attempts, (Fagerström Test of Nicotine Dependence; FTND (19)), intention to quit, length of time since last quit attempt, experiences of anxiety and depression, the Short Form Health Survey-36 (SF-36/RAND-36,^{27–29}) subscales, CO ppm, relief from smoking, study site, and trial treatment group (22 variables). Coding information for baseline variables is included in the supplemental materials.

To address bias due to missing data, baseline predictors with greater than 25% missing observations were discarded. Additionally, participants missing ≥ 7 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. This cautious approach was selected to help overcome bias introduced through non-random missingness.³⁰

Statistical Methods

The present work (1) uses latent class mixture models^{31–33} to uncover latent trends in cigarette usage in the parent clinical trials independent of treatment group assignment; (2) employs elastic net logistic regression to determine which baseline and demographic characteristics are most associated with these trajectories; (3) fits additional elastic net regression models to determine whether knowledge of a participant’s smoking trajectory improves prediction of cessation after the trial. See Figure 1 for details.

Analysis 1: Trajectories in CPD over time

A latent class mixture model^{31–33} was used to determine longitudinal trends in CPD from baseline assessed at trial weeks 2, 10, 18, and 26. The primary dependent variable was percent change in 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 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^{34,35} 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 (ROC AUC). Parameters from the model whose ROC AUC results were within one standard error of the optimal cross-validated results were selected.³⁶ For each model, ROC 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 ROC AUC’s equivalence with the Mann-Whitney U-statistic.³⁷ This procedure was repeated using each latent class as the target.

Analysis 3: Which trajectories in CPD predict smoking cessation?

Elastic net logistic regression was used to predict smoking abstinence at 52-week 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 biochemically verified smoking cessation. Participants

1
2
3 were counted as having quit smoking if they reported doing so and registered a CO value of <
4 6ppm.²⁵ All other participants were counted as continuing to smoke. Each regression model
5 was fit and evaluated using nested cross-validation procedures described in Analysis 2.
6
7

8 Quantitative Variables 9

10
11 Ordinal predictors (i.e., longest period without smoking, number of times tried to quit smoking,
12 intention to quit, length of time since last quit attempt, experience of anxiety in the last 24
13 hours, experience of depression in the last 24 hours, and relief from smoking) were treated as
14 numeric variables. These plus additional numeric predictors (i.e., SF-36 scores, CO, age at
15 intake) and outcomes (i.e., CO-confirmed abstinence at 52-week follow-up) in the training set
16 were normalized and scaled to have a mean of 0 and standard deviation of 1. Numeric
17 predictors in the testing set were normalized blindly using the same transformations as the
18 training set.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 5578 observations (weeks 2, 10, 19, and 26 had 1764, 1148, 1384, 1282 observations respectively) from 1783 participants (see flowchart in Figure 1).

Descriptive Data

Participants from five countries 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 and calming effects of smoking, 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 categorized 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 Supplemental Table 1, 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 greater than Class 3 (estimate (SE) = -1.31 (0.11), Wald = -12.32, $p < .001$), 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, $p = .701$). Within each trajectory, differences in CPD were observed for each week (p 's $< .001$).

Analysis 2: Predicting longitudinal trajectories in CPD using baseline variables

Regularized logistic regression models predicting membership to Classes 1, 2, and 3 vs. all performed better than chance (validation AUCs: Class 1=0.766, $p < .001$; Class 2=0.569, $p = .008$; Class 3=0.585, $p < .001$), but membership to Class 2 vs. all was closest to chance. The strongest predictive performance is seen predicting Class 1 vs. Class 3 (AUC=0.788, $p < .001$) and Class 1 vs. Class 2 (AUC=0.784, $p < .001$), while Class 2 vs. Class 3 was near chance (AUC=0.523, $p < .001$). See Supplemental Table 2 for cross-validated AUC scores and Supplemental Figure 3a for one-versus-all ROC curves.

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. These coefficients are presented as averages in Figure 3. Participants assigned to Class 1 – those who reduced smoking substantially – tended to be older 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 the greatest number of participants were assigned to Class 2, the model predicting membership to Class 2 showed low overall performance. Consequently, no distinct pattern of features describing Class 2 emerges.

Since study intervention was not a clear predictor of latent class, a post-hoc one-versus-all logistic regression model was used to predict membership to Class 1 within the 900 (50.5%) participants who received placebo NRT. These participants were more likely to be male, older at baseline, present with lower anxiety scores, and were more likely to have tried quitting before.

Large site effects were observed. Relative to participants in the USA, participants at the other 4 sites (Switzerland, Germany, Denmark, and Australia) 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?

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Of the 1783 participants assigned to latent classes, 122 (6.8%) met criteria for biochemically verified smoking cessation (Class 1: 70/186 [37.6%]; Class 2: 34/803 [4.2%]; Class 3: 18/776 [2.3%]) at the 52-week follow-up. Regularized logistic regression was used to predict smoking cessation using baseline characteristics alone ($AUC = 0.632 \pm 0.006$, $p < .001$), and baseline characteristics plus latent class ($AUC = 0.776 \pm 0.010$, $p < .001$). Each model outperformed classification using a permuted null distribution. Adding latent class as a predictor improved cessation prediction by 14.4%. See Supplemental Figure 3b for ROC curves.

As with the previous analysis, regression coefficients from each model’s validation folds were recorded to assess feature importance in predicting smoking cessation (see Figure 3). In the model using latent class as a predictor, participants in Classes 2 and 3 were approximately 90% less likely to achieve smoking cessation 6 months following the trial than Class 1 (Class 2 OR = 0.111 ± 0.013 , Class 3 OR = 0.070 ± 0.005). Participants who received active NRT were also more likely to quit smoking (OR = 1.90 ± 0.157). A large site effect was observed, with participants at sites in Denmark and Germany more likely to quit smoking than those at the USA site.

DISCUSSION

This study examined smoking patterns in a secondary analysis of five NRT trials and found three distinct repeated measures smoking trajectories among participants who were not motivated to quit smoking at baseline. Approximately 10% of participants initially reduced and nearly or completely eliminated their smoking (Class 1), 45% reduced by approximately half of their baseline CPD and remained at that level throughout the trial (Class 2), and the remaining 45% reduced initially but reverted to cigarette use similar to their baseline levels (Class 3).

Predictive modeling revealed that participants who reduced the most tended to be male, older and lower levels of anxiety. This trend was emphasized when comparing participants who reduced despite receiving placebo NRT versus those who failed to reduce substantially yet received active NRT. Additionally, regression models using baseline characteristics plus latent class assignment to predict smoking cessation at 52-weeks follow-up outperformed models including baseline characteristics alone. This suggests that smoking trajectories, including initial patterns of reduction, may have implications for predicting smoking cessation outcomes.

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.

Consistent with prior research^{38,39}, 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. Latent class assignment was approximately as strongly associated with follow-up smoking cessation as was random assignment to active or placebo NRT, 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 among those not ready to quit abruptly and who may need further support to get to this point. Our findings also indicate that those with lower anxiety achieved greater reductions in CPD, suggesting interventions that address anxiety could help to maximize smoking reduction.

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 LCA⁴⁰ which concluded that people who smoke and are not looking to quit are heterogeneous.

Directions for Future Research

The present study identified a substantial minority of people who were not looking to quit smoking but reduced substantially when asked to. That subset was far more likely to achieve complete cessation, suggesting that those who reduce smoking successfully may be more likely to quit. However, it remains unclear whether this initial reduction causes cessation or if those able to reduce are also more likely to achieve complete cessation.

Strengths and Limitations

This study has many strengths, including the use of a large sample across multiple countries, using ML-based predictive modeling to find robust associations, and objective, information-theoretic approaches to selecting LCA models. Additionally, associations with smoking cessation benefit from biochemical verification rather than relying solely on self-report.

Limitations include that latent classes were developed based on self-reported CPD. Although it is possible that participants inaccurately reported their CPD, it was not possible to biochemically confirm CPD using CO. Those who reduce their CPD may inhale their remaining cigarettes more deeply, which could mean that the CO reading does not decrease in-line with the reduction in CPD ⁴¹. Furthermore, CO readings are affected by the recency of smoking ²⁵. 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. Finally, missing data could confound results, although imputation was minimal and applied within relevant groups to minimize bias.

Conclusion

Our examination of latent trajectories in smoking behavior among a sample of people who were not motivated to quit revealed heterogeneity in smoking patterns. Specifically, three distinct smoking trajectories were identified, with one nearly twice as likely as the others to achieve subsequent smoking cessation. These findings establish that smoking reduction by 50% or more is associated with a substantially increased likelihood of smoking cessation among people who were not ready to quit at baseline and demonstrate the importance of reduction during the first two weeks of a smoking intervention.

FUNDING

This work was supported by the National Institute on Drug Abuse and the National Institute of General Medical Sciences at the National Institutes of Health (grant numbers T32DA045593 to AB, P30GM149331 to EK); the Prevention and Population Research Committee and a Population Researcher Postdoctoral Fellowship Award from Cancer Research UK (grant numbers PPRCPJT\100,023 to GT, C56067/A21330 to GT).

COMPETING INTERESTS

GT works for a scientific consulting company offering statistical services for projects unrelated to this research. The current authors were not involved in the design, conduct, analysis, write-up or dissemination of the original trials, and the funder of these trials (McNeil AB) had no involvement in the planning, analysis, or interpretation of this secondary analysis.

DATA AVAILABILITY

Analytical code is available at <https://github.com/ajbarrows/mcneil-lca>. Data are available from the corresponding author upon request.

REFERENCES

1. World Health Organization. WHO report on the global tobacco epidemic, 2011: warning about the dangers of tobacco. Published online 2011. Accessed January 20, 2023. <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;27(1):58-64. doi:10.1136/tobaccocontrol-2016-053305

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;381(9861):133-141. doi:10.1016/S0140-6736(12)61720-6

4. NICE. *National Institute for Clinical Excellence (NICE). Smoking: Harm Reduction Public Health Guideline [PH45].*; 2013. <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.*; 2005.

7. Fucito LM, Bars MP, Forray A, 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;16(7):909-914. doi:10.1093/ntr/ntu087

8. Pisinger C, Godtfredsen NS. Is there a health benefit of reduced tobacco consumption? A systematic review. *Nicotine Tob Res*. 2007;9(6):631-646. doi:10.1080/14622200701365327

9. Lindson N, Klemperer E, Hong B, Ordóñez-Mena JM, Aveyard P. Smoking reduction interventions for smoking cessation. Cochrane Tobacco Addiction Group, ed. *Cochrane Database Syst Rev*. 2019;2019(9). doi: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;14(7):849-856. doi:10.1093/ntr/ntr297

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;9(11):1177-1182. doi:10.1080/14622200701648441

12. Schmidt L, Reidmohr A, Harwell TS, Helgersson SD. Prevalence and Reasons for Initiating Use of Electronic Cigarettes Among Adults in Montana, 2013. *Prev Chronic Dis*. 2014;11:140283. doi:10.5888/pcd11.140283

13. Hughes J R. Smokers who choose to quit gradually versus abruptly. *Addiction*. 2007;102(8):1326-1327. doi:10.1111/j.1360-0443.2007.01948.x
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;22(12):2257-2261. doi:10.1093/ntr/ntaa123
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):388-399. doi:10.1037/0022-006X.63.3.388
16. Brockway BS, Kleinmann G, Edleson J, Gruenewald K. Non-aversive procedures and their effect on cigarette smoking. *Addict Behav*. 1977;2(2-3):121-128. doi:10.1016/0306-4603(77)90029-6
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*. Published online 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;78(6):689-696. doi:10.1016/j.clpt.2005.08.019
19. Bolliger CT. Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety. *BMJ*. 2000;321(7257):329-333. doi:10.1136/bmj.321.7257.329
20. Rennard S, Glover E, Leischow S, et al. Efficacy of the nicotine inhaler in smoking reduction: A double-blind, randomized trial. *Nicotine Tob Res*. 2006;8(4):555-564. doi:10.1080/14622200600789916
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;98(10):1395-1402. doi:10.1046/j.1360-0443.2003.00489.x
22. Shiyko MP, Li Y, Rindskopf D. Poisson Growth Mixture Modeling of Intensive Longitudinal Data: An Application to Smoking Cessation Behavior. *Struct Equ Model Multidiscip J*. 2012;19(1):65-85. doi:10.1080/10705511.2012.634722
23. Weyenberg G, Lindson N, Klemperer E, Taylor G. Smoking reduction trajectories, and their association with smoking cessation: A secondary analysis of longitudinal RCT data. Published online August 30, 2022. doi:10.17605/OSF.IO/QH378
24. Barrows A. mcneil-lca. Published online March 29, 2023. Accessed April 24, 2023. <https://github.com/ajbarrows/mcneil-lca>

25. Benowitz NL, Bernert JT, Foulds J, et al. Biochemical Verification of Tobacco Use and Abstinence: 2019 Update. *Nicotine Tob Res.* 2020;22(7):1086-1097. doi:10.1093/ntr/ntz132

26. Fagerstrom KO, Heatherton TF, Kozlowski LT. Nicotine addiction and its assessment. *Ear Nose Throat J.* 1990;69(11):763-765.

27. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med.* 2001;33(5):350-357. doi:10.3109/07853890109002089

28. Hays RD, Sherbourne CD, Mazel RM. The rand 36-item health survey 1.0. *Health Econ.* 1993;2(3):217-227. doi:10.1002/hec.4730020305

29. Hays RD, Prince-Embury S, Chen H. *RAND-36 Health Status Inventory*. Psychological Corporation San Antonio, TX; 1998.

30. Pepinsky TB. A Note on Listwise Deletion versus Multiple Imputation. *Polit Anal.* 2018;26(4):480-488. doi:10.1017/pan.2018.18

31. Proust-Lima C, Philipps V, Lique B. Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package **lcmm**. *J Stat Softw.* 2017;78(2). doi:10.18637/jss.v078.i02

32. Proust-Lima C, Philipps V, Diakite A, Lique B. *Lcmm: Extended Mixed Models Using Latent Classes and Latent Processes.*; 2022. <https://cran.r-project.org/package=lcmm>

33. Hagenaars JA, McCutcheon AL, eds. *Applied Latent Class Analysis*. Cambridge University Press; 2002.

34. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22. doi:10.18637/jss.v033.i01

35. 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. doi:10.18637/jss.v039.i05

36. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification And Regression Trees*. 1st ed. Routledge; 2017. doi:10.1201/9781315139470

37. 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;128(584):2145-2166. doi:10.1256/003590002320603584

38. Lindson-Hawley N, Shinkins B, West R, Michie S, Aveyard P. Does cigarette reduction while using nicotine replacement therapy prior to a quit attempt predict abstinence following quit date?: Does cigarette reduction predict abstinence? *Addiction.* 2016;111(7):1275-1282. doi:10.1111/add.13330

- 1
2
3 39. Klemperer EM, Hughes JR. Does the Magnitude of Reduction in Cigarettes Per Day Predict
4 Smoking Cessation? A Qualitative Review. *Nicotine Tob Res*. Published online March 5,
5 2015. doi:10.1093/ntr/ntv058
6
7
8 40. Borrelli B, Gaynor S, Tooley E, Armitage CJ, Wearden A, Bartlett YK. Identification of three
9 different types of smokers who are not motivated to quit: Results from a latent class
10 analysis. *Health Psychol*. 2018;37(2):179-187. doi:10.1037/hea0000561
11
12 41. Hughes JR, Carpenter MJ. The feasibility of smoking reduction: an update. *Addiction*.
13 2005;100(8):1074-1089. doi:10.1111/j.1360-0443.2005.01174.x
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TABLES

Table 1 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)

FIGURE LEGENDS

Figure 1 Overview of methods. **Analysis 1:** Trends in CPD are determined using Latent Class Mixture Models. **Analysis 2:** Class membership is predicted through elastic net regression using baseline and demographic variables. **Analysis 3:** Post-trial smoking cessation is predicted through elastic net regression using the same baseline and demographic variables as in Analysis 2, plus latent class membership from the results of Analysis 1.

Figure 2 Participant record availability flowchart. Participants with baseline CPD values at least one follow-up value were included in analysis of smoking trajectories. For smoking cessation analysis, participants missing CO values were assumed to have continued smoking.

Figure 3 Average smoking trajectories for the 3-class model (i.e., change from baseline smoking rate) ($n = 1783$). Error bars represent SD from average CPD at each trial follow-up point.

Figure 4 Feature importance from prediction of class membership using baseline characteristics (columns **A-D**) and smoking cessation 6 months following the trial (columns **E and F**). Each column represents a separate cross-validated binary logistic regression model. Values represent mean odds ratios for regularized binary logistic regression coefficients across five outer validation folds. Error bars represent the standard deviations from these averages. Dashed lines represent odds ratios of 1, or no effect. **Columns A-D:** Each class was predicted using one-versus-all classification. Values to the left of the dashed lines represent decreased odds of membership to a particular latent trajectory, while values to the right represent increased odds. Notably, participants in Class 1 – those who reduced smoking the most – tended to be slightly older, score lower on anxiety symptoms at baseline, and were more likely to be male. As expected, those in Class 1 were more likely to have received active NRT during the trial. In contrast, participants in Class 3 tended to have lower baseline CPD values, score lower on depression symptoms, were more likely to be female, and to have higher baseline nicotine dependence scores. These participants were also more likely to have received placebo NRT during the trial. **Column B:** Average feature importance for a binary logistic regression model using only the 900 (50.5%) who received placebo NRT, predicting membership to Class 1, the group which reduced their smoking the most. In this group, older participants presenting with lower anxiety scores, and those who have tried to quit before were more likely to reduce their smoking. **Columns E and F:** Values to the left of the dashed lines represent decreased odds of smoking cessation after the trial, while values to the right represent increased cessation odds. When using latent class as a predictor (left column), latent class becomes one of the predominant associations with smoking cessation (Class 2 $OR = 0.111 \pm 0.013$, Class 3 $OR = 0.070 \pm 0.005$), along with receiving active NRT ($OR = 1.90 \pm 0.157$). A large site effect was observed, with participants at sites in Denmark and Germany more likely to quit smoking than those at the USA site.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

Anthony Barrows, MS, Elias Klemperer, PhD, Hugh Garavan, PhD, Nicholas Allgaier, PhD, Nicola
Lindson, PhD, Gemma Taylor, PhD

Latent Class Analysis Model Fit Information

Smoking trajectories were fit to percent change in cigarettes per day (CPD) from baseline using the distributions shown in sFigure 1 for each trial follow-up point except Week 52, which was reserved for smoking cessation prediction.

Updated Smoking Cessation Verification Guidelines

Using ≤ 11 ppm CO as the threshold for biochemically verified smoking cessation, 135/1784 (7.5%) meet abstinence criteria (40.8% of Class 1, 4.7% of Class 2, and 2.6% of Class 3), compared with 122/1784 (6.8%) using < 6 ppm CO (37.6% of Class 1, 4.2% of Class 2, and 2.3% of Class 3). Using < 11 ppm CO as a verification threshold, participants in Classes 2 and 3 were substantially less likely to achieve smoking cessation 6 months following the trial compared to those in Class 1 (Class 2 OR = 0.112 ± 0.057 , Class 3 OR = 0.066 ± 0.007).

Given the relatively small impact of reducing the verification threshold, we used < 6 ppm for smoking cessation predictive modeling, which is consistent with the most recent guidance.⁽¹⁾

Variable Coding

Name		Type	Values
Age		Numeric	Years
Sex		Categorical	<ul style="list-style-type: none">MaleFemale
Age Started Smoking		Numeric	Years
Carbon Monoxide (CO)		Numeric	Parts Per Million (ppm)
Cigarettes Per Day (CPD)		Numeric	Self-report Average
Fagerstrom Test for Nicotine Dependence		Numeric	Summary Score (0-10)
Intention to Quit*		Categorical	<ul style="list-style-type: none">Not at allA littleSomewhatA lot
Last Cigarette Experience (Relief from Smoking Questionnaire)	Overall	Categorical	<ul style="list-style-type: none">Very unpleasantSomewhat unpleasantNeutralSomewhat pleasantVery pleasant
	Pepping-Up Effect		
	Calming Effect		
Longest Period Without Smoking		Categorical	<ul style="list-style-type: none"><1 week1 week – 1 month>1 month – 3 months>3 months
Number of Quit Attempts		Categorical	<ul style="list-style-type: none">NeverOnce2 to 5 times6 to 10 timesMore than 10 times
Time Since Last Quit Attempt		Categorical	<ul style="list-style-type: none">Never quit>12 months>12-24 months>6-12 months0-6 months
Smoking Cessation Quality of Life (SCQoL)	Anxiety	Categorical	<ul style="list-style-type: none">Not at allSomewhatModerately soVery much soExtremely so
	Depression		
RAND 36-Item short-form survey	Emotional Problems	Numeric	See (2)
	Emotional Wellbeing		

	Pain		
	Physical Health		
	Social Functioning		
Study Site		Categorical	<ul style="list-style-type: none"> • Australia • Denmark • Germany • Switzerland • USA
Treatment Group		Categorical	<ul style="list-style-type: none"> • Active • Placebo

*Four of the five included trials captured intention to quit smoking using continuous variables (i.e., “Rate the likelihood that you will quit smoking in the next 6 months,” or “How motivated are you to quit smoking?”), recorded from 0 to 10. In those cases, the continuous scales divided evenly into the four categories described above.

Supplemental Tables

sTable 1 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

Number of Classes	BIC	AIC	Maximum Log-Likelihood
5	-87.61	-169.90	99.95
6	-116.79	-215.54	125.77

***Table 2** *Average AUC across five-fold cross-validation \pm 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**	p
Class 1 vs. All	0.657 \pm 0.027	0.766	<.001
Class 2 vs. All	0.535 \pm 0.015	0.569	.008
Class 3 vs. All	0.567 \pm 0.020	0.585	<.001
Class 1 vs. Class 2	0.570 \pm 0.030	0.784	<.001
Class 1 vs. Class 3	0.670 \pm 0.047	0.788	<.001
Class 2 vs. Class 3	0.530 \pm 0.010	0.523	<.001
Class 1 vs. All (Placebo NRT Only)	0.610 \pm 0.046	0.717	<.001

Supplemental Figure Legends

Figure 1. Distributions of changes in cigarettes per day (CPD) as a percentage of baseline smoking rates (Total N = 1783). A value of 0 represent no change in smoking rate from baseline.

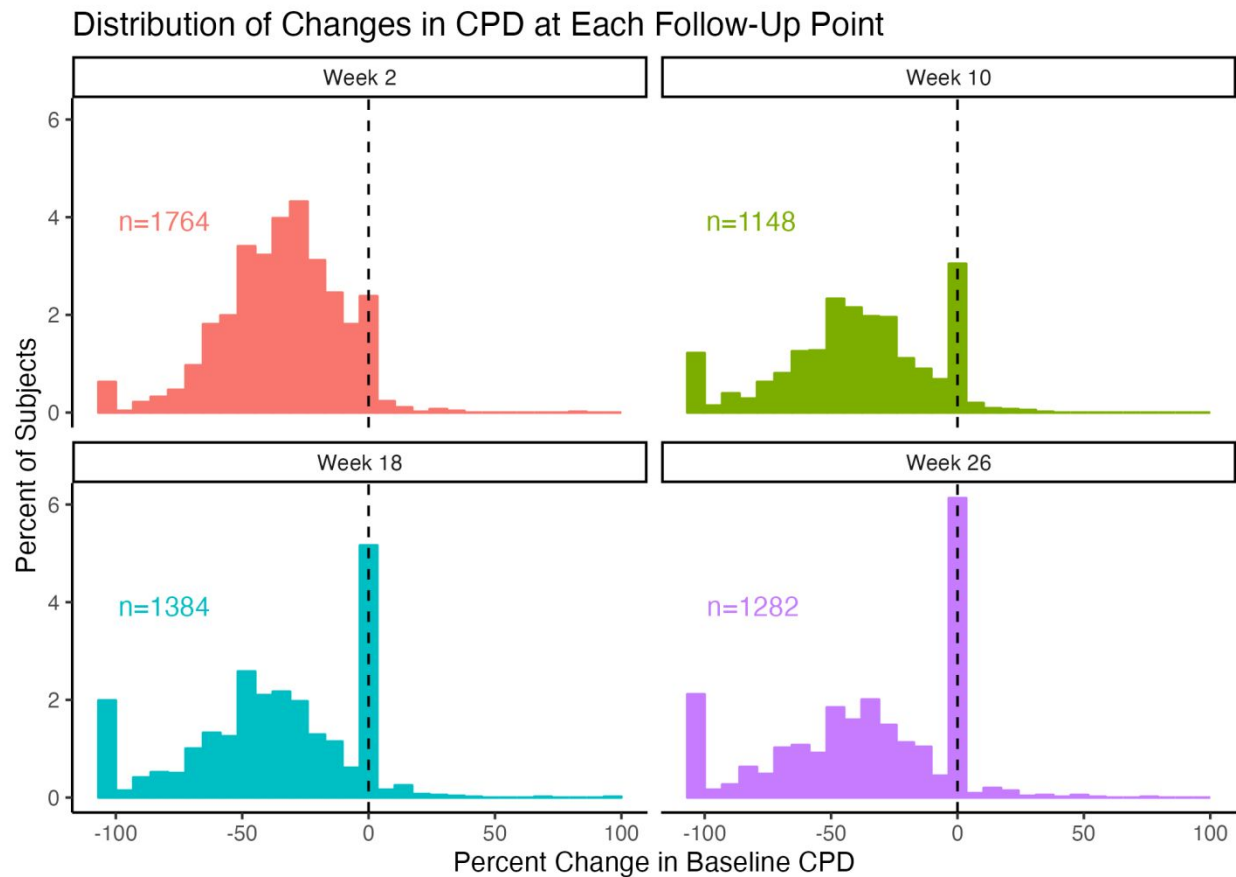


Figure 2. Latent class mixture model BIC curve (n = 1783).

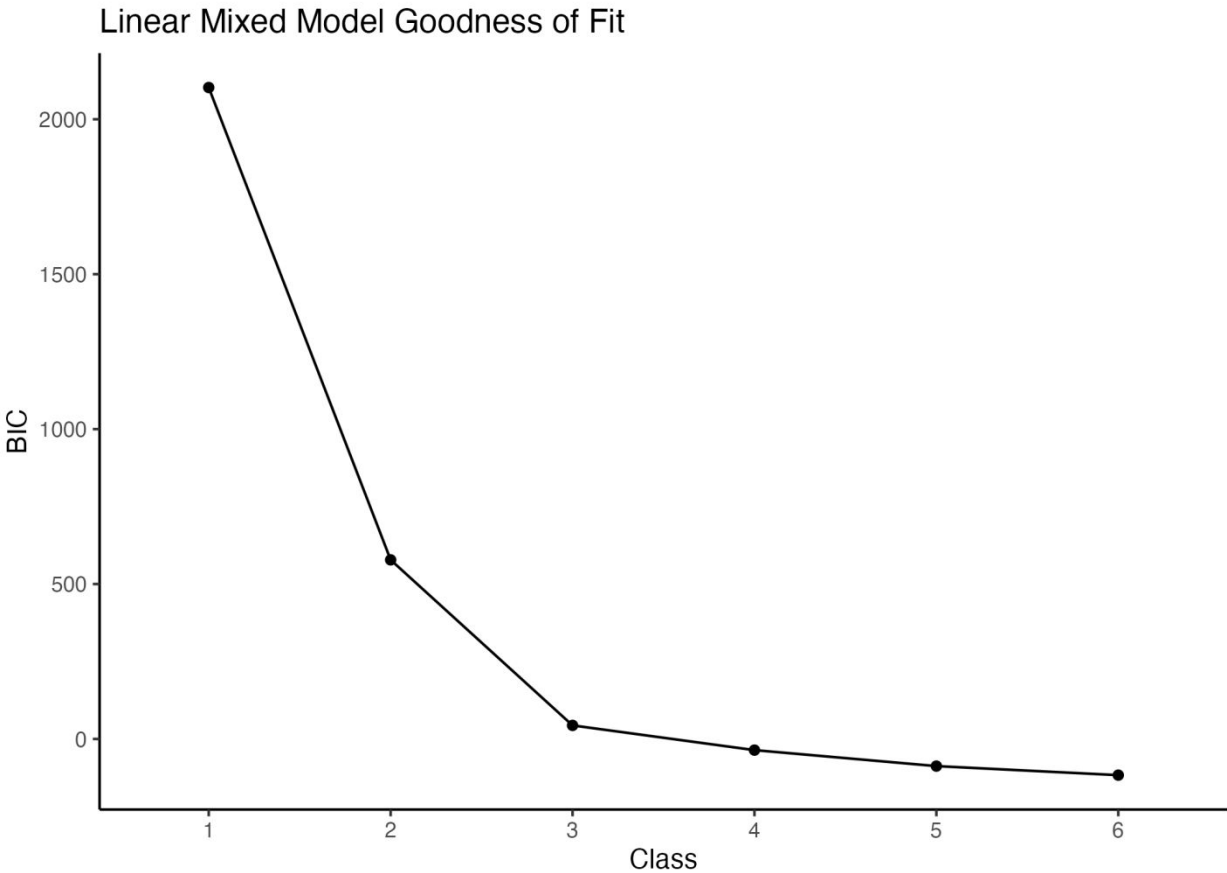
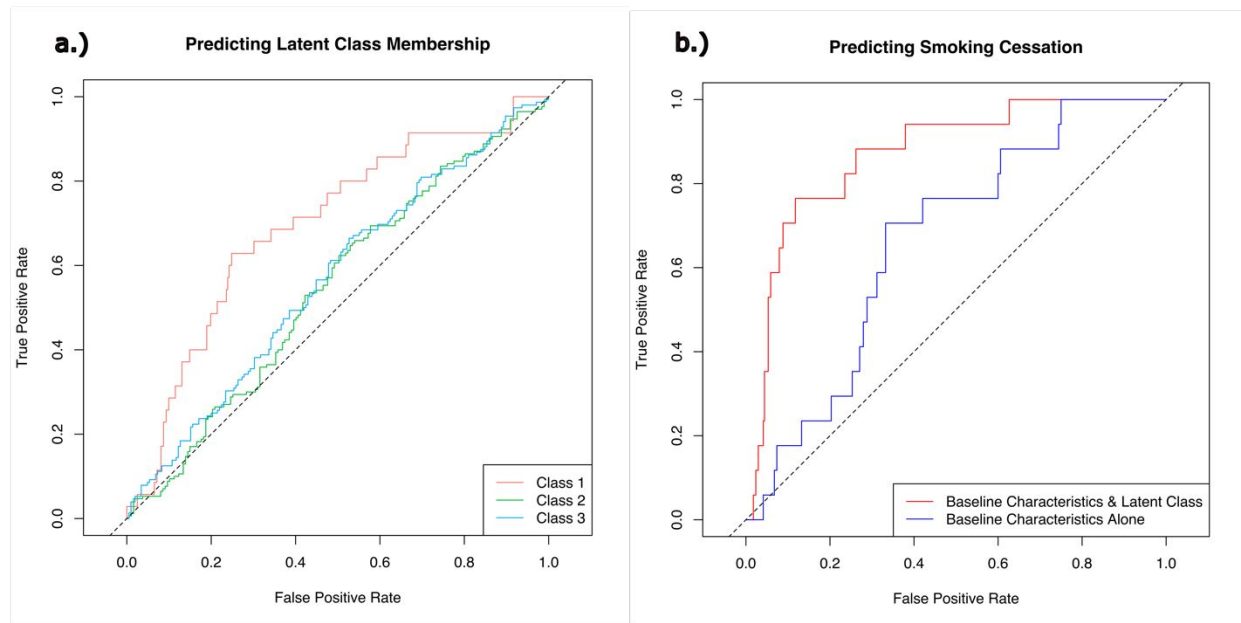


Figure 1. Receiver operating characteristic (ROC) curves for regularized binary logistic regression predictive models. **a.)** Prediction of smoking trajectories as a proportion of baseline smoking during the trial. Each curve represents one-versus-all prediction to latent trajectories 1, 2, and 3 in turn. The model predicting membership to Class 1, the group who reduced the most, performed best ($AUC = 0.657 \pm 0.027$). All three models performed better than classification using a permuted null distribution (p 's $<.001$). **b.)** Smoking cessation prediction 6 months after the trial using participant baseline characteristics alone (blue) and baseline characteristics plus latent class as a predictor (red). Adding latent class as a predictor improved classification performance by an average of 14.4% ($AUC = 0.776 \pm 0.010$, $p = 0.002$), suggesting smoking trajectories among people not looking to quit may be meaningful for long-term cessation outcomes. Each model performed better than classification using a permuted null distribution (p 's $<.001$).



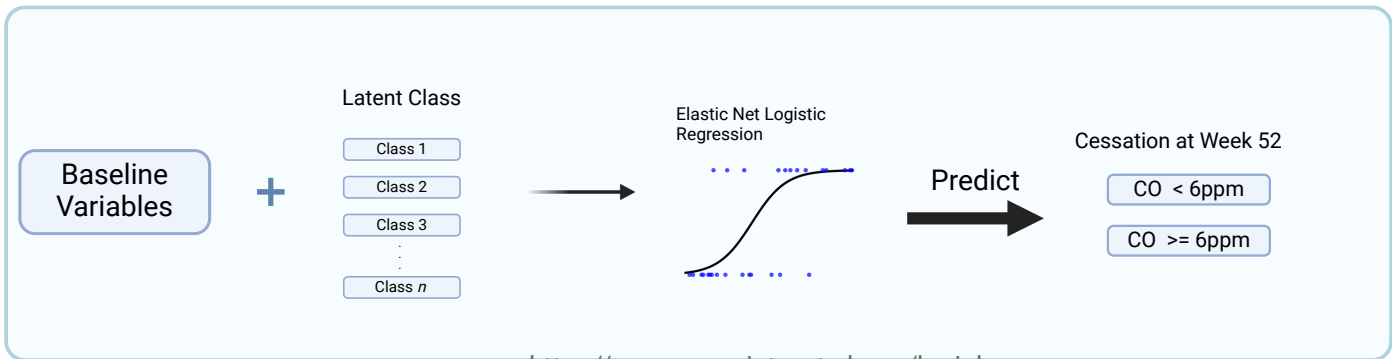
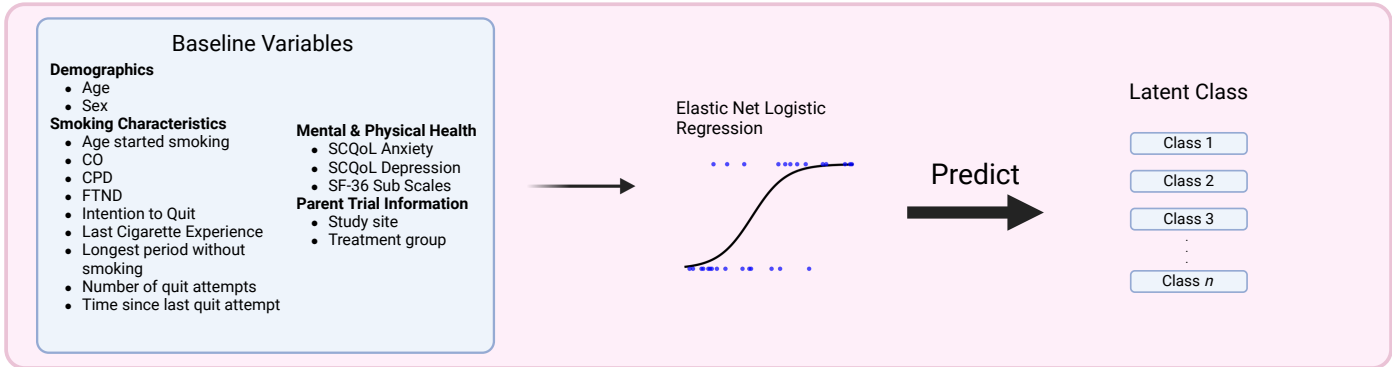
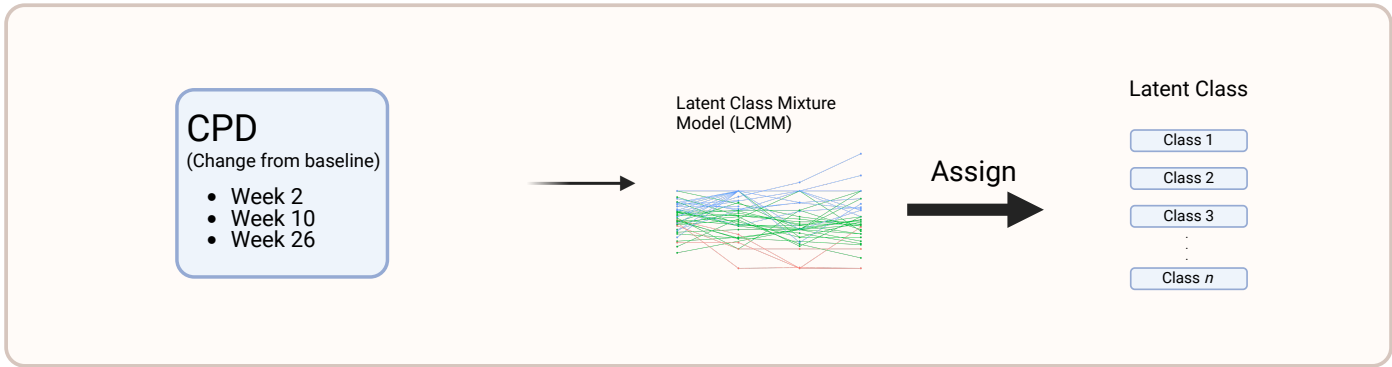
References

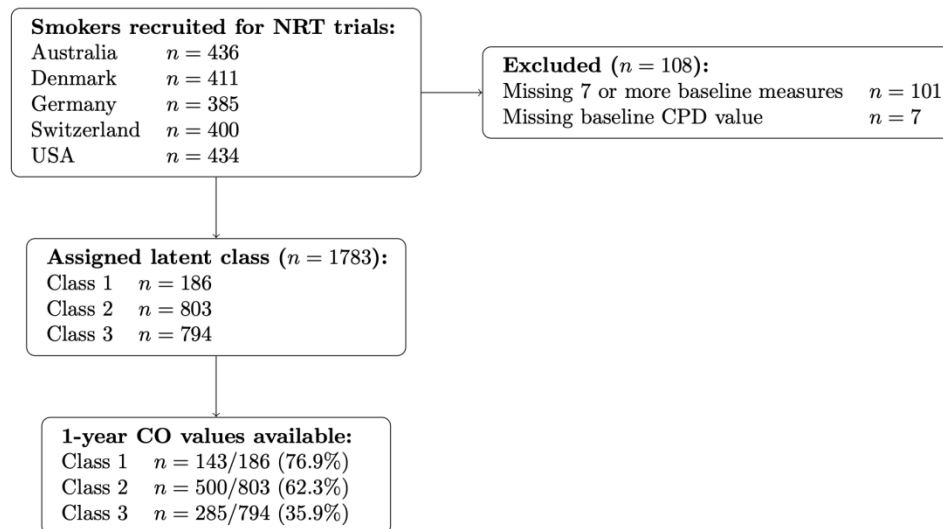
1. Benowitz NL, Bernert JT, Foulds J, Hecht SS, Jacob P, Jarvis MJ, et al. Biochemical Verification of Tobacco Use and Abstinence: 2019 Update. *Nicotine Tob Res.* 2020 Jun 12;22(7):1086–97.
2. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med.* 2001 Jan;33(5):350–7.

Analysis 1

Analysis 2

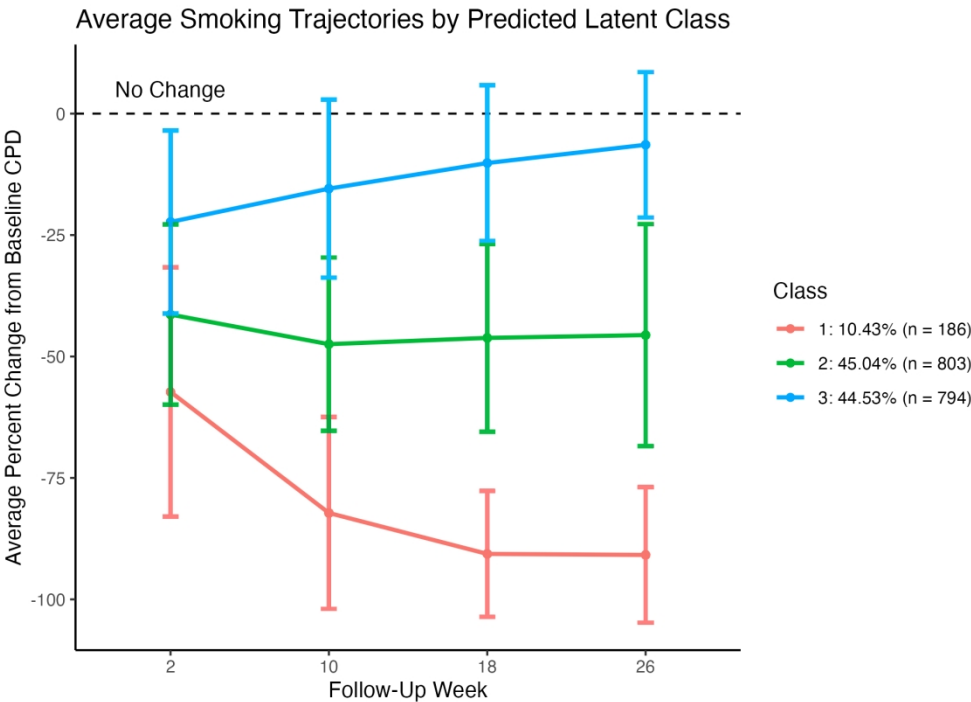
Analysis 3





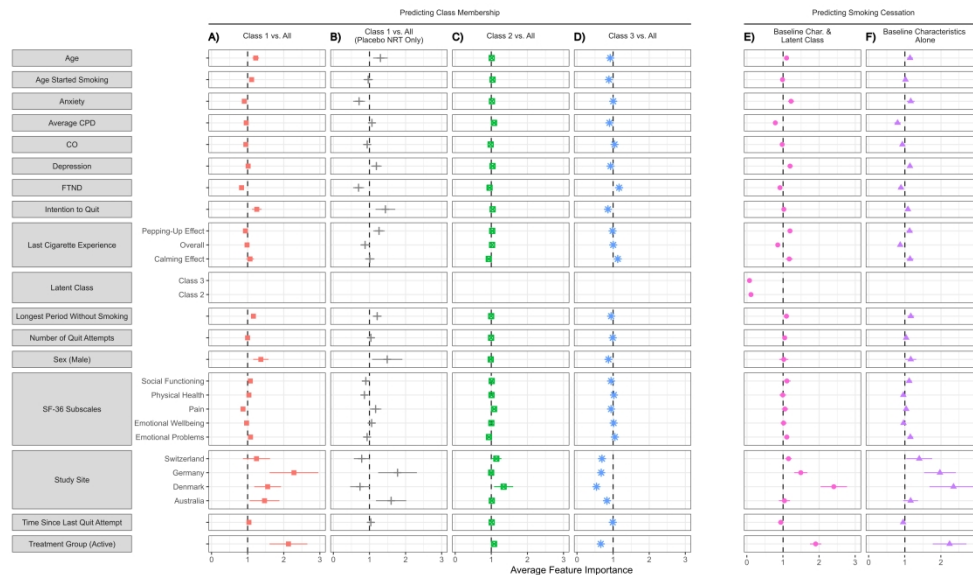
Participant record availability flowchart. Participants with baseline CPD values at least one follow-up value were included in analysis of smoking trajectories. For smoking cessation analysis, participants missing CO values were assumed to have continued smoking.

683x412mm (72 x 72 DPI)



Average smoking trajectories for the 3-class model (i.e., change from baseline smoking rate) (n = 1783). Error bars represent SD from average CPD at each trial follow-up point.

452x322mm (118 x 118 DPI)



Feature importance from prediction of class membership using baseline characteristics (columns A-D) and smoking cessation 6 months following the trial (columns E and F). Each column represents a separate cross-validated binary logistic regression model. Values represent mean odds ratios for regularized binary logistic regression coefficients across five outer validation folds. Error bars represent the standard deviations from these averages. Dashed lines represent odds ratios of 1, or no effect. Columns A-D: Each class was predicted using one-versus-all classification. Values to the left of the dashed lines represent decreased odds of membership to a particular latent trajectory, while values to the right represent increased odds. Notably, participants in Class 1 – those who reduced smoking the most – tended to be slightly older, score lower on anxiety symptoms at baseline, and were more likely to be male. As expected, those in Class 1 were more likely to have received active NRT during the trial. In contrast, participants in Class 3 tended to have lower baseline CPD values, score lower on depression symptoms, were more likely to be female, and to have higher baseline nicotine dependence scores. These participants were also more likely to have received placebo NRT during the trial. Column B: Average feature importance for a binary logistic regression model using only the 900 (50.5%) who received placebo NRT, predicting membership to Class 1, the group which reduced their smoking the most. In this group, older participants presenting with lower anxiety scores, and those who have tried to quit before were more likely to reduce their smoking. Columns E and F: Values to the left of the dashed lines represent decreased odds of smoking cessation after the trial, while values to the right represent increased cessation odds. When using latent class as a predictor (left column), latent class becomes one of the predominant associations with smoking cessation (Class 2 OR = 0.111 ± 0.013 , Class 3 OR = 0.070 ± 0.005), along with receiving active NRT (OR = 1.90 ± 0.157). A large site effect was observed, with participants at sites in Denmark and Germany more likely to quit smoking than those at the USA site.

1026x603mm (118 x 118 DPI)