[Cover Letter]

|  |  |
| --- | --- |
| Reviewer 1, General comment 1 | |
| I can be brief. This interesting manuscript from the Chinese Center for Disease Control and Prevention applies surveillance data from 539 HFMD outbreaks occurring between 2011 and 2018 in China to estimate the basic reproduction number (R0) for three enterovirus serotypes associated with this disease. The data suggest that all 3 viruses are highly communicable, that are relatively few differences in R0 among them, and that immunization in the case of E71 can reduce transmission rates. In making these calculations, the authors include useful observations on secular trends for each serotype and the influence of transmission within schools in facilitating HFMD outbreaks.  Overall, the paper benefits from the large number of outbreaks analyzed, the methodology and conclusions appear sound, and the paper is concise and well-written.  The observation that 89% of outbreaks occurred in children attending kindergarten, presumably the first year that children attend school in China, is most interesting and suggests that school attendance is the primary driver for HFMD outbreaks. This may be well known in China, but not more broadly and could be worth another report. | |
| Response | We thank the reviewer for their time and consideration. We appreciate the point raised by the reviewer and agree it is of value to point out the fact that most of the outbreaks we observed are among children attending kindergarten. We added some context to this observation. |
| Revisions | * Page 6, line 98-100, we revised the following sentence:   + In mainland China, HFMD outbreaks primarily occur in relatively closed units such as kindergartens (typically for 3 to 6 year olds) and primary schools(typically for 6 to 12 year olds) [9-10]. * Page 12, starting line 248, we added the following discussion point:   + The overwhelming majority of the HFMD outbreaks we identified (88.9%) were inkindergartens, implying that the impacts were predominantly on those between 3and 6 years of age. The HFMD Prevention and Control Handbook includes specificindividual and environmental measures designed to protect children within thisage group, including the promotion of hand hygiene and daily sanitation of toysand classrooms. Additionally, the handbook also suggests class suspension when 2 or more cases are reported within that class and within a week; suspensionof the entire childcare facility when more than 10 cases are reported within the facility or when more than 3 class units have met the class suspensioncriterion. |

|  |  |
| --- | --- |
| Reviewer #2 | |
| The authors estimate the basic reproduction number of the 3 main enterovirus serotypes causing HFMD (EV-A71, CVA16 and CVA6). They use data from outbreaks in China between 2011 and 2018, for which the serotype has been identified. The question and data are interesting and the results will contribute to the scarce literature on it. However, I have important concerns regarding the method used and the interpretation of the results that must be addressed. | |
| Response | We are grateful for the time and consideration. |

Major comments:

|  |  |
| --- | --- |
| Major comment #1 | |
| One of my main concerns is that the authors use a method that was developed for SARS (which was a new emergent virus) and assumes that close to 100% of the population is susceptible. However, the 3 viruses considered here are endemic in China, and therefore a proportion of children may have already been infected (and therefore, are probably immune). This proportion sharply increases with age (after maternal antibodies decline in approximately 6-12 months). The authors must address this issue. Considering detailed seroprevalence data across ages, and stratifying the estimates by the location of the outbreak (kindergartens, primary schools, etc.) as they affect different ages may help providing more robust results. The authors should perhaps also consider estimating the effective reproduction number as well, as a comparison. The are well established methods for that, for example the R-package Epiestim. | |
| Response | Thank you for this comment. We agree that the that the location of the outbreaks can be an appropriate proxy for the age of the underlying population and that estimate stratification could be meaningful. Following the suggestions, we first went back to the raw data, and assigned an age to each outbreak based on the description of the outbreak setting (e.g, first year kindergardeners are likely 3 year olds). We obtained age- and subtype-specific sero-prevalence from existing literature and calculated adjusted the raw R0 estimates under two scenarios, optimistic (using the lower limit for seroprevalence of CV-A16) and conservative (using the upper limit for seroprevalence of CV-A16). |
| Revisions | * Method is updated (page 9 middle paragraph). * Figure 4 is updated. The following revision has been made to the caption:   + Adjusted R0 estimates are based on seroprevalence data among Chinese children identified through existing literature. The seroprevalence of EV-A71 is based on the estimated mean in a meta-analysis; the seroprevalence of CV-A6 is based on the estimated mean in the only study identified. We identified several studies on the seroprevalence of CV-A16, yet no meta-analysis – so we include both the conservative (i.e., the highest mean seroprevalence found) and the optimistic (i.e., the lowest mean seroprevalence found) scenarios. * Age- and subtype specific histograms of R0 estimates are included in the technical appendix, Figure 1. |

|  |  |
| --- | --- |
| Major comment #2 | |
| My other main concern is that the method also assumes that there is no intervention that prevents the disease from spreading, however, I don't think this is the case for HFMD. It is not clear for a given outbreak, when do the authors stop including cases to the analysis. Is it when some control measures are put in place or do they use all the reported cases of the outbreak? Whichever the case, there is some right-censoring of the data that will bias the estimates of R0. The authors must explain how they address this issue. To address my main comments 1 and 2, I would encourage the authors to test the method on simulated data to show that the method works well and identify possible biases. | |
| Response | Thank you for pointing this out. We agreed our current description of the initial growth period may not be clear. We also completely agree that right censoring may occur, as do most studies that focus on exposures that may take time to show effects and that have to draw a line to assess the relevant statistics. To address these concerns, we first clarify the defintions of IGP. The end of the IGP is assumed to be the epidemic peak or the time of intervention, whichever comes first. Thus, outbreaks under consideration are all unmitigated.  Then, developed a stochastic SEIR model, tested a wide range of R0, simulate epicurves, and use the method in this study to estimate the R0s. This way, we approximate the bias caused by right-censoring. We also added discussion on this potential bais. |
| Revisions | * Page 7, top paragraph. * Page 16, middle paragraph. |

|  |  |
| --- | --- |
| Major comment #3 | |
| The authors must clarify what does it mean that the initial growth phase (IGP) is visually identifiable in quantitative terms. What are the criteria used? This reduces the number of outbreaks analysed from 1653 to 539, so it is an important selection criteria. The authors should also provide information on how their model fits the data. | |
| Response | 可以从流行曲线看出IGP。IGP要满足以下条件：有明确的首发病例，相邻代际病例发病间距介于2-10天，IGP是在干预措施采取之前首发病例至发病高峰之间的间隔。 |
| Revisions | * Page 8, middle section |

|  |  |
| --- | --- |
| Major comment #4 | |
| The authors should present a sensitivity analysis of the results to the assumption on the duration of the incubation period. | |
| Response | 目前关于手足口病的潜伏期尚无完全统一的结论，一般介于2-10天，平均3-7天，本研究取其中位值5天。如果需要的话可以尝试不同的潜伏期，但可以直接明确的是，在其他参数不变的情况下，假定的潜伏期越短，得出的再生数R0越小。 |
| Revisions |  |

|  |  |
| --- | --- |
| Major comment #5 | |
| The decrease in estimated R0 for EV-A71 in 2017 is unlikely to be a cause of the effect of the vaccine. EV-A71 vaccine coverage in China has been very low - mainly because the vaccine is not part of the routine immunisation programme and is only available through the private sector. In addition, how would this explanation be compatible with the estimated increase in 2018? The observed changes in R0 over time are likely to be due to the cyclical patterns observed for many enterovirus serotypes, i.e. a natural cycle in their dynamics resulting from the accumulation of serotype-specific immunity at the population level. These results must be presented in a completely different way, and abstract and discussion amended accordingly. In addition, I would encourage the authors to provide data on vaccine coverage for 2016 and 2017 in order to discuss the possible causality of the introduction of the vaccine. | |
| Response | 2017-2018年的再生数下降是我们观察到的一个现象，我们推测可能与疫苗有关，当然因果关系的确定还需要更多的证据。 |
| Revisions |  |

|  |  |
| --- | --- |
| Major comment #6 | |
| The data used must be better described:  a.The authors say that they use data from 2 surveillance systems, however they only mention PHEESS for the data collection - why? Please explain what is the difference between the two systems. Can a case be notified to both - i.e. is there some overlap?  b.Do both systems contain pathogen-specific data for each reported case? If not, which patients do have samples taken and typed?  c.What does structured and non-structured narratives mean? Systematically collected?  d.What is the epidemic curve exactly in this context: daily incidence of reported HFMD cases? Is it by day of symptom onset, day of sample taken or day of reporting?  e.How are the exposed and infected defined?  f.What does it mean that an outbreak is laboratory-confirmed? | |
| Response | We appreciate the suggestions from the reviewer that can help us include the clarity of our data description. We provide itemized response and revisions:   1. We used the PHEESS system in this study, not the NNIDSS. While NNIDSS has a lot more detailed information about specific patients, it does not have information about specific outbreaks. In other words, all reported in PHEESS were also reported in NNIDSS but not all NNIDSS can be found in PHEESS. Not all cases have confirmed infector/ cluster.   Additionally, HFMD is an infectious disease that disproportionally affect young children in China – which means the population is not equally vulnerable. Thus, calculating reproduction numbers in the context of outbreaks are more appropriate. This is consistent with XXXX and XXXX.   1. Both system requires patients to be subtyped. 2. There are structured data, and there are unstructured narratives. Structured data are existing items in the database – local CDC officials simply fill this online electronic form out. This database include items such as the timing of first and last cases.   Un-structure narrative is in the format of free text attached to the PHEESS reports (generated using the database we just described), and is designed to capture all information that do not fit into the electronic form. This could include pictures of epicurve and a detailed descriptives of the location where the outbreak occurred. There’s no standard template, which makes it challenging to consolidate.   1. Epicurves are all based on symptoms onset. 2. Infected individuals are those who are symptomatic and are confirmed based on lab results. Exposed individuals are all the close cantacts of infected individuals identified via outbreak investigation (as reported in PHEESS). 3. A lab-confirmed outbreak is an outbreak where all infected cases identified can be attributed to the same subtype. We did not include outbreaks in which two or more subtypes. To clarify these points, we revised the definition of inclusion criteria. |
| Revisions | 1. Page 7, surveillance system 2. Need to confirm 3. Revisions: bottom of page 7, top of page 8. 4. Top of page 8 5. Removed a sentence from the top of page 7 + added a paragraph in the middle of page 8. |

Minor comments:

|  |  |
| --- | --- |
| Minor comment #1 | |
| Figure captions need more detail:  a.Fig. 2. How is an outbreak attributed to a given month?  b.Fig. 4. How is an outbreak attributed to a given year? What is the vertical dashed line?  c.Fig. 5. What are the lines? What are the points? Why there are no lines in the first panel? | |
| Response | We thank the reviewers for raising these points. Below we provide itemized responses and reivisons:   1. An outbreak is attributed to a given month based on the onset time of the first case. 2. An outbreak is attributed to a given year based on the onset time of the first case. The dashed line is intend to indicate vaccine licensure time. However, we discovered a bug in the code that put this dashed line a year later than when it was supposed to be. This bug has since been fixed.   c.点代表539起暴发疫情以及它们对应的R0，线为趋势线。Why there are no lines in the first panel? |
| Revisions | 1. the Caption of figure 2 has been revised: 2. The Caption of figure 4 has been revised:   Additionally, figure 4 has been updated. |

|  |  |
| --- | --- |
| Minor comment #2 | |
| Table 1. Please explain how is the outbreak duration calculated. | |
| Response | 疫情持续时间指首末例发病时间间隔。 |
| Revisions | * Table 1, caption revised. |

|  |  |
| --- | --- |
| Minor comment #3 | |
| The authors must specify what are the values of R0 reported in the results (e.g. 5.46 and 5.11 reported in page 11). Are these the average of all the estimates for each outbreak? It would be good to report the uncertainty around these values (e.g. range or IQR). | |
| Response | These are the median R0s of all the estimates for each outbreak. |
| Revisions |  |

|  |  |
| --- | --- |
| Minor comment #4 | |
| Does the model assume that there are no asymptomatic infections (in other words, all infections are observed as cases of HFMD)? Please amend the methods section accordingly. | |
| Response | 无症状感染者是无法通过观察被发现的。一般来讲，基本再生数的估算是基于有症状病例进行的，手足口病基本再生数的估算也是如此。为了保持一致性和可比性，故本研究也是基于有症状病例进行估算。 |
| Revisions | * Page 8, top paragraph, last sentence. * Page 16, last paragraph, last sentence. |

ba