Original Investigation

Do point prevalence and prolonged abstinence measures produce similar results in smoking cessation studies? A systematic review

John R. Hughes, M.D.,¹ Matthew J. Carpenter, Ph.D.,^{2,3} & Shelly Naud, Ph.D.⁴

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

² Hollings Cancer Center, Medical University of South Carolina, Charleston, SC

³ Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC

⁴ Department of Medical Biostatistics, University of Vermont, Burlington, VT

Corresponding Author: John R. Hughes, M.D., Department of Psychiatry, University of Vermont, University Health Center, Mailstop 482, 1 South Prospect Street, Burlington, VT 05401, USA. Telephone: 802-656-9610; Fax: 802-847-1446; E-mail: john.hughes@uvm.edu

Received January 26, 2010; accepted April 22, 2010

Abstract

Introduction: Many smoking cessation trials report either prolonged abstinence (PA) rates (i.e., not smoking since a quit date, with or without a grace period) or point prevalence (PP) abstinence rates (i.e., no smoking one or more days prior to the follow-up), but how these two relate is unclear.

Methods: We located 28 pharmacotherapy trials that provided 76 within-study comparisons of PA versus PP. The first two authors independently coded all trials.

Results: The two measures were highly correlated ($r = .88$) and PA averaged 0.74 that of PP. Equations for converting PP to PA and vice versa produced estimations that, in 90% of cases, were within 4%–5% of actual PP or PA values. The odds ratio and the relative risk for active versus control were identical when PA and PP were used; however, the difference in proportion abstinent for active versus control was somewhat less when PA was used than when PP was used (8% vs. 10%).

Discussion: We conclude that PA and PP are closely related and can be interconverted with moderate accuracy. They also produce similar effect sizes when odds ratio and relative risk are used as effect sizes. When absolute difference in percent abstinent is used as an effect size, PA produces a smaller effect size than PP. We believe trials should continue to report both PA and PP outcomes to enhance comparisons across studies.

Introduction

The two most common outcome measures in clinical trials of smoking cessation are prolonged abstinence (PA) and point prevalence (PP) abstinence (Hughes et al., 2003). Both PA and

PP are typically tied to a follow-up that occurs a certain number of weeks after a designated quit date but can be tied to end of treatment or time prior to an assessment. PA—aka-sustained or continuous abstinence—is typically defined as not smoking for a period of several months after a quit attempt. Sometimes, this is for the entire period since the quit date; other times, it begins after an initial “grace” period. Point prevalence abstinence is typically defined as not smoking on the day of follow-up or for a few days before a follow-up (technically the later is period prevalence). Sometimes, PA and PP require “not even a puff” during the interval; other times, they allow small slips [e.g., no more than 5 cigarettes (cigs) smoked in the interval] or use a no-relapse definition (usually defined as never smoking for 7 consecutive days).

Several prior reviews and guidelines have discussed the pros and cons of PA versus PP (Hughes et al., 2003; Ockene et al., 2000; Ossip-Klein et al., 1986; Shipley, Rosen, & Williams, 1982; Velicer, Prochaska, Rossi, & Snow, 1992; West, Hajek, Stead, & Stapleton, 2005). The major benefits of PA are that it (a) is more stable, (b) is a better proxy for lifelong abstinence, (c) is a better proxy for health benefit, and (d) has a closer temporal relationship to intervention than PP. The major benefits of PP are that it (a) has less memory bias, (b) has less variability due to missing data, and (c) is able to detect delayed quitting. Since current methods to biochemically verify reports of abstinence can detect smoking only over the last few days or weeks, some have stated only PP can be verified; others believe that repeated biochemical verifications (e.g., at 1, 3, 6, and 12 months) can verify PA outcomes because smoking only between these measures is a very unusual outcome (Hughes et al., 2003).

Many clinical trials have reported only PA or PP outcomes (Hughes et al., 2003). This has caused problems in comparing results across trials or in collating results among trials for meta-analyses. Although one would expect PA and PP to be correlated, the empirical relationship of PA and PP is relatively

doi: 10.1093/ntr/ntq078

Advance Access published on May 26, 2010

© The Author 2010. Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

unstudied. One analysis of 41 estimates of PA and PP across four studies reported that PP and PA were highly correlated ($r = .82-.99$; Velicer & Prochaska, 2004). A summary (Hughes et al., 2003) of two prior quantitative reviews of nicotine patch treatments (Fiore, Smith, Jorenby, & Baker, 1994; Richmond, 1997) and one of worksite treatments (Fisher, Glasgow, & Terborg, 1990) noted that, in across-study comparisons, all three reviews found effect sizes in studies that used PP were less than in studies that used PA. The current study attempted to replicate these results using a larger more comprehensive dataset and, importantly, using within-study, rather than across-study, comparisons of PP versus PA. Specifically, the analysis determined (a) the relationship between PP and PA outcomes, (b) the ability to predict PP from PA and vice versa, and (c) whether PA and PP produce similar effect sizes for treatment.

Methods

Search for studies

To locate publications that reported both PA and PP outcomes, we searched the Cochrane, EMBASE, PsycInfo, and PubMed databases. The first inclusion criterion was that the study was a randomized control trial (RCT) testing a validated pharmacotherapy as defined by the Cochrane (www.cochrane.org), United States Public Health Service (USPHS) (Fiore et al., 2008), United Kingdom (West, McNeill, & Raw, 2000), or Society for Research on Nicotine and Tobacco (www.treatobacco.net) reviews. We limited our search to RCTs of proven treatments so that we could compare effect sizes of treatment when both PA and PP were used. We limited our analysis to trials of medications because their assessment methods appear to be more homogeneous across trials than those for psychosocial treatments. Validated pharmacotherapies were bupropion, clonidine, nicotine replacement therapies (NRTs), nortriptyline, and varenicline. Second, the study had to report both PA and PP results using the same sample size. Third, it had to be a study only of adults, that is, not a study of adolescent smokers because their smoking may be less stable and require different measures (Mermelstein et al., 2002). Fourth, the study had to be of smokers who were actively trying to initiate abstinence, for example, not a study of inducing quit attempts in those not ready to quit, nor a study of reduction only, nor a study of relapse prevention in those already abstinent. Fifth, the sample could not be of a special population, for example, not a study of pregnant smokers. Sixth, the study had to report PP and PA from follow-ups at least 6 months from the quit date because PA and PP abstinence rates usually do not stabilize till then (Hughes, Keely, & Naud, 2004). Seventh, since our analyses required detailed data, it could not be an abstract or brief report. Eighth, we excluded studies whose PP had a duration of >7 days to prevent blurring of the PA/PP distinction, that is, an exceptionally long period of PP could be redefined as PA. Non-English studies were not excluded but none met our inclusion criteria.

When multiple medication conditions (e.g., multiple doses) were tested against a placebo or control, or when medication was tested under different psychosocial conditions within a single study, we calculated an effect size for each active versus placebo/control comparison; thus, the number of comparisons is greater than the number of studies because a single control group could be used in multiple comparisons. When multiple

outcomes were reported in a study (e.g., with and without biochemical verification or 6 and 12 month follow-ups), we used the longest and most stringent outcome.

The first author identified studies that appeared to be relevant based on an initial reading of results and the second author examined these studies to verify appropriateness for inclusion.

Data extraction

The first and second authors independently coded all studies, compared ratings, and came to an agreement. On several occasions, we contacted authors to clarify outcomes. Although we did not record the exact incidence of disagreement, it was not small because many studies were unclear on methodological issues. Several studies were excluded because their definitions of PA or PP were missing or unclear. Many included studies were unclear whether biochemical validation applied to PP only or both PA and PP. These studies were excluded.

For each study/comparison, we recorded nine study characteristics: (a) publishing year, (b) sample size, (c) whether the follow-up occurred at 6 or 12+ months, (d) definitions of relapse (any smoking vs. 7 consecutive days), (e) number of weeks in the grace period before PA period began, (f) number of biochemical verifications on which PA was based, (g) type of control group (placebo or no drug), (h) length of treatment, and (i) varenicline versus other treatments. We used all participants randomized in the denominators whenever possible; however, this was often unclear. We recorded number abstinent (i.e., numerator) when possible based on actual numbers in tables or the text but often these had to be calculated from recorded percent abstinent. On rare occasions, these were obtained from relapse/survival curves or bar graphs using a Digimatic software that estimates numerical data from such graphs. We did not measure study quality with a formal scale because there is no consensus on whether this influences meta-analytic outcomes (Balk et al., 2002).

Data synthesis

The study had four aims. First, we examined the relationship between PA and PP abstinence rates by determining mean values for each, the correlation between the two, the difference between the two, and the ratio of the two. Second, we examined the ability to estimate one from the other using a metaregression analysis. Third, we determined "effect sizes" when PA versus PP outcomes were used. Effect sizes are different measures of the therapeutic effect of a medication, that is, the difference in outcomes between active and control conditions. For this aim, we used three effect sizes. The first two are the odds ratio (OR) and the relative risk (RR). The third is what we term "the difference in percent abstinent between the active and control conditions" (DIFF); for example, if the active quit rate was 30% and the control quit rate was 20%, the DIFF would be 10%; this has also been termed risk difference or absolute risk reduction (Fleiss, 1994; Shadish & Haddock, 1994). The pros and cons of using these measures have previously been reviewed (Hughes & Callas, 2007). The fourth aim was to determine whether study characteristics listed above moderated the relationship of PP to PA.

To examine the ability to estimate PA from PP and vice versa, we used Hierarchical Linear Modeling (HLM) software to conduct weighted, two level metaregression with study ($n = 28$)

as Level 2 factor (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004; Thompson & Higgins, 2002). HLM takes into account that some of the studies provided more than one comparison. To normalize outcomes, percent abstinences were transformed into logits ($p/1 - p$). Follow-up analyses explored the nine study characteristics listed above as possible causes of the heterogeneity. To illustrate the accuracy of our estimations, we report the 90% range of the residuals (residuals are the difference between the predicted and observed values).

Results

Description of included studies

We included 28 RCTs that provided 76 study conditions (42 active conditions and 34 control conditions) that recorded both PA and PP (see Appendix). Sixteen studies (57%) were of bupropion, 5 (18%) of nicotine gum, 5 (18%) of varenicline, 2 (7%) of nicotine patch, 2 of nortriptyline (7%), and 1 (4%) of nicotine inhaler (the total adds to more than 100% because four studies examined more than one medication). There were no studies of clonidine, combination medications, or nicotine nasal spray that met our inclusion criteria.

The longest follow-up was 6 months in 7 studies (25%), 12 months in 19 studies (68%), and >12 months in 2 studies (7%). The PA definition allowed some smoking (i.e., slips) in 3 (11%) studies and a grace period in 19 (68%) of studies. The mean sample size was 180 (124) for active conditions and 150 (85) for control conditions. The control was placebo in 67 (88%) comparisons and no medication in 9 (12%).

Relationship of PA and PP

The weighted percent abstinent for both the active and the control groups from the meta-analysis was consistent with those reported by USPHS, Cochrane, and other meta-analyses (Hughes, 2009; Table 1). By definition, the PP must either equal to or be greater than PA. In all but one comparison, PP was greater than PA (as illustrated by values above the line of unity in Figure 1). Across all 76 conditions, PP and PA outcomes were highly correlated (unweighted $r = .88$, $p < .0001$). The mean weighted difference between PP and PA across all 76 conditions was +5.7% (95% CI = 0%–12%). This difference appeared to increase as the value of PP increased. In contrast, the ratio of PA to PP did not change with increased PP; thus, we believe that the ratio is a preferable measure of the relationship of PP to PA. The mean weighted ratio of PA to PP was 0.74 (0.70–0.79).

Estimating of PA from PP and vice versa

Initial results from the metaregression indicated that the data were not homogenous. Exploratory work examining the study characteristics listed above found that including (a) whether the control group was a no-drug or placebo condition and (b) total sample size in the analyses, each reduced heterogeneity. These analyses indicated that studies with no-drug control conditions produced estimates of PP versus PA that were more discrepant than studies with placebo controls. Also, studies with smaller sample sizes had slightly lower PP rates than studies with larger sample sizes. The reverse was found for PA rates: smaller studies had higher PA rates. After adding these two moderators, the equations for estimating PP from PA and vice versa produced

Table 1. Weighted means (95% CI) for prolonged abstinence and point prevalence

Outcome	Adjusted ^a	
	Prolonged abstinence	Point prevalence abstinence
Percent abstinent		
All ($n = 76$)	18 (15, 20)	24 (21, 27)
Active ($n = 42$)	21 (18, 23)	28 (25, 31)
Control ($n = 34$)	12 (10, 14)	17 (15, 20)
Effect size		
Relative risk ($n = 42$) ^b	1.7 (1.5, 2.0)	1.6 (1.5, 1.8)
Odds ratio ($n = 42$) ^b	1.9 (1.7, 2.2)	1.9 (1.6, 2.2)
Difference ($n = 42$) ^c	8 (6, 10)	10 (8, 13)

Note. ^aAdjusted for nonindependence due to the 28 studies contributing several comparisons.

^bReferent = control.

^cDifference in percent abstinent in active minus control condition.

accurate estimations; that is, we would expect that 90% of estimations would fall within 4%–5% of observed values (Table 2).

Effect sizes with PA and PP

The heterogeneity test was significant for the OR and RR analyses but not the DIFF analyses. The meta-analytic mean OR and the meta-analytic RR were very similar when PA and PP were used (Table 1). In contrast, the DIFF (difference in percent abstinent between active and controls) when PA was used was about 80% of that when PP measures were used.

Figure 1. Correlation between point prevalence (PP) and prolonged abstinence (PA) across all 76 comparisons. The solid line represents values expected if PP = PA. The dotted line is the regression line of best fit for PP versus PA.

Table 2. Equations for estimation of point prevalence abstinence (PP) from prolonged abstinence (PA) or vice versa from metaregression

Equations ^a	Range of 90% of residuals ^b
logit PP = 0.00048n + 0.26 control + 0.82 logit PA	–5% to +4%
logit PA = –0.23 – 0.00044n – 0.26 control + 1.04 logit PP	–4% to +4%

Note. ^a n = sample size, control values are 0 = placebo control, 1 = nonplacebo control (e.g., no medication).

^bResiduals are the differences in predicted versus observed outcomes.

Discussion

Relationship of PA versus PP

PA and PP were highly correlated as found in prior analyses (Velicer & Prochaska, 2004). That PA is less than PP is a logical necessity (Hughes et al., 2003); however, the magnitude of this difference has not been well described (Velicer & Prochaska). We found that the relationship of PP versus PA is best thought of as a ratio in which PA is 0.74 that of PP.

A second conclusion is that it appears that one can accurately estimate PP versus PA and vice versa; however, our accuracy estimates and our equations are derived from the same sample; thus, our accuracy is likely to be overestimated. A test of our equations in a different sample, for example, among studies of psychosocial treatments, is needed to assess their true accuracy and external validity.

Estimates of the magnitude of therapeutic effects with PA versus PP

A third conclusion is that PA and PP produce very similar estimates of the magnitude of the efficacy of a treatment when OR or RR is used as effect sizes. In contrast, the three prior reviews (Fiore et al., 1994; Fisher et al., 1990; Richmond, 1997) found that PP produced smaller effect sizes than PA when OR effect sizes were used (Hughes et al., 2003). One possible reason for the discrepancy in our results and those of prior studies is that prior studies examined a smaller set of studies examining only one treatment. However, a more likely reason is that we compared PA and PP within the same study, whereas prior reviews compared across studies that used PA versus studies that used PP.

When the difference in percent abstinence between active and control groups (DIFF) was used as an effect size, this effect size was somewhat smaller when based on PA than when based on PP. This might appear to be contrary to our results using OR and RR effect sizes; however, in actuality, this outcome is expected given the mathematical relationship of DIFF outcomes OR and RR outcomes (Hughes & Callas, 2007). For example, assume in one study, the PA is 5% for the control condition and 10% for the active condition and assume the PP is 15% for the control condition and 30% for the active condition. Then, the RR using PA and PP are identical (2.0); however, the DIFF is 5% using PA but is 15% using PP.

Moderators of the relationship of PA versus PP

Given that PA almost always decreases over time and PP usually increases over time (Hughes et al., 2003), we anticipated that the relationship of PA and PP would change with different follow-up durations; however, we did not find that whether follow-ups were at 6 months versus longer influenced our results. This may be because relapse and recycling of new quit attempts are rare after 6 months (Hughes, Peters, & Naud, 2008). We believe that the relationship of PA and PP for follow-ups of less than 6 months may differ from those reported herein, but this hypothesis requires testing.

Another likely influence on the relationship of PA to PP is how soon after a quit date successful abstinence begins (Hughes et al., 2003). For example, most PA measures require abstinence to begin within the first month after a quit date. Almost no PP measures do so. Thus, if long-term abstinence is due to a “late quit,” this would not be a PA-defined success but would be a PP-defined success. Although late quits have thought to be unusual, recent experience with nicotine blocking agents such as varenicline suggest that they can induce late quit successes (Fagerstrom & Hughes, 2008). However, in our analyses, we did not find that the relationship of PP to PA differed between varenicline and other medication conditions. One possible explanation for this is that other study medications (in our case this was NRT and bupropion) also induce later quits or prevent a lapse from becoming relapse. For example, clinical studies have found that NRT (Shiffman et al., 2006) and bupropion (West, Baker, Cappelleri, & Bushmakina, 2008) also are nicotine blockers and thus can also result in delayed quitting.

Assets and limitations

The major assets of our analyses include our use of (a) within-study rather than between-study comparisons of PA versus PP, (b) a larger more comprehensive sample of studies, (c) meta-analytic techniques that allowed derivation of equations for estimating PA from PP and vice versa, and (d) direct tests whether PA and PP produce similar effect sizes for treatment outcomes.

One important liability of our analyses is that we only examined studies testing a pharmacological treatment. It is reasonable to hypothesize that the therapeutic effects of psychosocial treatments take longer to take effect than do pharmacological treatments. If this is the case (and we cannot find empirical verification of this assumption), then psychosocial treatments might especially produce late quitters, and thus, the relationship of PA to PP might differ for psychosocial versus medication trials.

A second liability is that many of the articles were not clear on the exact sample sizes used for PA and PP measures nor on whether biochemical verification was required for both PA and PP or for only one of the two. If studies consistently required biochemical validation for PA more than PP, then this could account for some of the reason why PA is usually less than PP.

Conclusions

Although our results suggest that results of studies are similar when PA and PP are used and that one can convert PA to PP

and vice versa, we believe replications of our results using other settings and treatments are needed. As a result, we believe studies should continue to report both PA and PP outcomes with clear descriptions of the numerators and denominators and the role of biochemical verification in their calculation.

Funding

This analysis was funded by Grant DA025089 (JRH), Senior Scientist Award DA000490 (JRH), and Mentored Patient-Oriented Research Career Development Award DA020482 (MJC) from the U.S. National Institute on Drug Abuse.

Declaration of Interests

JRH is currently employed by the University of Vermont and Fletcher Allen Health Care. Since 1 April 2007, he has received research grants from the National Institute on Health and Pfizer. Pfizer develops and sells smoking cessation medications. During this time, he has accepted honoraria or consulting fees from several nonprofit and for-profit organizations and companies that develop, sell, or promote smoking cessation products or services or educate/advocate about smoking cessation: Abbot Pharmaceuticals; Acrux; Aradigm; American Academy of Addiction Psychiatry; American Psychiatric Association; Begbies Traynor; Cambridge Hospital, Cline, Davis, and Mann; Constella Group; Consultants in Behavior Change; Dean Foundation, DLA Piper, EPI-Q, European Respiratory Society, Evotec; Exchange Limited; Fagerstrom Consulting; Free and Clear; Glaxo-Smith Kline; Golin Harris; Healthwise; Insyght; Informed, Invivodata; Johns Hopkins University; J L Reckner; Maine Medical Center; McNeil Pharmaceuticals; Novartis Pharmaceuticals; Oglivy Health PR, Ottawa Heart Institute, Pfizer Pharmaceuticals; Pinney Associates; Propagate Pharmaceuticals, Reuters; Scientia, Selecta; Temple University of Health Sciences; University of Arkansas; University of California-San Francisco; University of Cantabria; University of Kentucky, U.S. National Institutes on Health; Wolters Publishing; and Xenova. MJC and SN have no disclosures.

Appendix Studies used in the meta-analysis

Ahluwalia, J. S., Harris, K. J., Catley, D., Okuyemi, K. S., & Mayo, M. S. (2002). Sustained-release bupropion for smoking cessation in African Americans. A randomized controlled trial. *Journal of the American Medical Association*, 288, 468–474.

Aubin, H. J., Lebarry, F., Berlin, I., Bidaut-Mazel, C., Chemali-Hudry, J., & Lagrue, G. (2004). Efficacy of bupropion and predictors of successful outcome in a sample of French smokers: A randomized placebo-controlled trial. *Addiction*, 99, 1206–1218.

Blondal, T. (1989). Controlled trial of nicotine polacrilex gum with supportive measures. *Archives of Internal Medicine*, 149, 1818–1821.

Brown, R., Niaura, R., Lloyd-Richardson, E., Strong, D., Kahler, C., Abrantes, A., et al. (2007). Bupropion and cognitive-behavioral treatment for depression in smoking cessation. *Nicotine & Tobacco Research*, 9, 721–730.

Fossati, R., Apolone, G., Negri, E., Compagnoni, A., La Vecchia, C., Mangano, S., et al. (2007). A double-blind, placebo-controlled, randomized trial of bupropion for smoking cessation in primary care. *Archives of Internal Medicine*, 167, 1791–1797.

Gonzales, D., Nides, M., Ferry, L., Kustra, R., Jamerson, B., Segall, N., et al. (2001). Bupropion SR as an aid to smoking cessation in smokers treated previously with bupropion: A randomized placebo-controlled study. *Clinical Pharmacology and Therapeutics*, 69, 438–444.

Gonzales, D., Rennard, S. I., Nides, M., Oncken, C., Azoulay, S., Billing, C. B., et al. (2006). Varenicline, a $\alpha_4\beta_2$ nicotinic acetylcholine receptor partial agonist, vs sustained-release bupropion and placebo for smoking cessation. *Journal of the American Medical Association*, 296, 47–55.

Gourlay, S. G., Forbes, A., Marriner, T., Pethica, D., & McNeil, J. J. (1995). Double blind trial of repeated treatment with transdermal nicotine for relapsed smokers. *British Medical Journal*, 311, 363–366.

Hall, S. M., Humfleet, G. L., Reus, V. I., Munoz, R. F., Hartz, D. T., & Griffin, R. M. (2002). Psychological intervention and antidepressant treatment in smoking cessation. *Archives of General Psychiatry*, 59, 930–936.

Hjalmarson, A., Nilsson, F., Sjöström, L., & Wiklund, O. (1997). The nicotine inhaler in smoking cessation. *Archives of Internal Medicine*, 157, 1721–1728.

Hughes, J. R., Gust, S. W., Keenan, R. M., Fenwick, J. W., & Healy, M. L. (1989). Nicotine vs placebo gum in general practice. *Journal of the American Medical Association*, 261, 1300–1305.

Jarvis, M. J., Raw, M., Russell, M. A. H., & Feyerabend, C. (1982). Randomized controlled trial of nicotine chewing gum. *British Medical Journal*, 285, 537–540.

Jorenby, D., Hays, J., Rigotti, N., Azoulay, S., Watsky, E., Williams, K., et al. (2006). Efficacy of varenicline, an $\alpha_4\beta_2$ nicotinic acetylcholine receptor partial agonist, vs placebo or sustained-release bupropion for smoking cessation. *Journal of the American Medical Association*, 296, 56–63.

Jorenby, D. E., Leischow, S. J., Nides, M. A., Rennard, S. I., Johnston, J. A., Hughes, A. R., et al. (1999). A controlled trial of sustained-release bupropion, a nicotine patch, or both for smoking cessation. *New England Journal of Medicine*, 340, 685–691.

McCarthy, D., Piasecki, T., Lawrence, D., Jorenby, D., Shiffman, S., Fiore, M., et al. (2008). A randomized controlled clinical trial of bupropion SR and individual smoking cessation counseling. *Nicotine and Tobacco Research*, 10, 717–729.

Molyneux, A., Lewis, S., Leivers, U., Anderton, A., Antoniak, M., Brackenridge, A., et al. (2003). Clinical trial comparing nicotine replacement therapy (NRT) plus brief counselling, brief counselling alone, and minimal intervention on smoking cessation in hospital inpatients. *Thorax*, 58, 484–488.

Nakamura, M., Oshima, A., Fujimoto, Y., Maruyama, N., Ishibashi, T., & Reeves, K. (2007). Efficacy and tolerability of

varenicline, an $\alpha_4\beta_2$ nicotinic acetylcholine receptor partial agonist, in a 12-week, randomized, placebo-controlled, dose-response study with 40-week follow-up for smoking cessation in Japanese smokers. *Clinical Therapeutics*, 29, 1040–1056.

Oncken, C., Gonzales, D., Nides, M., Rennard, S., Watsky, E., Billing, C., et al. (2006). Efficacy and safety of the novel selective nicotinic acetylcholine receptor partial agonist, varenicline, for smoking cessation. *Archives of Internal Medicine*, 166, 1571–1577.

Pirie, P. L., McBride, C. M., Hellerstedt, W., Jeffery, R. W., Hatsukami, D. K., Allen, S., et al. (1992). Smoking cessation in women concerned about weight. *American Journal of Public Health*, 82, 1238–1243.

Prapavessis, H., Cameron, L., Baldi, J., Robinson, S., Borrie, K., Harper, T., et al. (2006). The effects of exercise and nicotine replacement therapy on smoking rates in women. *Addictive Behaviors*, 32, 1416–1432.

Richmond, R. L., Harris, K., & de Almeida Neto, A. (1994). The transdermal nicotine patch: Results of a randomised placebo-controlled trial. *Medical Journal of Australia*, 161, 130–135.

Rigotti, N., Thorndike, A., Regan, S., McKool, K., Pasternak, R., Chang, Y., et al. (2006). Bupropion for smokers hospitalized with acute cardiovascular disease. *American Journal of Medicine*, 119, 1080–1087.

Simon, J., Duncan, C., Carmody, T., & Hudes, E. (2004). Bupropion for smoking cessation: A randomized trial. *Archives of Internal Medicine*, 164, 1797–1803.

Tashkin, D., Kanner, R., Bailey, W., Buist, S., Anderson, P., Nides, M., et al. (2001). Smoking cessation in patients with chronic obstructive pulmonary disease: A double-blind, placebo-controlled, randomized trial. *Lancet*, 357, 1571–1575.

Tonnesen, P., Mikkelsen, K., & Bremann, L. (2006). Nurse-conducted smoking cessation in patients with COPD using nicotine sublingual tablets and behavioral support. *Chest*, 130, 314–316.

Tonnesen, P., Tonstad, S., Hjalmarson, A., Leborg, F., Van Spiegel, P., Hider, A., et al. (2003). A multicentre, randomized, double-blind, placebo-controlled, 1-year study of bupropion SR for smoking cessation. *Journal of Internal Medicine*, 254, 184–192.

Tonstad, S., Farsang, C., Klaene, G., Lewis, K., Manolis, A., Perruchoud, A., et al. (2003). Bupropion SR for smoking cessation in smokers with cardiovascular disease: A multicentre, randomised study. *European Heart Journal*, 24, 946–955.

Tsai, S., Cho, H., Cheng, H., Kim, C., Hsueh, K., Billing, C., et al. (2007). A randomized, placebo-controlled trial of varenicline, a selective $\alpha_4\beta_2$ nicotinic acetylcholine receptor partial agonist, as a new therapy for smoking cessation in Asian smokers. *Clinical Therapeutics*, 29, 1027–1039.

Wagena, E. J., Knipschild, P. G., Huibers, M. J. H., Wouters, E. F. M., & van Schayck, P. (2005). Efficacy of bupropion and nortriptyline for smoking cessation among people at risk for or

with chronic obstructive pulmonary disease. *Archives of Internal Medicine*, 165, 2286–2292.

Zellweger, J., Boeleskei, P., Carrozzi, L., Sepper, R., Sweet, R., & Hider, A. (2005). Bupropion SR vs placebo for smoking cessation in health care professionals. *American Journal of Health Behavior*, 29, 240–249.

References

Balk, E. M., Bonis, P. A. L., Moskowitz, H., Schmid, C. H., Ioannidis, J. P. A., Wang, C., et al. (2002). Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *Journal of the American Medical Association*, 287, 2973–2982.

Fagerstrom, K., & Hughes, J. R. (2008). Varenicline in the treatment of tobacco. *Neuropsychiatric Disease and Treatment*, 4, 1–11.

Fiore, M., Jaen, C., Baker, T., Bailey, W., Benowitz, N., Curry, S., et al. (2008). *Treating tobacco use and dependence: Clinical practice guideline*. Washington, DC: US Public Health and Human Services.

Fiore, M. C., Smith, S. S., Jorenby, D. E., & Baker, T. B. (1994). The effectiveness of the nicotine patch for smoking cessation: A meta-analysis. *Journal of the American Medical Association*, 271, 1940–1946.

Fisher, K. J., Glasgow, R. E., & Terborg, J. R. (1990). Worksite smoking cessation: A meta-analysis of long-term quit rates from controlled studies. *Journal of Occupational Medicine*, 32, 429–439.

Fleiss, J. (1994). Measures of effect size for categorical data. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 245–260). New York: Russell Sage Foundation.

Hughes, J. (2009). How confident should we be that smoking cessation treatments work? *Addiction*, 104, 1637–1640.

Hughes, J., Peters, E., & Naud, S. (2008). Relapse to smoking after 1 year of abstinence: A meta-analysis. *Addictive Behaviors*, 33, 1516.

Hughes, J. R., & Callas, P. W. (2007). Errors in interpreting abstinence curves in studies of smoking cessation. *Nicotine & Tobacco Research*, 8, 7–12.

Hughes, J. R., Keely, J., & Naud, S. (2004). Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction*, 99, 29–38.

Hughes, J. R., Keely, J., Niaura, R., Ossip-Klein, D., Richmond, R., & Swan, G. (2003). Measures of abstinence from tobacco in clinical trials: Issues and recommendations. *Nicotine & Tobacco Research*, 5, 13–25.

Mermelstein, R. J., Colby, S. M., Patten, C., Prokhorov, A., Brown, R., Myers, M., et al. (2002). Methodological issues in measuring treatment outcome in adolescent smoking cessation studies. *Nicotine & Tobacco Research*, 4, 395–403.

- Ockene, J. K., Emmons, K. M., Mermelstein, R. J., Perkins, K. A., Bonollo, D. S., Voorhees, C. C., et al. (2000). Relapse and maintenance issues for smoking cessation. *Health Psychology*, 19, 17–31.
- Ossip-Klein, D. J., Bigelow, G., Parker, S. R., Curry, S., Hall, S., & Kirkland, S. (1986). Classification and assessment of smoking behavior. *Health Psychology*, 5, 3–11.
- Raudenbush, S., Bryk, A., Cheong, Y., Congdon, R., & du Toit, M. (2004). *HLM 6: Linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International, Inc.
- Richmond, R. L. (1997). A comparison of measures used to assess effectiveness of the transdermal nicotine patch at 1 year. *Addictive Behaviors*, 22, 753–757.
- Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Russell Sage Foundation.
- Shiffman, S., Scharf, D., Shadel, W., Gwaltney, C., Dang, Q., Paton, S., et al. (2006). Analyzing milestones in smoking cessation: Illustration in a nicotine patch trial in adult smokers. *Journal of Consulting and Clinical Psychology*, 74, 276–285.
- Shipley, R. H., Rosen, T. J., & Williams, C. (1982). Measurement of smoking: Surveys and some recommendations. *Addictive Behaviors*, 7, 299–302.
- Thompson, S., & Higgins, J. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21, 1559–1573.
- Velicer, W. F., & Prochaska, J. O. (2004). A comparison of four smoking cessation outcome measures. *Addictive Behaviors*, 29, 51–60.
- Velicer, W. F., Prochaska, J. O., Rossi, J. S., & Snow, M. G. (1992). Assessing outcome in smoking cessation studies. *Psychological Bulletin*, 111, 23–41.
- West, R., Baker, C., Cappelleri, J., & Bushmakin, A. (2008). Effect of varenicline and bupropion SR on craving, nicotine withdrawal symptoms, and rewarding effects of smoking during a quit attempt. *Psychopharmacology*, 197, 371–377.
- West, R., Hajek, P., Stead, L., & Stapleton, J. (2005). Outcome criteria in smoking cessation trials: Proposal for a common standard. *Addiction*, 100, 299–303.
- West, R., McNeill, A., & Raw, M. (2000). Smoking cessation guidelines for health professionals: An update. *Thorax*, 55, 987–999.