

## RESEARCH ARTICLE

# Heterogeneous treatment effects of a text messaging smoking cessation intervention among university students

Marcus Bendtsen \*

Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

\* [marcus.bendtsen@liu.se](mailto:marcus.bendtsen@liu.se)

## Abstract

### Introduction

Despite tobacco being an important preventable factor with respect to ill health and death, it is a legal substance that harms and kills many of those who use it. Text messaging smoking cessation interventions have been evaluated in a variety of contexts, and are generally considered to have a positive effect on smoking cessation success. In order for text messaging interventions to continue to be useful as prevalence of smoking decreases, it may be necessary to tailor the interventions to specific individuals. However, little is known with regard to who benefits the most and least from existing interventions.

### Methods

In order to identify heterogeneous treatment effects, we analyzed data from a randomized controlled trial of a text messaging smoking cessation intervention targeting university students in Sweden. We used a Bayesian hierarchical model where the outcome was modelled using logistic regression, and so-called horseshoe priors were used for coefficients. Predictive performance of the model, and heterogeneous treatment effects, were calculated using cross-validation over the trial data.

### Results

Findings from the study of heterogeneous treatment effects identified less effect of the intervention among university students with stronger dependence of nicotine and students who smoke a greater quantity of cigarettes per week. No heterogeneity was found with respect to sex, number of years smoking, or the use of snuff.

### Discussion

Results emphasize that individuals with a more developed dependence of nicotine may have a harder time quitting smoking even with support. This questions the dissemination and development of text messaging interventions to university students in the future, as they may not be the optimal choice of intervention for those with a more developed

## OPEN ACCESS

**Citation:** Bendtsen M (2020) Heterogeneous treatment effects of a text messaging smoking cessation intervention among university students. *PLoS ONE* 15(3): e0229637. <https://doi.org/10.1371/journal.pone.0229637>

**Editor:** Wisit Cheungpasitporn, University of Mississippi Medical Center, UNITED STATES

**Received:** October 22, 2019

**Accepted:** February 10, 2020

**Published:** March 5, 2020

**Copyright:** © 2020 Marcus Bendtsen. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Ethical approval of the trial was given under the condition that data be stored only on Linköping University servers, thus we cannot share dataset freely. Request for data can be done via email to corresponding author [marcus.bendtsen@liu.se](mailto:marcus.bendtsen@liu.se) or to [registrator@liu.se](mailto:registrator@liu.se).

**Funding:** The author is funded by a grant from the Swedish Research Council for Health, Working Life and Welfare (FORTE, Dnr: 2018-01410).

**Competing interests:** MB owns a private company that develops and distributes lifestyle interventions to be used in health care settings. This does not

alter our adherence to PLOS ONE policies on sharing data and materials.

dependence. On the other hand, text messaging interventions may be useful to disseminate among university students that are at risk of developing a strong dependence.

## Trial registration

International Standard Randomized Controlled Trial Number (ISRCTN): 75766527; <http://www.controlled-trials.com/ISRCTN75766527>.

## Introduction

Among the factors considered in the Global Burden of Diseases, Injuries, and Risk Factors Study [1], smoking ranks number two when ordered by risk-attributable disability adjusted life years. Some of the non-communicable diseases for which the risk is higher among smokers include cardiovascular and respiratory diseases, diabetes, and cancer. Despite tobacco being an important preventable factor with respect to ill health and death [2], it is a legal substance that harms and kills many individuals when used as intended by manufacturers [3].

In Sweden, a steady decline in smoking prevalence has been recorded over the past decade by the Public Health Agency of Sweden [4]. The most recent data from 2018 indicate that the prevalence rate is as low as 7% in the general population. This means that we are closer than ever to eradicating one of the most important causes of disease in Sweden. However, adolescents and young adults still start smoking, and many will continue smoking into adulthood with a more severe nicotine dependence. There is therefore still a need for effective smoking cessation interventions that can scale to a national level and that are designed to reach adolescents and young adults.

One promising approach to help young individuals quit smoking is to use text messaging [5–7]. Text messaging interventions typically consist of a series of messages sent to participants' mobile phones over the course of 8–12 weeks. These messages motivate participants to make a quit attempt and then reinforce and support this decision throughout the intervention period. In addition, text messaging interventions may also increase access to education and support services that promote smoking cessation [6].

Several randomized controlled trials (RCTs) have been conducted to estimate the effect of text messaging interventions for smoking cessation [8–11], notably the txt2stop trial [9] ( $n = 5800$ ), which found strong evidence in favor of the intervention with respect to both biochemically verified abstinence (OR 2.20, 95% CI 1.80–2.68,  $P$ -value  $< .0001$ ) and self-reported abstinence (OR 1.47, 95% CI 1.40–1.66,  $P$ -value  $< .0001$ ). Three meta-analyses have concluded that text messaging interventions have a positive effect on smoking cessation: one reported a summary effect size of 0.25 (95% CI 0.13–0.38) [5], the second meta-analysis reported an overall summary OR of 1.37 (95% CI 1.25–1.51) of smoking cessation in favor of text messaging interventions [6], and the third analysis similarly found that quit rates were higher among those who had access to text messaging interventions (OR 1.36, 95% CI 1.23–1.51) [7]. Thus, there exists a relatively strong body of evidence for the average treatment effect of text messaging interventions.

As is often the case in RCTs, the effects of text messaging interventions are generally assessed at the group level, which is evident when inspecting the individual trials included in meta-analyses [5–7]. However, we do expect some individuals to benefit more than others from an intervention, and some may even be harmed, as not all individuals with access to

smoking cessation interventions successfully quit smoking. Treatment effects may therefore be considered heterogeneous within a group.

As prevalence rates of smoking in Sweden decrease [4], we may find that interventions need to become more targeted towards groups that persist with smoking despite current policy and interventions. It is therefore important to identify who benefits the most and least from existing interventions. While subgroup analyses are sometimes included in reports, these are problematic as they may fail to find effects in smaller subgroups, subgroups may be arbitrarily defined, and they are susceptible to false positives due to multiple hypothesis testing and confounding [12–16]. We therefore need to take a different approach and make predictions which estimate the outcome for an individual in both settings being contrasted. To do this, we need to create a prediction model from the data that we have collected in an RCT, and then assess this model's performance in order for any findings relating to heterogenous treatment effects to be useful.

This report communicates the results from a study of heterogenous treatment effects that aimed to identify which university students benefit the most and the least from a Swedish text messaging smoking cessation intervention called NEXit. The intervention consists of a 1- to 4-week motivational phase, during which participants can commit to a quit date, followed by a 12-week program with supportive and reinforcing text messages (a total of 157 messages). Data for the included study was taken from the NEXit trial, which estimated the average treatment effect of NEXit among Swedish university students in an RCT [11,17].

## Material and methods

The NEXit trial was a 2-arm randomized trial of a text messaging intervention (the NEXit intervention) [11,17–19]. The trial was conducted simultaneously at all universities and colleges in Sweden (except one which participated in the pilot study). Students were emailed invitations over a 3-week period (October 23 to November 13, 2014). The trial was approved by the Regional Ethical Committee in Linköping, Sweden (Dnr 2014/217-31), and was prospectively registered (ISRCTN75766527).

A total of 1590 students were randomized (eligible due to being daily or weekly smokers), of which 827 were allocated to the intervention group and 763 to the control group (waiting list). Approximately 70% were female, and the majority were between 21 to 25 years old.

At 3-months after randomization, follow-up data was collected from 1502 students (94.5%, 1502/1590), loss to follow-up was due to participants not being contactable by phone. Follow-up rates did not differ significantly between groups and attrition analyses did not identify any systematic difference between those who did and did not complete follow-up.

The primary outcome measure in the NEXit trial was subjective reporting of prolonged abstinence, following the Russel standard definition [20], as not having smoked more than 5 cigarettes the past 8 weeks. The intervention was found to have a statistically significant effect on prolonged abstinence, with an OR of 2.05 (95% CI 1.57–2.67, P-value < .001). Prevalence of prolonged abstinence in the intervention group was 25.9% and in the control group 14.6%, giving an average treatment effect estimate of 11.3 percentage points absolute difference. For more information regarding the NEXit intervention and the RCT we refer the interested reader to published reports [11,17–19].

## Model

For this study on heterogenous treatment effects, we used the same logistic regression model used in the original group comparison analysis, with the same covariates and primary outcome. However, a Bayesian approach to coefficient inference was taken. This was done in

order to: (1) calculate predictions using the joint posterior predictive distribution over coefficients, which may produce better predictions than approaches relying on point estimates [21]; (2) incorporate uncertainty regarding which coefficients to include in the model using shrinkage priors (see later description). Posterior distributions were calculated using Hamiltonian Monte Carlo (HMC) [22], a Markov chain Monte Carlo method particularly used in the probabilistic programming language STAN [23]. So called *horseshoe* priors were used for covariate coefficients, which favors shrinking coefficients towards zero, encoding an a-priori belief that single covariates do not dominate the odds of smoking cessation.

The full Bayesian model is presented in Eq 1, with priors as recommended by [24]. In the equation, the covariates are as follows:  $G$  = group allocation,  $S$  = sex,  $Y$  = the number of years smoked,  $W$  = the number of cigarettes smoked weekly,  $D$  = dependence according to Fagerström's test for nicotine dependence,  $U$  = the amount of snuff used. Coefficients (except for the intercept) are given normal priors with variance defined by a global parameter  $\tau$  and local parameters  $\lambda_{1-6}$ . Half-Cauchy priors for  $\tau$  and  $\lambda_{1-6}$  are set so that  $\tau$  shrinks all coefficients towards zero and  $\lambda_{1-6}$  allow for individual coefficients to avoid this shrinkage. The scalar  $\tau_0$  is defined by  $\tau_0 = \frac{p_0}{K-p_0} \times \frac{2}{\sqrt{N}}$ , where  $p_0$  represents the number of covariates that a-priori are believed to be effectively non-zero,  $K$  the number of covariates, and  $N$  the number of records in the dataset. In our case, we replaced categorical variables ( $S$ ,  $D$  and  $U$  in Eq 1) with dummy variables representing the levels of the original variables. This resulted in 11 columns in the dataset (ie.  $K = 11$ ), and we decided to set  $p_0 = 6$ , which equals the number of columns in our dataset divided by two (rounded up). The reasoning for this is that we do not believe that all covariates should be included in the model (or else we would not have included shrinkage priors), nor do we a-priori believe that none of the covariates should be included (or we would not have included them at all). Instead our best guess, prior to seeing any data, is to assume that half of them should remain. The inference is quite insensitive to the choice of  $p_0$ , especially when the number of records available far outweighs the number of covariates, as it does in our case. A detailed discussion about this insensitivity can be found in [24].

We used precompiled STAN code supplied in the R package `rstanarm`. R version 3.5.3 and `rstanarm` 2.18.2 was used for all inference. HMC was used to collect 2000 samples from the joint posterior distribution for each iteration of the cross-validation procedure. The procedure outlined in this Methods section took approximately 20 minutes to run on a 2017 MacBook Pro.

$$\begin{aligned}
 \text{prolonged abstinence} &\sim \text{Bernoulli}(p) \\
 \log\left(\frac{p}{p-1}\right) &= \beta_0 + \beta_1 G + \beta_2 S + \beta_3 Y + \beta_4 W + \beta_5 D + \beta_6 U \\
 \beta_0 &\sim t_7(0, 2.5) \\
 \beta_{1-6} | \tau, \lambda_{1-6} &\sim \mathcal{N}(0, \tau \lambda_{1-6}) \\
 \lambda_{1-6} &\sim \text{Half}\mathcal{C}(0, 1) \\
 \tau &\sim \text{Half}\mathcal{C}(0, \tau_0^2)
 \end{aligned} \tag{1}$$

Eq 1. Logistic regression model for smoking cessation using horseshoe priors.

## Predictive performance and cross-validation

In order to validly explore heterogeneous treatment effects, we have to make predictions that are accurate. If a model tells us that 15% of the individuals with access to a certain treatment will report prolonged smoking abstinence, then we also expect 15% of these individuals to

actually have been abstinent, not 5% or 25%. We refer to a model's capacity to produce accurate probabilities as the model's predictive performance.

A visualization of the procedure we use to assess predictive performance using a technique known as cross-validation is shown in Fig 1. Each circle in Step 1 represents a participant in the RCT. In Step 2, we randomly split the participants into two subsets, one which contains 90% of the participants, and one that contains the remaining 10%. We call these two subsets the *learning data* (the larger subset) and the *testing data* (the smaller subset).

Using the learning data, we run HMC to infer posterior distributions of the covariates in Eq 1, and thereby define a model  $M$  in Step 3. We do not involve the participants from the testing data when creating this model, but rather we treat them as if they were never part of the RCT at all. In Step 4, we input each participant withheld in the testing data into  $M$ , one at a time, and output a predicted probability that this individual will quit smoking. In Step 5, we place each participant from the testing data into a bin. The first bin is labelled 0%-10%, the next is labeled 10%-20%, and so on until we get the final bin labeled 90%-100%. We place each individual in the bin that corresponds to the prediction that we made for them, eg. if  $M$  predicted that an individual had a 15% probability of prolonged abstinence, then we place this individual in the bin labelled 10%-20%.

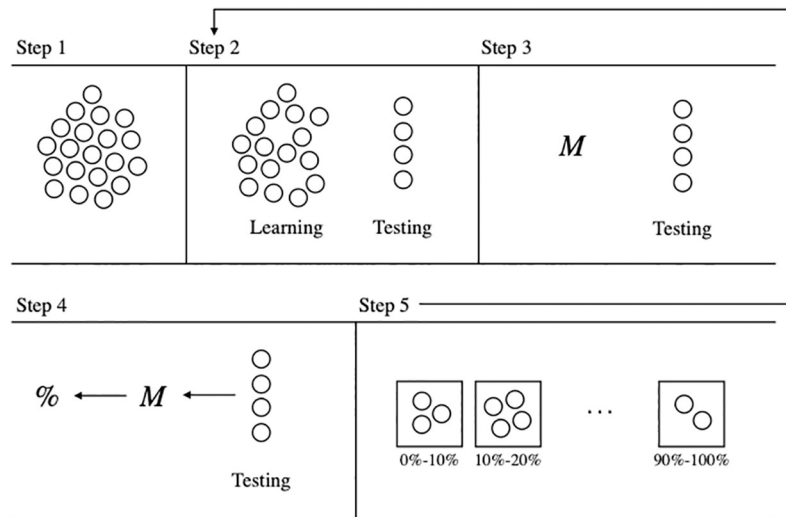
We then go back and repeat the procedure from Step 2 onwards in such a manner that each participant will be part of the testing data exactly once. This results in each participant being placed in one of the ten bins in Step 5.

Predictive performance of the model can be assessed by looking into the bins. The average prediction made within each bin should correspond to the actual ratio of occurrence of the outcome being predicted. For instance, assuming that we on average predicted a 15% probability of prolonged abstinence for the participants that we placed in the 10%-20% bin, then we expect 15% of these participants to actually have reported prolonged abstinence. We can make these calculations since we have follow-up data from the RCT, and a model's predictive performance can be judged by how well the predictions are aligned with the empirical findings.

Note, when this procedure is operationalized it is common to use locally weighted scatterplot smoothing (lowess) in order to visually show a model's predictive performance. Doing so allows us to examine the model's performance across a range of predicted values rather than within a set of predefined bins [25,26]. The visualization using bins is however still useful to understand the concept behind what we refer to as predictive performance, and one may think about using lowess as using an infinite number of bins.

## Individual treatment effects

Estimates of individual level treatment effects will be calculated during the cross-validation procedure described earlier. In Step 4 (see Fig 1), we predict the probability of prolonged abstinence (Russel standard [20], no more than 5 cigarettes the past week) for each individual given both the intervention and control setting. By subtracting these two probabilities from each other we get the individual level effect of the intervention in terms of absolute difference in percentage points. For instance, if an individual is predicted to have a 15% probability of prolonged abstinence if they are given access to the intervention, and a 5% without access, then the probability of prolonged abstinence increases with 10 percentage points if the individual has access to the intervention. In this context, access to the intervention is defined as being offered the intervention, as this reflects the intervention group in the NEXit trial [11].



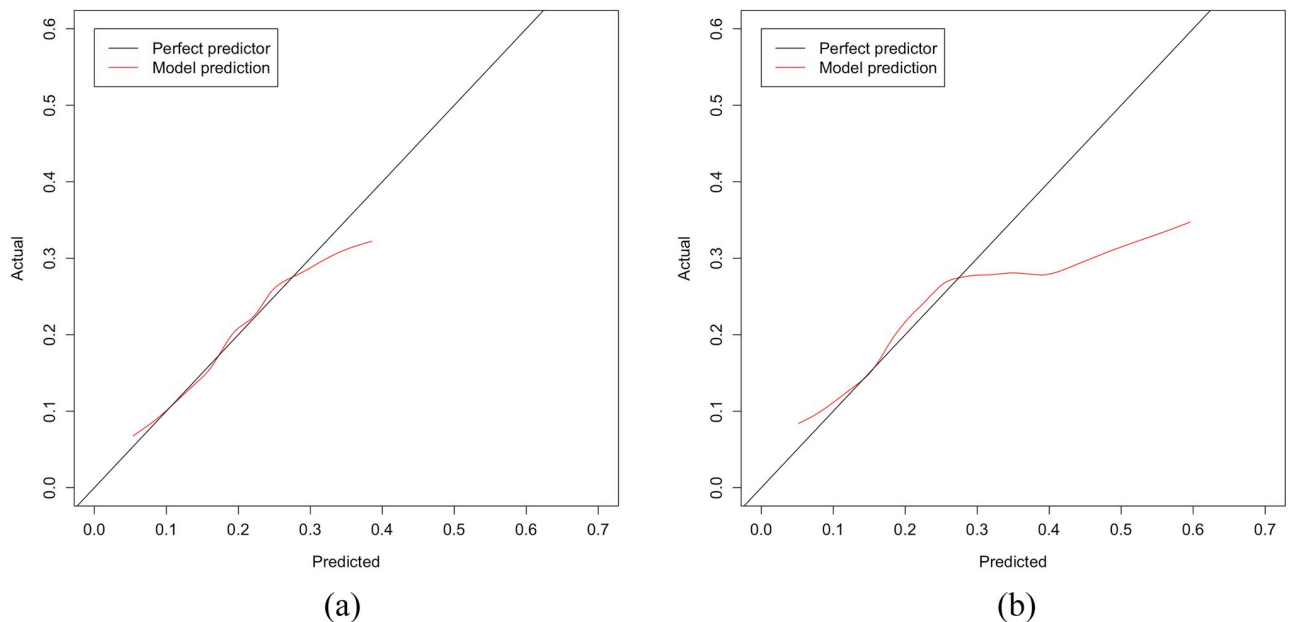
**Fig 1. Cross-validation procedure to assess the performance of a prediction model.**

<https://doi.org/10.1371/journal.pone.0229637.g001>

## Results

### Predictive performance

Fig 2(a) illustrates the predictive performance of the model in Eq 1 on the withheld testing data during cross-validation. The black line represents a perfect model which predicts probabilities that are accurate to what has been observed empirically, the closer we get to this line the better. The red line represents the model learnt using the procedure outlined in Fig 1. The line is drawn using lowess, and as described earlier, this can be understood as using an infinite



**Fig 2. The red line will overlap the black line when predictions are accurate.** (a) Predictive performance of model in Eq 1 on the withheld testing data during cross-validation using Bayesian inference. (b) Predictive performance of the same model using maximum likelihood estimates.

<https://doi.org/10.1371/journal.pone.0229637.g002>



number of bins rather than choosing a fixed number. The model's performance is close to the black line; thus, it tells us that when the model predicts that an individual has a certain probability of prolonged abstinence, then this is also the probability we should expect to see empirically.

For reference, we ran the procedure in Fig 1 using the model used in the original publication [11], ie. the same model as in Eq 1 but using maximum likelihood estimates rather than Bayesian inference. Fig 2(b) visualizes the predictive performance of this model. It is clear that this model produces erroneous predictions, as the red line indicates that optimistic predictions are being made, for instance some cases are predicted to have a 60% probability of prolonged abstinence despite this not being observed empirically. This should serve as a reminder that while it is common to report maximum likelihood estimates of coefficients when contrasting groups in trials, it is not necessarily the case that these models have good predictive performance.

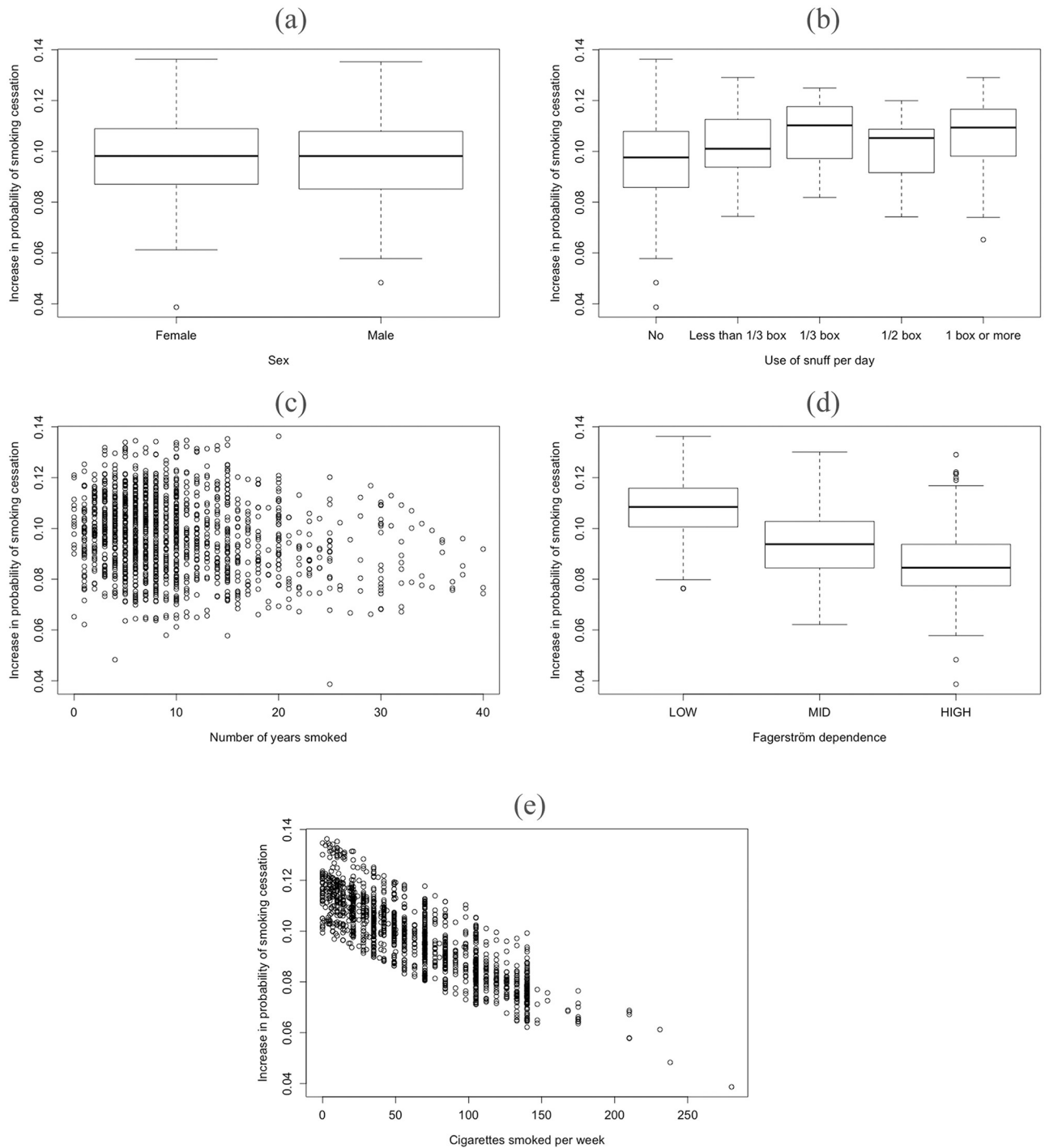
### Conditional average treatment effects

In Fig 3(a)–3(e) we have plotted the absolute difference in predicted prolonged abstinence (no more than 5 cigarettes the past 8 weeks) against each of the covariates used in the model in Eq 1. In Fig 3(a) we can see that the absolute difference in predicted prolonged abstinence among female students are similar to those among male students, on average 10 percentage point absolute difference between being offered the intervention and not offered the intervention. Thus, there does not seem to be any heterogeneity in treatment effects with respect to sex. Similarly, there does not seem to be much heterogeneity with respect to the use of snuff (Fig 3(b)) or the number of years smoked (Fig 3(c)).

Heterogeneity does however seem to exist when taking into consideration students' nicotine dependence (Fagerström's test for nicotine dependence), as can be seen in Fig 3(d). The absolute difference in predicted probability of prolonged abstinence is on average approximately 11 percentage points for those with low dependence, but closer to 9 percentage points for those with high dependence. The ranges are quite wide for these estimates, yet there does seem to exist a trend of reduced effect of the intervention as the severity of dependence increases. Put differently, the model predicts that the NEXit intervention is more effective among university students with low dependence to nicotine than among those with high dependence. This should however not be interpreted as a causal connection, thus we do not suggest that changing the nicotine dependence for a specific student will change the effectiveness of the intervention for this individual. Instead, the findings are observational in nature: among those with higher nicotine dependence the NEXit intervention was predicted to be less effective. A similar narrative can be given when looking at the number of cigarettes smoked per week (Fig 3(e)). As the number of cigarettes smoked increases, the absolute difference of predicted prolonged abstinence between intervention and control decreases, suggesting that the intervention is less effective among students who smoke more.

### Discussion

When studying heterogeneous treatment effects of the NEXit intervention, we found that the intervention's effect on university students was reduced among those with stronger nicotine dependence and among those who smoked more cigarettes per week. This should not be interpreted as a causal connection between nicotine dependence and the effectiveness of the NEXit intervention, but rather an associational one which we may be able to exploit when deciding which intervention to allocate to which subgroup in a population.



**Fig 3. Heterogeneous treatment effects of the NEXit smoking cessation intervention.** (a-c) No heterogeneity is apparent with respect to sex, the use of snuff, nor the number of years smoked. (d-e) The NEXit intervention is less effective among those with stronger nicotine dependence and those who smoke more per week.

<https://doi.org/10.1371/journal.pone.0229637.g003>

The findings from this study questions the purpose of future developments of digital smoking cessation aids. As prevalence rates of smoking drops in a population, it may be the case that among those who still smoke a great majority have a stronger dependence to nicotine. If this is the case, then the results enclosed suggests that text messaging interventions may not be



the right tool for further reduction of prevalence rates of smoking. However, this is only a hypothetical scenario which needs to be supported by future research and data aggregation.

The NEXit intervention was more supportive of smoking cessation among students who had not yet developed a strong nicotine dependence (defined by Fagerström's test for nicotine dependence). Thus, it may be an ideal tool to make available to university students to help them quit before they develop a stronger dependence. As data in Sweden suggest that smoking is still prevalent among young individuals, a digital intervention which is easily scaled to a national level could protect from increasing prevalence of smoking in the future. However, any decision of full-scale dissemination should be based on careful consideration of other factors, including other types of interventions, costs, time and resources.

There is a need to conduct heterogeneous treatment effect analyses of other types of smoking cessation interventions in order to get a better picture of which interventions are effective for which individuals. For instance, it was found that a web-based smoking cessation program as a supplement to nicotine patch therapy was effective in supporting smoking cessation after 12 weeks (OR = 1.33 of 10-week abstinence; 95% CI = 1.13–1.57; P-value < .001) [27], however, so was a digital intervention (using email, web-pages, interactive voice response, and text messaging) without the use of nicotine replacement after 12 weeks (OR = 2.93 of 7-day abstinence, 95% CI = 1.67–5.14, P < .001) [27]. Results from both trials may be masking important heterogeneous treatment effects, which could be compared to those enclosed, in order to better understand which individuals benefit the most (and least) from different smoking cessation interventions.

## Limitations

The results of the analyses included herein should be understood under several conditions. The population under consideration are university students, thus our findings may not be generalizable to other subpopulations or the general population as a whole. Also, readers should recall that in all RCTs we are formally testing the randomization component rather than one treatment against another, as individuals may decide to adhere and engage with an intervention at their own discretion. Thus, while being randomized may sometimes be a good proxy to treatment, the estimates in this analysis should be strictly understood as comparing access to the intervention rather than use of it.

## Conclusions

Using methods common in the machine learning field, we were able to get reliable estimates of heterogeneous treatment effects of the NEXit intervention. The findings suggest that subpopulations among university students were predicted to react differently to the NEXit intervention, and this was not evident from the original analyses when the intervention and control groups were compared directly. Thus, the heterogeneous treatment effects were masked in the original analysis [11].

This study is the first of its kind with respect to text messaging interventions for smoking cessation among university students. We recommend other researchers to pursue similar studies in order to strengthen the evidence and corroborate results. It should be noted that in trials where no average treatment effect was found, the exploration of heterogeneous treatment effects may reveal if there are subgroups for which the effect is being masked by the overall treatment effect. We also need a better understanding of heterogeneous treatment effects of other available interventions, eg. face-to-face and group counselling, in order to decide which individuals should be targeted by future intervention developments.

## Author Contributions

**Conceptualization:** Marcus Bendtsen.

**Data curation:** Marcus Bendtsen.

**Formal analysis:** Marcus Bendtsen.

**Investigation:** Marcus Bendtsen.

**Methodology:** Marcus Bendtsen.

**Project administration:** Marcus Bendtsen.

**Resources:** Marcus Bendtsen.

**Software:** Marcus Bendtsen.

**Validation:** Marcus Bendtsen.

**Visualization:** Marcus Bendtsen.

**Writing – original draft:** Marcus Bendtsen.

**Writing – review & editing:** Marcus Bendtsen.

## References

1. Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Stu. *Lancet*. 2018 Nov; 392(10159):1923–94. [https://doi.org/10.1016/S0140-6736\(18\)32225-6](https://doi.org/10.1016/S0140-6736(18)32225-6) PMID: 30496105
2. Gilljam H. Quitting smoking brings quick health benefits. *Lakartidningen*. 109(11):554–7. PMID: 22530423
3. WHO global report on trends in prevalence of tobacco smoking 2015. 2015.
4. Folkhälsomyndigheten. Nationella folkhälsoenkäten “Hälsa på lika villkor?” 2018.
5. Mason M, Ola B, Zaharakis N, Zhang J. Text Messaging Interventions for Adolescent and Young Adult Substance Use: a Meta-Analysis. *Prev Sci*. 2015 Feb; 16(2):181–8. <https://doi.org/10.1007/s11121-014-0498-7> PMID: 24930386
6. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text Messaging-Based Interventions for Smoking Cessation: A Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth*. 2016/05/22. 2016; 4(2):e49. <https://doi.org/10.2196/mhealth.5436> PMID: 27207211
7. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS Text Message Interventions for Smoking Cessation: A Meta-Analysis. *J Subst Abuse Treat*. 2015 Sep; 56:1–10. <https://doi.org/10.1016/j.jsat.2015.01.011> PMID: 25720333
8. Rodgers A. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. *Tob Control*. 2005 Aug; 14(4):255–61. <https://doi.org/10.1136/tc.2005.011577> PMID: 16046689
9. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *Lancet*. 2011 Jul; 378(9785):49–55. [https://doi.org/10.1016/S0140-6736\(11\)60701-0](https://doi.org/10.1016/S0140-6736(11)60701-0) PMID: 21722952
10. Haug S, Paz Castro R, Filler A, Kowatsch T, Fleisch E, Schaub MP, et al. Efficacy of an Internet and SMS-based integrated smoking cessation and alcohol intervention for smoking cessation in young people: study protocol of a two-arm cluster randomised controlled trial. *BMC Public Health*. 2014 Jan; 14(1):1140.
11. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. Effectiveness of Short Message Service Text-Based Smoking Cessation Intervention Among University Students. *JAMA Intern Med*. 2016 Mar 1; 176(3):321. <https://doi.org/10.1001/jamainternmed.2015.8260> PMID: 26903176
12. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet*. 2000 Mar; 355(9209):1064–9. [https://doi.org/10.1016/S0140-6736\(00\)02039-0](https://doi.org/10.1016/S0140-6736(00)02039-0) PMID: 10744093

13. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in Medicine—Reporting of Subgroup Analyses in Clinical Trials. *N Engl J Med*. 2007 Nov 22; 357(21):2189–94. <https://doi.org/10.1056/NEJMSr077003> PMID: 18032770
14. Hernández A V., Boersma E, Murray GD, Habbema JDF, Steyerberg EW. Subgroup analyses in therapeutic cardiovascular clinical trials: Are most of them misleading? *Am Heart J*. 2006 Feb; 151(2):257–64. <https://doi.org/10.1016/j.ahj.2005.04.020> PMID: 16442886
15. VanderWeele TJ, Knol MJ. Interpretation of Subgroup Analyses in Randomized Trials: Heterogeneity Versus Secondary Interventions. *Ann Intern Med*. 2011 May 17; 154(10):680. <https://doi.org/10.7326/0003-4819-154-10-201105170-00008> PMID: 21576536
16. Ranganathan P, Pramesh C, Buyse M. Common pitfalls in statistical analysis: The perils of multiple testing. *Perspect Clin Res*. 2016; 7(2):106. <https://doi.org/10.4103/2229-3485.179436> PMID: 27141478
17. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. SMS-based smoking cessation intervention among university students: study protocol for a randomised controlled trial (NEXit trial). *Trials*. 2015/04/16. 2015; 16(1):140.
18. Müssener U, Bendtsen M, McCambridge J, Bendtsen P. User satisfaction with the structure and content of the NEXit intervention, a text messaging-based smoking cessation programme. *BMC Public Health*. 2016/11/24. 2016; 16(1):1179. <https://doi.org/10.1186/s12889-016-3848-5> PMID: 27876031
19. Müssener U, Linderoth C, Bendtsen M. Exploring the Experiences of Individuals Allocated to a Control Setting: Findings From a Mobile Health Smoking Cessation Trial. *JMIR Hum Factors*. 2019 Apr; 6(2): e12139. <https://doi.org/10.2196/12139> PMID: 30938682
20. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction*. 2005 Mar; 100(3):299–303. <https://doi.org/10.1111/j.1360-0443.2004.00995.x> PMID: 15733243
21. McElreath R. *Statistical Rethinking*. Taylor & Francis Group; 2016.
22. Betancourt M. A Conceptual Introduction to Hamiltonian Monte Carlo. 2017 Jan 9; <http://arxiv.org/abs/1701.02434>
23. Gelman A, Lee D, Guo J. Stan: A Probabilistic Programming Language for Bayesian Inference and Optimization. *J Educ Behav Stat*. 2015 Oct; 40(5):530–43.
24. Piironen J, Vehtari A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electron J Stat*. 2017; 11(2):5018–51.
25. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med*. 2014 Feb 10; 33(3):517–35. <https://doi.org/10.1002/sim.5941> PMID: 24002997
26. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models. *Epidemiology*. 2010 Jan; 21(1):128–38. <https://doi.org/10.1097/EDE.0b013e3181c30fb2> PMID: 20010215
27. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. *Addiction*. 2005 May; 100(5):682–8. <https://doi.org/10.1111/j.1360-0443.2005.01093.x> PMID: 15847626