Pfizer Limited
Walton Oaks, Dorking Road, Walton on the Hill,
Tadworth, Surrey KT20 7NS
Telephone: (01304) 616161

Pfizer Global Pharmaceuticals

15th February 2007

Meindert Boysen
Associate Director - STA
Centre for Health Technology Evaluation
National Insitute for Health and Clinical Excellence
Peter House
Oxford street
Manchester
M1 5AN

Dear Meindert

Single technology Appraisal - Varenicline for smoking cessation

Please find attached responses to the queries raised in your letters dated 31st January 2007 and 12th February 2007.

Can you please clarify whether you are expecting to receive a revised submission document as part of the response?

Should you have any further queries regarding any of the answers provided please do not hesitate to contact me.

Yours sincerely

Chris O'Regan MSc RGN
Team Leader
Outcomes Research/Evidence Based Medicine

Enc.

Registered in England: No 526209
Registered Office: Ramsgate Road
Sandwich, Kent CT13 9NJ

A1. Please could you explain why Pfizer has not used the direct trial of varenicline versus NRT as the base case calculation?

Pfizer is mindful that any approach taken to use of data will be questioned by the Evidence Review Group (ERG) and the National Institute for Health and Clinical Excellence.

In this instance there was an option of presenting one of two efficacy values for the Pfizer product as well as for NRT. The decision to use the values derived from the indirect comparison was taken because they were a) the lower of the two efficacy values (the difference in efficacy between varenicline and NRT between the two approaches was not sufficient to modify the cost-effectiveness results) b) based on the results from randomised controlled double-blind studies and c) the NRT efficacy values were closer to those seen in the systematic.

In the interests of openness and transparency Pfizer also presented the results using the open-label varenicline versus NRT study.

You should also be aware that the results of the open-label study only became available in January of this year.

I'm unclear regarding your comment about including new wording in section 5.9. Do you want me to revise the submission document and re-submit?

Were Pfizer to do this, section 5.9.1 would now read:

> *Existing therapies for smoking cessation include Nicotine Replacement Therapy and bupropion.*
>
> [redacted]*. NICE specifies that in the absence of appropriate head-to-head trial data consideration be given to using the results from an appropriately conducted comparison.*
>
> *Based on this we have chosen to use the efficacy values for NRT from the results of a published systematic review and meta-analysis of smoking cessation therapies (Wu et al. 2006) for all comparative economic analyses. The comparison within the paper was an adjusted indirect one after the methods of Bucher et al. (1997) and Song et al. (2003). It is notable that these results conform closely to those from the wider evidence base. A summary of the main findings has been presented above.*

The decision to use the indirect comparison values rather than those from the open-label study does not impact the cost-effectiveness analysis (the results from using the open-label study values are presented as a sensitivity analysis to allow the ERG and NICE to reach their own conclusions regarding this).

Varenicline, NRT and bupropion all provide therapeutic effects in assisting with smoking cessation. The current evidence indicates varenicline has a superior therapeutic effect over the other interventions.

A2. Please provide reasons why you have not considered a 'mixed treatment comparison' approach to answering all the comparisons presented in your decision problem.

This was discussed at the meeting held between members of the Pfizer submission team and representatives of NICE in Manchester on November 23rd 2006. The conclusion was that if the findings of an appropriately conducted indirect comparison were available that these would be sufficient considering the *requirements* of people conducting a review as opposed to the most methodologically advanced approaches methods that may not have achieved widespread acceptance. Of not in this instance is that Mixed Treatment Comparisons are being promoted as the 'best' methodology by the Cochrane Methods group but that this has not been accepted by the mainstream of Colloquium for routine use.

A3. Please could Pfizer request Wu et al to make available all the analyses they present….

I have requested this information from the authors and will forward it on when it becomes available.
It should be noted that a principle difference between the Wu and other systematic reviews in this field is that Wu only included studies in analyses that confirmed the endpoint chemically, believing self report to be unreliable.

A4. Please quote in full the passages of the Wu review that were used and the source of any other data used.

(Page 20 of submission document)
70 trials of NRT versus control at 1 year were identified, (OR 1.71, 95% CI, 1.55–1.88). This was consistent when examining all placebo-controlled trials (49 RCTs, OR 1.78, 95% CI, 1.60–1.99), NRT gum (OR 1.60, 95% CI, 1.37–1.86) or patch (OR 1.63, 95% CI, 1.41–1.89). NRT also reduced smoking at 3 months (OR 1.98, 95% CI, 1.77–2.21). Bupropion trials were superior to controls at 1 year (12 RCTs, OR1.56, 95% CI, 1.10–2.21) and at 3 months (OR 2.13, 95% CI, 1.72–2.64). Two RCTs evaluated the superiority of bupropion versus NRT at 1 year (OR 1.14, 95% CI, 0.20–6.42).

Varenicline was superior to placebo at 1 year (4 RCTs, OR 2.96, 95% CI, 2.12–4.12) and also at approximately 3 months (OR 3.75, 95% CI, 2.65–5.30). Three RCTs evaluated the effectiveness of varenicline versus bupropion at 1 year (OR

1.58, 95% CI, 1.22–2.05) and at approximately 3 months (OR 1.61, 95% CI, 1.16–2.21).

Using indirect comparisons, varenicline was superior to NRT when compared to placebo controls (OR 1.66, 95% CI, 1.17–2.36) or to all controls at 1 year (OR 1.73, 95% CI, 1.22–2.45). This was also the case for 3-month data. Adverse events were not systematically different across studies.

Varenicline, NRT and bupropion all provide therapeutic effects in assisting with smoking cessation. The current evidence indicates varenicline has a superior therapeutic effect over the other interventions.

(Pages 55 to 66 of the submission document)
In order to assess inter-rater reliability on inclusion of articles, the *Phi* statistic (φ) was calculated. This provides a measure of inter-observer agreement independent of chance (Meade et. al. 2001) Odds Ratios [OR] and appropriate 95% Confidence Intervals [CIs] of outcomes were calculated according to the number of events of abstinence reported in the original studies or sub-studies. In circumstances of zero outcome events in one arm of a trial, 1 was added to each arm, as suggested by Sheehe (1966). All NRT interventions versus all controls were pooled using the DerSimonian-Laird (1986) random effects method, which recognises and anchors studies as a sample of all potential studies, and incorporates an additional between-study component to the estimate of variability. Table 17 below records odds ratios and relative risk reductions using both random and fixed effects. The I2 statistic was calculated for each analysis as a measure of the proportion of the overall variation that is attributable to between-study heterogeneity (Higgins and Thompson, 2002). Forest plots are displayed for each primary analysis, showing individual study effect measures with 95% CIs, and the overall DerSimonian-Laird pooled estimate.

A meta-regression analysis on the NRT studies was then conducted with predictors of heterogeneity including the following covariates: placebo control; reporting of sequence generation; reporting of allocation concealment; use of gum or patch; and, method of chemical confirmation of abstinence. When the meta-regression indicated heterogeneity, alternative sensitivity tests using z-tests were conducted to determine differences between the studies, reporting the covariates to the pooled all-studies effect size. Separate pooled analyses of NRT versus placebo, gum versus control and patch versus control were conducted. All analyses at 1 year and also at 3 months were conducted. For bupropion trials, all bupropion trials versus all controls were pooled and a meta-regression analysis was conducted using the following covariates: placebo control; reporting of sequence generation; reporting of allocation concealment; method of chemical confirmation of abstinence; and plans to quit.

Separate meta-regression analyses were conducted and the relevant ORs for the covariates as the exponent of the point estimates were calculated (Thompson and Higgins, 2002). All placebo-controlled trials were pooled and effect sizes at 1 year and at 3 months were evaluated. For head-to-head trials of bupropion versus NRT, pooled random-effects analyses at 1 year and at three months were conducted. For varenicline trials, pooled random-effects analyses of varenicline versus placebo

were conducted at 1 year and at three months and for head-to-head trials of varenicline versus bupropion at 1 year and at three months.

Results

70 RCTs examining NRT versus control interventions were found, 49 of which compared NRT to placebo. Thirty one studies compared NRT to other controlled groups, and one study used both placebo and no intervention as a control group. Thirty-three studies evaluated NRT gum, and 23 evaluated NRT patch. The remaining studies evaluated the efficacy of nicotine inhalers, nasal spray or lozenges. All of the studies provided sufficient details to evaluate NRT versus control at 1 year. Fifty-nine provided sufficient details to evaluate NRT versus control at or about 3 months.

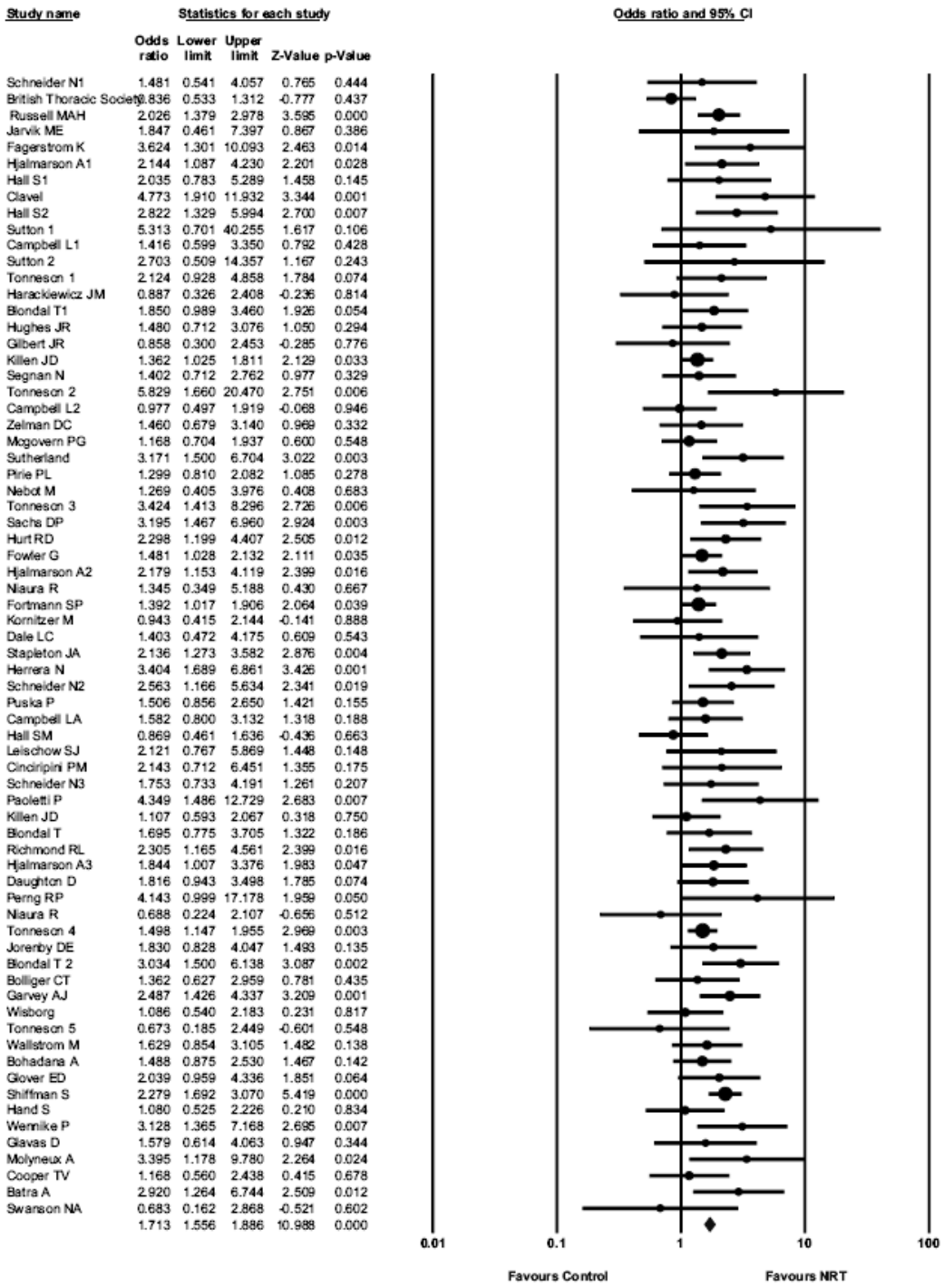
There were11 studies evaluating bupropion versus placebo and one RCT evaluating bupropion with no intervention. A further two of these evaluated bupropion versus NRT. Finally, 4 studies evaluating varenicline versus placebo were identified. Two of these also evaluated varenicline versus bupropion.

Meta-Analysis

NRT

70 trials (total n=28,343) assessing NRT versus controls at 1 year were combined. The pooled OR of smoking cessation favoured NRT over controls (OR 1.71, 95% CI, 1.55-1.88; P=<0.0001, I_2=26.5%, Heterogeneity P=0.02,) (Figure 5). This was consistent when evaluating only placebo controlled NRT trials (49 trials, n=21,512, OR 1.78, 95% CI, 1. 60-1.99; P=<0.0001, I_2 27.4%, Heterogeneity P=0.04) or when evaluating with cessation as sustained abstinence (52 trials, total n=22,704, OR 1.72, 95% CI, 1.54-1.93; P=<0.0001, I_2= 29.4%, Heterogeneity P=0.02) or point prevalence (31 trials, n=10,686, OR 1.53, 95% CI, 1.30-1.81; P=0.01, I_2= 46%, Heterogeneity P=0.01). This was also consistent whether one evaluated NRT gum (33 trials, total n=12,245, OR 1.60, 95% CI,
1.37-1.86; P=<0.0001, I_2= 35.8%, Heterogeneity P=0.02) or NRT patch (23 trials, total n=11,108, OR 1.63, 95% CI, 1.41-1.89; P=<0.0001, I_2=12.3%, Heterogeneity P=0.24).

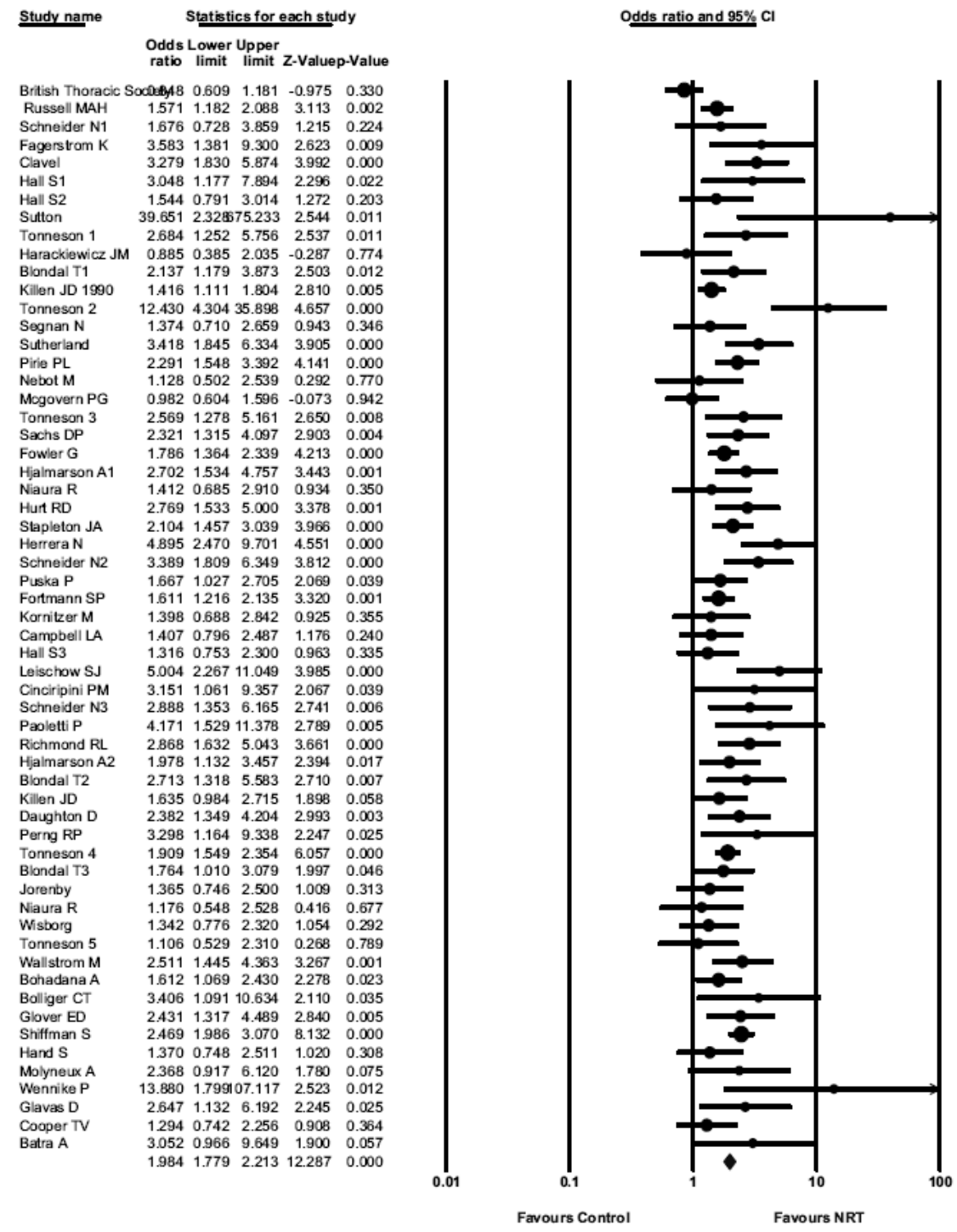
Figure 5: NRT versus Controls at 12 Months



Fifty-nine trials (total n=25,294) provided sufficient details to determine short-term effects of NRT on smoking cessation, as determined at 3 months. The pooled OR of the 59 trials was 1.98 (95% CI, 1.77-2.21; P=<0.0001, I2= 55.5%, Heterogeneity P=<0.0001, See Figure 6). The superiority of NRT over controls was consistent whether one evaluated placebo-controlled trials (42 trials, total n=19,216, OR 2.11, 95% CI, 1.86-2.40; P=<0.0001, I2= 57.6%, Heterogeneity P=<0.001), sustained abstinence (41 trials, total n=19,854, OR 2.04, 95% CI, 1.80-2.31; P=<0.0001, I2= 58%, Heterogeneity P=<0.0001) or point prevalence at 3 months (21 trials, total n=6,453, OR 1.78, 95% CI,1.47-2.14; P=<0.0001, I2=42.4, Heterogeneity

P=0.004). Studies assessing gum versus controls at 3 months (24 trials, total n= 9,347) yielded an OR of 1.71 (95% CI, 1.41-2.07; P=<0.0001, I2= 62%, Heterogeneity P=<0.0001) and studies assessing patch versus controls (21 trials, total n=10,957) yielded an OR of 1.93 (95% CI, 1.67-2.24, P=<0.0001; I2= 35%, Heterogeneity P=0.05).

Figure 6: NRT versus Controls at 3 Months

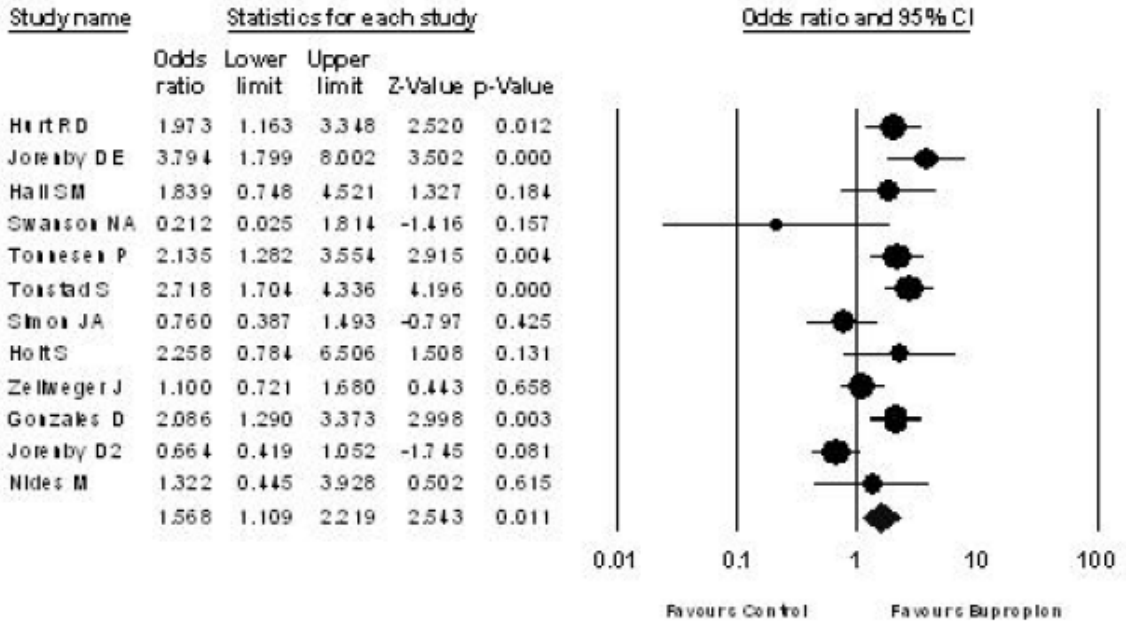


Bupropion

The effect of bupropion on smoking cessation relative to adequate controls at 1 year in 12 trials (total n = 5,228, See Figure 7). The pooled OR was 1.56 (95% CI, 1.10–

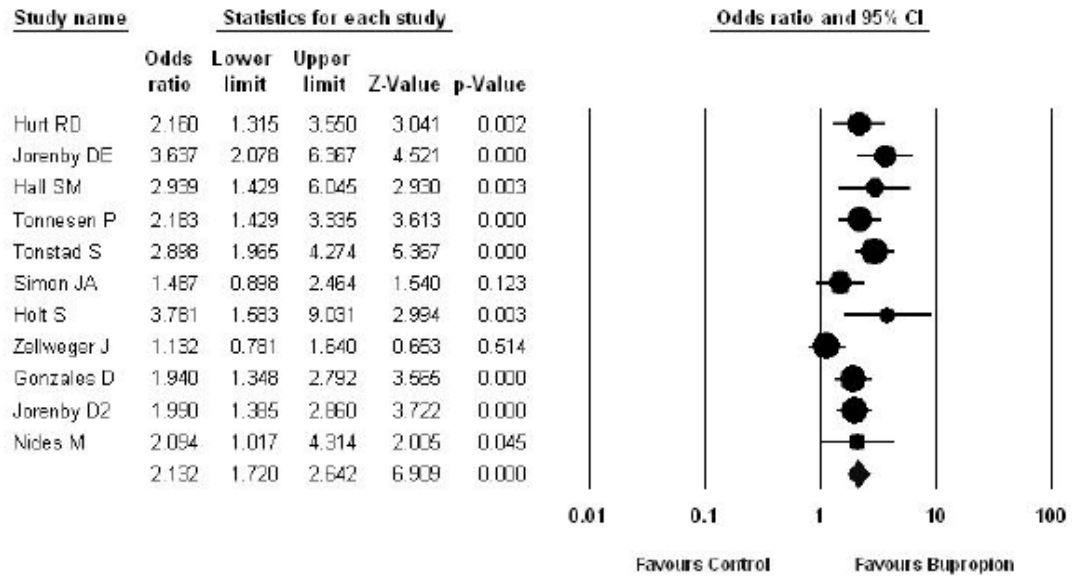
2.21; P = 0.01, I^2 = 71.5%, Heterogeneity P =< 0.001). This effect was consistent whether examining placebo-controls (11 trials, total n = 5,148, OR 1.64, 95% CI, 1.16–2.30; P =< 0.001, I^2 = 72%, Heterogeneity P = 0.001), sustained abstinence (11 trials, total n = 4,613, OR 1.52, 95% CI, 1.04–2.23; P =< 0.0001, I^2 = 73.6%, Heterogeneity P = 0.0001), or point prevalence (10 trials, total n = 4,845, OR 1.56, 95% CI, 1.13–2.16; P =< 0.0001, I^2 = 75.1%, Heterogeneity P =< 0.0001).

Figure 7: Bupropion versus Controls at 12 Months



In evaluating the effect of bupropion on placebo at 3 months (11 trials, total n = 5,148), the OR was 2.13 (95% CI, 1.72–2.64; P =< 0.0001, I^2 = 53.6%, Heterogeneity P = 0.01, See Figure 8). This effect was consistent across sustained abstinence measures (8 trials, total n = 4,143, OR 2.18, 95% CI, 1.67–2.86; P =< 0.0001, I^2 = 63.5%, Heterogeneity P = 0.008) and point prevalence measures (9 trials, total n = 4,765, OR 2.11, 95% CI, 1.77–2.52, P =< 0.0001; I^2 = 38.8%, Heterogeneity P = 0.10).

Figure 8: Bupropion versus Controls at 3 Months



Varenicline

Four studies assessing the effect of varenicline versus placebo at 1 year were pooled (total n=2,528, See Figure 9). The pooled OR is 2.96 (95% CI, 2.12-4.12; P=<0.0001, I2=20.5%, Heterogeneity P=0.20). This effect was consistent with short-term cessation effects (4 trials, total n=2,528, OR 3.75, 95% CI, 2.65-5.30; P=<0.0001, I2=57.7%, Heterogeneity P=0.06, (See Figure 10)).

Figure 9: Varenicline versus Placebo at 12 Months

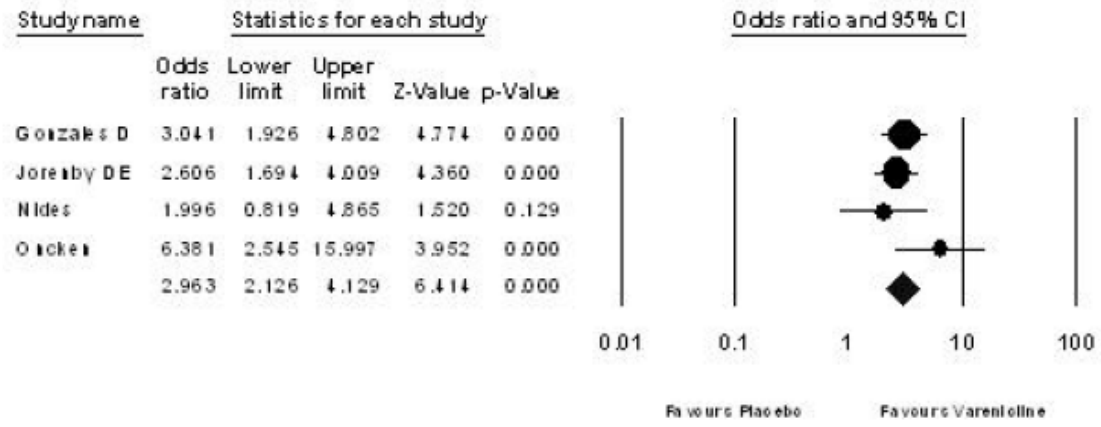
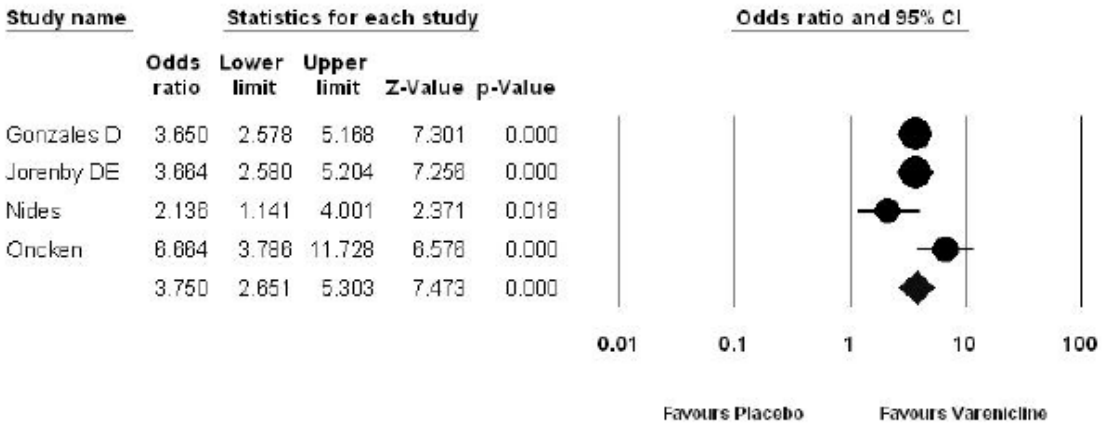


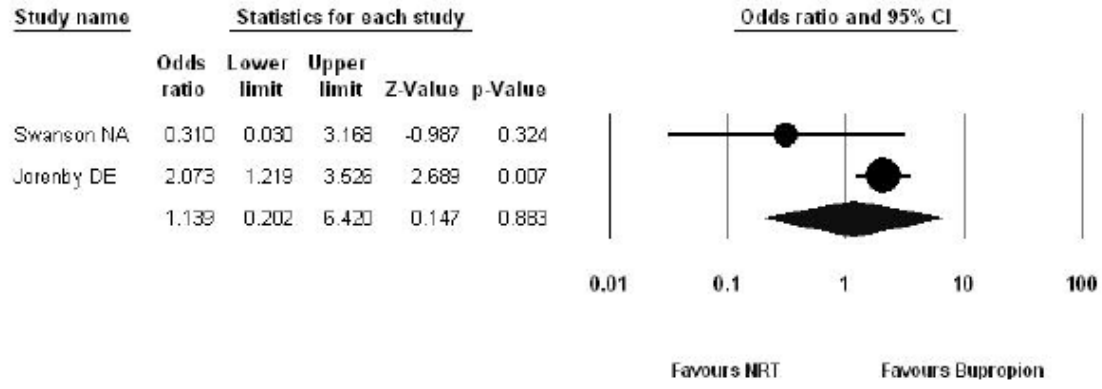
Figure 10: Varenicline versus Placebo at 3 Months



Comparisons

Two trials evaluated the superiority of NRT versus bupropion at 1 year (total n=548, See Figure 11) and found a pooled OR of 1.14 (95% CI, 0.20-6.42: P=0.88, I_2=59%, Heterogeneity P=0.11. Only 1 trial provided details on cessation rates at 3 months and favoured bupropion (OR 2.66, 95% CI 1.70-4.15; P=<0.001).

Figure 11: NRT versus Bupropion at 12 Months



Three trials evaluated the effectiveness of varenicline versus bupropion at 1 year and yielded a pooled OR of 1.58 (95% CI, 1.22-2.05; P=0.001, I_2=0%, Heterogeneity P=0.81, (See Figure 12)) in favour of varenicline. These same trials provided consistent data at 3 months (OR 1.61, 95% CI, 1.16-2.21; P=<0.0004, I2=56.1%, Heterogeneity P=0.10, (See Figure 13)).

Figure 12: Varenicline versus Bupropion at 12 Months

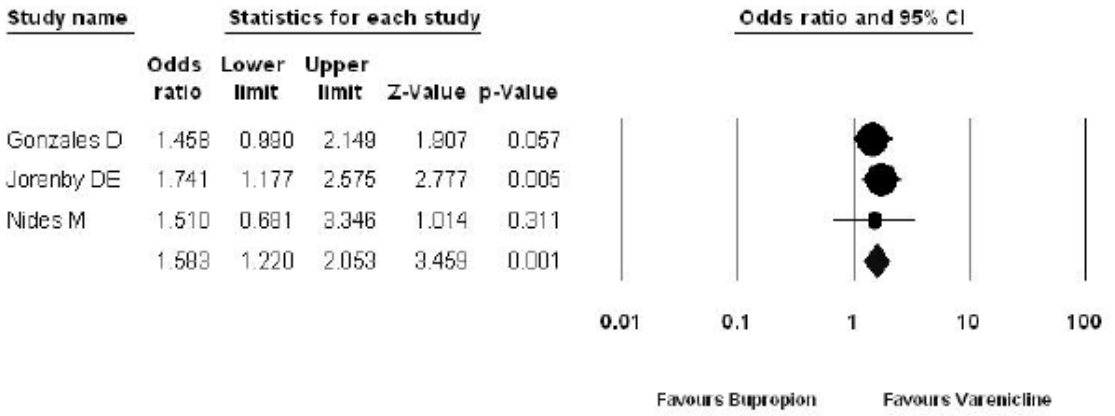
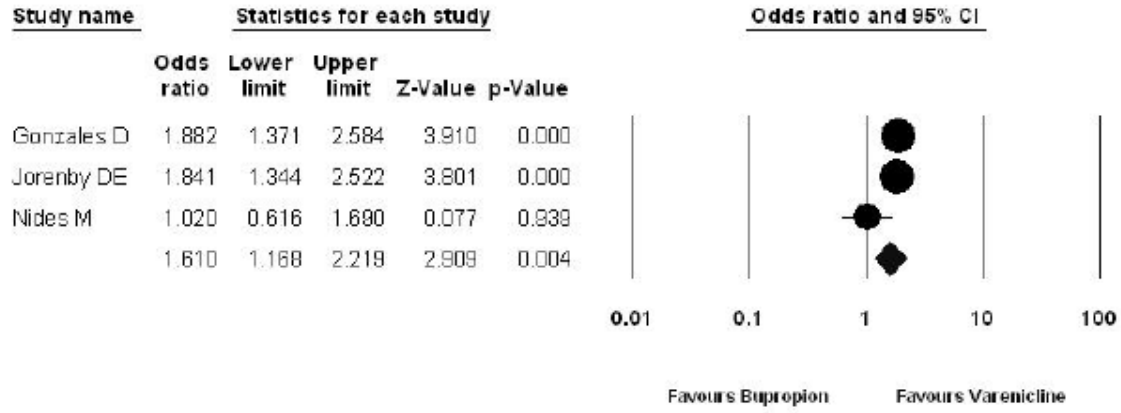


Figure 13: Varenicline versus Bupropion at 3 Months



Using indirect comparisons (Bucher et al. 1997) it was found that bupropion was not superior to NRT when compared to a placebo control at 1 year (OR 0.92, 95% CI 0.64-1.32;test for difference, P=0.65). This was similar for 3- month data (OR 1.01, 95% CI 0.79-1.29; test for difference 0.94). It was found that varenicline was superior to NRT when compared to placebo controls (OR 1.66., 95% CI 1.17-2.36; test for difference, P=0.004 (See Figure 14)) or to all controls at 1 year (OR 1.73, 95% CI 1.22-2.45, test for difference P=0.001). This was also the case when examining 3-month data for placebo controls (OR 1.78, 95% CI 1.23-2.57, test for difference P=0.002, See Figure 15) or all controls (OR 1.89, 95% CI 1.31-2.73, test for difference P=<0.0006).

Figure 14: Indirect Comparison between Varenicline and NRT versus Placebo at 12 Months (Bucher et al. indirect comparison methods)

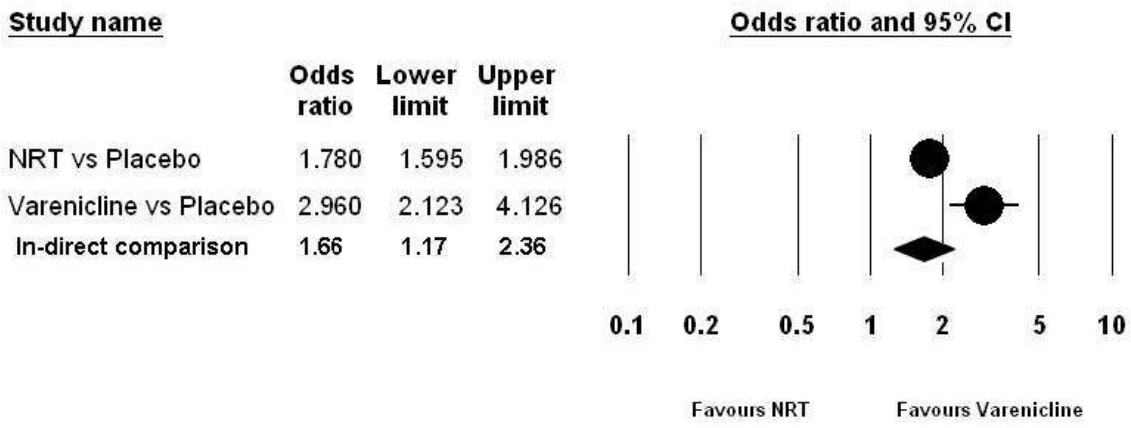
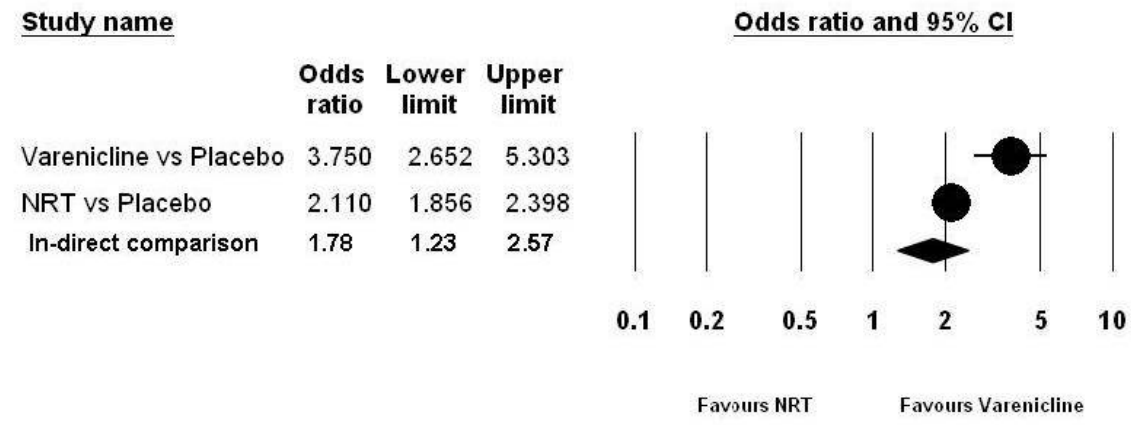


Figure 15: Indirect Comparison between Varenicline and NRT versus Placebo at 3 Months (Bucher et al. indirect comparison methods)



Meta-regression

Variability between study heterogeneity, considering the interventions, the methodological issues and the measurement tools was anticipated. Table 26 displays the covariates predicting heterogeneity in the primary outcomes of the NRT analysis using meta-regression. In this analysis, significant predictors of heterogeneity included: allocation concealment, use of NRT gum; and, methods of chemical confirmation (CO, cotinine, and urine markers). Using sensitivity analysis, only studies (n=3) using urine as a marker were significantly different from the pooled estimate (P=0.03), however, all but one of these studies also used CO as a chemical marker (P=0.5).

When examining covariates in the bupropion trials (See Table 24), only the sequence generation was a significant contributor to heterogeneity. Chemical covariates neared significance (P=0.06). Using sensitivity analyses, chemical confirmation and sequence generation did not predict heterogeneity. A meta-regression on the varenicline studies was not conducted, given the small number of studies.

Table 24: Univariable Meta-regression of Bupropion Studies

Covariates	Point estimate	Lower 95% limit	Upper 95% limit	OR	Lower 95% limit	Upper 95% limit	p-value
Placebo	2.00	-0.14	4.15	7.38	0.86	63.4	0.07
Sequence generation	-0.46	-0.84	-0.08	0.63	0.43	0.92	0.01
Allocation concealment	-0.20	-0.55	0.14	0.81	0.57	1.15	0.25
CO	0.76	0.06	1.46	2.13	1.06	4.30	0.03
Cotinine	-0.76	-1.46	-0.06	0.46	0.23	0.94	0.03
Planned to quit	0.13	-0.22	0.48	1.13	0.80	1.61	0.46

Legend. The point estimate and 95% CIs estimate the unit change in the effect size, whenever the predicted covariate is present. The OR for the point estimates and 95% CI denote the likelihood of covariate affecting the trial effect size.

Adverse events

For NRT trials, the following adverse events were reported significantly more often in active groups than control groups: mouth or throat irritation (n=12); skin irritation (n=11); nausea/vomiting (n=10); coughing (n=9); hiccoughs (n=6); dyspepsia (n=4); watering of eyes (n=3); headaches (n=3); heart palpitations (n=3); sneezing (n=3); sleep disturbances and dream abnormalities (n=2); insomnia (n=2); rhinitis (n=2); vertigo (n=1); taste disturbances (n=1) and muscle aches (n=1).

For bupropion trials, the following adverse events were reported significantly more in the active groups than control groups: dry mouth (9 trials), n=5,065, OR 1.86, 95% CI, 1.49-2.31, P=<0.0001); insomnia (9 trials), n=4,955, OR 1.93, 95% CI, 1.66-2.25, P=<0.0001); gastrointestinal upset (7 trials), n=4,206, OR 1.36, 95% CI, 1.07-1.73, P=0.01) and constipation (5 trials), n=3,373, OR 2.2, 95% CI, 1.53-3.16, P=<0.0001). Other severe events associated with trial participants in the active arms were: septic shock; grand mal seizure; sleep disorders; and anxiety. These were single cases and did not achieve significance.

For varenicline trials, the following adverse events were reported significantly more often than in the placebo groups: nausea (2 trials), n=1,379, OR 3.6, 95% CI, 2.75-

4.71P=<0.0001); flatulence (2 trials), n=1,379, OR 2.18, 95% CI, 1.29-3.68, P=<0.0001); and, constipation (2 trials), n=1,379, OR 2.66, 95% CI, 1.63-4.32, P=<0.0001). Other, severe events that occurred in the active group included: atrial fibrillation, pneumonia, possible stroke, chest pain, and elevated blood pressure. These were, however, single cases and did not achieve significance.

The indirect treatment comparison used in the meta-analysis was between NRT and varenicline using placebo as a reference. The results of this comparison have been discussed and presented in the previous section; the methodology is presented below.

Head-to-head trials provide the strongest inferences regarding intervention superiority (McAlister et al 1999). However, in the absence of head-to-head trials of varenicline versus NRT, indirect comparisons of these interventions versus placebo were conducted using methods described by Bucher et al (1999) and conducted z-tests to confirm. This method maintains the randomisation from each trial and compares the summary estimates of pooled interventions with CIs. Adverse events were calculated, where reported, using Peto's Odds Ratio [OR] with 95% CIs (Yusuf et al 1985). Analyses were conducted using StatsDirect (version 2.5.2, www.statsdirect.com) and Comprehensive Meta-analysis (version 2, www.meta-analysis.com).

A4 cont. Table 41 (p.95) presents efficacy rates, the source of which is not transparent…

The estimates were pooled by a statistician in Pfizer. We agree that there is not a legitimate method in literature to pooling rates, however we do recognise that the statistician was operating from the premise that, as the trial designs mirrored each other and the results were (therefore) markedly similar it was reasonable to pool. The reality of this is that the cost-effectiveness results are not impacted.

A4 cont. How was the efficacy value for NRT in Table 41 derived from the odds ratio values in Wu?

The Wu paper calculates the indirect comparison to find the probability of Champix vs NRT. This is estimated as being 1.66 .

The abstinence rate at 1 year for Champix is 22.5% (pooled analysis a3051028 and A3051036 studies).

We have used the formula below

$$(ODDS_{NRT.Champix} * P_{Champix}) / (1 - P_{Champix} + ODDS_{NRT.Champix*} P_{Champix})$$

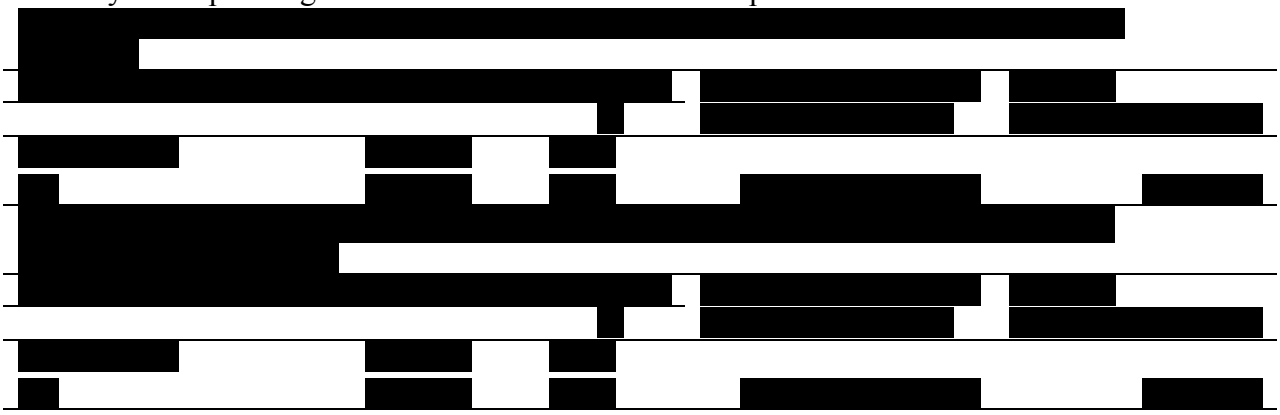
Imputing the odds ratio of NRT vs varenicline (0.66, inverse of 1.66) and the abstinence rate at 1 year for varenicline to retrieve the abstinence rate for NRT.

This gives an abstinence rate at 1 year of 14.9%

A5. The manufacturer's submission claims that the Pfizer analysts have used odds ratios to generate the probabilistic sensitivity analysis…..

The odds ratio together with the upper and lower confidence interval, and the random number generated from the lognormal distribution overimposed can be found in the spreadsheet PSAcalculation of the models we have submitted (range:B100:H153).

B1. Please make the correct event numbers available for tables 22 and 23….
Thank you for pointing this out. The corrected tables are presented below:



B2. The manufacturers submission suggests an efficacy rate of 15.7% for Bupropion, yet the model suggest this value is 15.5%. Which value is correct?
Thank you for pointing this out. The correct value should be 15.7% and we present a re-worked main analysis below:

Champix vs Bupropion (rate for Champix = 15.7%)

Model year	2	5	10	20	Lifetime
Champix Treatment Related Costs (Millions)	1,995	4,404	8,615	17,750	34,019
Bupropion Treatment Related Costs (Millions)	1,735	4,171	8,457	17,778	34,331
difference (Millions)	*260.2 [15%]*	*232.8 [5.6%]*	*158.6 [1.9%]*	*-28.1 [-0.2%]*	*-311.9 [-0.9%]*
Champix QALYs (Thousands)	5,059	11,677	20,411	31,782	42,135
Bupropion QALYs (Thousands)	5,059	11,675	20,403	31,755	42,066
difference (Thousands)	*0.3 [0%]*	*2.2 [0%]*	*8.1 [0%]*	*27.3 [0.1%]*	*69.3 [0.2%]*
Champix Life Years (Thousands)	6,204	15,041	28,346	50,530	86,711
Bupropion Life Years (Thousands)	6,204	15,039	28,338	50,493	86,546
difference (Thousands)	*0.1 [0%]*	*1.6 [0%]*	*8.3 [0%]*	*37.6 [0.1%]*	*165.6 [0.2%]*

Incremental Cost per additional QALY	**767,546**	**107,816**	**19,502**	**Dominates**	**Dominates**
Incremental Cost per LYG	**2,328,986**	**142,545**	**19,195**	**Dominates**	**Dominates**

A1. (From letter dated 12th February 2007) Please could you explain the apparent inconsistency in the Markov transition/population calculations?

Pfizer agrees that the population is 3,174,339 patients in the first year but according to our calculations, the number of the patient stay the same during the time horizon of the model.

We also have conducted a validation exercise (attached spreadsheet. 'BENESCO Model_NICE_validation').

To validate whether the number of patients add up to the same number, logically, we sum up, in each period of time, the patients in each state.

We have categorised the patients in

1) Patients still alive from year before/Smokers
2) Patients still alive from year before/Quitters

We have added up these two groups to produce a group called "still alive"

3) Patients dead

We have then added these groups together and they produce the number of 3,174,339 in each period of the time horizon.