Mealing, Stuart; Ghement, Isabella; Hawkins, Neil; Scott, David A; Lescrauwaet, Benedicte; Watt, Maureen; Thursz, Mark; Lampertico, Pietro; Mantovani, Lorenzo; Morais, Edith; +2 more... Bregman, Bruno; Cucherat, Michel; (2014) The importance of baseline viral load when assessing relative efficacy in treatment-naïve HBeAg-positive chronic hepatitis B: a systematic review and network meta-analysis. Systematic reviews, 3 (1). 21-. ISSN 2046-4053 DOI: https://doi.org/10.1186/2046-4053-3-21

Downloaded from: http://researchonline.lshtm.ac.uk/id/eprint/2534938/

DOI: https://doi.org/10.1186/2046-4053-3-21

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
The importance of baseline viral load when assessing relative efficacy in treatment-naïve HBeAg-positive chronic hepatitis B: a systematic review and network meta-analysis

Stuart Mealing1*, Isabella Ghement2, Neil Hawkins1, David A Scott1, Benedicte Lescrauwaet3, Maureen Watt1, Mark Thursz4, Pietro Lampertico5, Lorenzo Mantovani6, Edith Morais7, Bruno Bregman7 and Michel Cucherat8

Abstract

Background: To date no network meta-analysis (NMA) has accounted for baseline variations in viral load when assessing the relative efficacy of interventions for chronic hepatitis B (CHB). We undertook baseline-adjusted and unadjusted analyses using the same data to explore the impact of baseline viral load (BVL) on CHB treatment response.

Methods: We searched Embase, Medline, Medline in Process and the Cochrane CENTRAL databases for randomised clinical trials (RCTs) of monotherapy interventions at licensed doses for use in CHB. Search strategies comprised CHB disease and drug terms (a combination of controlled vocabulary and free text terms) and also a bespoke RCT filter. The NMA was undertaken in WinBUGs using fixed and random effects methods, using data obtained from a systematic review. Individual patient data (IPD) from an entecavir clinical trial were used to quantify the impact of different baseline characteristics (in particular undetectable viral load (UVL) at 1 year) on relative treatment effect. Study level mean baseline values from all identified studies were used. Results were generated for UVL and presented as relative risks (RRs) and 95% credible intervals (CrIs) using entecavir as reference treatment.

Results: Overall, for all eight relevant interventions we identified 3,000 abstracts. Following full text review a total of 35 (including the contents of six clinical study reports) met the inclusion criteria; 19 were in hepatitis B e antigen (HBeAg)-positive patients and 14 of the 19 contained outcome information of relevance to the NMA. Entecavir and tenofovir studies had heterogeneous patient populations in terms of BVL (mean values 9.29 and 8.65 log10 copies/ml respectively). After adjusting UVL for BVL using an informative prior based on the IPD analysis, the difference between entecavir and tenofovir was not statistically significant (RR 1.27, 95% CrI 0.96 to 1.47 - fixed effects). A similar conclusion was found in all sensitivity analyses. Adjusted tenofovir results were more consistent with observed clinical trial response rates.

Conclusions: This study demonstrates the importance of adjusting for BVL when assessing the relative efficacy of CHB interventions in achieving UVL. This has implications for both clinical and economic decision making.

Keywords: Network meta-analysis, Relative efficacy, Systematic review, Virologic response, Entecavir
Current European clinical guidelines recommend the following treatment options for CHB: entecavir, lamivudine, telbivudine, adefovir dipivoxil, tenofovir dipivoxil fumarate, peginterferon-alfa-2a, interferon-alfa-2a, and interferon-alfa-2b [2]. Information on their relative efficacy is important in order for healthcare professionals and payers to make evidence-based decisions on which treatments to prescribe. Because head-to-head comparisons of competing CHB treatment options via randomised clinical trials (RCTs) are not available for all comparators in HBeAg antigen-positive or -negative CHB, indirect evidence in the form of network meta-analyses (NMAs) has been used to estimate relative efficacy. NMAs extend conventional, pairwise meta-analysis, and are based on the principle that within trial estimates of relative treatment effects can be added and subtracted [5,6].

An important assumption with NMAs is that the studies used are sufficiently similar in terms of relative treatment effect modifiers [7] - that is, study-level factors that may influence the size of the treatment effect seen with a particular pair-wise intervention. These include patient characteristics, outcomes measured, and study design. Thus, to ensure a fair comparison of interventions, it is essential to control for differences between studies in terms of potential relative treatment effect modifiers. In particular, baseline differences in patient characteristics between different trials may distort between-trial comparisons if appropriate adjustments are not made.

In CHB, response to treatment varies according to the outcome of interest, the agent used, and the patient’s HBeAg status [2]. Patient/disease characteristics that have been shown to predict response to treatment in at least some categories include baseline viral load, serum ALT level, HBV genotype, and activity score on liver biopsy [2,4]. Ali and colleagues [8] analysed data from 1,353 patients in two RCTs of entecavir and found that higher baseline viral load was associated with reduced odds of response to treatment: when baseline viral load (by PCR) was treated as a continuous variable, the odds of achieving a response were reduced by a factor of 0.38 (62%) for every one unit increase in log_{10} PCR above a threshold of 400 copies/ml.

Given the absence of head-to-head RCTs for all interventions, the objective of the current study was to generate estimates of relative efficacy of achieving undetectable viral load (ULV) that take into account the potential of baseline viral load to act as a treatment effect modifier, in order to provide like-for-like comparisons between treatments for CHB that take into account the heterogeneity in baseline viral load across patient populations in different trials.

In order to compare the results with previously published NMAs, as well as demonstrate the implications for clinical and reimbursement decisions of using such estimates, we also generated unadjusted relative efficacy estimates using similar methodologies to those used in these previous analyses.

**Methods**

We carried out adjusted and unadjusted analyses, using the same trial data for each, to explore the impact of baseline viral load on treatment response at 1 year. The interventions analysed at licensed doses were interferon alfa, peginterferon alfa-2a/2b, lamivudine, adefovir dipivoxil, entecavir, tenofovir, telbivudine and also placebo. Trials for inclusion in the NMAs were identified through a systematic review of the literature.

The efficacy endpoints analysed in the unadjusted analysis were ALT normalization, histological improvement, HBeAg seroconversion and achievement of UVL at 1 year. Since Ali and colleagues [8] only generated results for one endpoint (achievement of UVL at 1 year) the adjusted analysis was necessarily restricted to this endpoint/time-point. In all analyses, UVL was defined as reduction in HBV DNA level (by PCR assay) below the trial specific lower level of quantification (LLOQ).

**Systematic review**

We carried out a systematic review of RCTs of the interventions listed above. The inclusion criteria were RCTs (phase II or III) of monotherapy interventions at licensed dose, adults with CHB, reporting any of the endpoints of interest, and published in English. Papers (full or otherwise) reporting interim results and studies using the interventions of interest at non-licensed doses were excluded.

Searches were carried out in the Embase, Medline, Medline in Process and Cochrane CENTRAL databases between March and April 2011. No restriction was placed on the earliest date of publication and all databases were searched from date of inception. Search strategies comprised CHB disease and drug terms (a combination of controlled vocabulary and free text terms), and also a bespoke RCT filter. A search was also made for abstracts from the European Association for the Study of the Liver and the American Association for the Study of Liver Diseases 2010 and 2011 annual conferences. Search syntax for all databases are available on request from the authors. The search strategy used to search the Embase database is presented in Additional file 1.

The studies were separated into four clinically distinct patient groups: treatment-naive HBeAg-positive or -negative, lamivudine refractory and ‘other’. Abstract screening was performed by two authors and included in the full paper review if one reviewer thought it relevant. Formal full paper review was undertaken by two reviewers against the prespecified inclusion criteria with a third acting as mediator in situations of disagreement. Three authors independently extracted study characteristics and the outcome data
required for the NMA using a standard form. Discrepancies were resolved by one of two other authors. Outcome data from weeks 48 and 52 were assumed to refer to 1 year. A risk of bias assessment was carried out using Cochrane methodology for those RCTs reported as full papers [9]. No formal protocol was created for this review.

Statistical analyses
Given the lack of head-to-head trial evidence estimating the relative efficacy of all licensed interventions, we used an NMA approach to synthesise the evidence. In the NMA methodology the difference in effect between treatments A and B is equal to the difference in effects between treatments A and C, and B and C. The analysis can be expanded to more complex networks of evidence, and can produce estimates of both mean effect and uncertainty [10].

Fixed effect models were used in the unadjusted analysis. For the adjusted analyses we used both fixed and random effects models, and final model choice in all analyses was based on deviance information criteria (DIC) [11].

Choice of prior distribution for parameters in NMA models is an important consideration, especially in the presence of sparse networks of evidence. Uninformative priors were used in the unadjusted analyses for all model parameters. In all covariate-adjusted analyses the results from Ali and colleagues [8] were used to inform the prior distribution on the regression coefficient associated with baseline viral load. Variations in baseline viral load in each study arm (where there were differences) were incorporated into the adjusted analyses via the use of the average baseline HBV DNA value across arms within a given RCT.

Baseline viral load was assumed to modify treatment effects relative to entecavir (0.5 mg), which was also used as the baseline against which all relative efficacy estimates were calculated. To make the results easier to interpret by a non-statistical audience, we represented relative efficacy as a relative risk (RR) of response instead of the more natural odds ratio. We reported the mean of the posterior probability distribution as well as the 95% credible interval (CrI) for each RR. When the 95% CrI did not include the value one, the RR was considered significantly different to that for entecavir.

In order to compare the results of the analyses with the input data, as well as presenting the output in an intuitive manner, we also generated the absolute predicted probabilities of response for each clinical outcome and treatment combination. In the adjusted NMA we also undertook a range of sensitivity analyses whereby in addition to the use of fixed and random effects models, the impact of adding or removing individual studies due to heterogeneity was assessed. Caterpillar, density and Brooks-Gelman-Rubin plots were examined in all analyses to ensure model convergence.

The analyses were conducted in WinBUGS Version 1.4 (MRC Biostatistics Unit, Cambridge, UK) [12] using Bayesian Markov Chain Monte-Carlo Gibbs sampling methods.

Results
Search results and summary of studies
Our search of clinical databases identified 3,000 abstracts; 179 articles, including clinical study reports (CSRs), were ordered/requested for review, of which 35 (six CSRs) met the inclusion criteria [13-48]. The contents of five of the CSRs had been reported in peer reviewed publications already captured by the search and hence the published data were used [18,26,37,42,44]. One CSR (BMS study AI463023) [13] and the Summary of Product Characteristics for telbivudine [22] were included in the review. In total, the review identified 29 unique trials. Of these, 19 contained information in HBeAg-positive patients, and 14 of the 19 reported enough information to warrant inclusion in a NMA, and 13 reported information on UVL at that timepoint [13-15,18,20-24,26,28,30,32,48].

The study selection process is presented as a PRISMA diagram in Figure 1. The PRISMA 2009 checklist is reported in Additional file 2. Study characteristics and reported UVL at 1 year (defined as either 48 or 52 weeks) are shown in Table 1. The assessment of study quality undertaken as part of the systematic review is reported in Additional file 1: Table S1. Studies identified by the systematic review used a range of LLOQ values from 1,000 to 200 copies/ml.

Unadjusted network meta-analysis
The network of evidence used to generate all results is presented in Figure 2. The results of the fixed effects analysis are presented as relative risks in Table 2 and absolute probabilities of response in Additional file 1: Table S2. There was only one instance where a treatment performed significantly better than entecavir; the RR for tenofovir achieving UVL was 1.43 (95% CrI 1.30 to 1.54). With the exception of telbivudine which demonstrated no statistically significant difference to entecavir (RR 0.88, 95% CrI 0.76 to 1.00) all other interventions performed significantly less well than entecavir.

Adjusted network meta-analysis
The primary adjusted analysis of achieving UVL at 1 year, when accounting for baseline viral load, was undertaken using materials available in the public domain (the “base case”). Thus, the material extracted from the CSR was excluded. In addition, the baseline rates for two studies were very different to the remainder in that they were assessed using a different assay with very different LLOQ definitions suggesting that baseline data were collected in a different manner to all other studies [16,24]. These studies were also excluded from the base case analysis. One study, TBVIG, reported median rather than mean
baseline viral load and was hence also excluded from the base case-adjusted analysis [30]. Information on this study is provided in Table 1. Data from the ten studies that reported baseline viral load were used in the adjusted analyses [14,15,18,20-23,26-28,31,48].

The results are presented as relative efficacy estimates in the second column of Table 3 and as absolute probabilities of UVL at 1 year in Figure 3 (fixed effects) and Figure 4 (random effects).

The relative risk estimates produced by the fixed and random effects base case analyses were very similar. In particular, entecavir produced significantly increased RRs of UVL at 1 year compared with all interventions except telbivudine and tenofovir, for which the likelihood was similar. In contrast to the unadjusted analysis, the relative efficacy of entecavir and tenofovir for achieving UVL was not significantly different. Thus, baseline viral load is a significant moderator of the effects of monotherapies for CHB. The DIC estimate for the fixed effect model was lower than that for the random effects model and was thus the preferred approach.

Adjusted network meta-analysis: sensitivity analyses

Exclusion of data from one adefovir study

The reported 1 year UVL rate for adefovir patients as reported by the 018 Study Group is approximately two to three times higher than reported for adefovir in all other studies (Table 1). In contrast, the absolute response rate for patients receiving telbivudine reported by this study was in line with that observed in the other studies. The impact of removing this study is presented in column three of Table 3, and in Figures 3 and 4.

The DIC statistics for both fixed and random effects models were similar, with the random effects analysis representing the best fitting model. While the results overall are similar to those in the base case analyses, the greatest impact is observed in the tenofovir results, with a relative risk value of 1.08 (95% CrI 0.22 to 1.52). Of note, the derived absolute response probabilities for entecavir and tenofovir in this scenario were 65.9% and 71.4%, respectively (random effects model). The corresponding values in the key regulatory trials were 66.7% and 76.5%, respectively.

Exclusion of data from the 018 Study Group, and inclusion of data from AI463023 and TBVIG

The systematic review identified two additional studies which contained information of potential interest: as yet unpublished data from BMS study AI463023 and median baseline values from the TBVIG study [13,29]. When these data were included, but the data from 018 remained
| Source                          | Treatment duration | Study design                  | Number of patients | Treatment | Age (years) | Male (%) | Endpoint timepoint | Method used to measure HBV DNA | LLOQ (log_{10} copies/ml) | Baseline viral load (log_{10} copies/ml) | Undetectable HBV DNA at 1 year (%) |
|--------------------------------|--------------------|-------------------------------|---------------------|-----------|-------------|-----------|-------------------|-------------------------------|---------------------------|--------------------------------|---------------------------------|
| 018 Study Group [13]           | 52 weeks           | Randomised, controlled, open label | 45                  | TBV 600 mg | 34          | 78        | 52 weeks          | Amplicor PCR Assay (Roche)   | 300 copies/ml             | 9.57                           | 60                              |
|                               |                    |                               | 44                  | ADV 10 mg  | 30          | 91        | -                | -                            |                          | 9.98                           | 40.9                            |
| ADV 437 Study Group [14]       | 48 weeks           | Randomised, single blind      | 167                 | Placebo   | 37          | 71        | 48 weeks          | Amplicor PCR Assay (Roche)   | 400 copies/ml             | 8.12                           | 0                               |
| AHLSG [15]                     | 52 weeks           | Randomised, double blind      | 72                  | Placebo   | 29          | 72        | 52 weeks          | Solution hybridising assay (Abbott) | 1.6 pg/ml               | 1.85                           | NR                              |
|                               |                    |                               | 143                 | LAM 100 mg| 31          | 74        | -                | -                            |                          | 1.8                            | NR                              |
| AI463023 [12]                  | 96 weeks           | Phase 3 randomised, double blind | 225                 | ETV 0.5 mg| -           | -         | 52 weeks          | PCR assay (company unspecified) | 300 copies/ml             | 8.80                           | 73.8                            |
|                               |                    |                               | 221                 | LAM 100 mg| -           | -         | -                | -                            |                          | 8.70                           | 37.6                            |
| BeHoLD_I [17]                  | 60 weeks           | Phase 3 randomised, double blind | 354                 | ETV 0.5 mg| 35          | 77        | 48 weeks          | Amplicor PCR Assay (Roche)   | 300 copies/ml             | 9.62                           | 66.7                            |
|                               |                    |                               | 355                 | LAM 100 mg| 35          | 74        | -                | -                            |                          | 9.69                           | 36.3                            |
| Globe study group [19-21]      | NR                 | Phase 3 randomised, double blind | 463                 | LAM 100 mg| 33          | 76        | 52 weeks          | Amplicor PCR Assay (Roche)   | 300 copies/ml             | 9.50                           | 40.4                            |
|                               |                    |                               | 458                 | TBV 600 mg| 32          | 73        | -                | -                            |                          | 9.50                           | 60                              |
| Hou [22]                       | 104 weeks          | Phase 3 randomised, double blind | 147                 | TBV 600 mg| 28          | 80        | 52 weeks          | Amplicor PCR Assay (Roche)   | 300 copies/ml             | 9.30                           | 66.7                            |
|                               |                    |                               | 143                 | LAM 100 mg| 29          | 75        | -                | -                            |                          | 9.70                           | 37.8                            |
| ILSG [23]                      | 52 weeks           | Randomised, partially double blind | 82                  | LAM 100 mg| 30          | 71        | 52 weeks          | Solution hybridising assay (Abbott) | 1.6 pg/ml               | 2.04                           | 60                              |
|                               |                    |                               | 69                  | IFNA      | 32          | 81        | -                | -                            |                          | 1.78                           | 29.1                            |
| Lau [47]                       | 72 weeks           | Phase 3 randomised, double blind | 271                 | PegIFNA   | 32.5        | 79        | 48 weeks          | Amplicor PCR Assay (Roche)   | 400 copies/ml             | 9.90                           | 25.1                            |
|                               |                    |                               | 272                 | LAM 100 mg| 31.6        | 79        | -                | -                            |                          | 10.10                          | 39.7                            |
| Leung [25]                     | Minimum 52 weeks   | Phase 3 randomised, open label | 33                  | ETV 0.5 mg| 37          | 61        | 48 weeks          | Amplicor PCR Assay (Roche)   | 300 copies/ml             | 10.30                          | 57.6                            |
|                               |                    |                               | 32                  | ADV 10 mg  | 32          | 66        | -                | -                            |                          | 9.88                           | 18.8                            |
| Marcellin [26]                 | 48 weeks           | Phase 3 randomised, double blind | 176                 | TDF 300 mg| 34          | 68        | 48 weeks          | Cobas Taq-Man PCR Assay (Roche) | 169 copies/ml        | 8.64                           | 76.1                            |
|                               |                    |                               | 90                  | ADV 10 mg  | 34          | 71        | -                | -                            |                          | 8.88                           | 13.3                            |
| Ren [27]                       | 48 weeks           | Randomised                    | 21                  | LAM 100 mg| 34          | 52        | 48 weeks          | PCR assay (company unspecified) | Unspecified             | 8.49                           | 38                              |
|                               |                    |                               | 21                  | ETV 0.5 mg| 31          | 57        | -                | -                            |                          | 8.52                           | 71.4                            |
Table 1 Study characteristics and 1-year outcomes of studies included in the network meta-analysis (HBeAg-positive patients only) (Continued)

| Study | Follow-up | Design | Group | Baseline | 12 Week | 52 Week | Virological Response | Viral Load (copies/ml) | N/R | 1 Year | N/R |
|-------|-----------|--------|-------|----------|---------|---------|---------------------|----------------------|-----|--------|------|
| TBVIG [29] | 52 weeks | Phase 2 randomised, double blind | 19 | LMV 100 mg | 34 | 74 | 52 weeks | Amplicor PCR Assay (Roche) | 200 copies/ml | N/R | 31.6 |
| TBVIG [29] | 22 | TBV 600 mg | 40 | 82 | N/R | 61.4 |
| USLIG [30] | 68 weeks | Prospective, randomised, double blind | 71 | Placebo | 38 | 80 | 52 weeks | Unspecified | Unspecified | 5.70 | 15.9 |
| USLIG [30] | 66 | LAM 100 mg | 40 | 86 | 10.20 | 44.4 |

The systematic review identified 21 studies reporting hepatitis B e antigen (HBeAg)-positive patients; this table only shows the 14 studies included in the network meta-analysis. The remaining 7 studies did not report outcomes at 1 year and so were not included in the network meta-analysis. *Patient numbers for overall population provided, baseline viral load and hepatitis B virus (HBV) DNA reported as HBeAg-positive/HBeAg-negative, outcomes for HBeAg-positive patients only. ADV, adefovir; ETV, entecavir; IFNA, interferon-alfa; LAM, lamivudine; LLOQ, lower level of quantification; NR, not reported; PegIFNA, pegylated interferon-alfa 2a; TBV, telbivudine; TDF, tenofovir.
excluded, the corresponding results from this analysis are presented in column four of Table 3, and in Figures 3 and 4.

The random effects analysis generated the lowest DIC and is therefore the preferred model. There was no significant difference between the relative efficacy of entecavir, telbivudine and tenofovir, but entecavir performed significantly better than all other interventions. These results are in contrast to the unadjusted results. The absolute probabilities derived using a random effects model for tenofovir, telbivudine and entecavir were similar to those observed in the landmark RCTs.

Discussion

NMA can be used to generate relative efficacy estimates of competing treatments in situations where more than two treatment options are available and direct head-to-head evidence from RCTs does not exist for all comparators. The NMA approach allows all relevant evidence to be considered and addresses research questions in the absence of direct comparative evidence, improving the precision of estimates by combining direct and indirect evidence.

Table 2 Unadjusted efficacy estimates relative to treatment with entecavir

|                       | Fixed effects model |
|-----------------------|---------------------|
| Entecavir 0.5 mg      | 1.00 (Baseline)     |
| Adefovir 10 mg        | 0.64 (0.42, 0.87)   |
| Lamivudine 100 mg     | 0.54 (0.46, 0.63)   |
| Placebo               | 0.10 (0.04, 0.19)   |
| Telbivudine 600 mg    | 0.88 (0.76, 1.00)   |
| Tenofovir 245 mg      | 1.43 (1.30, 1.54)   |
| Interferon alfa       | 0.21 (0.09, 0.38)   |
| Peginterferon alfa-2a/2b | 0.34 (0.23, 0.46) |

Results are shown as relative risk (95% credible interval). A relative risk <1 should be interpreted as a given treatment being less efficacious (that is, worse) than entecavir, and a value >1 being more efficacious (that is, better).

One of the key assumptions underpinning this method is that the studies included in the analysis are homogeneous (that is, the trials are sufficiently similar on study and patient characteristics). The similarity assumption is violated if one or more study-level covariates act as modifiers of the relative treatment effects and their distribution is not balanced across the studies being compared [49,50]. In this case, NMA may be affected by confounding bias, unless one explicitly controls for these covariates in the statistical analyses.

Controlling for covariates is particularly important in cases where response to treatment is defined in terms of post-treatment level of a measure, and when that baseline level of this measure is known to vary across studies. If one study recruits patients with worse levels of a variable that is known to modify the relative impact of treatment, then the level of response achieved is likely to be smaller compared with another study which primarily includes patients with better baseline levels, other things being equal.

The motivation for our work was the belief that such baseline covariate imbalances had occurred for patients recruited into studies looking at interventions for CHB. In particular, it was noted that there were differences in mean baseline viral load (expressed in terms of log10 copies/ml when measured using the PCR assay) with values for entecavir and tenofovir differing by approximately 1 log10 copies/ml (Table 1). We hypothesised that failure to account for these differences in previous analyses may have led to biased estimates of relative efficacy.

The work contained in this paper supports this hypothesis. When no adjustment was made to account for differences in baseline viral load among trials, tenofovir was shown to be significantly better than entecavir in terms of achieving UVL at 1 year (fixed effects RR 1.43, 95% CrI 1.30 to 1.54). However, when we accounted for the impact of baseline viral load the difference between the two treatments was not significant (fixed effects RR 1.27, 95% CrI 0.96 to 1.47; random effects RR 1.21, 95%
The fixed effects adjusted model best fitted the underlying data, although the difference was minor (fixed effects DIC, 35.56; random effects DIC, 35.86). Sensitivity analyses highlighted that the relative efficacy of tenofovir versus entecavir was contingent on the choice of studies included in the meta-analysis, and in particular whether or not data reported by one study group [13] were used. When these data were excluded, there is no significant difference between the two interventions (RR 1.08, 95% CrI 0.22 to 1.52). A subsequent sensitivity analysis, whereby this study was removed but two other studies were included (AI4463023 and TBVIG), generated similar non-significant results (RR 1.15, 95% CrI 0.39 to 1.50). In both sensitivity analyses the most appropriate model, based on DIC, consisted of random as opposed to fixed effects approaches. Close examination of the published paper [14] has identified no reason why this result should occur, and so there may be some other form of study level heterogeneity as yet unaccounted for that is influencing the results.

Our paper is the first to generate baseline viral load adjusted and unadjusted NMA results using data from the same set of studies, and the results from the unadjusted analyses are very similar to those generated by other research groups [51,52]. Accepting that NMA is based on relative efficacy, the results from all three unadjusted analyses for UVL appear to be at odds with those provided by the clinical trials included in the NMA. The systematic review identified one study of tenofovir [27] and the observed response rate was 76%. The corresponding value arising from our NMA was 93.2% (95% CrI 85.6% to 97.6%). Similar values were generated by two other research groups [51,52]. One other NMA has been recently published [53]. This analysis, however, contains a number of methodological flaws, the most notable being the pooling of data from HBeAg-positive and -negative individuals. We have therefore not extracted results from this paper for the purposes of discussion.

In contrast, with the exception of placebo and interferon-based therapies, the CrIs for the values derived in the

| Table 3 Adjusted relative risk estimates for virologic response, expressed as relative risk of achieving undetectable viral load |
|---------------------------------------------------------------|
| **Intervention** | **Base case** | **Base case without 018 study group** | **Base case without 018 study group; with AI4463023 + TBVIG** |
|------------------|---------------|-------------------------------------|-------------------------------------------------------------|
| **Fixed effects analyses** |               |                                     |                                                             |
| Entecavir 0.5 mg | 1.00 (Baseline) | 1.00 (Baseline) | 1.00 (Baseline) |
| Adefovir 10 mg   | 0.33 (0.14, 0.62) | 0.23 (0.05, 0.56) | 0.33 (0.10, 0.70) |
| Lamivudine 100 mg| 0.38 (0.22, 0.58) | 0.38 (0.22, 0.58) | 0.52 (0.41, 0.64) |
| Placebo          | 0.01 (0.00, 0.03) | 0.00 (0.00, 0.02) | 0.01 (0.00, 0.03) |
| Telbivudine 600 mg| 0.67 (0.43, 0.92) | 0.67 (0.44, 0.92) | 0.86 (0.71, 1.01) |
| Tenofovir 245 mg | 1.27 (0.96, 1.47) | 1.12 (0.61, 1.43) | 1.25 (0.84, 1.48) |
| Interferon alpha | 0.14 (0.05, 0.29) | 0.13 (0.05, 0.29) | 0.20 (0.08, 0.37) |
| Peginterferon alpha-2a/2b | 0.23 (0.11, 0.39) | 0.22 (0.11, 0.39) | 0.32 (0.21, 0.46) |
| Residual deviance | 19.48          | 17.10                               | 21.54 |
| DIC              | 35.56          | 32.14                               | 39.15 |
| **Random effects analyses** |               |                                     |                                                             |
| Entecavir 0.5 mg | 1.00 (Baseline) | 1.00 (Baseline) | 1.00 (Baseline) |
| Adefovir 10 mg   | 0.33 (0.07, 0.82) | 0.27 (0.03, 0.95) | 0.29 (0.04, 0.78) |
| Lamivudine 100 mg| 0.37 (0.09, 0.84) | 0.42 (0.07, 1.09) | 0.51 (0.23, 0.86) |
| Placebo          | 0.01 (0.00, 0.06) | 0.01 (0.00, 0.07) | 0.01 (0.00, 0.04) |
| Telbivudine 600 mg| 0.64 (0.18, 1.19) | 0.72 (0.12, 1.39) | 0.86 (0.41, 1.27) |
| Tenofovir 245 mg | 1.21 (0.48, 1.51) | 1.08 (0.22, 1.52) | 1.15 (0.39, 1.50) |
| Interferon alpha | 0.16 (0.01, 0.67) | 0.20 (0.01, 0.94) | 0.22 (0.03, 0.67) |
| Peginterferon alpha-2a/2b | 0.25 (0.03, 0.88) | 0.30 (0.02, 1.13) | 0.34 (0.07, 0.88) |
| Residual deviance | 18.23          | 16.12                               | 19.78 |
| DIC              | 35.86          | 32.10                               | 39.07 |

Results are shown as relative risk (95% credible interval). “Base case” refers to the adjusted analysis undertaken using the 10 studies listed in the document containing appropriate information. As with the unadjusted analyses, a relative risk <1 should be interpreted as a given treatment being less efficacious (that is, worse) than entecavir and a value >1 being more efficacious (that is, better). DIC, deviance information criteria.
adjusted analyses all contain the observed trial values, and the RR estimates are close to the trial values once the 018 Study Group data are removed (Figures 3 and 4). Hence, we would argue that the adjusted results are of greater clinical relevance than the unadjusted results.

Generating 'like-for-like' estimates of relative efficacy by controlling for covariates believed to be modifiers of relative treatment effects is not just of clinical interest but is essential for the purposes of reimbursement decisions. Such estimates are used by agencies such as the

---

**Figure 3** Absolute probability of undetectable viral load at 1 year (fixed effects). "Basecase" refers to the adjusted analysis undertaken using the ten studies listed in the document containing appropriate information.

**Figure 4** Absolute probability of undetectable viral load at 1 year (random effects). "Basecase" refers to the adjusted analysis undertaken using the ten studies listed in the document containing appropriate information.
National Institute for Health and Care Excellence in their appraisal processes when assessing the clinical efficacy in a given disease area [54]. In addition, such values are also used in economic models to evaluate the cost-effectiveness of interventions. A number of such models have been developed in CHB [55-59], of which one [57] used the results from their unadjusted analysis directly as model inputs. Another [59] used UVL as a surrogate variable for risk of cirrhosis using information from the REVEAL-HBV study [60], which quantified the relationship between HBV DNA and the likelihood of being diagnosed with cirrhosis. Overestimation of virologic response would thus correspond to underestimation of the likelihood of cirrhosis, which has been identified as a key driver of cost-effectiveness.

Despite the review finding a decent number of studies overall, as can be seen from Figure 2, the presence of a large number of treatment options means that the majority of the branches in the network are informed by the findings of a single study. This increases the uncertainty surrounding all results and means that baseline imbalances in other potential treatment effect modifiers may have influenced the results.

Further work is needed to complement the work contained in this paper in connection with the achievement of UVL at 1 year in order to explore the impact of other potentially clinically relevant covariates on the relative effects of comparators and the probability of achieving UVL. Exploring the impact of other areas of potential heterogeneity (for example, study design, impact of different LLOQ definitions) is also important. In addition, Ali and colleagues [8] identified the time of assessment as a treatment effect modifier in addition to baseline viral load. The studies included in this analysis were very similar in terms of assessment times and so the exclusion of this variable is likely to have had a modest effect. Nonetheless, it would be interesting to replicate the analyses contained in this paper when controlling for these slight differences. Furthermore, expanding this type of analysis to other clinically relevant endpoints is also worthwhile.

Conclusions
The analysis showed that baseline viral load is a treatment effect modifier in CHB and that failure to correct for this variable inflates the relative efficacy estimates for some interventions. Since these estimates are often used in economic models to generate cost-effectiveness estimates, failure to adjust for baseline viral load will generate erroneous ICERs, resulting in poor use of scarce healthcare resources. As such, reimbursement agencies should therefore only use covariate-adjusted relative efficacy estimates in their decision making surrounding treatments for CHB.

Additional files

Additional file 1: Supplementary material, including Embase search strategy, assessment of study quality for all included trials (Table S1) and comparison of current and previous NMA results (Table S2).

Additional file 2: PRISMA 2009 checklist.
4. Lok ASF, McMahon BJ: AASLD practice guidelines. Chronic hepatitis B: update 2009. Hepatology 2009, 50:561–662.

5. Caldwell D, Ades AE, Higgins JP: Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005, 331:897–900.

6. Lu G, Ades AE: Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004, 23:3105–3124.

7. Song F, Loke YK, Walld T, Glenny AM, Eastwood AJ, Altman DG: Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009, 338:b1147.

8. Ali S, Mealing S, Hawkins N, Lescauwat B, Bjerk S, Mantovani L, Lamperotto P: The use of individual patient data (IPD) to quantify the impact of pre-treatment predictors of response to treatment in chronic hepatitis B patients. BMJ Open 2013, 2:45.

9. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. http://www.cochrane.org/training/cochrane-handbook.

10. Haaglin DC, Hawkins N, Jansen JP, Scott DA, Inder R, Cappelleri JC, Boersma C, Thompson D, Lathoff RM, Diaz M, Barrett A: Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons. Good Research Practices: part 2. Value Health 2011, 14:429–437.

11. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A: Bayesian measures of model complexity and fit. Royal Stat Soc: Series B (Stat Methodology) 2002, 64:583–639.

12. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000, 10:337–355.

13. Bristol-Myers Squibb Ltd: Trial AI463023 Clinical Study Report. Data on file.

14. Chan HL, Heathcoe EJ, Marcellin P, Lai CL, Cho M, Moon YM, Chao YC, Myers RP, Minuk GY, Jaffers L, Sievert W, Zhai J, Qin XJ, NA, OA, BA, 98 Study Group: Treatment of hepatitis B e antigen positive chronic hepatitis with telbivudine or adefovir: a randomized trial. Ann Intern Med 2007, 147:745–754.

15. Marcellin P, Chang TT, Gish RG, de Man R, Gadano A, Sollano J, Chao YC, Lok AS, Han KH, Myer M, Cha G, Lloyd D, Brown NA, Telbivudine Phase II Investigator Group: A 1-year trial of telbivudine, lamivudine, and the combination in patients with hepatitis B e antigen-positive chronic hepatitis B. Gastroenterology 2005, 129:528–536.

16. Dienstag JL, Schiff ER, Wright TI, Perrillo RP, Han H, Renn FY, Diao DM, Piao XM: A one-year trial of entecavir treatment in patients with HBsAg-positive chronic hepatitis B. World J Gastroenterol 2007, 13:4624–4627.

17. Shindo M, Chayama K, Mochida S, Toyota J, Tomita E, Kumada H, Yokosuka O, Sata M, Hayashi N, Suzuki K, Okanoue T, Tsubouchi H, Ishikawa H, Seri T, Onnata M: Antiviral activity, dose-response relationship, and safety of entecavir following 24-week oral dosing in nucleoside-naive Japanese adult patients with chronic hepatitis B: A randomized, double-blind, phase II clinical trial. Hepatol Int 2009, 3:452–458.

18. Lai CL, Leung N, Teo EK, Tong M, Wong F, Fong HH, Han S, Poyntard T, Myers M, Chao G, Lloyd D, Brown NA, Telbivudine Phase II Investigator Group: A one-year trial of telbivudine, lamivudine, and the combination in patients with hepatitis B e antigen-positive chronic hepatitis B. Gastroenterology 2005, 129:528–536.
43. Bristol-Myers Squibb Ltd: Trial AI463014 Clinical Study Report. Data on file.
44. Chang TT, Gish RG, Hadziyannis SJ, Cianciara J, Rizzetto M, Schiff ER, Pastore G, Bacon BR, Poyntard J, Joshi S, Kleszewski KS, Thiy A, Rose RE, Colombo RJ, Hindes RG, BEHoLD Study Group: A dose-ranging study of the efficacy and tolerability of entecavir in lamivudine-refractory chronic hepatitis B patients. Gastroenterology 2005, 129:1198–1209.
45. Lieu VT, Sreen IS, Lee CM, Aruca US, Papatheodoridis GV, Suet-Hing Wong F, Chang TT, Horban A, Wang C, Kwan P, Buti M, Prieto M, Berg T, Kritinos K, Peschell K, Mondou E, Frederick D, Rousseau F, Schiff ER: Tenofovir disoproxil fumarate (TDF), emtricitabine/TDF, and entecavir in patients with decompensated chronic hepatitis B liver disease. Hepatology 2011, 53:62–72.
46. Lieu VT, Sung JJ, Chow WC, Farrell G, Lee CZ, Yuen H, Tanwarde T, Tao QM, Shue K, Keene ON, Dixon JS, Gray DF, Sabat J, Cirrosis Asian Lamivudine Mucibence Study Group: Lamivudine for patients with chronic hepatitis B and advanced liver disease. N Engl J Med 2004, 351:1521–1531.
47. Lai CL, Rosmawatt L, Mao L, Van Wier bergen H, Anderson FH, Thomas N, Dehertog DH: Entecavir is superior to lamivudine in reducing hepatitis B virus DNA in patients with chronic hepatitis B infection. Gastroenterology 2002, 123:1831–1839.
48. Lau GK, PratPassa T, Luo KK, Marcellin P, Thongsawat S, Cooksey G, Gane E, Fried MW, Chow WC, Paik SW, Chang WY, Berg T, Flisher R, McCloud P, Pluck N, Peginterferon Alfa-2a HBeAg-Positive Chronic Hepatitis B Study Group: Peginterferon alfa-2a, lamivudine, and the combination for HBeAg-positive chronic hepatitis B. N Engl J Med 2005, 352:2682–2695.
49. Jansen JP, Naci H: Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. BMC Med 2013, 11:59.
50. Salanti G: Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Meth 2012, 3:80–97.
51. Dakin H, Fidler C, Harper C: Mixed treatment comparison meta-analysis evaluating the relative efficacy of nucleos(t)ides for treatment of nucleos(t)ide-naive patients with chronic hepatitis B. Value Health 2010, 13:934–945.
52. Woy G, Tomlinson G, Nishikawa Y, Kowgier M, Sherman M, Wong DK, Pham B, Unger WJ, Enarson TR, Heathcote EJ, Krahm M: Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and Bayesian meta-analyses. Gastroenterology 2010, 139:1218–1229.
53. Wens A, LenT L, Venson R, Correr CC, Inajara I, Pedrosos ML, Pontarolo R: Comparative efficacy of oral nucleoside or nucleotide analog monotherapy used in chronic hepatitis B: a mixed-treatment comparison meta-analysis. Pharmacotherapy 2013, 33:144–151.
54. National Institute for Health and Clinical Excellence: Guideline to the Methods of Technology Appraisal, 12 http://www.nice.org.uk/aboutnice/howwework/devnicetech/guidetothemethodsoftechnologyappraisal.jsp.
55. Veenstra DL, Sullivan SD, Clarke L, Iloeje UH, Tafesse E, Di Bisceglier A, Kowdle KV, Gish RG: Cost effectiveness of entecavir versus lamivudine with adefovir salvage in HBeAg-positive chronic hepatitis B. Pharmacoeconom 2007, 25:963–977.
56. Veenstra DL, Sullivan SD, Dushko GM, Jacobs M, Aledort JE, Lewis G, Patel KH: Cost-effectiveness of peginterferon alpha-2a compared with lamivudine treatment in patients with HBe-antigen-positive chronic hepatitis B in the United Kingdom. European J Gastroenterol Hepatol 2007, 19:631–638.
57. Dakin H, Bentley A, Dushko G: Cost-utility analysis of tenofovir disoproxil fumarate in the treatment of chronic hepatitis B. Value Health 2010, 13:922–935.
58. Shepherd J, Gospodarevskaya E, Farampton G, Cooper K: Entecavir for the treatment of chronic hepatitis B infection. Health Technol Assess 2009, 135Suppl 831–8.
59. Bristol-Myers Squibb Ltd: Entecavir (Baraclude) for the Treatment of Chronic Hepatitis B: Single Technology Appraisal Submission to the National Institute for Health and Clinical Excellence: http://www.nice.org.uk/nicemedia/live/11833/40303/40303.pdf.
60. Iloeje UH, Yang HI, Su J, Jern CIL, You SJ, Chen CJ: Risk Evaluation of Viral Load Elevation and Associated Liver Disease/Cancer-In HBV (the REVEAL-HBV) Study Group: Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. Gastroenterology 2006, 130:678–686.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit