BMJ Open

Does smoking cessation result in improved mental health? A comparison of regression modelling and propensity score matching

Gemma Taylor,[1,2] Alan Girling,[3] Ann McNeill,[2,4] Paul Aveyard[2,5]

[1]School of Health & Population Sciences, University of Birmingham, Birmingham, UK
[2]UK Centre for Tobacco and Alcohol Studies, Birmingham, UK
[3]School of Health & Population Sciences, University of Birmingham, Birmingham, UK
[4]Institute of Psychiatry, King's College London, London, UK
[5]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

Correspondence to
Dr Gemma Taylor;
gmjtaylor@gmail.com

ABSTRACT

Objectives: Smokers report that smoking is therapeutic; a recent meta-analysis suggests the contrary. However, the association in that review may be explained by group-membership bias and confounding. Propensity score matching (PSM) aims to produce causal estimates from observational data. We examined the association between cessation and change in mental health before and after PSM.

Design: A secondary analysis of prospective data from 5 placebo-controlled randomised trials for smoking reduction.

Participants: All participants were adult smokers and had smoked for at least 3 years. Participants were excluded if they were pregnant, breast feeding, under psychiatric care, deemed to be unfit by a general practitioner or part of a cessation programme. In total, 937 participants provided smoking data at both 6-month and 12-month follow-ups. Of these, 68 were confirmed as abstinent at both 6 and 12 months and 589 as continuous smokers at both follow-ups.

Primary outcome: Change in mental health (36-item Short Form Survey (SF-36), scored 0–100) from baseline (while all participants were smokers) to 12-month follow-up (after cessation) was compared between quitters and continuing smokers with and without adjustment, and after PSM.

Results: Before matching, quitters' mental health scores improved compared with continuing smokers', the mean difference and 95% CI was 5.5 (1.6 to 9.4). After adjustment, the difference was 4.5 (0.6 to 8.5), and after PSM, the difference was 3.4 (−2.2 to 8.9).

Conclusions: Improvements in mental health after smoking cessation may be partly but not completely explained by group membership bias and confounding.

Strengths and limitations of this study

- The largest study to date examining the association between smoking cessation and change in mental health using propensity score matching.
- Use of a psychometrically sound mental health measure, which is sensitive to change.
- Use of propensity score matching to reduce confounding and bias from group membership.
- Presents a low risk of bias according to the Newcastle-Ottawa Scale for Quality Assessment of Observational Studies.
- Attrition was high, although the rate was similar to other studies of smoking interventions.

BACKGROUND

Most smokers want to quit[1 2] but report continuing to smoke because they feel that smoking helps them cope with stress and offers other mental health benefits.[3–9] Our recent systematic review found strong and consistent evidence that the opposite was true.[10] Smokers who quit showed marked

improvements in mental health over time, while smokers who continued smoking showed little change during the same period. We concluded that the strongest explanation for this finding was that cessation caused the improvement in mental health. However, critics countered that membership bias or confounding were possible explanations of the findings.[11] Very few studies in our review made any attempt to control confounding and none addressed membership bias.

As it is not feasible to assign participants randomly to continue smoking or to quit smoking, observational studies are the only source of data to assess the association between smoking and quitting on mental health. Regression modelling is commonly used to account for confounding by adjusting the association of interest for the effect of other variables associated with the outcome and the exposure variable. However, adjustment may not adequately account for membership bias arising from characteristics which differ by smoking status. An alternative method that may account for membership bias as well as confounding is propensity score matching (PSM). PSM involves matching individuals within a sample based on

1

their propensity to belong to an exposure group, or here, matching on the propensity to quit or continue smoking without considering the association of those variables with the outcome.[12] [13] Thus, by balancing covariate distribution between groups, confounding by those variables is eliminated. In addition, PSM can account for some unmeasured factors if they are correlated with observed covariates. Therefore, some unmeasured confounding associated with propensity to quit smoking may also be equalised by this process[13] further reducing bias and providing a more accurate estimate of the association between cessation and change in mental health.[12] [13]

One disadvantage of PSM is that it often reduces the size of the sample available to estimate the strength of the association between cessation and mental health because it requires participants to be matched. If the association between stopping smoking and mental health is influenced by membership bias or other confounding, effect estimates derived from a sample of participants matched on their propensity to quit may show a weaker association. The aim of this study was to estimate the strength of the association between cessation and change in mental health using a regression model adjusting for covariates, and compare this with the estimate derived from PSM.

METHODS

This study followed STROBE reporting guidelines for observational studies.[14] PSM procedures were conducted and reported following criteria outlined by Thoemmes and Kim.[15]

Study design and setting

This was a secondary analysis of prospective individual-level patient data from five merged placebo-controlled randomised trials of nicotine replacement therapy (NRT) for smoking reduction; these data were provided by McNeil pharmaceutical company (see reports of trials for further details: K Haustein. A double-blind, randomized, placebo-controlled multicentre trial of a nicotine chewing gum in smoking reduction. Study ID 980-CHC-9021-0013. Unpublished, 2001 and refs.[16–19]). Attrition rates in these trials were similar to other trials of NRT.[20]

Study size

The trials took place between 1997 and 2003. In total, 937 participants provided data at both 6-month and 12-month follow-ups. Of these, 68 were confirmed as abstinent at both 6 and 12 months and 589 as continuous smokers at both follow-ups. Participants who did not meet exposure criteria were excluded from the analysis.

Participants

All participants were adult smokers of least 3 years and were selected because they wanted to reduce but not

stop smoking. Participants were excluded if they were pregnant or breast feeding, under psychiatric care, deemed to be unfit by a general practitioner, or part of a cessation programme.

At baseline, investigators gathered data on participants' demographic details and age started smoking, cigarettes per day (CPD), nicotine dependence (Fagerström Test of Nicotine Dependency (FTND)),[21] intention to quit, intention to reduce, smoking history (eg, number of previous quit attempts, longest period without smoking), self-rated effects from smoking ('Relief from smoking questionnaire' K Haustein, unpublished data, 2001), mental health (36-item Short Form Survey, SF-36)[22] [23] (see online supplementary table S1). To preserve participant anonymity, some demographic data were unavailable for this secondary analysis.

Participants were followed up 8–10 times over 2 years, and on each occasion, they were encouraged to reduce smoking by using NRT or placebo and given behavioural strategies to assist. In addition, investigators collected data on CPD, 7-day point prevalence abstinence, recorded an expired air carbon monoxide reading (CO), and at baseline and some other follow-ups measured quality of life using the RAND-36. This scale is also known as the SF-36; however, the RAND-36 uses slightly different scoring algorithms.[23] Mental health was measured by the emotional well-being subscale. On this subscale, scores range from 0 to 100 and scores of ≤ 38 indicate a probable mental health problem.[23] In the general population, the subscale mean and SD are 70.4 (22.0). A minimally important difference on this subscale has been defined as a standardised effect size ranging between 0.09 and 0.28.[24]

Exposure

We classified a person as having achieved prolonged cessation if they were abstinent at both 6 and 12 months and were biologically verified to be so on both occasions by having a CO reading of <10 ppm. We classified a person as smoking continuously if they reported smoking at both times and had a CO reading of ≥ 10 ppm. We excluded from the analysis anyone not meeting either definition.

Outcome

The primary outcome was change in mental health from baseline (when all participants were smoking) to 12-month follow-up (after at least 6 months of continuous abstinence or continued smoking).

Analysis

If participants were missing any relevant data, they were excluded from the analysis. In the first model, we used a linear regression to examine the association between cessation and change in mental health. Because of regression to the mean when using within-person, repeated measures data, participants' mean change

scores were not used to measure change in mental health from baseline to follow-up; instead we used follow-up mental health scores, with adjustment for baseline mental health (see Vickers and Altman[25] for further explanation). We used a dummy variable representing cessation. We then adjusted for FTND score and treatment allocation (active/placebo), age, sex and trial ID. Second, we repeated this regression model using propensity score-matched groups (described below). To determine if the association was clinically important, we calculated Cohen's d and 95% CIs using a standard formula,[26–28] before and after matching.[24] Finally, we calculated Bayes factor to determine if the PSM analysis lost power to detect an association or if the PSM analysis indicated a true null result.[29] The Bayes calculation assumed a uniform distribution between zero, no association, and the coefficient for the unadjusted regression.

PSM procedure

PSM involved three steps towards building a logistic regression model to derive predictors of quitting. (1) Propensity scores were developed using covariates which predicted smoking status: nicotine dependency (FTND)[21] and treatment allocation[20] were forced into the model; and a forward stepwise procedure was used to determine whether baseline mental health, intention to quit, intention to reduce, sex, age, smoking history, SF-36, CPD, rated relief from smoking questionnaire, were also associated with quitting at the p<0.10 level. (2) This logistic regression model was combined with the PSMATCH2 command in Stata V.13 to calculate propensity scores representing the estimated probability of quitting contingent on each participant's baseline characteristics.[30] Quitters were matched to the continuing smoker with the closest propensity score on a ratio of 1:1 using a nearest neighbour greedy algorithm, with no replacement, and matching was restricted to within the common support region.[31 32] (3) We performed various checks to ensure the adequacy of the model. We checked the balance of means and variances of

covariates after matching[15 33] by examining the standardised mean differences between smokers and quitters, before and after matching; after matching bias should be ≤5%[34] to determine an adequate model. We calculated the achieved percentage of reduction in bias[34] and examined scatter plots comparing each covariate's standardised per cent bias before and after matching. We also examined the kernel density estimate of the probability distribution of propensity scores before and after matching.

Sensitivity analyses

We developed three sensitivity analyses using different PSM models. Our adjusted regression model was rerun for each sensitivity analysis, and we compared the regression coefficients between the sensitivity models. The trials measured key baseline variables consistently, and each trial also measured some variables in a different manner compared with the other trials. Therefore, we: (1) matched participants across trials including variables measured consistently; (2) matched participants within trials using all relevant variables. We repeated matching across or within trials, with and without common support restrictions. Table 1 summarises the PSM main model and sensitivity analyses.

Assessment of risk of bias

Risk of bias was assessed using an adapted version of the Newcastle-Ottawa Scale for risk assessment of observational studies.[10 35]

RESULTS
Unmatched participants

Sixty-eight participants were biologically validated as continuous quitters and 589 as continuous smokers. Three smokers and one quitter were excluded because of missing baseline data on mental health. Twenty-six smokers were missing outcome data. Smokers excluded for missing data were psychologically healthy at baseline[22]; mental health scores were M (SD)=74.0 (15.7),

Table 1 Summary of PSM main model and sensitivity analyses, and linear regressions conducted after matching

PSM method		Linear regression model in PSM sample		
Matching within/across trials	Within/without common support restriction	Exposure	Outcome	Adjusted for
Within trials	Within*	Smoking status	Mental health (SF-36)† at 12-month follow-up	Baseline mental health, age, sex
	Without	Smoking status	Mental health (SF-36)† at 12-month follow-up	Baseline mental health, age, sex
Across trials	Without	Smoking status	Mental health (SF-36)† at 12-month follow-up	Baseline mental health, age, sex, trial ID
	Within	Smoking status	Mental health (SF-36)† at 12-month follow-up	Baseline mental health, age, sex, trial ID

*Main model.
†Emotional well-being subscale.
PSM, propensity score matching; SF-36, 36-item Short Form Survey.

similar to included smokers. After exclusion for missing data, 67 participants were biologically validated as quit and 560 as smokers. Table 2 displays baseline characteristics of unmatched smokers and quitters. There were differences between groups' FTND scores (p<0.001) and the proportion who received active treatment (p=0.001).

The association between smoking cessation and change in mental health in the whole sample

Mental health scores improved in both groups. The mean change in the quit group was 4.9 (95% CI 1.1 to 8.7) compared with 1.0 (95% CI −0.4 to 2.4) in continuing smokers (figure 1). The difference between groups' follow-up mental health scores adjusted only for baseline mental health was 5.48 (95% CI 1.62 to 9.35, p<0.001). After further adjustment for FTND, treatment status, age, sex and trial ID, the difference between groups remained significant (B=4.53, 95% CI 0.56 to 8.49, p=0.025; table 3).

PSM main model

Table 4 presents ORs and 95% CIs for baseline covariates which predicted smoking status at p<0.10, after forced entry of FTND and treatment allocation. Covariates which predicted smoking status were different within trials and included: FTND scores, active treatment, age started smoking, report of calming effects from smoking, report of unpleasant symptoms from smoking, length of time to last cessation attempt, experience from last cigarette, longest period without smoking, SF-36 mental health (1 trial).

Main model adequacy checks

In all cases, variables which were significantly imbalanced between groups before matching were no longer significantly imbalanced after matching. There were no cases where variables became significantly imbalanced after matching.

Difference in bias between groups after matching was examined to determine which variables were adequately matched. For trial 1, 3/6 variables were adequately matched; trial 4, 0/2; and trial 5, 3/6 variables were adequately matched. In trials 2 and 3, all variables were adequately matched. In sum, matching led to a ≥90% reduction in bias for 13/20 variables (see online supplementary table S2).
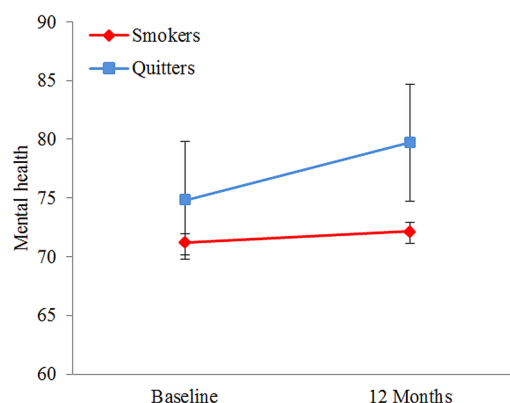
Figure 1 Unmatched groups' mental health scores M (SE) at baseline and follow-up.

As shown in figure 2, there was a common support area to perform PSM, and participants were predominately matched within the common region.

Participants after PSM

The main PSM model included 67 biologically validated continuous quitters who were matched to 67 smokers with similar propensity scores. Sixteen participants, eight per matched sample, were lost as they did not fall within the common support; those excluded did not differ from the included sample at baseline. Excluded participants' baseline mental health scores were: M (SD), 77.0 (14.6) for smokers and 73.5 (16.4) for quitters. FTND scores of excluded smokers were 4.0 (2.4) and 3.4 (1.9) for quitters. Six excluded smokers and seven excluded quitters received active NRT treatment. Before PSM, there were significant differences between the groups' nicotine dependency scores (FTND) and the number of people receiving active treatment (table 2). After matching, the sample became balanced on all baseline characteristics, as displayed in table 5.

The association between smoking cessation and mental health in the PSM sample

After matching, quitters showed an improvement in mental health 4.5 (95% CI 0.4 to 8.7) and smokers displayed a slight worsening in mental health −0.2 (95% CI −4.8 to 4.3; figure 3). The difference between groups after adjustment for baseline values and covariates was 3.37 (−2.15 to 8.90), p=0.229 (table 6).

Table 2 Baseline characteristics in the unmatched sample

	Smokers (n=560)	Quitters (n=67)	Test of significance	p Value
Age, M (SD)	45.6 (10.6)	46.2 (10.2)	T=−0.52	0.607
Sex, % male (N)	48 (258)	52 (35)	χ^2=0.42	0.515
FTND, M (SD)	6.2 (1.9)	5.3 (2.5)	T=3.42	<0.001
SF-36 mental health, M (SD)	71.2 (17.5)	74.8 (13.6)	T=−1.64	0.101
Treatment status, % received active (N)	50 (280)	72 (48)	χ^2=11.42	<0.001

FTND, Fagerström Test of Nicotine Dependency; SF-36, 36-item Short Form Survey.

Table 3 Adjusted linear regression model in whole sample

Predictors	Regression coefficient (B) and 95% CI	p Value
Smoking status	4.53 (0.56 to 8.49)	0.025
Baseline mental health	0.54 (0.46 to 0.61)	<0.001
NRT (active=1)	0.47 (−1.93 to 2.86)	0.703
FTND score	−0.57 (−1.19 to 0.06)	0.062
Sex (female=1)	−2.69 (−5.10 to −0.29)	0.028
Age	0.04 (−0.08 to 0.15)	0.54
Trial ID (reference=1)		
2	6.04 (2.39 to 9.69)	0.001
3	2.84 (−1.04 to 6.72)	0.151
4	1.68 (−1.69 to 5.04)	0.329
5	2.69 (−0.99 to 6.38)	0.152

FTND, Fagerström Test of Nicotine Dependency; NRT, nicotine replacement therapy.

Minimal clinically important difference

Cohen's d for the standardised effect size was d=0.42 (95% CI 0.16 to 0.67) for the unmatched sample, which suggests a clinically important association. After matching, the effect became imprecise 0.14 (95% CI −0.22 to 0.50); however, the minimal clinically important difference (MCID) was still included in the CI.

Bayes factor

Bayes factor was 1.41 which indicated that the PSM data were insensitive, and therefore the PSM results cannot be used to infer that the effect from the regression analysis completely attenuated after matching participants,

but it also does not provide strong evidence that the association is still present after matching.[29]

PSM sensitivity models

When matching was repeated across trials (see online supplementary table S3), variables that contributed to propensity scores were similar when matching was conducted within trials. When the models were run without restricting to the area of common support, smokers and quitters still presented balanced baseline mental health scores (see online supplementary table S4). At follow-up, smokers' mean scores showed little change, and in all analyses, quitters showed a moderate improvement in mental health scores. The coefficients for the difference in mental health at follow-up with baseline adjustment, between quitters and continuing smokers, ranged from: regression coefficient B=3.96 (95% CI −1.00 to 8.93) to 3.55 (−1.29 to 8.40; see online supplementary table S4).

Risk of bias

Risk of bias to the association was assessed using an adapted version of the Newcastle-Ottawa Scale.[10] This study scored 4/5 indicating a low risk of bias (see online supplementary table S5).

DISCUSSION

In this study, regression modelling showed evidence that cessation was associated with improved mental health compared with continuing to smoke, and the direction of the association was not altered by adjustment for confounding. PSM offers potential to control membership

**Table 4** Main PSM model: OR and 95% CI for baseline predictors of smoking status at follow-up

| Trial ID	Covariate	OR (95% CI) (quit=1)	p Value
1	FTND	0.52 (0.36 to 0.74)	<0.001***
	Active treatment (active=1)	0.41 (0.09 to 1.86)	0.250
	Age started smoking	1.38 (1.18 to 1.62)	<0.001***
	Calming effect from smoking	2.15 (1.15 to 4.00)	0.016*
	Number of unpleasant symptoms from smoking	2.17 (1.12 to 4.20)	0.021*
	Time to last cessation attempt	3.74 (0.85 to 16.38)	0.080*
2	FTND	0.87 (0.59 to 1.27)	0.460
	Active treatment (active=1)	11.65 (1.38 to 98.17)	0.024*
	Pleasant experience from last cigarette	2.36 (1.00 to 5.41)	0.043*
3	FTND	0.98 (0.74 to 1.30)	0.907
	Active treatment (active=1)	3.06 (0.77 to 12.14)	0.112
	Longest period without smoking	1.87 (1.07 to 3.26)	0.028*
4	FTND	0.80 (0.55 to 1.16)	0.241
	Active treatment (active=1)	2.91 (0.50 to 16.79)	0.232
5	FTND	0.85 (0.62 to 1.16)	0.316
	Active treatment (active=1)	6.07 (1.75 to 21.05)	0.004**
	Pleasant experience from last cigarette	0.43 (0.24 to 0.77)	0.005**
	SF-36 mental health	1.05 (1.01 to 1.09)	0.021*
	Age started smoking	0.77 (0.59 to 1.00)	0.046*
	Age	1.07 (1.00 to 1.15)	0.043*

See online supplementary table S1 for variable characteristics.
*p<0.1, **p<0.01, ***p<0.001.
FTND, Fagerström Test of Nicotine Dependency; PSM, propensity score matching; SF-36, 36-item Short Form Survey.

Before matching
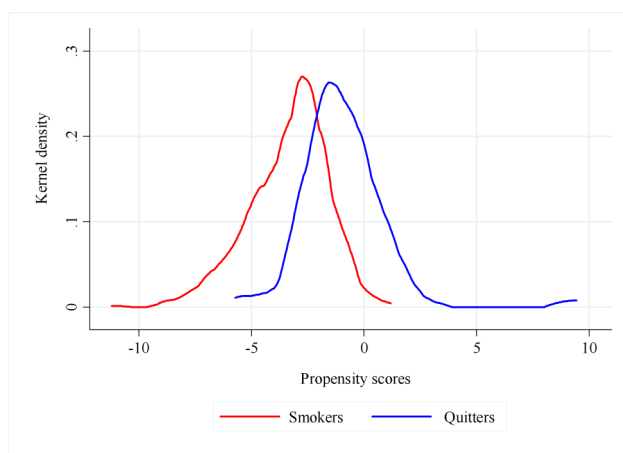
After matching
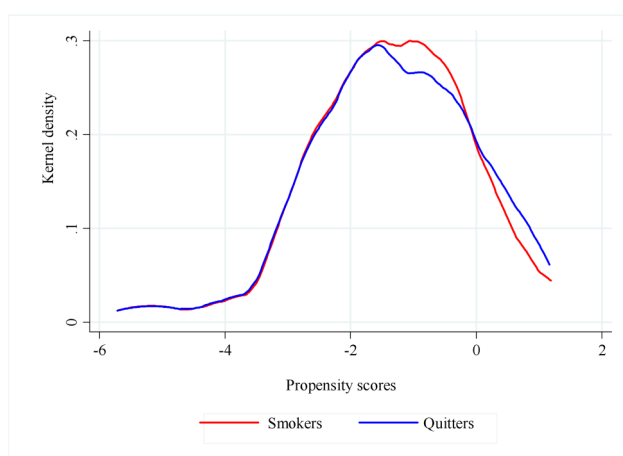
Figure 2 Overlay of Kernel density distributions of quitters' and smokers' propensity scores before and after propensity score matching.

bias as well as confounding. Using this technique, we achieved a good match between smokers that continued smoking and those that stopped. Doing so, the regression coefficient for the difference between smokers and quitters differed only slightly from that achieved by regression methods alone, but it was no longer significant. However, Bayes factor indicated that the PSM data were insensitive and cannot be used to infer that the

effect from the regression was completely attenuated by matching participants.[29]

Strengths and limitations

There were some important strengths of the study. Data were collected in a rigorous manner with clear biologically verified criteria to differentiate continuing smokers from quitters. We included all the key covariates that a systematic review reported were associated with achieving abstinence.[36] Mental health was assessed using a psychometrically sound tool,[22 24] and participants in the trials were not aware of our hypothesis, so there were no demand characteristics that might have biased the results. The Newcastle-Ottawa Score suggested that the results were unlikely to be subject to bias. After PSM, we achieved a good balance of covariates, and extensive sensitivity analysis showed no evidence that the results were sensitive to the methodological decisions we made.

There were some limitations. The regression analysis was based on a large sample and gave sufficient precision to give a statistically significant result. The analysis using PSM necessarily limited the sample size, and the estimate was no longer as precise and was not significant. Importantly though, the direction of the regression coefficient did not change and the size did not change greatly after matching. Although PSM theoretically balances covariates between groups, one cannot be certain, especially with regard to unmeasured variables.[12 13] The overlap of propensity scores in unmatched groups shows there was a small to medium region to conduct matching, and a small common support region may restrict the estimation of a causal effect by changing the observed population.[37] However, sensitivity analyses showed the association was similar regardless of matching within or outside the common region. Furthermore, those excluded during support restrictions had similar baseline characteristics to the entire cohort suggesting that restriction to the common area did not introduce bias. Multiple sensitivity analyses showed no evidence that change in the analysis method altered the effect estimate. The trials included in this analysis measured all key variables consistently; however, some variables were not measured consistently across trials. To investigate whether this inconsistency influenced the results, we conducted sensitivity analyses which indicated that this was unlikely. However, in an ideal

Table 5 Baseline characteristics after PSM

	Smokers (n=59) M (SD)	Quitters (n=59) M (SD)	Test of significance	p Value
Age	48.34 (10.39)	45.76 (8.97)	t=1.44	0.152
Sex (% male)	46	53	χ^2=0.54	0.461
FTND	5.58 (2.13)	5.56 (2.43)	t=0.04	0.968
SF-36 mental health	77.49 (15.07)	74.98 (13.28)	t=0.95	0.340
Treatment status (% received active)	29	31	χ^2=0.04	0.840

FTND, Fagerström Test of Nicotine Dependency; PSM, propensity score matching; SF-36, 36-item Short Form Survey.
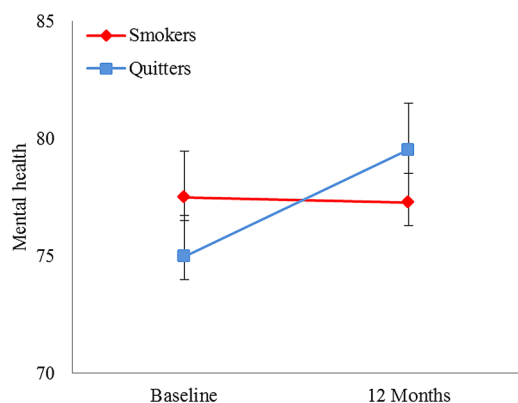
Figure 3 Matched groups' mental health scores M (SE) at baseline and follow-up.

analysis, all variables would have been measured consistently. To protect patient confidentiality, this analysis did not include certain demographic characteristics, such as ethnicity, social class and education, as covariates. These demographics are possible predictors of change in smoking status; however, a recent systematic review did not find consistent evidence to support this.[36] It is also possible that these demographics may predict the likelihood of experiencing change in mental health. However, for example, if ethnicity predicts change in mental health, this is likely to be true in both quitters and smokers. Therefore, although these characteristics may appear important, confounding occurs only if the strength or direction of association changes between baseline and follow-up and that change differs by group. For these reasons, it is unlikely that excluding these characteristics from the analysis influenced the association. However, interactions may still be possible.

Interpretation

The effect sizes reported here are similar to a recently reported systematic review[10] which examined studies similar to this in that mental health was measured before and after cessation in quitters and at corresponding time points in continuing smokers. None of the studies in the review had used PSM and few of them had used regression modelling to adjust for potential confounders. This study therefore adds to previous work in this field by allaying concerns that the apparent benefits of cessation on mental health are spurious and arise instead from differences between those who stop and those who do not stop smoking.

Table 6 Adjusted linear regression model in PSM sample

Predictors	Regression coefficient (B) and 95% CI	p Value
Smoking status	3.37 (−2.15 to 8.90)	0.229
Baseline mental health	0.41 (0.21 to 0.61)	<0.001
Sex (female=1)	−2.75 (−8.23 to 2.74)	0.324
Age	0.11 (−0.17 to 0.40)	0.435

In our systematic review,[10] we proposed that the improvement in mental health occurred because after cessation, regular smokers no longer experienced periods of withdrawal-induced negative affect between cigarettes. This would imply that smoking may be the cause symptoms of depression and anxiety. However, a Mendelian randomisation study found only scant evidence to suggest any association between the genetic instrument and anxiety and depression symptoms[38]; this might suggest that the improvement in mental health that appears to arise after cessation is not due to relief of the smoking-induced withdrawal symptoms.

Implications

The effect size reported here is similar in size to that reported in systematic reviews of the effects of antidepressants for anxiety and depression[39 40] and is larger than that deemed clinically important on the emotional well-being subscale of the RAND-36. This study adds to the growing evidence from observational research that cessation interventions in the general population and in those with mental health problems[10] at least do no psychological harm and may indeed be therapeutic. This evidence is supported by trials of cessation interventions in people with mental health problems, which show no evidence of harm and small suggestions of benefit to mental health from cessation interventions.[41 42]

CONCLUSION

In summary, this study used PSM to try to control membership bias and confounding, and found a similar effect size to that observed using regression modelling alone. This suggests that membership bias or confounding partly but not completely explains the apparent benefit of cessation on mental health and strengthens the case that cessation itself is the cause.

REFERENCES

1. Lader D. *Opinions survey report number 40: smoking related behaviour and attitudes 2008 to 2009*. London, UK, 2009.
2. Centers for Disease Control and Prevention (CDC). Quitting smoking among adults—United States, 2001–2010. *MMWR Morb Mortal Wkly Rep* 2011;60:1513–19.
3. Thompson B, Thompson L, Thompson J, *et al*. Heavy smokers: a qualitative analysis of attitudes and beliefs concerning cessation and continued smoking. *Nicotine Tob Res* 2003;5:923–33.
4. Kerr S, Watson H, Tolson D, *et al*. Smoking after the age of 65 years: a qualitative exploration of older current and former smokers' views on smoking, stopping smoking, and smoking cessation resources and services. *Health Soc Care Community* 2006;14:572–82.
5. Lawn S, Pols R, Barber J. Smoking and quitting: a qualitative study with community-living psychiatric clients. *Soc Sci Med* 2002;54:93–104.
6. Fidler J, West R. Self-perceived smoking motives and their correlates in a general population sample. *Nicotine Tob Res* 2009;11:1182–8.
7. McEwen A, West R, McRobbie H. Motives for smoking and their correlates in clients attending Stop Smoking treatment services. *Nicotine Tob Res* 2008;10:843–50.
8. Ferguson J, Bauld L, Chesterman J, *et al*. The English smoking treatment services: one-year outcomes. *Addiction* 2005;100:59–69.
9. Ng D, Jeffery R. Relationships between perceived stress and health behaviors in a sample of working adults. *Health Psychol* 2003;22:638–42.
10. Taylor G, McNeill A, Girling A, *et al*. Change in mental health after smoking cessation: systematic review and meta-analysis. *BMJ* 2014;348:g1151.
11. Sanderson S, Taylor A, Munafo M. Smoking cessation has not been shown to have an 'Effect' on mental health. *BMJ* 2014;348:g1151.
12. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
13. Rosenbaum P, Rubin D. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984;79:516–24.
14. Von Elm E, Altman D, Egger M, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Prev Med* 2007;45:247–51.
15. Thoemmes F, Kim E. A systematic review of propensity score methods in the social sciences. *Multivariate Behav Res* 2011;46:90–118.
16. Wennike P, Danielsson T, Landfeldt B, *et al*. Smoking reduction promotes smoking cessation: results from a double blind, randomized, placebo-controlled trial of nicotine gum with 2-year follow-up. *Addiction* 2003;98:1395–402.
17. Bolliger C, Zellweger J, Danielsson T, *et al*. Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety. *BMJ* 2000;321:329–33.
18. 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;8:555–64.
19. Batra A, Klingler K, Landfeldt B, *et al*. Smoking reduction treatment with 4-mg nicotine gum: a double-blind, randomized, placebo-controlled study. *Clin Pharmacol Ther* 2005;78:689–96.
20. Stead L, Perera R, Bullen C, *et al*. Nicotine replacement therapy for smoking cessation. *Cochrane Database Syst Rev*. 2012;11:CD000146.
21. Fagerstrom K, Heatherton T, Kozlowski L. Nicotine addiction and its assessment. *Ear Nose Throat J* 1990;69:763–5.
22. Hays R, Prince-Embury S, Chen H. *RAND-36 health status inventory*. San Antonio, TX: The Psychological Corporation, 1998.
23. Hays R, Sherbourne C, Mazel R. The RAND 36-item health survey 1.0. *Health Econ* 1993;2:217–27.
24. Hays R, Morales L. The RAND-36 measure of health-related quality of life. *Ann Med* 2001;33:350–7.
25. Vickers A, Altman D. Analysing controlled trials with baseline and follow up measurements. *BMJ* 2001;323:1123–4.
26. Huber C. http://blog.stata.com/2013/09/05/measures-of-effect-size-in-stata-13/:stata.com. 2013 (cited 2015).
27. STATA. esize—Effect size based on mean comparison. 2014. http://www.stata.com/stata13/effect-sizes/
28. Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
29. Dienes Z. Using Bayes to get the most out of non-significant results. *Front Psychol* 2014;5:781.
30. Leuven E, Sianesi B. *PSMATCH2: STATA module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Boston, MA: Boston College Department of Economics, 2003.
31. Augurzky B, Kluve J. Assessing the performance of matching algorithms when selection into treatment is strong. *J Appl Econ* 2007;22:533–57.
32. Gu XS, Rosenbaum PR. Comparison of multivariate matching methods: structures, distances, and algorithms. *J Comput Graph Stat* 1993;2:405–20.
33. Thoemmes F. Propensity score matching in SPSS. arXiv. 2012;1201 (6385).
34. Rosenbaum P, Rubin D. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat* 1985;39:33–8.
35. Wells G, Shea B, O'Connell D, *et al*. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa: 2010.
36. Vangeli E, Stapleton J, Smit E, *et al*. Predictors of attempts to stop smoking and their success in adult general population samples: a systematic review. *Addiction* 2011;106:2110–21.
37. King G, Zeng L. The dangers of extreme counterfactuals. *Polit Anal* 2005;14:131–59.
38. Bjorngaard J, Gunnell D, Elvestad M, *et al*. The causal role of smoking in anxiety and depression: a Mendelian randomization analysis of the HUNT study. *Psychol Med* 2013;43:711–19.
39. Fournier J, DeRubeis R, Hollon S, *et al*. Antidepressant drug effects and depression severity. *JAMA* 2010;303:47–53.
40. Excellence NIfHaC. *Generalised anxiety disorder in adults: management in primary, secondary and community care*. London, UK, 2011.
41. Ragg M, Gordon R, Ahmed T, *et al*. The impact of smoking cessation on schizophrenia and major depression. *Australas Psychiatry* 2013;21:238–45.
42. Banham L, Gilbody S. Smoking cessation in severe mental illness: what works? *Addiction* 2010;105:1176–89.