University of Sheffield

Varenicline for Smoking Cessation: A Single Technology Appraisal

March 2007

Danny Hind
Paul Tappenden
Jean Peters
Karligash Kenjegalieva

School of Health and Related Research (ScHARR)
Regent Court, 30 Regent Street, Sheffield S1 4DA. Tel: 0114 222 5454. Fax: 0114 272 4095

Acknowledgements

Dr Eva Kaltenthaler (Managing Director, ScHARR Technology Assessment Group), Jim Chilcott, (Technical Director, ScHARR Technology Assessment Group), Kate Cahill (Co-ordinator, Cochrane Tobacco Addiction Group), Hazel Pilgrim (Cost-effectiveness analyst, ScHARR Technology Assessment Group), ScHARR and Prashanth Kandaswamy (Technical Lead, NICE), commented on draft versions of the report. Kate Cahill, Sarah Plant and Sue Russell (Smoking Cessation Co-ordinators, Sheffield PCT) provided advice on the background, decision problem and assumptions behind the model. Edward Mills (research fellow, Department of Clinical Epidemiology and Biostatistics, McMaster University) provided unpublished data. The authors wish to thank all of the above. Responsibility for the accuracy of the report lies entirely with the authors. The authors also wish to thank Andrea Shippam for her help in preparing and formatting the report.

This report was commissioned by the NHS R&D HTA Programme on behalf of the National Institute for Health and Clinical Excellence. The views expressed in this report are those of the authors and not necessarily those of the NHS R&D HTA Programme or the National Institute for Health and Clinical Excellence. The final report and any errors remain the responsibility of the University of Sheffield. Jim Chilcott and Eva Kaltenthaler are guarantors.

Contents

List of Tables

List of Figures

List of Abbreviations

ACT	Abstinent-Contingent Treatment
BENESCO	Benefits of Smoking Cessation on Outcomes
BMC	BioMed Central
CHD	Coronary Heart Disease
CIC	Commercial-In-Confidence
COPD	Chronic Obstructive Pulmonary Disorder
EED	Economic Evaluation Database
ERG	Evidence Review Group
EU	European Union
HEA	Health Education Authority
HECOS	Health-Economic Model for Smoking-Related Morbidity
HRQoL	Health-Related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
ITT	Intention-To-Treat
LYGs	Life-Years Gained
MeSH	Medical Subject Headings
nAChR	Neuronal Nicotinic Acetylcholine Receptors
NNT	Number Needed to Treat
NRT	Nictotine Replacement Therapy
OR	Odds Ratio
PCT	Primary Care Trust
QALYs	Quality-Adjusted Life Years
QUOROM	Quality Of Reporting Of Meta-analyses
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SPC	Summary of Product Characteristics
VBA	Visual Basic for Applications

Definitions

Abstinent-Contingent Treatment. A programme in which the smoker makes a commitment to stop smoking on or before a particular date (the ‘target stop date’).

Point prevalence. The measure of a condition (in this case whether or not the individual is still a smoker) in a population at a given point in time

Sustained abstinence. Completely tobacco-free time from the delivery of the intervention to the time of follow-up.

Evidence Review Group's Report Template

This template should be completed with reference to NICEs 'Guide to the Methods of Single Technology Appraisal'

1 SUMMARY

1.1 Scope of the submission

The scope of the submission is appropriate.

1.2 Summary of submitted clinical effectiveness evidence

The selection and use of evidence in the manufacturer's submission exaggerates the effect size of varenicline when compared indirectly to nicotine replacement therapy (NRT); their incremental cost-effectiveness ratio (ICER) is artificially low as a consequence. However, varenicline is still likely to be superior to both comparators.

[redacted] The difference in treatment effects observed in a meta-analysis of two studies comparing varenicline (as licensed in the EU) and bupropion (OR 1.59, 95% CI 1.21 to 2.10) would be found by chance alone once in one thousand times (p=0.001).

1.3 Summary of submitted cost effectiveness evidence

The submission reports the methods and results of a state transition model (the Benefits of Smoking Cessation on Outcomes, or BENESCO model) to estimate the incremental cost-effectiveness of varenicline as compared against bupropion, NRT and placebo. The model suggests that varenicline dominates (I.e. is more effective and less expensive than bupropion, NRT and placebo. Treatment efficacy for each of the interventions is based on the results of a pooled analysis of 1-year quit rates sourced from the clinical trials of varenicline. Beyond this point, the model assumes that short-term efficacy translates into long-term health gains and associated cost savings. This assumption of sustained benefit is subject to a substantial degree of

uncertainty; shorter time horizons may be less uncertain, but may underestimate the benefits of varenicline. Longer time horizons provide more favourable cost-effectiveness estimates for varenicline, yet are subject to a much greater degree of uncertainty. The probabilistic sensitivity analysis suggests that the probability that varenicline produces the greatest amount of net benefit is estimated to be 0.70. However, this was restricted to a limited number of parameters and is inherently flawed. The true uncertainty surrounding the incremental cost-effectiveness of varenicline has not been appropriately addressed within the submission.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

- The manufacturers have recruited a highly experienced team of researchers from McMaster University to produce and publish a systematic review, which they have used as the basis for their analysis. In the clinical review, this resulted in an unusual degree of co-operation with respect to requests for additional data.

- The structural assumptions included in the submission model appear to be intuitively sensible, and the costs and consequences of most important smoking-related morbidities (lung cancer, COPD, asthma, CHD and stroke) are included in the analysis.

1.4.2 Weaknesses

- The manufacturer's use of indirect comparisons is inappropriate and their composition is problematic.

- The model assumes only a single quit attempt using a single smoking cessation intervention (varenicline, bupropion, NRT or placebo). In reality, smokers may attempt to quit more than once using several smoking cessation technologies. The costs and health outcomes of repeated quit attempts are not considered within the evaluation.

- The model extrapolates lifetime outcomes for subjects attempting to quit smoking (up to 81 years of extrapolated costs and consequences) based on a pooled analysis of 1-year efficacy outcomes from clinical trials.

- The model uses a large number of parameter values derived from US studies which may not reflect the smoking/abstinence behaviour of the population of England and Wales.

- Methods for identifying and selecting costs and health utilities associated with morbidities are not reported or justified within the sponsor submission.

- The presence of multiple computational errors should be borne in mind when considering cost-effectiveness results reported within the sponsor submission.

- The sensitivity analysis presented within the submission is very narrow and underestimates the true uncertainty surrounding the incremental cost-effectiveness of varenicline.

- The external validity of the model has not been demonstrated by the manufacturer.

1.4.3 Areas of uncertainty

The key area of uncertainty concerns the long-term experience of subjects who have remained abstinent from smoking beyond 12-months. The health economic model makes an assumption of sustained benefit for the remaining 81 years of the time horizon. The validity of the assumption of sustained benefit between treatment groups is unclear.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

For example, if the appraisal about metastatic hormone-refractory prostate cancer, and the submission gives details predominantly about prostate cancer in general, are those details relevant to the appraisal.

Section 4 (manufacturer's submission, page 23) provides a good succinct summary of the public health problem.[1]

2.2 Critique of manufacturer's overview of current service provision

For example, does the submission concord with opinions of clinical and patient experts? Has sufficient backing evidence been given about how often the comparators and intervention are used? Are the constraints of UK market authorisations considered.

Section 4 (manufacturer's submission, page 23)[1] summarises the existing options for treatment well, as validated by smoking cessation co-ordinators from Sheffield Primary Care Trust. There is no backing evidence about how often the comparators and intervention are used (presumably NICE means the manufacturer's own intervention here). They do say that 4 million people per year attempt to quit (bottom of page 23), but there is no indication of the level of usage of varenicline or its comparators. Pages 10-12 describe the market authorisations for varenicline; those for the comparator therapies are not discussed in any detail, but they do mention (top of page 24) that over the counter use of NRT is permitted although some GPs will prescribe.

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

3.1 Population

Description of relevant patient population(s) and comment on whether they are appropriately defined in the submission.

The manufacturer's population in the decision problem is the same as that in the Final Scope issued by NICE (October 2006): "Adults who smoke tobacco products and have indicated a desire to quit smoking".[2]

The manufacturer's submission does not present evidence for the clinical effectiveness of varenicline in any particular population subgroups.[1] The Evidence Review Group (ERG) notes that Cochrane reviews of smoking cessation interventions (other than varenicline) have been undertaken for the following subgroups: hospitalised[3] and preoperative[4] patients; pregnant women;[5] and, people with Chronic Obstructive Pulmonary Disorder (COPD).[6] Subgroup analysis is reported within the cost-effectiveness section (Section 6.3.2, p.117) according to age and sex; however, this analysis uses efficacy estimates which relate to the intention-to-treat populations within the clinical trials hence the results of the subgroup analysis should be treated with caution.

The manufacturer's submission notes (Section 3, p15) population-related restrictions (renal impairment, children and adolescents, pregnant or breast-feeding), contraindications (hypersensitivity to substance or excipients), special warnings and precautions for use (related predominantly to smoking cessation, rather than varenicline *per se*).[1]

3.2 Intervention

What is the technology and what is its relevant or proposed marketing authorisation/ CE mark?

Varenicline (Champix®, Pfizer Inc, UK) is described as a selective nicotinic receptor partial agonist. It was designed to selectively activate the *α*4*β*2 neuronal nicotinic acetylcholine receptors (nAChRs), mimicking the action of nicotine by causing a moderate and sustained release of mesolimbic dopamine.

Oral varenicline is indicated for smoking cessation in adults. The recommended dose is 1 mg varenicline twice daily following a 1-week titration as follows:

Days 1 – 3:	0.5 mg once daily
Days 4 – 7:	0.5 mg twice daily
Day 8 – End of treatment:	1 mg twice daily

The marketing authorisation recommends that the patient should set a date to stop smoking. Varenicline dosing should start 1-2 weeks before this date. Patients who cannot tolerate adverse effects of varenicline may have the dose lowered temporarily or permanently to 0.5 mg twice daily. Patients should be treated with varenicline for 12 weeks.

The marketing authorisation also recommends that patients who have successfully stopped smoking at the end of 12 weeks, may consider an additional course of 12 weeks treatment with varenicline at 1 mg twice daily. No data are available on the efficacy of an additional 12 weeks course of treatment for patients who do not succeed in stopping smoking during initial therapy or who relapse after treatment. In smoking cessation therapy, risk for relapse to smoking is elevated in the period immediately following the end of treatment. In patients with a high risk of relapse, dose tapering may be considered.

There are a variety of warnings and contraindications which appear in both the summary of product characteristics and the manufacturer's submission to NICE.[1]

3.3 Comparators

Relevant comparators in an NHS context, justification for choice of comparators. For example, where hard evidence is not available, has the manufacturer asked an unbiased clinical panel, or done its own survey, and does it agree with what the clinical experts for the appraisal say?

The manufacturer's comparators in the decision problem include three of those from the Final Scope issued by NICE (October 2006): bupropion; NRT; and, no therapy (placebo). The NICE final scope also mentioned, "other smoking cessation interventions, as appropriate without varenicline", however, our clinical advisors have not drawn our attention to any.[2]

NICE currently recommends ***either*** bupropion ***or*** NRT, as part of an abstinent-contingent treatment (ACT), in which the smoker makes a commitment to stop smoking on or before a particular date (target stop date) and is given advice and encouragement to do so.[7] GPs are supposed to offer brief interventions (opportunistic advice), but the ideal is that the patient will get intensive support, e.g. NHS Stop Smoking Services.[8] Health Education Authority guidelines make it clear that the optimal comparator is intensive behavioural support plus NRT or bupropion).[9]

A representative of Sheffield PCT advised us that they are currently implementing NICE guidance, giving: (1) counselling (mainly coping skills; either 1:1 or in a group) with either NRT or bupropion as preferred by the GP; (2) counselling alone where the individual refuses pharmacotherapy.

In summary, the manufacturer's submission represents current available NHS treatment options well. However, intensive support alone (without pharmacotherapy), which is still an option for the NHS, is not included as a comparator in the manufacturer's submission.

3.4 Outcomes

Including clinical effectiveness, adverse events quality of life and health economic outcomes and a discussion of appropriate mechanisms for measuring these outcomes? Critique of whether focus of submission is on the appropriate outcomes. Comment if whether the analysis has been limited to non-ideal outcomes.

The Final Scope issued by NICE (October 2006) requested the following outcome measures: survival; morbidity related to smoking; quit rates at 4 weeks; 6 months, 12 months and at longer periods; adverse effects of treatment; health-related quality of life.[2]

The manufacturer has excluded the 4 week quit rates as an outcome because the duration of treatment is 12 weeks. They excluded survival and smoking related morbidity as outcomes because, they state, these outcomes "are related to the giving up (or not) of smoking rather than the method used" (Manufacturer's submission, Section 2, page 13).[1]

The manufacturer's submission provides continuous abstinence rates at 52 weeks for all trials as per NICE's decision problem. For the manufacturer's trials this is reported as a secondary outcome (and their submission often refers to it as such), which is confusing given that this outcome is the most clinically meaningful and drives the manufacturer's own health economic model.

3.5 Time frame

Pfizer notes that 12 months is the maximum period for which they have data for varenicline. Trial assessments were made at 52 weeks after the course of varenicline began, which is common in trials of smoking cessation therapies.

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

A substantial proportion of the clinical effectiveness section of the manufacturer's submission,[1] is based on a separate piece of work, undertaken by researchers at McMaster University (Wu et al.[10]), funded by Pfizer, and previously published in a peer-reviewed journal (*BMC Public Health*). Throughout, we refer to this source study as, 'the McMaster review'. The McMaster review compares varenicline, bupropion and NRT to placebo, varenicline to bupropion and bupropion to NRT. Importantly, it then uses a common placebo comparator to perform a frequentist indirect comparison of varenicline versus NRT. The results are an important driver of cost-effectiveness within the manufacturer's model. For this reason, we sometimes distinguish criticisms of the design and conduct of the McMaster review from those specific to the manufacturer's submission.

Note that, aside from the indirect comparison, the McMaster review makes comparisons of clinical effectiveness previously undertaken in three (publicly-funded) Cochrane reviews, the latest versions of which are: Silagy 2004 (NRT);[11] Hughes 2007 (bupropion);[12] and, Cahill 2007 (varenicline).[13] The methods and the results of the McMaster review differ from these Cochrane reviews and these differences and their effects are discussed throughout this section. In critiquing the McMaster report, we have generally assumed that information in the Cochrane reviews is correct apart from where we have found or been made aware of errors in the reviews by Silagy[11] and Hughes[12]

4.1.1 Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

Databases and other sources including unpublished sources, any restrictions.

The manufacturer's submission

The manufacturers searched ten publicly accessible databases (to December 1, 2006): MEDLINE, EMBASE, Cochrane, AMED, CINAHL, TOXNET, Development and Reproductive Toxicology, Hazardous Substances

Databank, Psych-info and Web of Science. They also searched the bibliographies of published systematic reviews and Pfizer's own clinical trials database. Searches were not limited by language, sex or age.[1] As in most of the clinical effectiveness section, the content is virtually identical to that of the McMaster review.[10]

Manufacturer's submission: clinical evidence search strategy

The search utilises a combination of free-text and MeSH terms. However it is not clear from the reporting of the search strategy which terms are free-text and which are MeSH. Regarding the MeSH terms, it is not reported whether these were exploded or focused. Similarly it is not reported which fields were searched for the free-text terms – e.g. all fields, title and abstract, title only etc. Boolean operators are not reported so it is not possible to identify the relationship between the search terms. No methodological search filters have been used and the search utilised terms for the intervention only – no terms for population, outcome or comparator(s) were included in the search. In general, the search methodology is not sufficiently "transparent" to replicate exactly.

Manufacturer's submission: Cost-effectiveness Search Strategy

The terms used for the cost-effectiveness search appear to be exactly same as the clinical evidence search; therefore, all the issues surrounding the clinical effectiveness searches also apply to the cost-effectiveness searches. Four databases were searched to identify studies relating to the cost-effectiveness of varenicline. Two of these databases were the same as the clinical evidence search, so presumably the same results were retrieved. Two additional databases that had not been searched for clinical evidence were also searched for cost-effectiveness evidence. One of these was the NHS Economic Evaluation Database (EED); this was the only database where a different search strategy was applied. The search strategy reported for EED is very basic (searching for the term 'smoking') which at the time of writing would retrieve 355 references. If a more sensitive search strategy was used,

including cost-effectiveness terms, fewer references would be retrieved and these would be more specific to the topic.

The McMaster review

The search strategy for the McMaster review is identical to that of the manufacturer's submission,[1] but it does not specify search terms (presumably it included terms to identify NRT and bupropion trials, as well as trials for varenicline) and there is no sample search strategy as in the manufacturer's submission.[10]

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The manufacturer's submission

The manufacturer's evidence review eligibility criteria were any RCT of at least one year's duration which evaluated NRT (however delivered), bupropion or varenicline using chemical confirmation of smoking cessation, defined as either sustained abstinence or point-prevalence of abstinence (Manufacturer's submission, Section 5.2.2, page 29).[1] The manufacturer claims to have excluded dose ranging studies, non-RCTs, post-hoc analyses, maintenance therapy, and studies that reported outcomes as self-report were excluded (although see below, Section 4.1.3 on the McMaster review).

The exclusion of dose ranging studies results in the exclusion of the Phase II comparative trial, published by Nides and colleagues (study A3051002) although it does have bupropion and placebo arms.[14] The Nides study could also be excluded on the grounds of the duration of the treatment (6 weeks rather than 12, as per the EU marketing authorisation), or because it recruited participants with previous exposure to the active comparator. The exclusion of this study from the manufacturer's meta-analysis *would have* produced conservative rather than optimistic results, reduced statistical heterogeneity and would have resulted in a less favourable ICER. However, the McMaster team's indirect comparison of varenicline and NRT (which the manufacturer used) was, in fact, informed by the McMaster team's meta-analysis, which included the Nides study. *The effect of McMaster's inclusion of this study in*

their meta-analysis is to produce optimistic rather than conservative results, an increase in statistical heterogeneity and a more favourable incremental cost-effectiveness ratio for varenicline (see also this document, Section 4.1.4 on the Oncken study).

The McMaster review

The McMaster review (see the current document, top of Section 4.1) states the same inclusion criteria as the manufacturer's submission.[10]

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded.

The manufacturer's submission

Although the manufacturer's submission states that eligible studies were those evaluating NRT, bupropion or varenicline, they only report, tabulate and discuss studies which evaluate varenicline (regardless of comparator).[1] We have re-run what we believe might approximate the manufacturer's search strategy and confirm that the table of identified varenicline studies (manufacturer's submission, Section 5.2.1, pp28-29, Table 1) is complete. Our Table 1, below, is not meant to be exhaustive, but is intended to inform the Appraisal Committee of the key studies mentioned in the ERG's report.

The manufacturer's list, "published studies trials including varenicline" (Section 5.2.1, p27) is incomplete, as it omits the paper published by Reeves. We believe that this study is represented (as study "A3051037") in manufacturer's submission (Section 5.2.1, pp28-29, Table 1). No other publication is available from this study. Although it is a point-prevalence rather than continuous abstinence study, the manufacturer's submission claims these study designs were eligible for inclusion in the review. The study does not appear in the McMaster review,[10] although it does appear in Cahill's Cochrane review (where it is analysed separately from the continuous abstinence studies).[13] *The exclusion of this study from the manufacturer's meta-analysis produces conservative rather than optimistic results, reduces statistical heterogeneity and would results in a less favourable incremental cost-effectiveness ratio for varenicline.*

Table 1 Key varenicline studies mentioned in this report

Publication	Description
Gonzales 2006 (A3051028)[15]	Interventions: **varenicline** (1mg x2/day, n=352), **bupropion** (n=329) or **placebo** (n=344); Outcome used: continuous abstinence at 9-52 weeks; informs indirect comparison
Jorenby 2006 (A3051036)[16]	Interventions: **varenicline** (1mg x2/day, n=344), **bupropion** (n=342) or **placebo** (n=341); Outcome used: continuous abstinence at 9-52 weeks; informs indirect comparison
Nides 2006 (A3051002)[14]	Interventions: **varenicline** (0.3mg 1/d or 1.0mg 1/d or 1.0mg 2/d for 6w; n=382), **bupropion** (n=128) or **placebo** (n=127); Outcome used: continuous quit rate from wk 4 to wk 52; informs indirect comparison
Oncken 2006 (A3051007 or A3051008?)[17]	Interventions: **varenicline** (four regimens, titrated or non-titrated, 2-12w; n=518) or **placebo** (n=129); Outcome used: continuous verified abstinence at wks 9-52; informs indirect comparison
Reeves 2006 (A3051037?)[18]	Interventions: **varenicline** (1mg x2/d, 52 wks; n=251) or **placebo** (126); Outcome used: 7-day CO verified point-prevalence abstinence; does not inform indirect comparison
Tonstad 2006 (A3051035)[19]	Successful quitters following a 12 wk course of varenicline randomised to: **varenicline maintenance** (1mg x2/day; n=603) or **placebo** (n=607) for a further 12 wks. Outcome used: continuous validated abstinence at wk 52; does not inform indirect comparison
Unpublished CIC study (A3051044)	Interventions: **varenicline** (1mg x2/d, 12 wks; n=377) or **NRT** (n=378)
Wu 2006 (McMaster study)[10]	Systematic review and meta-analysis comparing **varenicline**, **bupropion** and **NRT** with placebo and each other.

The manufacturers (submission Section 5.2.3, page 30) state that they consider four Phase III trials to be relevant: those published by Gonzales (study A3051028), Jorenby (study A3051036) and Tonstad (study A3051035), along with a fourth unpublished study (A3051044).[1] The inclusion of the Tonstad study (study A3051035), which evaluates maintenance therapy, appears to contravene the manufacturer's own exclusion criteria (Section 5.2.2, p29), although *it does not contribute to the base case health economic model and so does not affect the base case cost-effectiveness estimates*.

The NICE rubric gives the manufacturer explicit instructions that, "Where studies have been excluded from further discussion, a justification should be provided to ensure that the rationale for doing so is transparent. A flow diagram of the numbers of studies included and excluded at each stage should be provided… as per the QUOROM statement". Neither of these conditions has been met, and the reader is left: (a) with no idea how many citations the literature search has retrieved; (b) with no record of the process of study selection; and, (c) to make inferences as to why eight studies in Table 1 of the Pfizer submission (not counting the three which are ongoing) have been omitted (as opposed to explicitly excluded with justification) from further discussion.

The McMaster review

The McMaster review team retrieved 70 NRT studies, 12 bupropion studies and 4 varenicline studies they considered eligible for inclusion in their meta-analyses.[10] The McMaster review included one study, Swanson, which the Cochrane NRT review[12] had excluded because they were unable to confirm the denominators. The McMaster team do not explain how they derived these denominators. *The impact of the inclusion of this study on the ICER is uncertain*.

Inclusion of ineligible studies

The McMaster review claims to have used only studies with chemical confirmation of smoking cessation although, in practice, they include the following studies: Wisborg 2000 (self-report confirmed by telephone); Russell

1983 (where only 66% of those claiming to have quit were chemically confirmed); Fagerstrom 1984 (who used a random subset); Killen 1997 (who only chemically confirmed those who lived in the area – percentage unknown); Clavell 1985 (unclear); Zelman (independent observer report); and, Perng 1998 (no validation at 12-months). *The inclusion of these studies is not likely to greatly impact either the results of the meta-analysis or the ICER.*

The McMaster meta-analysis includes two studies (Wennike 2003, Bolliger 2000) where the primary outcome was to assess smoking reduction (rather than cessation). These studies were recognised by Silagy's Cochrane review but excluded from his meta-analysis of 6- to 12-month smoking cessation. Silagy only identifies data on "reduction to <50% of baseline cigarette consumption at longest follow-up".[11] Whether or not these studies presented data on cessation which Silagy omitted, or the McMaster reviewers have confused reduction data for cessation data, we are unclear. *The impact of the inclusion of these studies on the ICER is uncertain but unlikely to be great*.

Open label studies and studies with concomitant therapies

The McMaster review includes studies which compare pharmacotherapy plus counselling (not necessarily intensive therapist-led) versus counselling alone *without placebo*: Harackiewicz 1988, Fagerstrom 1984, Hall 1985, Zelman 1992, McGovern 1992, Pirie 1992, Nebot 1992, Fortmann 1995, Niaura 1999, (NRT gum), Russell 1983, Gilbert 1989 ('offer of NRT gum') Cinciripini1996 (NRT patch); Molyneaux 2003 (choice of NRT). The McMaster review also includes studies which compare pharmacotherapy versus nothing (not even placebo): Sutton 1987, Sutton 1988, Niaura 1994 (all NRT gum). The McMaster review includes studies which compare different forms of NRT: Puska 1995 (NRT patch and gum versus gum alone; also a dose reduction study) and Dale 1995 (NRT patches – a dose comparison study); Tonnesen 2000 (low dose patch versus high dose patch versus NRT inhaler versus high patch and inhaler); Blondal 1999 (spray and patch versus patch). *The inclusion of these studies in the McMaster meta-analysis will result in the effect size of NRT being reduced, and the effect size of varenicline (when*

indirectly compared to NRT) being exaggerated (see this document, section 4.1.7 on the McMaster review for evidence).

4.1.4 Details of any relevant studies that were not included in the submission ?

Potentially eligible studies omitted by the manufacturer's submission

Of the eleven studies reported in Table 1 of the manufacturer's submission, (Section 5.2.1, pp28-29), and to which there is no further reference, three are included in the Cochrane Review by Cahill and colleagues.[13] One is the study by Nides,[14] which has already been discussed (Section 4.1.2).

Exclusion of the Oncken study

No reason is given for the exclusion of the Phase II RCT published by Oncken and colleagues (Probably either A3051007 or A3051008),[17] but it may relate to the study being a titration study (although this was not stated as an exclusion criterion). The exclusion of this study from the manufacturer's meta-analysis *would have* produced conservative rather than optimistic results, reduced statistical heterogeneity and kept the ICER high. However, the McMaster team's indirect comparison of varenicline and NRT (which the manufacturer used) was, in fact, informed by the McMaster team's meta-analysis, which included this study. *The effect of McMaster's inclusion of this study in their meta-analysis is to produce optimistic rather than conservative results, an increase in statistical heterogeneity and a reduction in the ICER* (see also this document, Section 4.1.2 on the Nides study).

Exclusion of the Reeves study[18]

The phase III RCT published by Reeves and colleagues (A3051037) compares varenicline with placebo.[18] The study may have been excluded due to the duration of the treatment (52 weeks, versus the proposed durations of 12 or 24 weeks, for which see the manufacturer's submission, Section 1.8, pages 11-12). *The exclusion of this study from the manufacturer's meta-analysis produces conservative rather than optimistic results, reduces statistical heterogeneity and would keep the ICER high.*

Eligible studies omitted by the McMaster review

Although the McMaster team claimed to have included all 12-month trials with chemical confirmation of cessation,[10] they omitted several eligible studies present in the Cochrane reviews. Some were available only in abstract or unpublished form: Brown 2006, Ferry 1994, Gonzales 2001, Rigotti 2006, Selby 2003, Tashkin 2001, Evins 2006, SMK 2001, Cooper 2003. We are not clear about the explanation for the following apparently eligible exclusions: Fee 1982, Jarvis 1982, Llivinia 1988, Richmond 1990, Abelin 1989, Ehrsam 1991, Hurt 1990, ICRF 1994 (all in the Cochrane NRT review[11]), Gorecka 2003, Simon 2004 and Evins 2006 (all in the Cochrane bupropion review[12]). All of these studies have 12 month follow-up and chemical confirmation. The McMaster team has also included a trial attributed to 'Fowler'. The Fowler study does not appear either in the McMaster team's reference list or in the extra files on the web version of the publication. There is no apparent pattern to the size or direction of the effect sizes found by these studies. *Whilst their exclusion will change the central estimate of effect in the McMaster team's meta-analysis, which will, in turn affect the indirect comparison, we do not believe it will considerably affect the ICER*.

4.1.5 Description and critique of manufacturers approach to validity assessment

Description

The manufacturer's submission reported the approaches of their included studies to the following elements associated with the reduction of bias in their controlled trials (varenicline trials only: Table 12, p 48): allocation concealment; randomisation technique; justification of sample size; adequate follow-up; blindness of outcome assessors; baseline comparability of study groups; appropriateness of statistical analysis; intention-to-treat analysis; comparability of dose to SPC recommendation; and, potential confounders.[1]

Blinding

Although, the manufacturers sponsored the remaining four studies they give a clean bill of health to only three, taking time to give a rather lengthy critique of the unpublished, commercial-in-confidence A3051044 study for its lack of blinding. The manufacturer's attempts to dismiss their own clinical trial on the grounds that the trial was open label are questionable (manufacturer' submission, pages 16 and 48). The purpose of blinding is to reduce the possibility of performance bias: systematic differences in the care provided to the participants in the comparison groups other than the intervention under investigation. Although, the effect size conferred by the varenicline could, feasibly, have been diluted by control arm participants accessing varenicline as well as / instead of NRT, it is more normal to think of unblinded studies exaggerating the effect size of the experimental intervention (by 17% in the empirical study by Schulz[20]) not the control intervention. Readers who have access to the commercial-in-confidence elements of this report should refer to material in Sections 4.1.7, 4.2.1 (Table 5) and 4.2.2 when interpreting the manufacturer's dismissal of this trial.

Loss to follow-up

The most important deficit in the manufacturer's 'critical appraisal' is the complete absence of any serious discussion of loss of trial participants to follow-up. As Table 1 (below) shows, the Gonzales[15] and Jorenby[16] trials, which are supposed to represent the sole basis for the manufacturer's estimates of clinical effectiveness (but see this document 4.1.2 and 4.1.4), were unable to include 43% and 35% of participants in their final analysis. There are three points to note here. First, the manufacturer has, correctly, presented quasi-intention-to-treat analyses, which assume that anyone lost to follow-up is still a smoker. Assuming that people lost to follow-up are smokers will ensure that actual quit rates are conservative, however it may not necessarily lead to conservative relative treatment effects (odds ratios), if loss to follow up is higher in the control group.[21] With this in mind, the Cochrane team conducted a sensitivity analysis to test the effect of including all randomized participants in the treatment group (quasi-ITT analysis) versus

only those who had follow-up data in the control group (available case analysis).[13] This has the effect of maintaining a conservative quit rate in the treatment group, but a more optimistic one in the control group. The clinical effect, favouring varenicline over placebo, remained statistically significant under these relatively extreme assumptions about differential distribution of missing data in treatment and control groups.

Table 2 Participants remaining at the end of the trial (from Cahill[13])

Study	Placebo [%]	Varenicline [%]	bupropion [%]	X2 and P value
Gonzales 2006	187/344 [54.4]	213/352 [60.5]	184/329 [55.9]	2.90, P=0.23
Jorenby 2006	204/341 [59.8]	240/344 [69.8]	221/342 [64.6]	7.42, P=0.02*
Nides 2006	68/127 [53.5]	77/127 [60.6]	68/128 [53.1]	1.83, P=0.40
Oncken 2006	40/129 [31.0]	146/253 [57.7]		24.32, P=0.0000008**
Reeves 2006	59/126 [46.8]	135/251 [53.8]		1.62, P=0.20
Tonstad 2006	463/607 [76.3]	494/603 [81.9]		5.83, P=0.016*

The second point is that treatment discontinuations were higher in the placebo group in all varenicline trials reported by the Cochrane review and this difference was statistically significant in the Jorenby,[16] Oncken[17] and Tonstad[19] trials. Systematic differences between comparison groups in the loss of participants from a study may result in attrition bias. The final point is the credibility of the Gonzales[15] and Jorenby[16] results given the proportion of participants (over one third in each case) they have lost to follow-up without reasonable explanation. Here there are two rules of thumb used by epidemiologists. One is that, while no loss to follow-up is desirable, those of less than 5% are considered relatively unimportant, whereas those of greater than 20% are considered extremely serious. The other rule is that, if the event rate is less than the loss to follow-up the validity must be considered suspect. Although neither rule of thumb reflects well on the included varenicline trials, it is worth noting that high attrition rates are considered a fact of life in some clinical settings more than others (psychiatric interventions being a well-known example). It may be that high loss-to-follow-up is common in or typical of smoking cessation studies.

4.1.6 Description and critique of manufacturers outcome selection

The manufacturer has used 12-month continuous abstinence rates in accordance with the decision problem.

4.1.7 Describe and critique the statistical approach used

The manufacturer's submission

Description

For all clinical endpoints, the manufacturer's submission presents absolute event numbers (except on one, unpublished study, marked commercial in confidence), rates, odds ratios and 95% confidence intervals with associated p-values. For each endpoint they present two analyses: one, which they call an 'intention-to-treat' analysis, assumes that those lost to follow-up are still smokers; the other, which they call a 'protocol intention-to-treat' analysis analyses only those who received at least one dose of study medication (manufacturer's submission, Section 5.4).[1]

Presentation of outcome data

Despite NICE's instructions that results should be pooled for "relative risk reduction and absolute risk reduction using both the fixed effects and random effects models" (rubric, Section 5.5), the manufacturers have presented only odds ratios, using a random effects model. The use of odds ratios rather than relative risks is, however, consistent with the majority of research in the field and is used by the Cochrane Tobacco Addiction Group (see the TAG's rationale, Appendix 2).

The indirect comparison

The manufacturer has used an indirect comparison, undertaken by the McMaster group (using methods described by Bucher[22]), to derive an estimate of the relative clinical effect of varenicline compared to NRT.[10] Criticisms specific to this work and its impact on the ICER are given below. Although a direct comparison exists (A3051044), the manufacturer has chosen to indirectly compare varenicline with NRT.

NICE's own rubric is explicit on this matter: "The Institute has a strong preference for evidence from 'head-to-head' randomised controlled trials (RCTs) that directly compare the technology and the appropriate comparator(s). Wherever such evidence is available, and includes relevant outcome evidence, this is preferred over evidence obtained from other study designs. *Where no head–to-head RCTs are available*, consideration will be given to indirect comparisons, subject to careful and fully described analysis and interpretation" (manufacturer's submission, Section 5, p26; emphasis our own).[1]

[redacted]

It is worth re-iterating at this point that, whilst the manufacturer's submission states that the estimates of the effectiveness of varenicline are derived from a synthesis of only two trials, Jorenby[16] and Gonzales,[15] the 12 month quit rates for NRT are derived from an indirect comparison, itself based on the McMaster review's meta-analysis,[10] which also pooled studies by Nides[14] and Oncken.[17] *The effect is to produce optimistic rather than conservative results and is likely to reduce the ICER* (see this document, Sections 4.1.2 and 4.1.4).

The McMaster review

Uncontrolled studies and concomitant therapies

The McMaster review combines the results of placebo-controlled studies with studies where the comparator was not placebo-controlled or where the bupropion and placebo were given in addition to another active treatment (e.g. bupropion + NRT versus NRT +/- placebo: see Section 4.1.3).[10] Arguably, this contaminates the indirect comparison, the consequences being unknowable, but definitely optimistic with regard to the efficacy of varenicline (which is only compared, on its own, to placebo). This is because the effect size of NRT versus non-placebo 'control' (OR 1.51, 95% CI, 1.25-1.84) is 29% less than that of NRT versus placebo (OR 1.80, 95% CI 1.61-2.01, Figure 1), making

the combination of all studies (OR 1.72, 95% CI, 1.57-1.90) 8% less than it would have been if NRT alone had been compared to placebo using the McMaster team's own data.

Bupropion is actually worse than non-placebo 'control' when combined with another active intervention (OR 0.61, 95% CI 0.23-1.58), whereas bupropion is superior to placebo (OR 1.76, 95% CI, 1.25-2.50, Figure 2), making the combination of all studies (OR 1.57, 95% CI, 1.11-2.22) 19% less than it would have been had bupropion alone been compared to placebo alone. These differences were not highlighted either by the McMaster review[10] or the Cochrane reviews by Hughes[12] and Silagy.[11]

Figure 1 NRT versus 'control' (data from McMaster team)

Review: Varenicline for smoking cessation
Comparison: 09 NRT versus control (Wu review)
Outcome: 01 Continuous cessation at 12 months

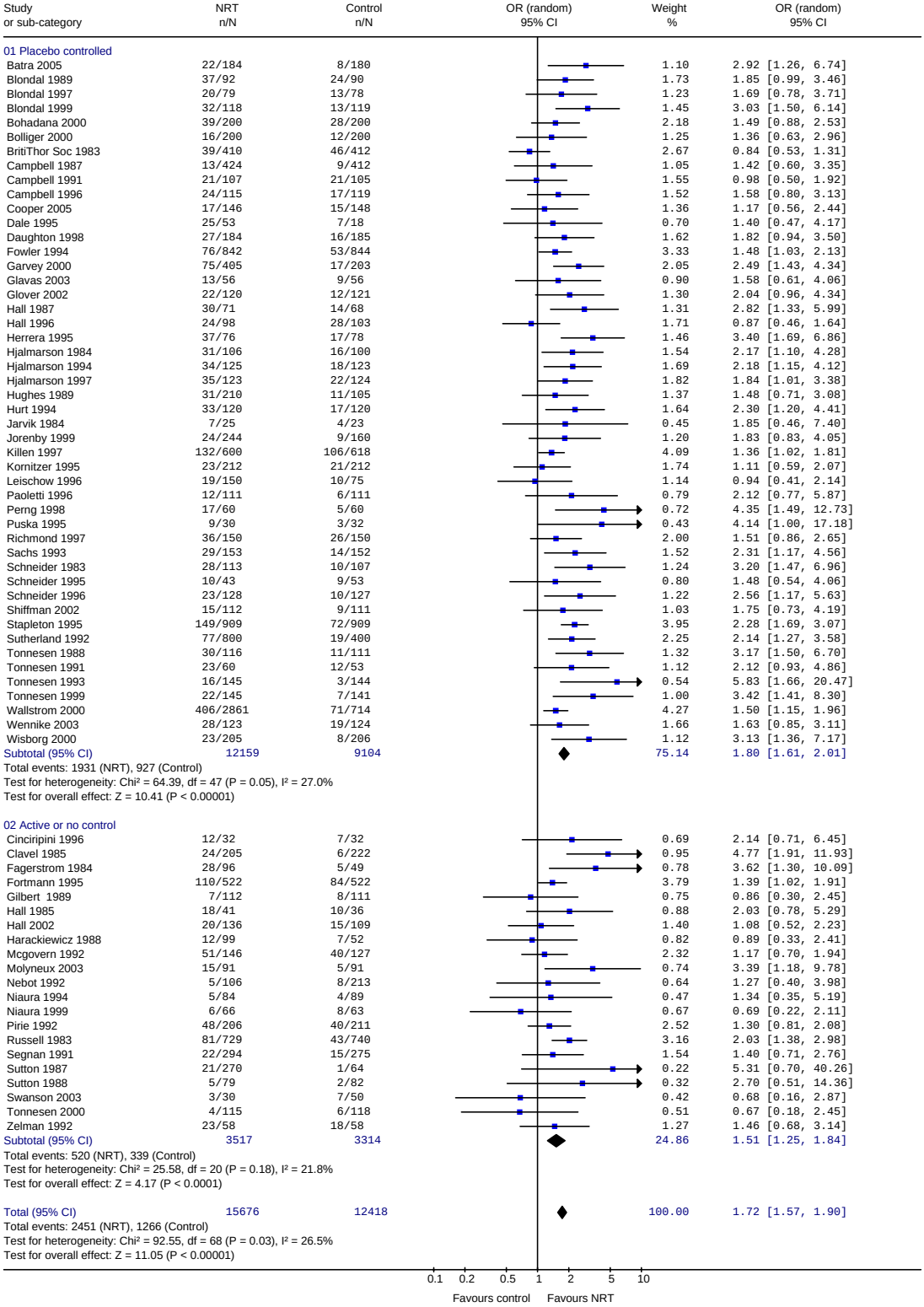
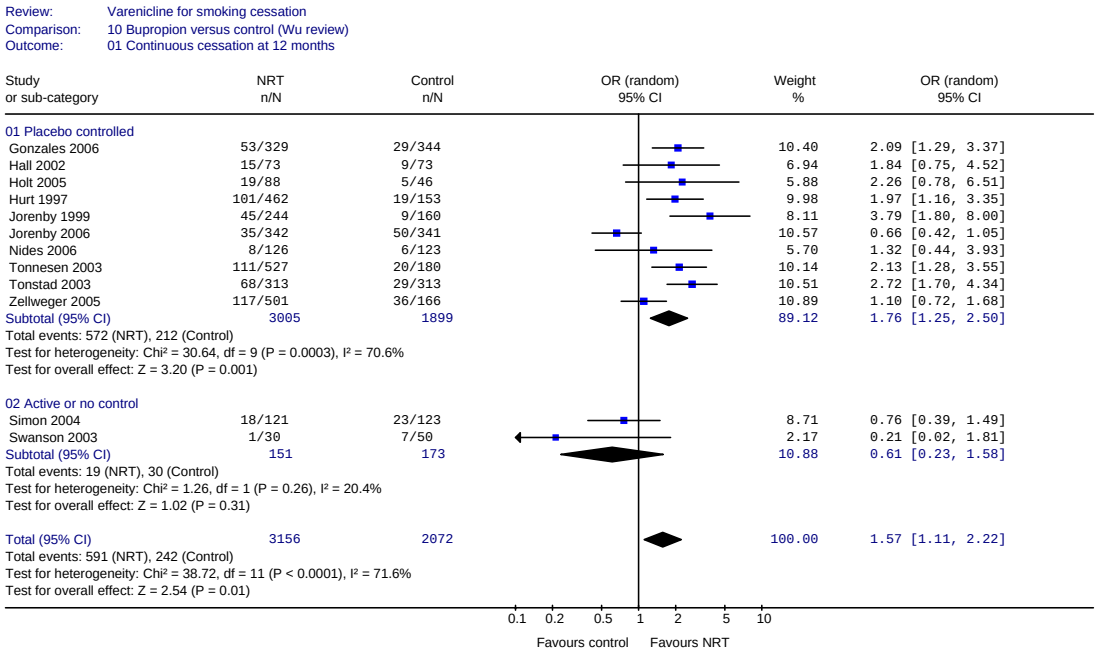


Figure 2 Bupropion versus 'control' (data from McMaster team)



Note that the central estimates of our recreation of the McMaster analysis (this document, Figures 1 and 2 above) differ by one percentage points to the published estimates, presumably due to the different software packages used.[10]

4.1.8 Summary statement

Describe the completeness of the submission with regard to relevant studies and relevant data within those studies. Reference should also me made concerning the extent to which the submitted evidence reflects the decision problem defined in the submission.

- The submitted evidence broadly reflects the 'decision problem' defined in the submission, in that: (1) the scope is the same; but, (2) the ERG believes that some studies may have been inappropriately included and excluded (Section 4.1.3) but are unable to establish this definitively, or assess the consequences for the ICER.

- The presentation of the clinical evidence in the manufacturer's submission is largely based on a systematic review of varenicline, bupropion and NRT which the manufacturer has commissioned ('the McMaster review'[10]).

- The McMaster review differs from Cochrane reviews on NRT,[11] bupropion[12] and varenicline[13] in the following respects:

    - The McMaster meta-analyses contain five trials either excluded or otherwise omitted by the equivalent Cochrane meta-analyses.

    - The Cochrane meta-analyses contain at least 20 trials (at least 10 in the bupropion review,[12] and 10 in the NRT review[11]) which are followed up for 12 months, have chemical validation and support, but which are not included in the McMaster meta-analyses.

    - The McMaster review includes studies with placebo control with studies which compare the addition of one pharmacotherapy to another or where the comparator is no treatment.

- The McMaster review presents direct evidence for the superiority of varenicline to bupropion and placebo, but does not include unpublished data from Pfizer's trial directly comparing varenicline to NRT.

4.2 Summary of submitted evidence

4.2.1 Summary of results

Tables 2 to 4 re-present the manufacturer's comparison of varenicline with placebo, varenicline with bupropion and varenicline with NRT. For clarity we have also presented relative risks, risk differences and numbers-needed-to-treat (NNTs).

Although other studies compare varenicline with placebo and with bupropion, (and these studies are included in the McMaster[10] and Cochrane[13] reviews), the studies presented here are the only studies which conform to the licensed indication and the outcomes specified in NICE's scope.[2]

Note that while the manufacturer's submission focuses on the two studies by Gonzales[15] and Jorenby,[16] both the McMaster[10] and Cochrane[13] groups also pooled two further studies (Nides[14] and Oncken[17]).

Table 3 Varenicline versus Placebo 12 month continuous quit rate

Study	Varenicline Non-smokers	Varenicline Randomized	Placebo Non-smokers	Placebo Randomized	Odds Ratio	Relative Risk	Risk Difference	Number Needed to Treat
A3051028 Gonzales	77 (21.9)	352	29 (8.4)	344	3.04 (1.93,4.80)	2.59 (1.74,3.87)	0.13 (0.08,0.19)	8
A3051036 Jorenby	79 (23.0)	344	35 (10.3)	341	2.61 (1.69,4.01)	2.24 (1.55,3.24)	0.13 (0.07,0.18)	8
Gonzales + Jorenby only pooled					Fixed: 2.81 (2.05,3.84) Random: 2.80 (2.05,3.83)	Fixed: 2.40 (1.83,3.15) Random: 2.39 (1.83,3.14)	Fixed and Random: 0.13 (0.09,0.17)	8
Cochrane Team's meta-analysis used for indirect comparison in manfacturer's submission (includes Nides and Oncken studies)					Fixed: 3.22 (2.43,4.27) Random: 3.19 (2.28,4.46)	Fixed: 2.74 (2.13,3.52) Random: 2.70 (1.98,3.68)	Fixed: 0.14 (0.11,0.17) Random: 0.14 (0.10,0.17)	7
McMaster Team's meta-analysis used for indirect comparison in manfacturer's submission (includes Nides and Oncken studies, but using different figures to Cahill's Cochrane team). Note that the ERG has acquired McMaster's event numbers and rerun their analyses with slightly different results: their original odds ratio (random effects model) was 2.96 (2.12 to 4.12)					Fixed: 3.06 (2.31,4.04) Random: 2.97 (2.12,4.16)	Fixed: 2.63 (2.04,3.38) Random: 2.54 (1.89,3.40)	Fixed: 0.12 (0.10,0.15) Random: 0.12 (0.06,0.17)	8

Table 4 Varenicline versus Bupropion 12 month continuous quit rate

Study	Varenicline Non-smokers	Varenicline Randomized	bupropion Non-smokers	bupropion Randomized	Odds Ratio	Relative Risk	Risk Difference	Number Needed to Treat
A3051028 Gonzales	77 (21.9)	352	53 (16.1)	329	1.46 (0.99,2.15)	1.36 (0.99,1.86)	0.06 (0.00,0.12)	17
A3051036 Jorenby	79 (23.0)	344	50 (14.6)	342	1.74 (1.18,2.57)	1.57 (1.14,2.17)	0.08 (0.03,0.14)	13
Pooled					Fixed: 1.59 (1.21,2.10) Random: 1.59 (1.21,2.10)	Fixed: 1.46 (1.17,1.83) Random: 1.46 (1.16,1.83)	Fixed and Random: 0.07 (0.03,0.11)	15

Table 5 Varenicline versus NRT 12 month continuous quit rate

Study	Varenicline Non-smokers	Varenicline Randomized	NRT Non-smokers	NRT Randomized	Odds Ratio	Relative Risk	Risk Difference	Number Needed to Treat
A3051044 Unpublished CIC trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
McMaster indirect comparison using pooled data					1.66 (1.17,2.36)	-	-	-

In terms of treatment harms, Cahill's Cochrane review states the following:

> "The predominant adverse effect for varenicline was nausea, reported at around 29% in Gonzales 2006 and Jorenby 2006, and at 40% in Reeves 2006, with attributable discontinuation rates from 2.5%to 7.6%. Both the Phase 2 trials found a dose-response relationship for the incidence of nausea: rates ranged from 17.5% (0.3 mg daily) to 52% (1.0 mg twice daily) in Nides 2006, while both the presence of titration and dosage levels affected the incidence and severity of nausea in Oncken 2006. Other adverse effects of varenicline across all six trials included insomnia, headache and abnormal dreams. In the two Phase 3 cessation trials, an average of 9.5% in the varenicline groups discontinued treatment but remained in the trial for follow up, compared with an average of 14% in the bupropion groups and 8% in the placebo groups. Discontinuation rates for any adverse effect were higher in Reeves 2006, at 28.3% in the varenicline group, and 10.3% in the control group."[13]

There were no treatment-related deaths in any of the intervention groups during treatment or follow-up phases. The Cochrane review identified non-fatal serious adverse events (SAEs) in six trials.[13,15,16,14,17,19] Only four individual SAEs were attributed to varenicline: atrial fibrilliation;[15] chest pain;[16] transient ischaemic attacks;[14] and, bilateral subcapsular cataracts.[18]

4.2.2 Critique of submitted evidence syntheses

The manufacturer's submission[1] is largely based on a systematic review of the RCTs evaluating varenicline, bupropion and NRT, which they have commissioned. The McMaster review is largely well designed and conducted although there may be inappropriate inclusions and exclusions (see this document, Sections 4.1.3 to 4.1.4). *The impact of this on the ICER is unknown*.

[redacted]

The selection of bupropion and NRT studies used in the indirect comparison may serve to exaggerate the treatment effect of varenicline when indirectly compared to other pharmacotherapies (see this document, Sections 4.1.3 and 4.1.7). The McMaster review pools the results from studies in which just one pharmacotherapy is compared with placebo (the type of studies the ERG believes are the least biased) with studies which compare the addition of one pharmacotherapy to a second. This will increase the effect size of varenicline in the indirect comparisons presented by the McMaster review. *This would decrease the ICER, making varenicline appear more cost-effective*.

The manufacturers claim that the studies published by Nides[14] and Oncken[17] were excluded from their analysis, but the McMaster meta-analysis pooled these studies with those by Jorenby[16] and Gonzalez[15] (see this document, Sections 4.1.2 and 4.1.4). The results of that meta-analysis were used for the indirect treatment comparison, the results of which were, in turn, used for the manufacturer's economic submission.[1] In the McMaster team's meta-analysis, the odds of smoking cessation at 12-months are 196% greater in the varenicline group than in the placebo group (OR 2.96, 95% CI 2.12 to 4.12). By just pooling the results of the Gonzalez[15] and Jorenby[16] trials (as the manufacturers claimed they had done), the odds of smoking cessation at 12 months are 180% greater in the varenicline group than in the placebo group (OR 2.80, 95% CI 2.05 to 3.83). The addition of the Nides[14] and Oncken[17] exaggerates the treatment effect by 16% (See this document, Sections 7.1 and 7.2 for further work by the ERG). *This would decrease the ICER, making varenicline appear more cost-effective*.

4.2.3 Summary

Does the submission contain an unbiased estimate of the technologies (relative and absolute) treatment effect in relation to relevant outcomes and the comparators of interest?

- The manufacturer's submission[1] is largely based on a systematic review of the RCTs evaluating varenicline, bupropion and NRT, which they have commissioned: the McMaster review. While that review is largely well

designed and conducted, certain studies may be inappropriately included or excluded with unknown effects on the ICER.

- The McMaster review also provides indirect treatment comparisons, which the manufacturer has used in its submission,[1] despite a direct comparison being available.

- The selection of studies used in the McMaster group's meta-analyses which, in turn, inform their indirect comparisons (the bases for the manufacturer's economic model), provides an optimistic basis for the assessment of varenicline's treatment effect because:

    - It allows the inclusion of phase II varenicline studies excluded by the manufacturer's submission and which improve the varenicline effect size;

    - It allows the inclusion of studies where bupropion or placebo are given with other active therapies (diluting the treatment effect).

5 ECONOMIC EVALUATION

5.1 Overview of manufacturer's economic evaluation

Summary of key points on the economic model

The cost-effectiveness results presented in the sponsor submission to NICE[1] should be interpreted in light of the following key methodological issues (these are presented in more detail in subsequent sections of this report):

- The model assumes only a single quit attempt using a single smoking cessation intervention. This may not be externally valid, as smokers may attempt to quit more than once using several smoking cessation technologies.
- The Benefits of Smoking Cessation on Outcomes (BENESCO) model is based upon a previously published smoking cessation model (the Health-Economic Model for Smoking-Related Morbidity or HECOS model).[23] Smoking-related morbidity health states differ between these models.
- The probability of short-term relapse to smoking is modelled using 1-year pooled quit rates from clinical trials of varenicline[15,16,19] and an indirect comparison.[10] Beyond 1-year, the annual probability of relapse to smoking is assumed to be independent of smoking cessation intervention, hence short-term benefits are assumed to be sustained in the long-term. Shorter time horizons may be subject to less uncertainty, but may underestimate the benefits of varenicline. Longer time horizons provide more favourable cost-effectiveness estimates for varenicline, yet are subject to a greater degree of uncertainty.
- The annual probabilities of relapse beyond 1-year of abstinence from smoking are sourced from US studies which may not reflect the smoking/abstinence behaviour of the population of England and Wales.
- Health utility scores used in the model do not consistently adhere to NICE's Reference Case.

- Methods for identifying and selecting costs and health utilities associated with morbidities are not reported or justified within the sponsor submission.[1]
- The presence of several computational errors should be borne in mind when considering cost-effectiveness results reported within the sponsor submission.
- The sensitivity analysis presented within the submission is very narrow and underestimates the true uncertainty surrounding the incremental cost-effectiveness of varenicline.
- The external validity of the model has not been demonstrated by the sponsor.

Scope of the economic evaluation of varenicline

The health economic evaluation presented within the sponsor submission to NICE[1] presents estimates of the incremental cost-effectiveness of varenicline as compared to other smoking cessation interventions which are routinely available on the NHS in the UK. Two health economic models are presented within the submission.

- The first model estimates the incremental cost-effectiveness of the standard regimen of varenicline as compared to bupropion, NRT, and placebo at the initial quit attempt (Group 1).
- The second model estimates the incremental cost-effectiveness of varenicline as compared to placebo for a population who have remained abstinent at the end of a 12-week course of varenicline (Group 2).

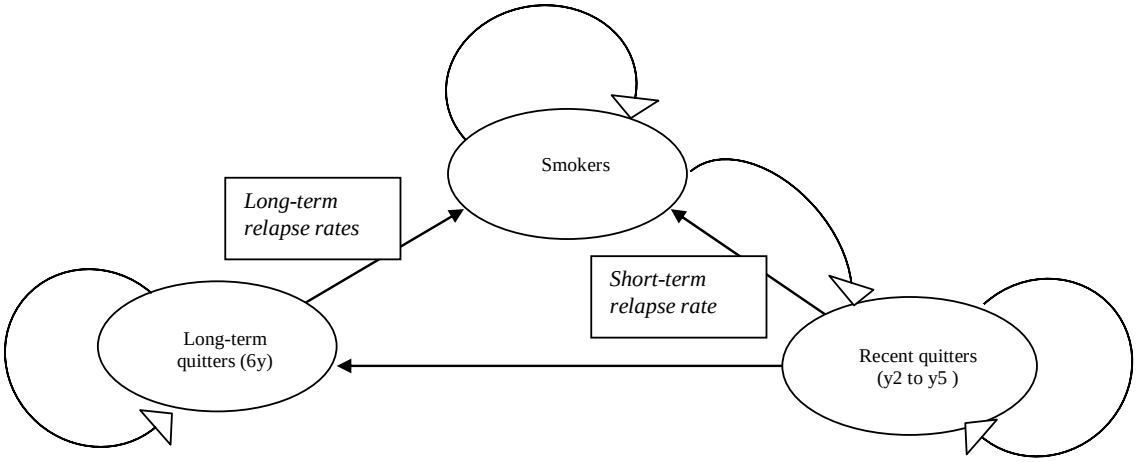
Counselling was also specified as a comparator for the analysis by NICE,[2] however this has not been included in the sponsor submission.[1] Both models are capable of estimating costs and health outcomes for individuals attempting unaided cessation without intervention, although results for this smoking cessation strategy are not presented in the submission. Both models employ similar structural and parametric assumptions; the key difference between the two models concerns the efficacy rates assumed for varenicline

and the comparator therapies. The primary health economic outcome for the evaluation is the incremental cost per quality adjusted life year (QALY) gained; the model also estimates the incremental cost per life year gained (LYG) and the incremental/net cost per quitter. The economic evaluation was undertaken from the perspective of the NHS only, as the sponsor states that the quantification of PSS resources relevant to smoking was not possible.[1] In the base case analysis, cost-effectiveness is evaluated over a lifetime horizon using an annual cycle length.

Overview of model structure

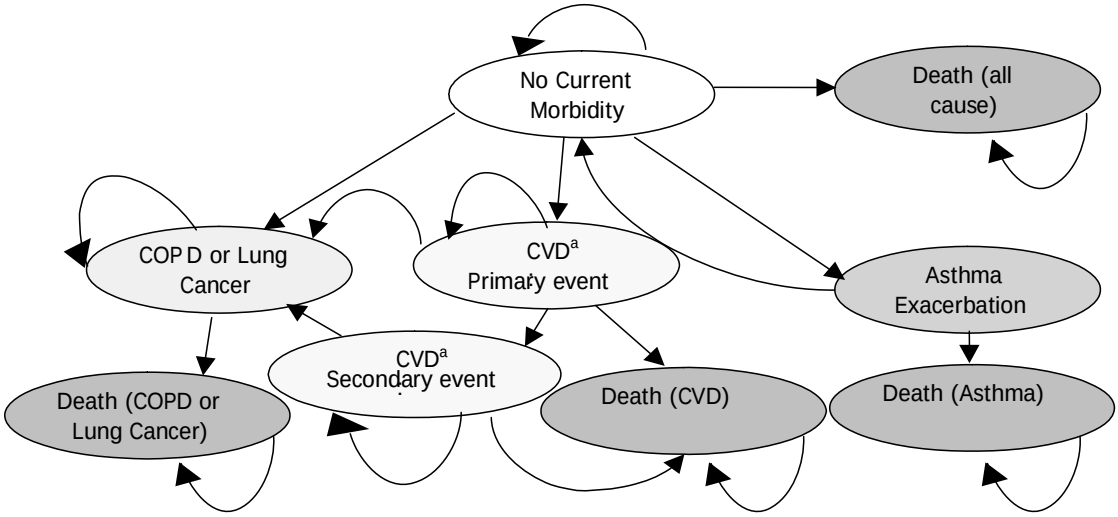
The varenicline model (hereafter referred to as the BENESCO model) is based upon an earlier smoking cessation model (the Health-Economic Model for Smoking-Related Morbidity, or 'HECOS' model) reported by Orme et al.[23] The model uses the state transition methodology to simulate the experience of individuals following an initial attempt to quit smoking. The model is complex, and a large number of health states are used to describe smoking status, acute and chronic morbidity, and mortality. All individuals enter the simulation as "quitters" and are initially distributed across the disease health states according to age- and sex-specific estimates of morbidity prevalence amongst the general population. The probability of relapse to smoking is assumed to follow a step-wise pattern of decreasing risk over time. During the first year following the initial quit attempt, the probability of relapse is modelled using pooled efficacy rates from individual arms of the clinical trials of varenicline and NRT (See sponsor submission, Table 41[1]). Following the first year of abstinence, the subsequent risk of relapse to smoking is modelled using data from two longitudinal studies of long-term abstinence from smoking undertaken in the US.[24,25] The relative risks of developing morbidities and mortality smoking are assumed to decrease according to time since smoking cessation based on two US prospective cohort studies (See sponsor submission, Appendix 1).[26,27] The methods used to derive these relative risks of morbidity and mortality is not reported within the sponsor submission. Figure 3 presents a simple schematic describing transitions between different states of smoking status.

Figure 3 Schematic describing overall BENESCO model structure[1]



The BENESCO model includes five morbidities which are related to smoking: chronic obstructive pulmonary disorder (COPD), lung cancer, coronary heart disease (CHD) events, asthma, and stroke. These morbidities were included in the model as they are reported by the sponsor to account for the greatest mortality, morbidity and cost associated with smoking. The original HECOS model reported by Orme et al[23] also included low birth weight pregnancy as a smoking-related morbidity although this has been excluded from the BENESCO model. Justification for the exclusion of this specific morbidity from the BENESCO model is not provided within the submission.[1] The BENESCO model has been structured such that subjects may experience one or more asthma exacerbations, CHD events or stroke events. Following the onset of chronic morbidity (COPD or lung cancer), the costs and HRQoL impact of subsequent acute morbidities are ignored for the remainder of the model as they are assumed to be superseded by chronic morbidities. The onset of smoking-related morbidities is associated with an increased risk of mortality. All Markov states are treated as mutually exclusive. Figure 4 shows the possible transitions through morbidity health states in the model.

Figure 4 Possible transitions through smoking-related morbidity states in the BENESCO model[1]



Costs and adjustments for health-related quality of life are applied to the number of subjects residing in each health state during each model cycle. These costs and outcomes are modelled using a half-cycle correction. Total costs and total QALYs gained for each smoking cessation intervention are estimated by summing the costs and QALYs accrued during each cycle over the entire model time horizon.

List of key model assumptions

A comprehensive list of model assumptions is presented on pages 98-100 of the sponsor submission.[1] The key assumptions underpinning the BENESCO model are summarised below.

Assumptions concerning morbidity

1. All subjects enter the model as smokers who are attempting to quit. The baseline prevalence of acute and chronic morbidities is assumed to be independent of the individual's intention quit smoking.

2. All health states are mutually exclusive, hence a subject cannot have two morbidities at the same time, although they may have a chronic morbidity with a previous history of one or more acute morbidities.
3. Transitions to health states are assumed to be independent of time spent in previous states (the Markovian assumption).
4. The costs and health outcomes associated with the onset of chronic morbidities supersede future costs and effects of acute morbidities.
5. Transitions between CHD and stroke health states are not possible.

Assumptions concerning the relationship between age and smoking duration

1. Age is used as a proxy to represent smoking duration (three age bands are modelled: 18-34 years, 35-64 years, > 65 years based on Thun et al[27]).
2. The risk of smoking-related morbidity varies with age and smoking status.

Assumptions concerning the relationship between age and morbidity and mortality

1. The incidence of smoking-related morbidity varies by age. The proportions of morbidity events are modelled using constant rates over time based on national population data.
2. All cause mortality is based on national mortality statistics.[28] For subjects aged 85 years and older, the all cause death rate was taken to be the midpoint between 1.00 (all subjects die) and the > 65 mortality rates. All model subjects who remain alive at age 99 are assumed to be absorbed into the mortality health state during the final model cycle.

Assumptions concerning the effect of smoking on morbidity

1. The model assumes zero morbidity or mortality in the 18-34 age-group except for asthma exacerbations.
2. The relative risk of developing a smoking-related morbidity for a subject who fails to abstain from smoking for at least one-year is assumed to be the same as that of a current smoker.

Assumptions concerning chronic morbidity

1. The risk of developing COPD, lung cancer, CHD and stroke increases with age irrespective of smoking status.
2. For former smokers, the relative risk of developing smoking related-morbidities is reduced over time compared to current smokers of the same age.
3. The number of deaths resulting from an asthma attack is assumed to be small and accounted for in the all cause mortality rate.
4. Disease events are assumed to be mutually exclusive. For example, a subject cannot acquire both CHD and lung cancer within same cycle.

Assumptions concerning health-related quality of life

Health utility values are assumed to be the same for smokers and non-smokers.

5.1.1 Natural history

The distribution of patients in the morbidity states over time is determined by prevalence, incidence and mortality parameters of these diseases estimated from different studies and/or databases. The parameters differ according to age group, sex and smoking status. The model considers the following health states for smokers and quitters groups:

- no smoking-related morbidity
- COPD
- lung cancer
- first non-fatal CHD event in this year
- recurrent non-fatal CHD event in this year
- non-fatal stroke event in this year
- previous stroke event
- asthma exacerbation
- death (all causes)
- death (COPD)
- death (lung cancer)
- death (post first CHD event)
- death (post recurrent CHD event)
- death (post stroke event)
- death (post secondary stroke event)
- death (asthma exacerbation).

The risks in the model vary according to the age and gender, and the calculations are based on the inputted parameters of particular events with the exception of the 'no smoking-related morbidity' state. The probability of remaining in the latter state during the next cycle is calculated using the following formula:

1 - death (all causes) – COPD – lung cancer – (first non-fatal CHD event in this year – death (post first CHD event)) – (non-fatal stroke event in this year - death (post stroke event))

It is noteworthy that the probability of experiencing an asthma exacerbation is not included into the above formula (See computational errors described in Section 5.3). Also, the model structure allows for the inclusion of recurrent non-fatal CHD and previous stroke events but they are not incorporated into the analysis. It is also important to point out that the death rate in the calculation of the 'no smoking-related morbidity' probability is overstated given the fact that it includes death from *all* causes, yet the definition of the health state excludes individuals with smoking-related morbidities.

The risks for the morbidity and death states are estimated taking into account the effect of ageing. The model assumes that 18-34 age group is not associated with a risk of developing morbidity except for asthma exacerbation.

5.1.2 Treatment effectiveness within the submission

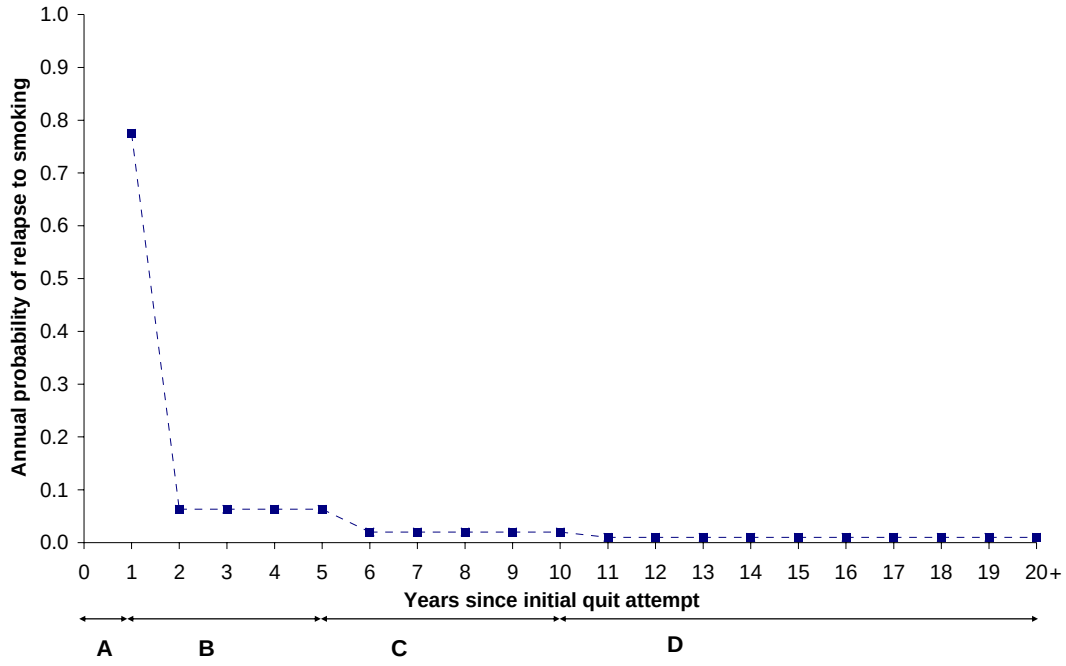
Treatment effectiveness is modelled using smoking cessation rates at 1-year sourced from the clinical trials of varenicline[15,16,19] and indirect analyses.[10] All smoking cessation intervention efficacy rates describe the probability of remaining abstinent from smoking at 1-year following the point of randomisation (See Table 6).

Table 6 One-year quit rates assumed within the health economic model

Intervention	1-year quit rate	Source
Model 1 – initial quit attempt		
Varenicline	22.5%	Pooled 1-year quit rate from studies A3051028[15] and A3051036[16]
Bupropion	15.5% (this is reported as 15.7% in the submission[1])	Pooled 1-year quit rate from studies A3051028[15] and A3051036[16]
NRT	14.9%	Indirect comparison based on work by the McMaster Group (Wu et al[10])
Placebo	9.4%	Pooled 1-year quit rate from studies A3051028[15] and A3051036[16]
Model 2 –sustained abstinence at the end of a 12-week course of varenicline		
Varenicline	43.6%	Study A3051035[19]
Placebo	36.9%	Study A3051035[19]

These quit rates are used to estimate the number of subjects who remain in the quitter health states at the end of the first Markov cycle (the end of the first year of the simulation). Following the first year, the probability of relapsing to smoking is described by annual relapse probabilities derived from Wetter et al and Krall et al.[24,25] Following the first model cycle, the annual probability of relapsing to smoking is common to all treatment strategies. As such, the model is underpinned by an assumption of sustained benefit over the remainder of the model time horizon. Figure 5 shows the annual step-wise probability of relapse to smoking assumed over the time horizon of the model.

Figure 5 Assumed annual probability of relapse to smoking



where A = 1-year trial quit rates for individual smoking cessation interventions (varenicline quit rate shown in figure)

B = Short-term relapse rate (Wetter et al)[25]

C = Medium-term relapse rates (Krall et al)[24]}

D = Long-term relapse rates (Krall et al)[24]

Owing to the absence of evidence concerning the long-term efficacy of varenicline, the appropriateness of the assumption of sustained benefit over the lifetime of the model cohort is unclear, and the extrapolation of 1-year abstinence outcomes over a subject's lifetime (up to 81 additional years) is highly uncertain.

For the standard treatment regimen model (Group 1 – initial quit attempt), efficacy rates for varenicline, bupropion and placebo were derived from a pooled analysis of two Phase III trials reported by Gonzales et al[15] and Jorenby et al.[16] Whilst direct trial evidence concerning the efficacy of varenicline versus NRT exists, the base case analysis presented in the submission instead uses an efficacy rate for NRT from a meta-analysis reported by the McMaster Group (Wu et al).[10] The submission attempts to

justify the substitution of indirect efficacy evidence in place of direct efficacy evidence for NRT stating that observed efficacy rates for varenicline and NRT were higher than those expected from previous varenicline trials and meta-analyses of NRT therapy.[1] As noted in Section 4.17 and following, the use of the indirect comparison estimate of relapse rate for NRT is unwarranted. Efficacy rates for varenicline and placebo in subjects who are abstinent at the end of a 12-week course of varenicline (See sponsor submission, Table 42) were derived from a Phase III trial reported by Tonstad et al.[19] It should be noted that the model does not use relative measures of treatment effectiveness such as odds ratios or relative risks to describe differential risks of relapse for each of the smoking cessation technologies versus a reference comparator. Instead, pooled cessation rates for individual treatment groups have been derived from the odds ratios reported in the clinical effectiveness chapter of the submission.[1] Methods for pooling efficacy estimates for common treatment arms across trials are poorly reported within the submission (see this document, Section 4 throughout) as are methods for deriving rates from pooled odds ratios (see this document, Section 6, question A4 below).

5.1.3 Health related quality of life

The number of QALYs gained for each smoking cessation intervention is estimated by multiplying the expected survival in each living health state by their respective utility scores over the model time horizon. Baseline HRQoL is assumed to be equivalent between smokers and non-smokers. Differential utility scores are applied to subjects without smoking-related morbidity according to age and sex. These baseline utilities were derived from a study by Fiscella et al;[29] utility weights by age group and sex are based on unweighted means of 5-year utility estimates for subjects who have quit smoking for 15 years or more. The use of these data is not ideal as these quality of life estimates are not fully preference-based,[29] and are intended to reflect the preferences of the US population. It is unclear why equivalent UK sources of preference data, such as the General Health Survey of England,[30] were not used to inform these baseline health utility values. Separate utility

scores are applied to subjects experiencing smoking-related morbidities. Health utilities for subjects with COPD are assumed to be constant over time. The health utility associated with lung cancer is assumed to be higher in the first year than subsequent years. Following initial stroke, subsequent stroke events are assumed to be associated with a lower utility score. CHD events are assumed to result in the same level of health utility irrespective of previous CHD events. In contrast to baseline health utilities for subjects without morbidity, utility scores for subjects experiencing morbidities were not differentiated by age or sex; this represents an inconsistency in the modelling approach. Table 7 shows the assumed utility values together with an outline of the methods and sources of preferences. Importantly, methods used to identify and select utility studies for inclusion in the BENESCO model are not presented within the sponsor submission.

Table 7 Health utility scores assumed within the varenicline cost-effectiveness model

Health utility score parameter	Utility score	Method used to elicit preferences	Source of utility estimate(s)	Source of preferences
Baseline utility according to age and sex	0.77-0.93	Healthy People 2000 years of healthy life (YHL) measure	Fiscella et al[29]	US
Lung cancer year 1	0.61	EQ-5D	Trippoli et al[31]	UK
Lung cancer subsequent years	0.51	EQ-5D	Trippoli et al[31]	UK
COPD	0.76	EQ-5D	Spencer et al[32]	UK
CHD first and subsequent event	0.76	Time trade-off	Hay et al[33]	US
Stroke first event	0.74	Meta-analysis of preference and non preference-based methods	Tengs et al[34]	Unclear
Stroke subsequent event	0.15	Time trade-off	Gage et al[35]	US
Asthma first event	0.52	EQ-5D	Szende et al[36]	UK

Adverse events and their impact upon a subject's HRQoL are not included in the economic evaluation. Whilst serious adverse events were rare and were not reported to have been statistically significant within the clinical trials of varenicline,[15,16,19] the omission of their impact upon health utility from the economic model may represent a minor bias in favour of the smoking cessation interventions.

5.1.4 Resources and costs

The health economic model includes two groups of costs: the short-term costs associated with the smoking cessation interventions, and the costs of managing smoking-related morbidities. The former set of costs is applied to the entire cohort of individuals attempting to quit smoking at the start of the simulation. The latter set of costs is applied to each year spent in the smoking-related morbidity health states. All costs in the model were updated to 2006 prices using inflation rates from the Personal Social Services Research Unit;[37] this is appropriate.

Acquisition costs for smoking cessation interventions

Table 8 shows the assumed acquisition costs for smoking cessation interventions included in the BENESCO model.

Table 8 Acquisition costs assumed within the health economic models

Smoking cessation intervention	Dosing schedule	Assumed cost per subject	Source of cost estimate
Varenicline (Groups 1 and 2)	0.5mg daily (3 days) 1.0mg daily (4 days) 2.0mg daily (77 days)	£165.66	BNF[38]
Bupropion	150mg daily (6 days) 300mg daily (63 days)	£81.56	BNF[38]
NRT	Not reported	£117.68	UK prescribing data (no reference provided within sponsor submission[1])

Within the BENESCO model, varenicline is reported to cost £165.66 per subject; this cost estimate was taken from the British National Formulary (BNF).[38] This cost estimate does not include the possibility of wastage. As varenicline would be prescribed either as a starter pack, which includes 0.5mg tablets and standard 1.0mg tablets, wastage would not be incurred unless a subject's dose is reduced due to adverse events. Dose reductions were observed in trials A3051028[15] and A3051036[16] for a small number of subjects although this was less than 5% of the intention-to-treat population.[1] These dose reductions for subjects receiving varenicline are not included in the model, but would have only a minor impact upon the cost-effectiveness of varenicline.

The cost of bupropion was assumed to be £81.56 per subject; this cost was derived from the BNF.[38] The sponsor submission assumes bupropion is given at a dose of 150mg once daily for 6 days then 150mg twice daily for a total of 9 weeks; this assumed treatment schedule would require 132 tablets; as the intervention is prescribed in 60-tab packs, each subject would require 3 packs (48 tablets would be wasted for each subject). Including this wastage, the cost

of bupropion should have been £119.55. Whilst underestimated within the model, bupropion acquisition costs are relatively minor in comparison to the costs associated with treating morbidities, and the ICER is unlikely to be substantially affected by this bias.

The cost of NRT is assumed to be £117.68 and is reported to be *"based on a basket for all NRT products prescribed in the UK at 2006 prescribing costs weighted basket of treatments."*[1] The methods and sources used to derive NRT cost estimates, and assumptions concerning the proportion of specific NRT interventions (gum, patches, inhalers, sublingual tablets) are not reported within the submission. The validity of this cost estimate is therefore questionable.

Costs associated with managing adverse events were not included in the BENESCO model. As noted above, serious adverse events were rare within the clinical trials, hence it is unlikely that this omission would represent a substantial bias in the cost-effectiveness results.

Costs of managing smoking-related morbidities

The submission includes estimates of the annual cost associated with each of the morbidities included in the model (See Table 9). The methods used to identify and appraise these cost estimates are not reported within the submission. In the absence of alternative cost estimates, it is difficult to gauge the validity of these parameter values.

Table 9 Annual costs of managing smoking-related morbidities

Cost	Annual cost of treating first event	Annual cost of treating first event	Data source
COPD	£819	N/A	Britton[39]
Lung cancer	£3,731	N/A	Parrott and Godfrey[40]
CHD	£980	£980	McMurray et al[41]
Stroke	£16,000	£16,000	Youman et al[42]
Asthma	£888	N/A	Hoskins et al[43]

5.1.5 Discounting

QALYs and costs were discounted at 3.5% per year within the BENESCO model; this is in line with NICE's Reference Case. Life years gained estimated within the model are not however subject to discounting.[1]

5.1.6 Sensitivity analyses

Simple parametric sensitivity analysis

Simple one-way and two-way sensitivity analysis was reported within the sponsor submission.[1] This sensitivity analysis explored the impact of varying parameter assumptions concerning discount rates for costs and health outcomes, exploring the impact of assuming the upper and lower confidence intervals of the placebo abstinence rates, and reducing the acquisition costs of NRT by 25% on the incremental cost-effectiveness of varenicline. The scope of this sensitivity analysis is very narrow, as no consideration is given to the impact of uncertainty surrounding the baseline and morbidity-specific utilities scores, costs of care, long-term smoking relapse rates or relative risks of developing morbidities.

Structural sensitivity analysis

The submission states that structural sensitivity analyses was not undertaken as *"the HECOS model was the most appropriate for measuring the cost-effectiveness of pharmacological therapies for smoking cessation."* This justification is inadequate, as the structure of the HECOS model[23] differs from BENESCO model in terms of which morbidities are included and how they are handled, assumptions concerning long-term risk of relapse, and the time horizon over which incremental costs and health outcomes are evaluated (See sponsor submission, Table 25). The assumption of sustained benefit following the first year of the simulation clearly indicates considerable uncertainty surrounding relapse rates beyond the first year of smoking abstinence. Despite such criticisms, the submission does in fact present cost-effectiveness results over a range of time horizons (2-, 5-, 10-, 20-years and lifetime); this analysis demonstrates the impact of long-term extrapolation on cost-effectiveness outcomes for the range of smoking cessation technologies.

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis was also undertaken within the sponsor submission.[1] However, this sensitivity analysis is not comprehensive, as probability distributions were used only to describe the uncertainty surrounding mean efficacy rates for smoking cessation interventions, mean treatment costs for morbidities and mean health utility scores. Other model parameters such as smoking relapse rates beyond the first year of smoking abstinence, the prevalence, incidence and severity of morbidities (particularly asthma), mortality estimates, and acquisition costs for NRT have been held constant at their mean values within the probabilistic sensitivity analysis. As a result, the true uncertainty surrounding the incremental cost-effectiveness of varenicline is likely to have been substantially underestimated within the submission.

In addition, for those parameters which are included in the probabilistic sensitivity analysis, the degree of uncertainty allowed for is narrow; standard errors surrounding mean morbidity treatment costs and utility values are

assumed to be +/-10% of the distributional parameter means. Furthermore, the selection of probability distributions is not justified in the submission. For the majority of uncertain model parameters, the choice of parametric distribution appears to be reasonable, however it is noteworthy that the use of beta distributions to describe health utility scores does not allow for states of health which are considered to be worse than death.

5.1.7 Model validation

The sponsor submission reports only a limited degree of model validation. Of the validation that is presented, this process was restricted only to tests of internal consistency to ensure that the model calculations performed as expected. The submission details seven verification checks that attempt to demonstrate that the model is internally logical and consistent. It should be noted that the Evidence Review Group identified several computational errors and inconsistencies within the submission models; these are discussed in more detail in the Critique of Approach section (See Section 5.3).

Importantly, the submission does not include any details of external validation against either relevant cohort studies of smoking cessation and relapse (for example Doll et al[44]), or against other the results of other mathematical models of smoking cessation interventions.[45,29,40] Whilst the submission claims that the majority of model assumptions are conservative, the external validity of the model has not been demonstrated within the sponsor submission.

5.2 Results included in manufacturer's submission

Central estimates of cost-effectiveness

The base case central estimates of cost-effectiveness for varenicline as compared to bupropion, NRT and placebo at the initial quit attempt are shown in Table 10. The sponsor submission reports costs and health outcomes on a national basis (although the sponsor submission does not state whether these relate to England and Wales or the UK). Presenting costs and health

outcomes on this basis provides no indication of the magnitude of treatment benefit or cost savings for individual subjects. Consequently, the ERG estimated per subject-level costs and health outcomes for each treatment option; these are shown in parentheses.

Table 10 Central estimates of cost-effectiveness assuming a lifetime horizon (Group 1 – standard varenicline treatment regimen)

Smoking cessation intervention	Costs	QALYs	Life years	Incremental cost-effectiveness results
Varenicline	£34,018,920,489 (£10,717)	42,135,027 (13.27)	86,711,276 (27.32)	Dominating
Bupropion	£34,347,878,880 (£10,820)	42,063,665 (13.25)	86,540,790 (27.26)	-
NRT	£34,514,466,202 (£10,873)	42,057,446 (13.25)	86,525,933 (27.26)	-
Placebo	£34,608,281,768 (£10,903)	42,001,477 (13.23)	86,392,224 (27.22)	-

The results suggest that the expected benefit resulting from the use of varenicline over the subject's lifetime is small. When compared to placebo (the least effective option), the incremental benefit of varenicline is estimated to be 0.0421 QALYs which equates to 15.4 quality adjusted days of survival. The model suggests that when compared against placebo, varenicline is expected to offer cost savings of around £185.66 over the lifetime of the subject. The equivalent cost-effectiveness results for subjects who have remained abstinent following a 12 week course of varenicline are shown in Table 11. The results for Group 2 again suggest that varenicline offers a small level of incremental benefit of 0.0215 QALYs and is cost savings compared to placebo in the lifetime model.

The base case lifetime analysis suggests that the standard varenicline treatment regimen (Group1) dominates all other smoking cessation interventions and placebo; that is, varenicline is expected to produce a greater number of QALYs than other smoking cessation interventions at a lower cost. Similarly, the analysis of subjects abstinent following 12 week course of varenicline (Group 2) suggests that varenicline is expected to dominate placebo.

Table 11 Central estimates of cost-effectiveness assuming a lifetime horizon (Group 2 – subjects abstinent following 12 week course of varenicline)

Smoking cessation intervention	Costs	QALYs gained	Life years gained	Incremental cost-effectiveness results
Varenicline	£32,222,646,438 (£10,151)	42,350,135 (13.34)	87,225,169 (27.48)	Dominating
Placebo	£32,267,166,299 (£10,165)	42,281,831 (13.32)	87,061,989 (27.43)	-

Subgroup analysis

Subgroup analysis was undertaken for the three age bands and gender individually for subjects attempting to quit smoking (Group 1 – standard varenicline treatment regimen). The results of this subgroup analysis are shown in Table 12. This subgroup analysis was not based upon adjusted estimates of efficacy derived from subgroup analyses within the varenicline trials, hence mean 1-year quit rates for each treatment group relate to the entire treatment group rather than the efficacy rates observed within the specified subgroups. Consequently, the results of this subgroup analysis should be interpreted with caution. This subgroup analysis was not repeated for Group 2 (subjects abstinent following 12 week course of varenicline).

Table 12 Results of subgroup analyses for the standard varenicline regimen (Group 1)

Smoking cessation intervention	Incremental results for subgroup				
	Age 18-34 subgroup	Age 34-64 subgroup	Age >64 subgroup	Male subgroup	Female subgroup
Varenicline	Dominant	Dominant	£12,426	Dominant	Dominant
Bupropion	-	-	Dominated by varenicline (extended)	-	-
NRT	-	-	Dominated by bupropion	-	-
Placebo	-	-	-	-	-

Sensitivity analysis results

The results of the one-way and two-way sensitivity analysis for the standard varenicline regimen (Group 1) over a lifetime horizon are summarised in Table 13. The results of the simple sensitivity analysis suggest that the cost-effectiveness of varenicline remains stable irrespective of assumptions concerning discount rates, baseline relapse risk for placebo, and costs for NRT. The analysis also suggests that varenicline is expected to dominate NRT irrespective of whether direct or indirect efficacy data are used.

Table 13 Simple sensitivity analysis results

Scenario	Sensitivity analysis results for Group 1 (standard varenicline treatment regimen)	Sensitivity analysis results for Group 2 (subjects abstinent following 12 week course of varenicline)
6% discount rate for costs, 1.5% discount rate for QALYs	Varenicline dominates all other smoking cessation interventions and placebo (analysis undertaken for lifetime horizon only)	ICER = £1,524 per QALY gained
Cost of NRT reduced to 25% of base case cost	Varenicline dominates all other smoking cessation interventions and placebo (analysis undertaken for lifetime horizon only)	Not applicable
Cost of NRT reduced to 0% of base case cost	Varenicline dominates NRT over lifetime horizon. ICER is less than £6,400 at 20-years. ICER is above £44,400 for 10-year horizons or shorter.	Not applicable
Lower 95% c.i. baseline risk for placebo	Varenicline produces more net benefit than other smoking cessation interventions (analysis undertaken for lifetime horizon only)	Not presented
Upper 95% c.i. baseline risk for placebo	Varenicline produces more net benefit than other smoking cessation interventions (analysis undertaken for lifetime horizon only)	Not presented
Efficacy rates taken from study A3051044 – open-label trial of varenicline versus NRT transdermal patch	Varenicline dominates NRT at 20-years and over subject lifetime. ICER is below £7,600 at 10-years. ICER is above £62,800 for horizons of 5-years or shorter.	Not applicable

Structural sensitivity analysis – results over different time horizons

Table 14 shows the incremental cost-effectiveness results over a range of different time horizons. Under the base case model assumptions, varenicline is estimated to dominate other smoking cessation interventions within 20-years. As NRT costs more and has a lower 1-year efficacy rate than bupropion, it is consistently dominated irrespective of the time horizon.

Table 14 Cost-effectiveness results over different time horizons (Group 1 – initial quit attempt)

Smoking cessation intervention	2-years	5-years	10-years	20-years	Lifetime
Varenicline	£745,046	£104,283	£18,564	Dominating	Dominating
Bupropion	£831,403	£117,843	£22,164	Dominated by varenicline	Dominated by varenicline
NRT	Dominated by bupropion	Dominated by bupropion	Dominated by bupropion	Dominated by bupropion and varenicline	Dominated by bupropion and varenicline
Placebo	-	-	-	-	

Table 15 shows the equivalent analysis of Group 2 (subjects abstinent following 12 week course of varenicline).

Table 15 Cost-effectiveness results over different time horizons (Group 2 – subjects abstinent following 12 week course of varenicline)

Smoking cessation intervention	2-years	5-years	10-years	20-years	Lifetime
Varenicline	£1,554,429	£231,370	£52,302	£8,735	Dominating
Placebo	-	-	-	-	-

The analysis suggests that the time horizon required in order for varenicline to dominate placebo is greater than 20 years.

Probabilistic sensitivity analysis

The results of the probabilistic sensitivity analysis are presented using cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs). The cost-effectiveness planes present the 95% confidence ellipses for marginal costs and QALYs as compared to placebo; these are presented only for Group 1 (subjects at the initial quit attempt). The CEACs are presented as both incremental and pairwise comparisons (only the former are appropriate); these are presented for both Groups 1 and 2.

The interpretation of the cost-effectiveness plane is problematic as multiple interventions are compared against placebo on the same plot. The cost-effectiveness plane suggests the following:

- There is a comparatively greater degree of uncertainty surrounding the incremental costs and QALYs associated with varenicline versus placebo, than for bupropion and NRT versus placebo.
- When compared marginally against placebo alone, varenicline, bupropion and NRT are expected to dominate placebo.
- There is a possibility that the true incremental cost-effectiveness of varenicline, bupropion, and NRT lies in any of the four quadrants (including a small probability that each of the three smoking cessation interventions is dominated by placebo). This phenomenon is likely to be

largely due to the use of independent efficacy rates from the clinical trials instead of modelling the relative efficacy of varenicline, bupropion and NRT against placebo using odds ratios or relative risks from the trials.[15,16,19]

Assuming a willingness-to-pay threshold of between £20,000 to £30,000 per QALY gained, the incremental CEACs suggest that the probability that varenicline produces more net benefit than bupropion, NRT and placebo for the standard varenicline treatment regimen is around 70% (Group 1). Assuming a willingness-to-pay threshold of between £20,000 to £30,000 per QALY gained, the CEACs suggest that the probability that giving varenicline to subjects who remain abstinent following a 12 week course of varenicline produces more net benefit than placebo is also around 70% (Group 2).

5.3 Critique of approach used

The model presented within the sponsor submission employs a structure which appears to be intuitively sensible, and attempts to include relevant costs and health outcomes associated over the lifetime of an individual attempting to quit smoking. Within the base case analysis, the standard varenicline treatment regimen is reported to dominate other smoking cessation interventions and placebo. The model also suggests that giving varenicline to subjects who have remained abstinent following a 12 week course of varenicline is expected to dominate placebo. A key limitation of the model however concerns the scope of the evaluation; the model attempts to predict costs and outcomes for an individual following a single quit attempt, based on a long-term extrapolation of 1-year quit rates from clinical trials of varenicline.[15,16,19] In reality however, individuals may repeatedly attempt to quit smoking in the long-term. Whilst the model attempts to capture the impact of subsequent relapse, it does not consider the impact of subsequent quit attempts on costs, morbidity or mortality. This may represent an important constraint on the external validity of the model structure.

As efficacy is modelled using 1-year quit rates alone,[1] there is considerable uncertainty surrounding the long-term relapse or abstinence experience of the model cohort. The BENESCO model employs an underlying assumption of sustained benefit following the first-year of abstinence. This means that the smoking cessation intervention with the highest 1-year quit rate will inevitably generate a greater number of QALYs and, unless the smoking cessation intervention cost is markedly high, will also produce greater cost savings than any other therapy it is compared against over a lifetime horizon. The uncertainty surrounding the long-term relapse probabilities has not been accounted for within the probabilistic sensitivity analysis, nor is its impact explored within the simple sensitivity analysis. The impact of this sustained benefit assumption is considerable: as the time horizon is increased, the economic attractiveness of the most effective option is also increased substantially. Whilst a lifetime horizon for the evaluation of costs and health outcomes may be intuitively valid, the uncertainty associated with the extrapolation of efficacy increases according to the duration over which costs and outcomes are evaluated. Shorter time horizons are subject to less uncertainty but may underestimate the true economic benefits of each of the smoking cessation interventions, and thus are likely to be overly conservative. Longer time horizons provide more favourable cost-effectiveness estimates, yet are subject to a greater degree of uncertainty as the impact of the assumption of sustained efficacy is increased.

It is important to note that only single parameter values for varenicline, bupropion and placebo are used from the clinical trials (the 1-year quit rate for each individual treatment group), whereas all other model parameter values have been drawn from other external sources. Although the sponsor submission reports 1-year quit rates for each smoking cessation intervention and odds ratios describing the odds of remaining abstinent between treatments based on the clinical trials,[15,16,19] the methods used to pool 1-year efficacy rates for individual treatments for use in the model are not described within the submission. Furthermore, the justification for excluding efficacy rates for NRT based on open-label study A3051044, and instead replacing

them with an indirect pooled efficacy estimate,[10] is entirely unclear and appears to be unjustified.

There are a number of other issues surrounding the data sources used to populate the BENESCO model. Whilst morbidity prevalence, incidence, cost and utility parameters are, for the most part, sourced from UK studies, long-term abstinence rates and relative risks of morbidity and mortality have been taken from cohort studies undertaken in the US.[26,24,27,25] It is however, unclear whether the smoking behaviour of a cohort of US subjects would reflect that of the population of England and Wales. Importantly, methods for identifying and selecting studies to inform model parameters are not presented within the submission.

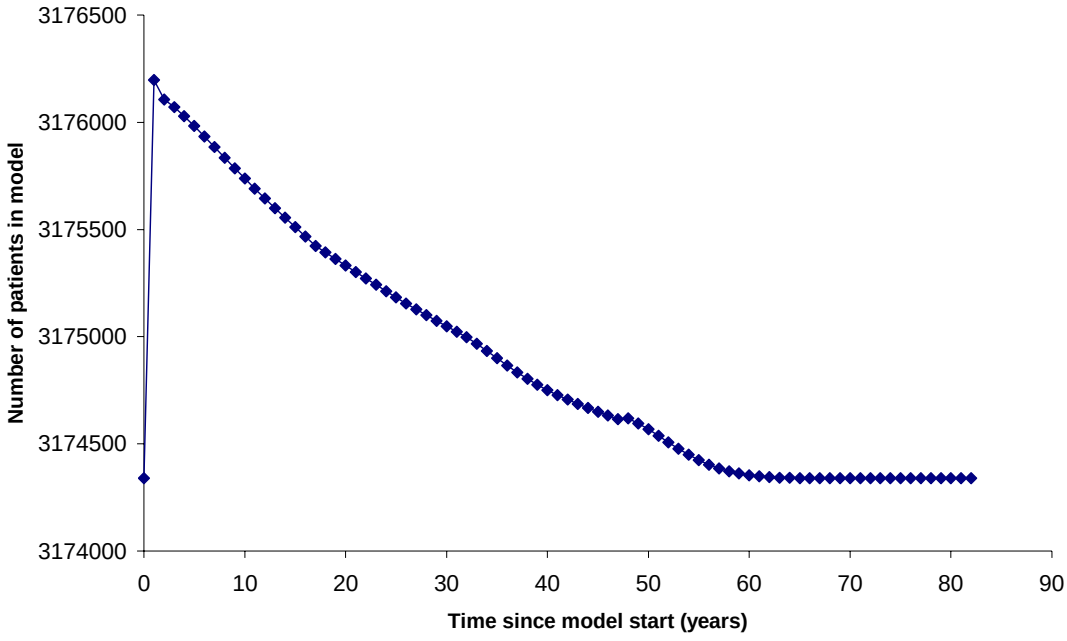
It should also be noted that the model is not transparent; during a detailed inspection of the model calculations, two computational errors were identified by the Evidence Review Group. Whilst the underlying structure of the model is relatively straightforward, much of the model has been programmed using Visual Basic for Applications (VBA) code which is virtually impenetrable to non-specialists. The first identified error relates to a underlying principle of state transition models - that at any point in time all subjects must exist within one of a finite number of health states;[46,47] this condition is violated within the BENESCO model. The risk calculations permit the model cohort to transit from the “no smoking-related morbidity” health states to the “asthma” health states, but subjects are not subsequently removed from the “no morbidity” state. Consequently, the probability of being in any health state at any point in time does not consistently sum to 1 over the duration of the model time horizon (See Figure 6). Whilst the sponsor submission presents several tests of logical consistency to validate the internal workings of the model, the validation exercise used to ensure that this principle is met was incorrectly programmed. The ERG alerted the sponsor to the presence of this error, yet the sponsor failed to acknowledge its presence hence the model remains incorrect. It should also be noted however, that the additional number of subjects entering the model at the end of year 1 is small in comparison to the total model cohort. As this computational error has not been resolved by the

sponsor, it is not possible to accurately gauge its impact on the resulting cost-effectiveness estimates presented within the submission.[1]

A second minor error was identified relating to relapse probabilities assumed during years 2-5; the model should use a rate of 7.5% instead of 6.3%.[25] The impact of this computational error on the incremental cost-effectiveness results for varenicline is likely to be minor.

A third potential error was also noted within the sponsor submission; the submission indicates that the annual probability of death for individuals without smoking-related morbidity is modelled using national all-cause mortality estimates from the Government Actuary's Department.[28] However, these annual probabilities should have been adjusted to reflect the probability of not dying from morbidity; that is, the probabilities should reflect "other cause" rather than "all cause mortality." It is likely that this omission would bias against the cost-effectiveness of more effective smoking cessation interventions.

Figure 6 Number of patients in health states over the time horizon of the model



A number of other minor problems may limit the validity of the cost-effectiveness results for varenicline. These include the exclusion of drug wastage, cost and health-related quality of life impacts associated with adverse events, and the absence of dose reductions for varenicline and bupropion.

Further work undertaken by the Evidence Review Group

The review and critical appraisal of the BENESCO model highlighted a number of issues surrounding the economic evaluation of varenicline for smoking cessation. Whilst some of the omissions could be elucidated through further analysis, the presence of computation errors within the models mean that such analysis would remain flawed and could present misleading results. Instead, the ERG focussed on the validation of the varenicline model against other comparative studies.

Validation against other smoking cessation studies

As noted within the methodological review, the BENESCO model presented within the sponsor submission uses a large number of structural and parametric assumptions to extrapolate health and cost outcomes beyond the duration of the clinical trials. The ERG undertook further analysis of the BENESCO model to examine the potential opportunity for clinical benefit underpinning the health economic analysis. Table 16 presents the expected health outcomes for a subject who reverts back to smoking after the first year as compared against the health outcomes for a lifetime quitter.

Table 16 Comparison of predicted health outcomes for immediate relapsers and lifetime quitters

Subject	LYGs	QALYs	Comment
Immediate relapser	27.14	13.20	Assuming smoking cessation efficacy rate is 0%
Lifetime quitter	28.36	13.69	Assuming smoking cessation efficacy rate is 100% and subsequent risk of relapse to smoking is 0%.
Additional health benefit	1.22	0.49	

The analysis of the BENESCO model presented in Table 16 suggests that the expected health gains resulting from successfully quitting smoking compared to immediate relapse to smoking are 1.22 additional life years and 0.49 additional QALYs.

Woolacott et al[48] report two estimates of life years saved for lifetime quitters: one based on the PREVENT simulation model, and another based on an analysis of Doll et al.[44] The PREVENT model estimates the number of life years saved per quitter to be 1.54,[48] whereas the analysis of Doll et al's data is reported to generate an estimate of 2.8 life years saved per quitter (discounted).[48] When compared against the results of the Doll et al study,[44] the above analysis suggests that the window of benefit for varenicline to achieve an "acceptable" level of cost-effectiveness is comparatively small, and that the long-term assumptions employed within the model appear to be conservative by comparison. The estimated life years saved within the BENESCO model appear to be broadly in line with those estimated by the PREVENT model.[40]

5.4 Summary of uncertainties and issues

The base case analysis of the BENESCO model presented in the submission suggests that varenicline dominates bupropion, NRT and placebo at the initial quit attempt. For individuals who have remained abstinent following a 12-week course of varenicline, varenicline is also reported to dominate placebo. However, several key issues should be borne in mind when considering the reliability of these results. The external validity of the model is questionable, as the analysis assumes only a single quit attempt using a single smoking cessation intervention; in reality smokers may attempt to quit more than once using several smoking cessation technologies. Within the model, the probability of short-term relapse to smoking is modelled using 1-year pooled quit rates and an indirect comparison. Beyond this point annual relapse probabilities are assumed to be independent of smoking cessation intervention, hence short-term benefits are assumed to be sustained in the

long-term. Shorter time horizons may be subject to less uncertainty, but may underestimate the benefits of varenicline. Longer time horizons provide more favourable cost-effectiveness estimates for varenicline, yet are subject to a considerable degree of uncertainty.

It is also noteworthy that many of the model parameters, specifically those describing the medium- to long-term probability of relapse to smoking, are based on US studies which may not reflect the smoking/abstinence behaviour of the smoking population of England and Wales. Methods for identifying and selecting costs and health utilities associated with morbidities are not reported or justified within the sponsor submission. It should also be noted that several computational errors were identified: the number of patients in the model is not constant over time, the risk of relapse between years 2 and 5 is incorrect, and all cause mortality appears to have been used for individuals who specifically do not experience smoking-related morbidities. The sensitivity analysis presented within the submission is very narrow and underestimates the true uncertainty surrounding the incremental cost-effectiveness of varenicline.

Finally, the external validity of the model has not been considered through comparison with other models or cohort studies.

6. COMMUNICATION WITH MANUFACTURER

6.1 Correspondence between NICE and Pfizer

The following questions were sent by NICE to Pfizer on 31 January 2007 and responded to by Pfizer on 15th Feburary 2007.

Section A: Clarifications

A1. Please could you explain why Pfizer has not used the direct trial of varenicline versus NRT as the base case calculation for clinical and cost effectiveness of varenicline versus NRT? It is unclear what reasoning is behind your decision to prefer an indirect comparison over a 'head-to-head' RCT that directly compares the technology and the appropriate comparator. See also the NICE reference case in the Guide to the Methods of Technology Appraisals – sections 3.2.2.1 and 5.4.1.3 – and the preamble to section 5 of the 'specification for manufacturer/sponsor submission of evidence.

Pfizer:

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 [review].

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

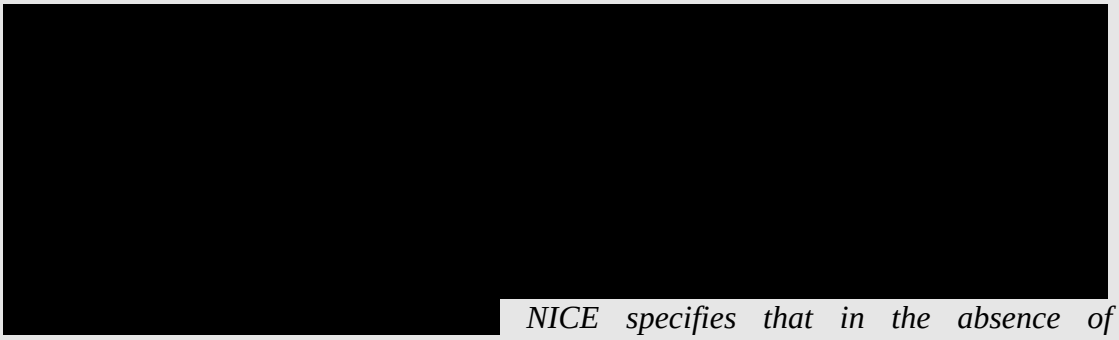
You should also be aware that the results of the open-label study only became available in January of this year.

Please also include in section 5.9 - 'interpretation of clinical efficacy' - your review of the earlier decision to exclude the direct comparison from the estimation of clinical efficacy and the impact this may have on the conclusions to be reached in 5.9.1.

Pfizer:

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.



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 using a 'multiple treatment comparison' approach to answering **all** the comparisons presented in your decision problem.

Pfizer:

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 that Wu and colleagues make the event numbers available for all the analyses they present, in order that we can check this work? Rather than using existing Cochrane reviews which compare bupropion and NRT to placebo, Pfizer has contracted a team from McMaster University, to produce a single systematic review of all three active interventions directly and indirectly compared. Some results presented by Wu and colleagues are different from those presented by Cochrane reviewers (Hughes for bupropion, Silagy for NRT and Cahill for varenicline) and validated by the ERG.

Pfizer:

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.

The ERG notes that: the McMaster team did in fact include several self report studies: see this document Section 4.1.3 for details.

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

The ERG notes that Pfizer responded to this by re-pasting 12 pages of text from their submission. We have not duplicated that response here.

Please can Pfizer also justify why treatment-specific efficacy rates were used instead of relative measures. Table 41 (p. 95) presents efficacy rates, the source of which is not transparent. There is no legitimate method in epidemiology for pooling rates. Pooling, or 'meta-analysis', produces a weighted average of the relative measures of effect (in this case, odds ratios) from individual studies.

Pfizer:

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.

The source given for the efficacy rate for NRT is the Wu systematic review (not in bibliography, but assumed to be *BMC Public Health* 2006, 6:300). The ERG notes that nowhere does Wu present rates, only odds ratios. Please make the full workings available that enabled Pfizer's analysts to convert pooled odds ratios presented in Wu to the efficacy rates presented in Table 41.

Pfizer:

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%

The ERG note that the manufacturer has made an error in reporting the inverse of the odds ratio for varenicline ('Champix®') versus NRT. The inverse of 1.66 (the results of McMaster's indirect comparison) is 0.60 and Pfizer appear to have used this figure to generate their NRT 12 month quit rate of 14.9 using the equation.

[redacted]

The ERG believes that this method of deriving rates via indirect comparisons is illegitimate and far from transparent and should not be used as the basis for a model, with preference given to models designed around relative measures (such as relative risks or odds ratios) rather than rates (and then, observed rates rather than rates derived from a problematic indirect comparison). This is backed up by Woolacott and colleagues in their HTA monograph (page 50) who write:

Smokers who participate in trials may be more motivated to stop smoking. If so, the quit rate in all groups (including the control group) would be higher than that when the same interventions are applied to the whole smoker population. Use of relative (rather than absolute) effectiveness for the different interventions may ameliorate this problem.[48]

However, the provenance for Pfizer's strange equation is found five pages later on page 55 of the same report and Woolacott's modellers seem to have used rates in their own model and, creating an unfortunate precedent.

A5. The manufacturer's submission claims (p. 109) that the Pfizer analysts have used odds ratios to generate the probabilistic sensitivity analysis (PSA). The ERG has thus far been unable to find any odds ratios in the model. Please can Pfizer tell us where the odds ratios and 95% CIs are contained in the model? If this information is in fact not correct, please

can they tell us how they have sampled from probability distributions for efficacy within the PSA without odds ratios and 95% CIs?

Pfizer:

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).

Section B: Tables

B1. Please make the correct event numbers available for tables 22 and 23 (p. 54). They probably present erroneous event numbers and rates for NRT (probably pasted without correction from tables 20 and 21).

Pfizer:

Thank you for pointing this out. The corrected tables are presented below:

Table 22: Continuous Abstinence Rate last 4 weeks of treatment through to week 52 – Full ITT

	n/N	CQR	Odds Ratio (95% CI)	p-Value
		%	Varenicline vs. NRT	Varenicline vs. NRT
Varenicline	[redacted]	[redacted]		
NRT	[redacted]	[redacted]	[redacted]	[redacted]

Table 23: Continuous Abstinence Rate last 4 weeks of treatment through to week 52 – Protocol Modified ITT

	n/N	CQR	Odds Ratio (95% CI)	p-Value
		%	Varenicline vs. NRT	Varenicline vs. NRT
Varenicline	[redacted]	[redacted]		
NRT	[redacted]	[redacted]	[redacted]	[redacted]

B2. A minor query in Table 41, the submission suggests an efficacy rate of 15.7% for bupropion, yet the model suggests this value is 15.5%. Which value is correct?

Pfizer:

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.

7 ADDITIONAL WORK UNDERTAKEN BY THE ERG

7.1 Meta-analyses

The first meta-analysis was of all placebo-controlled trials evaluating smoking cessation at 12 months (point prevalence or complete abstinence), with chemical validation, using any delivery method of NRT with intensive support (as in the NHS). We excluded: trials which did not placebo control or compared different delivery methods of NRT; trials which did not follow up for at least 12 months; and, trials in which cessation was not chemically validated for every individual. The results are shown in Figure 7.

The ERG's meta-analysis suggests that odds of smoking cessation at 12 months using NRT are 82% greater than using placebo (OR 1.82, 95% CI 1.60-2.08). Note that this estimate is 11% higher than the estimate derived by the McMaster team for NRT versus any control (OR 1.71, 95% CI 1.55-1.88). It is also 4% higher than the estimate derived by the McMaster team for NRT versus placebo (OR 1.78, 95% CI 1.60-1.99).[10]

The second meta-analysis was of all placebo-controlled trials evaluating smoking cessation at 12 months (point prevalence or complete abstinence), with chemical validation, using bupropion with intensive support (as in the NHS). We excluded: trials which were not placebo control or which evaluated bupropion with NRT against NRT alone; trials which did not follow up for at least 12 months; and, trials in which cessation was not chemically validated for every individual. The results are shown in Figure 8.

Figure 7 NRT versus placebo: ERG meta-analysis

Review: Varenicline for smoking cessation
Comparison: 11 NRT versus placebo (ERG selection)
Outcome: 01 Continuous cessation at 12 months

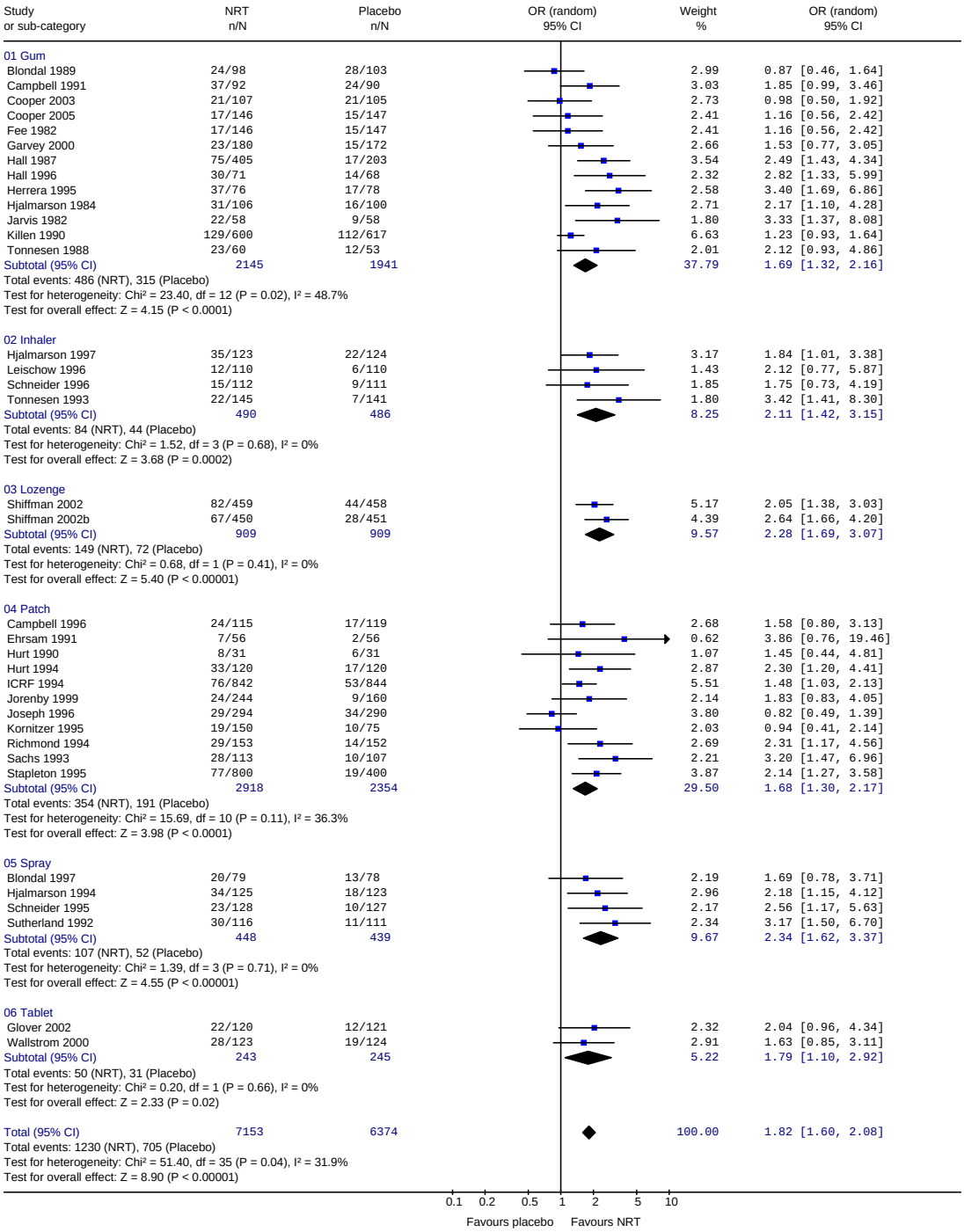
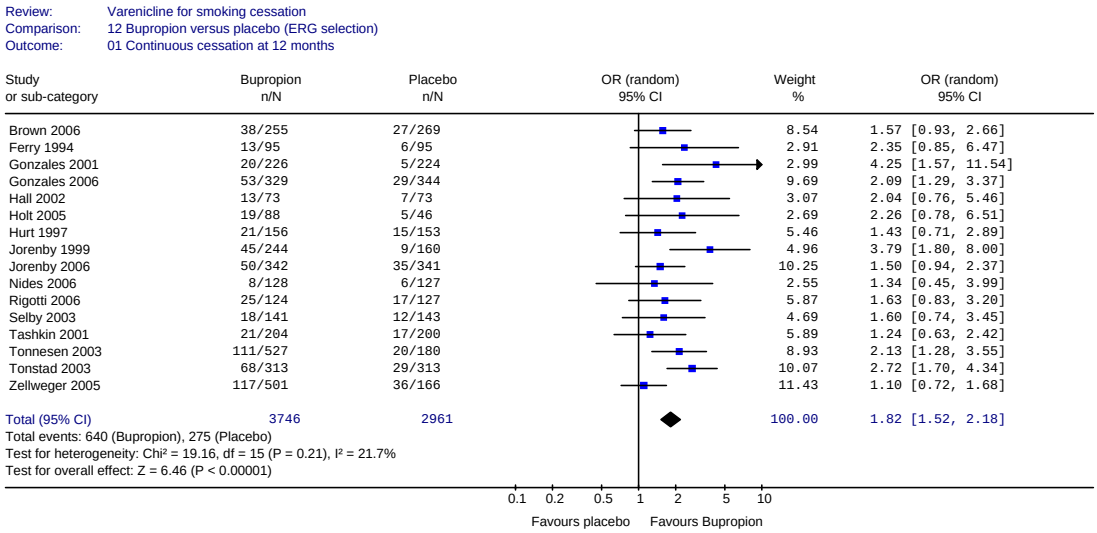


Figure 8 Bupropion versus placebo: ERG meta-analysis



The ERG's meta-analysis suggests that odds of smoking cessation at 12 months using bupropion are 82% greater than using placebo (OR 1.82, 95% CI 1.52-2.18). Note that this estimate is 26% higher than the estimate derived by the McMaster team for bupropion versus any control (OR 1.56, 95% CI 1.10-2.21 as reported in the published paper). It is also 18% higher than the estimate derived by the McMaster team for bupropion versus placebo (OR 1.64, 95% CI 1.16-2.30 as reported in the published paper).[10]

Note that the central estimates of our recreation of the McMaster analysis (this document, Figures 1 and 2 above) differ by one percentage points to the published estimates, presumably due to the different software packages used.[10]

7.2 Indirect comparisons

Because the composition of the McMaster meta-analyses creates an optimistic basis for the indirect comparison of varenicline with NRT (see this document, Section 4.2.2), we have attempted to rerun their indirect comparison using what we consider to be more balanced assumptions. Using the method used by Bucher,[22] we indirectly compared the NRT treatment effect derived through our own meta-analysis (Figure 7), which uses conservative assumptions, with the pooled effect size of the Gonzalez[15] and Jorenby[16] trials (Table 2).

We found that varenicline was still superior to NRT when compared to a placebo control at one year (OR 1.54, 95% CI 1.10 to 2.16, p=0.01). [redacted] The ERG estimate is 12% lower than the estimate derived from the McMaster indirect comparison (OR 1.66, 95% CI 1.17 to 2.36, p=0.004).

7.3 Further sensitivity analyses

Owing to computational errors in the model, no further sensitivity analyses were undertaken.

8 DISCUSSION

8.1 Summary of clinical effectiveness issues

The design and conduct of the manufacturer's submission on clinical effectiveness are compromised in a number of respects. The selection of studies to be pooled in a meta-analysis ensured an overly optimistic estimate of varenicline's clinical effect size due to at least three factors (unblinded studies; studies with concomitant therapies; inclusion of varenicline studies which did not meet the manufacturer's inclusion criteria). The results of this meta-analysis were used for (and, again, would have biased) an indirect treatment comparison of NRT and varenicline using placebo as a common comparator. The optimistic results of this indirect treatment comparison were used in the manufacturer's base case scenario.

[redacted].

Using data derived from systematic review teams based at McMaster and the Cochrane Tobacco Addiction Group, the ERG undertook their own meta-analyses and indirect comparisons using only placebo-controlled studies without concomitant therapies the design of which was externally valid (generalisable to NHS policy and practice). The results of the ERG's indirect comparison,

[redacted] were considerably more conservative than the manufacturer's indirect comparison.

Indirect comparisons are based on the framing and selection of evidence, which is rarely immediately transparent and can be difficult to assess even when, as in this case, the authors are from reputable research institutes with policies of openness. In the same way that review teams can correct optimistic assumptions in the parameters of manufacturer's economic models, the selection policies used to generate information for indirect comparisons can be challenged. However, such a robust appraisal may take more time,

clinical and methodological expertise than is typically available in the context of a single technology appraisal.

8.2 Summary of cost effectiveness issues

The model suggests that varenicline dominates bupropion, NRT and placebo within the base case analysis. However, the model uses pooled 1-year efficacy rates obtained from the clinical trials; for the remainder of the time horizon, the model assumes that short-term benefits are translated into long-term health gains and cost savings. For shorter time horizons which are, say, less than 10-years, the cost-effectiveness profile of varenicline appears to be considerably less favourable. In the absence of longer term evidence of efficacy, the validity of this assumption is difficult to gauge. This is a central consideration: whilst shorter time horizons may produce excessively conservative cost-effectiveness estimates for varenicline, longer time horizons are subject to a considerable degree of uncertainty.

The BENESCO model uses a complex methodology and a large number of health states to simulate the lifetime experience of subjects attempting to quit smoking. Importantly, the model is subject to at least two computational errors which compromise the validity of the results. Whilst the manufacturer was alerted to one of these errors during the review process, they failed to address or reconcile the problem. As such, the cost-effectiveness results should be interpreted with some degree of caution.

The external validity of the model should also be called into question. The model considers only a single quit attempt (subjects either quit or don't), whilst in reality smokers may attempt to quit several times using alternative smoking cessation interventions. Further, many of the model parameters describing the smoking/abstinence behaviour of subjects have been drawn from US studies which may not reflect that of the population of England and Wales. Further analysis by the ERG suggests that the BENESCO model assumptions are however broadly in line with other smoking cessation models.

8.3 Implications for research

The key uncertainty is the long-term efficacy of any smoking cessation intervention. The trials included in the sponsor submission report quit rates at one year's follow-up. The cost-effectiveness model projects forward over a time-horizon of 82 years and, as such, the cost-effectiveness of varenicline compared against other smoking cessation interventions is subject to substantial uncertainty. Future research should focus on longer term estimates of efficacy for all smoking cessation interventions.

E Evidence (What is the current state of the evidence?)

There is sound evidence for the short-term efficacy of a number of smoking cessation interventions, but little longer term evidence.

Population

Adults wishing to stop smoking.

Intervention

Varenicline in conjunction with intensive support.

Comparison

NRT or bupropion in conjunction with intensive support, or intensive support alone.

Outcome

Five-year smoking cessation rates.

Time stamp (Date of recommendation)

13 March 2007.

Appendix 1: The ERG ‘scope’.

Daniel Hind, Paul Tappenden and Jean Peters provided NICE with comments on their initial scope.

Appendix 2: Statement from the Cochrane TAG

'The Tobacco Addiction Group has always used the odds ratio as the primary summary measure, which is consistent with the majority of research in the field. The odds ratio has some convenient properties, although a case could certainly be made for reporting the relative risk.

Originally we used the Peto method for pooling studies, but this is not ideal when there are different numbers in experimental and control groups, as occurs quite frequently. We now use Mantel-Haenszel as the default method, as recommended by the Cochrane Statistical Methods Group.

Although a random effects model would give more *conservative* confidence intervals, in practice the weight given to smaller studies often results in a larger odds ratio and a similar lower confidence limit in our data sets. Where there was some evidence of heterogeneity and a confidence interval close to 1 we would probably do a sensitivity analysis of the effect of using a random effects model.'

References

1. Pfizer UK Ltd Varenicline: Single Technology Appraisal Submission, 17th January 2007. 2007.

2. National Institute for Health and Clinical Excellence Health Technology Appraisal: Varenicline for smoking cessation - Final scope. 2007.

3. Rigotti, N. A., Munafo, M. R., Murphy, M. F., and Stead, L. F. Interventions for smoking cessation in hospitalised patients.[update of Cochrane Database Syst Rev. 2001;(2):CD001837; PMID: 11406012]. *Cochrane Database of Systematic Reviews.(1):CD001837,* 2003.

4. Moller, A. and Villebro, N. Interventions for preoperative smoking cessation.[update of Cochrane Database Syst Rev. 2001;(4):CD002294; PMID: 11687156]. *Cochrane Database of Systematic Reviews.(3):CD002294,* 2005.

5. Lumley, J., Oliver, S. S., Chamberlain, C., and Oakley, L. Interventions for promoting smoking cessation during pregnancy.[update of Cochrane Database Syst Rev. 2000;(2):CD001055; PMID: 10796228]. *Cochrane Database of Systematic Reviews.(4):CD001055,* 2004.

6. van der Meer, R. M., Wagena, E. J., Ostelo, R. W., Jacobs, J. E., and van Schayck, C. P. Smoking cessation for chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews.(2):CD002999,* 2003.

7. National Institute for Health and Clinical Excellence Smoking cessation - bupropion and nicotine replacement therapy: Guidance (Technology Appraisal Guidance 39). 2007.

8. National Institute for Health and Clinical Excellence Brief interventions and referral for smoking cessation in primary care and other settings (Public Health Intervention Guidance No. 1). 2007.

9. West, Robert, McNeill, Ann, and Raw, Martin Smoking cessation guidelines for health professionals: an update. *Thorax* 1-12-2000; **55** 987-999.

10. Wu, P., Wilson, K., Dimoulas, P., and Mills, E. J. Effectiveness of smoking cessation therapies: a systematic review and meta-analysis. *Bmc Public Health* 2006; **6** 300.

11. Silagy, C., Lancaster, T., Stead, L., Mant, D., and Fowler, G. Nicotine replacement therapy for smoking cessation.[see comment][update of Cochrane Database Syst Rev. 2002;(4):CD000146; PMID: 12519537]. *Cochrane Database of Systematic Reviews.(3):CD000146,* 2004.

12. Hughes JR, Stead LF, and Lancaster T Antidepressants for smoking cessation. *Cochrane Database of Systematic Reviews* 2007; **Issue 1. Art. No.: CD000031. DOI: 10.1002/14651858.CD000031.pub3**

13. Cahill K, Stead LF, and Lancaster T Nicotine receptor partial agonists for smoking cessation. *Cochrane Database of Systematic Reviews* 2007; **Issue 1. Art. No.: CD006103. DOI: 10.1002/14651858.CD006103.pub2**

14. Nides, M., Oncken, C, Gonzales, D., Rennard, S., Watsky, E., Anziano, R., Reeves, K., and for the Varenicline Study Group Smoking Cessation With Varenicline, a Selective 42 Nicotinic Receptor Partial Agonist: Results From a 7-Week, Randomized, Placebo- and Bupropion-Controlled Trial With 1-Year Follow-up. *Archives of Internal Medicine* 2006; **166** 1561-1568.

15. Gonzales, D., Rennard, S. I., Nides, M., Oncken, C., Azoulay, S., Billing, C. B., Watsky, E. J., Gong, J., Williams, K. E., and Reeves, K. R. Varenicline, an a4beta2 nicotinic acetylcholine receptor partial agonist, vs sustained-release bupropion and placebo for smoking cessation: A randomized controlled trial. *JAMA: Journal of the American Medical Association* 2006; **296** 47-55.

16. Jorenby, D. E., Hays, J. T., Rigotti, N. A., Azoulay, S., Watsky, E. J., Williams, K. E., Billing, C. B., Gong, J., Reeves, K. R., Varenicline, Phase, Jorenby, Douglas E., Hays, J. Taylor, Rigotti, Nancy A., Azoulay, Salomon, Watsky, Eric J., Williams, Kathryn E., Billing, Clare B., Gong, Jason, Reeves, Karen R., and Varenicline, Phase Efficacy of varenicline, an alpha4beta2 nicotinic acetylcholine receptor partial agonist, vs placebo or sustained-release bupropion for smoking cessation: a randomized controlled trial.[see comment][erratum appears in JAMA. 2006 Sep 20;296(11):1355]. *Journal of the American Medical Association* 5-7-2006; **296** 56-63.

17. Oncken, C, Gonzales, D., Nides, M., Rennard, S., Watsky, E., Billing, CB., Anziano, R., Reeves, K., and for the Varenicline Study Group Efficacy and Safety of the Novel Selective Nicotinic Acetylcholine Receptor Partial Agonist, Varenicline, for Smoking Cessation. *Arch Intern Med* 2006; **166** 1571-1577.

18. Reeves K, Watsky E, Williams K, Azoulay S, Billing B, and Gong J The safety of varenicline taken for 52 weeks for smoking cessation [RPOS3-54]. *Society for Research on Nicotine and Tobacco 12th Annual Conference Orlando Fla, USA* 2006.

19. Tonstad, S., Tonnesen, P., Hajek, P., Williams, K. E., Billing, C. B., Reeves, K. R., Varenicline, Phase, Tonstad, Serena, Tonnesen, Philip, Hajek, Peter, Williams, Kathryn E., Billing, Clare B., Reeves, Karen R., and Varenicline, Phase Effect of maintenance therapy with varenicline on smoking cessation: a randomized controlled trial. *JAMA: Journal of the American Medical Association* 5-7-2006; **296** 64-71.

20. Schulz KF, Chalmers I, Hayes RJ, and Altman D Empirical evidence of bias. Dimensions of methodological quality associated with estimates

of treatment effects in controlled trials. *JAMA: Journal of the American Medical Association* 1995; **273** 408-412.

21. Hall SM, Delucchi KL, Velicer WF, Kahler CW, Ranger-Moore J, and Hedeker D Statistical analysis of randomized trials in tobacco treatment: longitudinal designs with dichotomous outcome. *Nicotine Tobacco Research* 2001; **3** 193-202.

22. Bucher, Heiner C., Guyatt, Gordon H., Griffith, Lauren E., and Walter, Stephen D. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997; **50** 683-691.

23. Orme, M. E., Hogue, S. L., Kennedy, L. M., Paine, A. C., and Godfrey, C. Development of the health and economic consequences of smoking interactive model. *Tobacco Control* 2001; **10** 55-61.

24. Krall, EA, Garvey, AJ, and Garcia, RI Smoking relapse after 2 years of abstinence: findings from the VA Normative Aging Study. *Nicotine Tob Res* 2002; **4** 95-100.

25. Wetter, D., Cofta-Gunn, L., Fouladi, R., Cinciripini, P., Sui, D., and Gritz, E. Late relapse/sustained abstinence among former smokers: a longitudinal study. *Preventive Medicine* 2004; **39** 1156-1163.

26. Cassino, C., Ito, K., Bader, I., Ciotoli, C., Thurston, G., and Reibman, J. Cigarette Smoking and Ozone-Associated Emergency Department Use for Asthma by Adults in New York City. *American Journal of Respiratory and Critical care Medicine* 1999; **159** 1773-1779.

27. Thun, M. J., Apicella, L. F., and Henley, S. J. Smoking vs. other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA: Journal of the American Medical Association* 2000; **284** 706-712.

28. Government Actuary's Department (GAD) Interim Life Tables. 2007.

29. Fiscella, K. and Franks, P. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians' smoking cessation counseling. *Journal of the American Medical Association* 1996; **275** 1247-1251.

30. Department of Health Health Survey for England. 2003. Volume 3. Methodology and documentation. 2004.

31. Trippoli, S, Vaiani, M, Lucioni, C, and et al Quality of life and utility in patients with non-small cell lung cancer. *Pharmacoeconomics* 2001; **19** 855-863.

32. Spencer, M, Briggs, A, Grossman, R, and Rance, L Development of an economic model to assess the cost effectiveness of treatment

interventions for chronic obstructive pulmonary disease. *Pharmacoeconomics* 2005; **23** 619-637.

33. Hay, JW and Sterling, KL Cost effectiveness of treating low HDL-cholesterol in the primary prevention of coronary heart disease. *Pharmacoeconomics* 2005; **23** 133-141.

34. Tengs, T and Lin, T A meta-analysis of quality of life estimates for stroke. *Pharmacoeconomics* 2003; **21** 191-200.

35. Gage, BF, Cardinalli, AB, and Owens, DK Cost-effectiveness of preference-based antithrombotic therapy for patients with nonvalvular atrial fibrillation. *Stroke* 1998; **29** 1083-1091.

36. Szende, A, Svensson, K, Stahl, E, Meszaros, A, and Berta, GY Psychometric and utility-based measures of health status of asthmatic patients with different disease control level. *Pharmacoeconomics* 2004; **22** 537-547.

37. Personal Social Services Research Unit, Curtis, L., and Netten, A. Unit costs of health and social care. 2006.

38. British National Formulary (BNF)

39. Britton, M The burden of COPD in the U.K.: results from the Confronting COPD survey. *Respir.Med* 2003; **97** S71-S79.

40. Parrott, S. and Godfrey, C. ABC of smoking cessation: Economics of smoking cessation. *British Medical Journal* 2004; **328** 947-949.

41. McMurray, J, Hart, W, and Rhodes, G An evaluation of the cost of heart failure to the National Health Service in the UK. *British Journal of Medical Economics* 1993; **6** 99-110.

42. Youman, P, Wilson, K, Harraf, F, and Kalra, L The economic burden of stroke in the United Kingdom. *Pharmacoeconomics* 2003; **21** 43-50.

43. Hoskins, G, McCowan, C, Neville, RG, Thomas, GE, Smith, B, and et al Risk factors and costs associated with an asthma attack. *Thorax* 2000; **55** 19-24.

44. Doll, R., Peto, R., Wheatley, K., Gray, R., and Sutherland, I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* 1994; **309** 901-911.

45. Cromwell, J., Bartosch, W. J., Fiore, M. C., Hasselblad, V., and Baker, T. Cost-effectiveness of the clinical practice recommendations in the AHCPR guideline for smoking cessation. *Journal of the American Medical Association* 1997; **278** 1759-1766.

46. Briggs, A. and Sculpher, M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; **13** 397-409.

47. Sonnenberg, F. A. and Beck, R. B. Markov models in medical decision making: A practical guide. *Medical Decision Making* 1993; **13** 322-338.

48. Woolacott, N. F., Jones, L., Forbes, C. A., Mather, L. C., Sowden, A. J., Song, F. J., Raftery, J. P., Aveyard, P. N., and Barton, P. M. The clinical effectiveness and cost-effectiveness of bupropion and nicotine replacement therapy for smoking cessation: a systematic review and economic evaluation. *Health Technology Assessmant* 2002; **6** 1-245.