Reviewer's report

Title: The epidemiology of hepatitis C virus in Egypt: A systematic review and data synthesis

Version: 2 Date: 8 September 2012

Reviewer: Leigh Anne Shafer

Reviewer's report:

Overall, the paper would make a very helpful addition to the literature on prevalence and incidence of HCV in Egypt. There are so many publications on small subgroups of the Egyptian population that it is difficult to see the big picture about the impact of the HCV epidemic in Egypt. This paper nicely summarizes the literature and is highly recommended for publication, after responding to those suggestions below that the authors feel are helpful.

Feedback for improvement are listed below under MAJOR and MINOR suggestions.

MAJOR

1. Two suggestions for the drivers of current HCV prevalence are given – the PAT theory with limited current incidence, and the theory that current risk factors may be driving the current epidemic. However, I think a third theory should be explored or discussed – that PAT initiated the high prevalence in Egypt, and that the high prevalence brought on initially by PAT is driving higher incidence in Egypt today than what might be seen in neighboring countries even if other exposures are similar. This would occur simply because prevalence is higher, which drives incidence because more infectious people exist (prevalence) so more people get infected (incidence). In your discussion (p. 15), you state that prevalence studies among women and children not exposed to PAT found substantial prevalence levels – and you reasonably suggest that this is evidence of other (non-PAT) sources of infection. This is true, but it could be explored or discussed whether we think this is because the risk factors (unclean needles, contaminated blood, etc) are higher in Egypt than in surrounding countries, or whether it is simply the residual effect of the earlier PAT campaigns resulting in currently higher prevalence and therefore currently higher incidence, even among people unexposed to PAT.

2. If a person wants to infer from your pooled study analysis a possibly less-biased trend in HCV prevalence than can be obtained from just one study, then I think: if the year of data collection is unknown for a given study, then a good estimate of this is certainly earlier than the year of study publication. I usually use 1 to 2 years earlier than the year of publication which is probably a “late” estimate for the actual year of data collection, but at least is more in line with how quickly studies get published than using the year of the publication.
3. How many of the studies fell into the category of those for which the year of data collection was unknown? This can give the reader an idea about how much they should worry about this concern – e.g., only 3 or 4 studies out of your 150 would likely have little impact on your estimated trends.

4. Figure 4, showing HCV trends among different subgroups, is nice! However: (a) It doesn’t seem to correspond completely with figure 2. For example, in figure 2, it appears that there are >5 general population studies, yet in figure 4, it looks like there are only 5. Which general population studies were selected for your figure 4? (b) I would love to see trends in a few more subgroups, those with a large enough sample of studies to warrant a trend – such as hemodialysis patients, multi-transfused patients, thalassemic patients, and viral hepatitis patients.

5. It was unclear from the methods as well as from the very short 1-paragraph of results, how you determined the trend in prevalence over time. You “adjusted for different subgroups”. How? Perhaps showing us the exact model and results of the model might help.

6. People previously exposed to PAT, such as can be obtained from the Egypt DHS, should be among those at high risk of HCV infection. This should be stated when you first define your groups, on p. 7. On p. 10, you do put schistosomiasis patients in the high risk group, but do not distinguish those treated during PAT and those treated recently (orally).

7. Figures 2 and 3 are great! I do wonder whether you would consider improving at least figure 2 even more by having it help us assess descriptively any possible trend? This could be done by dividing your study years into two groups as you see fit, for example pre-2000 and 2000+, or pre-1995 and 1995+. Each column of dots on figure 2 would then be replaced by two side-by-side columns, one for the prevalence estimates in the early period, and one for estimates in the late period. The IDU category would obviously only have it's dot in the relevant time-period column.

8. Define what you mean by ‘general population’. If you mean populations that should not be at elevated risk of HCV, should blood donors be in this group?

9. I found the risk factors of HCV section weaker than the prevalence and incidence sections. Can you find a way of quantifying the different risk factors, rather than just stating that increasing age, history of PAT, and rural areas were the “most common” risk factors (which feels subjective). I wonder if a regression can be performed with the outcome being prevalence (so you would have 146 observations – 1 from each prevalence study) and the independent variables being risk factors such as age, rural, injections, blood transfusions… The “noise” resulting from the fact that some people in a general population study in a village will also have had blood transfusions, etc, would dampen the estimated impact of the various risk factors, but that’s o.k. What you want to do is compare the risk factors to each other, so if all of their associations with HCV is dampened, that’s
o.k. … Or perhaps you can find a different way to quantify these associations. As it is, the lack of quantitative comparisons in this section makes it weak and hard to judge from a reader’s point of view whether the categories you give as the “most common” risk factors are accurate.

MINOR (mainly edits)
1. From what year does the prevalence estimate of 14.7% come? I’m guessing from 2008 (the DHS)? This could be stated in the first sentence of the manuscript.

2. Do all 4 refs given in the first sentence of the manuscript indicate an estimated 14.7% prevalence? Or do refs 1, 3, and 4 just indicate that Egypt has the highest HCV prevalence in the world?

3. For inclusion, a study did not have to have both prevalence and incidence. The sentence at the top of p. 6 should read, “A publication was considered eligible for inclusion in the review if it had data on at least one of the following outcomes of interest:...”

4. Although you state in the manuscript text that if a study reported both incidence and prevalence, it was counted as two studies in your count, it would still be of interest and easy to do to show how many studies this was. On your figure 1, at the bottom, you show 4 reports with incidence, and 146 reports with prevalence. Can you show, on this figure, how many of the 4 incidence reports were also counted among the prevalence reports?

5. The “HCV Incidence” sub-header on p. 8 doesn’t seem aligned right.

6. Did you have 5 incidence studies (as stated on p. 8 in the text, and shown in table 1), or 4 incidence studies (as in your figure 1 flow-diagram, and stated in your methods)?

7. The reference (ref 22) at the bottom of p. 9 is in the wrong spot.

8. Schistosomiasis is misspelled on p. 10.

9. Tables 2-4 are very long for a main manuscript and probably belong in an appendix.

10. Could hypothesize in discussion the reason for high HCV prevalence in the special population groups – e.g., because of more hospital exposure, injections.

11. You state (p. 12) that RNA prevalence is shown in tables 1-5. However, table 1 is incidence and there is no RNA prevalence. I did not find any table 5.

12. It was not clear which parts of the schistosomiasis section of table 3 was PAT (other than the last two studies of that section).

**Level of interest:** An article of importance in its field
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests.