Evidence Review Group Report commissioned by the NHS R&D HTA Programme on behalf of NICE

Telbivudine as treatment for Chronic Hepatitis B

Produced by Southampton Health Technology Assessments Centre

Authors Dr Debbie Hartwell, Research Fellow
Dr Jeremy Jones, Principal Research Fellow (Health Economics)
Ms Petra Harris, Research Fellow
Dr Keith Cooper, Research Fellow (Health Economics)

Correspondence to Debbie Hartwell
Southampton Health Technology Assessments Centre
Wessex Institute for Health Research and Development
University of Southampton
Mailpoint 728, Boldrewood
Southampton SO16 7PX

Date completed 5th February 2008

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.

Acknowledgements

We are grateful to Dr. Nick Sheron, University of Southampton/Southampton University Hospitals NHS Trust, who offered clinical advice to the ERG and comments on the draft report.

We also thank Alison Price, Information Scientist, Wessex Institute for Health Research and Development, for evaluating the search strategies; Jackie Bryant, Senior Research Fellow, SHTAC, for acting as internal editor for the ERG report; and Jonathan Shepherd, Principal Research Fellow, SHTAC, for comments on the draft report and support throughout the project.

Conflicts of Interest:

None

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AASLD	The American Association for the Study of Liver Diseases
ACP Journal Club	American College of Physicians Journal Club
ADV	Adefovir dipivoxil
ALT	Alanine Aminotransferase
BNF	British National Formulary
CCTR	Cochrane Central Register of Controlled Trials
CDSR	Cochrane Database of Systematic Reviews
CHB	Chronic Hepatitis B
CI	Confidence Interval
CIC	Commercial in Confidence
CRD	Centre for Reviews and Dissemination
DARE	Database of Abstracts of Reviews of Effectiveness
EASL	The European Association for the Study of the Liver
ERG	Evidence Review Group
ETV	Entecavir
HBeAb	Hepatitis B e Antibody
HBsAb	Hepatitis B Surface Antibody
HBeAg	Hepatitis B e Antigen
HBsAg	Hepatitis B Surface Antigen
HBV	Hepatitis B Virus
HCC	Hepatocellular carcinoma
HRQoL	Health Related Quality of Life Questionnaire
IFN	Interferon Alpha 2a
ITT	Intention to Treat
IVRS	Interactive Voice Response System
LAM	Lamivudine
LDT	Telbivudine
MEIP	Medline in Progress
MEq	Megaequivalents
Mg	Milligram
MITT	Modified Intention to Treat
mL	Millilitre
MS	Manufacturer's Submission
NHS	National Health Service
NICE	National Institute for Clinical Excellence
PEG	Pegylated Interferon Alpha 2a
PCR	Polymerase Chain Reaction
PSA	Probabilistic sensitivity analysis
RCT	Randomised Controlled Trial
SD	Standard deviation
SMPC	Summary of Product Characteristics
TTT	Time to treatment failure
ULN	Upper Limit of Normal
µmol	micromoles per litre

SUMMARY

Scope of the submission

The manufacturer's submission (MS) generally reflects the scope of the appraisal set by NICE, and is appropriate to the NHS. The intervention is telbivudine monotherapy in patients with compensated chronic hepatitis B (CHB) under the conditions specified in the marketing authorisation. The decision problem deviates from the scope in terms of the comparators and outcomes. Several comparator drugs outlined in the NICE scope were excluded by the MS as inappropriate due to their place in the treatment pathway (as recommended by previous NICE guidance{10091}). Three outcome measures specified in the NICE scope (time to treatment failure, health related quality of life and survival) were not reported in the MS, but all other outcomes are as appropriate and clinically meaningful as possible.

Summary of submitted clinical effectiveness evidence

- The MS presents clinical evidence for telbivudine in patients with compensated chronic hepatitis B based on one multi-centre, international, double blind RCT (known as the Globe trial).[1] This was the pivotal registration trial for telbivudine. The trial compares telbivudine with lamivudine in patients with HBeAg-positive and HBeAg-negative CHB for 104 weeks. The two-year data presented throughout the MS is unpublished, although publications of earlier results from the Globe trial are available.
- For the primary outcome of therapeutic response (suppression of HBV DNA <5 log copies/mL plus either clearance of detectable HBeAg or ALT normalisation), telbivudine was statistically superior to lamivudine at week 52 and 104 for HBeAg-positive patients, and significant at week 104 for HBeAg-negative patients.
- In terms of secondary outcomes, there were statistically significant differences in favour of telbivudine for HBV DNA reduction, HBV DNA non-detectability, ALT normalisation (though not for HBeAg-negative patients), virologic breakthrough and HBV resistance at two years. In HBeAg-positive patients, there was no significant difference between treatment groups for HBeAg loss or seroconversion at any time point. There were no significant differences in histologic response or changes in fibrosis score at one year, with the exception of histologic improvement in HBeAg-positive patients which was greater in telbivudine patients compared to lamivudine. In terms of adverse events, there appears to be no difference between treatments.

- In the elevated ALT sub-set analysis of the HBeAg-positive sub-group, telbivudine was statistically superior to lamivudine for most outcomes. In the ethnicity sub-group analysis, telbivudine was significantly more favourable than lamivudine in Asian patients, but there were no statistically significant differences between treatments for HBeAg-positive Caucasian patients, and few differences for HBeAg-negative Caucasian patients.
- In the indirect comparison, the MS reports that there were no statistically significant differences between telbivudine and entecavir for any efficacy outcome.

Summary of submitted cost effectiveness evidence

- The MS presents evidence on the cost effectiveness of telbivudine using two economic models, referred to as the viral load and seroconversion models. Evidence on the efficacy of telbivudine and lamivudine, in terms of reducing viral load, probability of normalising ALT and HBeAg seroconversion are taken from the Globe trial, for a sub-group of patients with ALT levels ≥ 2 x the upper limit of normal. The benefit of these outcomes is that they are associated with reduced probability of progression to advanced liver disease. Efficacy of adefovir is based on assumption.
- The viral load model, the manufacturer's preferred approach, stratifies response to treatment and the development of resistance by five viral load levels and regards reducing viral load as a key determinant of disease progression. This model is relevant both to patients with HBeAg-positive CHB and those with HBeAg-negative CHB. The viral load model incorporates a multivariate risk model to derive transition probabilities for the development of progressive liver disease based on viral load levels, the probability of ALT normalisation and HBeAg serological status (for HBeAg-positive patients).
- The seroconversion model is an attempt to replicate the model used in a recent NICE assessment[2] and is structured with HBeAg seroconversion as the key determinant of disease progression. By definition, this model is relevant only to patients with HBeAg-positive CHB.
- Both models adopt a lifetime horizon and extrapolate lifetime costs and QALYs for patients treated with telbivudine and each of the included comparators. Incremental cost effectiveness ratios against different comparators (depending on the model used) in the MS. The comparator in the viral load model is lamivudine, while in the seroconversion

model there are multiple, competing interventions (lamivudine, telbivudine and adefovir alone or in sequence as well as a no treatment (best supportive care) comparator).

- The MS concludes that telbivudine is a cost effective option compared with lamivudine using evidence from the viral load model (mean incremental cost of £19,087, mean QALY gain of 1.30 with an ICER of £14,665 per QALY gained for HBeAg-positive patients and mean incremental cost of £49,003, mean QALY gain of 4.67 with an ICER of £10,497 per QALY gained for HBeAg-negative patients).
- In response to a request for clarification from the ERG the manufacturer noted there were errors in the models originally submitted and therefore in the results reported in the MS. Resubmitted results gave less favourable ICERs, particularly for HBeAg-negative patients (mean incremental cost of £23,983, mean QALY gain of 1.56 with an ICER of £15,377 per QALY gained for HBeAg-positive patients and mean incremental cost of £41,910, mean QALY gain of 2.07 with an ICER of £20,256 per QALY gained for HBeAg-negative patients).
- The MS concludes that telbivudine is a cost effective option – on its own or followed by adefovir for patients who have developed resistance to first-line telbivudine treatment. The MS reported ICERs for seven treatment strategies, relative to best supportive care. This is not an ideal presentation of the results of competing treatment strategies. The ERG derived appropriate comparisons, based on the manufacturer's results, using the cost effectiveness frontier estimating ICERs of £7,887, £19,680 and £24,277 per QALY gained for lamivudine, telbivudine and telbivudine followed by adefovir repectively. The sequence of treatment options implied is problematic, since the strategy of using telbivudine followed by adefovir (for patients who develop resistance to telbivudine) is not accessible to patients who have lamivudine as their first line treatment. To provide the treatment strategy of telbivudine followed by adefovir (which yields the greatest QALY gain of all the strategies in the seroconversion model and which is optimal at a willingness to pay greater than £25,000 per QALY) telbivudine must be available as a first-line treatment.

Commentary on the robustness of submitted evidence

Strengths

- The MS conducted a systematic search for clinical-effectiveness studies of telbivudine. It appears unlikely that any additional trials would have met the inclusion criteria had the search been widened to include other databases.
- The Globe trial appears to be of reasonable methodological quality (with some limitations), and measured a range of outcomes that are as appropriate and clinically relevant as possible, although health related quality of life was not reported.
- On the whole, the MS appears to represent an unbiased estimate of the anti-viral treatment effect of telbivudine based on the results of one trial.
- The methods adopted for the economic evaluation of telbivudine were broadly consistent with those adopted for previous evaluations of anti-viral treatment of CHB, including the recent NICE assessment of adefovir and pegylated interferon.[2]

Weaknesses

- The MS does not include all the comparators specified in the scope.
- Despite a systematic search and screen of the literature, only one RCT was included. The MS is therefore largely dependent upon this one trial. Further high quality RCT evidence for the effectiveness of telbivudine in the patient group meeting the licensed indication would be beneficial.
- Literature searches were poorly documented, lacking clarity and transparency throughout. Search filters were extremely precise at the expense of sensitivity. The processes undertaken by the manufacturer for data extraction and applying quality criteria to the Globe trial are not detailed in the MS, and no formal quality assessment was undertaken on the comparator trials. These factors limit the robustness of the systematic review. In addition, one RCT[3] which appeared to meet the inclusion criteria was excluded from the MS.
- The indirect comparison with entecavir was poorly conducted and should be treated with caution. It was reported as a visual comparison, and then as a statistical comparison which the manufacturer deemed invalid. The MS provides an inadequate description of the methodology. The conclusions are based largely on a visual comparison of efficacy outcomes.

- The economic models use data from a sub-group of patients in the Globe study, which is not presented in detail in the Clinical Evidence section of the MS. There is no information on the baseline characteristics for the sub-group of patients with ALT levels ≥ 2 x ULN.
- The MS pays insufficient attention, in the Cost Effectiveness section of the MS, to appraising the data used to populate the economic models. Denominators used for calculation of some transition probabilities appear inconsistent and some input values (for example, resistance rates calculated using data reported in appendices are substantially lower than those reported for all patients in the Globe study). These discrepancies are not discussed in the MS.
- The electronic models submitted are complex and highly reliant on Visual Basic programming to produce any analyses. There is a large amount of reprocessing of data within the models that is not clearly documented or readily apparent to the user.
- There is little discussion in the MS of uncertainty around the mean estimates reported as the base case for both the viral load and seroconversion models. The NICE Guide to Methods of Technology Appraisal describe confidence ellipses and scatterplots on the cost effectiveness plane and cost effectiveness acceptability curves as the most appropriate ways of presenting uncertainty in PSA. These are not presented for all comparisons and were submitted in appendices, without commentary, rather than in the main body of the report.

Areas of uncertainty

- Results of the key efficacy outcomes were broken down by HBeAg status, study treatment and (i) race/ethnicity or (ii) ALT levels. It is not clear whether the Globe study was powered to detect differences in these sub-group sub-sets.
- Without confidence intervals and standard deviations in the reporting of the results, it is not possible to ascertain how much variance there was among the sub-groups/patients.
- The rates of viral resistance of entecavir were not reported in the MS and therefore do not allow for a comparison with the resistance rates for telbivudine.
- The adjustments to the Cox proportional hazards models used to estimate probability of developing compensated cirrhosis and HCC are inadequately reported as is the process of re-calibration (briefly reported in section 6.2.5.1 of the MS, p.88). These values enter the viral model deterministically – there is no assessment of parameter uncertainty for the risk models used in the viral model, nor of the methodological uncertainty around the adjustment or re-calibration.

- The lack of quality assurance of input data for both models introduces uncertainty – the impact of the prior value (zero or 0.5) on the model outcomes suggests that sparsity of data may be a problem, particularly for the model of HBeAg-negative patients. This is not surprising, given that data on around 250 patients has been stratified across viral load levels, ALT and serological status. The MS contains no discussion of alternative modelling strategies that might reduce the impact of sparsity of data nor does it clearly indicate which input variables are most affected by differences in prior values.

Key issues

- Whilst telbivudine is statistically superior to lamivudine for most anti-viral outcomes, the difference is not clinically significant, having an effectiveness advantage of only about 2% in patients treated between the two drugs. Viral breakthrough (>1 log increase over nadir) for telbivudine was 28.6% at two years; whilst this is significantly lower than lamivudine (45.5%), the ERG's clinical advisor asserts that it is still high at a clinical level.
- The conclusions from the indirect comparison are based largely on a visual comparison of efficacy outcomes, and a statistical indirect comparison which the MS states is not considered valid in the absence of any meta-analyses. Telbivudine seems to have approximately the same efficacy as entecavir for viral suppression but appears to have markedly higher rates of viral resistance (as per the rates for entecavir reported in the published trials).
- The exclusion of entecavir from all the economic models and the restricted comparison included in the viral load model – telbivudine versus lamivudine, with no follow up anti-viral treatments – means that the cost effectiveness evidence for telbivudine, presented in the MS, is limited. Lack of critical assessment and assurance of the quality of the data used to populate the model (apparent inconsistencies and incomplete data for lamivudine and telbivudine from the Globe trial along with the absence of systematic searches for evidence on the comparative effectiveness of adefovir) further limits the evidence reported in the MS.

1 Introduction to ERG Report

This report is a critique of the manufacturer's submission (MS) to NICE from Novartis Pharmaceuticals UK Limited on the clinical effectiveness and cost effectiveness of telbivudine for chronic hepatitis B. It identifies the strengths and weakness of the MS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the MS was requested from the manufacturer by the ERG via NICE on 12th December 2007. A response from the manufacturer via NICE was received by the ERG on 4th January 2008, with an updated economic model received on 8th January 2008. These responses have been annotated in the ERG report and can be seen in Appendix 1. The manufacturer declared that there was no commercial in confidence (CIC) data in the submission report as per their response on 4th January 2008.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer provides a clear and generally accurate overview of chronic hepatitis B. The overview covers the prevalence/incidence and natural history of the disease, the burden of disease to patients and health services, and the distinction between HBeAg-positive and negative patients. Current treatment goals and pharmacological management with the two major groups of therapies (immunomodulatory and anti-viral drugs) are also described.

The MS reports that chronic hepatitis B (CHB) affects about 326,000 people in the UK, being almost double the figure from five years ago. A Hepatitis B Foundation report is cited as the source of this figure.[4] However, according to the NICE technology appraisal of CHB in 2006,[2] the Department of Health estimates that the prevalence in the UK is approximately 180,000. This figure is in line with the prevalence estimate from the British Liver Trust of 150,000 to 200,000.[5] The overview notes that 15-40% of infected patients will go on to develop cirrhosis, liver failure and hepatocellular carcinoma (HCC).

2.2 Critique of manufacturer's overview of current service provision

The MS overview of current service provision is adequate. The MS provides a brief and generally accurate synopsis of treatment goals and the various pharmaceutical therapies available for the management of CHB (p.17-18), with the exception of telbivudine which is mentioned only to describe the mechanism of action. No summary of its efficacy, tolerability or rates of resistance is reported.

A number of guidelines are mentioned by the manufacturer in the submission report which are recognised in clinical practice:

- NICE guidance on adefovir and pegylated interferon for CHB[2]
- The American Association for the Study of Liver Diseases (AASLD) practice guidelines for CHB[6]
- The European Association for the Study of the Liver (EASL) consensus statement on hepatitis B[7]
- The Asian-Pacific consensus statement on CHB[8]

The NICE guidelines (2006)[2] are applicable to England and Wales and recommend first line treatment with pegylated interferon alpha (PEG) or interferon alpha (IFN). If contraindicated, or the patient fails to respond, lamivudine is recommended as a second-line option followed by adefovir (alone or in combination with lamivudine). Telbivudine and entecavir were not considered in this appraisal as neither was licensed at the time. The EASL guidelines (2003)[7] and the Asian-Pacific guidelines (2005)[8] similarly recommend treatment with IFN/PEG, followed by lamivudine or adefovir within their licensed indications. Neither telbivudine nor entecavir were considered in these guidelines. These recommendations are presented accurately in the MS report. The AASLD guidelines (2007)[6] recommend that treatment is initiated with any of the six FDA-approved drugs, although PEG/IFN, adefovir and entecavir were stated as the preferred options, as reported by the MS (p.21). However, the manufacturer fails to mention that these guidelines also state lamivudine and telbivudine are not preferred due to their high rate of drug resistance.

Regarding clinical practice, the MS states on p.5 that current usage of telbivudine in the UK is limited to two sites that participated in the Globe study (the key clinical study in the submission).[1] No further information is provided as to the location of these sites nor the

numbers of patients involved. The clinical advisor to the ERG asserts that there is variability in current practice as to the use of telbivudine, although it is not used locally in Southampton. Variability arises from the fact that although there are 12 or 13 specialist centres around the UK, the majority of patients are treated in district general hospitals by gastroenterologists who have limited or no training in hepatology.

The MS presents a treatment pathway diagram based on the NICE recommendations (which is incorrectly referred to as Figure 2, rather than Figure 1 (p.19-20)). It is clearly stated (as described above) that IFN/PEG is offered first, followed by lamivudine and adefovir as second and third line options respectively. According to the MS, it is intended that telbivudine will replace lamivudine treatment in this stepped care approach, and have proposed that telbivudine would be a first line oral therapy. The ERG's clinical advisor affirms that telbivudine would be used as second line therapy (but first line *oral* therapy) for patients who have failed other treatments.

2.3 Critique of manufacturer's definition of decision problem

2.3.1 Population

The final scope issued by NICE states that the population should be adults with compensated liver disease and active chronic hepatitis B - that is evidence of viral replication and active liver inflammation. The study population addressed in the manufacturer's submission is 'as per the NICE scope.' The population described reflects UK clinical practice for the treatment of CHB, and appears to be appropriate for the NHS. The MS does not include any further detail on the UK CHB population, such as mean age or HBV DNA level, or the proportions of patients who are HBeAg-positive or negative, against which to compare the characteristics of patients in the included clinical trial. The majority of patients in the included trial were Asian (64-83%), with approximately 50% being Chinese Asian (p.32). The ERG's clinical advisor concurs that the trial population are fairly representative of the UK CHB population given that the majority of new cases in the UK are immigrants from Eastern Europe and the Far East.

The decision problem makes reference to a sub-group analysis whereby HBeAg-positive and HBeAg-negative patients are considered separately 'according to their differing characteristics, responses and outcomes' (MS p.9). This was suggested as a possible analysis in the NICE scope. In addition, reference is made to a further sub-population of patients with alanine

aminotransferase (ALT) levels ≥ 2 x the upper limit of normal (ULN). This was in HBeAg-positive patients, which in itself is a sub-group. The scope for this appraisal does not specifically mention this sub-group, although it does permit analysis of sub-groups for whom the technology is particularly clinically and cost-effective, and where evidence allows. The manufacturer appears to have included this analysis based on the AASLD and Asian-Pacific guidelines which recommend that treatment be initiated in patients with ALT >2 x ULN. The clinical advisor to the ERG suggests that locally ALT tests may be used in the diagnosis/assessment of patients for treatment, but they are not a good indicator of liver damage. Liver biopsy is also used to stage the degree of fibrosis and inflammation and is a more useful measure to guide decisions on when to initiate treatment. Therefore, not all patients who are treated have elevated ALT levels. Patients with evidence of inflammation and fibrosis on biopsy would be treated irrespective of their ALT.

2.3.2 Intervention

The intervention specified in the MS decision problem is telbivudine monotherapy. According to section 3 (p.11) of the MS, it is administered as a 600mg tablet taken orally, once daily, in the population outlined in Section 2.3.1, with the addition of persistently elevated serum ALT levels and histological evidence of active inflammation and/or fibrosis. This is the licensed indication and is therefore appropriate for use within the NHS. The NICE scope states it will consider telbivudine alone or in combination with other therapies. The manufacturer reports that there is not enough evidence for combination therapy and that this indication is not within the licence. Clinical opinion suggests that telbivudine is not widely used in the UK, but where it is used, it is administered as second line monotherapy. Therefore the manufacturer's description of telbivudine in section 3 of the MS reflects its use in the UK.

Expert clinical opinion suggests that telbivudine is currently used in some parts of England and Wales, although generally not as a first line treatment. Clinical opinion also suggests that for those who have failed to respond to, or who have relapsed, following IFN or PEG, it would be advantageous to proceed directly to a combination of telbivudine and another nucleoside / nucleotide analogue. It is thought that this would lessen the risk of cross-resistance, a problem associated with the sequential use of nucleoside / nucleotide analogue monotherapies. This is also a problem that has been experienced in the HIV/AIDS and tuberculosis fields, where combination therapies are now commonplace.

2.3.3 Comparators

The MS does not include all the comparators specified in the NICE scope. In section 3 (p.11), the MS reports that 'telbivudine is a nucleoside analogue with activity against HBV DNA polymerase, and as such, it may be compared with other nucleoside/nucleotide analogues including lamivudine, adefovir and entecavir. The MS decision problem specifies that the intended comparator is lamivudine, first line oral anti-viral treatment. The comparators outlined in the NICE scope are:

- interferon alpha 2a
- interferon alpha 2b
- pegylated interferon alpha 2a
- lamivudine
- adefovir dipivoxil
- entecavir

The MS reports that lamivudine is the main comparator because it is the most widely used first-line oral anti-viral, and was the active comparator in the registration study for telbivudine (p.12). In addition, it is not the intention that telbivudine would replace IFN/PEG, but replace lamivudine in the treatment pathway. Entecavir is considered in an indirect comparison in the submission and is currently being appraised separately by NICE. The MS does not specifically state why adefovir was not included as a comparator in their review, despite it being a comparator in the inclusion criteria (MS Appendix M). The MS excluded a recent randomised controlled trial (RCT) comparing telbivudine with adefovir.[9] After seeking clarification, the manufacturer stated that adefovir was excluded because it was an inappropriate comparator given that the NICE guidance[2] recommends adefovir as third line therapy.

2.3.4 Outcomes

The NICE scope stated that the outcome measures to be considered include:

- HBeAg/HBsAg seroconversion
- virological response (HBV DNA)
- histological improvement (inflammation and fibrosis)
- biochemical response (e.g. ALT levels)
- development of viral resistance

- time to treatment failure (TTT)
- survival
- health-related quality of life (HRQoL)
- adverse effects of treatment

The manufacturer does not specify any outcomes *per se* in the decision problem, but states that they are 'as per the final scope.' The MS reports that clarification is needed on 'survival', to be discussed at a meeting prior to the submission date of the review, but no further information is provided. The outcome measures listed above are presented for the Globe study[1] in the MS report with the exception of TTT, HRQoL and survival. The primary efficacy endpoint was a composite measure - therapeutic response - defined as suppression of HBV DNA <5 $\log_{10}$ copies/mL plus either clearance of detectable HBeAg or ALT normalisation (depending on HBeAg status). Clinical opinion suggests that the chosen level of HBV suppression is appropriate to indicate effectiveness, but it would be less appropriate to use it to indicate any clinical differences in efficacy.

The outcomes are appropriate and clinically meaningful.

3 CLINICAL EFFECTIVENESS

3.1 Critique of manufacturer's approach

3.1.1 Description of manufacturer's search strategy

Overall, the search strategies are adequate for the clinical effectiveness searches, but the manufacturer's search contains some omissions (see below), and is poorly documented. It is thought unlikely to affect the identification of additional key studies, although it is unclear to the ERG why one RCT identified by the MS (Study 015[3]) was not included.

3.1.1.1 Clinical effectiveness searches

Databases, dates of searches and search strategies were reported by the manufacturer. The results of the searches were presented narratively in the MS report (p.24), with the tables of results and the search strategies presented in two appendices (L and M). The manufacturer ran searches in Ovid including Medline, Embase, Cochrane (CDSR and CCTR), DARE and the

American College of Physicians (ACP) Journal Club database, as well as Novartis' own in-house database. No evidence of searches of Medline in Progress (MEIP) is documented within the MS, and therefore the minimum database search criteria for undertaking clinical effectiveness searches as specified by NICE was not adhered to. The ERG requested clarification over the exclusion of MEIP from the manufacturer and the response only confirmed that 'MEIP was not included'. Other sources searched are described as a manual search of relevant publications, other in-house trials and the telbivudine registration dossier (p.23-24), but no results or details of the outcome of these searches are provided. In addition, it is not stated if the searches were restricted to English language. Additional databases that could have been searched to obtain clinical evidence include ISI proceedings and Biosis.

The searches were run in two stages. The main search (search strategies for the clinical effectiveness and indirect comparison searches) was run in January 2007 (MS Appendix L), whilst an updated replication of the searches was run from January to September 2007 in order to meet the requirements for an Australian submission (MS Appendix M). No further update searches from September 2007 were performed for this submission. Appendix M is referred to incorrectly as Appendix L in Section 5.2.1 (p.24) and again in Section 5.2.2 (p.26) of the MS report - in both cases the report should cite Appendix M.

The search terms used are the minimum suitable for precise searches, but not the sensitive results required for a systematic review. For the population group, use of 'Hepatitis B' rather than 'Chronic Hepatitis B' as an exploded MeSH term would have ensured greater comprehensiveness of a systematic search. A limited number of synonyms are also used in the text search. Searching on the word string 'chronic hepatitis B' as a phrase limits the results as the words may not be used in that order in the text. It is poor practice to search only with phrase searches. The drug search was for three interventions – telbivudine, lamivudine and adefovir. The MS does not state why adefovir was included in the search and the ERG find its inclusion inconsistent when it has not been considered in the submission. Entecavir was not included as a search term. In the Embase search, the Emtree subject headings for any of the drug names are not used, with only the text terms run in the mp field. An acceptable RCT filter was applied to the search strategy, and is sufficient for finding very precise and focussed results, however it does not meet the Cochrane standards of RCT sensitivity. The download is further limited by excluding all review papers that are not meta-analyses. Lines 32 and 33 in the Medline search are a dead end in the search strategy where the results do not appear to be used.

The ERG re-ran the Ovid Medline search from 1950 to Week 2 November 2007 and the numbers retrieved were similar to those of the manufacturer.

Indirect Comparison

Entecavir was not used as a search term in any of the search strategies in the publicly available databases. This means that any head-to-head trials of entecavir and lamivudine would have been identified in the lamivudine results only. It is the view of the ERG that the majority of papers identified in MS Section 5.6 would likely have been found within the internal company database and other in-house sources.

3.1.1.2 Cost effectiveness searches

The MS states (section 6.1.1, page 75) that no formal search of the cost effectiveness literature on treatments for CHB was undertaken given the recent date of the review reported in the HTA monograph.[10] Since the HTA monograph reports that the searches (including cost effectiveness and quality of life searches) for the review were conducted up to April 2005, more than two years prior to the manufacturer's submission to NICE, update searches would have been appropriate.

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

The MS specified the following inclusion criteria for the review of the literature (MS Section 1.4, Appendix M):

- Population: patients with chronic hepatitis B
- Intervention: telbivudine (Sebivo)
- Comparator: lamivudine (also include adefovir)
- Outcomes: primary and secondary outcomes (changes in HBV DNA, HBeAg, HBeAb, HBsAg, HBsAb, ALT)
- Study design: RCTs

These criteria do not specify several of the outcomes outlined in the decision problem, particularly viral resistance and histological improvement. In terms of the patient population, the criteria do not specify that patients should have persistently raised ALT levels and histological evidence of active inflammation and/or fibrosis as per the licensed indication, although these are stipulated in the patient inclusion criteria for the Globe trial. RCTs where telbivudine and the main comparator were not in separate trial arms were excluded, as were abstracts and studies reviewing quality of life data (MS Table 2, Appendix M). The MS did not specifically state whether systematic reviews would be considered, and neither is there discussion of whether conference proceedings would be included or excluded. Clarification of inclusion was sought from the manufacturer and the reply received stated that abstracts and systematic reviews were eligible for inclusion but not conference proceedings. However, abstracts were still listed in the table of exclusion criteria (see Appendix 1). The ERG therefore remain unclear on whether abstracts were included.

The patient inclusion/exclusion criteria for the Globe trial are clearly stated in the MS report, are appropriate and fulfil the specific criteria of the product licence. Setting was not used as an inclusion criterion, but it is unlikely that telbivudine would be administered outside secondary care. The patient inclusion criteria (p.30) stipulate that patients should have elevated ALT levels of 1.3 -10 x ULN at screening, although guidelines referred to several times by the MS state that treatment should be initiated when ALT levels are >2 x ULN.

The methodology adopted by the manufacturer for screening references for inclusion is given at the end of Appendix M, and appears to have been appropriate. Two reviewers assessed the citations at the title and abstract stage, ordered relevant full trial papers and screened them against the eligibility criteria. Disagreement was resolved through discussion. However, the MS provides no details about the data extraction and quality assessment procedure. Clarification was sought from the manufacturer and the response just reiterated the above information. The planned data extraction procedure therefore remains unclear.

The MS does not provide any inclusion/exclusion criteria for the indirect treatment comparisons, nor are there any details about the identification and selection of studies, the data abstraction strategy or quality assessment.

3.1.2.1 Identified studies

According to the MS, only one RCT met the inclusion criteria (brief details are shown in MS Table 2, p.27). The RCT, known as the Globe trial,[1] compared telbivudine to lamivudine in patients with HBeAg-positive or HBeAg-negative chronic hepatitis B. The manufacturer provided an electronic copy of the full trial report (>3,019 pages), as well as 13 appendices, on 28/11/07, two days after submission of the MS report. Other publications from the Globe trial are available (as abstracts), [11,12] though not for 104 week data. During the ERG's appraisal of the MS, the one year data from the Globe trial were published as a full paper.[13]

In Section 5.2.1, p.24, it is stated that 769 citations were identified on Medline, Embase and Cochrane databases, plus 22 from the trial registries and seven from the in-house database. Of these, there were seven relevant studies and two were included. A further five relevant RCTs were identified from the registration dossier but references were not provided nor an explanation as to the reasons for exclusion and how these relate to the above studies. The MS presents a table (Table 1, p.26) of retrieved excluded trials, detailing grounds for exclusions, which are not cited nor mentioned anywhere else. Furthermore, in Section 5.2.3, p.26, it is stated that four RCTs were selected for inclusion (Study 007 - Globe trial[1]), Study 015,[3] Study 018[9] and Study 019[14]), but the source of these studies is unknown and references were not provided. These are presented briefly in MS Table 2 (p.27) and there is no clear link to previous search results. The MS suggests that the Globe trial[1] is the only trial comparing telbivudine with lamivudine. However, Table 2 shows an additional two trials (Studies 015[3] and 019[14]) comparing these drugs and it remains unclear why these studies were not included. In Table 2, references offered for the Globe trial were in the form of two abstracts[11,12], but the information presented throughout the MS is from the full clinical trial report.[1] A further reference in the table was an abstract incorrectly cited as Hwang *et al*, which should have been Gane *et al* (2006).[14]

The MS states that with only one included trial it was not possible to present a QUOROM style flow diagram. The ERG requested a clearer explanation of study exclusions, with reasons, which the manufacturer provided in their response (see Appendix 1). The ERG also requested a flow chart for clarity of numbers, but this was not provided. The MS responded with reasons for exclusion of studies 015, 018 and 019. The exclusion of studies 018 and 019 was due to telbivudine being used outside the licensed indication as they include patients who have previously been treated with lamivudine. The reason given for exclusion of study 015 was that 'it

was a relatively small, 2 year study (n=332), exclusively enrolling Chinese patients and is only partially reported with results at one year'. The ERG do not consider this response satisfactory as the study meets the inclusion criteria stipulated in Appendix M of the MS and reports relevant outcome measures at week 52 for HBeAg-positive and HBeAg-negative patients (as a whole group), as does the Globe study, although the patient population may not be generalisable to the UK population. After excluding the trial, the MS nevertheless inappropriately provides some details about the trial and refers to its results in the executive summary (p.14).[3]

The only included MS RCT was the Globe telbivudine registration trial. Most summary details and a consort flow-chart were provided. However, the number of eligible patients is not reported and the number of drop-outs is unclear – in MS Figure 4 (p.33) the proportion of patients dropping out for each of the six reasons do not match the total numbers (n=18 at week 52 and n=56 at week 104 for telbivudine; n=32 at week 52 and n=88 at week 104 for lamivudine). After seeking clarification from the manufacturer it became apparent that a seventh reason for discontinuations (patients discontinuing medication at their own request) had been omitted due to a transcription error. However, the numbers given for patients discontinuing treatment still do not match the total numbers (see Appendix 1).

Table 1 Characteristics of the included RCT

Methods	Participants	Outcomes
007, GLOBE[1] *Design:* phase III, multi-centre double-blind RCT *Interventions*: Grp1: LDT 600mg daily + matching LAM placebo Grp2: LAM 100mg daily + matching LDT placebo *Number of centres*: 112 centres, 3 in UK; 20 countries worldwide *Duration*: 104 weeks *Length of follow-up:* none (pts could continue for a further 2 yrs)	*Participant numbers:* Grp1: *n* = 680 Grp2: *n* = 687 *Key Inclusion criteria*: Adults (16-70) with clinical history compatible with a diagnosis of CHB and at screening: • detectable serum HBsAg • HBeAg-positive or negative • elevated serum ALT levels (1.3 - 10 x ULN) • liver biopsy within last 12mths (with compatible histology of CHB)	*Primary composite endpoint:* Suppression of HBV DNA < 5 $\log_{10}$ copies/mL + <u>either</u> clearance of detectable HBeAg or ALT normalisation *Secondary antiviral endpoints:* HBV DNA suppression, HBV DNA PCR negativity *Other efficacy endpoints:* E antigen response (HBeAg loss and seroconversion), histologic response, serum ALT changes

LDT = entecavir; LAM = lamivudine

The MS provides details of the trial design, intervention, population, patient numbers, outcomes, and analysis of the Globe trial[1] (MS p.29-38). Data were presented separately for HBeAg-positive and negative patients for each treatment in the MS. However, safety data were

presented for telbivudine and lamivudine groups. Summary information of the Globe trial is provided in Table 1.

The MS also identified two studies for an indirect comparison and summary information on study characteristics were tabulated (MS Tables 11-16). The ERG considers the information provided to reflect the information given in the trial publications. One study[15] compared entecavir and lamivudine in HBeAg-positive patients, whilst the other study[16] compared the same drugs in HBeAg-negative patients. Summary information is provided in Table 2.

Table 2 Characteristics of the included indirect comparison RCTs

Methods	Participants	Outcomes
Chang *et al,* 2006[15] (ETV 022) *Design:* phase III double-blind, double dummy RCT *Interventions*: Grp1: ETV 0.5gm daily Grp2: LAM 100mg daily *Number of centres*: 137 centres worldwide (Europe 41; North America 40; Asia 26; Australia 12; South America 18) *Duration*: treatment 52 wks *(*clinical management decision), serum samples obtained wk 48 *Length of follow-up:* 24 wks if pts responded at wk 48 & ceased treatment Patients with virologic response (HBV DNA level < 0.7 MEq/ millilitre & no HBeAg loss) could continue therapy for up to 96 wks.	*Participant numbers:* Grp1: *n* = 354, *mean age: 35* (± 13) Grp2: *n* = 355, *mean age: 35* (± 13) *Key Inclusion criteria*: HBeAg-positive CHB nucleoside analogue naive CHB adults (≥ 16) • Compensated liver (total bilirubin level of 2.5 mg/decilitre [42.8 µmol/litre] or less, a prothrombin time ≤ 3 secs longer than normal or international normalised ratio ≤ 1.5, ALT at least 3.0 g/decilitre, and no history of variceal bleeding or hepatic encephalopathy) • Detectable HBsAg for min. 24 wks prior screening • evidence of CHB on baseline liver biopsy obtained within 52 wks prior randomisation • evidence of HBV DNA (any commercial assay) min. 4 wks prior screening • HBV DNA level of min. 3 MEq/ millilitre at screening • ALT 1.3 – 10 x ULN at screening	*Primary endpoint wk 48:* proportion of pts with histologic improvement (improvement by at least 2 points in the Knodell necroinflammatory score, no worsening in the Knodell fibrosis score at wk 48 relative to baseline) *Secondary endpoints wk 48:* • Baseline reduction in HBV DNA • proportion of pts with undetectable HBV DNA • decrease in the Ishak fibrosis score • HBeAg loss • HBeAg seroconversion (HBeAg loss & appearance of HBe antibody) • ALT normalisation (< 1.25 x ULN*) *data was reanalysed according to a more stringent definition of normalisation (ALT no greater than ULN)
Lai *et al*, 2006[16] (ETV 027) *Design:* phase III double-blind RCT *Interventions*: Grp1: ETV 0.5gm daily Grp2: LAM 100mg daily *Number of centres*: 146 centres worldwide (Europe & Middle East 68; Asia 25; Australia 11; North America 30, South America 12) *Duration*: treatment 52 wks *(*clinical	*Participant numbers:* Grp1: *n* = 325, *mean age: 44* (± 11) Grp2: *n* = 313, *mean age: 44* (± 11) *Key Inclusion criteria*: HBeAg-negative CHB nucleoside analogue naive CHB adults (≥ 16) • Compensated liver (total bilirubin level of 2.5 mg/decilitre [42.8 µmol/litre] or less, a prothrombin time ≤ 3 secs longer than normal or international normalised ratio ≤ 1.5, ALT at least 3.0 g/decilitre, and no history of variceal bleeding or hepatic encephalopathy)	*Primary endpoint wk 48:* proportion of pts with histologic improvement (improvement by at least 2 points in the Knodell necroinflammatory score, no worsening in the Knodell fibrosis score at wk 48 relative to baseline) *Secondary endpoints wk 48:* • Baseline reduction in HBV DNA • proportion of pts with undetectable HBV DNA • decrease in the Ishak fibrosis score • normalisation of serum ALT (<

management decision), serum samples obtained wk 48 *Length of follow-up:* 24 wks if pts responded at wk 48 & ceased treatment Patients with virologic response (HBV DNA < 0.7 MEq/mL & ALT ≥ 1.25 x ULN) could continue therapy for up to 96 wks.	• detectable HBsAg min. 24 wks prior screening • evidence of CHB on a baseline liver biopsy obtained within 52 wks prior randomisation • evidence of HBV DNA (any commercial assay) min. 2 wks prior screening • undetectable HBeAg • detectable anti-HBe • serum HBV DNA level ≥ 0.7 MEq/ millilitre at screening • ALT 1.3 – 10 x ULN at screening	1.0 x ULN)

ETV = telbivudine; LAM = lamivudine

3.1.2.2 Details of any irrelevant studies that were included in the submission

The MS has not included any inappropriate RCTs.

3.1.2.3 Ongoing studies

The manufacturer states that the US National Institutes of Health and the Australian Clinical Trials Registries were searched because 'UK registries are not available and the former are comprehensive data sources' (p.23). The UK database of NHS funded trials, the National Research Register (NRR), was freely available to January 2008. In addition, the Current Controlled Trials (CCT) database also contains UK and international trials. The manufacturer did not search either of these databases. Searches undertaken by the ERG did not identify any trials for telbivudine currently listed.

The manufacturer lists seven ongoing trials in a table in Section 5.2.5 (p.28), but does not cite the source of the references. In an email to the ERG via NICE on 18/12/07, the manufacturer declared all seven references as CIC. However, in the manufacturer's response to the ERG's clarification questions received on 4/1/08, the manufacturer declared that none of the information in the MS is CIC.

Three trials are investigating telbivudine combination therapy with adefovir (two trials) or PEG (one trial), which are currently outside the licensed indication. Another trial is comparing 12 weeks treatment with telbivudine versus entecavir in HBeAg-positive patients. One trial is a paediatric pharmacokinetic study, whilst another is in adults with decompensated CHB and cirrhosis. The seventh trial is an open label, longer-term dosing study of telbivudine and

includes patients who completed the two year Globe trial regardless of previous treatment assignment (Study 022). No references or further details were provided in the MS.

3.1.2.4 Additional studies

As a minimum, searches for conference proceedings abstracts should have been included. By not searching for conference abstracts, a useful source of new evidence has been ignored. The ERG has identified seven abstracts in the American Association for the Study of Liver Disease (AASLD 108th meeting), published in Gastroenterology April 2007 (Supplement 1, Volume 132, Number 4), of which three (numbers I, 4 and 7) may be of relevance (the other references relate to the Globe trial):

1. Bzowej N, Marcellin P, Chan HLY, Lai CL, Cho M, Heathcote EJ *et al*. A randomized trial of telbivudine vs adefovir for HBeAg-positive chronic hepatitis B: Efficacy through week 76, predictors of response and effects of switching to telbivudine. *Gastroenterology* 2007;132, (4):A764.

2. Han SH, Lai CL, Gane EJ, Liaw YF, Thongsawar S, Wang YM *et al*. Telbivudine globe trial at year two: Efficacy, safety, and predictors of outcome in patients with chronic hepatitis B. *Gastroenterology* 2007;132, (4):A765.

3. Heathcote EJ, Gane EJ, Lai CL, Min AD, Poynard T, Kurdas OO *et al*. Salvage therapy with adefovir for virologic breakthrough in telbivudine-treated patients from the globe study. *Gastroenterology* 2007;132, (4):A765.

4. Rajendra A, Wong JB. Cross trial comparison of genotypic response rates with adefovir, entecavir or telbivudine for chronic hepatitis B. *Gastroenterology* 2007;132, (4):A766.

5. Seifer M, Patty A, Chapron C, Van Doorn LJ, Belanger B, Brown NA *et al*. Genotypic analysis of patients with evaluable HBV DNA after 1 year of telbivudine therapy in the globe registration trial. *Gastroenterology* 2007;132, (4):A729.

6. Standring DN, Patty A, Chapron C, Van Doorn LJ, Belanger B, Brown NA *et al*. Resistance determination in patients experiencing virologic breakthrough following telbivudine or lamivudine therapy in the international globe trial. *Gastroenterology* 2007;132, (4):A766.

7. Wang Y, Jia JD, Hou JL, Yin YK, Xu DZ, Tan DM *et al*. A phase III comparative trial of telbivudine vs lamivudine in Chinese patients with chronic hepatitis B: Two-year results. *Gastroenterology* 2007;132, (4):A763.

Indirect comparison

In addition, one entecavir study was identified which was a 96 week trial of 519 nucleoside naïve Chinese patients with CHB. The majority of Globe patients are also of Chinese origin. Results are reported separately for HBeAg-positive and negative patients, based on week 48 data.

1. Yao G, Chen C, Lu W, Ren H, Tan D, Wang Y *et al*. Efficacy and safety of entecavir compared to lamivudine in nucleoside-naïve patients with chronic hepatitis B: a randomized double-blind trial in China. *Hepatology International* 2007;1, (3):365-372.

3.1.3 Description and critique of manufacturer's approach to validity assessment

The MS provides a formal appraisal of the validity of the included trial using the quality assessment criteria developed by NICE. The process of applying quality criteria was not reported in the MS. The ERG queried this with the manufacturer but no further details were provided. Also, no formal quality assessment was undertaken on the comparator trials. The ERG's comments on the MS appraisal of validity assessment of the Globe trial can be seen below, although there is no published trial paper with which to compare and verify. The recent Globe publication[13] does include some additional information about the methodology of the trial, but does not add anything further regarding the MS quality assessment. Furthermore, there are some minor differences in results reported for some of the outcomes.

- How was allocation concealed?

The MS provides details of how concealment of treatment allocation was achieved. The trial used a double-blind, double-dummy procedure, with identical capsules and placebo tablets in indistinguishable packaging. To maintain blinding throughout the trial and to facilitate study drug dispensation through the interactive voice response system (IVRS), study drugs were packaged into uniquely numbered kits, each containing two bottles: one with active or placebo telbivudine tablets and one with active or placebo lamivudine capsules. The study drug was dispensed to patients on a schedule that would ensure uninterrupted dosing throughout the treatment interval. The investigators and personnel involved in monitoring remained blinded throughout all the study periods. If a patient's safety was at risk, an emergency coded break using the IVRS could be performed by either the global or the sponsor medical monitor. The ERG assessment of the methods of allocation concealment is adequate.

- What randomisation technique was used?

Using the IVRS for randomisation, patients were stratified by positive or negative HBeAg status and serum ALT levels (above or below 2.5 x ULN). Stratums were randomised using block sizes of four, with a 1:1 ratio across the two treatment groups. This method appears to be appropriate.

- Was a justification of the sample size provided?

Justification for the sample size and power percentages are provided (p.37). The MS states that the trial was adequately powered overall and for the subpopulations separately for the primary outcome of therapeutic response. Calculations appear to be standard and appropriate. The MS states that the RCT was not powered to detect differences in all histological parameters (p.49) and also acknowledges that there was an over-representation of HBeAg-positive patients (p.43).

- Was follow-up adequate?

The MS states that follow-up was adequate. However, treatment was for 104 weeks, with no follow-up period after treatment had ceased. (The MS reported that patients were subsequently enrolled into an extension study for a further 2 years to assess long-term efficacy and safety (ongoing Study 022)). A consort flow chart is provided in the MS on p.33 giving the numbers of drop-outs with reasons, and the numbers analysed. The MS states that discontinuations were clearly documented at each stage. However, totals of treatment discontinuations are unclear as reported in Section 3.1.2.1.

- Were the individuals undertaking the outcome assessment aware of allocation?

The MS does not report *per se* if outcome assessors were independent or blinded, but it would appear that those undertaking the assessments of outcomes were unaware of the allocation to the randomised groups. Conversely, patients who received the wrong study medication were to be analysed according to the group to which they were randomized. The MS does not explain how these patients would be identified.

- Was the design parallel-group or crossover?

The Globe trial was of a parallel design. Two crossover trials (Studies 018[9] and 019[14]) identified in searches were excluded.

- Was the RCT conducted in the UK?

The Globe trial was conducted in 112 centres in 20 countries, with three centres in the UK. The MS does not address differences in clinical practice and it is therefore difficult to establish how this may differ between countries in the Far East and the UK, for example. The clinical advisor to the ERG states that there are vast differences in the management of CHB in different centres and different countries.

- How do the included RCT participants compare with patients who are likely to receive the intervention in the UK?

The MS reports that patients in the trial were largely of Asian origin and that this population is relevant to the UK, as high proportions of CHB cases occur in migrant communities within the UK. The ERG clinical expert advisor concurred that many new cases in the UK are immigrants from Eastern Europe and the Far East. The ERG therefore agrees that the patients of this trial are largely representative of UK CHB patients. The report provides a breakdown of efficacy results by treatment and race, but there is no statistical comparison of results between races. In addition, the MS does not report whether the ethnicity sub-group analysis was planned and hence whether the study was powered for this analysis.

- Were the study groups comparable?

The MS states that there were no significant differences between study groups, but omits to provide *p* values for baseline characteristics (although these are reported in the full trial report). Furthermore it is not stated whether the 'age in years' is given as mean or median (MS Figure 3, p. 32). The ERG requested clarification from the manufacturer who responded that age was reported as mean values. Baseline ALT levels were not reported in the MS and given that a sub-group analysis was subsequently reported, the ERG requested this information from the manufacturer. These were provided along with *p* values for comparison between treatments (see Appendix 1). The proportion of patients with ALT ≥ 2 x ULN was 64% in the HBeAg-positive group and 57% in the HBeAg-negative group. It is acknowledged that patients in the HBeAg-positive sub-groups are around 10 years younger, have higher HBV DNA levels and a shorter mean duration of years since diagnosis than patients in the HBeAg-negative sub-group. The MS suggested that the age and HBV DNA difference is 'consistent with the natural history of CHB' (p. 31).

- Were the statistical analyses used appropriate?

According to the MS, the RCT was intended to demonstrate effects in both HBeAg subpopulations or in the pooled population (p. 37). No analysis for the pooled population was presented due to 'statistical interaction between treatment effect and HBeAg sub-group'. The analysis tested for non-inferiority first, followed by treatment superiority within each population. The ERG asserts that this would appear to be appropriate. The MS performed a sub-set analysis of a sub-group of key efficacy parameters in the 'elevated ALT population' (all patients in the ITT population with ALT screening values ≥ 2 x ULN). The justification of this sub-set analysis was that it would 'allow for comparisons to historical results from interferon treatment studies' (p.38). It is unclear whether the trial was powered for this analysis. Section 3.1.4 of the ERG report contains further detail on the statistical analyses.

- Was an intention-to-treat (ITT) analysis undertaken?

An ITT analysis was carried out for the 104-week efficacy data, restricted to randomised patients with at least one dose of study medication and a minimum of one observation after baseline. The histologic response analysis used a modified ITT (MITT) which was defined as all patients in the ITT population with evaluable pre-treatment liver biopsies at week 52.

- Where there any confounding factors that may attenuate the interpretation of the results of the RCT(s)?

The MS states there were a number of confounding factors. Baseline histologic changes compared at one and two-year intervals may have been confounded by histological disease progression, due to baseline biopsies taken up to 12 months prior to the start of the trial. Furthermore, the RCT was not powered to detect differences in all histological parameters (MS p. 49). The over-representation of HBeAg-positive patients may have influenced the results of the HBeAg-negative sub-group (p. 43), whilst lower than expected numbers of HBeAg-negative patients (with their limited ability to meet the primary efficacy endpoint) may have impacted on the efficacy results. All these factors are acknowledged in the MS report.

- For pharmaceuticals, what dosage regimens were used in the RCT? Are they within those detailed in the Summary of Product Characteristics?

Patients received 600mg telbivudine (and 100mg lamivudine), as per the license.

3.1.4 Description and critique of manufacturer's outcome selection

Overall, the outcomes would appear to be appropriate. As outlined in Section 2.3.4, the majority of outcome measures listed in the decision problem are present in the only included study in the MS report, apart from TTT, HRQoL and survival. The MS notes that clarification about survival is needed (p.8), but no further information is provided in the report. TTT is not a commonly reported outcome in hepatitis B studies.

The primary outcome is 'composite therapeutic response', defined as suppression of HBV DNA <5 $\log_{10}$ copies/ml plus either clearance of detectable HBeAg or ALT normalisation. Outcome measures are reported separately for HBeAg-positive and HBeAg-negative patients throughout the MS, with the exception of the safety results. However, results for telbivudine versus lamivudine for the whole population are presented in the full clinical trial report (Table 14.2.1.2, p.409), although the ERG has not systematically checked or analysed these.

There appeared to be a discrepancy for secondary outcomes, in that in the power calculation (MS Section 5.3.6, p.39) histological response is the key secondary efficacy endpoint. However, on p.35 it is stated that antiviral efficacy (HBV DNA level) is the key secondary efficacy endpoint. Clarification was sought from the manufacturer and the response stated that the term "key" was used to describe multiple secondary endpoints and results, and may have been used inappropriately, as is it "probably not helpful to assign one of these endpoints as more important than the other" (Appendix 1).

The MS reports HBeAg seroconversion rates for a sub-set of patients with ALT ≥ 2 x ULN. The MS suggests that they represent 'the majority of HBeAg patients selected for treatment in the UK' and that these patients are recommended for treatment in the AASLD, APASL and EASL guidelines. The MS states that this sub-group represented nearly 70% of the HBeAg-positive ITT population in the Globe study. However, this proportion was actually 64% based on baseline ALT levels provided by the manufacturer in their response to clarification questions.

Adverse events are reported, but limited to selected events.

3.1.5 Description and critique the statistical approach used

Achievement of therapeutic response (the primary outcome) was reported as proportion of patients (numbers and %), % absolute difference with 95% CI and p values. The majority of the secondary outcomes were presented as proportions of patients (%) only, with *p* values. The difference (%) and 95% confidence intervals are reported in the full trial report but not the MS.

The ITT population was defined as comprising all randomised subjects, presumed to have had at least one dose of study medication and at least one baseline observation. A true ITT analysis should include all randomised patients, regardless of having received treatment. This analysis excluded six randomised patients that failed to return for baseline visits and therefore did not receive any study drug. Three other patients are reported as not having any baseline data. Missing data was treated as “no response” in the ITT analysis.

The data of patients having used prohibited medication was censored at the day it was taken, but included in the ITT analysis. Clarification from the manufacturer stated that 21 patients were censored (12 for lamivudine & nine for telbivudine) (Appendix 1). Post-treatment endpoint values for patients discontinuing treatment for efficacy were to be summarised separately.

The Globe study had an over-representation of HBeAg-positive patients, which ‘may have influenced the power of the study to show statistically significant differences in the HBeAg-negative sub-group, as stated in the MS (p. 43). Sample size calculations were based on 600 (minimum 600, maximum 800) HBeAg-positive patients and 400 (minimum 400, maximum 600) HBeAg-negative patients, but the study recruited 900 HBeAg-positive and 467 HBeAg-negative patients (p.43). The ERG would question whether the power of the HBeAg-negative sub-group was affected as the minimum number of HBeAg-negative patients was exceeded.

The results of the key efficacy points were broken down for sub-groups (HBeAg-positive and negative patients) by treatment and race/ethnicity and were reported in percentages. *P* values were reported, but no statistical comparison was carried out between Asian and Caucasian patients. The sub-group analysis on key efficacy parameters for HBeAg-positive patients with ALT ≥ 2 x ULN was undertaken, but it is unclear whether the trial was powered for this analysis. No statistical comparison between treatment groups was made for the safety data/adverse events.

Indirect comparison

The indirect comparison with entecavir (also compared to lamivudine) carried out by the manufacturer is methodologically poor and should be treated with caution. As only one telbivudine study was included in the MS, no meta-analysis was undertaken. If the MS had included Study 015[3], a meta-analysis may have been possible. However, heterogeneity between this study and the Globe trial may have prevented the studies being combined.

The two identified entecavir trials had different patient populations (HBeAg-positive in one trial[15] and HBeAg-negative in the other[16]) and there were some differences between the telbivudine and entecavir studies (see Section 3.3.2 for further details). The manufacturer states (p.65) that since only one trial is available for each comparator (the Globe study for telbivudine; and only one study in HBeAg-positive patients and one in HBeAg-negative patients for entecavir), no meta-analysis could be undertaken for either and therefore a formal indirect comparison is not valid (citing the Glenny and colleagues HTA Monograph on Indirect Comparisons[17]). However, if the RCT of entecavir vs lamivudine by Yao and colleagues (2007)[18] reporting results for HBeAg-positive and negative Chinese patients separately had been included, a formal indirect comparison may have been possible.

Despite the manufacturer stating that a formal indirect comparison would not be valid given the lack of meta-analyses, a statistical indirect comparison is nonetheless conducted (MS Appendix A). It is the opinion of the ERG that this approach is not valid. Results are reported as log relative risks between entecavir and telbivudine for histologic improvement, HBV DNA detectability, ALT normalisation, HBeAg loss and seroconversion.

A random effects model was used for most outcomes analysed 'to allow for heterogeneity between studies'. However, for HBeAg loss and seroconversion a fixed effects model was used due to there 'only being two trials (with four arms) to provide data'. The ERG considers this inappropriate.

3.2 Summary statement of manufacturer's approach

The manufacturer's approach identified all relevant trials which met their inclusion criteria, albeit there were some discrepancies between the scope and the decision problem. Study inclusions

and exclusions were very unclear in places and difficult to follow. Search strategies were provided in appendices and were geared towards precision rather than sensitivity. However, the ERG did not identify any additional relevant RCTs.

According to the MS, the only trial to meet the inclusion criteria was the registration trial for telbivudine, although the RCT by Hou and colleagues[3] (Study 015) would appear to meet the inclusion criteria. There was no inclusion criteria set for the entecavir comparator trials. Although both comparator trials were randomised comparisons of entecavir versus lamivudine, there were some differences in the study outcomes and baseline characteristics of the populations compared to the telbivudine trial (see Section 3.3.2).

Results were broken down by HBeAg status (positive or negative), study treatment (telbivudine or lamivudine) and (i) race/ethnicity (Asian or Caucasian or Other) or (ii) ALT levels (< or ≥ 2 x ULN). It is not clear whether the study was powered to detect differences in these sub-group sub-sets.

The ERG have assessed the MS for its quality as a systematic review using the questions in Centre for Reviews and Dissemination (CRD) report 4[19] (see Table 3). The process of applying quality criteria was not reported in the MS. Also, no formal quality assessment was undertaken on the comparator trials. Although the manufacturer's quality assessment of the Globe trial was not adequate for some parameters (see Section 3.1.3), overall quality appears to be reasonable, although not very clear in some areas.

Table 3 Quality assessment (CRD criteria) of MS review of telbivudine study

CRD Quality Item: score Yes/ No/ Uncertain with comments	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes. However, the criteria are in an appendix (M) and not in the main submission report. Due to the treatment pathway, the Globe trial only addresses the review question in relation to telbivudine and lamivudine, excluding other comparators issued by the NICE scope such as adefovir.
2. Is there evidence of a substantial effort to search for all relevant research?	Partial. Overall search strategies were adequate but not exhaustive or clear and were poorly documented. NICE's minimum search criteria were not met (PreMedline was not searched) and other databases could have been searched.
3. Is the validity of included studies adequately assessed?	Partial. The Globe trial was adequately assessed, but not the comparator trials included in the indirect comparison.
4. Is sufficient detail of the individual studies presented?	Yes.
5. Are the primary studies summarised appropriately?	N/A – only 1 RCT.

There appeared to be no inclusion criteria for selecting studies for the indirect comparison and details of whether the process was performed by two independent reviewers were lacking. The indirect comparison was weak and should be treated with caution. Even so the MS considers using a statistical approach invalid, it nevertheless proceeds with using just such an approach. Conclusions were drawn from a visual comparison of numbers without statistical evidence. The ERG has reservations about the conclusions drawn.

3.3 Summary of submitted evidence

3.3.1 Summary of results

Results are reported separately for HBeAg-positive and HBeAg-negative patients throughout the manufacturer's submission, apart from safety data. Results are presented for week 52 and week 104 but the MS focuses on the 2 year data 'for consistency' (p.28). On the whole, telbivudine is significantly more favourable than lamivudine for the majority of efficacy outcomes.

3.3.1.1 Outcome 1 - Therapeutic response

Therapeutic response was the primary endpoint in the Globe study, and was defined as suppression of HBV DNA <5 log copies/mL plus *either* clearance of detectable HBeAg *or* ALT normalisation. For HBeAg-positive patients, telbivudine demonstrated statistical superiority over lamivudine at week 52 (75.3% vs 67%, absolute difference 8.3% (95% CI 2.4, 14.2), p=0.0047). At week 104, the number achieving a therapeutic response was less, but the difference between treatment groups was greater and the statistical superiority of telbivudine was maintained (63.3% vs 48.2%, absolute difference 15.1% (95% CI 8.6, 21.6), p<0.0001). For HBeAg-negative patients, there was no statistical difference between treatment groups at week 52 (74.8% vs 77.2%, absolute difference -2.4% (95% CI -10.6, 5.7), p=0.5433), but telbivudine was significantly better than lamivudine at week 104 (77.5% vs 66.1%, absolute difference 11.4% (95% CI 2.9, 19.9), p=0.0069). In both of the tables (MS Table 5, p.44 and the second table on p.47 lacking a table heading), the word '*p* value' is followed by an asterisk, but there is no footnote to offer an explanation.

There is no discussion in the MS of whether the difference in therapeutic response is clinically significant to CHB patients.

3.3.1.2 Outcome 2 – Viral response

The mean reduction from baseline in HBV DNA level was reported to be significantly better for telbivudine patients compared to lamivudine patients at week 52 (-6.45 vs -5.55 $\log_{10}$ copies/mL, p<0.0001) and week 104 (-5.74 vs -4.42 $\log_{10}$ copies/mL, p<0.0001) for HBeAg-positive patients. The same trend was seen in HBeAg–negative patients (-5.22 vs -4.40 $\log_{10}$ copies/mL, p<0.0001 at week 52; -5.00 vs -4.17 $\log_{10}$ copies/mL, p=0.0002 at week 104).

Similarly, the proportion of patients with non-detectable HBV DNA (<300 copies/mL) was greater in patients receiving telbivudine compared to those receiving lamivudine at week 52 and week 104 for both HBeAg-positive and HBeAg–negative patients.

3.3.1.3 Outcome 3 – Biochemical response

ALT normalisation was defined as the proportion of patients with elevated ALT levels at baseline (>ULN) who return to ALT within normal limits. In HBeAg-positive patients, ALT normalisation at year one was similar for telbivudine and lamivudine patients (76.8% vs 74.3%, p=0.3776). At year two, there was a statistically significant difference in favour of telbivudine in the proportion of patients achieving ALT normalisation (69.5% vs 61.7%, p=0.0135). In HBeAg-negative patients, there was no difference between treatment groups at either year one (72.9% telbivudine vs 77.8% lamivudine, p=0.2466) or year two (77.8% telbivudine vs 70.1% lamivudine, p=0.0725).

3.3.1.4 Outcome 4 – HBeAg loss/seroconversion

This includes HBeAg loss and HBeAg seroconversion and is only applicable to HBeAg-positive patients. On p.45 the MS reports that for both the HBeAg loss and HBeAg seroconversion outcome measures ‘telbivudine shows proportional advantages over lamivudine with results meeting non-inferiority criteria’. However, this statement is unsupported statistically as there

was no significant difference between telbivudine and lamivudine patients at either week 52 or week 104.

The MS also presents data for a sustained HBeAg loss of 84% vs 89% for telbivudine and lamivudine respectively, and likewise a sustained HBeAg seroconversion of 83% vs 88% for telbivudine and lamivudine respectively. These are not presented in the table of results (MS Table 6, p.45) and no *p* values or confidence intervals are provided. Sustained response normally refers to six months post-treatment. The ERG sought clarification from the manufacturer on their definition of 'sustained', and requested statistical support of these values. In their response the manufacturer defined a sustained response as that being 'documented on at least two consecutive post-treatment visits and at the last post-treatment study evaluation with no two intervening, consecutive, disqualifying values' (see Appendix 1). However, the time-scale relating to these post-treatment visits was not given. The manufacturer stated that statistical analysis was not possible due to the small total patient numbers and the fact that durability was not a secondary endpoint. The ERG is therefore unable to critique the data presented.

3.3.1.5 Outcome 5 - Virologic breakthrough

Virologic breakthrough was reported in two ways: per protocol and 1 log above nadir. Per protocol virologic breakthrough was defined as an increase in HBV DNA to ≥ 5 $\log_{10}$ copies/mL on two consecutive occasions in patients who had previously achieved post-baseline virological response (i.e. 2 values < 5 $\log_{10}$ copies/mL). 1 log above nadir virologic breakthrough was defined as a confirmed HBV DNA increase of ≥1 $\log_{10}$ copies/mL above nadir HBV DNA (the lowest post-baseline HBV DNA level achieved) in those patients with a confirmed treatment response (i.e. ≥ 1 log reduction in HBV DNA). This outcome was measured at week 48 and week 104.

Virologic breakthrough (both types) was significantly less common with telbivudine compared to lamivudine at weeks 48 and 104 for both HBeAg-positive and HBeAg-negative patients.

3.3.1.6 Outcome 6 - Viral resistance

Treatment emergent HBV resistance was defined as virologic breakthrough with evidence of genotypic resistance-associated mutations. As for virologic breakthrough, HBV resistance was

reported in two ways - per protocol and 1 log above nadir, and was measured at weeks 48 and 104.

HBV resistance occurred in significantly less patients who received telbivudine compared to lamivudine at weeks 48 and 104 for both HBeAg-positive and HBeAg-negative patients. The data reported for HBV resistance in Table 6 (p.45) and Table 8 (p.48) of the MS have superscripts which refer to the source of the data (IDX-07-206), but these are not provided in the reference list and no further details are given. The MS acknowledged that genotyping is not conducted in routine clinical practice and therapeutic management is often based solely on viral rebound (p.46), which in the opinion of the ERG's clinical expert is unacceptably high.

3.3.1.7 Outcome 7 - Histologic response

Histologic assessment encompassed two outcome measures: (a) histologic response defined as ≥ 2 point reduction from baseline in Knodell necroinflammatory score without a worsening in Knodell fibrosis score; (b) Ishak fibrosis score where improvement was defined as a ≥ 1 point reduction from baseline in Ishak fibrosis score. Histologic assessment was conducted at baseline and week 52 only owing to the invasive nature of the liver biopsy procedure, and was performed in the mITT population (see Section 3.1.5). Results for both HBeAg populations are presented in Table 9 (p.50).

The proportion of patients showing an improvement in histologic response was statistically significantly greater in the telbivudine group compared to lamivudine for HBeAg-positive patients (64.7% vs 56.3% respectively, p=0.01), with little apparent difference between the HBeAg-negative patient groups (66.6% vs 66.0%, no p value provided). There was no significant difference between treatment groups in the number of patients showing no improvement. The MS states that the study was not powered to detect differences in all histological parameters and p values are only presented for improvement in Ishak fibrosis score (p.49). However, a p value was reported for the proportion of patients showing a histologic improvement. 95% CI are reported in the full clinical trial report but not in the MS.

For both HBeAg-positive and HBeAg-negative patients, there were no statistically significant differences between treatment groups for the proportion of patients showing an improvement, no change or worsening of Ishak fibrosis score. The MS state that the impact of anti-viral

therapy on liver pathology is not as immediate as it would be on viral load and 52 weeks may be too early for observation of clinically relevant changes in pathology. The ERG's clinical advisor would agree with this.

3.3.1.8 Outcome 8 - Elevated ALT sub-group

An analysis was carried out on a sub-set of the HBeAg-positive patient subgroup with ALT elevated to ≥ 2 x ULN. Results are presented for week 104 only, except for histologic response which was week 52 only (MS Table 7, p.47). Telbivudine was statistically superior to lamivudine for therapeutic response, HBV DNA reduction, HBV DNA non-detectability, ALT normalisation, virologic breakthrough (>1 log above nadir) and histologic response. HBeAg seroconversion in this sub-group at week 52 was presented in Table 6 (p.45) and showed no statistically significant differences between treatment groups. At week 104, statistical superiority was achieved for telbivudine where no significant difference had been evident in the overall HBeAg population. HBeAg loss and HBV resistance were not reported in the MS for this sub-set of a sub-group. Results should be treated with caution, as it is not clear whether the study was powered for a sub-set analysis of a sub-group.

3.3.1.9 Outcome 9 - Ethnicity sub-group

The MS presents key efficacy results by treatment and race/ethnicity (Asian, Caucasian and Other) at week 104 in HBeAg-positive (MS Table 3, p.40-41) and HBeAg-negative patients (MS Table 4, p.41-42 (not Table 3 as incorrectly reported by the MS)). The MS states that 'results from Caucasian patients were similar to Asian patients' (p.40). This statement is unsupported by any statistical comparison, with the results between ethnicities compared only visually.

The MS further adds that within the Asian and Caucasian sub-groups, in both HBeAg-positive and HBeAg-negative patients, 'telbivudine exhibited consistently better outcomes for therapeutic response, HBV DNA reduction, HBV non-detectability and virologic breakthrough, with similar or better results for histologic response' (p.40). However, the ERG disagrees with this statement as, with the exception of HBV DNA non-detectability, there was no statistically significant difference between telbivudine and lamivudine for any efficacy outcome in HBeAg-positive Caucasian patients. In HBeAg-negative Caucasian patients, telbivudine was statistically

superior for HBV DNA reduction, HBV DNA non-detectability and virologic breakthrough, but there was no significant difference for the primary outcome of therapeutic response, nor ALT normalisation or histologic response.

In the third ethnicity sub-group entitled 'Other', there were no statistically significant differences between telbivudine and lamivudine for any outcome measure in HBeAg-positive patients; significant differences were observed for only two outcome measures (HBV DNA non-detectability and virologic breakthrough), and borderline significance for therapeutic response (p=0.0532), in HBeAg-negative patients.

3.3.1.10 Adverse events

Telbivudine appears similar to lamivudine in terms of adverse events. The MS presents data from the Globe study[1] on the incidence of selected grade 2-4 (moderate to severe) clinical adverse events (MS Table 21, p.68), selected grade 3-4 laboratory abnormalities (MS Table 22, p.69) and on-treatment ALT flares (MS Tables 24 & 25, p.71). Table 21 has a footnote which states that 'upper respiratory infection, pharyngitis/ nasopharyngitis, post-procedural pain, influenza and influenza-like symptoms and laboratory abnormalities that were considered adverse events were excluded', but provides no explanation as to why. The proportion of patients experiencing any moderate to severe adverse event was the same in both treatment groups (22%), with no differences between groups for specific adverse events either. Neither Table 21 or 22 nor the accompanying text state at what time point (week 52 or week 104) these results allude to. No statistical comparison is provided, and the MS does not provide any narrative discussion of the results in Table 21.

The MS reports that creatine kinase (CK) elevations were more frequent among telbivudine subjects (9% vs 3% telbivudine vs lamivudine respectively). The MS also reports that the incidence of ALT flares was slightly higher in the lamivudine arm (5.1%) compared to telbivudine (3.2%), but no p values are provided so it is not clear whether this difference is significant.

Table 23 (p.70) lists the frequency of adverse reactions at week 52 from the Summary of Product Characteristics (SPC) of the drug, however without numerical data. Therefore, this has not been checked by the ERG.

In the executive summary (MS Section 3, p.14), the MS reports that the incidence of serious adverse events was 4.9% for telbivudine patients compared to 6.4% in lamivudine patients. The telbivudine adverse event profile was stated to be similar to lamivudine with the majority of patients in each group reporting at least one adverse event (81% telbivudine vs 77% lamivudine). No statistical comparison is provided.

The MS states that patient discontinuations for adverse events, clinical disease progression or lack of efficacy [at week 104] were 0.6% for telbivudine and 2.0% for lamivudine (p.67). However, according to the values in Figure 4 (p.33), the discontinuation rates should be 1.6% (11/680) for telbivudine and 4.1% (28/687) for lamivudine. Upon highlighting this error in the ERG's clarification questions, the manufacturer acknowledged that the rates had been incorrectly stated in the original submission. The MS does not report whether there were any deaths (related or unrelated to study drugs) during the Globe trial. However, the full clinical trial report states that there was one death (lamivudine) during the first 52 weeks, but this was unrelated to the study drug.

The manufacturer has made no attempt to compare the data on safety between telbivudine and the comparator (entecavir) either narratively or quantitatively. From the data presented, it is impossible to compare the safety of telbivudine with entecavir, because for telbivudine results are presented by treatment group only (regardless of HBeAg status), whilst for entecavir results are reported for HBeAg-positive and HBeAg-negative patients separately.

The MS does not provide an overall summary of the adverse event profile of telbivudine treatment. The recent publication of week 52 results provides a greater breakdown of adverse events, giving patient numbers, but again no statistical comparison is attempted.[13]

3.3.1.11 Health related quality of life

The MS did not report any data for quality of life.

3.3.2 Critique of submitted evidence synthesis

Given that there were no head to head trials of telbivudine versus entecavir, the manufacturer carried out an indirect comparison (entecavir was also compared with lamivudine), which in the opinion of the ERG, had a number of methodological limitations and should be treated with caution.

The MS provide an inadequate description of the methodology of the indirect comparison. There is no QUOROM flow chart, but a table is presented (MS Table 10, p.52-53) listing eight identified studies with a comment regarding their suitability, or not, for inclusion. No inclusion/exclusion criteria are reported, and the strategy for the identification and selection of studies, and the data abstraction, are not described.

Two RCTs were identified for inclusion in the indirect comparison: (1) Chang and colleagues comparing entecavir versus lamivudine in HBeAg-positive patients;[15] (2) Lai and colleagues comparing entecavir versus lamivudine in HBeAg-negative patients.[16] A comparative summary of the characteristics of the two included RCTs, the eligibility criteria, the participants and the outcomes are provided in Tables 12-15 of the MS. Those data have been checked by the ERG and appears to be in line with the original published papers. On p.55, the MS states that the comparator arms in the entecavir studies were lamivudine 600mg once daily – this should be 100mg once daily.

Differences between the telbivudine and entecavir trials include:

- Primary outcome – this was histologic response in the entecavir trials, and therapeutic response in the telbivudine trial. Histologic response was a secondary outcome in the Globe study;
- Primary analysis end point – this was 48 weeks in the entecavir trials and 52 weeks (and also 104 weeks) in the telbivudine study, although histologic response was measured at week 48 in the telbivudine trial;
- Racial composition of patients – the proportion of Asian patients was higher in the telbivudine trial, with the entecavir studies having a higher proportion of Caucasian patients;
- HBV genotype – there was a greater proportion of patients with HBV genotype C and fewer with HBV genotype A in the telbivudine trial compared to the entecavir studies;

- Previous IFN therapy – fewer patients in the HBeAg-positive group of the telbivudine study had previously received IFN therapy.

The MS acknowledges that 'these differences highlight the potential pitfall of between trial comparisons' (p.65). The higher proportion of Caucasians in the entecavir studies is noteworthy given that, in the Globe trial, the efficacy of telbivudine was not so pronounced in Caucasians, with no statistically significant differences between telbivudine and lamivudine for most outcomes in this subgroup (MS Tables 3 & 4, p.40-42).

A 'naïve indirect comparison' is reported by the MS in the form of key efficacy outcomes from the telbivudine and entecavir trials presented side by side in a table (MS Table 19, p.64) with conclusions drawn from a visual comparison of numbers. Results for histological improvement are sourced from the 'Telbivudine SPC, June 2007' and are higher than the values reported in the Globe study and presented in Table 9 (p.50) of the MS. There is no explanation as to why the SMPC values were reported rather than those from the Globe study. The MS concludes that telbivudine performs as well as entecavir, is superior to lamivudine and is therefore a cost-effective option. It is the view of the ERG that this statement is unfounded given that the comparison is merely visual and no cost-effective data has been reported at this stage in the submission report. However, the clinical advisor to the ERG asserts that telbivudine does appear to be as effective in viral suppression as entecavir.

The MS then presents a statistical comparison having previously stated that this would be inappropriate given the lack of meta-analyses (this is discussed further in Section 3.1.5). It is the view of the ERG that this statistical approach is invalid. For each outcome there were no statistically significant differences between telbivudine and entecavir.

The MS reports that 'it is impossible to compare resistance rates of telbivudine and entecavir at two years because the data for entecavir is not based upon the ITT population' (p.64). The ERG would highlight that the Chang[15] and Lai[16] published papers both report that no viral resistance to entecavir was detected at week 48. It is the opinion of the clinical advisor to the ERG that this is clinically very relevant given the high resistance rates reported for telbivudine.

3.4 Summary

On the whole, the manufacturer's submission report appears to represent an unbiased estimate of the anti-viral treatment effect of telbivudine. These findings are based on the two-year unpublished results of a single RCT, generally judged to be of reasonable quality when using NICE quality assessment criteria.

In their interpretation of the clinical evidence (MS Section 5.9, p.73), the manufacturer states that 'the evidence from the Globe study demonstrates the significant benefits of telbivudine on the key outcomes of interest'. The clinical advisor to the ERG suggests that whilst the results of the Globe study are statistically significant, they are not clinically significant. When the proportions of patients who discontinued treatment due to disease progression or lack of efficacy are considered (0.8% vs 2.6% for telbivudine and lamivudine respectively), there is an effectiveness advantage of only about 2% in patients treated between the two drugs. Furthermore, no mention is made of the high viral resistance rate and the clinical impact this has on CHB patients. The ERG's clinical advisor stressed that a virological breakthrough of 28.6% (for telbivudine vs 45.5% lamivudine) at two years is unacceptably high at a clinical level.

The MS reports that 'the Globe study provides evidence for 24-week HBV DNA as an early predictive marker of treatment success on nucleoside monotherapy' (p.74). However, no data are provided in the report to support this statement.

The ERG have particular reservations regarding the indirect comparison. The conclusions are based on a visual comparison of efficacy outcomes, and a statistical indirect comparison which the MS states is not considered valid in the absence of any meta-analyses. Resistance rates at one year reported in the published comparator studies are omitted from the MS.

There is additional supporting evidence presented in the MS executive summary (p.14), which was excluded by the MS when identifying relevant studies for inclusion. The supporting evidence for telbivudine was an RCT (Hou and colleagues[3]), which would appear to meet the manufacturer's inclusion criteria. The trial reported that treatment with telbivudine for 52 weeks provided significantly greater antiviral and clinical efficacy than lamivudine. Viral resistance was approximately half that observed with lamivudine but the difference was not statistically different. The findings are supportive of the efficacy and safety of telbivudine.

4 ECONOMIC EVALUATION

4.1 Overview of manufacturer's economic evaluation

The manufacturer's submission to NICE includes a report of an economic evaluation undertaken by the manufacturer, for the NICE STA process. The cost-effectiveness of telbivudine for patients with CHB whose ALT ≥ 2 x ULN, but who have not developed cirrhosis, is estimated using two different economic models. The first, referred to as the viral load model, is the manufacturer's preferred approach and is used for both HBeAg-positive and HBeAg-negative patients. The second – seroconversion model – has been included for consistency with the recent HTA report of adefovir and pegylated interferon[10] (used to underpin the NICE appraisal) and is for HBeAg-positive patients only. The comparators used differ between the two models and are discussed in section 6.2.3, page 78-79 of the MS. Neither model includes all comparators identified in the scope issued by NICE – in particular neither model includes entecavir. The comparators included in the seroconversion model are lamivudine and adefovir dipivoxil (alone or with adefovir, lamivudine or telbivudine as a salvage strategy for patients who develop resistance) and best supportive care, while lamivudine is the only comparator considered in the viral load model. The results of the economic evaluation are presented as incremental cost per QALY gained for telbivudine relative to lamivudine in the viral load model, and for all treatment strategies relative to best supportive care in the seroconversion model.

The ERG requested clarification regarding tables that were reporting errors in the viral load models submitted by the manufacturer. In response the manufacturer submitted a revised set of results and updated versions of the electronic models. The original results are summarised in section 4.2.8 of this report, with the revised results in Appendix 1. All ERG analyses presented in this report, from section 4.4 onward, are based on the updated versions of the electronic models.

4.2 CEA Methods

Both the viral load and seroconversion models are based on state transition models. The structure of the models and the methodology used to evaluate the cost-effectiveness are similar to those used in previous economic evaluations of anti-viral treatments for patients with chronic hepatitis B[10,20-23] – although the disease progression model adopted in the viral load model is

more complex than in previous evaluations (see section 4.2.1 and sections 4.4.1 and 4.4.1.1 later in this report).

No deterministic results are reported in the MS. In response to a request for clarification from the ERG the manufacturer cited NICE methods guidance section 5.9.2 to support their decision to report only the probabilistic evaluation of the models (see section C in Appendix 1 of this report). The results of probabilistic analyses are reported as the base case estimates in the main body of the MS, tabulating mean incremental costs and QALYs and a mean ICER (with confidence intervals). Viral load model results are reported in Section 6.3.1, tables 26 and 27 pages: 96-98 of the MS; seroconversion model results are reported in Section 6.3.1, tables 29 to 31 pages: 99-102 of the MS. The main report contains no representation of uncertainty in the probabilistic cost effectiveness estimates (other than percentile-based 95% confidence intervals and Jackknife confidence intervals). Scatterplots on the cost effectiveness plane and CEACs are reported in appendices (MS Appendix J).

4.2.1 Natural history

Patients with HBeAg-positive and HBeAg-negative CHB are modelled separately, with no transitions between these two cohorts (this assumption is common to most models including both patient groups). Transitions from HBeAg-positive to HBeAg-negative CHB are clinically plausible.

The model of disease progression is similar to that adopted in the HTA report[10] and in previous economic evaluations of anti-viral treatment for CHB[20-23]. Five main health states are defined:

1. Chronic hepatitis B
2. Compensated cirrhosis
3. Decompensated cirrhosis
4. Hepatocellular carcinoma
5. Death

A proportion of patients with decompensated cirrhosis will receive liver transplant, resulting in improved life expectancy and quality of life. However they are likely to suffer reactivation of disease following liver transplantation. The chronic hepatitis B and compensated cirrhosis states are further divided by patients' serological status (for both HBsAg and HBeAg), giving a total of ten states in the seroconversion model (including two absorbing states, one for deaths related to

CHB and another for deaths from other causes). In the viral load model these states are further stratified by ALT and viral load, giving a total of 111 states in the viral load model.

Both decompensated cirrhosis and HCC are associated with significantly raised risk of death, attributed to CHB infection. Five year survival for patients with CHB has been estimated at 99-100%, compared with 80-86% for patients with compensated cirrhosis and 28% for patient with decompensation.[24] Median survival for patients with HCC has been estimated at 2.1 to 10 months.[10123} The models are structured to reflect the greater disease-specific risk of death associated with advanced liver disease.

The seroconversion model is structured so that the effects of treatment are to induce HBeAg seroconversion (in HBeAg-positive patients only, by definition). The viral load model is structured so that the effects of treatment are to induce HBeAg seroconversion (in HBeAg-positive patients only), reduce viral load and to normalise ALT. The viral load model is therefore applicable to both HBeAg-positive and HBeAg-negative patients. The benefit of these outcomes are that they are each (to a greater or lesser extent) associated with reduced progression to advanced liver disease.

4.2.2 Treatment effectiveness

Direct evidence of the effectiveness of telbivudine and lamivudine, and of the development of resistance to each drug, was taken from a sub-group of patients with ALT ≥ 2 x ULN in the Globe study. Data applied in the models are reported in Appendix B (for the viral load model) and Appendix F (for the seroconversion model) of the MS. In the seroconversion model the effectiveness of adefovir was assumed to be the average of lamivudine and telbivudine, while data on resistance was taken from the HTA monograph.[10]

The risks of progressive liver disease (compensated/ decompensated cirrhosis and HCC) applied in the seroconversion model were taken directly from the HTA monograph.[10] The risks of developing compensated cirrhosis and hepatocellular carcinoma applied in the viral load model were derived from equations estimated in a Taiwanese population of 3,653 CHB patients, followed for a mean of 11.4 years[25,26] – this study is commonly referred to as the REVEAL-HBV study. The models predict higher risks of developing compensated cirrhosis in male patients with higher viral loads, with elevated ALTs and for patients who have not seroconverted the e

antigen. The models predict higher risks of HCC for male patients with higher viral loads and for patients who have developed compensated cirrhosis. These models are discussed further in section 4.4.1.2.2.

Adverse events are not included in either model.

4.2.3 Health related quality-of-life

Changes in health state were assumed to be associated with changes in quality of life, with progressive liver disease being associated with poorer quality of life. Chronic Hepatitis B, without cirrhosis or other progressive liver disease, was assumed to be associated with minimal reduction in quality of life. Health state valuations (for CHB and progressive liver disease states) were taken from the HTA monograph,[10] although the method of applying these differs between the two models. In the seroconversion model health state utility decrements are applied to age-specific utilities for the general population (as in the model used in the HTA report[10]) whereas multipliers are applied to a constant health state utility value of one (assumed for patients who lose the surface antigen (HBsAg negative), referred to as the post-hepatitis in the model) in the viral load model (discussed in section 4.4.1.2.3).

4.2.4 Resources and costs

Drug costs are based on licensed dosages using unit costs from the British National Formulary (BNF), Number 54, (September 2007).[27] Non–drug costs for patients receiving anti-viral treatment were derived using treatment protocols reported in the HTA monograph.[10] Health state costs were taken from the HTA monograph[10], inflated to 2005/06 costs using the HCHS inflator.[28]

4.2.5 Discounting

An annual discount rate of 3.5% was applied to both costs and outcomes.

4.2.6 Sensitivity analyses

The MS does not report one-way sensitivity analyses for either the viral load or seroconversion models. The results of probabilistic sensitivity analyses are treated as base case results, with limited discussion of uncertainty in the body of the MS.

4.2.7 Model validation

Approaches to validating the viral load model are described in MS section 6.2.12, page 96. The MS does not report on attempts to establish the model's internal validity (no discussion of consistency checks or checks of coding accuracy). The validation reported is limited to checking the outputs of the model at two years against the trial results, but no detail is provided.

The MS does not report on attempts to establish the seroconversion model's internal validity (no discussion of consistency checks or checks of coding accuracy). The approach to establishing external consistency of this model was to compare results with those in the HTA monograph[10] for the assessment group model and those reported by the manufacturer of adefovir (reported in the MS Appendix K).

4.2.8 Results

Results from the economic models are presented as incremental cost per QALY gained (see MS, section 6.3.1.1, tables 26 and 27 (pages 96-98) for the viral load model and tables 29 to 31 (pages 99-102) for the seroconversion model).

Table 4 below summarises the results from the probabilistic evaluation of the viral load model, including 95% confidence intervals derived using the 2.5 and 97.5 percentiles of the ICER distribution and Jackknife estimates, as reported in the main body of the MS (more details are included in Appendix J of the MS). See Appendix 1 of this report (Table 26 and Table 27) for updated results submitted by the manufacturer in response to the ERG's request for clarification regarding errors in the models originally submitted.

Table 4 Cost-effectiveness results from the viral load model presented in the MS (original submission)

	Telbivudine versus lamivudine	
	HBeAg-positive patients	HBeAg-negative patients
Uninformative prior probability of 0.0		
Mean incremental QALYs	1.30	4.67
Mean incremental cost	£ 19,087	£ 49,003
Mean ICER (percentile based 95% confidence interval)	£ 14,665 (£ 4,345 – Dominated)	£ 10,497 (£ 7,980 – Dominated)
Jackknifed ICER (Jackknife 95% confidence interval)	£ 14,660 (£ 14,184 – £ 15,136)	£ 10,497 (£ 10,401 – £ 10,592)
Uninformative prior probability of 0.5		

Mean incremental QALYs	1.36	0.94
Mean incremental cost	£ 12,664	£ 31,255
Mean ICER (percentile based 95% confidence interval)	£ 9,332 (Dominating – Dominated)	£ 33,300 (Dominating – Dominated)
Jackknifed ICER (Jackknife 95% confidence interval)	£ 9,321 (£ 8,611 – £ 10,031)	£ 33,239 (£ 30,292 – £ 36,186)

Table 5 below summarises the results from the probabilistic evaluation of the seroconversion model, including 95% confidence intervals derived using the 2.5 and 97.5 percentiles of the ICER distribution, reported in the main body of the MS (more details are included in Appendix J of the MS). The mean ICERs reported in the MS (column 4, Table 5) assume that best supportive care is the most appropriate comparator for all anti-viral treatment strategies. Alternative calculations using lamivudine as comparator for telbivudine and using telbivudine as comparator for telbivudine followed by adefovir (identified as their next-best strategies) yield ICERs of £19,680 per QALY gained and £24,277 per QALY gained respectively.

Table 5 Cost-effectiveness results from seroconversion model presented in the MS (results based on a cohort size of 100 patients).

	Compared with BSC			Compared with next best strategy
Treatment strategy	Mean incremental cost £	Mean incremental QALY	Mean ICER £/ QALY (95% CI)	Mean ICER £/ QALY
Best supportive care				
Lamivudine	503,059	63.78	7,887 (3,924 – 16,717)	7,887
Lamivudine then adefovir	1,667,090	113.75	14,655 (8,599 – 25,242)	
Telbivudine	1,529,867	115.96	13,193 (7,788 – 25,194)	19,680
Adefovir dipivoxil	2,136,201	117.63	18,160 (11,490 – 30,160)	
Adefovir dipivoxil then lamivudine	2,247,279	129.17	17,398 (11,063 – 28,322)	
Adefovir dipivoxil then telbivudine	2,512,060	136.61	18,388 (11,707 – 30,357)	
Telbivudine then adefovir	2,345,968	149.58	15,684 (9,491 – 28,151)	24,277
Notes: Next best strategy defined using the cost-effectiveness frontier, calculated by ERG. The order of strategies in this table is different from that in the MS – in this table strategies are ordered in terms of increasing effectiveness (in mean incremental QALY terms, compared with best supportive care).				

4.3 Critical appraisal of the manufacturer's submitted economic evaluation

4.3.1 Critical appraisal of economic evaluation methods

The ERG have considered the methods applied in the economic evaluation in the context of the critical appraisal questions listed in Table 6 below, drawn from common checklists for economic evaluation methods (e.g. Drummond and colleagues[29]).

Table 6 Critical appraisal checklist of economic evaluation

Item	Critical Appraisal	Reviewer Comment
Is there a well defined question?	Yes	
Is there a clear description of alternatives?	Yes	Comparators are clearly described for both models. However the comparators do not match those in the scope for this assessment (see discussion below).
Has the correct patient group / population of interest been clearly stated?	Yes	Patients with (HBeAg +ve or HBeAg –ve) CHB with ALT at ≥ 2 x ULN. Patients with ALT levels ≥ 2 x ULN are the sub-group indicated for anti-viral treatment, but there is limited information on this group in the MS. Patients with ALT at ≥ 2 x ULN comprise 588/921 (64%) of HBeAg-positive patients and 255/446 (57%) of HBeAg-negative patients in Globe study, according to MS – does not agree with data in Table 7 in MS where sub-group of HBeAg-positive patients with ALT ≥ two times the upper limit of normal is 637 (320 LAM and 317 TBV). No discussion of baseline characteristics or proportion of these randomised to TBV or LAM.
Is the correct comparator used?	?	• Not considered interferon (pegylated or non-pegylated) – patients suitable for telbivudine defined as "intolerant" of Peginterferon Alfa-2a (see section 6.2.1, page 77 of MS). 5.6% of HBeAg-positive and 11.2% of HBeAg-negative patients in Globe had previously been treated with interferon[1] • Entecavir not included in models – based on an indirect comparison showing equality of effectiveness. But have ignored issue of likely better resistance profile of entecavir. • In the seroconversion model all treatment options are compared withbest supportive care – see Table 5 for comparisons against next-best alternatives. • Lamivudine is an appropriate comparator for first line treatment of patients who are intolerant of, or have failed on, interferon treatment. Also for patients unwilling to accept interferon treatment.
Is the study type reasonable?	Yes	Cost-utility model is appropriate as quality of life differences are important as well as life expectancy differences. Previous evaluations have shown small differences life expectancy for anti-viral treatment strategies.
Is the perspective of the analysis clearly stated?	Yes	Study perspective stated as that of NICE reference case (MS section 6.2.4, page 79)

Is the perspective employed appropriate?	Yes	• Costs from NHS and PSS perspective. • Outcomes from patient perspective –quality-adjusted life expectancy.
Is effectiveness of the intervention established?	?	• Model used observed data directly from the Globe study for lamivudine and telbivudine. • Adefovir was not included in indirect comparison and its effectiveness in the seroconversion model is based on assumption. No search for effectiveness of adefovir reported and no discussion or quality assessment of evidence that was presented in HTA monograph.[10] • The effectiveness of telbivudine versus entecavir has not been established (particularly, no account was taken of resistance) and was not considered in models.
Has a lifetime horizon been used for analysis (has a shorter horizon been justified)?	Yes	Lifetime horizon is appropriate as CHB is chronic lifetime condition and treatment may be required for lifetime (assuming no complete clearance of virus and/or non-development of resistance).
Are the costs and consequences consistent with the perspective employed?	Yes	• Health state costs and on-treatment management costs taken directly from HTA monograph[10] (uprated to 2004/05 prices). Drug costs are from BNF. • Outcomes from patient perspective – quality-adjusted life expectancy using utility weights from HTA monograph[10]
Is differential timing considered?	Yes	Discounting incorporated in model at appropriate rates (check how they've discounted)
Is incremental analysis performed?	Yes	• Lack of information in MS on total costs and QALYs for both lamivudine and telbivudine in the viral load model makes interpretation of the differences between results with prior of zero and 0.5 very difficult. • Presentation of the seroconversion model does not take into account 5.9.2.2 of NICE methods guidance "Standard decision rules should be followed … any situation where dominance or extended dominance exists". All comparisons are made with best supportive care and not the best alternative comparator. See ERG addition to Table 5 of this report.
Is sensitivity analysis undertaken and presented clearly?	No*	• No deterministic sensitivity analyses are reported for any of the models. • Base case results are means of probabilistic evaluations of the model – very little of the PSA is reported in main body of report. • Discussion of uncertainty in PSA, as presented in main body of report, does not meet NICE reference case. PSA output relegated to appendices.
Note: * PSA was undertaken but not reported in main body of report.		

4.3.2 NICE reference case

Table 7 NICE reference case requirements

NICE reference case requirements (see detail in NICE report)**:**	Included in submission
Decision problem: As per the scope developed by NICE	†
Comparator: Alternative therapies routinely used in the UK NHS	✓
Perspective on costs: NHS and PSS	✓
Perspective on outcomes: All health effects on individuals	✓
Type of economic evaluation: Cost effectiveness analysis	✓
Synthesis of evidence on outcomes: Based on a systematic review	‡
Measure of health benefits: QALYs	✓
Description of health states for QALY calculations: Use of a standardised and validated generic instrument	?
Method of preference elicitation for health state values: Choice based method (e.g. TTO, SG, not rating scale)	?
Source of preference data: Representative sample of the public	?
Discount rate: 3.5% pa for costs and health effects	✓

Notes:
N/A=not applicable
† comparators listed in scope include interferon (pegylated alfa-2a, non-pegylated alfa-2a and alfa-2b), entecavir and adefovir:

- interferon is not included in either of the manufacturer's economic models – stating that "telbivudine is assumed to be a treatment option for patients with Hepatitis B that are intolerant of Peginterferon Alfa-2a".
- entecavir is not included in either of the manufacturer's economic models – based on assumption of equality of efficacy in terms of seroconversion, normalisation of ALT and reduced viral load (adequacy of indirect comparison discussed in section 3.1.5). However comparison does not consider impact of resistance.

‡ no systematic review. Indirect comparison was based on a search for clinical evidence for telbivudine/entecavir/lamivudine. but not adefovir or interferon. Outcomes in the viral load model based on Globe trial. Outcomes in seroconversion model based on Globe and HTA monograph[10] supplemented by assumption on efficacy of adefovir which is not justified by discussion in body of MS
? health state utilities were taken from HTA monograph[10]. No search for updated sources on health state utility for CHB patients. No discussion of source of valuations, with respect to NICE reference case – see section 4.4.1.2.3 this report.

4.4 Modelling methods

An outline critical review of modelling methods has been undertaken. The review has used the framework for good practice in modelling presented by Philips and colleagues[30] as a guide, addressing issues of model structure, structural assumptions, data inputs, consistency, and assessment of uncertainty.

4.4.1 Modelling approach / Model Structure

The MS provides little commentary on the development of the model structure and does not report any clinical validation. However the model structures adopted are consistent with current clinical understanding of the underlying biological process of progressive liver disease associated with CHB infection. The model structures incorporate the effects of anti-viral treatment (inducing HBeAg seroconversion (where relevant), reducing viral load and normalising ALT) in decreasing the probability of progressive liver disease in an appropriate and clinically meaningful way.

The modelling approach and structure adopted for the seroconversion model was based on the approach used in the HTA monograph (in that the MS was seeking to replicate the model). The basic structure of the viral load model is presented in section 6.2.5.1 of the MS, pages 80 to 83, and has briefly been discussed in section 4.2.1 of this report. Schematics for the viral load model (for patients with HBeAg-positive and with HBeAg-negative disease) were not presented clearly in the MS (see Figures 5 and 6, pages 81 and 82 of the MS) and have been reproduced in appendices to this report (see Appendix 2).

While there is some discussion on the relative merits of discrete event simulation and state transition models, there is no discussion of alternative approaches that could have been adopted for populating the model. In particular, there seems to have been no consideration of the possibility of developing statistical risk models (similar to those adopted for modelling progression to compensated cirrhosis and hepatocellular carcinoma) rather than the direct use of events observed in the Globe study. A statistical modelling approach may have reduced the impact of the sparsity of data for some transitions included in the model.

Both models have adopted a lifetime horizon which is appropriate given the chronic nature of the disease and the long-term nature of treatment (particularly for HBeAg-negative patients, since current treatment guidelines suggest that treatment for HBeAg-positive patients could stop following HBeAg seroconversion). No sensitivity analysis is reported for time horizon, although the electronic viral load model contains results using one year and two year time horizons. In both the one year and two year time horizon models lamivudine is dominant for HBeAg-negative patients, while ICERs for HBeAg-positive patients are greater than would conventionally be

considered cost effective (approximately £9 million and £3 million per QALY gained for time horizons of one and two years, respectively).

The horizon in the seroconversion model is fixed at 100 years – since patients enter the model aged 31, this implies an age of 131 years at the termination of the process. However, reducing the time horizon to seventy years (i.e. terminate with maximum age of 100) or forty seven years (average life expectancy assuming 75% of cohort are male and 25% female) had little impact on the cost effectiveness estimates (see Table 8).

Table 8 Impact of time horizon on seroconversion model outcomes (ERG analysis)

	Model horizon = 100 years (Original submission)		Model horizon = 70 years		Model horizon = 47 years	
	Total cost	Total QALYs	Total cost	Total QALYs	Total cost	Total QALYs
Best supportive care	£15,211	15.18	£15,211 (0.00%)	15.18 (0.00%)	£15,092 (0.78%)	14.90 (1.85%)
Lamivudine	£20,438	15.74	£20,438 (0.00%)	15.74 (0.00%)	£20,312 (0.62%)	15.45 (1.88%)
Telbivudine	£30,808	16.25	£30,808 (0.00%)	16.24 (0.00%)	£30,655 (0.50%)	15.93 (1.95%)
Adefovir dipivoxil	£36,983	16.25	£36,983 (0.00%)	16.25 (0.00%)	£36,792 (0.51%)	15.93 (1.96%)
Lamivudine then adefovir	£32,288	16.21	£32,288 (0.00%)	16.20 (0.00%)	£32,078 (0.65%)	15.89 (1.97%)
Telbivudine then adefovir	£39,105	16.56	£39,105 (0.00%)	16.56 (0.00%)	£38,870 (0.60%)	16.22 (2.02%)
Adefovir dipivoxil then lamivudine	£38,121	16.36	£38,121 (0.00%)	16.36 (0.00%)	£37,919 (0.53%)	16.03 (1.99%)
Adefovir dipivoxil then telbivudine	£40,776	16.43	£40,776 (0.00%)	16.43 (0.00%)	£40,543 (0.57%)	16.10 (2.01%)

Notes
The table shows the total cost and total QALYs estimated in the seroconversion model for different model horizons. Figures in parentheses are the percentage change from the values in columns 2 and 3 – deterministic evaluations of the model using a 100 year time horizon.

The viral load model terminates after 135 cycles (equivalent to an age of 98.5 years for HBeAg-positive patients and 107.5 years for HBeAg-negative patients). The reason for this duration is not discussed in the MS. However reducing the horizon, for HBeAg-negative patients, to a maximum age of 100 had little effect (see Table 9). Reducing the time horizon to the average remaining life expectancy for each cohort (assuming 75% of HBeAg-positive cohort and 90% of HBeAg-negative cohort are male) has a disproportionate effect for telbivudine treated patients. However the effect is still small (see Table 9).

Table 9 Impact of time horizon on viral load model outcomes: prior = 0 (ERG analysis of the resubmitted model)

HBeAg-positive cohort	Model horizon = 67.5 years (Original submission)				Model horizon = 47 years	
	Total cost	Total QALYs			Total cost	Total QALYs
Lamivudine	34,214	14.60			34,165 (0.14%)	14.58 (0.14%)
Telbivudine	56,669	16.43			56,519 (0.26%)	16.35 (0.49%)
HBeAg-negative cohort	Model horizon = 67.5 yrs (Original submission)		Model horizon = 60 years		Model horizon = 38 years	
	Total cost	Total QALYs	Total cost	Total QALYs	Total cost	Total QALYs
Lamivudine	£36,417	13.05	£36,417 (0.00%)	13.05 (0.00%)	£36,218 (0.55%)	13.02 (0.23%)
Telbivudine	£77,428	15.06	£77,428 (0.00%)	15.06 (0.00%)	£75,932 (1.93%)	14.84 (1.46%)

The cycle lengths used in the two models differ, with the viral load model using a six month cycle and the seroconversion model using annual cycles. The MS does not discuss the relative merits of these two cycle lengths. One drawback with using a six month cycle for the viral load model is that transitions in the final six months of the trial are used to extrapolate treatment effects beyond the trial duration, therefore subject to maximum data sparsity. A half-cycle correction has been used.

Sources of data used to develop/populate the model structure are specified (primarily HTA monograph[10] and Globe trial for effectiveness of and resistance to telbivudine and lamivudine). Key data on transitions to compensated cirrhosis and HCC are derived from unpublished analyses of data from the REVEAL-HBV study,[25,26] but are not adequately described or critically appraised in the MS (see section 4.4.1.2.2 of this report for further discussion of this).

4.4.1.1 Structural Assumptions

In common with models adopted in previous economic evaluations, loss of the surface antigen (HBsAg) was assumed to be a permanent cure. Patients in this state are assumed to have no risk of spontaneous reactivation of disease – though this does not imply that they have immunity from re-infection with hepatitis B. In contrast, patients losing the e antigen (i.e. HBeAg seroconversion) are at risk of spontaneous reactivation of disease. On the basis of these assumptions treatment (with lamivudine or telbivudine) in the viral load model is continued until patients lose the surface antigen, develop resistance or progress to advanced liver disease.

This approach is clinically plausible, but contrasts with clinical guidelines for patients with HBeAg-positive disease which suggest treatment can be withdrawn four to six months following response (HBeAg seroconversion[6] or virologic response[7]). In the seroconversion model reactivation of disease is more likely in the year following a treatment-induced HBeAg seroconversion, based on assumptions adopted in the HTA monograph.[10] This assumption does not appear to hold in the viral load model, but is not discussed in the MS.

In both models HBeAg-positive patients can e seroconvert and clear the surface antigen in the same cycle. These transitions are based on data observed in the Globe study. However previous evaluations have assumed patients must first clear e antigen and then may subsequently clear the surface antigen.

The MS does not report whether clinical advice was sought during the development of the viral load model and does not discuss whether the complexity of the model is warranted. For example, as discussed in section 4.2.1, each viral load level (1 through 5) is further stratified by ALT (elevated or normal) and HBeAg serological status. This stratification appears to be primarily driven by the statistical models used for estimating probabilities of developing compensated cirrhosis and HCC, while the literature briefly reviewed in section 6.1.2 of the MS only refers to “high viral load as a prognostic indicator for morbidity and mortality from CHB”.

Neither model assumes any continuing treatment effect once treatment has ceased - due to development of resistance or (in the seroconversion model) HBeAg seroconversion. However the models do not assume that patients who have achieved treatment-related seroconversion or ALT normalisation immediately reactivate disease on cessation of treatment.

Patients enter the viral load model with a diagnosis of CHB, without cirrhosis, after initial use of peginterferon alfa-2a (if appropriate) and are distributed across viral levels according to the proportions observed in the Globe trial (see table and text on page 86 of MS). Patients receive anti-viral treatment with lamivudine or telbivudine until they lose the surface antigen, develop treatment resistance, progress to a non-treated state (decompensated cirrhosis, hepatocellular carcinoma, and liver transplant) or die. Once patients develop resistance they are assumed to revert to transition probabilities applied to patients with the greatest viral load.

Patients enter the seroconversion model with a diagnosis of CHB, without cirrhosis, after initial use of peginterferon alfa-2a (if appropriate). Patients receive anti-viral treatment until they lose the surface antigen, achieve HBeAg seroconversion, develop treatment resistance or die. Once patients develop resistance they are assumed to follow disease progression of untreated patients. The seroconversion model assumes no benefit of treatment, in terms of transitions to the inactive carrier (HBeAg seroconverted) state, for patients with compensated cirrhosis – the rate applied here is the spontaneous seroconversion rate of 9%. In the base case model, the benefit of anti-viral treatment for cirrhotic patients is a reduced probability of developing decompensation. The only states in the seroconversion model where patients do not receive anti-viral treatment are the HBeAg seroconverted (other than six months' consolidation treatment in the year following seroconversion) and HBsAg negative ("cured") states. This differs from the viral load model and is not an appropriate assumption for telbivudine which currently has no marketing authorisation for treatment of patients with decompensated disease (see Appendix 4). In contrast, the marketing authorisation for lamivudine and adefovir state that treatment can continue for decompensated patients (see Appendix 4).

4.4.1.2 Data Inputs

4.4.1.2.1 Patient Group

The patient group modelled in the submission are those with CHB whose ALT levels are greater than or equal to two times the upper limit of normal. The MS states that this is the standard criterion for treatment, citing AASLD[6], APASL[8] and EASL[7] guidelines. This population has been used to model the cost-effectiveness of anti-viral treatment for CHB in previous economic evaluations.[20,21] However, as discussed in section 2.3.1 of this report, the scope for this appraisal does not specifically mention this sub-group.

These patients are a sub-group of those recruited to the Globe trial and the MS does not report whether the trial was powered for this sub-group. Section 6.1.1 of the MS reports the number of patients identified as having ALT levels ≥ 2 x ULN who provide data to populate the model (see Table 10), but does not report the breakdown between the lamivudine and telbivudine arms of the trial. The MS does not report baseline characteristics for this sub-group.

Table 10 Globe study patients in sub-group with ALTs ≥ 2 x ULN

	ALT ≥ 2 x ULN	Total
HBeAg-positive CHB	588 (63.8%)	921
HBeAg-negative CHB	255 (57.2%)	446

The number of HBeAg-positive patients providing data for populating the model (as described in section 6.1.1 of the MS and in Table 10 of this report) does not match those identified in the pre-defined elevated ALT population (also referred to as the interferon eligible population, see CSR[1]) discussed elsewhere in the submission (see Table 7 of the MS, page 47 which reports that 320 telbivudine and 317 lamivudine (Total 637) had ALTs ≥ 2 x ULN). There is no discussion in the MS regarding the discrepancy between the number of patients with ALTs ≥ 2 x ULN identified in the analysis reported on pages 46-47 of the MS and the number of patients identified as providing data for the economic models.

The MS reports that randomisation was stratified by serum ALT level, as well as HBeAg status. However the ALT level used (above or below 2.5 times upper limit of normal) does not correspond to that used to define the sub-group providing data to populate the model and would not therefore ensure balance between treatment groups (assuming complete data were available from both groups). Given the discrepancy discussed above, over the number of patients defined as having ALT levels ≥ 2 x ULN in Table 7 of the MS compared with section 6.1.1 of the MS, and concerns over missing data discussed in section 4.4.1.2.2 the ERG feel there is real uncertainty over the completeness of data (from the Globe study) used to populate the model.

The seroconversion model only includes patients with HBeAg-positive disease, with baseline characteristics taken from the HTA monograph.[10] Patients with HBeAg-positive and HBeAg-negative CHB were both included in the viral load model, but were modelled separately with differences in the proportion of males and age at treatment initiation assumed for the two cohorts based on the HTA monograph.[10] Viral load models for patients with HBeAg-positive and HBeAg-negative CHB were structurally identical, except that HBeAg-negative patients could not seroconvert the e antigen. No account seems to have been taken of differences in clinical profile of patients presenting with HBeAg-positive and HBeAg-negative disease. Patients with HBeAg-negative disease tend to present with more advanced disease, with 29-38% presenting with cirrhosis in some studies.[24]

No other sub-groups were modelled. Patients co-infected with HCV, HDV or HIV were excluded from the Globe trial as were patients who had previously received lamivudine or other anti-HBV nucleoside or nucleotide analogue. This is consistent with scope from NICE which stated that the STA would not consider co-infected patients, in line with Technology Appraisal no. 96.

4.4.1.2.2 Clinical Effectiveness

As mentioned, direct evidence of the effectiveness of telbivudine and lamivudine in inducing HBeAg seroconversion (in HBeAg-positive patients), loss of the surface antigen (HBsAg), reducing viral load and normalising ALT were taken from the Globe trial. Transition probabilities were estimated from these data using beta distributions, with parameters as the number of events (α) and the number of non-events (β). The mean of these distributions (α/(α+β)) is therefore the proportion of patients making a given transition between states observed in the trial. A similar approach was taken to derive risks of developing resistance to telbivudine and lamivudine and for reactivating disease (in patients who achieved HBeAg seroconversion).

Transition probabilities in the viral load model were estimated using data for six month periods, within the trial duration of two years (0-6 months, 7-12 months, 13-18 months, 19-24 months), with the values observed in the last six months assumed to apply for any patients treated for greater than two years. There is no discussion in the MS of the appropriateness of a six month cycle, nor of the effect this cycle length has on the robustness of the viral load model.

In the seroconversion model annual transition probabilities were estimated for telbivudine and lamivudine, with the values observed in the second year carried forward for all subsequent years. Transition probabilities for loss of the surface antigen and HBeAg seroconversion with adefovir were estimated as the mean of the modelled values for telbivudine and lamivudine, without reporting any justification for this assumption. Probability of developing resistance to adefovir was taken from the HTA monograph[10] and entered the modelled deterministically, while the probability of reactivating disease was modelled probabilistically.

In the seroconversion model patients who cease treatment, due to resistance, revert to the risks for patients receiving best supportive care (taken from the HTA monograph[10]). In the absence of data on disease progression for patients who cease treatment following the development of resistance transition probabilities for these patients are based on those for patients with the

highest viral load (it is not stated whether they also assumed they would be raised ALT or HBeAg-negative (if appropriate) which are included in disease progression model). This assumption is not discussed in detail in the MS. However, it may questionned whether this group of patients - patients (in the trial) whose viral levels remain high, despite treatment – is an appropriate population to use for estimating transitions for patients who initially responded to treatment but subsequently developed resistance.

In general, there is inadequate discussion in the MS regarding the completeness and reliability of data used to populate the model. Closer inspection of data from the Globe study used in both models reveals some concerning issues over data completeness. For example, while the MS states that 588 HBeAg-positive and 255 HBeAg-negative patients had ALT levels ≥ 2 x ULN transition probabilities for resistance in the first year of treatment appear to have been estimated using total populations of 544 and 236, respectively, in the viral load model (based on data from Tables B29, B32 and B37 in Appendix B of the MS) and a total population of 575 HBeAg-positive patients in the seroconversion model (based on data from Table F2 in Appendix F of the MS). These discrepancies are not discussed in the MS nor is there any consideration of the implications of missing data on the robustness or validity of model outputs. Using data reported in Appendix B and Appendix F of the MS to calculate the proportion of patients developing resistance to lamivudine and telbivudine in the first year of treatment yield different values to those reported in the MS, for all patients in the Globe study (see Table 11). Since the MS did not include efficacy results by treatment arm for patients with ALT levels ≥ 2 x ULN the ERG cannot determine whether these discrepancies are the result of real differences for this sub-group of patients or an artefact of missing data.

Table 11 Viral resistance reported for Globe study and values estimated by ERG from tables in Appendix B and F

	Globe (MS)	Viral load model	Seroconversion model
HBeAg-positive cohort			
Lamivudine	11.0%	9.7%	9.1%
Telbivudine	5.0%	3.6%	3.5%
HBeAg-negative cohort			
Lamivudine	10.7%	5.2%	
Telbivudine	2.2%	0.8%	

It appears, from tables reported in Appendix F of the MS, that no adjustments have been made to denominators in the calculation of probabilities of treatment effect for patients developing treatment resistance during the period of observation. Table 12 shows the input values used to estimate HBeAg seroconversion, in the seroconversion model, for lamivudine and telbivudine. As discussed above, transition probabilities have been estimated using beta distributions parameterised using the number of events (α) and the number of non-events (β), with mean equal to α/(α+β). This assumes that all patients treated during the year can achieve the indicated transition. However it is unlikely that patients who develop treatment resistance will achieve HBeAg seroconversion (or other treatment aims) at the same rate as those who are not resistant. To examine the influence of this assumption on transition probabilities, the denominators were reduced by half the number of patients who were assumed to become resistant during the year – i.e. using figures from Table 12, if 283 patients receive lamivudine in year 1, with 25.73 (283 * 0.0909) becoming resistant in year 1, the denominator is reduced to 270.14 (283 – (270.14 ÷ 2)). Table 12 shows the percentage change in the transition probability brought about by this adjustment – failure to adjust transition probabilities for treatment resistance will only bias the cost effectiveness results if anti-viral agents differ significantly in their resistance profiles. This adjustment is predicated on the assumption that patients who continue on treatment to the year end will not achieve a treatment-induced seroconversion once resistance develops.

Table 12 Adjusting HBeAg seroconversion transition probabilities for resistance

	Year	Parameter		Annual Resistance	TP	$TP_{adjusted}$	% change
		α	β				
Lamivudine	1	68	215	0.0909	0.2403	0.2517	4.8%
	2	21	194	0.2154	0.0977	0.1095	12.1%
Telbivudine	1	75	207	0.0346	0.2660	0.2706	1.8%
	2	33	174	0.1470	0.1594	0.1721	7.9%

Notes:
These “adjusted” transition probabilities are applied in re-estimated seroconversion models in section 4.4.1.4.2 (ERG sensitivity analysis) 4.4.1.4.4 (ERG scenario analysis).
This concern also applies in the viral load model. However the stratification of input data across viral load levels and complexity of the model makes similar analyses unfeasible within the constraints of this review.

The risks of progressive liver disease (compensated/ decompensated cirrhosis and HCC) applied in the seroconversion model were taken directly from the HTA monograph[10]. The risks of developing compensated cirrhosis and HCC applied in the viral load model were derived from equations estimated in a Taiwanese population of 3,653 CHB patients, followed for a mean of

11.4 years[25,26] – this study is commonly referred to as the REVEAL-HBV study. The models predict higher risks of developing compensated cirrhosis in male patients with higher viral loads, with elevated ALTs and for patients who have not seroconverted the e antigen. The models predict higher risks of developing HCC for male patients with higher viral loads and for patients who have not seroconverted the e antigen. The MS reports that these equations were "adjusted" to include patient age in both equations and cirrhosis in the HCC model and "re-calibrated" so that the average probability of developing compensated cirrhosis and HCC matched the original data. However no further details of this adjustment and calibration were provided and the analysis does not appear to have been published or peer reviewed so the ERG can make no judgement on the appropriateness or validity of the adjustment procedure.

As discussed in section 4.4.1.1, beneficial effects of continued telbivudine treatment have been assumed, in the seroconversion model, for non-resistant patients with decompensated cirrhosis and for those undergoing liver transplantation, based on values applied for adefovir and lamivudine in the HTA monograph.[10] However telbivudine does not currently have marketing authorisation for treatment of patients with decompensated disease. The impact, on the cost effectiveness of telbivudine, of removing these effects has been examined in ERG sensitivity analyses reported in sections 4.4.1.4.2 and 4.4.1.4.4.

4.4.1.2.3 Patient outcomes

The cost effectiveness models assumed that health states defined by disease progression (HBsAg loss, HBeAg seroconversion, CHB, compensated cirrhosis, decompensated cirrhosis, HCC, and liver transplantation) determine the patients' quality of life and that quality of life was not directly affected by viral load. This is consistent with previous economic evaluations of anti-viral treatment for CHB.[10,20,21,23]

Utility values applied to health states in both the viral load and the seroconversion models were taken from the HTA monograph.[10] These are calculated as state-specific decrements (see Table 13) from either an age-related average utility for the general population (applied in the serconversion model) or a constant value of one (applied in the viral load model) which were assumed to apply to patients who had cleared the surface antigen (HBsAg) or had seroconverted the e antigen (for patients in the HBeAg-positive cohort).

Table 13 Health state utility decrements – from HTA monograph{10095]

State-specific decrement	Health state
0.04	CHB
0.44	Compensated cirrhosis
0.54	Decompensated cirrhosis, HCC, liver transplant
0.32	Post-liver transplant

For the seroconversion model the state-specific decrements were estimated using beta distributions, with parameters taken from the HTA monograph.{10095} The MS states that the viral load model was "constructed in a manner that required utility multipliers rather than decrements." It is not clear what this statement means. However examination of the electronic model shows that the method used has maintained the average decrements applied in the HTA monograph, since the mean utility for CHB equals 0.96 (i.e. a utility reduction of 0.04 as in Table 13). The mean utility for each of the other health states in the model also reflects the state-specific utility reductions shown in Table 13.

The utility values adopted in the HTA monograph[10] come from a range of sources, including previously published economic evaluations of anti-viral treatment for CHB[20,21,23] (which have generally based their utility estimates on values elicited from expert panels of clinicians rather than patients or the general population), a UK trial of HCV patients[31] (for advanced liver disease states) and a UK study of liver transplant patents.[32] The HTA monograph included a review of the evidence on quality of life and CHB and utility studies, concluding that there were no robust estimates for this patient population. The validity of the values adopted in the HTA monograph is not discussed in the MS (other than to raise concern over the size of decrement for compensated cirrhosis) nor are any literature searches or other attempts to identify more appropriate health state valuations for patients with CHB or advanced liver disease associated with CHB infection reported.

The MS briefly discusses the impact of assuming a constant underlying utility of one in the viral load model and suggests that this will tend to favour lamivudine, at the expense of telbivudine, by reducing the utility losses associated with advanced liver disease. It seems equally plausible that adopting a constant utility of one for the "cured" (HBsAg negative) and inactive carrier states may overstate the benefits achieving these outcomes. A sensitivity analysis applying age-related utility values in the viral load model is reported in 4.4.1.4.2 and 4.4.1.4.4.

4.4.1.2.4 Resource use

Three groups of resource were identified and costed in the MS (see section 6.2.8, pages 93-94 for details):

1) Drug acquisition
2) On-treatment monitoring and management
3) Health state costs – associated with post-treatment surveillance of patients with chronic disease as well as symptomatic management of advanced liver disease states

The MS does not report undertaking a systematic search for data on resource use for CHB patients receiving anti-viral treatment or on resource use associated with symptomatic management of advanced liver disease states. As discussed earlier (section 3.1.1.2) the MS states that no searches of the cost-effectiveness literature were attempted and that the HTA monograph treated as the "gold standard" for economic evaluations relevant to England and Wales. It appears this assumption has been applied to estimating resource use, other than drug costs. However the list of identified resource groups seems to be comprehensive and agree with categories of resource use identified in previous economic studies of anti-viral treatment for patients with CHB.[20-23,33,34]

Treatment costs have been calculated using the licensed dosage for each anti-viral agent (see Table 14). Treatment has been assumed to continue while patients remain non-resistant, hence models have calculated anti-viral drug use and costs per cycle. The duration of continuation of treatment for resistant patient is therefore dependent on the cycle length in each model – a patient developing treatment resistance during a given cycle in the viral load model will continue treatment until the end of that six months cycle, while in the seroconversion model they will continue to the end of the annual cycle.

Table 14 Resource use assumptions for anti-viral drugs in MS

Anti-viral drug	Licensed dosage	Packaging	Price per pack	Cost per year of treatment†
Lamivudine	100 mg/ day	28 x 100mg tablets	78.09	1,018.66
Adefovir	10 mg/ day	30 x 10mg tablets	315.00	3,835.13
Telbivudine	600 mg/ day	28 x 600mg tablets	290.33	3,787.25
Notes † assumes average year of 365.25 days				

Resource use assumptions for monitoring patients while on treatment in both models have been taken directly from the HTA monograph.[10] This assumed, based on review of clinical guidelines and expert opinion, that patients treated with lamivudine or adefovir would be seen eleven times during a year of treatment, with more detailed assessments every quarter, and that all contacts would be with specialist nurses. The MS has assumed that similar protocols would apply for telbivudine – this seems reasonable. However the MS does not contains any reference to clinical advice received in adopting these assumptions nor any critical appraisal of the assumptions adopted in the HTA monograph.

Health state costs adopted in the models are also taken from the HTA monograph. The MS does not discuss the resource use assumptions underlying the health state costs, nor do they consider the appropriateness of the assumption that costs for advanced liver disease (compensated cirrhosis, decompensated cirrhosis and HCC) estimated for patients with chronic hepatitis C can be applied to CHB patients. However, this seems a reasonable approach - in the absence of searches for alternative data on resource use by disease stage, and to retain compatibility with assumptions adopted for the previous NICE assessment.

4.4.1.2.5 Costs

Unit costs for all anti-viral drugs are taken from the British National Formulary (BNF 54).[27] Other unit cost data (cost of patient assessments while on-treatment and health state costs) were taken from an HTA monograph[10] and have been uprated to 2005/06 prices using the Hospital and Community Health Services Pay and Prices Index. [28]

4.4.1.3 Consistency

Internal consistency

All the submitted electronic models were developed in MS Excel. Random checking has been conducted for some of the key equations in the models, for example on the worksheets named *'Calculations'* (which contains the CC and HCC risk models), *'Markov_Tel'* and *'Markov_Lam'* (which contain the transition models for telbivudine and lamivudine respectively) in the viral load model and the *'HBeAg+ Model'* worksheet in the serconversion model. However, the ERG has not undertaken a comprehensive check of all cells in each model. The models are fully executable. However both are heavily reliant on Visual Basic to produce any analyses. Neither

model is structured for simple display of deterministic outputs which has hampered the ERG's ability to test the models' internal consistency using simple sensitivity analyses.

The MS does not report on attempts to establish the models' internal validity – there are no reports or discussion of consistency checks or checks of coding accuracy. This is surprising given the complexity of the models and the large volume of reprocessing of data required in the viral load model. There is no report in the MS, nor any evidence in the submitted models of any on checks conducted of the accuracy of input data in the models. This again is surprising given the large quantity of input data to the models. The ERG have checked samples of the input data and have checked the means of the probabilistic samples, but cannot verify the accuracy of all input data for each of 2,500 input sets for the viral load model or the 10,000 input sets for the seroconversion model.

In response to a request for clarification from the ERG regarding a number of tables on the *'Trees_DNA'* worksheet in the viral load model that were reporting errors, the manufacturer noted that the model that was originally submitted contained errors (see Appendix 1 for details) and submitted a new model with updated results. The ERG has examined the new models and found further discrepancies:

- the calibration factors for the compensated cirrhosis and HCC risk equations used in the submitted models do not match those reported in Appendix C of the MS. Values reported in Appendix C of the MS and those applied in the original and resubmitted models are shown in Table 15.

Table 15 Calibration factors reported in MS and used in electrronic models

	Compensated cirrhosis		HCC	
	Appendix C	Electronic model†	Appendix C	Electronic model†
HBeAg Positive Cohort	0.5984	0.6872	0.3182	0.5009
HBeAg Negative Cohort	0.6500	0.7739	0.3010	0.4427
Notes † values applied in original and resubmitted electronic model				

- Table 16 reports deterministic results for the re-submitted viral load model (prior=0) using the calibration factors reported in Appendix C of the MS rather than the values in the re-submitted models.

Table 16 Cost effectiveness results from resubmitted model using calibration factors in Appendix C of MS

	Total cost	Total QALYs	Incremental cost	Incremental QALYs	ICER
HBeAg-positive Cohort					
Lamivudine	34,837	15.41			
Telbivudine	57,583	17.19	22,746	1.78	12,791
HBeAg-positive Cohort					
Lamivudine	37,241	13.93			
Telbivudine	78,611	15.68	41,370	1.75	23,678

- This has a small effect on incremental costs (increasing in both cohorts by around 1%, compared with the manufacturer's re-submitted results). The effect on incremental QALYs is more marked with a reduction of around 3% in the HBeAg-positive cohort and 13% in the HBeAg-negative cohort, compared with the manufacturer's re-submitted results. Given that the analyses used to derive these risk models appears not to have been published the ERG cannot state which values are correct.

External consistency

There is very limited discussion of the external validity of the viral load model. There is no systematic comparison of modelled outcomes against the Globe study data. The MS states (section 6.1.12, page 96) that the Globe data were "predicted in the initial 2 years of the model", but do not discuss the meaning of this statement nor do they provide any documentation to support the statement. Other discussion of the model's validity is limited to a description of the method used to estimate risk of cirrhosis and HCC.

The approach to establishing the external consistency of the seroconversion model was to compare results with those from the SHTAC model and those submitted by the manufacturer of adefovir, reported in the HTA monograph[10] and in the NICE guidance.[2] Only the results for lamivudine and adefovir were relevant to this analysis, given that telbivudine was not included in the previous NICE assessment. When adopting similar assumptions regarding the efficacy of lamivudine and adefovir, and using similar discount rates (6% for costs and 1.5% for outcomes as were required in assessments conducted for NICE at the time the research reported in the HTA monograph[10] was undertaken) ICERs estimated in the serconversion model were consistent with those estimated by the SHTAC model and the manufacturer of adefovir.

4.4.1.4 Assessment of Uncertainty

General comments on the assessment of uncertainty in the MS:

- No deterministic sensitivity analyses are reported for any of the submitted models. As a result it has been difficult to identify, on the basis of the submitted evidence, what are key drivers of cost effectiveness in the models
- The mean values from probabilistic evaluations of the models have been used in the base case results, in accordance with NICE methodological guidance, but there is very limited discussion on uncertainty in PSA. The MS reports percentile-based and Jackknife confidence intervals, but cost effectiveness plots and CEACs are only included in appendices and are not discussed.
- The MS contains no brief summary of the variables, or categories of variables in the PSA – hence it is not always clear what has or has not been included. The reliance on Visual Basic coding to produce any analyses from the models also hampers the identification of variables included in the PSA.

4.4.1.4.1 One-way sensitivity analyses

The MS does not report any one-way sensitivity analyses for either the viral load model or the seroconversion model. However the submitted viral load model included a worksheet that contained one-way sensitivity analyses and these are reproduced in this section by the ERG.

Table 17 shows the one–way sensitivity analyses from the viral load model as reported in the manufacturer's model, including the range in the ICER between the higher and lower values of the parameter inputs. As mentioned above, the MS does not include any information on the one-way sensitivity analyses, and so there is no explanation of the results, rationale for the choice of variables included or excluded and explanation on the choice of the ranges used for these variables. The following variables were subjected to sensitivity analyses: patient sex, patient age, telbivudine cost and resistance, utilities, and the mortality rates in different health states.

The inputs appear to be chosen arbitrarily, rather than being based upon 95% confidence interval ranges of the parameters. For example, the input ranges for the mortality rates include the mean as either the higher or lower input. Generally the model results were not sensitive to

changes in parameters. The exception is the HBeAg-negative disease model (with a prior of 0.5) where the results are highly sensitive to changes in model parameters. The results are most sensitive to changes in quality of life (ICER range -£135,000 to £108,000), the patient start age, gender and cox model hazards for HCC and cirrhosis. For the other models, results are most sensitive to telbivudine cost, discount rate and the resistance with telbivudine. For example, for the HBeAg-positive disease model with a prior of 0.0, the ICER varies between £7,700 and £20,300 for changes in the cost of telbivudine.

Table 17 One – Way sensitivity analyses from the viral load model. Results show the range of the ICER between the higher and lower values of the parameter inputs (from submitted model).

Variable		Prior = 0.5		Prior = 0.0	
		HBeAg-positive	HBeAg-negative	HBeAg-positive	HBeAg-negative
Telbivudine Cost	(20% lower, 50% higher)	11808.10	61033.00	12583.26	7857.21
Annual Discount Rate	(0, 0.05)	10616.86	32815.10	10654.98	2155.17
Resistance with Telbivudine	(50% higher, 50% lower)	4138.01	15637.38	13253.22	2308.03
Cox Model Hazards for HCC and Cirrhosis	(50% higher, 50% lower)	6578.46	134493.50	5484.24	981.35
Sex	(Men, Women)	4292.75	126576.92	3150.35	612.24
Start age	(40,60)	2753.86	140975.11	3286.98	519.13
Quality of Life	(Canadian VAS, US)	1162.72	243042.52	1414.35	3641.69
Hepatitis B Disease Cost	(50% lower, 50% higher)	654.25	6655.63	412.82	1149.82
Probability of Developing Decompensated Cirrhosis	(0.095, 0.038)	229.32	6084.68	104.16	46.79
Probability of Liver Transplant for Developing Decompensated Cirrhosis	(0.01, 0.25)	50.89	1656.14	52.91	3.93
Annual Excess Mortality for Decompensated Cirrhosis	(0.39, 0.11)	89.64	1892.38	14.98	23.59
Annual Excess Mortality for Hepatocellular Carcinoma (1st year and subsequent year)	(0.71, 0.43)	178.33	807.85	217.59	34.93
Annual Excess Mortality for Liver Transplant 1st Year	(0.21, 0.057)	2.41	87.86	2.38	0.23
Annual Excess Mortality for Liver Transplant beyond 1st Year	(0.14, 0.01)	16.83	331.30	15.56	1.27

Table 18 shows the amended one–way sensitivity analyses from the resubmitted viral load model. Generally the sensitivity analyses in the amended results are similar to those reported in the original analysis, i.e. the model results were not sensitive to changes in parameters except the HBeAg-negative disease model with a prior of 0.5, where the results are very sensitive to changes in model parameters.

Table 18 One – Way sensitivity analyses from the resubmitted viral load model. Results show the range of the ICER between the higher and lower values of the parameter inputs.

Variable		Prior = 0.5		Prior = 0.0	
		HBeAg-positive	HBeAg-negative	HBeAg-positive	HBeAg-negative
Telbivudine Cost	(20% lower, 50% higher)	10598	49323	12289	18560
Annual Discount Rate	(0, 0.05)	9166	26750	8888	10084
Cox Model Hazards for HCC and Cirrhosis	(50% higher, 50% lower)	4679	90243	2811	14499
Sex	(Men, Women)	3009	81695	1517	10536
Quality of Life	(Canadian VAS, US)	934	2538348	2954	7218
Start age	(40,60)	3330	91193	5672	1850
Resistance with Telbivudine	(50% higher, 50% lower)	4077	15224	5246	129
Hepatitis B Disease Cost	(50% lower, 50% higher)	592	4340	723	1289
Probability of Developing Decompensated Cirrhosis	(0.095, 0.038)	166	4324	78	271
Annual Excess Mortality for Hepatocellular Carcinoma (1st year and subsequent year)	(0.71, 0.43)	176	547	128	22
Probability of Liver Transplant for Developing Decompensated Cirrhosis	(0.01, 0.25)	31	1133	31	91
Annual Excess Mortality for Decompensated Cirrhosis	(0.39, 0.11)	77	1191	10	7
Annual Excess Mortality for Liver Transplant 1st Year	(0.21, 0.057)	12	224	9	18
Annual Excess Mortality for Liver Transplant beyond 1st Year	(0.14, 0.01)	1	60	1	4

4.4.1.4.2 ERG sensitivity analysis

All sensitivity analyses reported in section 4.4.1.4.2 and section 4.4.1.4.4 are deterministic analyses – given the difference between the deterministic and probabilistic results for the viral load model using a prior of 0.5, deterministic sensitivity analyses will not be reported for that model. In contrast the deterministic results for the viral load model with a prior of zero match the probabilistic means reasonably closely and will be reported in the following sections.

ERG sensitivity analyses for viral load model

The viral load model applies utility decrements (based on those adopted in the HTA monograph[10]) to a constant value of one – the health state utility assumed for patients who have lost the surface antigen (HBsAg), irrespective of age (see section 4.4.1.2.3). This may not be the most appropriate approach when extrapolating to a lifetime horizon for patients entering the model aged between 30 and 40. To investigate the impact of applying non-constant utilities, the age-related utility values adopted in the seroconversion model (derived from the HTA

monograph[10]) were applied in the viral load model. Table 19 shows the total and incremental QALYs calculated using the age-specific utilities and the associated ICERs. Using state-specific decrements from age-related utility values yields a smaller QALY gain for telbivudine, compared with lamivudine, than using state-specific decrements from a constant value of one.

Table 19 Applying age-specific utility values in the viral load model (prior = 0)

	Total cost	Total QALYs	Incremental cost	Incremental QALYs	ICER
HBeAg Positive Cohort					
Lamivudine	34,214	12.7			
Telbivudine	56,669	14.2	22,456	1.50	14,930
HBeAg Negative Cohort					
Lamivudine	36,417	10.9			
Telbivudine	77,429	11.5	41,012	1.60	25,557

In the submitted viral load model patients are all assumed to enter the model in the chronic hepatitis state, without cirrhosis. However surveys of the literature suggest that 10-24% of HBeAg-positive and 29-38% of HBeAg-negative patients have cirrhosis at the time of presenting for treatment. To investigate the impact of the stage of disease at initiation of treatment a proportion of patients in both the HBeAg-positive and HBeAg-negative cohorts were assumed to have developed cirrhosis prior to entering the model – this required some amendments to be made to the viral load model. It was assumed that patients with cirrhosis were distributed across viral load levels in the same proportion as non-cirrhotic patients.

Table 20 Cost effectiveness of telbivudine compared with lamivudine, varying the proportion of cohort who are cirrhotic at initiation of treatment in the viral load model (prior = 0)

Proportion cirrhotic	Patient cohort	Incremental cost	Incremental QALYs	ICER
0.0	HBeAg Positive Cohort	22,456	1.83	12,278
	HBeAg Negative Cohort	41,012	2.01	20,383
0.1	HBeAg Positive Cohort	21,772	1.74	12,502
	HBeAg Negative Cohort	38,045	1.91	19,934
0.2	HBeAg Positive Cohort	21,089	1.65	12,751
	HBeAg Negative Cohort	35,078	1.81	19,433
0.3	HBeAg Positive Cohort	20,406	1.57	13,027
	HBeAg Negative Cohort	32,111	1.70	18,872
0.4	HBeAg Positive Cohort	19,722	1.48	13,336
	HBeAg Negative Cohort	29,144	1.60	18,238

Assuming that a proportion of the cohort are cirrhotic at initiation of treatment reduces total costs and total QALYs for each treatment. It also reduces the incremental cost and incremental QALYs for telbivudine compared with lamivudine. The impact on the cost effectiveness estimate depends on the proportionate changes in incremental cost and incremental QALYs, which differs between patient cohorts. The proportionate reduction in incremental QALYs is greater than the proportionate reduction in incremental cost in the HBeAg-positive cohort, hence the ICERs increase as the proportion of the cohort that are cirrhotic at initiation of treatment increases (see Table 20). The reverse is the case for the HBeAg-negative cohort.

ERG sensitivity analyses for seroconversion model

The MS did not contain one-way sensitivity analyses from the seroconversion model and there were none included in the manufacturer's model. The ERG used the seroconversion model to generate one-way sensitivity analyses. The ERG used input parameters and ranges based upon those previously used in the HTA monograph.[10] The analyses show that the model results were most sensitive to changes in the probability of patients progressing from CHB or compensated cirrhosis to the inactive carrier state, patient start age and changes in the cost of the treatments.

Table 21 ERG one–way sensitivity analyses from the seroconversion model. Results show the ICER for selective parameter inputs for different treatment combinations.

	Lamivudine vs supportive care	Telbivudine vs lamivudine	Lamivudine followed by adefovir vs Lamivudine	Telbivudine followed by adefovir vs lamivudine followed by adefovir
Baseline	9,208	20,722	25,738	19,243
Start Age 20	9,663	18,975	23,093	17,670
Start Age 50	12,385	29,958	37,845	27,686
Drug costs 20% higher	10,489	25,262	30,601	23,731
Drug costs 20% lower	7,926	16,181	20,874	14,755
Compensated cirrhosis state cost of £2,220 (Base case = £1,117)	8,986	20,485	25,773	18,935
Probability of CHB patient becoming inactive carrier of 0.21 (Base case = 0.16)	9,208	12,997	21,503	10,211
Probability of CHB patient becoming inactive carrier of 0.11 (Base case = 0.16)	9,208	44,295	32,029	51,694

	Lamivudine vs supportive care	Telbivudine vs lamivudine	Lamivudine followed by adefovir vs Lamivudine	Telbivudine followed by adefovir vs lamivudine followed by adefovir
Double progression from compensated cirrhosis to decompensation (Base case = 1.8%)	9,208	23,399	25,738	23,046
Annual excess mortality for decompensation of 0.29 (Base case = 0.19)	9,208	20,672	25,738	19,133
Liver transplant annual excess mortality halved (Base case = 0.21, 0.057)	9,208	20,735	25,738	19,258
Compensated cirrhosis utility decrement is 0.07	9,990	22,610	25,421	21,592
Telbivudine resistance same as lamivudine	9,208	19,117	25,738	18,090
Telbivudine resistance same as adefovir	9,208	22,524	25,738	21,018
No treatment with telbuvidine for decompensated patients	9,208	19,783	25,738	16,727
Remove resistant patients from denominators for transition probabilities	8,687	21,413	25,192	20,419
15% cohort have cirrhosis prior to start of treatment	9,283	23,139	25,548	22,685
Treated cirrhotic patients seroconvert at same rate as treated non-cirrhotic	8,870	14,045	17,326	12,712

4.4.1.4.3 Scenario Analysis

The MS does not report any deterministic scenario analyses.

4.4.1.4.4 ERG scenario analysis

ERG scenario analyses for viral load model

Applying age-specific utilities and increasing the proportion of the cohort who are cirrhotic at treatment initiation reduces the total QALYs for both drugs and reduces the QALY gain for telbivudine compared with lamivudine, with little or no effect on total or incremental costs. Applying the calibration factors reported in Appendix C of the MS, rather than those in the submitted models has a minimal effect on costs and QALYs for patients with HBeAg-positive CHB, in the viral load model.

Table 22 Cumulative effect of alternative assumptions on health state utilities, proportion of cohort cirrhotic and calibration factors for risk of advanced liver disease. HBeAg-positive patients in viral load model (prior = 0)

	Total cost	Total QALYs	Incremental cost	Incremental QALYs	ICER
Apply age-specific utilities					
Lamivudine	34,214	12.7			
Telbivudine	56,669	14.2	22,456	1.50	14,930
15% of cohort is cirrhotic at initiation of treatment					
Lamivudine	33,296	11.3			
Telbivudine	54,732	12.7	21,436	1.39	15,419
Use calibration factors (for risk of advanced liver disease) from Appendix C					
Lamivudine	34,057	11.9			
Telbivudine	55,768	13.3	21,711	1.35	16,139

Applying age-specific utilities leads to reduction in the total QALYs for both drugs and reduces the QALY gain for telbivudine compared with lamivudine. In contrast with the analysis above, if the proportion of the cohort that is cirrhotic at treatment initiation is increased (for patients with HBeAg-negative CHB), incremental cost and QALYs for telbivudine compared with lamivudine are reduced improving the incremental cost effectiveness ratio. Applying the calibration factors reported in Appendix C of the MS, rather than those in the submitted models marginally increases incremental costs and reduces the QALY gain leading to a less favourable cost effectivness estimate (see Table 23).

Table 23 Cumulative effect of alternative assumptions on health state utilities, proportion of cohort cirrhotic and calibration factors for risk of advanced liver disease. HBeAg-negative patients in viral load model (prior = 0)

	Total cost	Total QALYs	Incremental cost	Incremental QALYs	ICER
Apply age-specific utilities					
Lamivudine	36,417	10.9			
Telbivudine	77,429	12.5	41,012	1.60	25,557
15% of cohort is cirrhotic at initiation of treatment					
Lamivudine	31,910	8.3			
Telbivudine	64,021	9.7	32,111	1.35	23,766
Use calibration factors (for risk of advanced liver disease) from Appendix C					
Lamivudine	32,793	8.9			
Telbivudine	65,331	10.1	32,537	1.24	26,164

ERG scenario analyses for seroconversion model

Removing treatment effects (and costs) for decompensated patients when the cohort is treated with telbuvidine has a comparatively small effect, but improves the cost effectivness of telbivudine compared with treatment strategies including lamivudine. If resistant patients are removed from the denominators when calculating transition probabilities for treatment effects (as discussed in section4.4.1.2.2 of this report), gives slightly more favourable ICERs for lamivudine-containing strategies and slightly less favourable ICERs for telbivudine-containing strategies, though the effect is small (see Table 24).

Assuming that a proportion of patients have compensated cirrhosis at treatment initiation has a small effect on cost effectiveness estimates. However assuming that cirrhotic patients can HBeAg seroconvert at the same rate as non-cirrhotic patients has a large impact on cost effectiveness estimates, producing more favourable ICERs for all treatment strategies.

Table 24 Cumulative effect of alternative assumptions on continuation of treatment, removing resistant patients from denominators, proportion of patient cohort who are cirrhotic at start of treatment and seroconversion rate in cirrhotic patients, in the seroconversion model

	Lamivudine vs supportive care	Telbivudine vs lamivudine	Lamivudine followed by adefovir vs lamivudine	Telbivudine followed by adefovir vs lamivudine followed by adefovir
No treatment with telbuvidine for decompensated patients	9,208	19,783	25,738	16,727
Remove resistant patients from denominators for transition probabilities	8,687	20,453	25,192	17,814
15% cohort have cirrhosis prior to start of treatment	8,836	22,477	25,074	20,186
Treated cirrhotic patients seroconvert at same rate as treated non-cirrhotic	8,152	10,760	15,809	8,371

4.4.1.4.5 Probabilistic Sensitivity Analysis

Viral load model

The probabilistic sensitivity analysis in the viral load model can be run by clicking on the 'Refresh probabilistic results' button on the '*PSA*' worksheet (for a specified number of trials,

outputting results to the '*PSA*' worksheet) or by clicking on the 'Refresh base case results' button on the '*Results*' worksheet (running 2,500 trials, outputting results to the '*PSA Results*' worksheet). The PSA takes approximately 3 days to run (on a computer with 1.86 GHz dual core processor, 1 Gb RAM) for 1000 simulations. The results of the PSA (incremental cost effectiveness ratio calculated at the mean incremental cost and mean incremental QALY for telbivudine relative to lamivudine in the PSA) are presented in Table 26 and Table 27 in the MS, with more detail (including cost effectiveness scatterplots and CEACs) in Appendix J of the MS. These are summarised in Table 4 of this report.

In response to a request for clarification from the ERG the manufacturer noted there were errors in the models originally submitted and therefore in the results included in the MS. The manufacturer submitted a revised set of results (see Appendix 1 of this report) and updated versions of the electronic models. The manufacturers did not submit updated cost effectiveness scatterplots or CEACs – the ERG have produced these (scatterplots are included in Appendix 3 of this report). See Figure 1 and Figure 2 for the CEACs for telbivudine from the viral load model for patients with HBeAg-positive CHB, with prior of zero and 0.5 respectively, as calculated by the ERG.

Figure 1 CEAC for telbivudine (HBeAg-positive CHB cohort - prior = 0.0)

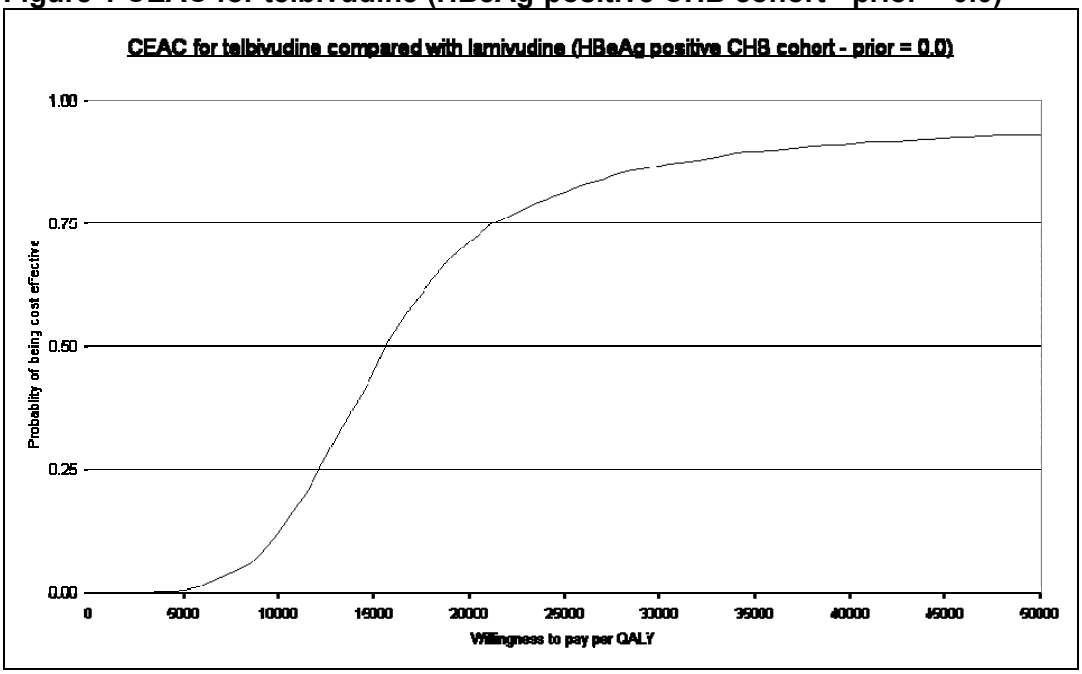


Figure 2 CEAC for telbivudine (HBeAg-positive CHB cohort - prior = 0.5)

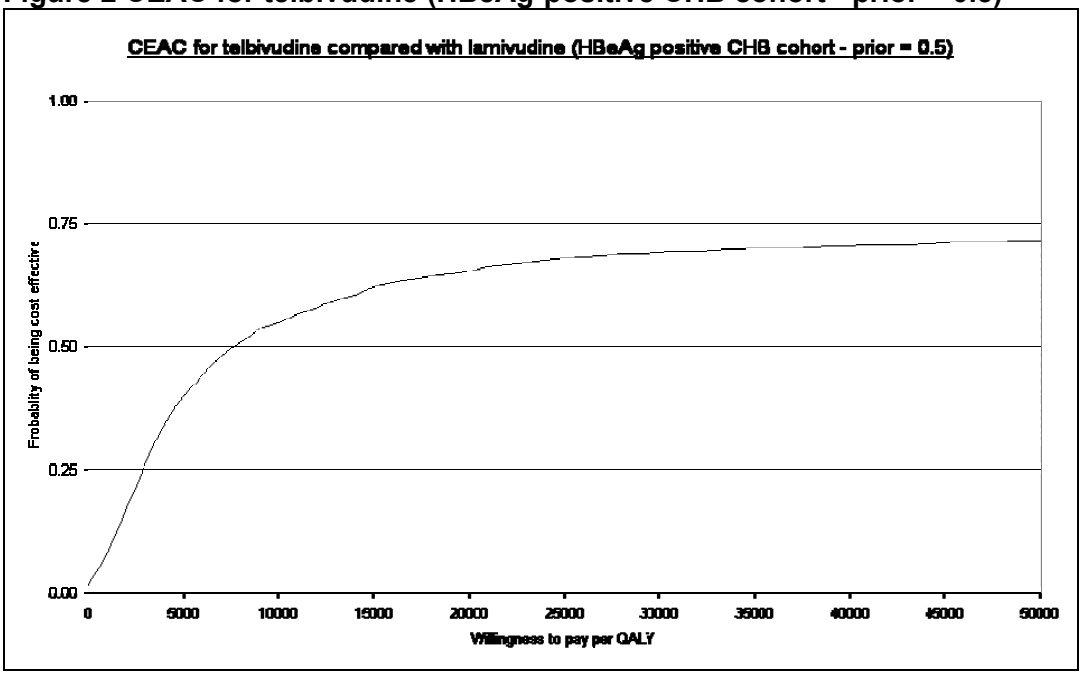


Figure 3 and Figure 4 show the CEACs for telbivudine from the viral load model for patients with HBeAg-negative CHB, with prior of zero and 0.5 respectively, as calculated by the ERG.

Figure 3 CEAC for telbivudine (HBeAg-negative CHB cohort - prior = 0.0)

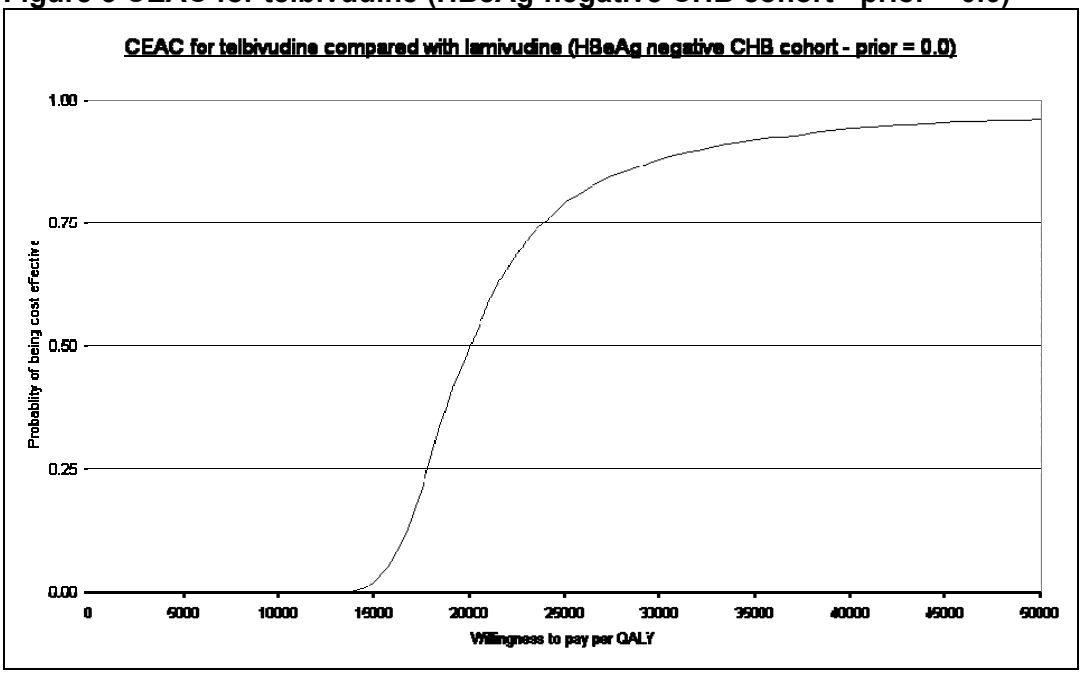
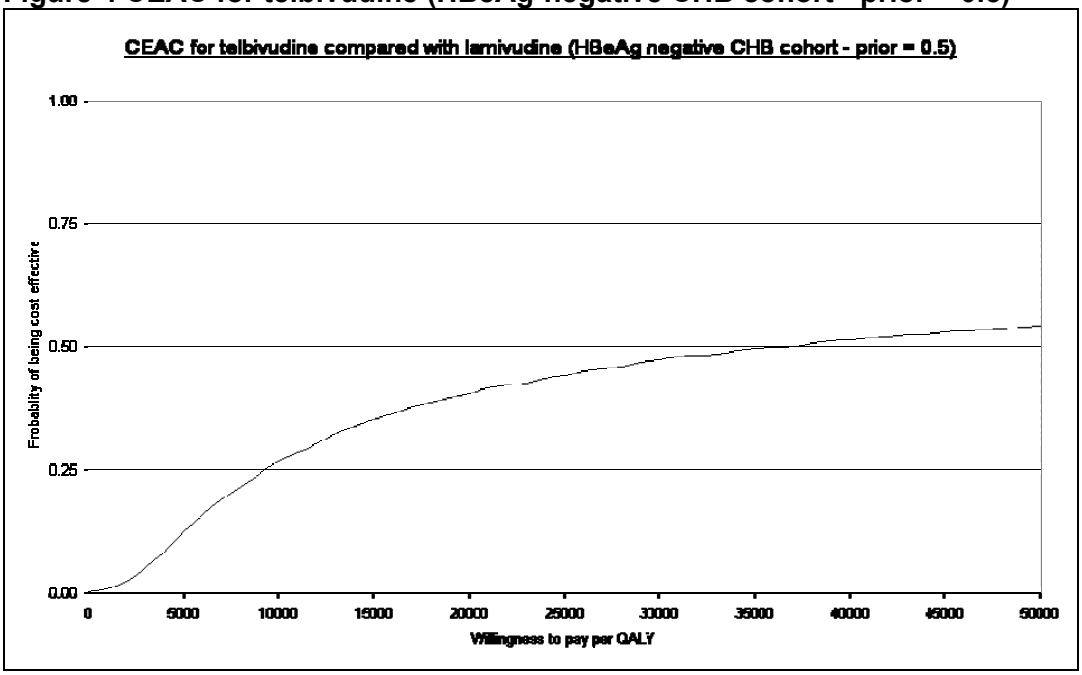


Figure 4 CEAC for telbivudine (HBeAg-negative CHB cohort - prior = 0.5)



For HBeAg-positive patients, telbivudine has a 71% probability of being cost-effective at a willingness to pay threshold of £20,000 per QALY gained and 87% at a threshold of £30,000 per QALY gained when using a prior of 0.0. The equivalent values, with a prior of 0.5, are 65% and 69%. For HBeAg-negative patients, telbivudine has a 49% probability of being cost-effective at a willingness to pay threshold of £20,000 per QALY gained and 88% at a threshold of £30,000 per QALY gained when using a prior of 0.0. The equivalent values, with a prior of 0.5, are 40% and 47%.

It appears that all variables in the model, other than drug costs and on-treatment/ post-treatment monitoring costs, have been included in the PSA. Baseline characteristics (patients' mean age and the proportion of cohort that were male) were also fixed. Male sex and age are both associated with greater risk of developing cirrhosis and HCC in the disease progression risk equations used in the model. Uncertainty around these could have been included in PSA or a justification offered for excluding them. Details of the variables, distributions and parameters used are included in Appendix B of the MS (Appendix B1 for variables using data form the Globe study with prior=0.5; Appendix B2 for variables using data form the Globe study with prior=0.0; Appendix B3 for other variables included in the PSA, generally using parameter values derived from the HTA monograph[10]). There is no discussion in the MS of the choice of variables to include, the distributions chosen, or of appropriate ranges for the data. The

distributions chosen seem appropriate. However key inputs (such as the risk equations for developing compensated cirrhosis and HCC) have been entered deterministically so that the PSA does not fully reflect the uncertainty in the model (see summary and discussion below).

Summary of assumptions for manufacturer's PSA:

1. All values sampled from distributions are hard-coded into worksheet rather than being sampled using Excel's built-in distribution functions. It seems likely that this approach was adopted as a mechanism to manage memory in such a large model.
2. The majority of transition probabilities are estimated using beta distributions, using the number of events (α) and non-events (β) for each transition observed in the Globe trial as parameters. The exception to this is the calculation of transitions between viral load levels, which were estimated using Dirichlet distribution. In response to a request for clarification the manufacturer stated that the Dirichlet distributions were sampled using an Excel Add-in developed by the Centre for Health Economics and Bayesian Statistics.[35] The input parameters for the Dirichlet distributions were the number of patients in each viral load level at the start of the cycle.
3. It appears that the probability of developing compensated cirrhosis and hepatocellular derived using the REVEAL-HBV risk equations were not included in the probabilistic sensitivity analysis, but were fixed at the values reported in MS Appendix C. It is unclear what effect this would have on the ICER results but would under-estimate the uncertainty in the model, both due to uncertainty in parameter estimates, but also due to methodological uncertainty (for example the "calibration" is entered deterministically).
4. Costs were assumed to follow gamma distribution. It appears that the mean and standard errors for health state costs (reported in Appendix 17, Table 55 of the HTA monograph[10]) were both uprated to 2005/6 prices and were then used to derive parameters for gamma distributions, using the method of moments.[36]
5. It is not clear from the MS what distribution was assumed for the health state utilities. It appears that values for the utility decrements may have been sampled using beta distributions with parameters from the HTA monograph[10] and that values stored in the spreadsheet models are 1 minus the sampled values.

Seroconversion model

The probabilistic sensitivity analysis in the seroconversion model can be run by clicking on the 'Stochastic results' button on the '*Model Input*' worksheet. The PSA takes about 2 hours to run

(on a computer with 1.86 GHz dual core processor, 1 Gb RAM) for 1000 simulations. The results of the PSA (incremental cost effectiveness ratio calculated at the mean incremental cost and mean incremental QALY for treatment algorithms including telbivudine, lamivudine and adefovir) are presented in Table 29 in the MS, with more detail (including CEACS) in Appendix J of the MS. The mean ICER, with a 95% confidence interval (using the 2.5th and 97.5th percentiles), are reported in Table 5 of this report.

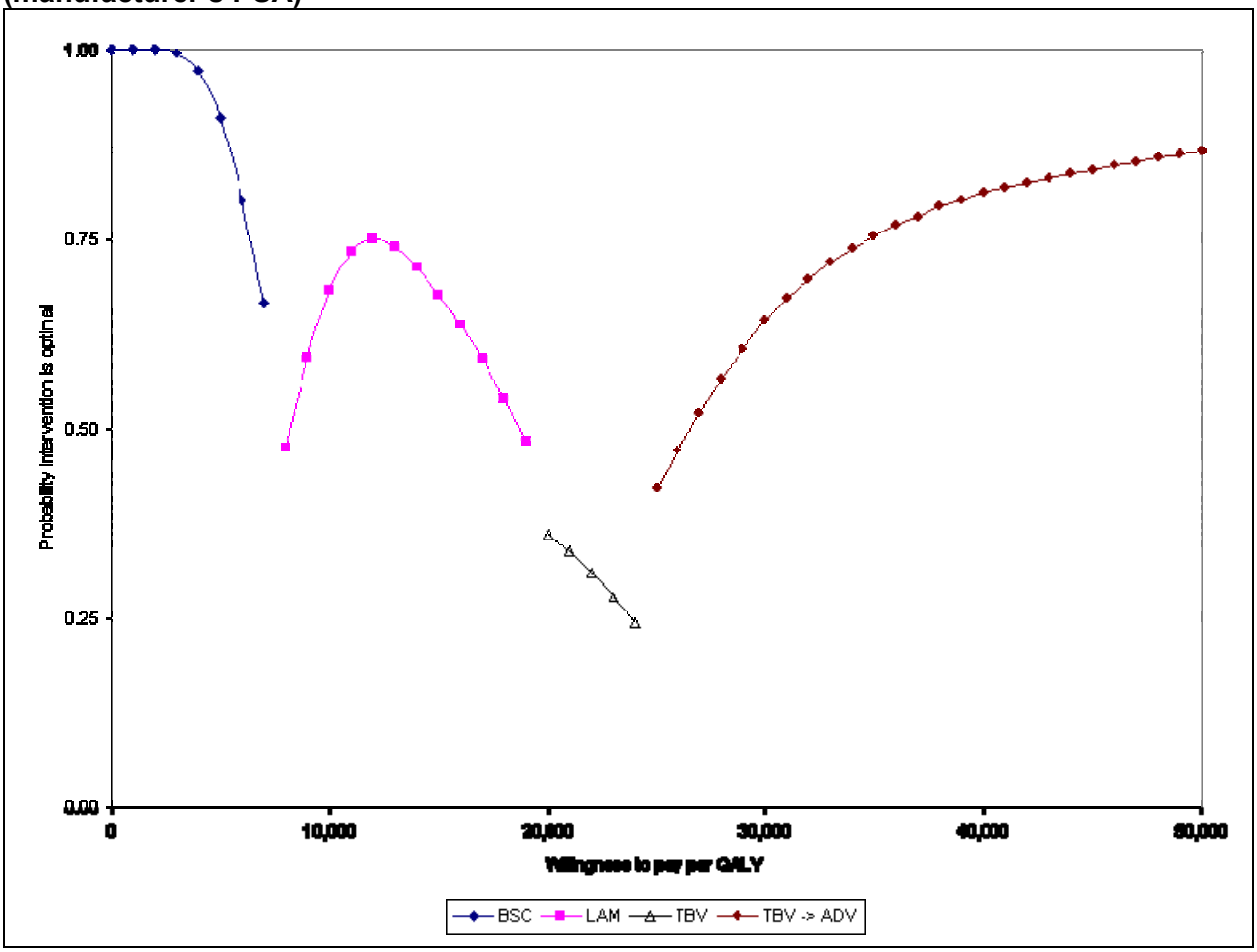
Interpretation of the results of the PSA, in the body of the MS, is limited to tables presenting the mean incremental net benefit for all treatment strategies compared with best supportive care. This suggests that telbivudine is the optimal intervention at a threshold willingness to pay of £20,000 per QALY and that telbivudine followed by adefovir is optimal at a threshold of £30,000 per QALY. The MS does not report the probability of interventions being cost effective at the same willingness to pay thresholds (see Table 25). As can be seen from the table, the strategy with the maximum incremental next benefit is not necessarily that with the greatest probability of being cost-effective.[37]

Table 25 Probability intervention is cost effective at selected willingness to pay thresholds (manufacturer's PSA)

Treatment strategy	Threshold willingness to pay	
	£20,000 per QALY	£30,000 per QALY
Lamivudine	43%	8%
Telbivudine	36%	8%
Lamivudine followed by adefovir	7%	19%
Telbivudine followed by adefovir	14%	64%

By focussing on these two threshold values, the MS does not consider strategies that may be optimal at other threshold values. For example, lamivudine is the optimal strategy (using the net benefit criterion) in the range from £8,000 to £20,000 per QALY threshold (see Figure 5 for the cost effectiveness acceptability frontier[37]) which shows the portions of the CEAC where interventions are deemed optimal using the maximum net benefit criterion over a range of willingness to pay values).

Figure 5 Cost effectiveness acceptability frontier from seroconversion model (manufacturer's PSA)



It appears that all variables in the model, other than drug costs, on-treatment/ post-treatment monitoring costs and resistance to adefovir, have been included in the PSA. Details of the variables, distributions and parameters used are included in Appendix F to the MS. There is no discussion in the MS of the choice of variables to include, the distributions chosen, or of appropriate ranges for the data. Nevertheless the choice of variables included in the PSA appears reasonable and distributions chosen seem appropriate (see summary below).

Summary of assumptions for manufacturer's PSA:

1. All values sampled from distributions are hard-coded into worksheet – not sampled in worksheet.
2. The majority of the disease progression transition probabilities are estimated using beta distributions with the parameters taken directly from the HTA monograph.[10]

3. Effectiveness (in inducing HBeAg seroconversion), drug resistance and reactivation of disease for patients treated with lamivudine or telbivudine have been estimated using beta distributions, with the number of events (α) and non-events (β) for each transition observed in the Globe trial as parameters. Reactivation of disease for patients treated with adefovir has been estimated in the same way. The effectiveness of adefovir, in inducing HBeAg seroconversion, was estimated as the average of the effectiveness of lamivudine and telbivudine in each simulation.
4. Costs were assumed to follow gamma distribution. It appears that the mean and standard errors for health state costs (reported in Appendix 17, Table 55 of the HTA monograph[10]) were both inflated to 2005/6 prices and were then used to derive parameters for gamma distributions, using the method of moments.[36]
5. Health state utilities were estimated in the model using the same method as the HTA monograph, [10] using state-specific decrements applied to age-specific utilities reported by Kind and colleagues.[38] The state-specific decrements were sampled from beta distributions with the parameters taken directly from the HTA monograph.[10]

4.4.1.4.6 ERG probabilistic sensitivity analysis

Viral load model

The ERG conducted a probabilistic sensitivity analyses using the viral load model (prior = 0.0) after:

- replacing the constant health state utility for the "cured" (lost HBsAg) health state with age specific utilities for a general population, as in the HTA monograph.[10] This is discussed in section 4.4.1.4.2 of this report.
- applying the calibration factors for the advanced liver disease risk equations reported in Appendix C of the MS, rather than the values in the submitted electronic models This is discussed in section 4.4.1.4.2 of this report.

Due to the time taken to run the PSA using the viral load model the ERG limited the number of iterations to 1,000 and only ran the PSA for the model with a prior of zero.

Table 26 reports the mean total costs and QALYs for telbivudine and lamivudine, mean incremental cost and QALYs for telbivudine compared with lamivudine and the ICER for patients with HBeAg-positive CHB from the ERG's PSA.

Table 26 Mean results from ERG PSA, using the viral load model (HBeAg-positive cohort - prior = 0.0)

	Total		Incremental		ICER
	Cost	QALYs	Cost	QALYs	
Lamivudine	£35,595	13.45			
Telbivudine	£60,026	14.74	£24,445	1.29	£18,984

Table 26 reports the mean total costs and QALYs for telbivudine and lamivudine, mean incremental cost and QALYs for telbivudine compared with lamivudine and the ICER for patients with HBeAg-negative CHB from the ERG's PSA.

Table 27 Mean results from ERG PSA, using the viral load model (HBeAg-negative cohort - prior = 0.0)

	Total		Incremental		ICER
	Cost	QALYs	Cost	QALYs	
Lamivudine	£37,600	11.68			
Telbivudine	£79,609	13.10	£42,009	1.42	£29,496

Cost effectiveness acceptability curves from the ERG's PSA are shown in Figure 6 and Figure 7.

Figure 6 CEAC for telbivudine (HBeAg-positive cohort – prior = 0.0) ERG Analysis

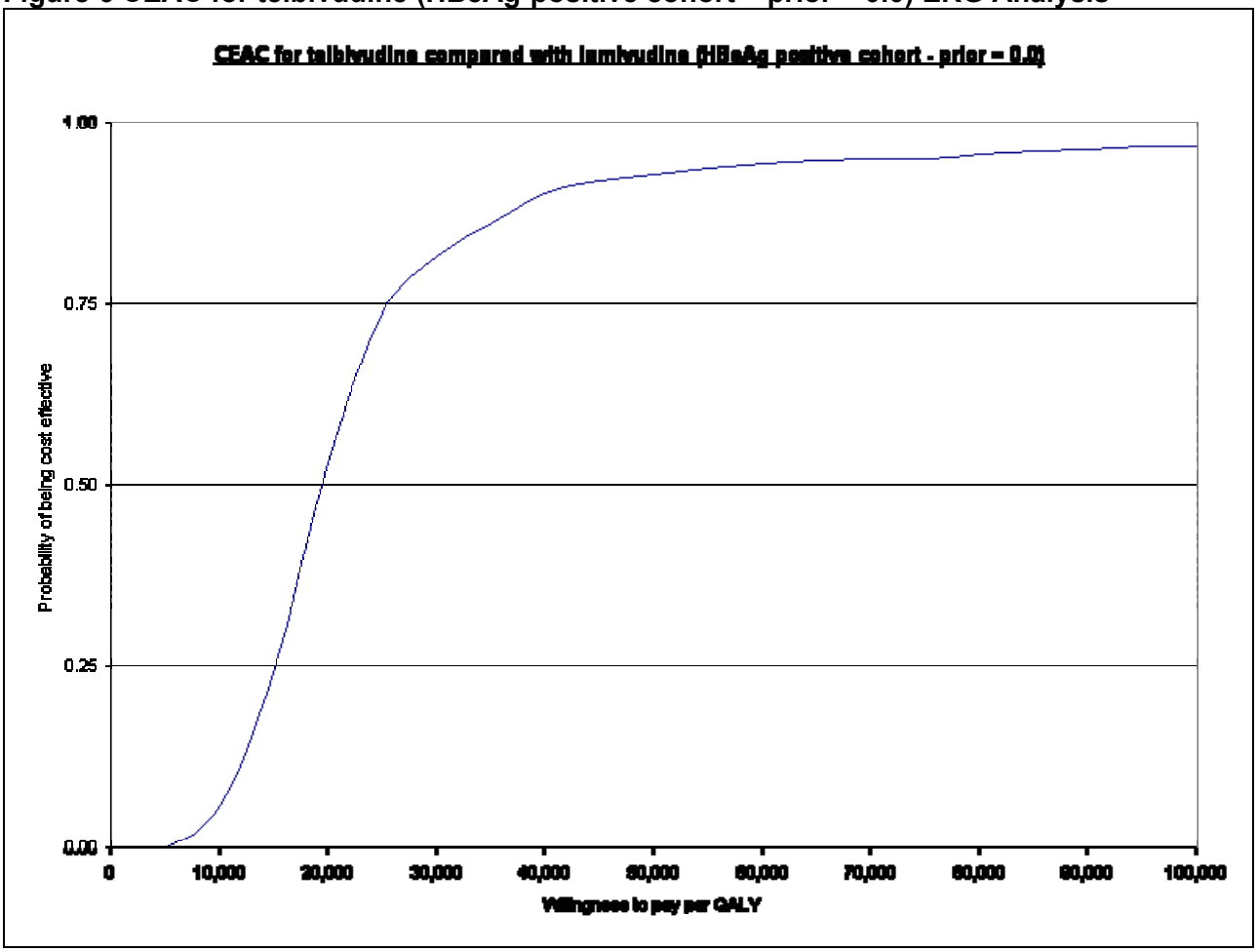
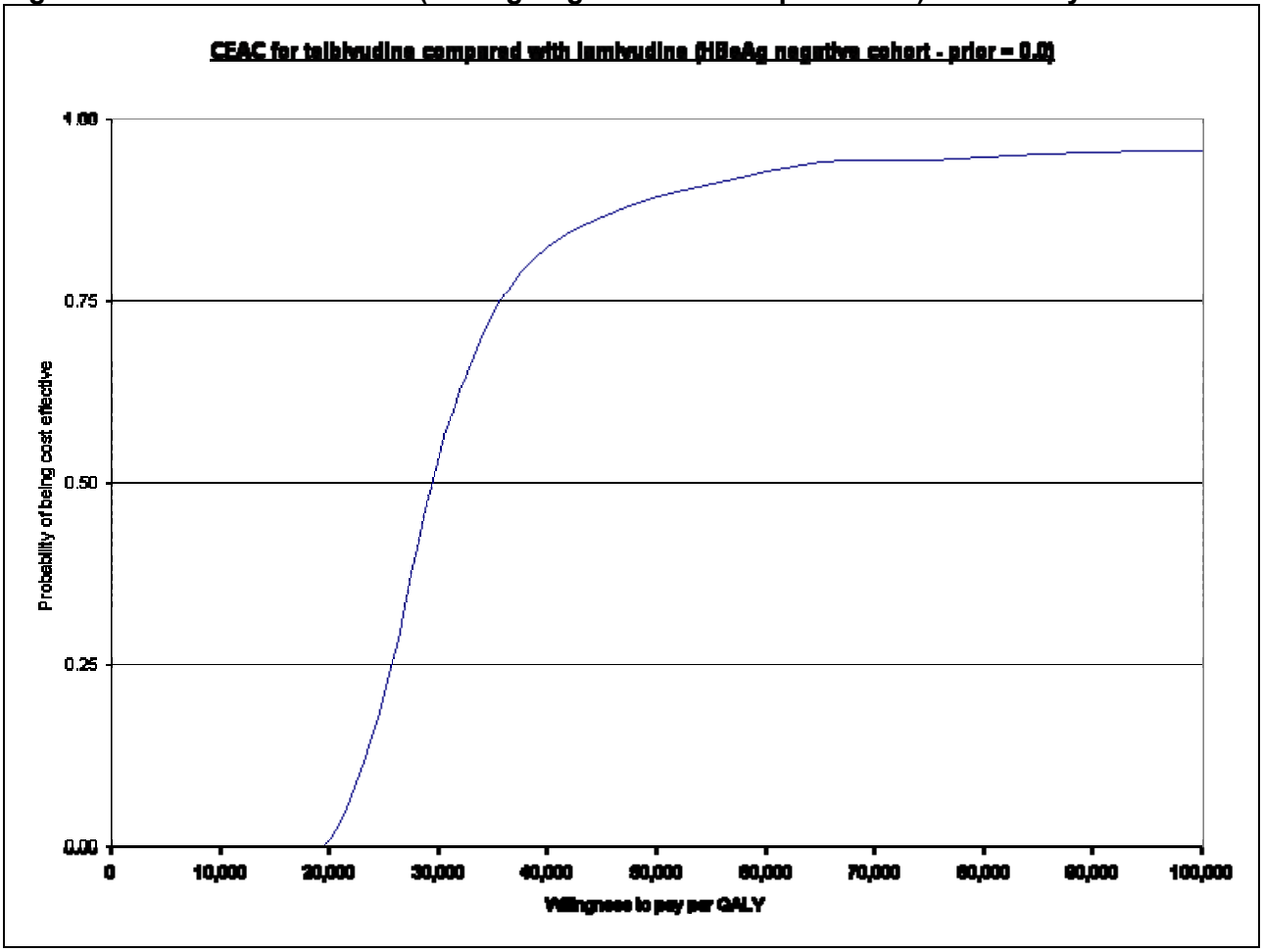


Figure 7 CEAC for telbivudine (HBeAg-negative cohort – prior = 0.0) ERG Analysis



The effect of changes in assumptions adopted by the ERG is to reduce the cost-effectiveness of telbivudine compared with lamivudine. In these analyses for the positive HBeAg model, the probability that telbivudine was cost effective at £20,000 and £30,000 per QALY gained was 0.53 and 0.82 respectively. For patients with HBeAg-negative CHB the probability that telbivudine was cost effective, compared with lamivudine, at £20,000 and £30,000 per QALY gained was 0.01 and 0.54 respectively.

Seroconversion model

The ERG conducted a probabilistic analysis after removing the benefits (and costs) of telbivudine treatment for patients with decompensation (discussed in 4.4.1.1 and illustrated in section 4.4.1.4.4). Additional assumptions in the ERG probabilistic sensitivity analysis are:

- Removing resistant patients from denominators of treatment transition probabilities (see section 4.4.1.2.2

- The mean effectiveness of adefovir (in terms of HBeAg seroconversion) was modelled at the mean of the values for telbivudine and lamivudine (as in the MS), but was sampled independently using the mean events and non-events for telbivudine and lamivudine (based on Globe data) as parameters for beta distributions (Year 1 values: α=72, β=202; Year 2 values: α=27, β=165).
- Adefovir resistance was included in the probabilistic sampling. Values from the HTA monograph[10] adopted for the MS (see Table F3, Appendix F of the MS) were used as mean values. Observed events (α) and populations at risk (α+β) were taken from Locarini and colleagues[39] and used as parameters for beta distributions.

Cost effectiveness acceptability curves from the ERG's PSA are shown in Figure 8 and the cost effectiveness acceptability frontier is shown in Figure 9 (derived using the same method as for Figure 5, discussed in the previous section of this report).

Figure 8 Cost effectiveness acceptability curves, seroconversion model (ERG's PSA)

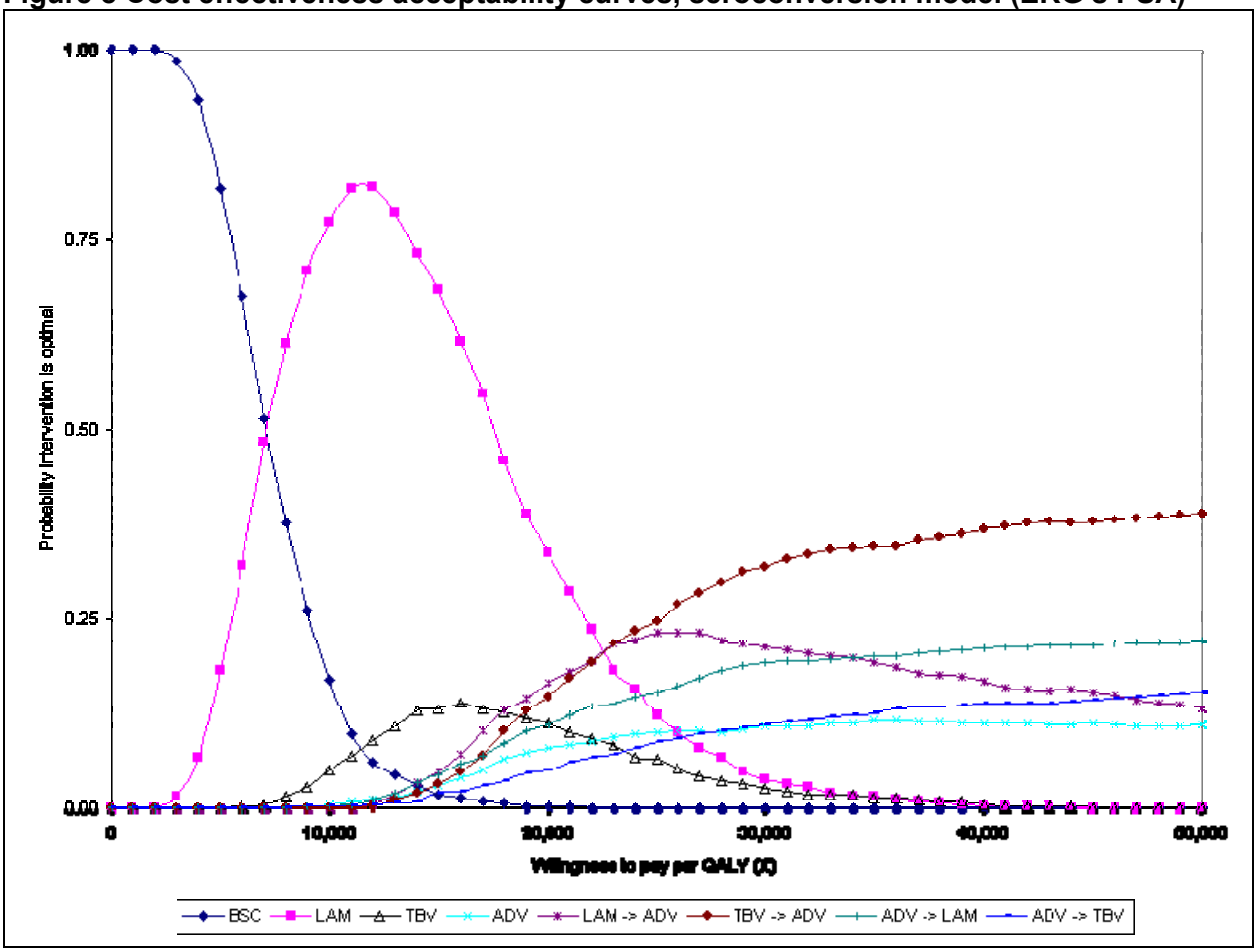
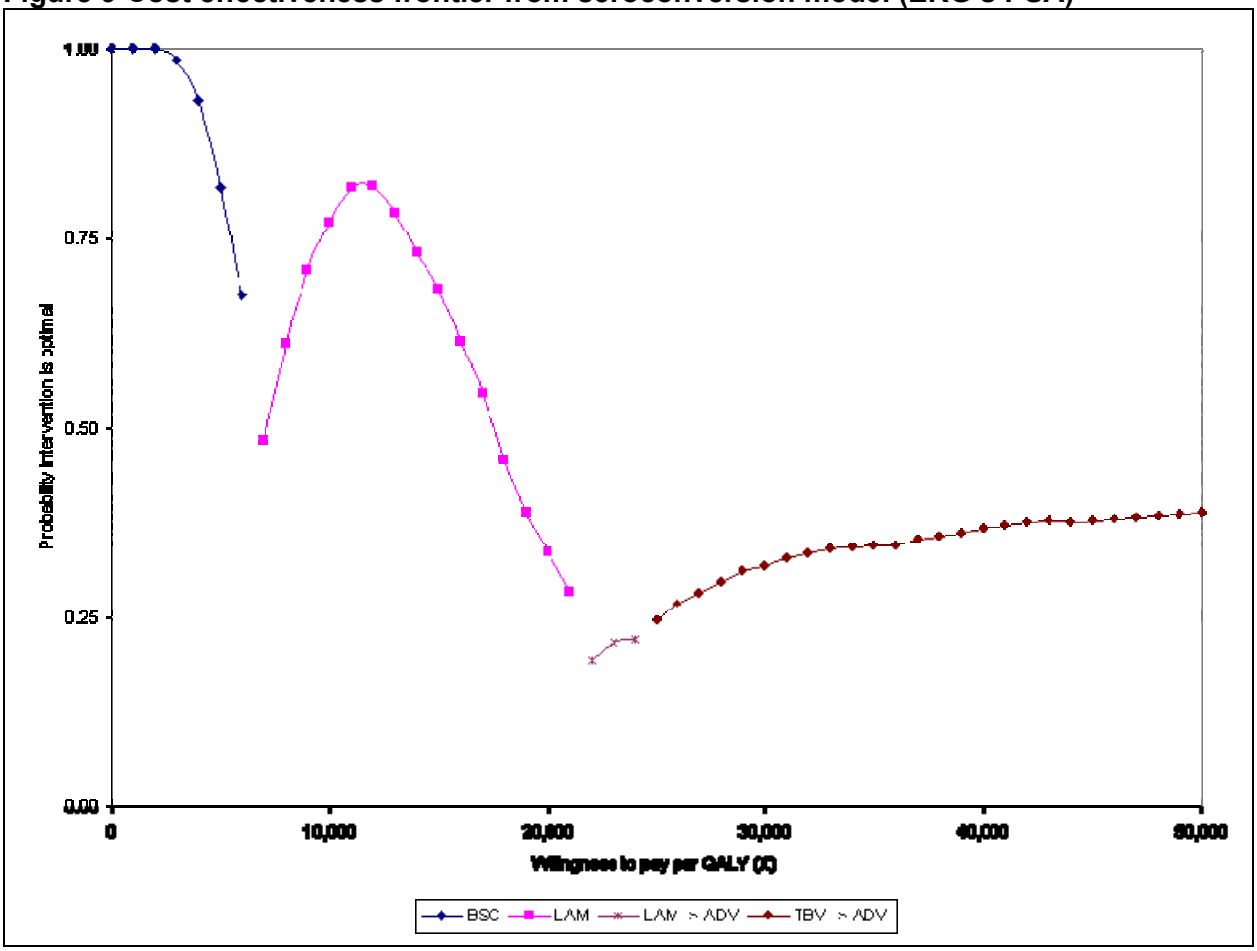


Figure 9 Cost effectiveness frontier from seroconversion model (ERG's PSA)



The effect of changes in assumptions adopted by the ERG is to reduce the cost-effectiveness of telbivudine compared with other treatment options. In particular, lamivudine is optimal (using the maximum net benefit criterion) over a wider range of willingness to pay, with lamivudine followed by adefovir being optimal over the range £22,000 to £24,000 per QALY (rather than telbivudine, as in the manufacturer's PSA). Telbivudine followed by adefovir remains the optimal strategy at higher values of willingness to pay.

Table 28 Probability intervention is cost effective at selected willingness to pay thresholds (ERG's PSA)

Treatment strategy	Threshold willingness to pay	
	£20,000 per QALY	£30,000 per QALY
Lamivudine	34%	4%
Telbivudine	11%	2%
Lamivudine followed by adefovir	17%	21%
Telbivudine followed by adefovir	15%	32%

Conclusions

The ERG applied limited changes in assumptions in the probabilistic sensitivity analysis of the viral load model. Due the time taken to complete the PSA in the viral load model the ERG only attempted this analysis for the model with a prior of zero – to allow a comparison with the results from the deterministic sensitivity analyses reported in sections 4.4.1.4.2 and 4.4.1.4.4. The ERG PSA using the viral load model reports results of probabilistic evaluation of the model after replacing the constant health state utility for the "cured" health state with age specific utilities, as in the HTA monograph,[10] and after applying the calibration factors for the advanced liver disease risk equations reported in Appendix C of the MS, rather than the values in the submitted electronic models. The probabilistic means are similar to the results from the deterministic analysis using the model with a prior of zero. The effect of applying the ERG's alternative assumptions is to shift the CEACs to the right (reducing the probability of telbivudine being cost effective at each given willingness to pay threshold) rather than changing the shape of the curve.

The results of the manufacturer and ERG's PSA using the seroconversion model are robust to changes in assumptions, with regard to:

- best supportive care (no anti-viral treatment is the optimal strategy at low values of willingness to pay per QALY)
- lamivudine (optimal over willingness to pay range £8,000 to £20,000 per QALY)
- telbivudine followed by adefovir (optimal over willingness to pay range £25,000 to £50,000 per QALY)
- sequential treatment strategies with adefovir as first-line treatment (never optimal over the range up to £50,000 per QALY).

The results of the manufacturer and ERG's PSA were not robust to changes in assumptions, with regard to the optimal strategy over the range £20,000 to £25,000 per QALY. In the manufacturer's PSA telbivudine was optimal over this range, while in the ERG's PSA lamivudine followed by adefovir was optimal.

The sequence of treatment options implied in both sets of PSA results is problematic since the strategy of using telbivudine as first-line treatment followed by adefovir (for patients who develop resistance to telbivudine) is not accessible to patients who have lamivudine as their first line treatment. In order to provide the treatment strategy of telbivudine followed by adefovir (which yields the greatest QALY gain of all the strategies in the seroconversion model and

which is optimal at a willingness to pay greater than £25,000 per QALY) telbivudine must be available as a first-line treatment.

4.4.2 Comment on validity of results presented with reference to methodology used

In general, the approach taken to modelling cost-effectiveness in this patient group seems reasonable and is consistent with methods adopted in previous economic evaluations of anti-viral treatment for CHB. A number of concerns have been raised by the ERG, with respect to:

- the risk equations applied for developing advanced liver disease adopted in the viral load model
- the complexity and lack of transparency of the submitted economic models, particularly in the viral load model
- the lack of discussion in the MS over data quality (primarily completeness), the powering of the trial for the sub-group of patients included in the economic model (ALT ≥ 2 x ULN) and potential biases (for example randomisation was not stratified by ALT at 2 x ULN and no baseline characteristics were presented for this sub-group of patients in the MS).
- the lack of discussion in the MS regarding the role of the prior values in the analysis.

A key concern is the selection of comparators included in the economic models (principally the exclusion of entecavir from the economic models) and the approach to including evidence of effectiveness of adefovir in the seroconversion model. An additional concern is the lack of discussion of uncertainty in the results, in the main body of the submission.

4.4.3 Summary of uncertainties and issues

- The economic analysis is not consistent with the scope for this assessment. Placing telbivudine as an alternative to lamivudine (for patients intolerant of or unwilling to accept conventional or pegylated interferon) seems reasonable. Exclusion of entecavir on the basis of either the “naïve” comparison or the indirect comparison reported in Appendix A of the MS does not seem justified. Even if there were no methodological questions around the indirect comparison exclusion of entecavir from the economic model would not be justified:

  - Firstly, lack of statistically significant differences in treatment outcomes is not a justification for excluding a valid comparator from the economic model;
  - Secondly the submission ignores the impact of the resistance profile for entecavir, though MS indicates uncertainty as to the resistance to entecavir. This could be addressed by clear statement of assumptions regarding resistance and careful sensitivity analyses.
- Evidence of the comparative effectiveness of adefovir is not adequately addressed in the MS. No searches for evidence to include adefovir in the indirect comparison were undertaken and there was no critical assessment of data taken from HTA report. The MS contains no attempts to justify the use of assumptions in the seroconversion model and there is no assessment of how these assumptions might affect the outputs from the seroconversion model.
- The disease progression model for estimating transitions to compensated cirrhosis and HCC does not appear to have been published, in form used in the economic model. The risk models are not adequately described or presented in the MS and insufficient detail is provided of the method of adjusting for age and the impact of cirrhosis on risk of developing HCC. The MS reports no evidence of searches to find alternative (validation) sources to justify the use of the adjusted risk models.
- The submitted electronic models are heavily reliant on Visual Basic, which reduces transparency and makes them difficult to check. The viral load model is unwieldy, with data requirements that preclude the inclusion of relevant comparators. The MS does not consider whether a simpler model, based on viral load, could have been developed that would retain clinical validity. The MS contains no evidence of clinical validation of the viral load model.
- There is very limited discussion in the MS on the differences between viral load model with zero prior and with 0.5 prior. While the ICERs are not substantially different, and application of priors of zero or 0.5 do not produce contradictory results, the distribution of costs and outcomes (and the correlation between costs and outcomes) differ markedly. The MS does not report any investigation of the differences between models with different priors nor does it indicate which input variables are most affected by differences in prior values.
- There is insufficient attention given, in the MS, to data quality. Given that the model relies on data observed in the Globe study to derive transition probabilities, more consideration should have been given to assuring consistency of the data used to

populate the model and to explain apparent inconsistencies in denominators and missing data. The ERG cannot be certain that some differences are due to missing data rather than real differences in efficacy.

5 Discussion

5.1 Summary of clinical effectiveness issues

The clinical evidence for telbivudine comes from a single RCT[1] which was the licensing trial for telbivudine in patients with HBeAg-positive and HBeAg-negative CHB. Results on the efficacy of telbivudine are mainly focussed on 104 week unpublished data (results for 52 weeks are also provided), and are largely reported separately for the HBeAg sub-groups. It is not clear whether the study was powered to detect differences in the race/ethnicity or ALT sub-group sub-sets, and an over-representation of HBeAg-positive patients in the RCT may have influenced the results of the HBeAg-negative sub-group. Not all the comparators issued in the NICE scope were included in the MS, being excluded as inappropriate due to their place in the treatment pathway. Not all outcome measures specified in the scope were included in the MS.

Although statistically different, the difference between telbivudine and lamivudine is not clinically significant, having an effectiveness advantage of only about 2% in patients treated between the two drugs. No mention is made in the MS of the high viral resistance rate and the clinical impact this has on CHB patients. Viral breakthrough (>1 log increase over nadir) for telbivudine was 28.6% at two years; whilst this is lower than lamivudine (45.5%), it is still high at a clinical level.

The indirect comparison with entecavir was poor and should be treated with caution. The MS provide an inadequate description of the methodology, a systematic review of the two entecavir trials was omitted, and the conclusions are largely based on a visual comparison of efficacy outcomes. An indirect statistical comparison is presented in an Appendix, but as no meta-analysis could be undertaken the ERG would question its validity. Viral resistance rates of entecavir were not reported in the MS. Overall, telbivudine appears to be approximately the same efficacy as entecavir for viral suppression but appears to have markedly higher rates of viral resistance (as per the rates for entecavir reported in the published trials).

5.2 Summary of cost effectiveness issues

The model structures adopted for the cost effectiveness analysis are based on those adopted for previous economic evaluations of anti-viral treatment of CHB and seem appropriate for this analysis. However the economic models submitted by the manufacturer are large, complex and are highly reliant on Visual Basic for reprocessing of input data and to generate results. As a result it is difficult to test the models and to ensure that they are error free. Two sets of results have been submitted for the viral load model – the original results are reported in Table 4 (section 4.2.8 of this report) and a corrected set of results, submitted following a request for clarification from the ERG, are included in Appendix 1 of this report.

Evidence of the effectiveness of telbivudine comes from a clinical trial comparing telbivudine with lamivudine.[1] The economic model uses data from a sub-set of patients, with ALT ≥ 2 x ULN, which have not been discussed or critically appraised in the clinical effectiveness section of the MS. Since these data are not presented in the MS, the key clinical effectiveness data in the economic model could not be critically appraised by the ERG.. There were no searches for evidence on the effectiveness of adefovir, which was included as a comparator in one of the economic models. Effectiveness of adefovir was taken from a HTA monograph or was based on assumptions. Entecavir was not included in any of the economic models. The MS concluded that telbivudine is a cost effective option when compared with lamivudine, in the viral load model – though the incremental cost effectiveness ratio is highly variable depending on the prior value adopted. The MS also concluded that telbivudine (alone or followed with adefovir salvage for patients who develop resistance to telbivudine) is cost effective using the seroconversion model.

The ERG has concerns over the selection of comparators in the models, particularly the exclusion of entecavir, and the methods used to incorporate evidence for treatments not included in the Globe trial. There is insufficient discussion of the risk equations used to model progression to advanced liver disease in the viral load model and of the impact of values adopted for the prior in the viral load model. The ERG found discrepancies between input values for the viral load model reported in appendices to the MS and values used in the submitted electronic models. Replacing values in the models with those from the appendices lead to a lower QALY gain for telbivudine compared with lamivudine and a less favourable incremental cost effectiveness ratio for telbivudine compared with lamivudine.

There is insufficient critical discussion of the data (from the Globe trial) used to populate the models. There is no discussion of the baseline comparability of patients in the ALT ≥ 2 x ULN sub-group randomised to each treatment or of potential biases due to missing data. Sensitivity analyses undertaken by the ERG have been able to address a limited number of these concerns.

6 References

1. Globe study NV-02B-007. *Week 104 Clinical Study Report.* 2007.

2. National Institute for Health and Clinical Excellence. Adefovir dipivoxil and peginterferon alfa-2a for the treatment of chronic hepatitis B. NICE Technology Appraisal TA96. 2006.

3. Hou J, Yin Y-K, ., Xu D, Tan D, Niu J *et al*. Telbivudine versus lamivudine in Chinese patients with chronic hepatitis B: Results at 1 year of a randomized, double-blind trial. *Hepatology* 2007;46:(early online viewing prior to publication).

4. The Hepatitis B Foundation. Rising Curve: Chronic Hepatitis B Infection in the UK. http://www.hepb.org.uk/news/current/new_report_from_hepatitis_b_foundation_uk/rising_curve.pdf. Accessed 29th January 2008

5. British Liver Trust. How prevalent is Hepatitis B? http://www.britishlivertrust.org.uk/home/health-professionals/literature-for-professionals/a-professionals-guide-to-hepatitis-b/how-prevalent-is-hbv.aspx Accessed 17th December 2007

6. Lok A, McMahon B. Chronic hepatitis B: AASLD Practice Guidelines. *Hepatology* 2007;45:507-39.

7. The EASL Jury. EASL International Consensus Conference On Hepatitis B: 13-14 September, 2002 Geneva, Switzerland. Consensus statement (Long version). *Journal of Hepatology* 2003;39:S3-S25.

8. Liaw Y, Leung N, Guan R, Lau G, Merican I, McCaughan G *et al*. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2005 update. *Liver International* 2005;25:472-89.

9. Chan HLY, Heathcote EJ, Marcellin P, Lai CL, Cho M, Moon YM *et al*. Treatment of Hepatitis B e Antigen Positive Chronic Hepatitis with Telbivudine or Adefovir: A Randomized Trial. *Ann Intern Med* 2007;147:745-54.

10. Shepherd, J, Jones, J, Takeda, A, Davidson, P, and Price, A. Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation. *Health Technol Assess* 10(28). 2006.

11. Lai CL, Gane E, Liaw YF, Thongsawat S, Wang Y, Chen Y *et al*. Telbivudine (LDT) vs. Lamivudine for chronic hepatitis B: First-year results from the international phase III globe trial. *Hepatology* 42(4 suppl. 1), 748A. 2005.

12. Lai CL, Gane E, Liaw YF, Thongsawat S, Wang Y, Chen Y. *et al*. Maximal early HBV suppression is predictive of optimal virologic and clinical efficacy in nucleoside-treated hepatitis B patients: scientific observations from a large multinational trial (the Globe study). *Hepatology* 42(4 suppl. 1), 232A-233A. 2005.

13. Lai CL, Gane E, Liaw YF, Hsu C-W, Thongsawat S, Wang Y *et al*. Telbivudine versus Lamivudine in Patients with Chronic Hepatitis B. *N Engl J Med* 2007;357:2576-88.

14. Gane E, Safadi R, Xie Q, Chen Y, Yin Y, Wei L *et al*. A Randomized Trial of Telbivudine (LdT) versus Lamivudine in Lamivudine Experienced Patients - Week 24 Primary Analysis. (Abstract). American Association for the Study of Liver Diseases Conference 2006.

15. Chang TT, Gish RG, de Man R, Gadano A, Sollano J, Chao YC *et al*. A comparison of entecavir and lamivudine for HBeAg-positive chronic hepatitis B. *N Engl J Med* 2006;354:1001-10.

16. Lai CL, Shouval D, Lok AS, Chang TT, Cheinquer H, Goodman Z *et al*. Entecavir versus Lamivudine for Patients with HBeAg-Negative Chronic Hepatitis B. *N Engl J Med* 2006;354:1011-20.

17. Glenny AM, Altman DG, Sakarovitch C, Deeks JJ, Amico RD, Bradburn M *et al*. Indirect comparisons of competing interventions. *Health Technol Assess* 2005;9:1-148.

18. Yao G, Chen C, Lu W, Ren H, Tan D, Wang Y *et al*. Efficacy and safety of entecavir compared to lamivudine in nucleoside-naïve patients with chronic hepatitis B: a randomized double-blind trial in China. *Hepatol Int* 2007;1:365-72.

19. Centre for Reviews and Dissemination (CRD). *Undertaking Systematic Reviews of Research on Effectiveness*. Report No. 4. York, 2001.

20. Crowley S, Tognarini D, Desmond P, Lees M, Saal G. Introduction of lamivudine for the treatment of chronic hepatitis B: expected clinical and economic outcomes based on 4-year clinical trial data. *J Gastroenterol Hepatol* 2002;17:153-64.

21. Crowley SJ, Tognarini D, Desmond PV, Lees M. Cost-effectiveness analysis of lamivudine for the treatment of chronic hepatitis B. *Pharmacoeconomics* 2000;17:409-27.

22. Dusheiko GM, Roberts JA. Treatment of chronic type B and C hepatitis with interferon alfa: an economic appraisal. *Hepatology* 1995;22:1863-73.

23. Wong JB, Koff RS, Tine F, Pauker SG. Cost-effectiveness of interferon-alpha 2b treatment for hepatitis B e antigen-positive chronic hepatitis B. *Ann Intern Med* 1995;122:664-75.

24. Fattovich G. Natural history of hepatitis B. *Journal of Hepatol* 2003;39:50-8.

25. Iloeje UH, Yang HI, Su J, Jen CL, You SL, Chen CJ *et al*. Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology* 2006;130:678-86.

26. Chen CJ, Yang HI, Su J, Jen CL, You SL, Lu SN *et al*. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA* 2006;295:65-73.

27. Joint Formulary Committee. *British National Formulary*. No. 54. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 2007.

28. Curtis L, Netten A. *Unit Costs of Health and Social Care*. Canterbury: Personal Social Services Research Unit, University of Kent, 2006.

29. Drummond M, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 1997.

30. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R *et al*. A review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8.

31. Wright, M., Grieve, R., Roberts, J., Main, J., Thomas, H. C., and on behalf of the UK Mild Hepatitis C Trial Investigators. Health benefits of antiviral therapy for mild chronic hepatitis C: randomised controlled trial and economic evaluation. *Health Technol Assess* 10(21). 2006.

32. Ratcliffe J, Longworth L, Young T, Bryan S, Burroughs A, Buxton M *et al*. Assessing health-related quality of life pre- and post-liver transplantation: a prospective multicentre study. *Liver Transpl* 2002;8:263-70.

33. Orlewska E. The cost-effectiveness of alternative therapeutic strategies for the management of chronic hepatitis B in Poland. *Value in Health* 2002;5:405-21.

34. Kanwal F, Gralnek IM, Martin P, Dulai GS, Farid M, Spiegel BM. Treatment alternatives for chronic hepatitis B virus infection: a cost-effectiveness analysis. *Ann Intern Med* 2005;142:821-31.

35. Centre for Bayesian Statistics in Health Economics. Sampling from multivariate distributions. http://www.shef.ac.uk/chebs/software . 2008.

36. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press, 2006.

37. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* 2001;10:779-87.

38. Kind P. *Population norms for EQ-5D*. York: Centre for Health Economics, University of York; 1999. CHE Discussion Paper 172

39. Locarini S, Qi X, Arterburn S, Snow A, Brosgart CL, Curfie G *et al*. Incidence and predictors of emergence of adefovir resistant HBV during 4 years of adefovir dipivoxil (ADV) therapy for patients with chronic hepatitis B (CHB). *Journal of Hepatol* 2005;42:17.

7 Appendices

7.1 Appendix 1 Response to ERG clarification questions

Section A. Clarification on systematic review

A1. What were the inclusion/exclusion criteria? Please provide a description of the processes undertaken in applying the inclusion and exclusion criteria, the data extraction and the quality assessment of the trials. Please provide a flow chart (for example as outlined in the QUORUM statement flowchart) to clarify the destination of the trials retrieved and explain why trials/ papers were excluded. .

A2. In section 5.2.1, p.24, it is stated that 798 citations were identified on Medline, Embase and Cochrane databases of which there were 7 relevant studies and 2 were included. Please provide details of these 2 and the other 5 references.

A further 5 relevant RCTs, 2 relevant abstracts and one 'in-house' trial are also mentioned – please provide these references as well as an explanation as to the reasons for exclusion and how these relate to the above studies.

For the purpose of clarity, response to questions A1-2 are compiled in the following section

Description of search strategies

Literature searches were conducted using the Medline, EMBASE, Cochrane and Novabase (eNova) databases as well as of the trial registries. An examination of the telbivudine registration dossier was also conducted as well as a manual search of relevant publications. The search strategies, inclusion and exclusion criteria are provided below and all unique citations are included in Appendix A.

Inclusion Criteria

Population: Patients with chronic hepatitis B
Intervention: telbivudine (Sebivo)
Comparator: lamivudine (also include Adefovir)
Outcomes: primary and secondary outcomes (changes in HBV DNA, HBeAg, HBeAb, HBsAg, HBsAb, ALT)
Study Design: RCTs

Exclusion Criteria

a)	Not a randomised trial;
b)	Randomised trial does not include the proposed drug and the main comparator in separate arms;
c)	Characteristics of the recruited participants do not overlap with the main indication;
d)	Preclinical or pharmacological trial (e.g. pharmacokinetic or pharmacodynamic study without clinical endpoints) or mechanistic/*in-vitro* studies with biochemical, cellular or molecular endpoints;
e)	Dose finding, dose ranging or early phase study, of relevant drug(s), or very small patient numbers;
f)	Studies or meta-analyses where the outcome measure is not relevant to the analysis (e.g.);
g)	Duplicate publication of a study;
h)	Study published in abstract form only;

i) Review of economic data;
j) Review of quality of life data;
k) Trial results not available.

Identification and Selection of Studies

- Two reviewers independently screen all titles/abstracts
- Full manuscripts of potentially relevant studies were ordered
- All the full potentially relevant studies using inclusion/exclusion criteria were assessed
- Disagreement resolved through discussion

The Medline, EMBASE and Cochrane literature searches were performed using the Ovid platform and the search outputs were combined in Reference Manager. The total number of citations from each database, the number of duplicates and the number of unique citations are shown in Table 1.

Telbivudine registration dossier

An examination of the Summary of Clinical Data was conducted for any further phase III randomised trials involving telbivudine. This search resulted in eight clinical trials eligible for the inclusion in the systematic review.

Manual search for other relevant citations

A manual search for other relevant citations was performed. This resulted in 2 citation retrieved: these were publications of the identified trials in other databases.

Listing of all direct randomised trials

Table 1 documents the total number of citations retrieved from the respective databases and the number of citations excluded, for each of the reasons listed in the table. The search of the EMBASE, Medline and Cochrane databases retrieved five relevant citations for inclusion. The eNOVA database (a database internal to Novartis) retrieved six citations which were abstracts of relevant randomised trials. The Trial registry search retrieved no relevant randomised trials. Detailed information on reasons why trials were not included can be found in Appendix A.

Table 1: Summary of identification of direct randomised trials from the search of published literature

	EMBASE	MEDLINE	COCHRANE	Trial registries	eNOVA
• Number of citations retrieved by search	497	372	215	22	7
• Number of duplicates removed	0	181	134	0	
• Number of unique citations	769			22	7
Citations excluded after title/abstract review:					
a) Not a randomised trial;	516			5	
b) Randomised trial does not include the proposed drug and the main comparator in separate arms;	211			15	
c) Characteristics of the recruited participants do not overlap with	18			2	

	EMBASE	MEDLINE	COCHRANE	Trial registries	eNOVA
the main indication;					
d) Preclinical or pharmacological trial (e.g. pharmacokinetic or pharmacodynamic study without clinical endpoints) or mechanistic/*in-vitro* studies with biochemical, cellular or molecular endpoints;	5				
e) Dose finding, dose ranging or early phase study, of relevant drug(s), or very small patient numbers;	0				1
f) Studies or meta-analyses where the outcome measure is not relevant to the analysis;	2				
g) Duplicate publication of a study;	5				
h) Study published in abstract form only;	0				
i) Review of economic data;	7				
j) Review of quality of life data;	0				
k) Trial results not available.	0				
TOTAL exclusions after title/abstract review	764			22	1
Citations excluded after full text review:					
a) Not a randomised trial;					
b) Randomised trial does not include the proposed drug and the main comparator in separate arms;					
c) Characteristics of the recruited participants do not overlap with the main indication;					
d) Preclinical or pharmacological trial (e.g. pharmacokinetic or pharmacodynamic study without clinical endpoints) or mechanistic/*in-vitro* studies with biochemical, cellular or molecular endpoints;					
e) Dose finding, dose ranging or early phase study, of relevant drug(s), or very small patient numbers;					
f) Studies or meta-analyses where the outcome measure is not relevant to the analysis;					

	EMBASE	MEDLINE	COCHRANE	Trial registries	eNOVA
g) Duplicate publication of a study;					
h) Study published in abstract form only;					
i) Review of economic data;					
j) Review of quality of life data;					
k) Trial results not available.					
TOTAL exclusions after full text review					
• Number of citations of direct randomised trials included from each database	5			0	6
• Consolidated number of citations of direct randomised trials (removing exact duplicates across different databases)	7 (a)				
• **Number of published direct randomised trials included**	**2 (b)**				

Note: Present columns that correspond with submitted printouts (e.g if the printouts combine MEDLINE and EMBASE, theses results can be combined in the table)

Detailed information on reasons why trials were not included can be found in Appendix A.

(a) The list of the 7 citations of direct randomised trials is as follow:

No.	Articles
1.	Title: Maximal early HBV suppression is predictive of optimal virologic and clinical efficacy in nucleoside-treated hepatitis B patients:Scientific observations from a large multinational trial (The GLOBEStudy)Author(s): **Lai C-L** ; Gane E ; Liaw Y-F ; Thongsawat S ; Wang Y ; Chen Y ;Heathcote EJ ; Rasenack J ; Bzowej N ; Naoumov N ; Chao G ;Constance BF ; Brown NA ;Source: *Hepatology 2005; 42 (4 SUPPL. 1), 232A-233A [ISSN0270-9139]*
2.	Title: A 1-year trial of telbivudine, lamivudine, and the combination inpatients with hepatitis B e antigen-positive chronic hepatitis B. Author(s): **Lai C-L** ; Leung N ; Teo E-K ; Tong M ; Wong F ; Hann H-W ; Han S ;Poynard T ; Myers M ; Chao G ; Lloyd D ; Brown NA ;Source: *Gastroenterology 2005; 129 (2), 528-536 [ISSN0016-5085]*
3.	Title: International multicenter trial of LDT (telbivudine), alone and in combination with lamivudine, for chronic hepatitis B: An interim analysisAuthor(s): **Lai C-L** ; Leung N ; Teo EK ; Tong M ; Wong F ; Hann H-W ; Han S ;Poynard T ; Myers M ; Chao G ; Lloyd D ; Brown NA ;Source*: Hepatology 2002; 36 (4 PART 2), 301A [ISSN0270-9139]*
4.	Title: A phase IIb comparative trial of LdT, lamivudine, and the combination in hepatitis B patients: Greater antiviral effect with LdT Author(s): **Lai CL** ; Leung N ; Teo EK ; Tong M ; Wong F ; Hann HW ; Han S ;Poynard T ; Myers M ; Chao G ; Lloyd D ; Brown N ;Source: *Abstracts of the Interscience Conference on Antimicrobial Agents and Chemotherapy 2003; 43 502 [ISSN1532-0227]*
5.	Title: Results of a one-year international phase IIB comparative trial of telbivudine, lamivudine, and the combination, in patients with chronic hepatitis B Author(s): **Lai C-L** ; Leung NWY ; Teo E-K ; Tong M ; Wong F ; Hann H-W ; Han S; Poynard T ; Myers M ; Chao G ; Lloyd D ; Brown N ;Source*: Hepatology 2003; 38 (4 SUPPL. 1), 262A [ISSN0270-9139]*

6.	Title: Results of a one-year international phase IIB trial of LDT, and LDT plus lamivudine, in patients with chronic hepatitis B Author(s): **Han SH** ; Leung NWY ; Teo EK ; Tong M ; Wong F ; Hann HW ;Poynard T ; Brown NA ; Myers M ; Chao G ; Lloyd D ; Lai CL ;Source: Journal of *Hepatology 2004; 40 (SUPPL. 1), 16 [ISSN0168-8278]*
7.	Title: Improving antiviral therapy in chronic hepatitis B: maximal viral suppression at week 24 correlates with better clinical efficacy at one year [EASL abstract]. **Poynard T.** *Journal of Hepatology 2004; 40(Suppl 1):130.*

(b) The above citations are representative selected from the above list to be considered for the systematic review are as follow:

No.	Articles
1	Title: Maximal early HBV suppression is predictive of optimal virologicand clinical efficacy in nucleoside-treated hepatitis B patients:Scientific observations from a large multinational trial (The GLOBEStudy)Author(s): Lai C-L ; Gane E ; Liaw Y-F ; Thongsawat S ; Wang Y ; Chen Y ;Heathcote EJ ; Rasenack J ; Bzowej N ; Naoumov N ; Chao G ;Constance BF ; Brown NA ;Source: Hepatology 2005; 42 (4 SUPPL. 1), 232A-233A [ISSN0270-9139]
2	Title: A 1-year trial of telbivudine, lamivudine, and the combination inpatients with hepatitis B e antigen-positive chronic hepatitis B. Author(s): **Lai C-L** ; Leung N ; Teo E-K ; Tong M ; Wong F ; Hann H-W ; Han S ;Poynard T ; Myers M ; Chao G ; Lloyd D ; Brown NA ;Source: *Gastroenterology 2005; 129 (2), 528-536 [ISSN0016-5085]*

Table 2 presents the citations identified from the telbivudine registration dossier, other “in house” trials and a manual search. There were five direct randomised trials that were identified from the search of the telbivudine TGA registration dossier as being relevant for potential inclusion in the submission. The manual search retrieved two further abstracts of a relevant randomised trial. One “in house” trial was identified. These 8 citations were considered for potential inclusion in the literature review.

Table 2: Summary of identification of sponsor's direct randomised trials and information from the manual search of retrieved citations

	TGA dossier	Other "in-house" trials	Manual search	Total
Number of reports or citations of randomised trials retrieved	7	1	2	10
Randomised trials excluded:				
b) randomised trial does not include the proposed drug and the main comparator in separate arms	1			1
c) randomised trial does not include the proposed drug and the main comparator in separate arms	1			1
TOTAL excluded:	2			2
• Number of direct randomised trials included from these searches	5	1	2	8
• Number of these direct randomised trials identified in Table	0		0	0
• Number of other direct randomised trials identified in Table	0		0	0
• **Total direct randomised trials considered for potential inclusion in the submission**	**5**	**1**	**2**	**8 (c)**

TGA = Therapeutic Goods Administration

a For the purposes of the search for relevant randomised trials, 'sponsor' includes any original sponsor (including head office and all subsidiaries) and/or any co-licensing sponsor of the proposed drug in addition to the sponsor lodging the submission.

b Separately list and identify each of these trials using the identifying nomenclature used for the trials in the TGA evaluation reports to enable a cross-check against the trials considered by the TGA.

(c)The list of the 8 citations of direct randomised trials considered for potential inclusion in the submission is as follow:

	Title
1.	Study No. NV-02B-003: A randomized, double-blind study of treatment with telbivudine (LdT), lamivudine, or the combination of both agents in adults with HBeAg-positive chronic hepatitis B
2.	Study No. NV-02B-007: A randomized, double-blind trial of LdT (telbivudine) versus lamivudine in adults with compensated chronic hepatitis B – Primary analysis of week 52 data
3.	Study No. NV-02B-010: A phase IIb extension study of LdT (telbivudine), lamivudine, or LdT plus lamivudine in patients with chronic hepatitis B who have completed study NV-02B-003 - Interim clinical study report
4.	Study No. NV-02B-011: A randomized, double-blind trial of telbivudine (LdT) versus lamivudine in adults with decompensated chronic hepatitis B and evidence of cirrhosis
5.	Study No. NV-02B-015: A phase III randomized, double blind trial of LdT (telbivudine) versus lamivudine, in chinese adults with compensated chronic hepatitis B
6.	Study No. NV-02B-018: A randomized, open label trial of telbivudine (LdT) versus adefovir dipivoxil in adults with HBeAg-positive, compensated chronic hepatitis B
7.	Study No. NV-02B-019: A randomized trial of switching antiviral therapy from lamivudine to telbivudine (LdT) vs. continued lamivudine treatment in adults with chronic hepatitis B
8.	Study No. NV-007/015 pooled A combination of the data from the global Phase III trial (Study 007) and the China specific Phase III trial (Study 015) to achieve 600 patients from China for an eventual regulatory submission in China.

In the submission only study NV-007 has been included. The reasons for excluding studies NV-003, NV-010, NV-011 are presented in the table below

Table 3: Reasons to exclude each trial from further detailed assessment

Trial ID	Ground(s) for seeking exclusion
Study 003	• Phase IIb trial: preliminary data from this trial were used for planning the phase III trial, and Study 003 consequently has been superseded by Study 007. • Small patient numbers: 104 patients were randomised in a 1:1:1:1:1 fashion over 5 treatment arms. This resulted in only 19 and 22 patients in the lamivudine 100 mg and telbivudine 600 mg treatment groups, respectively.
Study 010	• Phase IIb extension study (of Study 003): the secondary objective of this trial was to gather preliminary data regarding the clinical efficacy of the drugs (telbivudine, lamivudine/telbivudine combination) compared to lamivudine monotherapy, prior to obtaining data from Phase III clinical trials. Again, as for Study 003, Study 010 has been superseded by Study 007. • The study was not powered for its endpoints (maintenance of clinical benefits in Study 003 in the longer term) as Study 003 was originally intended as a 52 week study and no adjustments could be made in the sample size to accommodate loss of patients in the extension study. • Small patient numbers: the sample size was too small to prove that the differences in efficacy between the drugs that were observed in the study achieved statistical significance.
Study 011	• The study was performed in patients with decompensated CHB • Characteristics of recruited participants do not overlap with the main indication

A3. In section 5.2.3, p.26, it is stated that 4 RCTs were selected for inclusion (studies 007, 015, 018 and 019). Please provide the references for the latter 3 studies and reasons for exclusion (particularly of studies 015 and 019) and how these relate to the aforementioned included studies.

Error! Reference source not found. 4 provides the grounds for excluding Studies 015, 018 and 019 from further detailed assessment.

Study 015 was a relatively small, 2 year study (n=332), exclusively enrolling Chinese participants and is only partially reported with results at 1 year.

The justification for exclusion of 018 and 019 hinges on the fact that in neither study was telbivudine used in the context described in the Marketing Authorisation and Summary of Product Characteristics. Study 018, compares the efficacy of adefovir and telbivudine used alone or in sequence. Adefovir is inappropriate comparator since TA 096 only recommends adefovir as third-line therapy (after PegIFN and lamivudine), according to the telbivudine licence the drug cannot be use following lamivudine.

Study 019 examines the responses to telbivudine in patients that have been previously treated with lamivudine. Due to issues of cross-resistance between lamivudine and telbivudine, the use of these drugs in sequence is not recommended, is outside the licensed indication and is therefore not relevant to this appraisal.

Table 4: Reasons for excluding each trial from further detailed assessment

Trial ID	Ground(s) for seeking exclusion
Study 015 Appendix B	• **Phase III trial: telbivudine vs lamivudine in Chinese patients.** • Characteristics of recruited participants and HBV genotypes do not adequately represent the UK population under consideration in this appraisal • Study incomplete: 24 month study but only 12 month data available
Study 018 Appendix C	• **Randomised controlled open-label study: telbivudine vs adefovir vs adefovir followed by telbivudine** • Adefovir is inappropriate comparator since TA 096 only recommends adefovir as third-line therapy (after PegIFN and lamivudine) • Telbivudine would thus be used outside licensed indication (ie in patients previously using lamivudine) • Limited statistical power: small participant numbers (n<50 per arm)
Study 019 Appendix D	• **Randomised controlled open label:** telbivudine vs lamivudine in lamivudine experienced patients. • Telbivudine used outside licensed indication (ie in patients previously using lamivudine) • Results to week 24 only

A4. Medline was reported as having been used in clinical searches; did this include Medline in Progress? Also were systematic reviews, abstracts and conference proceedings eligible for inclusion or not?

Medline in Progress was not included. Abstracts and systematic reviews were eligible for inclusion but not conference proceedings.

Section B. Clarification on clinical effectiveness data

B1. In the power calculation (section 5.3.6, p.39) histological response is the key secondary efficacy endpoint. However, on p.35 it is stated that antiviral efficacy (HBV DNA level) is the key secondary efficacy endpoint. Please clarify this discrepancy.

To be precise, histological response was described as “the” key secondary endpoint on p 39 while HBV DNA level was described on p35 as “a” key antiviral efficacy endpoint measured at 52 and 104 weeks. Moreover the term “key” has been used to describe multiple secondary endpoints and results in tables 3, 4 and 7.

On reflection, the use of the word “key” in describing results and endpoints is the context of this submission is, perhaps, inappropriate. Each one of the secondary endpoints is important in reflecting a different aspect of Hepatitis B disease activity or clinical status of the patient.

Thus, the histological response endpoint is an important and objective measure reflecting the changes in liver pathology and hence a direct measure of disease progression. However its applicability and value are diminished because of the invasive nature of the biopsy procedure which precludes frequent or repeated assessments. Therefore histological response was only assessed at week 52 of the study and compared against a baseline biopsy which could be taken up to 12 months prior to study entry. Furthermore, there was a significant likelihood that

patients would decline to undergo the week 52 biopsy rendering a number of patients non-evaluable for this endpoint.

In contrast, the HBV DNA levels represent viral load and activity of virus replication, measures which directly reflect the mode of action of the drugs under scrutiny and also correlate with risk of disease progression and occurrence of hepatocellular carcinoma. Moreover, HBV DNA levels are easy to determine in blood samples drawn at every visit, and thus allow for continuous monitoring of viral activity over the entire 104 week period of the study in virtually every participant.

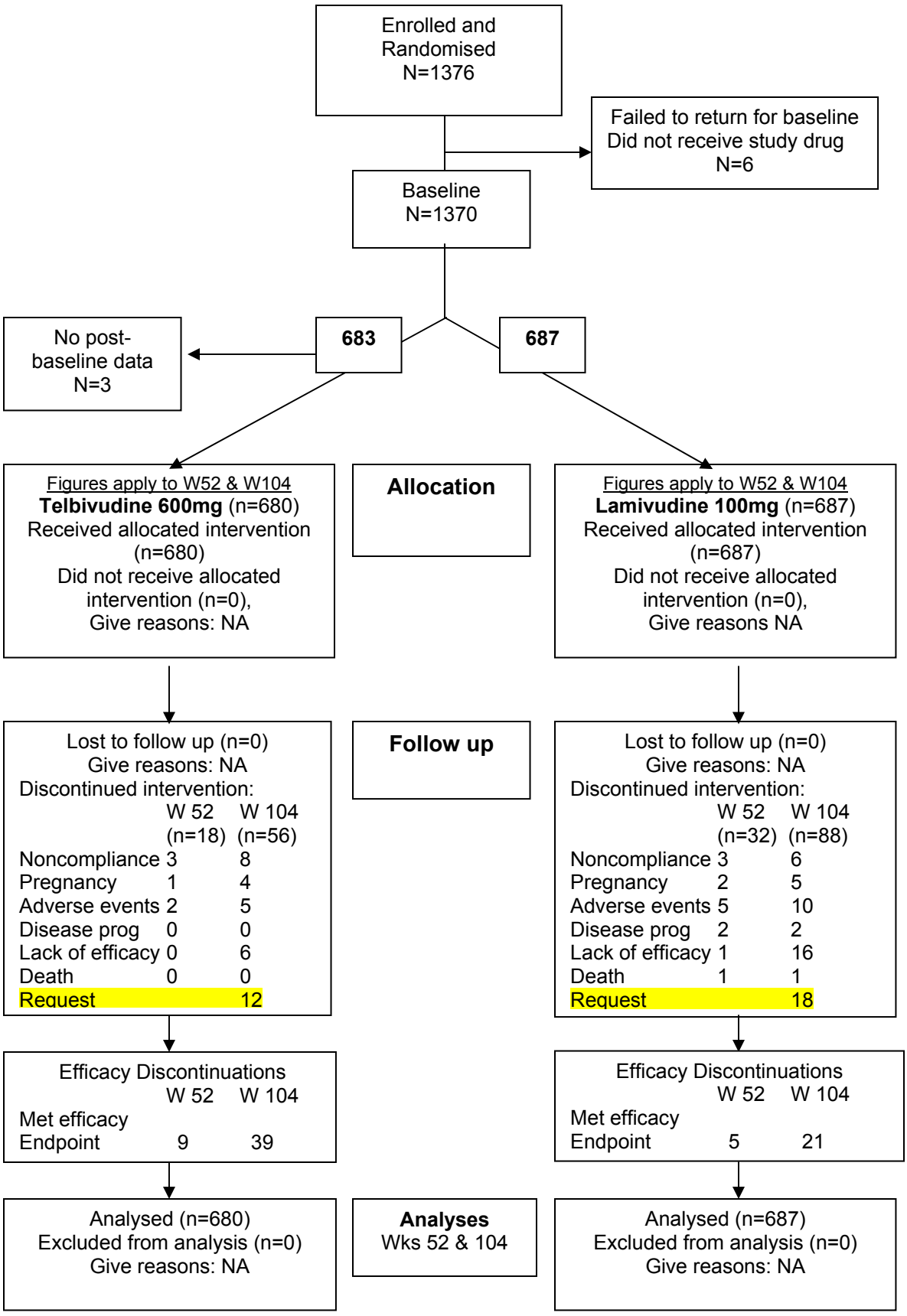
In summary, both histology and viral endpoints provide key information but from different perspectives, at different time-points, and in different patient populations in the study. Histology reflects the patient pathology while viral load reflects viral activity and impact of the drugs on viral replication. It is probably not helpful to assign one of these endpoints as more important than the other.

B2. In Fig. 4 (follow-up box), p.33, the numbers given for each of the 6 reasons for patients discontinuing treatment do not match the total numbers (n=18 wk 52, n=56 wk 104 telbivudine and likewise for lamivudine). Also on p.67, it is stated that discontinuations for adverse events, clinical disease progression or lack of efficacy were 0.6% for telbivudine and 2.0% for lamivudine. However, according to the numbers given in Fig 4, these values would be 1.6% (11/680) and 4.0% (28/687). Please clarify this discrepancy?

Figure 4 is reproduced below. Regrettably, due to a transcription error, the details of patients who discontinued medication at their own request was omitted from the original submission. The missing data are highlighted below:

The proportions of patients discontinuing for adverse events, clinical disease progression or lack of efficacy were incorrectly stated in the original submission. As pointed out in the question above, the values should indeed be 1.6% (11/680) for telbivudine and 4.0% (28/687) for lamivudine. The values of 0.6% and 2% actually represent the discontinuation rates at 52 weeks derived from the 1 year CSR and used in drafting the submission.

Figure 4: Consort Flow Chart for GLOBE Study



B3. In Fig 3, p.32, are the values presented for age mean or median? Please provide the baseline ALT levels? Also, can you provide *p* values for comparison of baseline data?

In Fig 3 the age data were presented as mean values (with range).

Baseline ALT values and the p values for the comparison are presented below for eAg positive and eAg negative populations.

Table 11-13 Baseline liver function parameters – HBeAg-positive ITT population

Laboratory parameter	Lamivudine (N=463)	Telbivudine (N=458)	Total (N=921)	*P*-value*
ALT, IU/L				0.1269
Mean (SE)	158.9 (6.30)	146.2 (5.36)	152.6 (4.14)	
Median	111.0	110.5	111.0	
25%, 75%	74.0, 191.0	74.0, 181.0	74.0, 186.0	
Min, Max	25, 1133	19, 1137	19, 1137	
ALT multiples of ULN, n (%)				0.5993
<1 x	17 (4)	14 (3)	31 (3)	
1 x - <2 x	153 (33)	149 (33)	302 (33)	
2 x - <5 x	204 (44)	219 (48)	423 (46)	
≥5 x	89 (19)	76 (17)	165 (18)	
Mean (SE) multiple of ULN	3.52 (0.140)	3.27 (0.117)	3.39 (0.092)	0.1694

Table 11-14 Baseline liver function parameters – HBeAg-negative ITT population

Laboratory parameter	Lamivudine (N=224)	Telbivudine (N=222)	Total (N=446)	*P*-value*
ALT (IU/L)				0.5477
Mean (SE)	143.7 (8.74)	137.0 (6.94)	140.4 (5.58)	
Median	98.5	99.0	99.0	
25%, 75%	67.5, 175.0	68.0, 169.0	68.0, 174.0	
Min, Max	12, 982	31, 569	12, 982	
ALT multiples of ULN, n (%)				0.6157
<1 x	15 (7)	18 (8)	33 (7)	
1 x - <2 x	84 (38)	74 (33)	158 (35)	
2 x - <5 x	88 (39)	98 (44)	186 (42)	
≥5 x	37 (17)	32 (14)	69 (15)	
Mean (SE) multiple of ULN	3.14 (0.185)	2.98 (0.144)	3.06 (0.117)	0.5010

B4. On p.36, it is stated that the proportion of patients experiencing virologic breakthrough were evaluated according to pre-specified protocol definitions and post hoc definitions – please explain what the post-hoc definitions were?

Definitions of virologic breakthrough are given on p.36: the pre-specified protocol definition is numbered 1 and the post-hoc definition numbered 2.

1. Protocol defined Virologic Breakthrough; An increase in HBV DNA to ≥ 5 $\log_{10}$ copies/ml on 2 consecutive occasions in patients who had previously achieved post baseline

virologic response (i.e. 2 values < 5 log_{10} copies/ml)

2. "1 log above nadir" Virologic Breakthrough: defined as a confirmed HBV DNA increase of ≥1 log_{10} copies/ml above nadir HBV DNA (the lowest post baseline HBV DNA level achieved) in those patients with a confirmed treatment response (i.e. ≥1 log reduction in HBV DNA).

B5. In section 5.3.5, p.38, please explain what is meant by censoring data for patients who received prohibited medication?

Censoring is a statistical technique used in time-to-event / survival analyses (eg Kaplan Meier plots) and is commonly applied to efficacy evaluations of ITT populations, but not generally used for analyses of safety. In this instance the purpose of censoring data is to avoid attributing a therapeutic effect to one of the study medications where, in fact, the patient is taking other (prohibited) medications which are thought, or known, to have anti-HBV effects.

In the case of a patient censored for prohibited medication commenced at timepoint X, all efficacy endpoints occurring up until timepoint X would be included in the analysis of group results in the normal way. After timepoint X, efficacy endpoints which occur are disregarded and the patient is no longer considered to be evaluable for efficacy endpoints.

As explained in the clinical study report for the GLOBE study the efficacy status of the patient at the time of censoring is fixed and carried forward (section 9.7.1 p 75).

"Missing data were to be treated in one of the following ways in the efficacy analyses. For patients who met the protocol criteria for treatment discontinuation due to efficacy, any missing data after the date of treatment discontinuation were to be considered as for treatment responders in the intent-to-treat (ITT) and efficacy-evaluable (EE) analyses.
With the exception of the above criteria for treatment discontinuation, protocol-defined Virologic Breakthrough, HBV resistance, and treatment failure, missing categorical data were imputed by LOCF in the EE analyses and treated as no response (i.e., missing = failure) in the ITT analyses. For the efficacy parameters of protocol-defined Virologic Breakthrough, HBV resistance and treatment failure, the LOCF method was used for missing data in both the ITT and EE analyses. For missing continuous variables, the LOCF method was to be applied in the EE analyses and excluded in the ITT analyses. When patients met the protocol criteria for treatment discontinuation for disease progression or lack of efficacy or for protocol-defined Virologic Breakthrough, any missing data thereafter were to be considered as treatment failure in both ITT and EE analyses. In the off-treatment summaries of patients who discontinued treatment due to efficacy, missing data was not imputed."

Prohibited medications accounted for data censoring in only 21 patients (12 Lamivudine; 9 Telbivudine) in the entire study. Most (18/21) occurred in year 2 of the study. A breakdown is detailed in the table below.

Table 10-5 Patients whose efficacy data were censored after on-treatment initiation of protocol-prohibited medication – overall Safety population

Assigned treatment	Site-patient ID	HBeAg status*	Time of prohibited medication initiation Day (Week)	Prohibited antiviral medication
Lamivudine	030-006	Negative	396 (56)	Adefovir dipivoxil
	030-009	Positive	262 (37)	Adefovir dipivoxil
	035-001	Positive	679 (97)	Adefovir dipivoxil
	035-019	Positive	421 (60)	Adefovir dipivoxil
	079-011	Positive	434 (62)	Adefovir dipivoxil
	080-002	Positive	489 (69)	Adefovir dipivoxil
	080-014	Negative	365 (52)	Adefovir dipivoxil
	084-002	Negative	540 (77)	Adefovir dipivoxil
	095-001	Positive	378 (54)	Lamivudine
	105-008	Positive	542 (77)	Adefovir dipivoxil
	106-022	Positive	727 (103)	Adefovir dipivoxil
	130-011	Negative	673 (96)	Adefovir dipivoxil
Telbivudine	008-042	Negative	168 (24)	Methylprednisolone
	016-012	Positive	536 (76)	Entecavir
	027-001	Positive	201 (28)	Lamivudine/zidovudine
	035-007	Negative	589 (84)	Adefovir dipivoxil
	035-015	Negative	526 (75)	Adefovir dipivoxil
	035-023	Positive	600 (85)	Adefovir dipivoxil
	057-001	Negative	493 (70)	Dexamethasone & prednisone
	106-024	Positive	653 (93)	Adefovir dipivoxil
	106-030	Positive	702 (100)	Adefovir dipivoxil

*All patients HBeAg status matched between IVRS and laboratory data.

B6. Section 5.4, p.45, gives values for sustained HBeAg loss and sustained HBeAg seroconversion. Please define these terms, along with any statistical comparison (e.g. *p* values) to support these values?

Sustained endpoint response: (for HBV DNA suppression, ALT normalization, HBeAg loss or seroconversion, or Virologic Response in patients who discontinue treatment due to efficacy)

Definition: A response documented on at least 2 consecutive post-treatment visits and at the last post-treatment study evaluation with no 2 intervening, consecutive disqualifying values.

Thus “HBeAg loss” refers to loss of detectable HBeAg where HBeAg was detected at baseline.

And “HBeAg seroconversion” reflects loss of detectable HBeAg (where present at baseline) together with gain or appearance of detectable antibodies to HBeAg (HBeAb) .

There are no p values to support this descriptive comparison. Durability of eAg loss and eAg seroconversion was assessed only in those patients who, at the investigators discretion, discontinued medication having achieved the efficacy endpoint. Statistical analysis was not

possible due to the small total numbers of patients. In addition, durability of response was not a secondary endpoint.

B7. Please define the mITT population as presented in section 5.3.6, tables 3, 4 and 7, p.41, 42 and 47?

Histologic Response populations
The Histologic Response populations were to include all patients in the ITT population who had evaluable pre-treatment liver biopsies. Patients from the ITT population meeting this criterion comprised the modified ITT (mITT) population for assessment of Histologic Response.

The definition of a 'modified ITT' (mITT) population for the histologic analyses, based on ITT patients with evaluable baseline liver biopsies, is similar to the methodology used in the adefovir and entecavir Phase III trial programs.

Section C – clarification on the Health Economic Data

1. The following tables provide the deterministic answers from the Viral Load model, having made changes in response to these comments that are discussed later. All results are presented per person. The deterministic analyses are contained in rows 8 and 9, when the 0.5 prior and when the 0.0 prior is used, respectively, in either of the final models.

Positive Disease (Prior = 0.0)

	Lifetime Costs (£)	Lifetime QALYs	Incremental Cost (£)	Incremental QALY	Incremental Cost per QALY (£)
Telbivudine	56,669	16.43	22,456	1.83	12,278
Lamivudine	34,214	14.60			

Positive Disease (Prior = 0.5)

	Lifetime Costs (£)	Lifetime QALYs	Incremental Cost (£)	Incremental QALY	Incremental Cost per QALY (£)
Telbivudine	32,333	20.01	11,961	1.38	8,669
Lamivudine	20,372	18.63			

Negative Disease (Prior = 0.0)

	Lifetime Costs (£)	Lifetime QALYs	Incremental Cost (£)	Incremental QALY	Incremental Cost per QALY (£)
Telbivudine	77,429	15.06	41,012	2.01	20,383
Lamivudine	36,417	13.05			

Negative Disease (Prior = 0.5)

	Lifetime Costs (£)	Lifetime QALYs	Incremental Cost (£)	Incremental QALY	Incremental Cost per QALY (£)
Telbivudine	43,823	18.82	26,683	0.46	57,419
Lamivudine	17,141	18.35			

2. We did not use the PSA facilities that came with the 3rd party model in order that we fully understood the derivation of the distribution for each sensitivity analyses. Instead, we sampled the PSA configurations outside of the model and loaded them one configuration at a time and 'deterministically' calculated results to provide the full PSA analyses.

3. Deterministic results were not reported as these can be misleading if not carefully interpreted, and the NICE reference case clearly states that probabilistic sensitivity analyses (PSA) are the preferred methodology. Deterministic analyses incorporate neither non-linearities in the model nor interactions between parameters. Deterministic results using the most-likely value for each parameter have now been provided (see above) but the assessment group are urged to use caution when interpreting these values. Deterministic sensitivity analyses were not provided, as these are inferior to PSA.

4. The assessment group's interpretation is correct. Calculating the random numbers in advance is also a generally beneficial modelling technique as it allows a reduction both in the size of the model and in the computational time required. The stored 'random numbers' meant that the same results would always be produced and would not, for an individual run, be subject to the random numbers sampled, which could be beneficial were debugging required. Where appropriate, Excel's distributions and random number generator were used for the sampling (and then re-pasted as values to increase speed). For dirichelet distributions, an add-in to Excel written by the Centre for Health Economics and Bayesian Statistics was used to ensure that the probabilities were correctly proportioned. As previously stated, this approach was used purely to increase the running speed of the model. If a formulae had been left within the model then, at each iteration, the random numbers would be recalculated introducing significant delay. As the model takes greater than 1 week of computational time to run, any methodology which would increase the speed of the model, without bias, was incorporated.

5. The best summary of the Jackknife methodology is in Law AM, "Simulation Modelling and Analysis" (4th Edition, McGraw-Hill (2007)). The technique (which removes bias associated with ratios) allows an estimation in the uncertainty of the incremental cost-effectiveness ratios produced by the simulated data set and indicates whether the number of PSA configurations was sufficient. As expected, the more data points that have been simulated the more robust the central estimate of cost-effectiveness. The jackknife methodology is different to that of calculating a 95% confidence interval for the mean from a percentile methodology; the former describes the uncertainty in the average incremental cost-effectiveness ratio taken from all the runs, whilst the latter reports the uncertainty associated in individual incremental cost-effectiveness ratios due to the sampled parameters. The report provides confidence intervals from both a jackknife and a percentile perspective.

6. We thank the assessment group for this comment as there was, indeed, an error in the model. Due to limited data points some of the time; viral load level and resistance status combinations had no observations and initially (in the zero prior model) the transition probabilities were left blank. Additionally the transition probabilities in the last observed time cycle were assumed to continue indefinitely. This caused a problem as, in the last time period of the GLOBE trial, some patients entered a previously 'unused' state and since the transition probabilities were blank, these patients, and other patients who entered that state in forthcoming cycles, would be lost to the model. This has been corrected within the model by setting the probability of remaining in the same viral load level as 100% in combinations where there has been no observed data. It is believed that this adjustment (made for both Lamivudine and Telbivudine) is unbiased. The check that all patients that begin the model are accounted for is provided in Row 1252 in the Markov_Tel and Markov_Lam sheets where all cells within this row should equal 1.000.

This is achieved. Tables where errors are flagged still exist. However, these 'errors' that have been investigated by the 3rd party that constructed the model can be safely ignored. Reasons for these errors being flagged include: inappropriate table – for example patients who begin in the e-negative state cannot lose their 'e' antigen; that patients do not begin the model in some states such as decompensated cirrhosis; and that Excel works to a finite precision level that introduces very small rounding errors once a number of low values are multiplied.

Having corrected the model as detailed in part 6, the analyses were re-undertaken.

Table 26 and 27 from the initial STA submission have been re-calculated and are provided below.

Whilst the individual numbers have changed, the key messages have remained. Namely that, for patients with 'e' positive disease, using telbivudine as a first line treatment has a cost per QALY of less than £20,000 when compared with lamivudine regardless of the prior used. For patients with 'e' negative disease, the cost per QALY is larger and lies between and £20,000 and £30,000, with the larger figures associated with using a prior of 0.5. It is unlikely that the cost per QALY will lie above £30,000 even when the 0.5 prior is included.

Table 26: Results from the viral load model after the application of an uninformative prior probability distribution of 0.0. The ICER reported is that of telbivudine followed with BSC where appropriate compared with lamivudine followed by BSC where appropriate. Results presented per individual patient.

	Mean incremental costs from PSA analyses	Mean incremental QALYs from PSA analyses	Mean ICER. (95% CI)	Jackknifed ICER (95% CI of integrated values)
HBeAg-positive patients	£23,983	1.56	£15,377 (£6,643 – £432,748)	£15,376 (£15,114 - £15, 638)
HBeAg-negative patients	£41,910	2.07	£20,256 (£15,237 – £66,459)	£20,255 (£20,084 - £20,427)

Table 27: Results from the viral load model after the application of an uninformative prior probability distribution of 0.5. The ICER reported is that of telbivudine followed with BSC where appropriate compared with lamivudine followed by BSC where appropriate. Results presented per individual patient.

	Mean incremental costs from PSA analyses	Mean incremental QALYs from PSA analyses	Mean ICER. (95% CI)	Jackknifed ICER (95% CI of integrated values)
HBeAg-positive patients	£12,479	1.46	£8,542 (£291 – Dominated)	£8,533 (£7,910 - £9,156)
HBeAg-negative patients	£26,883	0.97	£27,801 (£2,000 – Dominated)	£27,751 (£25,304 - £30,198)

.

7.2 Appendix 2 Viral Load Schematics

Figure 10 (app - 2) Viral load model for HBeAg-positive cohort

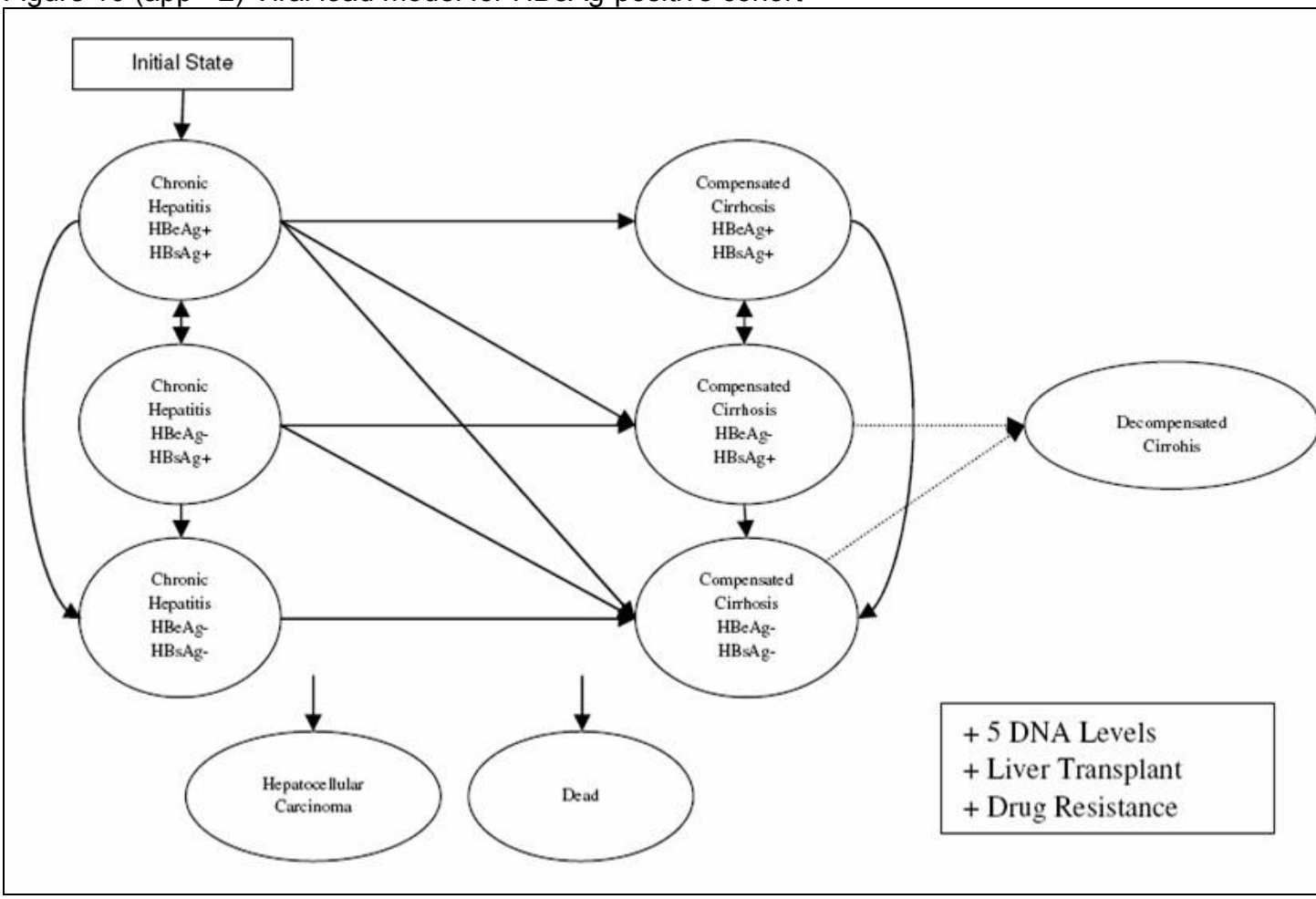
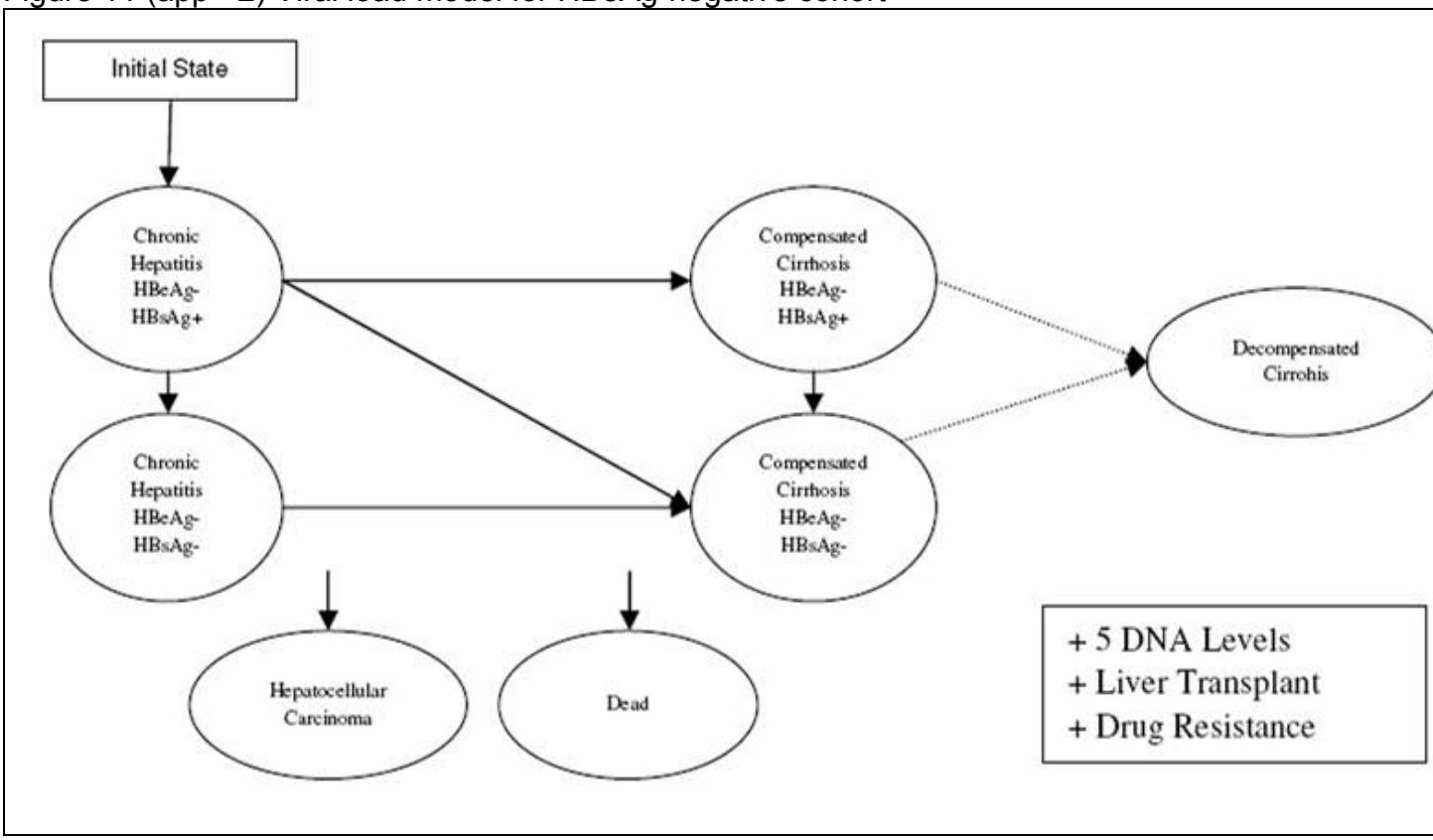


Figure 11 (app - 2) Viral load model for HBeAg-negative cohort



7.3 Appendix 3 Cost effectiveness plots and CEACs from resubmitted viral load model. Telbivudine compared with lamivudine

Figure 12 (app – 3) Incremental costs and QALYs for telbivudine relative to lamivudine - HBeAg-positive patients (prior = 0.0)

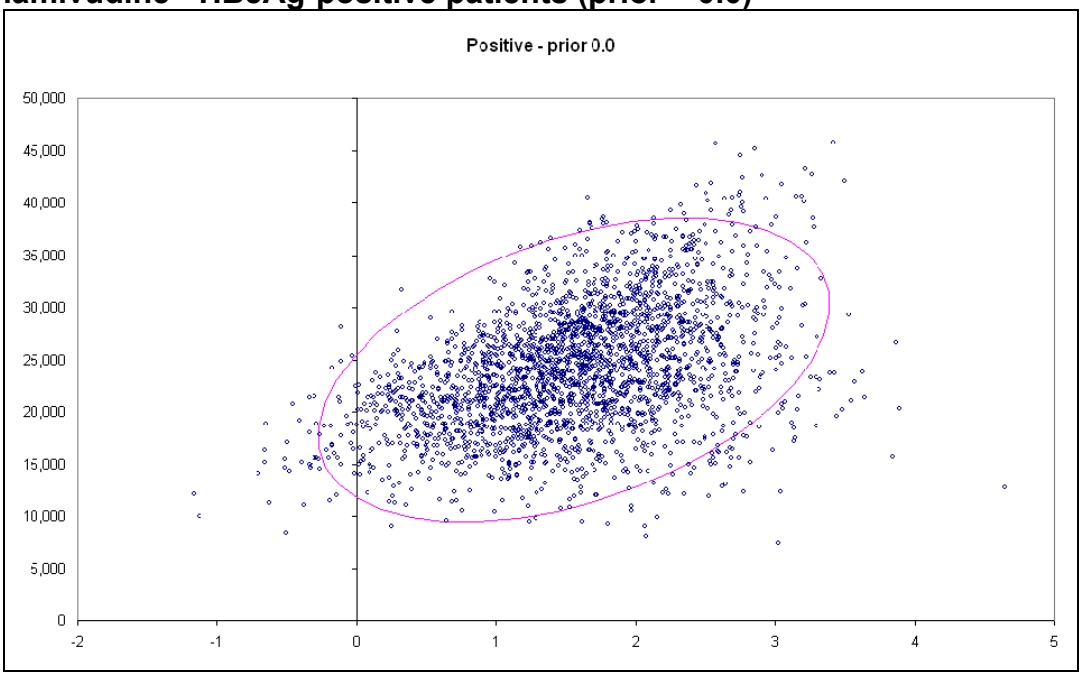


Figure 13 (app – 3) Incremental costs and QALYs for telbivudine relative to lamivudine - HBeAg-positive patients (prior = 0.5)

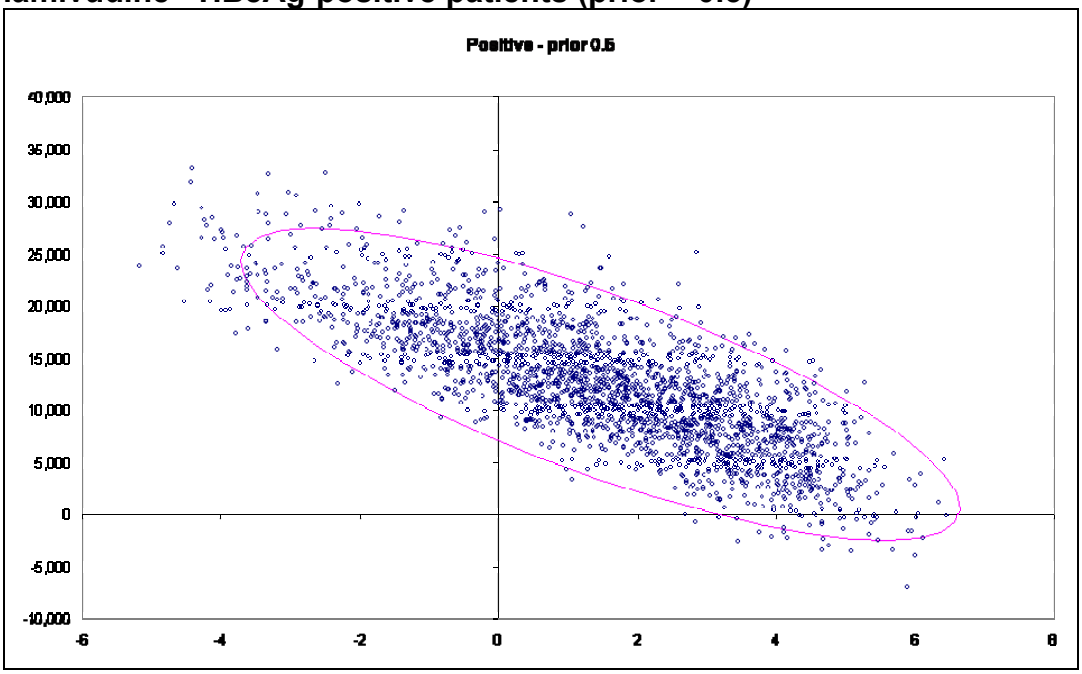


Figure 14 (app – 3) Incremental costs and QALYs for telbivudine relative to lamivudine - HBeAg-negative patients (prior = 0.0)

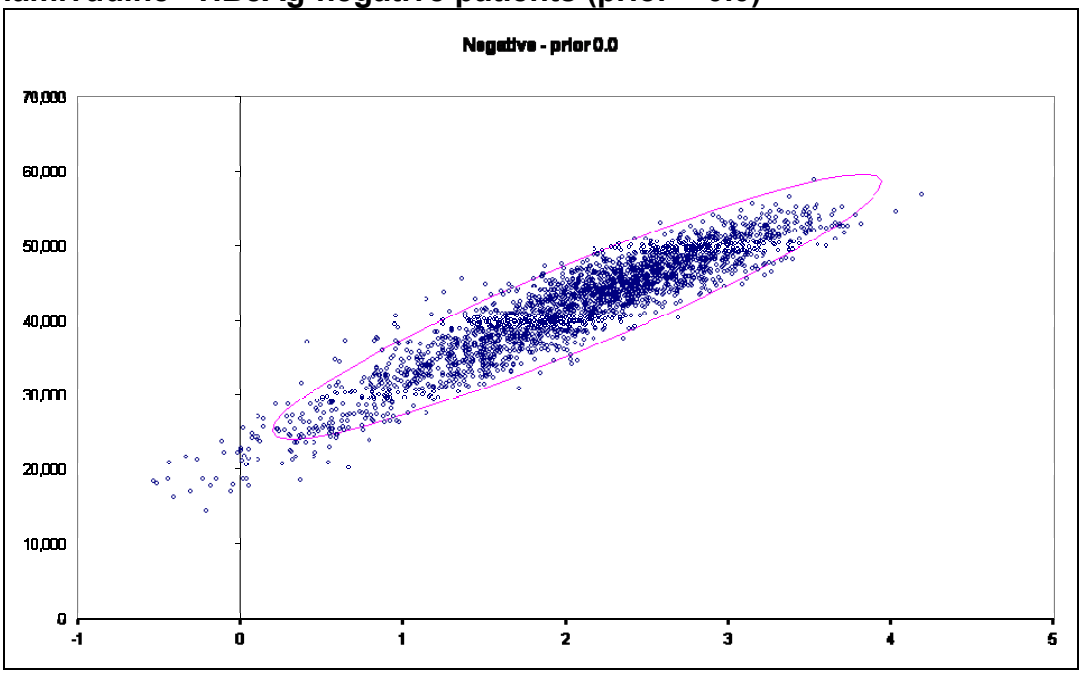
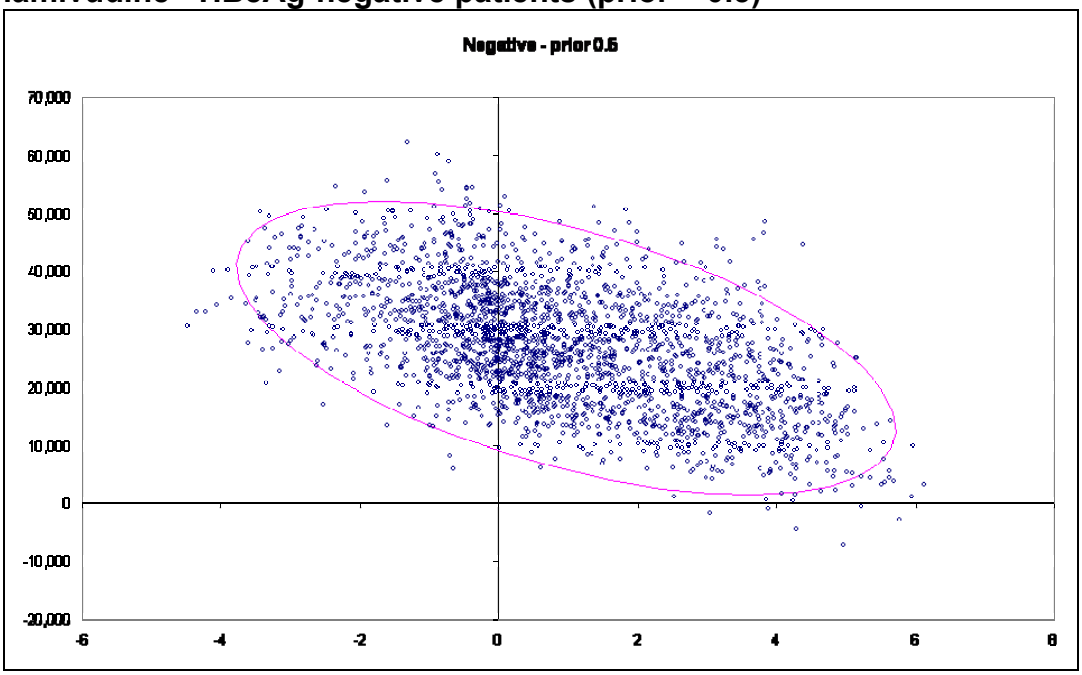


Figure 15 (app – 3) Incremental costs and QALYs for telbivudine relative to lamivudine - HBeAg-negative patients (prior = 0.5)



7.4 Appendix 4 Therapeutic indications for lamivudine (Zeffix), adefovir (Hepsera) and telbivudine (Sebivo).

Therapeutic Indication (Zeffix)
Zeffix is indicated for the treatment of chronic hepatitis B in adults with: - Compensated liver disease with evidence of active viral replication, persistently elevated serum alanine aminotransferase (ALT) levels and histological evidence of active liver inflammation and / or fibrosis. - Decompensated liver disease.
Source: http://www.emea.europa.eu/humandocs/Humans/EPAR/zeffix/zeffix.htm

Therapeutic Indication (Hepsera)
Treatment of chronic hepatitis B in adults with: - compensated liver disease with evidence of active viral replication, persistently elevated serum alanine aminotransferase (ALT) levels and histological evidence of active liver inflammation and fibrosis - decompensated liver disease
Source: http://www.emea.europa.eu/humandocs/Humans/EPAR/hepsera/hepsera.htm

Therapeutic Indication (Sebivo)
Treatment of chronic hepatitis B in adult patients with compensated liver disease and evidence of viral replication, persistently elevated serum alanine aminotransferase (ALT) levels and histological evidence of active inflammation and/or fibrosis. See section 5.1 *of the Summary of Product Characteristics* for details of the study and specific patient characteristics on which this indication is based.
Source: http://www.emea.europa.eu/humandocs/Humans/EPAR/sebivo/sebivo.htm