Author's response to reviews

Title: Incubation periods of viral gastroenteritis: a systematic review

Authors:

Rachel M Lee (rlee@jhsph.edu)
Justin Lessler (jlessler@jhsph.edu)
Rose A Lee (ralee@jhsph.edu)
Kara E Rudolph (krudolph@jhsph.edu)
Nicholas G Reich (nick@schoolph.umass.edu)
Trish M Perl (tperl@jhmi.edu)
Derek AT Cummings (dcumming@jhsph.edu)

Version: 4 Date: 17 July 2013

Author's response to reviews: see over
Dr. Nokes –
Please consider out attached revision of “Incubation periods of viral gastroenteritis: a systematic review” (1357234219371887) for publication in BMC Infectious Disease. We thank the reviewers for their valuable suggestions, and we believe the revised manuscript is greatly improved. The response to reviewers comments is included below, and author responses are shown in bold.

Please do not hesitate to contact me if you have any questions or concerns regarding this submission.

Sincerely,
Justin Lessler
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
jlessler@jhsph.edu

Response to the reviewers:

This is a thorough, well-constructed and well-written paper on an important topic. The authors are absolutely correct to highlight the fact that the latent period estimates used in dynamic models of GI pathogen transmission are too often arbitrary and not grounded in data. Accurate estimates of the distribution of latent periods for viral pathogens that cause gastroenteritis is important for obtaining meaningful transmission parameter estimates when applying dynamic models to data. I believe that with some additional analysis, the authors will have made a very strong contribution towards closing this gap in the literature.

Major Compulsory Revisions:
1. My concerns relate to the choice of a log-normal distribution for analysis of all of the latent period data. Although the authors cite several studies that show that log-normal distributions can explain incubation periods of acute infectious diseases, it would be more convincing if they demonstrated this via quantitative (AIC, BIC) comparisons of the goodness-of-fit for a log-normal distribution versus, e.g., a gamma or weibull distribution. The authors should present a comparison with at least 1 (but preferably more) alternative distributions.

As per the reviewer’s suggestion we have fit log-normal, gamma, Weibull and Erlang distributions to the incubation period data. Because all are two parameter distributions, models can be compared directly using their log-likelihood as in the new supplemental table. While in some cases (astrovirus and norovirus) the gamma distribution outperformed the log-normal, predicted times by which 5%, 25%, 50%, 75% and 95% of infections developed symptoms are the same within a 10th of a day in all cases. We continue to use the log-normal distribution as our primary analysis for consistency and
ease of interpretation. We have amended the methods to point interested readers to this comparison and analysis.

2. In addition, providing estimates of an Erlang distribution (gamma with an integer shape parameter) fit to the latent period data would make these findings invaluable for researchers constructing transmission models for these pathogens. Since the authors state this as a goal of their analysis, this area deserves additional attention. An Erlang distribution with shape parameter k and mean duration x can be included in a compartmental model as k compartments, each with expected (exponentially distributed) sojourn time x/k. Because transmission models used for these analyses are typically compartmental, estimates of the latent period based on the Erlang distribution are easy to plug in to these models. In addition, comparisons of the quality of fit of these model-ready distributions against the best-fitting distribution would provide a guide for interpreting modeling results based on these estimates.

We thank the reviewer for this excellent suggestion. We have included parameters and measures of fit for the Erlang distribution in the new supplemental table.

Minor Essential Revisions:
1. Figure 2 would be more informative if it contained another row showing the fit of these distributions to the data, i.e. plotting the fitted log-normal distribution against the histogram of the latent period data. If space limitations preclude this from being in the main text, it should be included in the supplementary materials.

Figure 3 showing the cumulative log-normal distributions superimposed onto a histogram of latent period data has been added.

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.

Major comments
This paper by Lee and colleagues presents a systematic review and pooled-analysis of the incubation periods of key agents of viral gastroenteritis. The paper is clearly written and presented. The motivation for the paper and the need for improved evidence-based estimates of the incubation period are nicely laid out. One of the simpler, but striking conclusions of the paper is to demonstrate the oft-cited estimated of the incubation period are based on
judgment or authoritative texts, with little supportive data. My main reservation concerns the data for the pooled analysis. It seems (though it’s not stated here) that the systematic review included search terms like ‘incubation period’. This would severely limit the search to papers where the authors specifically reported on the incubation period, though this is not strictly necessary because papers reporting time of exposure and time of onset would provide valuable data. From looking at the bibliography, it is clear that many papers with those types of data were not included. It may be that the authors considered this approach, but determined it was too non-specific and therefore unwieldy. It could also be argued that the limited data they did find was sufficient to generate robust estimates, though not for all pathogens, and certainly not enough to perform subgroup analysis (e.g. by age).

Please clarify how the data were extracted for the pooled analysis? Did these studies all report individual level data, and those data were abstracted. Following that is a question about the statistical analysis: assuming that it was in fact individual level data that were used, it seems that they were all counted as equally weighted observation, not accounting for the fact that they come from different studies. Why not use a random effect model that allowed for between study variation?

The studies from which data were extracted all individual level data, reported either in the text, in a table, or graphically. Only data for which a reasonable exposure interval (24 hours or less) could be determined and with which onset of symptoms could be linked were extracted. We recognize that the inclusion of “incubation period” limited our search results and prevented a comprehensive review of the entirety of the literature for each virus, a task that would be unwieldy. This limitation has been added to the discussion (lines 361-363).

I am actually surprised how few articles were identified. Matthews et al, Epi & Inf 2012 identified ~2400 articles just on norovirus outbreaks from a more limited search just on norovirus using PCR diagnostics in a defined time frame. It is not stated in the methods here if the terms related to incubation period were used (is stated in Lessler 2009), but that may have severely restricted the numbers of papers identified.

The following was added to the methods section to clarify our use of search terms (lines 131-134): “on PubMed we searched for the words “incubation”, “period”, and the virus name, on Google Scholar we searched for the phrases “incubation period of [virus name]” and “incubation period for [virus name]”, and on ISI Web of Knowledge we searched for “incubation period” and the virus name.”
To acknowledge the limitations our search terms placed on article return, the following was added to the discussion (lines 361-363):

“This review was limited by our inclusion only of published data and by our search terms. Due to our inclusion of some form of “incubation period” in searching for articles, the entirety of the literature on each virus was not reviewed. Our estimates for astrovirus and rotavirus are each based on three studies and fewer than 20 observations. Due to difficulties in studying these diseases experimentally, careful observational studies are needed to provide more evidence to support the incubation period and its distribution. “

In addition to the outbreak studies reviewed by Matthews are the challenge (i.e. volunteer) studies, mainly for norovirus. There about ~15 of these that I am aware of’ only 3 were included in this review.

Eleven challenge studies for norovirus were included in this review, of which eight did not contain data suitable for abstraction. Four additional articles published within our search period were identified and reviewed, three of which would have also been excluded. The fourth article contained abstractable data (Atmar 2008), but not a sufficient quantity to affect our final estimates.

The eleven challenge studies identified in using our search terms were as follows:


The four additional studies identified were the following:

Minor comments
Introduction
Line 83: The characterization of astrovirus being the second most common cause of AGE in children is outdated. Norovirus is the second most common cause in most places, and perhaps the most common in places with rotavirus vaccine use. See for example Payne et al, NEJM 2013 or Amar Eur J Clin Micro 2007. The characterization of astrovirus was changed from “the second most common cause” to “an important cause” (line 84).
Ln 87 + elsewhere: May also be worthwhile to distinguish the human caliciviruses into noro and sapoviruses. To distinguish noroviruses from sapoviruses, “in particular noroviruses” was added to line 87.
Ln 97: In the absence of laboratory diagnostics, ‘Kaplan’s criteria’ are frequently used to ascribe viral etiology in outbreaks. The incubation period is one key feature of these criteria. Kaplan et al AJPH, 1982 We agree that this is a very helpful addition to the introduction, “Kaplan’s criteria were developed and are frequently employed to determine whether an outbreak was caused by norovirus; the incubation period is one of the key elements of these criteria.” was added (line 97).
Methods
Ln 153: Why consider breakfast between 0h and 10h; or dinner between 14h and 0h? Would it I am not clear how the individual incubation period was determined using a wide (and variable window of possible exposure. The sentence starting “We report the range of incubation...” I think it meant to clarify this, but I did not follow.
To clarify why we determined standardized exposure intervals for meals the following was added to the methods section (line 155):

“Because a large number of foodborne outbreaks described in the literature did not report exact meal times, we established standard exposure intervals that were used in abstracting individual-level data for studies in which mealtimes were reported as just ‘breakfast’, ‘lunch’, or ‘dinner’”

Ln 165: For each disease? Does this mean for each pathogen? [they all cause the same, or similar disease: gastroenteritis]

Thank you for pointing out this oversight, we replaced disease with pathogen (line 170).

Ln 166: clarify the meaning of “doubly interval centered observations”

To clarify, the following was added (line 170):

“For each pathogen all observations were pooled together to form a single set of doubly interval censored observations; each data point contained a range of possible exposure times, for example “dinner”, and a range of possible times of symptom onset. Because times of exposure and symptom onset are rarely reported exactly, the minimum time frame we considered was a one hour range. If the time of symptom onset was reported to be 5:00PM, we recorded the time of symptom onset to between 5:00PM and 6:00PM. ”

Ln 169: mean each norovirus genogroup – calicivirus includes sapovirus. In general, better to say human caliciviruses (HuCV) as the caliciviruses are a diverse group, mainly affecting animals.

The following edit was made beginning in line 178: “Pooled data for each norovirus genogroup were analyzed individually, data from genogroups I and II were analyzed together, and finally all human calicivirus data (both norovirus genogroups and sapoviruses) was pooled and analyzed.” Additionally, throughout the document when caliciviruses were mentioned the sentence was edited to read human caliciviruses.”

Results
Ln 192, 214, 234, 291: this is background material. Should it be included in the results section? It gives the impression that these citations are from the lit review. In fact, the citations for these sections come from the same textbooks that are criticized for making statements that are not well-grounded in data on the incubation period. I think the authors would find the same of the other characteristics (e.g. frequency of vomiting for specific pathogens), so it may make sense to boil down these clinical description to aspect strictly relevant to estimating incubation and to move them to the introduction.
While we recognize that some of this information is not strictly results, we think it is highly useful to have this information tied with the disease sections of the results section. In addition to reading better overall, we believe this allows these sections to be read independently by those with a particular interest in a certain virus. For this reason we have kept the original structure.

Ln 251: Again, 15 observations from 3 experimental studies were included, but there are many more. Most are for GI noroviruses.

Please see the response to the second major comment.

Discussion
Ln 334: There are relatively few rota and astro outbreaks in healthcare facilities, compared with norovirus, for which there are hundreds, or even thousand per year, just in the US.

This was changed to just say “outbreaks” rather than “healthcare associated outbreaks” (line 348). Our objective was to indicate that there were many outbreaks described in the literature that we were unable to abstract data from, thus explaining the low numbers of observations we used to determine the incubation period distributions for astrovirus and rotavirus.

May be worth adding a comment on the serial interval, of which the incubation period makes one of the two components. While the incubation period may be a biological entity and may not differ greatly between settings, the serial interval may. It is hard to distinguish the two, when the timing of exposure cannot be defined (e.g. health care facility outbreaks).

We have added the following statement to highlight this relationship and it’s important for infectious disease dynamics (Line 375):
“Furthermore, the incubation period is an important component of the serial interval (difference in symptom onset times in a case and those that case infects), which is one of the fundamental determinants of how quickly epidemics spread in a population.”

Ln 359. It not exactly clear why knowledge of the incubation period is crucial for understanding vaccine failure, etc. Also, why will vaccine efficacy improve?

We agree with the reviewer that this is somewhat of a stretch. We have removed this statement from the revised manuscript.

Level of interest: An article of outstanding merit and interest in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests:
I declare that I have no competing interests