PEER REVIEW HISTORY

This paper was submitted to the BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open where it was re-reviewed and accepted.

ARTICLE DETAILS

TITLE (PROVISIONAL)	ANTIVIRAL THERAPY FOR PREVENTION OF HEPATOCELLULAR CARCINOMA AND MORTALITY IN CHRONIC HEPATITIS B: SYSTEMATIC REVIEW AND META-ANALYSIS
AUTHORS	Thiele, Maja; Gluud, Lise; Dahl, Emilie; Krag, Aleksander

VERSION 1 - REVIEW

REVIEWER	**Schwarzinger, Michael** INSERM U912, SE4S
REVIEW RETURNED	27-Apr-2013

GENERAL COMMENTS	* Originality - does the work add enough to what is already in the published literature? If so, what does it add? If not, please cite relevant references. Partially. 7 meta-analyses were conducted so far (references: 16-18, 20,21,23,24), although inconsistent results were found. The originality of the presented meta-analysis (based on aggregate data) is to incorporate studies of different designs (RCT, longitudinal studies, and case-control series). In my opinion, a more detailed meta-analysis could be performed to clarify the role of other sources of heterogeneity than study design (see Results section below). * Importance of work to general readers - does this work matter to clinicians, patients, teachers, or policymakers? Is a general journal the right place for it? No. As it stands, teachers would mostly benefit from new material showing how study design may affect study results. However, the meta-analysis mostly reveals that too limited numbers of patients were enrolled in RCT and longitudinal studies as compared to the numerous case-control series: even meta-analysis cannot show any convincing effect of HBV treatment on long-term outcomes. In my opinion, it would be of interest for clinicians and policy makers that the authors elaborate on the last statement pages 14-15: "Previous evidence shows that interventions supported by surrogate markers may in fact have no benefit or even harmful effects on clinical outcome measures68. Still, our findings are not sufficiently convincing and do not allow for changes in clinical practice. As it would be unethical for future randomised trials on antiviral treatment to include an untreated control group, this problem can probably not be overcome." It seems a bit paradoxical to call for noninferiority

	randomized trials for future drugs, while ongoing treatment did not show significant improvement on long-term outcomes versus no treatment. * Scientific reliability Research Question - clearly defined and appropriately answered? Partially (see Results section below). * Overall design of study - adequate? Yes. * Participants studied - adequately described and their conditions defined? Partially. The meta-analysis is based on aggregate data. Besides study design, there are many confounding factors other than those tested (page 9) that may explain discrepant results on long-term outcomes. Some patient characteristics are presented in Tables 1 & 2, but other factors contributing to liver fibrosis progression may be of interest (HIV, alcohol) as well as the context of study (high-income vs. lower income countries). * Methods - adequately described? Complies with relevant reporting standard – Eg CONSORT for randomised trials? Ethical? Yes * Results - answer the research question? Credible? Well presented? Partially. Although the authors had no access to individual data, they performed a careful meta-analysis. In my opinion, it would be of interest if the authors could test in their random effects inverse variance meta-analysis previous variables to take into account the heterogeneity of studies (and not only study design). * Interpretation and conclusions - warranted by and sufficiently derived from/focused on the data? Message clear? Yes * References - up to date and relevant? Any glaring omissions? Yes, up to date (2012) and relevant. *Abstract/summary/key messages - reflect accurately what the paper says? Yes.

REVIEWER	**Cammà, Calogero** Sezione di Gastroenterologia, Di.Bi.M.I.S.
REVIEW RETURNED	02-May-2013

GENERAL COMMENTS	The authors have attempted a meta-analysis of the available literature, in order to evaluate whether IFN, lamivudine or IFN plus NA combination therapy, compare to no treatment, reduces the risk of developing HCC or all-cause mortality in patients with chronic HBV infection. Thirty-five studies (8 RCTs, 8 prospective cohort studies and 19 case control studies) were assessed. Grouped rather than individual patient data were used. Authors concluded that

	architecture research clearly influences the estimate of treatment effect on HCC development and mortality. IFN and first generation NA have no proven effect on these true outcomes. Although this paper was conducted with rigorous methodology and there is some valuable information contained in this paper regarding the issue of HCC chemoprevention which is of interest to hepatologists, it does not add any novel observations. The authors do not make a sufficiently convincing case that this study will add significantly to clinical practice and to future research. I am not sure whether the data presented in this meta-analysis is necessarily stronger than data of the several meta-analyses that have been previously published . The studies pooled in this meta-analysis are heterogeneous in terms of country of origin, design, baseline features of patients included, schedules of treatment, surveillance/screening methodology and diagnostic criteria for HCC, making it more difficult to draw firm conclusions and limiting the conclusions that may be drawn regarding the effectiveness of antiviral therapy in preventing HCC development. Ideally, a meta-analysis would include individual patient data with up-dated follow-up. Unfortunately, the overall interest of the study is limited by the clinical heterogeneity among the included studies, which limits the conclusions that may be drawn regarding the chemoprevention efficacy of IFN or first-generation NA treatments. Another problem with this meta-analysis is that there are limitations in the data, which are beyond the authors' control, but nevertheless compromise the value of the study. In particular, the lack of data according to treatment response and the inclusion of studies with different length of follow-up hamper the validity of the pooled estimates of the effect size and the authors are very fair in pointing out the resultant limitations on the conclusion that may be drawn from these . Several important methodological issues must also be reconsidered. 1. A major concern of this meta-analysis is that the studies selected are heterogeneous from a clinical point of view. I agree that it is correct to test for quantitative heterogeneity. The significant statistical heterogeneity observed in many of the performed analyses is a valid reason for choosing a random effect over a fixed effect model. The random-effect model should take care of the quantitative heterogeneity. However, whichever statistical method is chosen one needs to be confident that clinical and methodological diversity is not so great that the studies should not be combined at all. If there is substantial clinical heterogeneity, it is preferable not to pool the studies. Therefore, when a significant heterogeneity of baseline risk is found, more detailed treatment comparisons could be achieved by a meta-analysis of individual patient data only. Referee suggests that the authors may be looking to perform a meta-regression to see if both study-level as well as patient-level covariates could explain the observed heterogeneity and if they are effect modifiers. Please try to include in the meta-regression model study- and patient-level covariates as well as data on the quality assessment of the studies. 2. Any meta-analysis is susceptible to publication bias. One can sympathize with the authors when they included studies published as full paper only. In this heterogeneous setting, the authors restricted the literature search excluding abstracts. However, many of those performing meta-analysis include abstracts in the literature search in order to contain all the work in the field. The authors may

	need to justify themselves. 3. Nonrandomized trials may experience many problems that could reduce their internal and external validity. Their lack of precision and reliability causes inherent biases towards false positive results. When assessing nonrandomized trials, the most important bias is the likelihood of inappropriate selection of patients for treatment, which can lead to incorrect results and spurious associations. Therefore, therapeutic guidelines cannot be definitively derived from nonrandomized trials unless the observed benefit of treatment is relevant and the clinical course of untreated patients predictable by reliable prognostic models. None of the studies included in this meta-analysis fulfils these methodological conditions. 4. The unavailability of individual data hampers the analysis of the HCC prevention benefit as a time-dependent variable. Furthermore, it is becoming recognized that the results of meta-analyses of time-to-event outcomes are likely to be affected by censoring and by the duration of follow up of individual trials (Vale CL, et all Effects of adjusting for censoring on meta-analyses of time-to-event outcomes. Int J Epidemiol 2002;31:107-111). These limitations are particularly important when the follow up across trials are heterogeneous, as were the mean follow up periods of the studies included in this meta-analysis. 5. Another relevant issue of this meta-analysis is the lack of data on viral load, expressed as serum HBV-DNA levels, as well as on HBeAg seroconversion after therapy and on HBV genotypes. 6. We know very little about surveillance and diagnostic procedures of HCC in this meta-analysis. Studies conducted in Europe, Asia, Africa, North and South America that included HBV infected patients with or without cirrhosis, were pooled. A large variability of surveillance/screening procedures and diagnostic criteria among studies could be expected. We need the maximum of information concerning surveillance/screening procedures and diagnostic criteria of the included studies.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Recommendation:

Comments:

* Originality - does the work add enough to what is already in the published literature? If so, what does it add? If not, please cite relevant references.

Partially. 7 meta-analyses were conducted so far (references: 16-18, 20,21,23,24), although inconsistent results were found.

The originality of the presented meta-analysis (based on aggregate data) is to incorporate studies of different designs (RCT, longitudinal studies, and case-control series). In my opinion, a more detailed

meta-analysis could be performed to clarify the role of other sources of heterogeneity than study design (see Results section below).

1.1. Author reply:

In addition to study design, we have included subgroup and sensitivity analyses on parameters that could account for between-study heterogeneity: i) type of intervention (interferon, nucleos(t)ides or combined therapy; ii) low risk of bias trials; iii) whether HCC screening was performed. We have additionally performed subgroup analyses in patients with cirrhosis.

We have sought to further investigate heterogeneity in the HCC incidence analysis of RCTs and observational studies and showing high statistical heterogeneity (I^2 = 63%) by performing a post-hoc meta-regression analysis. The following variables were included in the meta-regression analysis: proportion of men (coefficient -0.074; P=0.08), mean age of treated patients at inclusion (coefficient 0.020; P=0.94), mean age of untreated patients at inclusion (coefficient 0.121; P=0.65), proportion of patients with cirrhosis at inclusion (coefficient -0.001; P=0.76), and region of trial (coefficient -0.394; P=0.55).

* Importance of work to general readers - does this work matter to clinicians, patients, teachers, or policymakers? Is a general journal the right place for it?

No. As it stands, teachers would mostly benefit from new material showing how study design may affect study results. However, the meta-analysis mostly reveals that too limited numbers of patients were enrolled in RCT and longitudinal studies as compared to the numerous case-control series: even meta-analysis cannot show any convincing effect of HBV treatment on long-term outcomes.

In my opinion, it would be of interest for clinicians and policy makers that the authors elaborate on the last statement pages 14-15: “Previous evidence shows that interventions supported by surrogate markers may in fact have no benefit or even harmful effects on clinical outcome measures68. Still, our findings are not sufficiently convincing and do not allow for changes in clinical practice. As it would be unethical for future randomised trials on antiviral treatment to include an untreated control group, this problem can probably not be overcome.” It seems a bit paradoxical to call for noninferiority randomized trials for future drugs, while ongoing treatment did not show significant improvement on long-term outcomes versus no treatment.

1.2 Author reply:

The authors agree that the number of RCTs is limited. We have stated the reason for including observational studies in the methods section (“Due to the expected prognosis and the duration of follow up necessary to evaluate intervention effects on clinical outcome measures in HBV, observational studies were included in sensitivity analyses.”). **We have adressed the limited number of RCTs in the discussion.**

Meta-analyses assessing the effect of various types of medication (i.e. statins[1]) on HCC incidence and including observational studies are currently being published. We believe our findings of detection and ascertainment bias in the HCC incidence estimate provide valuable new knowledge for clinicians and researchers when appraising results of both meta-analyses and single observational studies on HCC incidence. The topic of bias is adressed in detail in

the discussion.

We agree that the statement on pages 14-15 can be sen as paradoxical and have accordingly limited the statement to: "Previous evidence shows that interventions supported by surrogate markers may in fact have no benefit or even harmful effects on clinical outcome measures. Still, our findings are not sufficiently convincing and do not allow for changes in clinical practice."

* Scientific reliability Research Question - clearly defined and appropriately answered?

Partially (see Results section below).

* Overall design of study - adequate?

Yes.

* Participants studied - adequately described and their conditions defined?

Partially. The meta-analysis is based on aggregate data. Besides study design, there are many confounding factors other than those tested (page 9) that may explain discrepant results on long-term outcomes. Some patient characteristics are presented in Tables 1 & 2, but other factors contributing to liver fibrosis progression may be of interest (HIV, alcohol) as well as the context of study (high-income vs. lower income countries).

1.3 Author reply:

We agree that many confounding factors exist for HCC and mortality in HBV. We have tried adress such factors by performing a meta-regression analysis (see point 1.1. above). We did extract data on co-infections (HIV, HCV, HDV), genotypes, concurrent alcohol abuse, HBeAg-status, treatment response, seroconversions of HBsAg and HBeAg, and inflammatory and viral activity at inclusion (by alanine amino transferase and HBV-DNA level). There was unfortunately not enough data to allow for meta-analyses. We have now included a section in the discussion adressing this particular subject. We have additionally referred the raw data regarding treatment response and HCC risk with a discussion of possible conclusions that can be drawn from the data.

* Methods - adequately described? Complies with relevant reporting standard – Eg CONSORT for randomised trials? Ethical?

Yes

* Results - answer the research question? Credible? Well presented?

Partially. Although the authors had no access to individual data, they performed a careful meta-analysis. In my opinion, it would be of interest if the authors could test in their random effects inverse variance meta-analysis previous variables to take into account the heterogeneity of studies (and not only study design).

1.4 Author reply:

Please see answer above, 1.1.

* Interpretation and conclusions - warranted by and sufficiently derived

from/focused on the data? Message clear?

Yes

* References - up to date and relevant? Any glaring omissions?

Yes, up to date (2012) and relevant.

*Abstract/summary/key messages - reflect accurately what the paper says?

Yes.

Additional Questions:

Please enter your name: Michaël Schwarzinger

Job Title: Researcher

Institution: INSERM, France

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may

in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way

gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://bit.ly/VW8GVB'target='_new'>(please see BMJ Group policy) please declare them here:

Reviewer: 2

Recommendation:

Comments:

The authors have attempted a meta-analysis of the available literature, in order to evaluate whether IFN, lamivudine or IFN plus NA combination therapy, compare to no treatment, reduces the risk of developing HCC or all-cause mortality in patients with chronic HBV infection. Thirty-five studies (8 RCTs, 8 prospective cohort studies and 19 case control studies) were assessed. Grouped rather than individual patient data were used. Authors concluded that architecture research clearly influences the estimate of treatment effect on HCC development and mortality. IFN and first generation NA have no proven effect on these true outcomes.

Although this paper was conducted with rigorous methodology and there is some valuable information contained in this paper regarding the issue of HCC chemoprevention which is of interest to hepatologists, it does not add any novel observations. The authors do not make a sufficiently convincing case that this study will add significantly to clinical practice and to future research. I am not sure whether the data presented in this meta-analysis is necessarily stronger than data of the several meta-analyses that have been previously published .

The studies pooled in this meta-analysis are heterogeneous in terms of country of origin, design, baseline features of patients included, schedules of treatment, surveillance/screening methodology and diagnostic criteria for HCC, making it more difficult to draw firm conclusions and limiting the conclusions that may be drawn regarding the effectiveness of antiviral therapy in preventing HCC

development. Ideally, a meta-analysis would include individual patient data with up-dated follow-up. Unfortunately, the overall interest of the study is limited by the clinical heterogeneity among the included studies, which limits the conclusions that may be drawn regarding the chemoprevention efficacy of IFN or first-generation NA treatments.

Another problem with this meta-analysis is that there are limitations in the data, which are beyond the authors' control, but nevertheless compromise the value of the study. In particular, the lack of data according to treatment response and the inclusion of studies with different length of follow-up hamper the validity of the pooled estimates of the effect size and the authors are very fair in pointing out the resultant limitations on the conclusion that may be drawn from these .

2.1 Author reply:

We agree with the reviewer that a meta-analysis including individual patient data would be optimal, especially when many competing risk factors for HCC development exist. Individual patient data could regrettably not be obtained. We are thankful that the reviewer acknowledges our discussion of study limitations. Regarding novelty, we refer to our answers above (1.2.)

Several important methodological issues must also be reconsidered.

1. A major concern of this meta-analysis is that the studies selected are heterogeneous from a clinical point of view. I agree that it is correct to test for quantitative heterogeneity. The significant statistical heterogeneity observed in many of the performed analyses is a valid reason for choosing a random effect over a fixed effect model. The random-effect model should take care of the quantitative heterogeneity. However, whichever statistical method is chosen one needs to be confident that clinical and methodological diversity is not so great that the studies should not be combined at all. If there is substantial clinical heterogeneity, it is preferable not to pool the studies. Therefore, when a significant heterogeneity of baseline risk is found, more detailed treatment comparisons could be achieved by a meta-analysis of individual patient data only.

Referee suggests that the authors may be looking to perform a meta-regression to see if both study-level as well as patient-level covariates could explain the observed heterogeneity and if they are effect modifiers. Please try to include in the meta-regression model study- and patient-level covariates as well as data on the quality assessment of the studies.

2.2. Author reply:

Please see our answer above (1.1.). We agree with the reviewer that clinical heterogeneity is a common and difficult problem in meta-analyses. We have sought to adress heterogeneity by performing I-square statistics and comparing random- and fixed effects meta-analyses. We respectfully point out that only one analysis (HCC incidence combining RCTs and observational studies) had I-square above 25%, and that no analyses differed in fixed and random effects analyses.

2. Any meta-analysis is susceptible to publication bias. One can sympathize with the authors when they included studies published as full paper only. In this heterogeneous setting, the authors restricted the literature search excluding abstracts. However, many of those performing meta-analysis include abstracts in the literature search in order to contain all the work in the field. The authors may need to justify themselves.

2.3. Author reply:

We agree with the reviewer that a meta-analytic approach requires searching the full body of available evidence. As stated in the MOOSE checklist, no abstract publications were eligible for inclusion; largely due to lack of relevant outcome data.

3. Nonrandomized trials may experience many problems that could reduce their internal and external validity. Their lack of precision and reliability causes inherent biases towards false positive results. When assessing nonrandomized trials, the most important bias is the likelihood of inappropriate selection of patients for treatment, which can lead to incorrect results and spurious associations. Therefore, therapeutic guidelines cannot be definitively derived from nonrandomized trials unless the observed benefit of treatment is relevant and the clinical course of untreated patients predictable by reliable prognostic models. None of the studies included in this meta-analysis fulfils these methodological conditions.

2.4 Author reply:

We fully agree with the reviewer's comments. We have tried to adress the subject of bias in observational studies in detail in the discussion section. Accordingly, we do not recommend changes of practice even though some of our sensitivity analyses including observational studies (on HCC in cirrhosis, and on over-all mortality) show a positive effect of treatment.

4. The unavailability of individual data hampers the analysis of the HCC prevention benefit as a time-dependent variable. Furthermore, it is becoming recognized that the results of meta-analyses of time-to-event outcomes are likely to be affected by censoring and by the duration of follow up of individual trials (Vale CL, et all Effects of adjusting for censoring on meta-analyses of time-to-event outcomes. Int J Epidemiol 2002;31:107-111). These limitations are particularly important when the follow up across trials are heterogeneous, as were the mean follow up periods of the studies included in this meta-analysis.

2.5 Author reply:

Please see above (2.1. and 2.2.)

5. Another relevant issue of this meta-analysis is the lack of data on viral load, expressed as serum HBV-DNA levels, as well as on HBeAg seroconversion after therapy and on HBV genotypes.

2.6. Author reply:

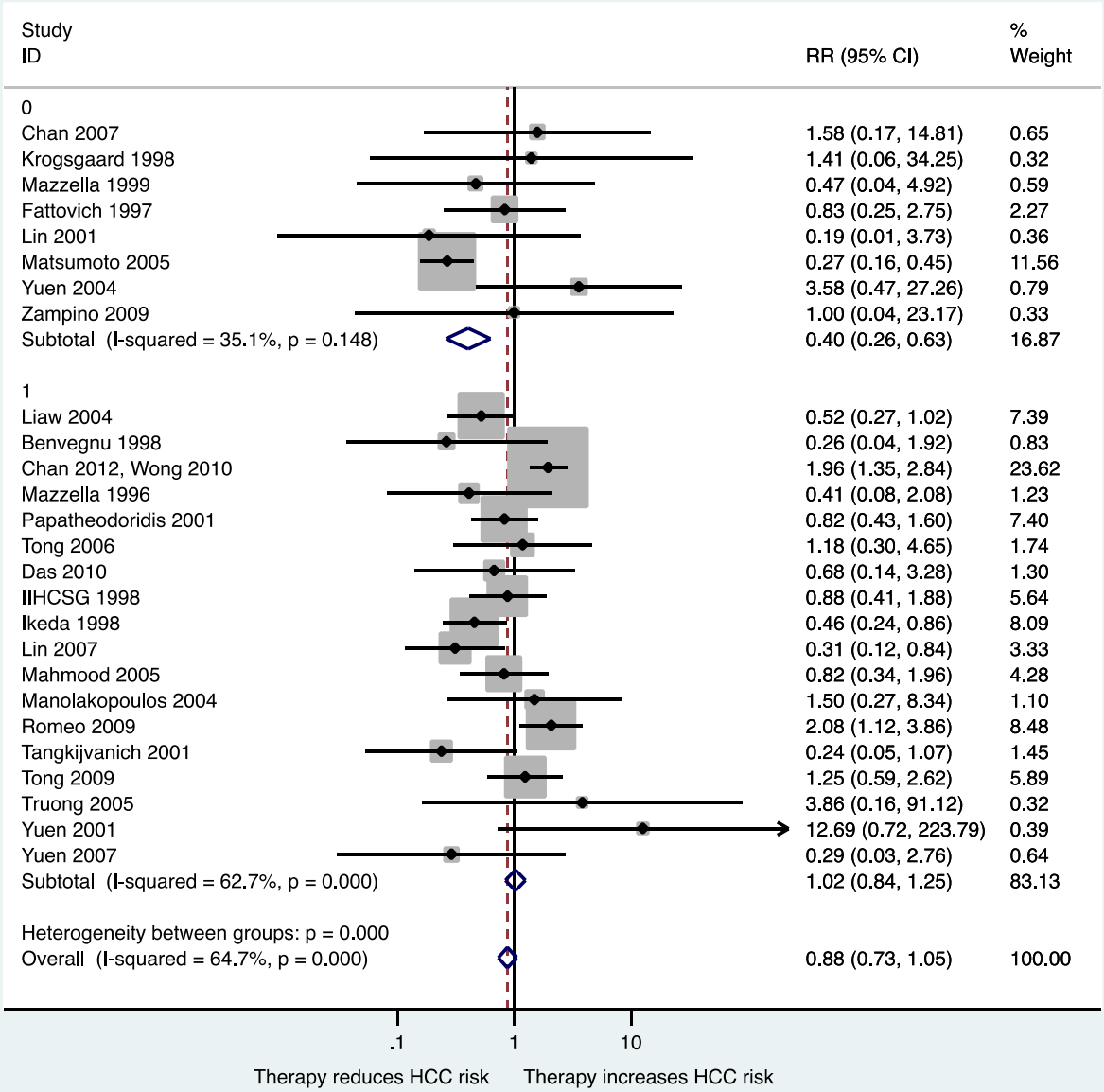
Please see answer above (1.3.)

6. We know very little about surveillance and diagnostic procedures of HCC in this meta-analysis. Studies conducted in Europe, Asia, Africa, North and South America that included HBV infected patients with or without cirrhosis, were pooled. A large variability of surveillance/screening procedures and diagnostic criteria among studies could be expected. We need the maximum of information concerning surveillance/screening procedures and diagnostic criteria of the included studies.

2.7. Author reply:

The authors agree with the reviewer that HCC screening is vital to correct detection of HCC. In the primary outcome analysis including only RCTs we did not find any subgroup differences when adjusting for HCC screening, most likely due to lack of power. We did however find a subgroup difference in the sensitivity analyses including observational studies (*P*<0.001). Trials not performing HCC screening showed a benefit of treatment (8 trials, RR 0.40, CI 0.26-0.63), whereas trials performing HCC screening did not (18 trials, RR 1.02, CI 0.84-1.25).

In the manuscript we choose to focus on the observed subgroup differences regarding trial design, as we believed this to be the most clinically relevant subgroup difference. For length purposes, the subgroup difference regarding HCC screening was therefore omitted from the text.



Additional Questions:

Please enter your name: calogero camma

Job Title: full professor of gastroeterology

Institution: University of palermo

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may

in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way

gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://bit.ly/VW8GVB'target='_new'>(please see BMJ Group policy) please declare them here:

If you elected during submission to send your article on to another journal the article will be transferred in 5 working days. If you intend to appeal against this decision please notify us before then.

The journal(s) (if any) you have selected at submission are:

If you want to speed up or stop this onward transmission please email the editorial office: papersadmin@bmj.com

References

1. Singh S, Singh PP, Singh AG, et al. Statins Are Associated With a Reduced Risk of Hepatocellular Cancer: A Systematic Review and Meta-analysis. Gastroenterology 2013;**144**(2):323-32

VERSION 2 – REVIEW

REVIEWER	**Schwarzinger, Michael** researcher, Inserm, France
REVIEW RETURNED	13-Jun-2013

- The reviewer completed the checklist but made no further comments.

REVIEWER	Hinde, Andrew University of Southampton, Southampton Statistical Sciences Research Institute
REVIEW RETURNED	01-Jul-2013

THE STUDY	On p. 7, ll. 1-2 you write '[t]o avoid prevalent cases of HCC the outcomes were assessed after at least 12 months of follow up'. I think you should emphasise that you mean you want to avoid patients who has HCC at initial contact, and the 12 months is conservative because once a patient has developed HCC, the condition is generall deadly well before 12 months. 'Two authors extracted data in an independent manner' (p. 7, l. 4). Do you mean 'independently'? If so, just say 'independently'.
RESULTS & CONCLUSIONS	Egger's test is not very powerful with small numbers of studies, and you do not have many studies here as you acknowledge. So the fact that you cannot demonstrate small-study bias (p. 10, l. 11) may not mean it does not exist. Is Egger's test really appropriate with so few studies, especially when you consider just the RCTs. When reporting the results there are apparent inconsistencies. On p. 7, ll. 9-10 you write that 'HCC was diagnosed in 22 of 840 patients in the treatment group versis 19 of 447 controls (relative risk 0.58 ...' However, (22/840)/(19/447) = 0.62 not 0.58. Similarly, on p. 11, ll. 12-13 you write 'one of 20 patients in the treatment group and two of 12 controls developed HCC (relative risk 0.75 ...' but (1/20)/(2/12) = 0.3 not 0.75. Clearly in the second of these there is an enormous confidence interval which does include 0.3, but the inconsistency needs explaining. On p. 14, ll. 6-7 do you have any explanation to suggest for the apparent finding that 'the study design was closely related to the estimated treatment effects'? You mention 'detection and

	ascertainment bias' and 'confounding by indication', but could you give specific examples of how these mechanisms might operate? I think you could expand the discussion of this point. This is where you might make some genuinely novel and useful observations, but you shy away from them!
GENERAL COMMENTS	This version of the paper appears to be a revision of an earlier submission. As I did not review the earlier submission I have read this version as a new paper. However, I was provided with the comments of the reviewers of the original submission and the authors' responses. The main criticism of the original paper seems to have been the lack of consideration of the clinical heterogeneity of the studies included. The authors have made some effort to address this be computing I-squared statistics and performing a meta-regression in the case where I-squared was high. My view is that they have done about as much as the data will allow. The first conclusion of the study is a negative one, and this is correct. The effect of antiviral therapy 'remains to be established'. The second conclusion that 'research design plays an essential role in the overall assessment; is interesting, and I was disappointed that you did not make more of this in the discussion and conclusions, to think through and suggest reasons why this might be so in order to generate hypotheses for future research to address.

VERSION 2 – AUTHOR RESPONSE

On p. 7, ll. 1-2 you write '[t]o avoid prevalent cases of HCC the outcomes were assessed after at least 12 months of follow up'. I think you should emphasise that you mean you want to avoid patients who has HCC at initial contact, and the 12 months is conservative because once a patient has developed HCC, the condition is generall deadly well before 12 months.
AUTHOR REPLY: We agree with the reviewer and the text has been clarified accordingly (page 7, lines 2-4).

'Two authors extracted data in an independent manner' (p. 7, l. 4). Do you mean 'independently'? If so, just say 'independently'.
AUTHOR REPLY: We agree with the reviewer and the text has been changed accordingly (page 7, lines 5).

Egger's test is not very powerful with small numbers of studies, and you do not have many studies here as you acknowledge. So the fact that you cannot demonstrate small-study bias (p. 10, l. 11) may not mean it does not exist. Is Egger's test really appropriate with so few studies, especially when you consider just the RCTs.
AUTHOR REPLY: Egger’s test is recommended in the Cochrane Handbook (section 10.4.3) as the preferred statistical test for funnel plot assymetry. We have not been able to find an alternative test that performs better with few studies. We do agree that the number of studies limit the inferences we can make based on the result of the Egger’s test. We have mentioned this in the discussion.

When reporting the results there are apparent inconsistencies. On p. 7, ll. 9-10 you write that 'HCC was diagnosed in 22 of 840 patients in the treatment group versis 19 of 447 controls (relative risk 0.58 ...' However, (22/840)/(19/447) = 0.62 not 0.58. Similarly, on p. 11, ll. 12-13 you write 'one of 20 patients in the treatment group and two of 12 controls developed HCC (relative risk 0.75 ...' but (1/20)/(2/12) = 0.3 not 0.75. Clearly in the second of these there is an enormous confidence interval

which does include 0.3, but the inconsistency needs explaining.
AUTHOR REPLY: In our review we report relative risks generated in a meta-analysis. This number does not correspond to crude relative risks. Part of the discrepancy is due to the fact that trials without events were excluded from the meta-analysis. We have therefore changed the wording and only report the number of events and patients in trials included in the meta-analysis.

On p. 14, ll. 6-7 do you have any explanation to suggest for the apparent finding that 'the study design was closely related to the estimated treatment effects'? You mention 'detection and ascertainment bias' and 'confounding by indication', but could you give specific examples of how these mechanisms might operate? I think you could expand the discussion of this point. This is where you might make some genuinely novel and useful observations, but you shy away from them!
This version of the paper appears to be a revision of an earlier submission. As I did not review the earlier submission I have read this version as a new paper. However, I was provided with the comments of the reviewers of the original submission and the authors' responses. The main criticism of the original paper seems to have been the lack of consideration of the clinical heterogeneity of the studies included. The authors have made some effort to address this be computing I-squared statistics and performing a meta-regression in the case where I-squared was high. My view is that they have done about as much as the data will allow. The first conclusion of the study is a negative one, and this is correct. The effect of antiviral therapy 'remains to be established'. The second conclusion that 'research design plays an essential role in the overall assessment; is interesting, and I was disappointed that you did not make more of this in the discussion and conclusions, to think through and suggest reasons why this might be so in order to generate hypotheses for future research to address.
AUTHOR REPLY: We agree and believe that the observed difference in intervention effects according to study design is novel and useful knowledge. In order to stress the importance of the finding we have added possible explanations to the discussion section; this includes a few lines on the subgroup difference regarding HCC screening (pages 11 and 14).