The missing=smoking assumption: a fallacy in internet-based smoking cessation trials?

Blankers, M.; Smit, E.S.; van der Pol, P.; de Vries, H.; Hoving, C.; van Laar, M.

DOI
10.1093/ntr/ntv055

Publication date
2016

Document Version
Final published version

Published in
Nicotine & Tobacco Research

Nicotine & Tobacco Research, 2016, 25–33
doi:10.1093/ntr/ntv055
Original investigation
Advance Access publication March 5, 2015

Original investigation

The Missing=Smoking Assumption: A Fallacy in Internet-Based Smoking Cessation Trials?

Matthijs Blankers PhD[1,2,3], Eline Suzanne Smit PhD[4,5], Peggy van der Pol PhD[6], Hein de Vries PhD[5], Ciska Hoving PhD[5], Margriet van Laar PhD[1,6]

[1]Netherlands Expertise Centre on Tobacco Control (NET), Trimbos Institute, Utrecht, the Netherlands; [2]Department of Psychiatry, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands; [3]Department of Research, Arkin, Amsterdam, the Netherlands; [4]Amsterdam School of Communication Research/ASCoR, Department of Communication Science, University of Amsterdam, Amsterdam, the Netherlands; [5]CAPHRI School for Public Health and Primary Care, Department of Health Promotion, Maastricht University, Maastricht, the Netherlands; [6]Department of Drug Monitoring, Trimbos Institute, Utrecht, the Netherlands

Corresponding Author: Matthijs Blankers, PhD, Trimbos-instituut, PO Box 725, 3500 AS, Utrecht, the Netherlands. Telephone: +31-30-297-11-00; Fax: +31-30-297-11-11; E-mail: mblankers@trimbos.nl

Abstract

Introduction: In this study, penalized imputation (PI), a common approach to handle missing smoking status data and sometimes referred to as "missing=smoking," is compared with other missing data approaches using data from internet-based smoking cessation trials. Two hypotheses were tested: (1) PI leads to more conservative effect estimates than complete observations analysis; and (2) PI and multiple imputation (MI) lead to similar effect estimates under balanced (equal missingness proportions among the trial arms) and unbalanced missingness.

Methods: First, the outcomes of 22 trials included in a recent Cochrane review on internet-based smoking cessation interventions were reanalyzed using only the complete observations, and after applying PI. Second, in a simulation study outcomes under PI, complete observations analysis, and two types of MI were compared. For this purpose, individual patient data from one of the Cochrane review trials were used. Results of the missing data approaches were compared with reference data without missing observations, upon which balanced and unbalanced missingness scenarios were imposed.

Results: In the reanalysis of 22 trials, relative risks (*RR* = 1.15 [1.00; 1.33]) after PI were nearly identical to those under complete observations analysis (*RR* = 1.14 [0.98; 1.32]). In the simulation study, PI was the only approach that led to deviations from the reference data beyond its 95% confidence interval.

Conclusions: Analyses after PI led to pooled results equivalent to complete observations analyses. PI also led to significant deviations from the reference in the simulation studies. PI biases the reported effects of interventions, favoring the condition with the lowest proportion of missingness. Therefore, more sophisticated missing data approaches than PI should be applied.

Introduction

In this study, the validity of common approaches to handle missing smoking status data, such as penalized imputation (PI)[1]—sometimes referred to as "missing=smoking"—and multiple imputation (MI)[2] is evaluated for internet-based smoking cessation trials. Missing smoking status data are often caused by participant drop-out before follow-up measurements took place. Randomized controlled trials of internet-based interventions including those for smoking cessation typically face higher drop-out rates than face-to-face psychological intervention trials or drug trials.[3] High drop-out rates are sometimes even considered a typical characteristic of internet interventions,

especially of internet-based unguided self-help interventions.[3] A recent Cochrane review on the effectiveness of 28 internet-based smoking cessation interventions in comparison to no-intervention controls, or to different (non-)Internet interventions[4] indicated that on average nearly half of the randomized participants dropped out of the study prematurely, which led to missing smoking status data and potentially invalidated results. This potentially jeopardizes the validity of the conclusions regarding effectiveness of the interventions when not addressed in a principled manner.

Among traditional methods to address missing data, historically the most commonly used method was to remove all participants from the analysis whose information is incomplete: complete observations analysis.[5] There are some (older) publications indicating that valid estimates can be obtained under certain specific conditions using complete observations analysis[6] whereas others have argued against the use of this approach.[2,5] Nowadays, it is a convention in trial research to address drop-out according to the intention-to-treat principle.[7] This entails to analyze all allocated participants according to their original trial arm assignment, regardless of what subsequently occurred.[7,8] To apply intention-to-treat analysis, one should either estimate likely values for missing observations, or apply statistical methods that can handle longitudinal data with missing observations at collection waves (eg, generalized linear mixed modeling).[9] This manuscript will focus on the first family of solutions (imputation).

The use of imputation techniques based on simple heuristics such as last-observation-carried-forward or PI is argued against in the Consolidated Standards of Reporting Trials elaboration.[8] In last-observation-carried-forward, a missing observation in a longitudinal study is replaced by a previous observation (of the same variable). In PI, any participant for whom smoking status data are missing at follow-up is assumed to be smoking. According to the Consolidated Standards of Reporting Trials elaboration, imputation methods based on heuristics (such as last-observation-carried-forward) may introduce bias, and no allowance is made for the uncertainty of imputation (p. 17).[8]

As an alternative to imputation techniques based on heuristics such as PI or last-observation-carried-forward, model-based imputation techniques have been developed, often based on regression or expectation maximization.[10] Using those techniques, the most likely value given the available data in the dataset is calculated for each missing observation separately, and this most likely value is imputed in the dataset. After this imputation process, all analyses can be carried out as if no missingness has occurred while collecting the data.[2,11] Model-based imputation techniques typically have a number of underlying assumptions regarding the patterns of missingness: missing completely at random (missing data are a random sample of all the cases), missing at random (MAR: whether data are missing or not is dependent on observed data) or missing not at random (missingness is dependent on unobserved data). A number of papers present an accessible introduction to these concepts.[1,5,11,12] For the remainder of this manuscript, it is important to keep in mind that the missing data pattern is an assumption underlying missing data approaches, and that the model-based imputation techniques presented all assume MAR missingness patterns. These approaches will also perform well under missing completely at random missingness patterns, but may perform suboptimal under missing not at random missingness—although the sensitivity of the final results to misspecification of the missingness pattern is a matter of debate.[5,12]

Estimating and imputing the most likely value for each missing value can be performed a single time (single imputation) or multiple times (MI). An advantage of imputing each missing value multiple times is that the uncertainty of each imputation is reflected in the variance between the MIs of a given missing value. If the imputation model has high predictive validity, the estimation of each missing value will be precise, and the variance between the MIs of a single variable will be small. Otherwise, if the model is imprecise, variance between the imputations will be high(er). This variance can be taken into account in the analyses,[2,11] which marks a methodological advantage of MI over PI and other single imputation approaches.

In MI, typically around 10 estimated versions of each missing value are calculated.[5] As a consequence of applying MI, missing values are replaced with more than one estimated value, which will result in more than one completed dataset. The analysis phase of MI data comprises two steps. In the first step, each of the MI datasets is analyzed as if it were an "ordinary" complete dataset with no missing values. After the analysis has been performed on each of the MI datasets separately, the results for each dataset can then be combined in the second step, into one final result. To combine the results for each dataset, specific calculation rules ("Rubin's rules"[2]) have been formulated. The combined mean (or regression coefficient) is the mean of the means obtained from the imputed datasets; the combined standard error is based on both the standard errors of the means, and the variance between these means.[2,11] Detailed discussions and studies on addressing missing data using MI have been published over the last two decades, which can be consulted for a more technical discussion.[2,5,11–14] There are a number of MI software applications available for often used statistical programmes.[15] SPSS has a build in MI module, STATA has the "mi" command with a variety of methods, including univariate regression-based imputation, predictive mean matching, (o)logit (specifically for ordered/binary variables such as smoking status), and offers multivariate methods as well. SAS also offers a variety of methods through the proc mi command. R offers a couple of MI solutions from which two widely used packages (mice and Amelia). These last two will be evaluated in the manuscript. Multivariate Imputation by Chained Equations (MICE)[16] was selected because its use is common as Multivariate Imputation by Chained Equations is available in both R and STATA (as the "ice" module). Amelia 2[17] was selected because it provides many options especially for longitudinal datasets, and because in a previous application, good results were achieved using non-normally distributed outcome data.[11]

In smoking cessation trials, a common approach to dealing with missing smoking outcomes is PI: to classify all participants lost to follow-up as smokers.[18] PI is presented in the Russell Standard[18] as a conservative approach. This assumption usually remains untested in trials applying PI, and may be unjustified. Some previous research has concluded that PI may lead to more biased estimates in comparison to model-based missing data approaches when attrition rates are unbalanced, that is, attrition rates differ between the compared groups (eg, in control group, 20% nonresponse at follow-up, in the experimental group, 40% nonresponse at follow-up).[1,9,19,20] Yet, because of the technical nature of the publications which address this issue, many researchers working on smoking cessation intervention studies may not be reached sufficiently yet. A clear sign of the latter is the fact that PI is still the de facto standard of handling participants lost to follow-up in smoking cessation trials (see for examples the trials reviewed in a Cochrane review on internet-based smoking cessation interventions).[4]

Therefore, this article tests the assumption that PI yields conservative effect estimates, compared to two other missing data approaches often used outside of the smoking cessation research field: complete observations analysis and MI. The following two hypotheses will be tested:

1. PI leads to more conservative effect estimates in recently published internet-based trials, in comparison to complete observations analysis;
2. PI and MI lead to similar effect estimates under balanced and unbalanced missingness scenarios.

Methods

Meta-Analysis

The first hypothesis was tested using meta-analytic techniques. Trials included in the mentioned review[4] with follow-up periods of at least 6 months were reanalyzed under the missing=smoking assumption and using only data from complete cases. Relative risk (*RR*) estimates obtained with the two approaches were compared. In order to reanalyze the *RR* estimates, the original publications included in the review were retrieved. From these publications, follow-up data on the numbers of smoking study participants per trial arm, the numbers of participants who successfully quit smoking, and the numbers of participants for whom smoking statuses were missing at follow-up were collected.

The review[4] comprised 28 comparison trials with in total over 45 000 participants. Of these 28, 23 comparisons reported smoking status at least 6 months after starting the intervention. Of these 23, one trial (by Rabius and colleagues[21]) did not provide information on the number of drop-outs for each trial arm separately at times of the follow-up which was necessary for our analyses. Therefore, a total of 22 comparisons[22–40] obtained from the review[4] were included in the analysis. For all of the 22 comparisons, the *RR* on 7-day abstinence of smoking, at least 6 months post-randomization was calculated twice. First by including those participants that did not provide their smoking status at follow-up in the *RR* calculation and labeling them as smokers (PI), then by excluding those for whom smoking statuses were missing at follow-up from the *RR* calculation (complete observations analysis). A *RR* significantly larger than 1.0 indicated an advantageous effect of internet-based smoking cessation support compared to controls. All analyses were performed using meta-analysis package "metaphor"[41] for R version 3.0+.[42] Using a random effects model, the overall *RR* of the 22 comparisons combined was calculated. In addition, using a fixed effects model, the four subgroup analyses for which the Cochrane review[4] presented pooled *RR* estimates were replicated under PI and complete observations analysis. These four subgroups were:

1. Studies with an interactive, tailored internet interventions with phone contact compared to a nonactive control group (two studies, subgroup 1.1.1);
2. An interactive, tailored internet interventions without phone contact compared to a nonactive control group (three studies, subgroup 1.1.2);
3. A tailored/interactive versus not tailored/interactive internet intervention (two studies, subgroup 4.1);
4. Other comparisons between internet interventions (four studies, subgroup 5.1).

Simulation Study

To test the second hypothesis, the smoking status estimates (with 95% CI) after complete observations analysis, PI, and MI were compared. Two often used MI algorithms were applied, in order to provide an indication of the stability of the results obtained using different algorithms. In a simulation study, individual patient data from one of the trials[22] presented in the review[4] were used. In this trial, the effects of an internet-based, computer-tailored smoking cessation program on smoking outcomes were tested in a sample of adult smokers from the Netherlands (n = 1123). The internet-based, tailored smoking cessation program was found to have a significant effect on abstinence reported after a 6-week period. At the 6-month follow-up, however, no intervention effects were identified.[22]

In the first step of this simulation study, all the cases with missing follow-up data at 6 weeks or 6 months were removed from the dataset. The resulting dataset with 0% missing smoking statuses was the reference dataset for the purpose of this simulation. This first step resulted in a dataset with 122 participants per trial arm (total n = 244) and with 0% missing data on the key outcome variable of interest for this study (7-day continuous smoking abstinence) at 6 weeks and 6 months post-randomization. The advantage of creating a reference set first, and imposing missingness on this reference set in a second step is that it was known what the results of the analysis on the created datasets with missing values should ideally be: The most optimal imputation strategy would lead to results closest to those of the reference dataset.

In the second step, missingness was imposed on the main outcome variable (7-day continuous abstinence) in the reference set at baseline, 6 weeks post-randomization, and 6 months post-randomization. Four datasets with missing observations were created (scenarios A–D, see Table 1). The probability of missingness was made dependent on other data in the dataset (observed data), which is the missingness pattern earlier referred to as MAR. The missingness percentages were chosen to reflect the actual variance in missingness found in internet-based smoking cessation trials. Both balanced and unbalanced missingness scenarios were created. In the balanced scenarios, missingness rates were the same for the two trial arms,

Table 1. Four Missing Data Scenarios

Scenario	Type	Amount	Baseline (%)	6 weeks (%)	6 months (%)
A	Balanced	Low	10	17.5	25
B	Balanced	High	15	35	55
C	Unbalanced	Low	Exp: 7 | Ctr: 13	Exp: 12 | Ctr: 23	Exp: 18 | Ctr: 33
D	Unbalanced	High	Exp: 11 | Ctr: 20	Exp: 25 | Ctr: 46	Exp: 39 | Ctr: 72

Exp. = experimental condition; Ctr = control condition. The percentages under baseline, 6 weeks post-randomization and 6 months post-randomization indicate the imposed percentages of missing smoking status data; Balanced indicates that the missingness proportion was the same for the two trial arms, Unbalanced indicates that the missingness proportion varied over the trial arms.

and were low at baseline, intermediate at 6 weeks, and highest at 6 months post-randomization. In the unbalanced scenarios, missingness rates were made dependent on both the trial arm and the time of measurement. Compared with the balanced scenarios, the missingness rates in the unbalanced scenarios were multiplied by a factor 0.7 in the control arm, whereas in the experimental arm the missingness rates were multiplied by a factor 1.3. Missingness in scenario A and in scenario B was dependent on time. Missingness in scenario C and in scenario D was not only dependent on time (as in scenario A and B), but also on trial arm. In the control arm, a higher missingness rate was induced than in the experimental arm. An overview of the actual induced missingness percentages is presented in Table 1.

In a final step of this simulation, each of the four datasets for the scenarios A–D were analyzed separately after applying different missing data approaches: complete observations analysis, PI, and two forms of MI. One MI algorithm was based on predictive mean matching (Multivariate Imputation by Chained Equations[16]), the other algorithm was based on bootstrapping and expectation maximization (Amelia 2,[17]). With both MI software packages, five datasets with imputed missing values were created. All available data were used to predict and impute missing observations. The five outcomes of each MI algorithm were combined using Rubin's rules.[2] The results obtained by applying these missing data approaches were compared to the reference dataset before missingness was simulated. In the analysis step, the *RR* to be smoking was calculated as: (number of cases smoking) / (all cases with [measured or imputed] smoking status data). Confidence intervals were calculated using the "binom.confint" function in the "binom" package[43] for R, with confidence intervals constructed using the Clopper-Pearson method ("exact"). For each approach, the deviation of the *RR* from the reference dataset *RR* will be presented (with the 95% CI of the *RR* for each approach). All analyses were performed using R version 3.0+.[42]

Results

Meta-Analysis

The averaged (weighted) missingness rate across the trials was 51% at follow-up. On average, the 22 comparisons had 22% difference in missingness proportion between the two compared trial arms. So clearly, drop-out imbalance was common. In nine comparisons, missingness was higher in the experimental arm and in 12 other comparisons, missingness was higher in the control arm. In one comparison, missingness was exactly balanced. Altogether, the missingness proportion was slightly (4%) higher in the experimental arm ($\chi^2(1) = 7.86$, $P = .005$). Intervention effects under PI and complete observations analysis of the 22 trials are presented in Figure 1.

Based on the overall forest plot (random effects model) in Figure 1, PI results (*RR* = 1.15 [1.00; 1.33]) were nearly identical to results under complete observations analysis (*RR* = 1.14 [0.98; 1.32]). Based on the meta-analysis performed on these 22 comparisons, there was no indication that PI led on average to more conservative estimates than complete observations analysis. For the fixed effects subgroup analyses presented in the work by Civljak and colleagues (2013),[4] no support for the hypothesis that PI leads to more conservative estimates than complete observations analysis was found either. Subgroup analyses 1.1.1, 4.1 and 5.1 led to identical conclusions under PI and complete observations analysis. Subgroup analysis 1.1.2 even led to the conclusion that there was a significant difference between the compared intervention conditions under PI, whereas under complete observations analysis, one would have concluded that there was no significant difference between the compared intervention conditions. This means that in this situation, PI was less conservative than complete observations analysis.

In addition, in individual comparisons the differences between PI and complete observations analyses can be large. For example, in the study by Elfeddali[26] the missingness was notably larger in the control group (523/636 = 0.82) than in the experimental group (1031/1395 = 0.74). This has led to a large difference in the *RR* calculated under PI (1.18 [0.84; 1.64]) and complete observations analysis (0.80 [0.61; 1.05]). As the missingness in this study was larger in the control group, the intervention effects under PI appear larger than under complete observations analysis and these results are thus not conservative. The largest difference in missingness between the two trial arms among the 22 included comparisons was found in the study by Humfleet and colleagues.[29] In the experimental arm of this study, the proportion of missingness was 15/58 = 0.26. In the control arm, the missingness proportion was 9/82 = 0.11. This amount of drop-out imbalance is reflected in the difference between the *RR*s calculated under PI (*RR* = 1.11 [0.54; 2.27]) and complete observations (*RR* = 1.33 [0.67; 2.67]).

Simulation Study

Figure 2 presents deviations of the four missing data approaches from the reference set in scenarios A–D. What can be observed from Figure 2 is that only PI led to significant deviations from the reference value in scenario B and D, indicated by the 95% CIs which did not enclose the reference value (the black horizontal line). In the balanced missingness scenario B, this deviation may be conservative: the experimental condition (indicated by the black triangle) has a somewhat stronger increase in proportion smoking than the control condition relative to the horizontal reference line, thus reducing the difference between the proportion smoking in the control and experimental condition. As a consequence, the power of the trial to find true differential trial arm effects would be reduced. In the unbalanced missingness scenario D, the effect of applying PI appears anticonservative. A stronger increase in the proportion smoking is seen in the control condition than in the experimental condition, relative to the reference data. This is a direct consequence of the fact that the missingness proportion was higher in the control condition than in the experimental condition in the scenarios B and D. In a secondary analysis (data not presented in this manuscript) it was confirmed that the opposite was also true: if the proportion of missingness is higher in the experimental condition, effects of the intervention in the experimental arm were underestimated in comparison to the control condition when PI is applied.

Figure 3 presents the *RR* estimates 6 months post-randomization (with 95% CI) under a high proportion of balanced (scenario B, 55% missing in both arms) or unbalanced missingness (scenario D, experimental condition 39%, control condition 72% missing). For the reference dataset, no statistically significant effect for the experimental intervention is found (*RR* = 1 is within the 95% CI of the reference dataset). This is also true for all missing data approaches under the balanced scenario B, and for all missing data approaches in the unbalanced scenario D, except for PI. Under the high unbalanced missingness scenario D, applying PI leads to the erroneous conclusion that the experimental intervention has favorable effects on smoking, in comparison to the control condition. For both PI *RR*s in Figure 3, it can be observed that the CIs are small relative to all other datasets. This is a consequence of consistently imputing missing data as if all missing observations were positive smoking statuses. As MI

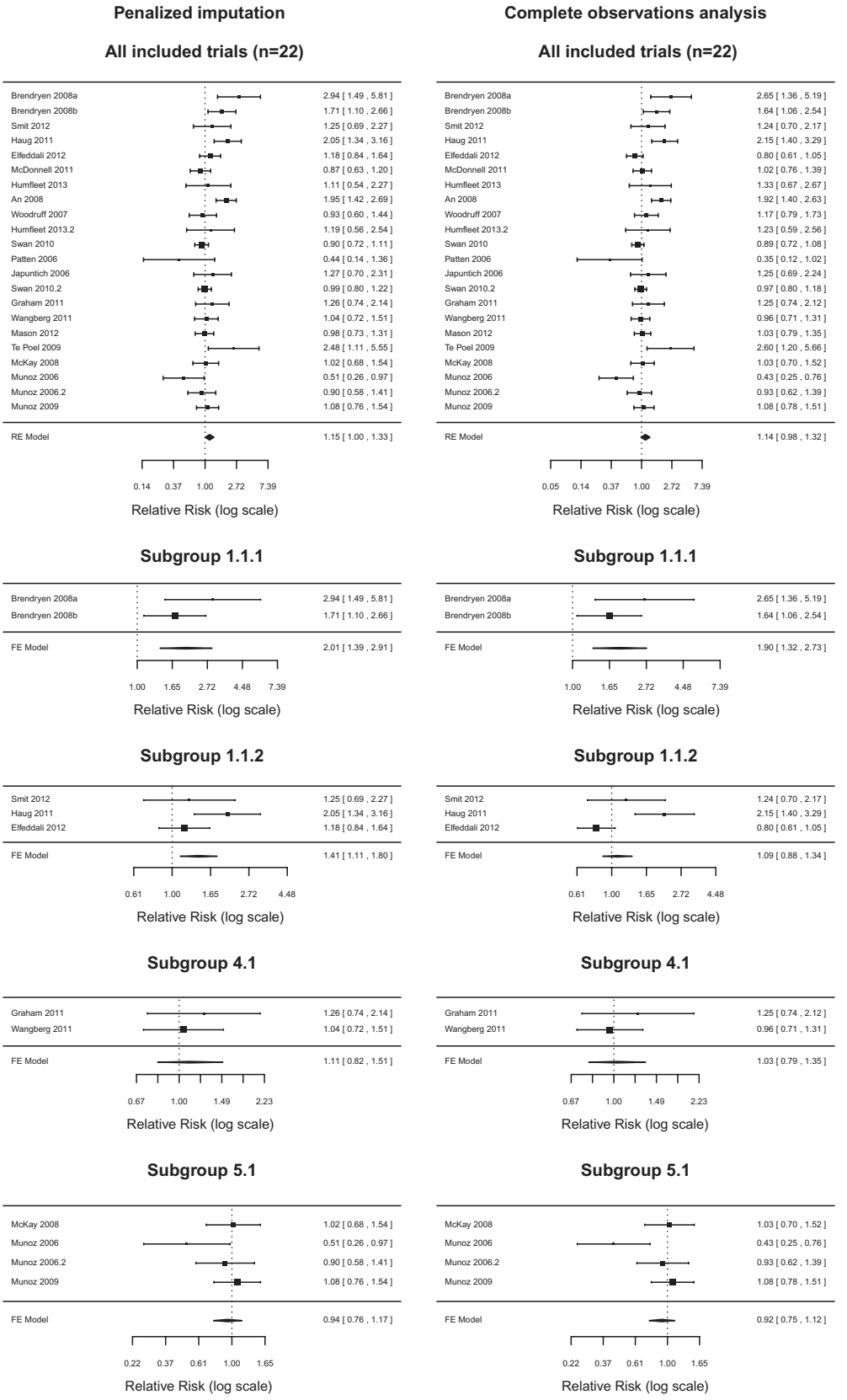

Figure 1. Outcomes of 22 trials included in the Cochrane review,[4] reanalyzed under penalized imputation and complete observations analysis. FE = fixed effects; RE = random effects. *Notes*. For Mason et al.[31], all participants were included, whereas the review[4] excluded recent ex-smokers for this trial. However including these participants did not change the overall results. The review[4] did not present the overall RE model ($n = 22$), but did present the pooled *RR* FE estimates for the four subgroup analyses presented in this figure; numbers for the FE models may differ slightly from those presented in[4] due to differences in software used.

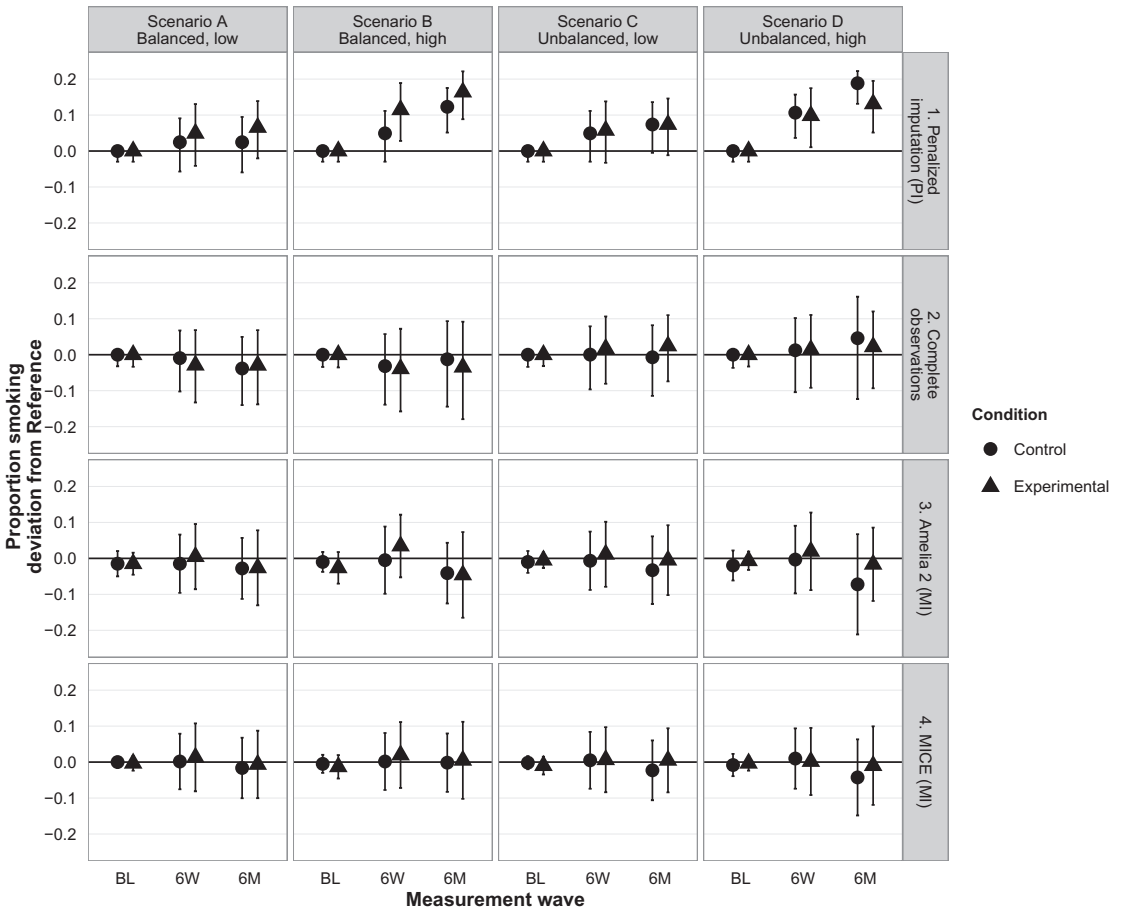

Figure 2. Deviations from reference after applying missing data approaches in scenarios A–D. BL = Baseline; MICE = Multivariate Imputation by Chained Equations; MI = multiple imputation; PI = penalized imputation; 6W = 6-week follow-up; 6M = 6-month follow-up. *Notes.* The first column of plots presents the deviations from the reference data under a low proportion of balanced missingness, scenario A. The second column of plots presents the deviations under a high proportion of balanced missingness, scenario B. The third and fourth column of plots presents deviations under low (scenario C) and high (scenario D) proportions of unbalanced missingness, respectively. The first row presents the results under PI. The second row presents the results under the four scenarios when only the complete observations are analyzed (complete observations analysis). The third and fourth rows present the results after applying MI, either by using Amelia[17], or MICE[16]. Error bars indicate 95% confidence interval around the estimated cessation proportion. Numerical data on which this figure is based are available here: Supplementary Material 1.

takes the uncertainty of the imputations into account, the CIs of the MI analyses are large compared to the reference set. Numerical data support these observations (Supplementary Material 2).

Discussion

In this study, PI ("missing=smoking") has been compared to other common approaches to missing data, to assess its validity when applied in internet-based smoking cessation effect trials. Those trials tend to face high attrition rates. It is often assumed that PI leads to conservative effect estimates in cessation trials, but apparently this is not the case.

In the meta-analysis performed on 22 studies included in a recent review,[4] PI did not lead to more conservative overall estimates than complete observations analysis. In fact, the estimates and CIs were almost identical for the two approaches. The simulation study showed that applying PI significantly affected the results of trials in case of high proportions of missingness. If the proportion missingness is unbalanced, and higher in the control condition than in the experimental condition, analysis under PI leads to an overestimation of the effectiveness of the intervention in the experimental condition. If the proportion missingness is higher in the experimental condition, the effects of the intervention are falsely attenuated. This effect is more prominent when the proportion missingness is higher.

All in all, based on the results in our study it can be concluded that PI biases the reported effects of interventions, favoring the condition with the lowest proportion of missingness. This imprecision as a consequence of applying PI in unbalanced missingness scenarios is undesirable, as it may either lead to elevated Type I error rates (falsely concluding there is a significant effect), or to elevated Type II error rates (falsely concluding that there is no significant effect).

On average, a slightly higher missingness rate was found in the experimental condition of the 22 trials. This finding is in line with a recent systematic review of a random selection of health behavior change trials.[44] In this review it has been shown that over 52 included studies, the relative missingness proportions (missingness proportion in the experimental condition / missingness proportion in the control condition) varied widely and ranged from 0.14 to 4.80. On average, a 10% higher proportion of missingness is found in the intervention conditions of health behavior change trials.[44]

In the simulation study, all tested missing data approaches led to some degree of error in comparison to the reference dataset.

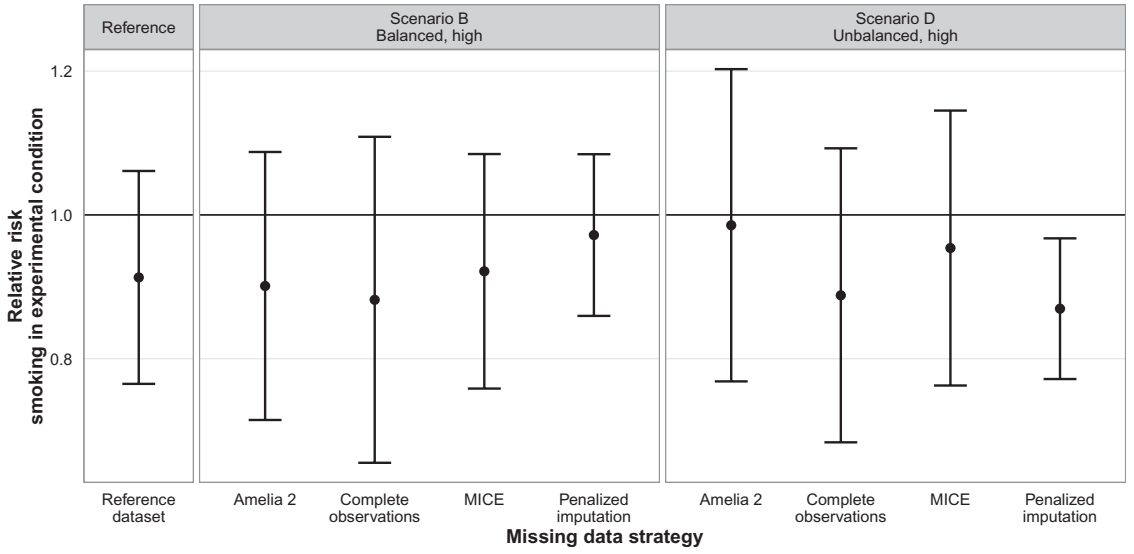

Figure 3. Relative risk estimates under scenarios B and D after applying missing data approaches. MICE = Multivariate Imputation by Chained Equations; RR = relative risk. *Notes.* Error bars indicate 95% confidence interval around the estimated *RR*. *RR* < 1 indicates favorable clinical results for the experimental condition, whereas *RR* > 1 indicates favorable clinical results for the control condition. When *RR* = 1 is within the 95% CI of the estimate, one would conclude that there is no statistically significant difference between the two conditions. Numerical data on which this figure is based are available here: Supplementary Material 2.

However, the two missing data approaches based on MI were not as systematically affected by unbalanced missingness as when PI was applied, indicating an advantage of MI over PI. Although in the simulation study it was found that the complete observations analysis did not lead to biased results to the extent that PI did, use of complete observations analysis should not be advocated. This is mainly considering the reduced power as a consequence of including fewer cases in the analysis than if some form of imputation would be used, and the notion that under MAR missingness, complete observations analysis has been shown to lead to biased results.[5,11,12] Another argument why not to use complete observations analysis is that is does not adhere to the intention-to-treat principle. These conclusions align well with what has previously been discussed in the literature with regard to the use of PI in the analysis of trial data.[1,9,19,20] The effect is even more pronounced for internet-based trials, as drop-out rates tend to be higher in internet-based trials than in face-to-face trials. These conclusions are also in line with Consolidated Standards of Reporting Trials[7] and Consolidated Standards of Reporting Trials-EHEALTH,[45] in which the use of sophisticated missing data approaches is encouraged, but not in line with the Russell standard,[18] nor with what is common practice in the smoking cessation trial literature.[4]

In the simulation study, we created five imputed datasets for the two applications of MI. There is, however, some debate regarding the number of imputations necessary to obtain reliable estimations using MI. Some authors indicate that 5–10 imputations are (often) sufficient,[2,5] while others [46,47] indicate that more imputations may be necessary under certain conditions in order to obtain estimates with a minimum of Monte Carlo variability, for example under high proportions of missingness. White and colleagues also present methods for determining whether an adequate number of datasets have been produced.[46] To assess whether the results presented in this manuscript are sensitive to the number of imputations, we have replicated our findings using a larger number of imputations ($m = 50$). As this has led to comparable results, raising the number of imputations in the current application of MI would not have made any difference for the conclusions drawn.

Strengths and Limitations

The results of this study should be considered in the light of its strengths and limitations. A strength of this study may be that the implications of missingness approaches have been demonstrated in a comprehensive and accessible way, and that the use of much statistical or technical terminology has been avoided, in order to get the key message across to an audience beyond researchers keenly interested in statistical research methodology. Regarding the meta-analysis, the fact that not all studies with at least 6 months follow-up included in the original Cochrane review were included could be regarded as a limitation. The trial that was excluded[21] did not provide the necessary details to calculate the *RR* estimates under the two missing data approaches (number of missing observations was not provided), which may have hampered some comparisons between the original review and our work. However, considering the systematic nature of the effects of PI on the *RR* estimates, including this trial would not have changed any of the conclusions. Another limitation is the fact that in the simulation study the missingness was induced by an algorithm. This had the advantage that four different missingness scenarios could be created and that a reference to compare the imputed results against was available. A disadvantage was however that the missing data approaches have not been tested on naturally occurring missing data in this study. On the other hand, MI has been extensively evaluated and can be considered a valid missing data strategy under the commonly assumed MAR missing data pattern.[5,12]

Conclusion

In order to optimize the methodological quality of future reports on internet-based smoking cessation intervention trials, researchers should refrain from assuming those missing are smoking, stop using PI in case of missing outcome data, and use more sophisticated missing data approaches such as MI instead. Ideally, in sensitivity analyses the impact of the chosen missing data strategy on the reported results should be assessed and presented.[14] For the wider smoking cessation research community, it may be considered to evaluate

current standards of handling nonresponse and missing outcome data in cessation trials, especially the use of PI, with the advent of internet-based intervention trials which tend to face high attrition rates.

Supplementary Material

Supplementary Materials 1 and 2 can be found online at http://www.ntr.oxfordjournals.org

Funding

This work was supported by the Dutch Cancer Society (UM 2007-3834), as they funded the trial from which data were used in the simulation study.

Declaration of Interests

HdV is the scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools. None of the other authors have any conflicts of interest to declare.

Acknowledgments

Time to perform this study was provided by the Trimbos Institute, the Netherlands Institute of Mental Health and Addiction. Since 2013, the Netherlands Expertise Centre for Tobacco Control (NET) is hosted by the Trimbos Institute. NET is funded by the Netherlands Ministry of Health, Welfare and Sport.

References

1. Nelson DB, Partin MR, Fu SS, Joseph AM, An LC. Why assigning ongoing tobacco use is not necessarily a conservative approach to handling missing tobacco cessation outcomes. *Nicotine Tob Res*. 2009;11(1):77–83. doi:10.1093/ntr/ntn013.
2. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: J. Wiley & Sons; 1987. doi:10.1002/9780470316696.
3. Eysenbach G. The law of attrition. *J Med Internet Res*. 2005;7(1):e11. doi:10.2196/jmir.7.1.e11.
4. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. *Cochrane Database Syst Rev*. 2013;7:CD007078. doi:10.1002/14651858.CD007078.pub4.
5. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods*. 2002;7(2):147–177. doi:10.1037/1082-989X.7.2.147.
6. Graham JW, Donaldson SI. Evaluating interventions with differential attrition: the importance of nonresponse mechanisms and use of follow-up data. *J Appl Psychol*. 1993;78(1):119–128. doi:10.1037//0021-9010.78.1.119.
7. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. doi:10.1136/bmj.c332.
8. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg*. 2012;10(1):28–55. doi:10.1016/j.ijsu.2011.10.001.
9. Hall SM, Delucchi KL, Velicer WF, et al. Statistical analysis of randomized trials in tobacco treatment: longitudinal designs with dichotomous outcome. *Nicotine Tob Res*. 2001;3(3):193–202. doi:10.1080/14622200110050411.
10. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Stat Methodol*. 1977;39(1):1–38. doi:10.1.1.133.4884.
11. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. *J Med Internet Res*. 2010;12(5):e54. doi:10.2196/jmir.1448.
12. Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol*. 2009;60:549–576. doi:10.1146/annurev.psych.58.110405.085530.
13. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996;91(434):473–489. doi:10.1080/01621459.1996.10476908.
14. Liublinska V, Rubin DB. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial [published online ahead of print May 20, 2014]. *Stat Med*. doi:10.1002/sim.6197.
15. Yucel RM. State of the Multiple Imputation Software. *J Stat Software*. 2011;45(1):1–7. www.jstatsoft.org/v45/i01/paper. Accessed December 24, 2014.
16. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Software*. 2011;45(3):1–67. www.jstatsoft.org/v45/i03. Accessed October 8, 2014.
17. Honaker J, King G, Blackwell M. Amelia II: a program for missing data. *J Stat Software*. 2011;45(7):1–47. www.jstatsoft.org/v45/i07. Accessed October 8, 2014.
18. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction*. 2005;100(3):299–303. doi:10.1111/j.1360-0443.2004.00995.x.
19. Hedeker D, Mermelstein RJ, Demirtas H. Analysis of binary outcomes with missing data: missing=smoking, last observation carried forward, and a little multiple imputation. *Addiction*. 2007;102(10):1564–1573. doi:10.1111/j.1360-0443.2007.01946.x.
20. Barnes SA, Larsen MD, Schroeder D, Hanson A, Decker PA. Missing data assumptions and methods in a smoking cessation study. *Addiction*. 2010;105(3):431–437. doi:10.1111/j.1360-0443.2009.02809.x.
21. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. *J Med Internet Res*. 2008;10(5):e45. doi:10.2196/jmir.1008.
22. Smit ES, de Vries H, Hoving C. Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers. *J Med Internet Res*. 2012;14(3):e82. doi:10.2196/jmir.1812.
23. An LC, Klatt C, Perry CL, et al. The RealU online cessation intervention for college smokers: a randomized controlled trial. *Prev Med*. 2008;47(2):194–199. doi:10.1016/j.ypmed.2008.04.011.
24. Brendryen H, Drozd F, Kraft P. A digital smoking cessation program delivered through internet and cell phone without nicotine replacement (happy ending): randomized controlled trial. *J Med Internet Res*. 2008;10(5):e51. doi:10.2196/jmir.1005.
25. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. *Addiction*. 2008;103(3):478–484. doi:10.1111/j.1360-0443.2007.02119.x.
26. Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. *J Med Internet Res*. 2012;14(4):e109. doi:10.2196/jmir.2057.
27. Graham AL, Cobb NK, Papandonatos GD, et al. A randomized trial of Internet and telephone treatment for smoking cessation. [Erratum appears in Arch Intern Med 2011;171(5):395]. *Arch Intern Med*. 2011;171(1):46–53. doi:10.1001/archinternmed.2010.451.
28. Haug S, Meyer C, John U. Efficacy of an internet program for smoking cessation during and after inpatient rehabilitation treatment: a quasi-randomized controlled trial. *Addict Behav*. 2011;36(12):1369–1372. doi:10.1016/j.addbeh.2011.08.004.
29. Humfleet GL, Hall SM, Delucchi KL, Dilley JW. A randomized clinical trial of smoking cessation treatments provided in HIV clinical care settings. *Nicotine Tob Res*. 2013;15(8):1436–1445. doi:10.1093/ntr/ntt005.
30. Japuntich SJ, Zehner ME, Smith SS, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. *Nicotine Tob Res*. 2006;8(suppl 1):S59–S67. doi:10.1080/14622200601047900.
31. Mason D, Gilbert H, Sutton S. Effectiveness of web-based tailored smoking cessation advice reports (iQuit): a randomized trial. *Addiction*. 2012;107(12):2183–2190. doi:10.1111/j.1360-0443.2012.03972.x.

32. McDonnell DD, Kazinets G, Lee HJ, Moskowitz JM. An internet-based smoking cessation program for Korean Americans: results from a randomized controlled trial. *Nicotine Tob Res*. 2011;13(5):336–343. doi:10.1093/ntr/ntq260.
33. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. *J Med Internet Res*. 2008;10(5):e40. doi:10.2196/jmir.993.
34. Muñoz RF, Lenert LL, Delucchi K, et al. Toward evidence-based Internet interventions: a Spanish/English Web site for international smoking cessation trials. *Nicotine Tob Res*. 2006;8(1):77–87. doi:10.1080/14622200500431940.
35. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Perez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. *Nicotine Tob Res*. 2009;11(9):1025–1034. doi:10.1093/ntr/ntp090.
36. Patten CA, Croghan IT, Meis TM, et al. Randomized clinical trial of an Internet-based versus brief office intervention for adolescent smoking cessation. *Patient Educ Couns*. 2006;64(1–3):249–258. doi:10.1016/j.pec.2006.03.001.
37. Swan GE, McClure JB, Jack LM, et al. Behavioral counseling and varenicline treatment for smoking cessation. *Am J Prev Med*. 2010;38(5):482–490. doi:10.1016/j.amepre.2010.01.024.
38. Te Poel F, Bolman C, Reubsaet A, De Vries H. Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months. *Health Educ Res*. 2009;24(6):930–940. doi:10.1093/her/cyp036.
39. Wangberg SC, Nilsen O, Antypas K, Gram IT. Effect of tailoring in an internet-based intervention for smoking cessation: randomized controlled trial. *J Med Internet Res*. 2011;13(4):e121. doi:10.2196/jmir.1605.
40. Woodruff SI, Conway TL, Edwards CC, Elliott SP, Crittenden J. Evaluation of an Internet virtual world chat room for adolescent smoking cessation. *Addict Behav*. 2007;32(9):1769–1786. doi:10.1016/j.addbeh.2006.12.008.
41. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Software*. 2010;36(3):1–48. www.jstatsoft.org/v36/i03. Accessed October 8, 2014.
42. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014. www.r-project.org/. Accessed October 8, 2014.
43. Dorai-Raj S. binom: Binomial Confidence Intervals For Several Parameterizations. R package version 1.1-1; 2014. http://CRAN.R-project.org/package=binom. Accessed December 24, 2014.
44. Crutzen R, Viechtbauer W, Spigt M, Kotz D. Differential attrition in health behaviour change trials: a systematic review and meta-analysis. *Psychol Health*. 2015 Jan;30(1):122–134. doi:10.1080/08870446.2014.953526.
45. Eysenbach G; CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res*. 2011;13(4):e126. doi:10.2196/jmir.1923.
46. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30(4):377–399. doi:10.1002/sim.4067.
47. Bodner TE. What improves with increased missing data imputations? *Struct Equ Modeling*. 2008;15(4):651–675.