

Peer Review File

Manuscript Title: Climate change drives novel cross-species viral transmission

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

This is an ambitious and potentially useful project. I would like to understand it better.

The main text does not do a good job of outlining the work. The authors need to clearly define terms like "viral sharing" and "spillover" and outline the broad strokes of what they did (distribution models with and without dispersal, a GAMM to predict sharing ...) and what assumptions they made (no adaptation to new environments, no constraints due to ecological interactions, sharing predicted by proportional range overlap and phylogenetic distance with multi-member REs ...)

In addition to saying more in general terms about methods in their main section, the authors could say more specifically about methods in their Methods. I am particularly concerned with dichotomization vs.

the propagation of uncertainty. I was also confused how the authors arrived at numbers of shared viruses, since their main GAMM seems to be about the probability of viruses being shared.

DETAILS

The authors could choose a friendlier word than "sympatric" for a broad-interest journal.

The text from L50-90 should be integrated better and sharpened. The "fundamental rules" come as something of a surprise, and their links to the concerns outlined in the previous ¶ are not immediately made clear. The s. at L80 is weird. What could it mean to treat dispersal as a layer of realism?

I guess I'm supposed to assume that the X.Y in RCPX.Y refers to degrees of warming by a certain time?, but the authors don't say so, and the C stands for "concentration", so who knows? The authors also don't promptly explain their "paired" scenarios.

The authors need to define "sharing events" promptly. By L125 they are "staggering" concerns for conservation, but I still don't know what they are (nor how they relate to "spillover").

Is there a reason to use "eutherian" instead of the less-technical "placental"? Aren't they equivalent except for hair-splitting about extinct animals? The authors might also mention somewhere in the main text that they are actually focused on this subgroup.

"The original study" is insufficiently clear on L132

The qualification "without known reservoirs" on L172 does not really apply to the main logic of this s. – that host jumps complicate tracing. If there is a separate point to be made about the ZEBOV mystery, it is not coming through and the authors should make it separately.

The antecedent on L194 should be cited and less-speculative: "Although mitigation is likely to minimize extinctions and impacts [citations], our results..."

"does not imply" L209 is a bit too strong, technically speaking.

"should not be interpreted as" should cover it. The authors should probably leave "disease emergence" out of the following line as well.

I don't really understand the claim at L214.

"no-analog" is a weird piece of jargon that the authors don't need.

"most new exposures will be among mammal species" apparently really means "most mammalian exposures will be among non-human mammals".

L261. The authors "removed spatial and environmental outliers" from the occurrence records. No details of this process are discussed nor justification provided, beyond a cite to a generic methods paper from 1950.

It is not completely clear how the authors chose suites of predictors with correlation < 0.7 (presumably, absolute correlation). Nor is it clear why. Excluding correlated predictors has a long but controversial history. I've never heard a convincing justification for it in the context of regularized, cross-validated prediction exercise and don't believe the authors should probably be doing it.

"folds were assigned based on spatial clusters" needs to be elaborated.

I was not able to understand what was meant by the fifth percentile threshold, but I imagine this should be justified as well.

The acronym PPM is used only once (and not defined).

The github repo pointed to has no obvious documentation or other clues how it might be navigated, or how it links to the work done in the paper.

Referee #2 (Remarks to the Author):

The paper is on an important and topical subject: the extent to which climate change (and land use

changes) may increase the risk of emerging zoonotic diseases, via novel species interactions. The authors use a statistical model based on linking climate data to current species extent data and making projections to future scenarios. We have some reservations over how the motivation and then results are presented. Detailed comments below:

Main comments

1) The focus on only mammalian viruses warrants a brief justification in the context of considering how "Climate change will drive novel cross-species viral transmission" – the potential role of non-mammalian reservoirs should be at least mentioned, e.g., many viruses can have complex paths of circulation through mammalian and non-mammalian hosts (flaviviruses of zoonotic concern are well documented in reptiles: *Emerg Infect Dis.* 2004 Dec;10(12):2150-5); and influenza in birds, etc.

2) We think the focus on RCP2.6 is misleading. There are four representative concentration pathways and at best they should be interpreted with equal probability of occurring, and at worst one of either RCP4.5/RCP6 should be chosen as the reference case given these are the medium stabilization scenarios [Van Vuuren 2011]. Even RCP8.5 has more frequently been used as a "business-as-usual" scenario (although we recognize there has been some push back against this in recent years). We commend the authors for explicitly stating in lines 200-203 that their results are strongest for RCP2.6. and milder for other scenarios. However, we think selectively showing the strongest results, e.g. Fig 1 and stated numeric changes, in the main text still seems somewhat biased, given there is no justification for presenting only RCP2.6. One option would be to include ED3 in the main text, to make the range of outcomes clear. This should be done in addition to amending single stated numbers, such as in line 107, to a range of values and updating map plots, at least Figure 1, to show the upper and lower range of outcomes.

3) We think the authors could better present the baseline scenario. For instance, in line 94, 300,000 novel species interactions compared to what, over which time period? Similar issues occur in several places e.g. lines 107, 122. We think it would be helpful for the reader to understand what the baseline looks like, how many species interactions take place in the current climate with current land use? 300,000 sounds large, but what is the context? Baselines could also be better expressed in the methods, what year is the baseline climate? Is the baseline land use map from the same period?

4) ED3 and lines 200-203. We think the authors need to state their main statistics/results more clearly. In lines 201, 75% and 28% seem to be being compared, yet the former is described as the range that species gain (presumably a change in area?) and the other is described as the percent of species that experience an expansion (a change in number of animals). Reading ED3 doesn't make this any clearer – the figure caption needs to be updated to state clearly what these numbers mean (and again, compared to what baseline?).

5) Following on from this, are there species that see a decline in range? How are they factored in? If in RCP8.5, 28% of species see an increase in range, does that mean 72% see a decline? I think some more transparency on whether there are species that have seen a decline in interactions should be included in the supplement. As viruses are continually evolving, a decline in existing interactions is scientifically important.

6) Line 364. Why HadGEM2? We believe WorldClim provides output from the full suite of CMIP5 climate models. It is standard to either show the full range of model output, or use a multi-model mean. If the authors are not using WorldClim for the climate projection data (which needs to be clearly stated in the methods), then was the model output bias-corrected?

7) Line 424, section "Explaining spatial patterns". What is the outcome variable here? Are we correct that the authors are here running a regression on the predicted results from another regression?

8) Finally, the introduction might benefit from re-framing to include discussion of some relevant complexities in projecting range shifts. For example (i) communities that are displaced in concert may move to a new geographic area but without necessarily forming novel assemblages (which is mentioned later on in lines 101-102) (ii) the ability to shift or expand ranges for many species is limited by lack of access (e.g., if no remaining forest corridors exist), (iii) it is difficult to predict when in-situ adaptation to changing climate regimes will allow local persistence for a host species, (iv) vector and host communities may displace at different scales and interrupt transmission. Invasive and introduced species seem to be worthy of mention when describing ongoing "global reassortment of animal biodiversity" Discussing all would be beyond the scope of the manuscript but some mention of these complexities may be beneficial

Minor comments

Abstract: Is the observation that most projected viral sharing is driven by rodents and bats surprising given that together bats and rodents contribute over 60% of mammalian species?

Line 59 "Some viruses may spread environmentally between spatially-proximate species with no direct behavioral contact" - mention of transmission by vectors seems warranted as well here

Line 111 "If species survive a changing climate by aggregating in high elevation refugia, this suggests emerging viruses may be an increasing problem for their conservation" - again this includes an assumption that populations would have access to higher elevation refugia. Highly fragmented, disjoint lowland forests or other primary habitats often have no remaining suitable habitat connecting them to higher elevation refugia through which species could pass.

Line 119, "SDM" needs to be defined.

Figure 4: We don't find this color scheme particularly intuitive, and probably not the best for color-blind readers either?

A cited "Data" section in the methods would be helpful.

References

Van Vuuren, Detlef P., et al. "The representative concentration pathways: an overview." *Climatic change* 109.1-2 (2011): 5.

Referee #3 (Remarks to the Author):

Review of "Climate Change will drive novel cross-species viral transmission" by Carlson et al.

Overview

In this section I explicitly address the criteria requested of me (*italics*); I then include some general comments before addressing specific major/minor comments.

Summary of the key results: The authors find thousands of opportunities for novel virus sharing between mammals if the climate continues to warm over this century. Interestingly, RCP2.6 (the most limited warming scenario), may produce the largest increases in novel sharing.

Originality and significance: This is not my field, but the authors argue convincingly for the novelty.

Data and methodology: As a non-specialist in that field, I cannot speak for the statistical modelling, but as outlined in my major comments, the treatment of the climate data is not detailed enough for the work to be reproducible.

Appropriate use of statistics and uncertainties: See above cf. statistical modelling; and to follow from the point on climate modelling – uncertainty stemming from the choice of climate model (see major comments) is not acknowledged.

Conclusions: They seem novel (but recall I am not a specialist in the field), but due to the treatment of the climate data, I am sceptical that they are robust (major comments).

Suggested improvements: See major comments. Those issues need to be addressed.

References: Seem fine, but the normal caveat about being a non-specialist also applies here.

Clarity and context: I found the writing to be clear throughout, but there are too many normative statements (see minor comments). These would be very easy to remove during a revision. I would also expect the authors to add some more context about the Covid-19 outbreak in the introduction – if given the chance to submit a revised paper.

General Comments

As a climate scientist reviewing this during the Covid-19 outbreak, I found the topic fascinating, and the paper was generally well written and accessible to a non-specialist in virology/ecology. However, the "climate" element of the paper is communicated in insufficient detail for me to recommend

publication. From the information included, I cannot ascertain whether just the level of detail needs to be improved in the text, or whether the study is flawed in that its conclusions are unlikely to be robust. I hope the authors can clarify and respond to my concerns – outlined in the major comments section below.

Major comments

Detail of the climate change projections and their processing is entirely absent. When inferring the response of a predictand to future climate change simulated by climate models, it is necessary to: (1) establish how the modelled climate for the present diverges from historical reality; (2) to assess how sensitive the results are to the model used. I explain these below.

(1) The changes you see projected in the future could be due to biases in the model's ability to resolve the current climate. For example, the drying you see projected in the future at some point relative to your observations could be because the model already simulates a climate that is too dry – not because the climate is expected to change in that way. There is an enormous body of work dedicated to dealing with these problems, and it may be that you are applying appropriate methods from that literature, but I see no mention of this in the paper. I stress that this is particularly important here because your most impactful scenario of RCP2.6 has relatively little climate change, so biases in the present climate may have a heavy weight on your results. To be clear: we need details of how the climate data were processed (including how they were interpolated(?) to the 25 km grids).

2) There are well over 20 climate models forced with the same data to simulate possible future climates. The different models do project different climates, and this reflects irreducible uncertainty at this stage. It is therefore good practice to sample across the range of models to see how sensitive your results are to the choice of model. It is indeed common for climate change studies to use all available models, or at least a double-digit ensemble. The choice of HadGEM-ES as the only model here is particularly unfortunate because it has one of the highest climate sensitivities of all the models (it produces one of the largest increases in global mean air temperature for a doubling of CO₂ – the so-called “equilibrium climate sensitivity”: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter09_FINAL.pdf -- see p. 818). We need to see how sensitive the results are to climate modelling uncertainty.

Minor comments

L27: what do you mean by “sharing novel viruses between 3,000 and 13,000 times”? This metric (sharing) does not yet mean anything to the reader.

L122: replace “mildest” with proper terminology – suggest “least radiative forcing” or “climate most similar to recent decades”.

L124: avoid hyperbole – please remove “still-staggering”

L144: hyperbole – avoid “staggering” (I’m not being pedantic, but scientific writing should be free of normative statements)

L159-160: Without the background, I don't follow your logic here: where is the wildlife-wildlife host jump with simian immunodeficiency? I think the sentence here needs re-writing to articulate more clearly where the wildlife-wildlife host jumps are.

L209: Same issue with hyperbole: remove "catastrophic" and replace with the amount of global warming you're referring to (more than 2C? More than 3C?). Please check for, and remove, all instances of hyperbole throughout the text.

L371: The W/m^2 does not refer to solar radiation, but radiative forcing in general.

Response to reviewers

Carlson *et al.*, submission to *Nature*

June 2020

Response to Editor

We are very grateful to the three referees and the editor for thoughtful suggestions and the chance to revise the manuscript. We have spent the last few months following up on the detailed suggestions of the referees, and substantially expanded our methodology to address them.

The most significant change in the study is a shift from using a single CMIP5-generation climate model to a total of nine general circulation models for CMIP6 (all of the models currently available on WorldClim). In the total pairwise design with the RCP-SSP scenarios and dispersal assumptions, this produces a total of 136 future scenarios, a full quantification of the uncertainty propagated in our modeling study.

In addition to a clearer and more transparent presentation of uncertainty, we've worked to follow the reviewer comments and more clearly present the terminology used in the study, the limitations of the work, and the finer details of the methodology. That includes an entirely rewritten Data section that starts the Methods, making the attribution of particular datasets much easier for a reader. We believe this has substantially improved the study we are returning today, and are grateful for these suggestions!

Referee 1

This is an ambitious and potentially useful project. I would like to understand it better.

The main text does not do a good job of outlining the work. The authors need to clearly define terms like “viral sharing” and “spillover” and outline the broad strokes of what they did (distribution models with and without dispersal, a GAMM to predict sharing ...) and what assumptions they made (no adaptation to new environments, no constraints due to ecological interactions, sharing predicted by proportional range overlap and phylogenetic distance with multi-member REs ...)

We appreciate this suggestion, and have worked hard to clarify these terms early on, as the reviewers agreed they needed a straightforward definition. We have substantially rewritten the Introduction of the paper in service of that goal, particularly focusing on early definitions of our methodology:

“This poses a measurable threat to global health, particularly given several recent epidemics and pandemics of viruses that originate in wildlife (zoonotic viruses, or zoonoses). Most frameworks for predicting cross-species transmission therefore focus on the steps that allow animal pathogens to make the leap to human hosts (a process called spillover). However, zoonotic viruses are a small fraction of total viral diversity, and viral evolution is an undirected process, in which humans are only one of over 5,000 mammal hosts with over 12 million possible pairwise combinations....

“Because only a few species are common worldwide, most hosts have no *opportunity* to exchange pathogens: of all possible pairs of mammal species, only ~7% share any geographic range, and only ~6% are currently known to host one or more of the same virus species (hereafter ‘viral sharing’). As host geographic ranges shift, some interactions will become possible for the first time, and a subset will lead to viral establishment in a previously-inaccessible host (novel viral sharing).”

In addition to saying more in general terms about methods in their main section, the authors could say more specifically about methods in their Methods. I am particularly concerned with dichotomization vs. the propagation of uncertainty. I was also confused how the authors arrived at *numbers* of shared viruses, since their main GAMM seems to be about the *probability* of viruses being shared.

We appreciate the suggestion to expand the Methods and have made an effort both to better detail the process in the Methods proper, and to explain the Methods early on in the paper. For example, with regard to the probability vs. number issue, we have added two passages early in the paper that we believe clarify this:

“We then used a recently-developed model to predict the probability of a novel viral sharing event—here defined as the cross-species transmission of at least one virus, between a pair of hosts that are newly in contact—based on novel geographic overlap and host phylogenetic similarity, a first order approximation of opportunity and compatibility....Based on the total number and distribution of first encounters, we used cumulative viral sharing probabilities to estimate the total number of novel sharing events that are expected (each of which describes the cross-species transmission of at least one virus). ”

“Accounting for geographic opportunity and phylogenetic compatibility, we projected that a total of 316,426 (+/- 1,719.0) first encounters in RCP 2.6 would lead to 15,311 novel sharing events (+/- 139.6)—that is, a minimum of at least ~15,000 cross-species transmission events of at least one novel virus (but potentially many more) between a pair of naive host species.”

Throughout we have focused on making the methods similarly more transparent and clear, particularly by using multiple GCMs to quantify model uncertainty, and by presenting both extreme RCPs as bookends in-text to quantify scenario uncertainty.

Details

The authors could choose a friendlier word than “sympatric” for a broad-interest journal.

We have changed the first appearance from “even if species are nominally sympatric” to “even if species’ ranges nominally overlap”, and simply removed the second appearance: “...especially after a recent study implicating reduced Arctic sea ice in novel viral transmission between pinnipeds and sea otters.”

The text from L50-90 should be integrated better and sharpened. The “fundamental rules” come as something of a surprise, and their links to the concerns outlined in the previous ¶ are not immediately made clear.

We appreciate this feedback, and have taken care to rewrite the first four paragraphs of the study in much clearer language, that lays out the logic of our study much more linearly.

The s. at L80 is weird. What could it mean to treat dispersal as a layer of realism?

We have entirely rewritten this sentence, in the hopes it is now clearer:

“Because many species are unlikely to be biologically suited for rapid range shifts, and will therefore move slower than the local velocity of climate change, we constrained the speed of range shifts based on inferred allometric scaling of animal movement, and compared scenarios that assumed limited dispersal against “full dispersal” (no dispersal limitation).”

I guess I’m supposed to assume that the X.Y in RCPX.Y refers to degrees of warming by a certain time?, but the authors don’t say so, and the C stands for “concentration”, so who knows? The authors also don’t promptly explain their “paired” scenarios.

We have made the definition of RCPs that we included in the original draft much clearer in the methods, in a now-consolidated section of text:

“We considered four possible scenarios for the year 2070 each based on a pairing of the Representative Concentration Pathways (RCPs) and the Shared Socioeconomic Pathways (SSPs). RCP numbers (e.g., 2.6 or 4.5) represent Watts per square meter of additional radiative forcing due to the greenhouse gas effect by the end of the century, while SSPs describe alternate possible pathways of socioeconomic development and demographic change. As pairs, SSP-RCP scenarios describe alternative futures for global socioeconomic and environmental change. Not all SSP-RCP scenario combinations in the “scenario matrix” are realistically possible. For example, in the vast majority of integrative assessment models, decarbonization cannot be achieved fast enough in the SSP5 scenario to achieve RCP 2.6.

We used four SSP-RCP combinations: SSP1-RCP2.6, SSP2-RCP4.5, SSP3-RCP7.0, and SSP5-RCP8.5. We selected these four scenarios because they span a wide range of plausible global change futures, and

serve as the basis for climate model projections in the Scenario Model Intercomparison Project for the newest generation of global climate models (CMIP6). SSP1-RCP2.6 is a scenario with low population growth, strong greenhouse gas mitigation and land use change (especially an increase in global forest cover), which makes global warming likely less than 2°C above pre-industrial levels by 2100; SSP2-RCP4.5 has moderate land use change and greenhouse gas mitigation with global warming of around 2.5°C by 2100; SSP3-RCP7.0 has high population growth, substantial land use change (especially a decrease in global forest cover) and very weak greenhouse gas mitigation efforts with global warming of around 4°C by 2100; and SSP5-RCP8.5 is the highest warming scenario with less decrease in forest cover than SSP3 but more substantial increases in coal and other fossil fuel usage leading to more than 4°C warming by 2100.”

In the main text, we now explain this in brief detail as well:

“We built species distribution models (SDMs) for 3,870 placental mammal species, and projected geographic range shifts based on four paired scenarios for climate change (Representative Concentration Pathways, RCPs) and land use change (Shared Socioeconomic Pathways, SSPs) by 2070. These scenarios characterize alternative futures for global environmental change, from sustainable land use change and a high chance of keeping global warming under 2°C (SSP1-RCP2.6), to a high chance of 4°C warming, continued fossil fuel reliance, and rapid land degradation and change (SSP5-RCP8.5; see “Climate and land use futures” for a detailed explanation). We present results for SSP1-RCP 2.6 in the main text because this scenario is most in line with the goals of the Paris Agreement to keep global warming “well below” 2°C”

The authors need to define “sharing events” promptly. By L125 they are “staggering” concerns for conservation, but I still don’t know what they are (nor how they relate to “spillover”).

As we mention above, we define viral sharing and novel viral sharing events now in the second paragraph of the manuscript:

“Because only a few species are common worldwide, most hosts have no *opportunity* to exchange pathogens: of all possible pairs of mammal species, only ~7% share any geographic range, and only ~6% are currently known to host one or more of the same virus species (hereafter ‘viral sharing’). As host geographic ranges shift, some interactions will become possible for the first time, and a subset will lead to viral establishment in a previously-inaccessible host (novel viral sharing).”

Is there a reason to use “eutherian” instead of the less-technical “placental”? Aren’t they equivalent except for hair-splitting about extinct animals? The authors might also mention somewhere in the main text that they are actually focused on this subgroup.

The distinction is indeed trivial; we have switched to using “placental” in the text, and now identify this early on: “We built species distribution models for 3,870 placental mammal species...”

”The original study” is insufficiently clear on L132

This sentence has now been rephrased:

“This viral sharing hotspot is driven disproportionately by bats, because their dispersal was left unconstrained; we made this choice given their exclusion from previous research characterizing the dispersal capacity of range-shifting mammals....”

The qualification “without known reservoirs” on L172 does not really apply to the main logic of this s. – that host jumps complicate tracing. If there is a separate point to be made about the ZEBOV mystery, it is not coming through and the authors should make it separately.

We have added an explanation to this sentence which we hope clarifies this point:

“Moreover, for zoonoses like Zaire ebolavirus without known reservoirs, future host jumps—and therefore, the emergence of a larger pool of potential reservoirs covering a greater geographic area (e.g., potential introduction of Zaire ebolavirus to east African mammals)—would only complicate ongoing efforts to trace the sources of spillover and anticipate future emergence.”

The antecedent on L194 should be cited and less-speculative: “Although mitigation is likely to minimize extinctions and impacts [citations], our results. . .”

We have updated this sentence, and added five citations:

“Whereas most studies agree that climate change mitigation through reducing greenhouse gas emissions will prevent extinctions and minimize harmful ecosystem impacts...”

“does not imply” L209 is a bit too strong, technically speaking. “should not be interpreted as” should cover it.

We have made this change as reviewers have requested.

The authors should probably leave “disease emergence” out of the following line as well.

We are hesitant to remove this climate impact from the list (“...catastrophic warming, which will be accompanied by mass defaunation, devastating disease emergence, and unprecedented levels of human displacement and global instability.”), given relevance to subject matter and high confidence from IPCC reports that climate change will lead to the emergence of novel infectious diseases. We have added several citations to this statement to increase the clarity that we are referring to the work of other studies, which have shown these upcoming patterns with a high degree of certainty and societal relevance.

I don’t really understand the claim at L214.

We are uncertain how the reviewer would like to have this sentence altered, and would perhaps appreciate further guidance if it still requires clarification after revision, as that sentence is the core claim of the study:

“Our study establishes a macroecological link between climate change and cross-species viral transmission.”

”no-analog” is a weird piece of jargon that the authors don’t need.

We have removed that phrase from this sentence:

“These efforts are being undertaken during the greatest period of global ecological change recorded in human history, and in a practical sense, the rapid movement of species (and their virome) poses an unexpected challenge for virological research.”

”most new exposures will be among mammal species” apparently really means “most mammalian exposures will be among non-human mammals”.

We appreciate this comment, and for clarity have updated the sentence accordingly (using wildlife, instead of non-human mammals, as we also do not consider livestock in this study):

“While several studies have addressed how range shifts in zoonotic reservoirs might expose humans to novel viruses, few have considered the fact that most new exposures will be among wildlife species.”

L261. The authors “removed spatial and environmental outliers” from the occurrence records. No details of this process are discussed nor justification provided, beyond a cite to a generic methods paper from 1950.

We have added a more detailed explanation to the Methods:

“Spatial and environmental outliers were removed based on Grubb outlier tests. To implement the Grubb outlier tests we defined a distance matrix between each records and the centroid of all records (in both environmental or geographic space, respectively) and determined whether the record with the largest distance is an outlier with respect to all other distances, at a given statistical significance ($p = 1e - 3$, in order to exclude only extreme outliers). If an outlier was detected it was removed and the test was repeated until no additional outliers were detected.”

It is not completely clear how the authors chose suites of predictors with correlation < 0.7 (presumably, absolutely correlation). Nor is it clear why. Excluding correlated predictors has a long but controversial history. I’ve never heard a convincing justification for it in the context of regularized, cross-validated prediction exercise and don’t believe the authors should probably be doing it.

We chose the maximum number of predictors with global correlation < 0.7 . Broadly the 19 candidate Bioclim predictors clustered into 4 groups that respectively included temperature, precipitation, and their seasonalities, so we chose these as representative of their clusters. Additionally, precipitation in the

driest quarter met this correlation threshold. We chose to remove these correlated predictors in spite of using regularization to maintain interpretability of marginal response curves. While this was not a major focus of this study we wanted to maintain this option to explore species-specific patterns. Further, it is an approach that C.M. has used in a wide range of studies including range maps for all plants and vertebrates (see Maitner et al., “The BIEN R package: A Tool to Access the Botanical Information and Ecology Network (BIEN) Database.”). Because we consider multiple features for each predictor, we wanted to be sure that, e.g., from a cluster of 3 correlated variables, that linear, quadratic or product features were not chosen based on idiosyncrasies in each variable, but rather were associated with a single variable representative of the cluster. With this approach, marginal response curves/surfaces can be more readily interpreted. Furthermore, we note that if the excluded variables had high correlation with variables we retained, they are unlikely to contain additional information on species occurrence, and therefore would not affect the results of our study.

”folds were assigned based on spatial clusters” needs to be elaborated.

We have substantially expanded this methods section:

“To reduce the possibility of overfitting patterns due to spatial aggregation, we used spatially stratified cross validation. Folds were assigned by clustering records based on their coordinates and splitting the resulting dendrogram into 25 groups. These groups were then randomly assigned to five folds. (If species had fewer than 25 records, a smaller number of groups was used based on sample size, and these were split into five folds.) This approach flexibility accounts for variation in the spatial scale of of aggregation among species by using the cluster analysis. By splitting into 25 groups initially (rather than 5) we obtain better environmental coverage (at least on average) within a fold and minimize the need to extrapolate for withheld predictions.”

I was not able to understand what was meant by the fifth percentile threshold, but I imagine this should be justified as well.

We chose as a threshold the 5% quantile of predicted values at training presences, which corresponds to a 5% omission rate of the known presences. This is a common threshold used with PPMs/Maxent that balances predicting overly large ranges with marginal habitat included with overly conservative ranges that miss fringe populations. For example, see:

Phillips, S. J., and M. Dudik. (2008) Modeling of Species Distributions with Maxent: New Extensions and a Comprehensive Evaluation. 31 (2): 161–75.

The acronym PPM is used only once (and not defined).

We now use the full name for clarity: Poisson point process model.

“For 3,088 species with at least 10 unique presence records, Poisson point process models (PPM), a method closely related to maximum entropy species distribution models (MaxEnt), were fit using regularized downweighted Poisson regression....”

The github repo pointed to has no obvious documentation or other clues how it might be navigated, or how it links to the work done in the paper.

We have organized the repository much more thoroughly and added a README.md annotating what each file corresponds to in the analysis and how they should be navigated.

Referee 2

The paper is on an important and topical subject: the extent to which climate change (and land use changes) may increase the risk of emerging zoonotic diseases, via novel species interactions. The authors use a statistical model based on linking climate data to current species extent data and making projections to future scenarios. We have some reservations over how the motivation and then results are presented. Detailed comments below:

Main comments

1) The focus on only mammalian viruses warrants a brief justification in the context of considering how “Climate change will drive novel cross-species viral transmission” – the potential role of non-mammalian reservoirs should be at least mentioned, e.g., many viruses can have complex paths of circulation through mammalian and non-mammalian hosts (flaviviruses of zoonotic concern are well documented in reptiles: *Emerg Infect Dis.* 2004 Dec;10(12):2150-5); and influenza in birds, etc.

We appreciate this feedback, and have added this justification explicitly. We now say:

“We focused on mammals because they have some of the most complete biodiversity data, the highest proportion of viral diversity described, and the greatest downstream relevance to human health and zoonotic disease emergence of any vertebrate class.”

We also now acknowledge the possible role of birds in emerging zoonoses (noting that alligators as amplifiers of West Nile Virus are more of an exception to the rule; very few zoonotic viruses have significant hosts in reptiles, amphibians, or fish):

“Birds have the best documented virome after mammals, and account for the majority of non-mammalian reservoirs of zoonotic viruses; changing bird migration patterns in a warming world may be especially important targets for prediction.”

2) We think the focus on RCP2.6 is misleading. There are four representative concentration pathways and at best they should be interpreted with equal probability of occurring, and at worst one of either RCP4.5/RCP6 should be chosen as the reference case given these are the medium stabilization scenarios [Van Vuuren 2011]. Even RCP8.5 has more frequently been used as a “business-as-usual” scenario (although we recognize there has been some push back against this in recent years). We commend the authors for explicitly stating in lines 200-203 that their results are strongest for RCP2.6. and milder for other scenarios. However, we think selectively showing the strongest results, e.g. Fig 1 and stated numeric changes, in the main text still seems somewhat biased, given there is no justification for presenting only RCP2.6. One option would be to include ED3 in the main text, to make the range of outcomes clear. This should be done in addition to amending single stated numbers, such as in line 107, to a range of values and updating map plots, at least Figure 1, to show the upper and lower range of outcomes.

We appreciate the reviewers’ concern on this point. We have gone through the entire manuscript and now present 2.6 and 8.5 as bookends capturing a fuller range of uncertainty than we would have from just presenting 4.5 or 6 as a reference case. In essentially all cases, that range of uncertainty is very narrow. For example:

“...we predict that the vast majority of mammal species will overlap with at least one unfamiliar species somewhere in their future range, regardless of emissions scenario (mean across GCMs +/- s.d. here and after; RCP 2.6: 98.6% +/- 0.2%; RCP 8.5: 96.6% +/- 0.8%). At the global level, community turnover would permit over 300,000 novel species interactions in every scenario (RCP 2.6: 316,426 +/- 1,719; RCP 8.5: 313,973 +/- 2,094.”

We note that the range in amount of viral sharing across emissions scenarios is similarly narrow, as highlighted in Extended Data Figure 3. For example, there is a 1% difference between RCP 2.6 and 8.5 in terms of viral sharing with limited dispersal, and a 1% difference without dispersal limits. As a result, we suggest that RCP selection is largely trivial in terms of transparency.

This allows us to keep the current framing in which we present RCP 2.6 (and associated land-use change) in the main text explicitly because of the global policy relevance of this scenario with regard to the Paris Agreement goal to keep global warming below 2 degrees Celsius. Multiple authors are involved in the IPCC process, and have made this decision with the explicit aim of presenting the information most relevant to that process and associated governance work. We are now more explicit about this decision in the text:

“These scenarios characterize alternative futures for global environmental change, from sustainable land use change and a high chance of keeping global warming under 2°C (SSP1-RCP2.6), to a high chance of 4°C warming, continued fossil fuel reliance, and rapid land degradation and change (SSP5-RCP8.5; see “Climate and land use futures” for a detailed explanation). We present results for SSP1-RCP 2.6 in

the main text because this scenario is most in line with the goals of the Paris Agreement to keep global warming “well below” 2°C.”

3) We think the authors could better present the baseline scenario. For instance, in line 94, 300,000 novel species interactions compared to what, over which time period? Similar issues occur in several places e.g. lines 107, 122. We think it would be helpful for the reader to understand what the baseline looks like, how many species interactions take place in the current climate with current land use? 300,000 sounds large, but what is the context?

We have tried to make this more explicit in two locations; first in the introduction:

“Because only a few species are common worldwide, most hosts have no *opportunity* to exchange pathogens: of all possible pairs of mammal species, only ~7% share any geographic range, and only ~6% are currently known to host one or more of the same virus species (hereafter ‘viral sharing’). As host geographic ranges shift, some interactions will become possible for the first time, and a subset will lead to viral establishment in a previously-inaccessible host (novel viral sharing).”

And second when we introduce the idea of first encounters:

“At the global level, geographic range shifts would permit over 300,000 first encounters in every scenario (RCP 2.6: $316,426 \pm 1,719$; RCP 8.5: $313,973 \pm 2,094$). Compared to a present-day baseline, in which we calculated 345,850 current pairwise overlaps among the 3,870 species (~ 7%), this essentially represents a doubling of potential species contact.”

Baselines could also be better expressed in the methods, what year is the baseline climate? Is the baseline land use map from the same period?

We have now clarified this point in much more explicit detail:

“The Bioclim variables for the historical climate are the mean from 1970-2000, and those for the future climate are the mean from 2060-2080. ’

“Historical land-use data for 2015 and projected land-use data for 2070 were obtained from the Land Use Harmonization 2 (LUH2) project at 0.25 degree spatial resolution.”

4) ED3 and lines 200-203. We think the authors need to state their main statistics/results more clearly. In lines 201, 75% and 28% seem to be being compared, yet the former is described as the range that species gain (presumably a change in area?) and the other is described as the percent of species that experience an expansion (a change in number of animals). Reading ED3 doesn’t make this any clearer – the figure caption needs to be updated to state clearly what these numbers mean (and again, compared to what baseline?).

We agree that these were unclear before, and have standardized the language used here, as well as switching the statistics from mean to median for a more appropriate metric:

“Accounting for dispersal limits, species experienced a median loss of 0.3% ($\pm 2.5\%$) of their range in RCP 2.6, with 49.8% ($\pm 3.8\%$) experiencing a net increase in range; in contrast, species experienced a 26.2% ($\pm 13.2\%$) median loss in RCP 8.5, and only 30.8% ($\pm 5.45\%$) gained any range. (ED Figure 3)”

5) Following on from this, are there species that see a decline in range? How are they factored in? If in RCP8.5, 28% of species see an increase in range, does that mean 72% see a decline? I think some more transparency on whether there are species that have seen a decline in interactions should be included in the supplement. As viruses are continually evolving, a decline in existing interactions is scientifically important.

A few species gain substantial range but the majority of species lose a small to medium amount of range. In response to this comment, we have decided to shift from using the mean to the median (see above comment):

“Accounting for dispersal limits, species experienced a median loss of 0.3% ($\pm 2.5\%$) of their range in RCP 2.6, with 49.8% ($\pm 3.8\%$) experiencing a net increase in range; in contrast, species experienced a 26.2% ($\pm 13.2\%$) median loss in RCP 8.5, and only 30.8% ($\pm 5.45\%$) gained any range. (ED Figure 3)”

We note that we don’t spend much more time unpacking what this loss of connections means for viruses, largely because we feel anything we could say here would essentially be speculation. Other work has

explored loss of connectivity in the context of parasite coextinction and disruption of complex life cycles, e.g., Pickles, R.S., Thornton, D., Feldman, R., Marques, A. and Murray, D.L., 2013. Predicting shifts in parasite distribution with climate change: a multitrophic level approach. *Global Change Biology*, 19(9), pp.2645-2654.

We hope to continue exploring these implications for parasites in a future study we've planned. But for viruses, it's difficult to say whether loss of overlap would have similar impacts, largely because coextinction of viruses has been rarely, if ever, documented. It is possible that once viruses are already shared among hosts, the loss of host overlap could lead to divergence and speciation over time (a vicariance event) or simply have no effect. As a result, we choose not to speculate too much about this at this time.

6) Line 364. Why HadGEM2? We believe WorldClim provides output from the full suite of CMIP5 climate models. It is standard to either show the full range of model output, or use a multi-model mean. If the authors are not using WorldClim for the climate projection data (which needs to be clearly stated in the methods), then was the model output bias-corrected?

We appreciate reviewer 2 and 3's comments about the need to better identify the contributions of model uncertainty and climate model selection to our manuscript. These have motivated the most substantial change in this iteration, which has been to replace the single CMIP5-generation HadGEM2-ES model with all nine currently-available CMIP6 GCMs available on WorldClim (which has its own internal calibration process for current climates). We then present multi-model averages for all in-text numbers (including standard deviations by GCM, to highlight the comparatively small model uncertainty) and figures. In doing so we have been able to highlight the fairly small contribution of climate model uncertainty to our final results.

7) Line 424, section "Explaining spatial patterns". What is the outcome variable here? Are we correct that the authors are here running a regression on the predicted results from another regression?

We appreciate the reviewer's concern. The outcome is first encounters, which are generated by novel pairwise overlap of the SDMs. We use linear models only to interpret and visualize patterns of our predicted outcomes, and contributor variables (elevation, population, or land use type) are not included in the species distribution models. To clarify the outcome variable and the methods being implemented, this section opens "we used linear models which predicted the number of first encounters (novel overlap of species pairs) at the 25km level ($N = 258,539$ grid cells)."

8) Finally, the introduction might benefit from re-framing to include discussion of some relevant complexities in projecting range shifts. For example (i) communities that are displaced in concert may move to a new geographic area but without necessarily forming novel assemblages (which is mentioned later on in lines 101-102) (ii) the ability to shift or expand ranges for many species is limited by lack of access (e.g., if no remaining forest corridors exist), (iii) it is difficult to predict when in-situ adaptation to changing climate regimes will allow local persistence for a host species, (iv) vector and host communities may displace at different scales and interrupt transmission. Invasive and introduced species seem to be worthy of mention when describing ongoing "global reassortment of animal biodiversity" Discussing all would be beyond the scope of the manuscript but some mention of these complexities may be beneficial

We appreciate these suggestions! We have removed reference to no-analog communities, which reduces the focus on novel assemblages (though we would suggest that the 300,000+ first encounters are by definition the real-time formation of novel assemblages). We have also added the remaining points to the discussion paragraph:

"Our study establishes a macroecological link between climate change and cross-species viral transmission. The patterns we describe are likely further complicated by uncertainties in the species distribution modeling process, including local adaptation or plasticity in response to changing climates, or lack of landscape connectivity preventing dispersal. The projections we make are also likely to be complicated by several ecological factors, including the temperature sensitivity of viral host jumps; potential independence of vector or non-mammal reservoir range shifts; the possibility that defaunation especially at low elevations might interact with disease prevalence through biodiversity dilution and amplification effects, not captured by our models; or temporal heterogeneity in exposure (hosts might exchange viruses in passing but not overlap by 2070, especially in warmer scenarios). "

Minor comments

Abstract: Is the observation that most projected viral sharing is driven by rodents and bats surprising given that together bats and rodents contribute over 60% of mammalian species?

No, not surprising at all! In fact this is why we include an extended data figure showing the tight correlation of species richness and contribution to first encounters (and therefore viral sharing). This is also consistent with previous work by Albery *et al.* in *Nature Communications* and Mollentze *et al.* in *PNAS*, the latter of which we have added as a citation.

Line 59 “Some viruses may spread environmentally between spatially-proximate species with no direct behavioral contact” - mention of transmission by vectors seems warranted as well here

We appreciate this suggestion and have added this: “Although some viruses spread environmentally or by arthropod vector between spatially-proximate species with no direct behavioral contact”

Line 111 “If species survive a changing climate by aggregating in high elevation refugia, this suggests emerging viruses may be an increasing problem for their conservation” - again this includes an assumption that populations would have access to higher elevation refugia. Highly fragmented, disjoint lowland forests or other primary habitats often have no remaining suitable habitat connecting them to higher elevation refugia through which species could pass.

We appreciate this suggestion, noting that the connectivity issue is an important one that characterizes all of results and the vast majority of SDM studies, projected over 50 years, and have rephrased the sentence to be a conditional:

“In contrast, if species track thermal optima along elevational gradients, they will aggregate in the most novel combinations in mountain ranges, especially in tropical areas with the highest baseline diversity, matching prior predictions.”

We also mention this qualification in our discussion paragraph: “The patterns we describe are likely further complicated by uncertainties in the species distribution modeling process, including local adaptation or plasticity in response to changing climates, or lack of landscape connectivity preventing dispersal.”

Line 119, “SDM” needs to be defined.

We now define this early on: “We built species distribution models (SDMs) for 3,870 placental mammal species,”

Figure 4: We don’t find this color scheme particularly intuitive, and probably not the best for color-blind readers either?

We agree with the reviewers that colorblind friendly schemes are important - and note that this scheme was chosen specifically for that reason (it’s red-blue, not red-green).

A cited “Data” section in the methods would be helpful.

We appreciate this very helpful suggestion, which has greatly improved the clarity of the methods. We now open the Methods with a Data section that pulls several subsections out of the original Methods, and makes an effort to better reference all existing datasets. That now includes a detailed treatment of the underlying host-virus association data, which have not been modified since their use in a previous study we build on (Albery *et al.* *Nature Communications* 2020), but should still be discussed upfront.

Referee 3

Review of “Climate Change will drive novel cross-species viral transmission” by Carlson et al.

Overview

In this section I explicitly address the criteria requested of me (*italics*); I then include some general comments before addressing specific major/minor comments.

Summary of the key results: The authors find thousands of opportunities for novel virus sharing between mammals if the climate continues to warm over this century. Interestingly, RCP2.6 (the most limited warming scenario), may produce the largest increases in novel sharing.

Originality and significance: This is not my field, but the authors argue convincingly for the novelty.

We appreciate the positive feedback from the reviewer!

Data and methodology: As a non-specialist in that field, I cannot speak for the statistical modelling, but as outlined in my major comments, the treatment of the climate data is not detailed enough for the work to be reproducible.

We have substantially expanded our treatment of the climate data, including the complete overhaul of scenarios used and a detailed rewrite of the Methods.

Appropriate use of statistics and uncertainties: See above cf. statistical modelling; and to follow from the point on climate modelling – uncertainty stemming from the choice of climate model (see major comments) is not acknowledged.

We have worked hard to itemize throughout how little uncertainty we found was attributable to climate model uncertainty, starting by including a standard deviation for every main text statistic by GCM, and including a new Extended Data figure.

Conclusions: They seem novel (but recall I am not a specialist in the field), but due to the treatment of the climate data, I am sceptical that they are robust (major comments).

Suggested improvements: See major comments. Those issues need to be addressed.

References: Seem fine, but the normal caveat about being a non-specialist also applies here.

We hope that our expanded treatment of climate model uncertainty and other broader updates throughout the paper will help clarify the salient points for the reviewer and editor.

Clarity and context: I found the writing to be clear throughout, but there are too many normative statements (see minor comments). These would be very easy to remove during a revision. I would also expect the authors to add some more context about the Covid-19 outbreak in the introduction – if given the chance to submit a revised paper.

Our paper was written long before the COVID-19 outbreak, and in our other work we have been advocates for ecological research not “overselling” relevance to current outbreaks. However, given the obvious relevance to particular mechanisms we discuss, we have added a mention of this to the last paragraph of the study:

“....While several studies have addressed how range shifts in zoonotic reservoirs might expose humans to novel viruses, few have considered the fact that most new exposures will be among wildlife species. The relevance of this process is reinforced by the COVID-19 pandemic, which began only weeks after the completion of this study; the progenitor of SARS-CoV-2 likely originated in southeast Asian horseshoe bats (*Rhinolophus* sp.), and may have spread to humans through an as-yet-unknown bridge host. While we caution against overinterpreting our results as indicative of future pandemic risk—which is largely the product of global health governance, capacity, and preparedness—we note that the global reassortment of mammalian viruses will undoubtedly have a downstream impact on human health (though attribution to climate change will be difficult in any individual case). Tracking spillover into humans is paramount, but so is tracking of viral sharing in wildlife. Targeting surveillance in hotspots of future sharing like southeast Asia, and developing norms of open data sharing for the global scientific community, will help researchers identify host jumps early on, ultimately improving our ability to respond to potential threats.”

General Comments

As a climate scientist reviewing this during the Covid-19 outbreak, I found the topic fascinating, and the paper was generally well written and accessible to a non-specialist in virology/ecology. However, the “climate” element of the paper is communicated in insufficient detail for me to recommend publication. From the information included, I cannot ascertain whether just the level of detail needs to be improved

in the text, or whether the study is flawed in that its conclusions are unlikely to be robust. I hope the authors can clarify and respond to my concerns – outlined in the major comments section below.

Major comments

Detail of the climate change projections and their processing is entirely absent. When inferring the response of a predictand to future climate change simulated by climate models, it is necessary to: (1) establish how the modelled climate for the present diverges from historical reality; (2) to assess how sensitive the results are to the model used. I explain these below.

(1) The changes you see projected in the future could be due to biases in the model’s ability to resolve the current climate. For example, the drying you see projected in the future at some point relative to your observations could be because the model already simulates a climate that is too dry – not because the climate is expected to change in that way. There is an enormous body of work dedicated to dealing with these problems, and it may be that you are applying appropriate methods from that literature, but I see no mention of this in the paper. I stress that this is particularly important here because your most impactful scenario of RCP2.6 has relatively little climate change, so biases in the present climate may have a heavy weight on your results. To be clear: we need details of how the climate data were processed (including how they were interpolated(?) to the 25 km grids).

We used downscaled and bias-corrected climate projection data at 10 minutes spatial resolution (approximately 18.5km at the equator) downloaded from WorldClim2. WorldClim is very widely used in ecology, biodiversity, and agricultural projections of climate change impacts. We regridded the WorldClim data to 0.25 degree spatial resolution (approximately 25km at the equator) using bilinear interpolation to match with the LUH2 land-use data resolution. We have added details on the WorldClim “calibration” process that is intended to address bias correction issues in current climates:

“Temperature and precipitation for future climates have been downscaled and bias-corrected by WorldClim 2 using a change factor approach where the difference in the GCM output between the simulated historical and future period is calculated, resulting in climate anomalies which are then applied to the 10-minute resolution observed historical dataset. This approach makes the assumption that the change in climate is relatively stable across space (that is, has high spatial autocorrelation). We downloaded the five pre-processed Bioclim variables for all nine GCMs at 10 minutes spatial resolution from WorldClim 2, and aggregated with bilinear interpolation to 0.25 degree spatial resolution (approximately 25km at the equator) to match with the LUH2 land use data resolution.”

2) There are well over 20 climate models forced with the same data to simulate possible future climates. The different models do project different climates, and this reflects irreducible uncertainty at this stage. It is therefore good practice to sample across the range of models to see how sensitive your results are to the choice of model. It is indeed common for climate change studies to use all available models, or at least a double-digit ensemble. The choice of HadGEM-ES as the only model here is particularly unfortunate because it has one of the highest climate sensitivities of all the models (it produces one of the largest increases in global mean air temperature for a doubling of CO₂ – the so-called “equilibrium climate sensitivity”: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter09_FINAL.pdf – see p. 818). We need to see how sensitive the results are to climate modelling uncertainty.

We appreciate this comment, which has pointed us towards the biggest change we have made in this revision. We agree that our previous version did not adequately differentiate model uncertainty from scenario uncertainty, and particularly given questions about propagation of uncertainty, this was important to address.

In the new version of our manuscript, we have shifted from using a single CMIP5-generation GCM (HadGEM-ES) to using all nine currently-available CMIP6 GCMs, which span a range of equilibrium climate sensitivities. Every statistic in the paper is now given with a +/- standard deviation across GCMs, which all show that maybe 1-10% of the variation in our study is attributable to these sources. For example:

“...we predict that the vast majority of mammal species will overlap with at least one unfamiliar species somewhere in their future range, regardless of emissions scenario (mean across GCMs +/- s.d. here and after; RCP 2.6: 98.6% +/- 0.2%; RCP 8.5: 96.6% +/- 0.8%). At the global level, community turnover

would permit over 300,000 novel species interactions in every scenario (RCP 2.6: 316,426 +/- 1,719; RCP 8.5: 313,973 +/- 2,094). ”

As additional steps to present this uncertainty in a transparent way that strengthens the analysis, we have added these standard deviations to all of the tile figures in the Extended Data, creating a full presentation of RCP x GCM x Model Assumption uncertainty in key outcomes. We also observe that our results are largely robust to the sensitivity of GCMs, and have included a new extended data figure (ED Fig. 5) that shows where the most dramatic effect of those patterns is by comparing the highest-and lowest sensitivity GCM. Finally, we also present every GCM’s map of first encounters and novel viral sharing in a Supplementary Figures document. We have added a section to the methods about “Climate model uncertainty” contextualizing this:

“To identify the contribution of climate model uncertainty and its propagation through our analysis, we used all nine selected GCMs from CMIP6 and produced multi-model averages for all main text figures. For all of the main text statistics, we present each multi-model mean with a standard deviation across the nine global climate models. We also compared the first encounters from the two models with the highest (CanESM5) and lowest (MIROC6) effective climate sensitivity in the available CMIP6 set on WorldClim. We also present the map of first encounters and novel viral sharing in each GCM run for each RCP, accounting for both climate and land use change, with the full dispersal and limited dispersal scenario, in Supplementary Figures 1-18.”

And two brief references in the main text:

“In contrast, if species track thermal optima along elevational gradients, they will aggregate in the most novel combinations in mountain ranges, especially in tropical areas with the highest baseline diversity, matching prior predictions This pattern was very robust to climate model uncertainty (Supplemental Figures 1-9). The most notable model variation is in the Amazon basin, as well as a small portion of the central African basin, Botswana, and parts of the Indian subcontinent. These areas become essentially devoid of first encounters in the most sensitive climate models and warmest pathways, presumably because all are high-endemism basins of homogenous climate that may warm too fast for species to “escape” into high-elevation refugia (a fairly well-documented pattern). ”

“Across scenarios, the spatial pattern of expected sharing events was nearly identical, and was dominated more by the extent of range shifts than by underlying community phylogenetic structure (Supplemental Figures 10-18).”

Minor comments

L27: what do you mean by “sharing novel viruses between 3,000 and 13,000 times”? This metric (sharing) does not yet mean anything to the reader.

In the abstract we have changed this phrase to: “Range-shifting mammal species are predicted to aggregate at high elevations, in biodiversity hotspots, and in areas of high human population density in Asia and Africa, driving the cross-species transmission of novel viruses at least 4,000 times.”

L122: replace “mildest” with proper terminology – suggest “least radiative forcing” or “climate most similar to recent decades”.

We appreciate this suggestion, and the passage now reads “Even in RCP 2.6 (the scenario with the least warming)”

L124: avoid hyperbole – please remove “still-staggering”

We have removed this phrase: “...and reduced the associated viral sharing events by 70.1% ($\pm$ 0.1%) to 4,584 ($\pm$ 51.9) projected viral sharing events.”

L144: hyperbole – avoid “staggering” (I’m not being pedantic, but scientific writing should be free of normative statements)

We have removed this word: “Several studies have also identified ongoing rapid range expansions in bat species around the world...”

L159-160: Without the background, I don't follow your logic here: where is the wildlife-wildlife host jump with simian immunodeficiency? I think the sentence here needs re-writing to articulate more clearly where the wildlife-wildlife host jumps are.

We apologize for the lack of clarity: the jump was from Old World monkeys to chimps and gorillas along the way to humans. We now clarify this and have added a citation:

“Just as simian immunodeficiency virus making a host jump from monkeys to chimpanzees and gorillas facilitated the origins of HIV, or SARS-CoV spillover into civets allowed a bat virus to reach humans, these kinds of wildlife-to-wildlife host jumps may be evolutionary stepping stones”

L209: Same issue with hyperbole: remove “catastrophic” and replace with the amount of global warming you're referring to (more than 2C? More than 3C?). Please check for, and remove, all instances of hyperbole throughout the text.

Though the IPCC frequently uses this language, we have replaced it here with “unmitigated” and left the broader statement about extreme climate change:

“We caution that this should not be interpreted as a possible upside to unmitigated warming, which will be accompanied by mass defaunation, devastating disease emergence, and unprecedented levels of human displacement and global instability.”

L371: The W/m^2 does not refer to solar radiation, but radiative forcing in general.

We have corrected this - thank you!

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors' revisions are much appreciated and the paper is improved. In my opinion, further explanations are required, including better integration and explanation of the authors' earlier work. The viral sharing model relies heavily on an unpublished paper by Albery, which in turn relies heavily on a Nature paper by Olival. The methods and assumptions of those papers need to be unpacked here.

The authors now define viral sharing explicitly, but they define it as “cross-species transmission of at least one virus”. It is not clear how this could be practical (can we ever know if not even one virus has been “transmitted”), nor why it is biologically relevant. The authors' claim that they have operationalized this definition certainly depends on deeper engagement with the papers mentioned above. The Olival paper seems to rely on _published_ associations, and talks a great deal about publication bias -- an idea that seems absent from this paper (and, on a quick read, from Albery). The unclear chain Olival → Albery → Carlson also _seems to_ conflate the original idea of sharing (A and B _have_ the same virus) with transmission (A _passed_ the virus to B, or vice versa). H7 flu, for example, is probably shared between people and Mongolian wild fowl sensu Olival, but probably not sensu Carlson.

The exact meaning, operationalization, and relevance of viral sharing is at the heart of the paper, and affects the ability to interpret many of the conclusions. Of particular note is the observation at L152 that phylogeny is less important than expected.

Minor

L450 Authors should say in what sense they are being conservative here.

The authors apparently analyzed 3139 species in their final analysis. They mention the intermediate number 3870 8 times before they finally get to 3139 on p13. 3139 is the more relevant number.

Referee #2 (Remarks to the Author):

We appreciate this resubmission which has addressed many of the issues we previously raised, and clearly represents a huge effort. A few minor points emerged in this one:

L109 "This model framework has previously allowed inference..." inference of what?

L129 "However, our findings show that species tend to shift together along latitudinal gradients with others in their community, leading to few first encounters." - presumably even a small amount of variance in relative speed will modulate this?

L131 Is there an intuition as to why mountains are so different from latitudes? Is this because mountains force clustering? (which is suggested in L149)

L133 ... and robust in variance in relative tracking time across species?

L276 "Because hotspots of cross-species transmission are predictable" - how do we know that this is true, given ascertainment bias in almost everything?

Referee #3 (Remarks to the Author):

Summary

I thank the authors for their very detailed response. I am pleased to say that most of my comments from the last round of review have been addressed. My remaining concerns are outlined below. Most of the points I raise are minor, but there is one issue which I think might be more substantial. If they can address it (or clarify any misunderstanding), I would be happy to recommend their work for publication.

Major comments

The authors note that the largest changes in species' overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the "present" – the 1970-2000 mean (i.e. 1985). That is, if "less (climate change) is more" when it comes to novel species' interaction, how do we know that we're not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth's mean air temperature has risen by ~0.7-0.8C.

There are two ways to resolve this:

(1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5

reanalysis)

(2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study.

In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the “present climate”. If not, please comment on how this affects your modelling assumptions.

Minor comments

L142 (and elsewhere): Showing uncertainty to 1 d.p. seems excessive; the nearest integer would suffice.

L 239: “climate velocity” is ambiguous – please use another phrase.

L277-278: unclear what point the “... billions could be spent...” sentence is trying to make. It reads as though your research is enabling billions to be spent (which in itself is not a reason for celebration). Please rephrase to clarify your point.

L329: If “the present”? is 1970 – 2000 (on line 339), then I would content that it is not the “present”, but the recent past (see major comments).

L338: I do not understand what you mean by “global” correlation here. Please clarify/rephrase.

L346: Please justify your assertion that the models used span a wide range of equilibrium climate sensitivities. It would be easiest to do this by quoting the CMIP5 range (across all models), and then clarifying how your sample of 9 GCMs from CMIP6 compares.

L348: A little more detail is needed here. Are the change factors applied monthly? Is it just the difference in means that is accounted for (vs. a quantile-dependent correction)? These details are important because you include seasonality in your five Bioclim variables – and depending on how its applied, the change factor approach may not be able to get the GCM to capture the observed climate.

L479: It isn’t just “the greenhouse gas effect”. I suggest editing this out of the sentence (so you have “... of additional radiative forcing, while SSPs...”)

Referee #4 (Remarks to the Author):

Based on the fact that this manuscript has already been reviewed by several other reviewers, I decided to focus on my review only to the part where I have most expertise, namely statistical aspects of species distribution modelling. This is clearly a very critical part of the work, as the

conclusions are based on predicting how species distributions will change as a response to climate change. If the predictions are not reliable, the results will remain speculative.

Making such predictions on how species respond to climate change is technically relatively easy to do, but what is very hard to do is to evaluate how reliable the predictions are. The authors do a relatively good job in validating their models with spatially stratified cross-validation. This however validates only the current distributions, and not the future distributions. The big open question is thus whether the predicted future distributions are even any better than one would expect e.g. by chance. To address this question, the authors should validate their models not only for the current distribution, but also for the change of a distribution, e.g. based on using some historical data for model fitting, and testing if the models predict the changes in distributions that have already taken place.

A more rigorous model validation (i.e., validating that the models correctly predict changed species is probably outside the scope of this manuscript (and is generally rarely done in climate change research). As the models are not validated in terms of predicting changes in species distributions, the research remains necessarily highly speculative. To what extent this is a problem concerning the publication of the results in Nature is essentially an editorial decision. What the authors could in case do, is at least to discuss even more openly to what level the results are speculative.

Furthermore, I note that the authors now test the robustness of the results to different climate change predictions, as well as to accounting for dispersal limitation or not including dispersal limitation. They however do not account for structural model uncertainty, as all results are based on the same species distribution modelling framework as well as the same choices of environmental predictors. It would be possible to examine the robustness of the results to these choices; though even in this case, the fundamental problem of the models remaining not validated in terms of their future predictions would remain.

Let me finally note that the above criticism is not very specific to this manuscript, but holds to most work that uses species distribution modelling to predict how biodiversity responds to climate change. Thus, what the authors do in this manuscript largely follows the state-of-the-art in this field.

Response to reviewers

Carlson *et al.*, submission to *Nature*

April 2021

Response to Editor

We are very grateful to the four referees and the editor for thoughtful suggestions and the chance to revise the manuscript. As part of this revision, we have added an entirely novel analysis showing that the world today could already experience higher (but probably not peak) viral sharing rates due to climate change impacts (at the suggestion of reviewer 3); clarified key definitions foundational to our study (at the request of reviewer 1); aimed to more transparently present our species distribution modeling approach, including a new sensitivity analysis (at the request of reviewer 4); and throughout, made additional improvements at all reviewers' suggestions that improve the clarity and transparency of the manuscript. As such, we believe our second revision is substantially improved, and hope the editors and reviewers agree we have fully addressed all outstanding concerns.

Referee #1 (Remarks to the Author)

The authors' revisions are much appreciated and the paper is improved. In my opinion, further explanations are required, including better integration and explanation of the authors' earlier work. The viral sharing model relies heavily on an unpublished paper by Albery, which in turn relies heavily on a Nature paper by Olival. The methods and assumptions of those papers need to be unpacked here.

The authors now define viral sharing explicitly, but they define it as "cross-species transmission of at least one virus". It is not clear how this could be practical (can we ever know if not even one virus has been "transmitted"), nor why it is biologically relevant. The authors' claim that they have operationalized this definition certainly depends on deeper engagement with the papers mentioned above. The Olival paper seems to rely on *published* associations, and talks a great deal about publication bias – an idea that seems absent from this paper (and, on a quick read, from Albery). The unclear chain Olival to Albery to Carlson also *seems to* conflate the original idea of sharing (A and B *have* the same virus) with transmission (A *passed* the virus to B, or vice versa). H7 flu, for example, is probably shared between people and Mongolian wild fowl *sensu* Olival, but probably not *sensu* Carlson.

The exact meaning, operationalization, and relevance of viral sharing is at the heart of the paper, and affects the ability to interpret many of the conclusions. Of particular note is the observation at L152 that phylogeny is less important than expected.

We appreciate this point, and aim here to clarify some remaining points, which we have also added to the text at the required locations.

- The three studies all use the same dataset describing known associations of hosts and viruses (the HP3 dataset compiled by the Olival study).
- Despite the data overlap, the Olival study overlaps minimally with Albery's study or ours, as it does not concern viral sharing (instead, it focuses on predicting a host's total viral diversity and zoonotic viral diversity, and conversely, the zoonotic potential of a given virus). Zoonosis, as the sharing of a virus from an animal to a human, is a kind of viral sharing - but not one that we predict in this manuscript. Both Albery's study and this one exclude *Homo sapiens* from the sharing model.
- We have now more clearly defined viral sharing in the main text as (and differing from the reviewer's interpretation): when "species... are currently known to host one or more of the same virus species

(hereafter *viral sharing*)” (lines 63-64). We define a **novel viral sharing event** in the text as “future cross-species transmission of at least one virus, between a pair of hosts that are newly in contact” (lines 110-111) - a future change in the *status quo* of the viral sharing network.

- While all three studies use the same dataset, and correct for sampling bias similarly, the Albery and Carlson studies each incorporate additional different data (the Albery study incorporates an external validation dataset; our study incorporates the species distribution modeling workflow, and the underlying climate data, species occurrence data, and habitat preference data). Moreover, the Albery study identifies the conditions that allow hosts to share viruses at equilibrium (**viral sharing**), while here, we identify which species will experience those conditions in the future, and predict **novel viral sharing events**. Unlike Olival, we do not predict zoonosis (viral sharing between animals and humans).
- In the Albery study’s viral sharing model, which we use in this study, both phylogeny and geography are important. However, we find here that when that model is integrated with range shift predictions, the phylogenetic structure of communities is a relatively minor contributor to geographic patterns. We have added a sentence that explains this in slightly greater depth: “..., in our framework, opportunity drives spatial patterns more than compatibility. Given that phylogeny is a strong determinant of viral sharing in the underlying model, this difference from previous studies can probably be explained by evolutionary scale, where prior work focused on primates, and our study includes all mammals. At this broader scale, predicted viral sharing patterns mostly track total host richness (see Figure 3b in Albery *et al.*), whereas at finer scales, phylogeny has a stronger effect (see Extended Data Figure 8 for an example).”
- Finally, the final published version of the Albery study — Albery, Gregory F., et al. “Predicting the global mammalian viral sharing network using phylogeography.” *Nature Communications* 11.1 (2020): 1-9 — discusses sampling bias, including in regards to methodological correction for it, by using the same methods as the Olival study (correction for citation counts), and uses this approach to make predictions in a way that minimizes the effect of this bias. This study’s discussion of sampling bias is reproduced below:

“...there are certain inherent biases in the training data which must be considered when interpreting our findings. Most notably, viral sharing estimates in our dataset may be affected by the fact that zoonotic discovery efforts commonly search limited geographic regions for a specific virus or group of viruses, artificially increasing the likelihood of detecting these viruses in the same region compared to a geographically random sampling regime. Moreover, when a mammal species (e.g., a bat) is found with a focal virus (e.g., an ebolavirus), it is logical for researchers to then investigate similar, closely related species in nearby locales. These sampling approaches could disproportionately weight the network towards finding phylogeographic effects on viral sharing probability. However, it is highly encouraging that our model predicted patterns in the external EID2 dataset, which was constructed using different data compilation methods but also comprises global data covering several decades of research. In sum, we believe that our approach is a conservative method for minimising the biases inherent in the data.”

Based on this explanation, we have added these sentences to the Methods section, to link this more up-front: “We note that this model uses citation counts to correct for sampling bias, an imperfect method but one that leads to strong validation performance on an independently-compiled dataset of host-virus associations, which carries a different set of biases. However, it is still possible that sampling bias in host-virus datasets like the Olival *et al.* dataset could artificially inflate the signal of phylogeography in viral sharing, if researchers investigating a noteworthy viral detection then preferentially sample closely-related host species in the immediate area. It is unlikely these effects would bias our results in a particular direction, but accounting for these biases should at least involve some acknowledgement that cross-species transmission is challenging to predict. (See the Albery *et al.* study’s Discussion for a more in-depth treatment of sampling bias effects.)”

In conclusion, we appreciate the reviewer’s encouragement to engage with these issues in more depth, and we hope that these additional dimensions of framing have better clarified the overlap with prior work, the definitions we operationalize throughout the study, and how those differ from previous studies.

Minor

L450 Authors should say in what sense they are being conservative here.

We have clarified this sentence: “Filtering based on habitat was done as permissively as possible: species were allowed in current and potential future ranges to exist in a pixel if any non-zero percent was assigned a suitable habitat type; almost all pixels contain multiple habitats.”

The authors apparently analyzed 3139 species in their final analysis. They mention the intermediate number 3870 8 times before they finally get to 3139 on p13. 3139 is the more relevant number.

We appreciate this point and have changed this number in the abstract, and have also added it on line 115 (“Based on the total number and distribution of first encounters among a subset of 3,139 species...”) and at the head of the Methods section (“For a final subset of 3,139 species, we predict the...”). These numbers are also presented in Extended Data Figure 10 for clarity.

Referee #2 (Remarks to the Author)

We appreciate this resubmission which has addressed many of the issues we previously raised, and clearly represents a huge effort.

Thanks for this note! It has indeed been an enormous effort, so this is lovely to hear.

A few minor points emerged in this one:

L109 “This model framework has previously allowed inference...” inference of what?

We have updated this sentence to read “This model framework has previously provided insights into viral macroecology and zoonotic risk based on the 1% of the global mammalian virome that has been described.”

L129 “However, our findings show that species tend to shift together along latitudinal gradients with others in their community, leading to few first encounters.” - presumably even a small amount of variance in relative speed will modulate this?

Indeed, some variance will introduce more novel overlaps at range boundaries; however, this constitutes a relatively small portion of species’ ranges, concentrated right at the edges. Throughout much of a species’ range (i.e., not near the edge), the potential shifts are such that species are projected to co-occur in very similar assemblages in the future compared to the present (or subsets of those assemblages, meaning no novel overlaps). To clarify, we have phrased this sentence “... we find that when species shift along latitudinal gradients, they travel in the same direction as others that are already included in their assemblage, leading to few first encounters.”

L131 Is there an intuition as to why mountains are so different from latitudes? Is this because mountains force clustering? (which is suggested in L149)

The reviewer correctly interprets this result, and we have added another phrase for clarification: “In contrast, when species track thermal optima along elevational gradients (allowing them to come from different directions; i.e., mountains force species to cluster)...”

L133 ... and robust in variance in relative tracking time across species?

Yes; we have now added this to the sentence: “This pattern was robust to climate model uncertainty (Supplemental Figures 1-9) and to differences in dispersal capacity (e.g., Figure 2C).”

L276 “Because hotspots of cross-species transmission are predictable” - how do we know that this is true, given ascertainment bias in almost everything?

We appreciate this point, and have removed this clause: “Our study provides the first template for how surveillance could target future hotspots of viral emergence in wildlife.”

Referee #3 (Remarks to the Author)

Summary

I thank the authors for their very detailed response. I am pleased to say that most of my comments from the last round of review have been addressed. My remaining concerns are outlined below. Most of the points I raise are minor, but there is one issue which I think might be more substantial. If they can address it (or clarify any misunderstanding), I would be happy to recommend their work for publication.

We thank the reviewer very much for the supportive feedback!

Major comments

The authors note that the largest changes in species' overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the "present" – the 1970-2000 mean (i.e. 1985). That is, if "less (climate change) is more" when it comes to novel species' interaction, how do we know that we're not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth's mean air temperature has risen by 0.7-0.8C.

There are two ways to resolve this:

(1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5 reanalysis) (2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study.

In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the "present climate". If not, please comment on how this affects your modelling assumptions.

We appreciate this suggestion, and have chosen to address it with a new analysis (the first suggestion, as recommended by the Editor) that allows us to state the world today could be experiencing higher viral sharing than a pre-warming baseline. To accomplish this, we have rerun the entire analysis using the ERA5 reanalysis product as the reviewer suggested, and have included a new methods section and Extended Data Figure reporting the results. However, we caution that because we use the ERA5 product as specified by the reviewer's suggestion, these analyses cannot necessarily be considered comparable to those calibrated with the WorldClim data (given methodological differences, e.g., in the bias correction algorithm), and so we cannot make strong inference on whether viral sharing has "peaked" today relative to the future. Descriptive statistics like the number of first encounters can be compared by readers between the main analysis (2070) and the ERA5 analysis (2020), but we have been careful to caution how they might be interpreted: "We caution that these results [using the ERA5 data]—particularly the number of encounters and sharing events—should not be interpreted as the same "units" as the main analysis, given that they are calibrated to an entirely different climate reconstruction."

In addition to the new section "The effect of recent warming" in the Methods, we include these results in the section "Climate change mitigation is insufficient to prevent undesirable outcomes":

"...there were 5.4% ($\pm 1.7\%$) fewer potential first encounters in RCP 8.5 compared to RCP 2.6, and unexpectedly, a 1.9% ($\pm 0.3\%$) predicted reduction in the connectivity of the future global viral sharing network (ED Figure 3B,D). Overall, our results indicate that a mild perturbation of the climate system could create thousands of new opportunities for viruses to find new hosts. Finally, in a supplemental analysis comparing the present climate to the near past (see Methods and ED Figure 9), we found that if species are already tracking shifting habitats, present-day Africa and the Amazon might already be hotspots of novel cross-species viral transmission, given the warming that has taken place over the last 25 years ($\sim +1^\circ\text{C}$)."

Finally, we allude back to these results at the end of the main text:

"While several studies have addressed how range shifts in zoonotic reservoirs might expose humans to novel viruses, few have considered the fact that most new exposures will be among wildlife species. The relevance of this process is reinforced by the COVID-19 pandemic, which began only weeks after the

completion of this study; the progenitor of SARS-CoV-2 likely originated in southeast Asian horseshoe bats (*Rhinolophus* sp.), and may have spread to humans through an as-yet-unknown bridge host. While we caution against overinterpreting our results as explanatory of the current pandemic or indicative of future pandemic risk—which is largely the product of global health governance, capacity, and preparedness—we note that the global reassortment of mammalian viruses will undoubtedly have a downstream impact on human health (though attribution to climate change will be difficult in any individual case).”

Minor comments

L142 (and elsewhere): Showing uncertainty to 1 d.p. seems excessive; the nearest integer would suffice.

We appreciate this and have made the change throughout where appropriate.

L 239: “climate velocity” is ambiguous – please use another phrase.

We have rephrased this clause “...when warming is slower, species can successfully track shifting climate optima...”

L277-278: unclear what point the “... billions could be spent...” sentence is trying to make. It reads as though your research is enabling billions to be spent (which in itself is not a reason for celebration). Please rephrase to clarify your point.

We have rephrased this sentence: “In the next decade alone, it may cost at least a billion dollars to comprehensively identify and counteract zoonotic threats before they spread from wildlife reservoirs into human populations.” We cite the study that estimates this price point as well (this is a specific cost estimate, not a general conjecture, as it previously read).

L329: If “the present”? is 1970 – 2000 (on line 339), then I would content that it is not the “present”, but the recent past (see major comments).

We have added a detailed qualifier that is as transparent as possible and links this framing to the new analysis: “Climate and land use data were compiled from WorldClim 2 and the Land Use Harmonization 2 (LUH2) project respectively, for both baseline conditions (operationalized as 1970-2000 for the climate data, 2015 for land use, and 2020 for dispersal limits; see “The effect of recent warming” for an interrogation of the difference between climate baselines and actual present-day climate) and a half-century in the future (operationalized as 2061-2080 for climate, 2070 for land use, and 2070 for dispersal).”

L338: I do not understand what you mean by “global” correlation here. Please clarify/rephrase.

We have rephrased this sentence to read “This is the largest set of Bioclim variables possible that keeps their correlation over a global extent suitably low ($r < 0.7$)...”

L346: Please justify your assertion that the models used span a wide range of equilibrium climate sensitivities. It would be easiest to do this by quoting the CMIP5 range (across all models), and then clarifying how your sample of 9 GCMs from CMIP6 compares.

The range of effective climate sensitivities is 1.8-5.6K for 27 global climate models in CMIP6. This is, on average, a higher climate sensitivity than for CMIP5 for which the range of effective climate sensitivities was 2.1-4.7K. The nine models from CMIP6 we use in our analysis span the range from 2.6K (MIROC6) to 5.6K (CanESM5). We now provide the effective climate sensitivity ranges calculated by Zelinka et al. 2020 for the 27 CMIP6 models and for CMIP5 in the methods text when we list the climate models we use in our analysis: “These nine GCMs encompass a wide range of effective climate sensitivities from 2.6K (MIROC6) to 5.6K (CanESM5) compared with a range of 1.8-5.6K across 27 CMIP6 models and 2.1-4.7K for CMIP5.”

L348: A little more detail is needed here. Are the change factors applied monthly? Is it just the difference in means that is accounted for (vs. a quantile-dependent correction)? These details are important because you include seasonality in your five Bioclim variables – and depending on how its applied, the change factor approach may not be able to get the GCM to capture the observed climate.

Yes, the change factors are applied for each month. The steps used by the WorldClim dataset are as follows: “(1) calculation of 30-year averages for each month for the present-day simulations and 20-year averages for each month for the future period; (2) calculation of anomalies as the absolute difference between future and present day values in temperatures (minimum and maximum) and proportional

differences in total precipitation; (3) interpolation of these anomalies using centroids of GCM grid cells as points for interpolation; and (4) addition of the interpolated gridded data to the current climates from WorldClim” (see Navarro-Racines et al. 2020 *Scientific Data*, which we now cite). The resulting temperature and precipitation layers are then used by WorldClim to calculate the the Bioclim variables. We now state in the method text that the change factors are applied monthly and describe the method used by WorldClim in more detail:

“Temperature and precipitation for future climates have been downscaled and bias-corrected by WorldClim 2 using a change factor approach. The multi-year average of the GCM output for minimum temperature, maximum temperature and total precipitation is calculated for each month of the simulated historical and future period, and the absolute (for temperature) or proportional (for precipitation) difference in these values is then calculated, resulting in climate anomalies which are then applied to the 10-minute spatial resolution observed historical dataset. WorldClim 2 then calculates Bioclim variables based on these downscaled and bias-corrected data. This approach makes the assumption that the change in climate is relatively stable across space (that is, has high spatial autocorrelation).”

L479: It isn’t just “the greenhouse gas effect”. I suggest editing this out of the sentence (so you have “... of additional radiative forcing, while SSPs...”)

We have made the suggested change and the passage now reads “RCP numbers (e.g., 2.6 or 4.5) represent Watts per square meter of additional radiative forcing by the end of the century...”

Referee #4 (Remarks to the Author)

Based on the fact that this manuscript has already been reviewed by several other reviewers, I decided to focus on my review only to the part where I have most expertise, namely statistical aspects of species distribution modelling. This is clearly a very critical part of the work, as the conclusions are based on predicting how species distributions will change as a response to climate change. If the predictions are not reliable, the results will remain speculative.

Making such predictions on how species respond to climate change is technically relatively easy to do, but what is very hard to do is to evaluate how reliable the predictions are. The authors do a relatively good job in validating their models with spatially stratified cross-validation. This however validates only the current distributions, and not the future distributions. The big open question is thus whether the predicted future distributions are even any better than one would expect e.g. by chance. To address this question, the authors should validate their models not only for the current distribution, but also for the change of a distribution, e.g. based on using some historical data for model fitting, and testing if the models predict the changes in distributions that have already taken place.

We appreciate this healthy skepticism of SDM transfer applications. The reviewer proposes two alternatives to address their concerns: (1) validate the models through transfer in time or (2) (mentioned below) to better caveat our model’s limitation and their appropriate interpretation. We have chosen option (2), and just briefly point out the key challenges with option (1) (which it seems the reviewer is likely familiar with, based on comments below). First, we note that using only historical data (e.g., collected before 1985) would dramatically reduce the sample sizes for most species, making models intractable for many of our 3,000+ species. Second, considerably greater sampling bias is likely to exist in these earlier and much sparser data sets, leading to significant practical challenges in fitting robust models for such a large group of species. Third, models built in this way may well differ from the models we fit using all existing data because they would depend on different and significantly-reduced datasets, and hence, we would not actually be evaluating the models we forecast with.

Therefore, we have aimed throughout to better explain the strengths and limitations of interpretation of the models we have already built, including a number of changes in language detailed in the other responses. Most importantly, we highlight that our future projections should not be interpreted as realized species distributions (and understand that our previous phrasing could be misinterpreted); rather, they reflect potentially suitable locations, conditional only on species climatic niches. Throughout, we refer to future distributions as potential future distributions, and are up-front about this in the definition of our study: “Here, by projecting potential geographic range shifts (that is, newly suitable habitat, which a species may or may not migrate to) and applying mechanistic biological rules for cross-species

transmission, we predicted how and where global change could potentially create novel opportunities for viral sharing, with particular attention to the potential connections between these risks and human health.” We also now clarify in the Methods: “...while many factors besides climate are ignored by our models, such as biotic interactions or animal social behavior, our models are tailored to our aim: predicting hotspots of elevated risk under climate change. In our application, correctly predicting presences is more important than incorrect prediction of absences, because we are focused on the potential for novel species overlap. We cannot say whether that overlap *will* happen, based on the multiple factors besides climate that influence distributions and range shifts, but we can say with confidence - based on robust current niche estimates, validated with spatially stratified cross-validation, and biologically-grounded estimates of dispersal capacity - where risk would be elevated in accordance with our simulations.”

A more rigorous model validation (i.e., validating that the models correctly predict changed species is probably outside the scope of this manuscript (and is generally rarely done in climate change research). As the models are not validated in terms of predicting changes in species distributions, the research remains necessarily highly speculative. To what extent this is a problem concerning the publication of the results in *Nature* is essentially an editorial decision. What the authors could in case do, is at least to discuss even more openly to what level the results are speculative.

As noted above, we have chosen to take the approach of explaining how our results represent predictions of potential risk rather than definitive expectations for the future. We have clarified that our interpretations should not be interpreted as future realized distributions but rather as newly suitable habitat; that is, we cannot say whether a location will be occupied, but instead, whether it is newly within a species’ climate niche, and whether species are likely to be capable of the dispersal required to reach it. This distinction is critical, but we do not believe that it reduces the value of our forecasts - we aim to ensure that we capture all the potential locations for new overlaps between species (true positives are more important than false positives). Based on robust estimates (discussed below) of species present-day climate niches, we can project all the places that newly fall within a species’ niche in the future, and these places therefore have elevated risk of cross-species transmission. Rather than interpreting such predictions as ‘highly speculative,’ we believe our predictions are critical as an early warning system to target enhanced monitoring to detect new transmissions. We also note that we are not aware of any alternative approach that could or should instead be used to generate species distribution maps across so many taxa at a global extent for our modelling purposes.

Furthermore, I note that the authors now test the robustness of the results to different climate change predictions, as well as to accounting for dispersal limitation or not including dispersal limitation. They however do not account for structural model uncertainty, as all results are based on the same species distribution modelling framework as well as the same choices of environmental predictors. It would be possible to examine the robustness of the results to these choices; though even in this case, the fundamental problem of the models remaining not validated in terms of their future predictions would remain.

Regarding structural model uncertainty, we note that recent research has shown that a carefully-built and well-tuned model from a single algorithm outperforms multi-algorithm ensembles of untuned models (e.g. Hao *et al.* 2020 *Ecography*: “Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models”). We believe that our models are well-tuned because: (1) we cleaned presence records for environmental or geographic outliers using the Grubb’s outlier test (we know of no other SDM study that has automated outlier removal); (2) we only used predictors with low correlation, to reduce the dependence on correlation structures remaining the same between the present and future climates; (3) we used penalized regression to arrive at an optimal model (i.e., an alternative approach to model selection); (4) we use a comprehensive search of the full regularization path to determine the optimal model complexity; (5) we used spatially stratified cross-validation to both accurately assess model performance and combined predictions across the folds to avoid overfitting (no other studies have used the conservative spatial fold assignment algorithm); (6) we do not allow positive coefficients for quadratic features in the models (i.e., all have an upper bound of 0 during fitting) to avoid the undesirable effect of increasing suitability predictions at range edges (no other SDM studies have used this approach); and (7) for future projections, we avoid extrapolation more than 1 standard deviation along each climate variable axis beyond the range of climate conditions in the training presences (such limitation is rare in SDM studies, and none have used this conservative rule). Together, this approach is designed to be incredibly stringent and tracks the most up-to-date information on SDM calibration in

the literature.

As for other potential abiotic drivers we could have used, we note that our focus is on climate change induced potential range shifts, and so only climate variables are relevant to our study. We have chosen variables that characterize the main dimensions of climate that can be reasonably forecasted into the future (temperature, precipitation, and their seasonalities); other variables are likely to have diminishing importance relative to these core dimensions of change; and in any case, many alternatives would be highly correlated with them, and would therefore contribute little predictive ability to our correlative models.

Let me finally note that the above criticism is not very specific to this manuscript, but holds to most work that uses species distribution modelling to predict how biodiversity responds to climate change. Thus, what the authors do in this manuscript largely follows the state-of-the-art in this field.

We appreciate this set of concerns and summarize our changes to address them as follows:

First, we have rephrased the text comprehensively to avoid implying that our future projections represent future realized distributions. Instead we refer exclusively to newly suitable or unsuitable habitat, based on species climate niches. (In all instances where we use the phrase “future distribution” or “future range” this is now preceded with the word “potential.”) Second, we have changed the section “Model validation” to “Model validation and limitations”, and added both a new analysis demonstrating the sensitivity of our methods to one exemplar user-end choice (thresholding) and a broader passage encouraging caution in the interpretation of the results and their dependence on model assumptions:

“Although our models performed well, we note that researchers should approach the interpretation of species distribution models (SDMs) with a certain degree of caution. Even adhering to best practices, many SDM methods are sensitive to subjective user-end choices that influence model performance, transferability, and interpretability. Some of those choices may have marginally affected the patterns we document in this study. For example, [...analysis removed for ease of reading.] These results demonstrate that the choice of thresholding had an impact on the results of our analysis, but an extremely marginal one, and we expect similar results would be found for other choices like variable set reduction, model calibration, the resolution of predictor data, and the processing of point occurrence data.”

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

Major comment:

The novel element in this manuscript (adding predicted species distributions in possible futures) seems to remain rather speculative due to limitations including: (1) the currently poor knowledge of viral diversity in nonhuman mammal communities, (2) uncertainty in how populations of diverse mammal species will move (see Reviewer 4 comment "what is very hard to do is to evaluate how reliable the predictions are"), and (3) how hypothetical new communities may interact and the relevance to pathogen sharing of those hypothetical new interactions. While this study grapples with a truly interesting topic and presents an enormous body of work, given the above listed limitations, it is difficult to interpret the strength of the conclusions, and how much further they take us than existing studies using the same core data set (e.g., Albery et al Nature Communications). However, as also indicated by the other reviewers, it is close to being the current state of the art.

Minor comments:

Some findings are presented in the abstract in a way that prevents the reader from appreciating the novelty of the results:

For example, "Range-shifting mammal species are predicted to aggregate at high elevations, in biodiversity hotspots" is presented as the first result of this study but that is fairly intuitive: It is difficult to think of how range-shifting mammal species could aggregate at low elevation, low diversity areas

Again, a key result presented in the abstract is "Most projected viral sharing is driven by diverse hyperreservoirs (rodents and bats)": which is expected? Rodents and bats contribute approx. half of mammal species so contributing a lot of the viral sharing is unsurprising

[L52] Very minor point but current estimates are 6,000+ mammal species. See: How many species of mammals are there? Burgin et al. Journal of Mammalogy, Volume 99, Issue 1, 1 February 2018, Pages 1–14, <https://doi.org/10.1093/jmammal/gyx147>

[L85] "Here, by projecting potential geographic range shifts (that is, newly suitable habitat, which a species may or may not migrate to)" - Is it more accurate (see discussion with Reviewer 4), and more straightforward, to simply say "projecting newly suitable habitat, which a species may or may not migrate to"

The "Dispersal drives bats' disproportionate importance" section may warrant some qualifying

statements that reflect the uncertainties around the author's conclusions

For example,

[L170] While this is mentioned in L283 in the conclusion, it would be appropriate to note here that not only is dispersal potential likely to vary but the degree to which dispersal potential is realized is also likely to vary among taxa (e.g., corridors of remaining habitat that would allow access to future suitable habitats no longer exist in some areas, some taxa are more likely to be targeted by conservation efforts and intentional relocation, etc)

[L199] While bats in theory could rapidly shift their range, their preferred food sources, preferred roosting sites etc, may not rapidly shift in which case the effect of their flight ability may be attenuated

The focus on RCP2.6 implies that countries will meet the Paris goal - which is perhaps unlikely - it seems a bit as if this scenario was chosen because it maximizes the results, which perhaps does a disservice to the analysis as a whole? Extended Data Fig. 3 might perhaps be usefully added to the main paper?

It wasn't always obvious if the numbers represented a change on the baseline (i.e. present-day) model run and when the results were the absolute numbers for scenario X. For instance Fig. 2C/D/E, are these the change in predicted encounters (vs the present) or the total predicted encounters? The same for Fig. 1. The key question is how much difference does climate change make on the baseline scenario. It seems from Extended Data Fig. 3D that the change is actually very small?

Referee #3 (Remarks to the Author):

Please see the attachment.

Re-review of Climate change will drive novel cross-species viral transmission by Carlson et al.

I thank the authors for their considerable efforts and have only a few points to raise after this revision, mostly related to the major issues I highlighted last time (Referee #3). I said:

The authors note that the largest changes in species' overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the "present" -- the 1970-2000 mean

(i.e. 1985). That is, if "less (climate change) is more" when it comes to novel species' interaction, how do we know that we're not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth's mean air temperature has risen by 0.7-0.8°C.

There are two ways to resolve this:

(1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5 reanalysis)

(2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study.

In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate". If not, please comment on how this affects your modelling assumptions.

The authors have dealt with this by adding a new section: The effect of recent warming, which is a very welcome inclusion. However, the authors note in their response (and in the paper) that the results from this analysis cannot be compared directly with analysis based on WorldClim due to the fundamental differences in how these datasets are constructed (e.g., bias correction). The remaining question is why these datasets cannot be made equivalent. In other words, why can you not 'adjust' (i.e., bias correct) ERA5 to WorldClim? In essence, this allows you to "get" WorldClim from ERA5 for the real present. I understand that this would be demanding new analysis, and am not insisting on it, but I highlight here that such an adjustment would be possible – not least because ERA5 data are now available from 1950- present (which may not have been true at the time of the authors' last revision, when I think data started in 1979). It would be useful to hear the authors' response to this suggestion. Also, what may be a quick question to answer – the authors did not confirm whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate?". If not, please comment on how this affects your modelling assumptions. I am not familiar with the species data (and may have missed key details in the paper), but can you confirm that the species' presence records are for the same period as the 1970-2000 interval used to define the 'present' climate? I ask because, if they do not align, don't you risk incorrectly specifying the climate-species dependency (because the climate has changed considerably since 1970)? Finally, given the new analysis with ERA5, I suggest that the title should change from: Climate change will drive novel cross-species viral transmission

to: Climate change drives novel cross-species viral transmission

(because it is/has likely occurred already)

Referee #4 (Remarks to the Author):

In my original comments, my main criticism was that the results of the ms remain somewhat speculative because of the lack of validation on the models ability to predict changes over time. I proposed that the authors would include such validation, or at least discuss more openly the limitations of their approach. I fully understand that such validation is not easy to do, partially because of sparsity of the appropriate data. I think that the authors have done a very good job in openly discussing the limitations of their approach and how the results can/cannot be interpreted. I find the manuscript highly interesting, and support its publication.

Response to reviewers

Carlson *et al.*, submission to *Nature*

April 2021

Response to Editor

We are very grateful to the three referees on revision #3 and the editor for thoughtful suggestions and the chance to revise the manuscript. We have taken care to address the referees' remaining points, and believe these have brought the manuscript close to a finalized state.

Referee #1 (Remarks to the Author)

No comments provided in revision #3.

Referee #2 (Remarks to the Author)

Major comment:

The novel element in this manuscript (adding predicted species distributions in possible futures) seems to remain rather speculative due to limitations including: (1) the currently poor knowledge of viral diversity in nonhuman mammal communities, (2) uncertainty in how populations of diverse mammal species will move (see Reviewer 4 comment "what is very hard to do is to evaluate how reliable the predictions are"), and (3) how hypothetical new communities may interact and the relevance to pathogen sharing of those hypothetical new interactions. While this study grapples with a truly interesting topic and presents an enormous body of work, given the above listed limitations, it is difficult to interpret the strength of the conclusions, and how much further they take us than existing studies using the same core data set (e.g., Albery et al Nature Communications). However, as also indicated by the other reviewers, it is close to being the current state of the art.

We agree with the reviewer about the predictive challenges that are embedded in this study—and many others about the future impacts of climate change on health and biodiversity—and have aimed to discuss all three of these uncertainties extensively. We appreciate the reviewer's point that our approach is as close as possible to the current state of the art (see also Referee #4's comments below), and also note that throughout, we have carefully framed our predictions as what *could* happen rather than what *will*, given the uncertainties identified by the reviewer.

As discussed in revisions 1 and 2, our study is the first to address the potential impacts of climate change on the future global virome. In that regard, it represents an advance on the Albery study and others that only explore the patterns underlying the present-day viral sharing network, and do not address climate change impacts. As such, we feel that the challenges of the predictive exercise are worth the potential insights into important climate change impacts, with relevance to wildlife and human health. Moreover, our study is the first to suggest where wildlife disease surveillance could be increased in response to this aspect of global change, and—by identifying areas that may experience elevated disease risks in the future *before* these climate impacts become more severe (including in specific cases that require further consideration, like the potential movement of Ebola virus reservoirs into east Africa)—we believe that our predictions have the potential to be an asset to outbreak preparedness efforts.

Minor comments:

Some findings are presented in the abstract in a way that prevents the reader from appreciating the novelty of the results:

For example, “Range-shifting mammal species are predicted to aggregate at high elevations, in biodiversity hotspots” is presented as the first result of this study but that is fairly intuitive: It is difficult to think of how range-shifting mammal species could aggregate at low elevation, low diversity areas

We appreciate the reviewer’s point here, and agree that these results are intuitive and generally supported by prior assumptions. There are plausible counter-results that are discussed in the main text that were not necessarily evident in the abstract: for example, species are also predicted to aggregate in low-diversity sub-Arctic communities at higher richness than today; these communities, however, have fewer first encounters. To set up this finding better, we have rephrased this point: “Range-shifting mammal species are predicted to aggregate *in new combinations...*”; this highlights an aspect of our study that is methodologically unique.

Again, a key result presented in the abstract is “Most projected viral sharing is driven by diverse hyper-reservoirs (rodents and bats)”: which is expected? Rodents and bats contribute approx. half of mammal species so contributing a lot of the viral sharing is unsurprising

We appreciate the reviewer’s note here. The full result in the abstract (“Most projected *novel* viral sharing is driven by diverse hyperreservoirs (rodents and bats) and large-bodied predators (carnivores).”; insertion marked in italics) is indeed intuitive, but not necessarily unimportant or uncontroversial; several recent high-profile studies have been published specifically making the point that host diversity begets centrality in viral sharing and diversity, including the Albery study and: Mollentze, Nardus, and Daniel G. Streicker. “Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts.” *Proceedings of the National Academy of Sciences* 117.17 (2020): 9423-9430.

[L52] Very minor point but current estimates are 6,000+ mammal species. See: How many species of mammals are there? Burgin et al. *Journal of Mammalogy*, Volume 99, Issue 1, 1 February 2018, Pages 1–14, <https://doi.org/10.1093/jmammal/gyx147>

Thanks for spotting this error; we have made the suggested change: “...in which humans are only one of at least ~6,5000 mammal hosts with over 21 million possible pairwise combinations.”

[L85] “Here, by projecting potential geographic range shifts (that is, newly suitable habitat, which a species may or may not migrate to)” - Is it more accurate (see discussion with Reviewer 4), and more straightforward, to simply say “projecting newly suitable habitat, which a species may or may not migrate to”

We appreciate this suggestion and have made the suggested change; the sentence now reads “Here, by projecting newly suitable habitat, which a species may or may not migrate to, and applying mechanistic biological rules for cross-species transmission, we predicted...”

The “Dispersal drives bats’ disproportionate importance” section may warrant some qualifying statements that reflect the uncertainties around the author’s conclusions

For example,

[L170] While this is mentioned in L283 in the conclusion, it would be appropriate to note here that not only is dispersal potential likely to vary but the degree to which dispersal potential is realized is also likely to vary among taxa (e.g., corridors of remaining habitat that would allow access to future suitable habitats no longer exist in some areas, some taxa are more likely to be targeted by conservation efforts and intentional relocation, etc)

We thank the reviewer for this suggested point and agree; we have added this point to the sentence:

“Species’ intrinsic dispersal capacity is likely to constrain their ability to move to newly suitable locations, and therefore to limit novel viral sharing (as are harder to predict extrinsic factors, such as landscape connectivity or facilitated migration by conservation efforts).”

[L199] While bats in theory could rapidly shift their range, their preferred food sources, preferred roosting sites etc, may not rapidly shift in which case the effect of their flight ability may be attenuated

We agree, and have made the suggested change at this location:

“If flight does allow bats to undergo more rapid range shifts than other mammals, we expect they should drive the majority of novel cross-species viral transmission, and likely bring zoonotic viruses into new regions. However, their ability to move rapidly might be attenuated by the other biotic constraints to species distributions (e.g., social behavior and food availability, which are unaddressed by the current approach). This uncertainty adds an important new dimension to ongoing debate about whether bats are unique....”

We also finally note that this question is of key interest to the authors, and is addressed in work currently in preparation, which explores the potential impacts of food web changes on bats and other zoonotic reservoirs as a result of climate and land use change.

The focus on RCP2.6 implies that countries will meet the Paris goal - which is perhaps unlikely - it seems a bit as if this scenario was chosen because it maximizes the results, which perhaps does a disservice to the analysis as a whole? Extended Data Fig. 3 might perhaps be usefully added to the main paper?

We appreciate the reviewer’s point here, and refer back to the first revision, where we addressed the same concern by adding summary statistics from RCP2.6 and RCP8.5 throughout the manuscript, aiming to show readers both a best-case (Paris aligned) scenario and a worst-case scenario for mitigation. At the time we also noted that the variation between the scenarios is actually quite small (e.g., from revision 1’s response: “For example, there is a 1% difference between RCP 2.6 and 8.5 in terms of viral sharing with limited dispersal, and a 1% difference without dispersal limits.”) and added our explanation in the main text that we “...present results for SSP1-RCP 2.6 in the main text because this scenario is most in line with the goals of the Paris Agreement to keep global warming “well below” 2°C.” While we appreciate the global challenges that stand in the way of achieving this scenario, we feel that this is still an important framing choice, particularly if our manuscript can be finalized in time to stay cited in the IPCC’s sixth Assessment Report, which will discuss climate change impacts on health in Paris-compatible scenarios extensively. Finally, as suggested by the reviewer here and in that original comment, we have now moved Extended Data Figure 3 to the main text of the paper, so that readers can now see the effect of SSP-RCP scenario choice more transparently; we appreciate this suggestion very much, and believe it will make the complexity of the study easier for readers to navigate.

It wasn’t always obvious if the numbers represented a change on the baseline (i.e. present-day) model run and when the results were the absolute numbers for scenario X. For instance Fig. 2C/D/E, are these the change in predicted encounters (vs the present) or the total predicted encounters? The same for Fig. 1. The key question is how much difference does climate change make on the baseline scenario. It seems from Extended Data Fig. 3D that the change is actually very small?

We have clarified the text on line 107 to explain that the first encounters are all a future event by definition: “We used projections of newly suitable habitat to identify where novel range overlap among currently non-overlapping species could happen *in the future* (hereafter *first encounters*).” There are no cases where raw future overlaps are used as a statistic. In extended data figure 3D - explained in the caption as “the overall change in sharing probability (connectance) across the viral sharing network between the present day and the future scenarios” - the values are given as absolute changes in raw connectance between present and future. Connectance is the number of species with predicted viral sharing out of all possible pairs; a change of 0.4% connectance in a network with millions of possible pairwise interactions represents a substantial real-world change. We also now explain this more up-front in the caption: “The overall change in sharing probability (connectance) across the viral sharing network between the present day and the future scenarios; absolute changes may appear small, but an 0.4% increase in connectivity is notable on the scale of millions of possible pairwise combinations of species. Change is positive across all scenarios....”

Referee #3 (Remarks to the Author)

I thank the authors for their considerable efforts and have only a few points to raise after this revision, mostly related to the major issues I highlighted last time (Referee #3). I said:

The authors note that the largest changes in species’ overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This

all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the “present” – the 1970-2000 mean (i.e. 1985). That is, if “less (climate change) is more” when it comes to novel species’ interaction, how do know that we’re not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth’s mean air temperature has risen by 0.7-0.8C.

There are two ways to resolve this: (1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5 reanalysis) (2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study. In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate”. If not, please comment on how this affects your modelling assumptions.

The authors have dealt with this by adding a new section: The effect of recent warming, which is a very welcome inclusion. However, the authors note in their response (and in the paper) that the results from this analysis cannot be compared directly with analysis based on WorldClim due to the fundamental differences in how these datasets are constructed (e.g., bias correction). The remaining question is why these datasets cannot be made equivalent. In other words, why can you not ‘adjust’ (i.e., bias correct) ERA5 to WorldClim? In essence, this allows you to “get” WorldClim from ERA5 for the real present. I understand that this would be demanding new analysis, and am not insisting on it, but I highlight here that such an adjustment would be possible – not least because ERA5 data are now available from 1950-present (which may not have been true at the time of the authors’ last revision, when I think data started in 1979). It would be useful to hear the authors’ response to this suggestion.

We appreciate this suggestion from the reviewer very much. Rather than bias correct one historical dataset to another (which is uncommon), another way to address the same concern would have been, from the start of the study, to use ERA5 as the historical climate product and to bias correct the future projections from each run of each global climate model for each scenario to the ERA5 historical product; this would make the recent past and future analyses more directly comparable. However, we feel this—which would require a total rerun of our entire multi-year study, start to finish—is beyond the present scope of our study, noting that the recent past analysis was added in the second revision at the suggestion of the reviewer (including the use of ERA5 specifically), and is largely secondary to the aim of our study, which is to anticipate potential future impacts. In future work, we hope to answer the exact questions the reviewer has nodded to here and with the suggestion for the ERA5 analysis now that ERA5 data back to 1950 has become available since we revised this manuscript, including the relative impacts of climate change in the present versus future, and more broadly the timing of potential abrupt changes over the coming century; for similar work, see: Trisos, Christopher H., Cory Merow, and Alex L. Pigot. “The projected timing of abrupt ecological disruption from climate change.” *Nature* 580.7804 (2020): 496-501.

Also, what may be a quick question to answer – the authors did not confirm whether *the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate?* If not, please comment on how this affects your modelling assumptions. I am not familiar with the species data (and may have missed key details in the paper), but can you confirm that the species’ presence records are for the same period as the 1970-2000 interval used to define the ‘present’ climate? I ask because, if they do not align, don’t you risk incorrectly specifying the climate-species dependency (because the climate has changed considerably since 1970)?

We appreciate this point and apologize for missing it in the previous revision. Our dataset does include post-2000 records; we now provide this information in the methods, as well as an explanation of how the pipeline proactively limits the potential downsides of using these data. Three points that are particularly salient:

1. Had we restricted to only the 1970-2000 sightings, the sample size would have been reduced by roughly ~70%, leading to poorer-calibrated predictions. Moreover, limiting sightings this way would also disproportionately eliminate species in ways that would be biased both geographically (small-ranged species) and phylogenetically (less easily documented species, particularly bats and rodents).
2. The vast majority of post-2000 sightings are occurrences within species’ baseline ranges, and are often more accurate than pre-2000 sightings due to both better GIS technology and more standard-

ized data curation and quality control. Moreover, multiple steps are taken in the model pipeline to prevent recent sightings from reducing model accuracy. In addition to the use of strict geographic and environmental outlier detection and removal (the Grubb test), we removed all records in the training data that occurred more than 10km beyond the boundary of IUCN expert range maps. This 10km buffer is rather conservative; Hurlbert and Jetz (2007 Species richness, hotspots, and the scale dependence of range maps in ecology and conservation) have suggested that expert map boundaries have a roughly 150km accuracy. Because we performed this spatial filtering step based on vetted expert maps, we felt it was justified to include both historical records and post-2000 records to increase sample size.

3. Finally, we note that if species geographic ranges or climatic niche are overestimated even after these layers of quality assurance, this will increase the current baseline rate of species geographic overlap, decreasing future first encounters, and biasing our study mostly towards more conservative predictions.

We agree with the reviewer that it is important to be particularly transparent about each of these points, and have now added this explanation to the methods:

“We downloaded Global Biodiversity Informatics Facility (GBIF: gbif.org) occurrence records for all mammals based on taxonomic names resolved by the IUCN Red List. Records spanned the years 1658 to 2019 and were intended to provide a comprehensive baseline of species distributional data for as many mammal species as possible. These records were filtered to those within a 10 km buffer of species’ IUCN range maps, and further filtered using a Grubb outlier test (see below), which is likely to eliminate most signal of climate-driven geographic range shifts from recent sightings. If any of that signal persists after these steps, there are two possible impacts on models: extending the geographic extent attributed to a species’ current range, and potentially, expanding the climatic envelope that a model identifies as suitable for its presence. Both of these would increase the predicted rate of overlap among species pairs in their current (modeled) distributions, and reduce the predicted number of first encounters, making our predictions more conservative.”

Finally, given the new analysis with ERA5, I suggest that the title should change from: *Climate change will drive novel cross-species viral transmission* to: *Climate change drives novel cross-species viral transmission* (because it is/has likely occurred already)

We appreciate this suggestion, but suggest this title change might be an over-reach, given that our study presents no original field data suggesting that novel viral transmission has occurred to date in a climate-linked setting.

Referee #4 (Remarks to the Author)

In my original comments, my main criticism was that the results of the ms remain somewhat speculative because of the lack of validation on the models ability to predict changes over time. I proposed that the authors would include such validation, or at least discuss more openly the limitations of their approach. I fully understand that such validation is not easy to do, partially because of sparsity of the appropriate data. I think that the authors have done a very good job in openly discussing the limitations of their approach and how the results can/cannot be interpreted. I find the manuscript highly interesting, and support its publication.

We thank the reviewer for their positive comments!

Referee #3 (Remarks to the Author):

Please see the attachment -- I have typed a response that includes a figure.

Re-review of “*Climate change will drive novel cross-species viral transmission*”, by Carlson et al.

I thank the authors for once again attending thoughtfully to the issues raised last time in my review (I was Reviewer #3). There is one issue I raise based on their response (in blue) to an issue I raised last time (in green).

The authors note that the largest changes in species' overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the “present” –the 1970-2000 mean (i.e. 1985). That is, if “less (climate change) is more” when it comes to novel species' interaction, how do we know that we're not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth's mean air temperature has risen by 0.7-0.8C. There are two ways to resolve this: (1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5 reanalysis) (2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study. In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate”. If not, please comment on how this affects your modelling assumptions.

*We appreciate this suggestion from the reviewer very much. Rather than bias correct one historical dataset to another (which is uncommon), another way to address the same concern would have been, from the start of the study, to use ERA5 as the historical climate product and to bias correct the future projections from each run of each global climate model for each scenario to the ERA5 historical product; this would make the recent past and future analyses more directly comparable. However, we feel this—which would require a total rerun of our entire multi-year study, start to finish—is beyond the present scope of our study, noting that the recent past analysis was added in the second revision at the suggestion of the reviewer (including the use of ERA5 specifically), and is largely secondary to the aim of our study, which is to anticipate potential future impacts. In future work, we hope to answer the exact questions the reviewer has nodded to here and with the suggestion for the ERA5 analysis now that ERA5 data back to 1950 has become available since we revised this manuscript, including the relative impacts of climate change in the present versus future, and more broadly the timing of potential abrupt changes over the coming century; for similar work, see: Trisos, Christopher H., Cory Merow, and Alex L. Pigot. “The projected timing of abrupt ecological disruption from climate change.” *Nature* 580.7804 (2020): 496-501.*

I am very sympathetic to the position expressed by the authors, but contend that re-running the analysis with ERA5 would not be beyond the “scope” of the study; it would require a big investment of effort in a piece of work that has already taken a long time, but the scope would be the same: assessing the impact of climate change on the opportunity for novel cross-species virus sharing events. Is this investment of time worth it? As an interested reader (not encumbered by the substantial new work

that would be needed), I would say yes, it is. Why? Because I do not think the paper can presently assert that future climate change will lead to an increase in virus sharing.

To be clear for the benefit of the editor, I understand your key conclusion to be that, relative to a 1970-2000 baseline, relatively little climate change (i.e., the RCP2.6 pathway) will lead to the greatest increase in virus sharing probability. But, as pointed out earlier (and explored in your new analysis), the climate has already changed since the preindustrial (by about 1C). So, could it be that future climate change does not lead to an increase in virus sharing at all – and that instead the changed climate of *now* represents the ideal for new virus sharing? In the plot below, I show your results for changes in range due to climate (y-axis) vs. global mean warming amount under the SSP-RCPs (digitised approximately from the IPCC AR6 for the period 2061-2080). The y-axis intercept (1C) is the amount of warming we have currently experienced and, clearly, the strong linearity in your results would indicate “range change due to climate” is likely higher now than it would be in the future under the RCP2.6 scenario.

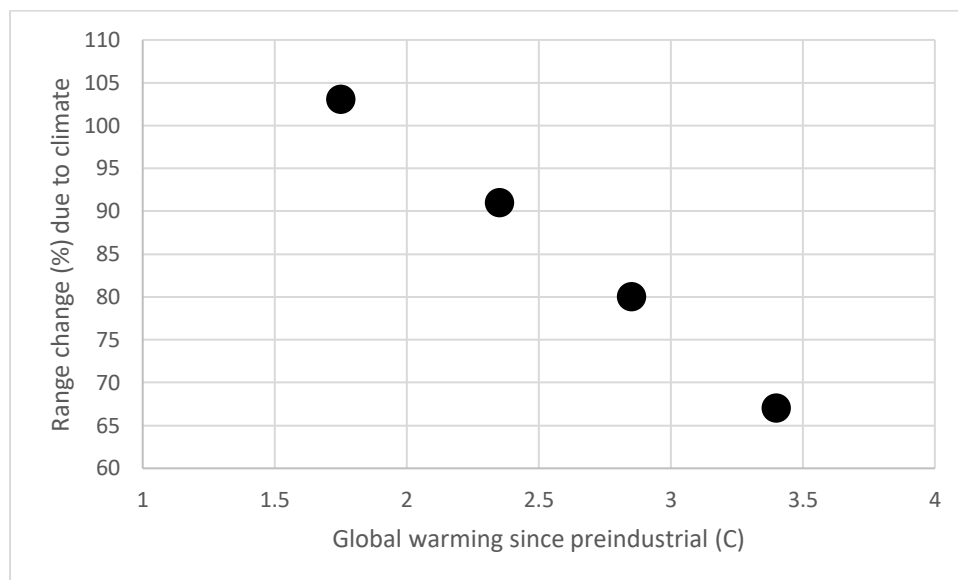


Figure 1. See text for details. Y-axis coordinates taken from Figure 2A in the submitted manuscript.

Because it's not an apples vs. apples comparison, your new analysis with ERA5 does not allow us to conclude if the climate of now is “better” or “worse” than RCP2.6 for novel virus transmission. This appears to be a critical point, because, in short, we cannot rule out that future climate change may be “good” from a virus transmission point of view (if we are *already* at the peak). Re-running your analysis with ERA5 would, I think, address this question very neatly. It would tidy up other issues in the manuscript too (the misalignment in time between the species baseline and your climate baseline).

This point also gets to the heart of my comment about your title. Your analysis firmly outlines that changes in climate drive changes in virus transmission (from a mechanistic perspective). Inclusion of “will” implies (I think), that it is the future changes that are important, when that may not be the case.

To summarise, I sympathize with the amount of time and effort invested by the authors. However, this is clearly a very important study that will generate lots of interest, and if they cannot explain why my concerns are unfounded, I request that they please make this last push – remove any ambiguity by re-running the analysis with the best tool for the job (ERA5).

Response to reviewers: revision 4

Carlson *et al.*, submission to *Nature*

April 2021

Dear editor,

Enclosed is the fourth revision of our study “Climate change drives novel cross-species viral transmission,” following an appeal to extend our third revision and add an analysis requested by the editor and reviewer #3. Before including the reviewer’s comments in response to revision 3, we briefly summarize progress up to the present point.

Three out of the four reviewers of previous versions have decided our revisions are satisfactory. The fourth, Reviewer #3, has asked us to more carefully engage with a counter-hypothesis revolving around the timing of climate change impacts. (Reviewer 3 statement from revision 2: “Most of the points I raise are minor, but there is one issue which I think might be more substantial. If they can address it (or clarify any misunderstanding), I would be happy to recommend their work for publication.”)

In our study, we present the hypothesis that in the future, when species meet for the first time due to climate-linked geographic range shifts, they will have the opportunity to exchange new viruses for the first time. Given that we find a negative relationship between the amount of projected warming and the magnitude of this outcome, Reviewer #3 asks us to consider the possibility that much of these impacts might already have happened (i.e., if futures with lower increases in global warming have worse impacts, might this also be the case for global warming up to the present?):

Reviewer 3 statement from revision 2: Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the “present” – the 1970-2000 mean (i.e. 1985). That is, if “less (climate change) is more” when it comes to novel species’ interaction, how do know that we’re not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth’s mean air temperature has risen by 0.7-0.8C.

Reviewer 3 statement from revision 3: ...I suggest that the title should change from: *Climate change will drive novel cross-species viral transmission* to: *Climate change drives novel cross-species viral transmission* (because it is/has likely occurred already)

In revision 2, we added a new analysis with a separate climate product (ERA5, specifically suggested by the reviewer) that allowed us to compare a recent past (1981-1995) to the present (1991-2015). We found that substantial climate impacts might already have happened, but noted that the number of first encounters could not be compared between this analysis and our main analysis, as they used different climate baselines (ERA5, at the time when we started our revisions, had yet to fully release their dataset for the second half of the 20th century). The reviewer felt that this was insufficient to address their concerns; their response to revision 3 is included below, minus a figure they generated by reproducing our results.

I thank the authors for once again attending thoughtfully to the issues raised last time in my review (I was Reviewer 3). There is one issue I raise based on their response to n [sic] issue I raised last time.

[reviewer comments from revision 3] The authors note that the largest changes in species’ overlap and opportunities for viral shedding are expected under RCP2.6. Greater warming scenarios (RCP8.5) are expected to see more modest changes because it is hypothesised that the rate of climate change may be too rapid for species to adjust. This all sounds very plausible, but it begs the question whether we may have already seen bigger changes since the period designated by the authors as the “present” –the

1970-2000 mean (i.e. 1985). That is, if “less (climate change) is more” when it comes to novel species’ interaction, how do know that we’re not currently living in a climate more conducive to viral sharing than projected by 2070 under RCP2.6? Note that since 1985, the Earth’s mean air temperature has risen by 0.7-0.8C. There are two ways to resolve this: (1) Run your analysis on Bioclim variables representative of the current climate (perhaps using ERA5 reanalysis) (2) Explicitly highlight this caveat to your research in the text. It is important because it does limit what can be concluded from your study. In dealing with the above (noting that the climate has already changed), please clarify whether the biodiversity data can be considered as representative of the same 1970-2000 period that characterises the present climate”. If not, please comment on how this affects your modelling assumptions.

[author comments from revision 3] We appreciate this suggestion from the reviewer very much. Rather than bias correct one historical dataset to another (which is uncommon), another way to address the same concern would have been, from the start of the study, to use ERA5 as the historical climate product and to bias correct the future projections from each run of each global climate model for each scenario to the ERA5 historical product; this would make the recent past and future analyses more directly comparable. However, we feel this—which would require a total rerun of our entire multi-year study, start to finish—is beyond the present scope of our study, noting that the recent past analysis was added in the second revision at the suggestion of the reviewer (including the use of ERA5 specifically), and is largely secondary to the aim of our study, which is to anticipate potential future impacts. In future work, we hope to answer the exact questions the reviewer has nodded to here and with the suggestion for the ERA5 analysis now that ERA5 data back to 1950 has become available since we revised this manuscript, including the relative impacts of climate change in the present versus future, and more broadly the timing of potential abrupt changes over the coming century; for similar work, see: Trisos, Christopher H., Cory Merow, and Alex L. Pigot. “The projected timing of abrupt ecological disruption from climate change.” *Nature* 580.7804 (2020): 496-501.

[response from reviewer 3]: I am very sympathetic to the position expressed by the authors, but contend that re-running the analysis with ERA5 would not be beyond the “scope” of the study; it would require a big investment of effort in a piece of work that has already taken a long time, but the scope would be the same: assessing the impact of climate change on the opportunity for novel cross-species virus sharing events. Is this investment of time worth it? As an interested reader (not encumbered by the substantial new work that would be needed), I would say yes, it is. Why? Because I do not think the paper can presently assert that future climate change will lead to an increase in virus sharing.

To be clear for the benefit of the editor, I understand your key conclusion to be that, relative to a 1970-2000 baseline, relatively little climate change (i.e., the RCP2.6 pathway) will lead to the greatest increase in virus sharing probability. But, as pointed out earlier (and explored in your new analysis), the climate has already changed since the preindustrial (by about 1C). So, could it be that future climate change does not lead to an increase in virus sharing at all – and that instead the changed climate of now represents the ideal for new virus sharing? In the plot below, I show your results for changes in range due to climate (y-axis) vs. global mean warming amount under the SSP-RCPs (digitised approximately from the IPCC AR6 for the period 2061-2080). The y-axis intercept (1C) is the amount of warming we have currently experienced and, clearly, the strong linearity in your results would indicate “range change due to climate” is likely higher now than it would be in the future under the RCP2.6 scenario.

Because it’s not an apples vs. apples comparison, your new analysis with ERA5 does not allow us to conclude if the climate of now is “better” or “worse” than RCP2.6 for novel virus transmission. This appears to be a critical point, because, in short, we cannot rule out that future climate change may be “good” from a virus transmission point of view (if we are already at the peak). Re-running your analysis with ERA5 would, I think, address this question very neatly. It would tidy up other issues in the manuscript too (the misalignment in time between the species baseline and your climate baseline).

This point also gets to the heart of my comment about your title. Your analysis firmly outlines that changes in climate drive changes in virus transmission (from a mechanistic perspective). Inclusion of “will” implies (I think), that it is the future changes that are important, when that may not be the case.

To summarise, I sympathize with the amount of time and effort invested by the authors. However, this is clearly a very important study that will generate lots of interest, and if they cannot explain why my concerns are unfounded, I request that they please make this last push – remove any ambiguity by re-running the analysis with the best tool for the job (ERA5).

First, we want to express a sincere overall thanks to the reviewer for encouraging us to consider this aspect of our study again. After working on this study for almost five years, it’s been a delight to further explore this system, and we think that presenting these results does, as the reviewer suggested, make our study even more interesting.

Up front: as we explained in our appeal, while we agree on overall mechanism (“**Your analysis firmly outlines that changes in climate drive changes in virus transmission (from a mechanistic perspective).**”), we disagree with one key aspect of the logic presented by the reviewer. Cross-species transmission operates as a cumulative process (i.e., even if some species lose some geographic overlaps in the future, we don’t simulate that they will stop sharing viruses previously acquired from each other; this might happen in very rare cases of dead-end transmission, but these will be the exception to the rule, and are beyond our ability to predict with the viral sharing model.) No matter how much warming happens, some species will move, new species pairs will still meet for the first time, and they will still share new viruses. Therefore, in any scenario, we should expect that future climate change will lead to *some* increase in virus sharing among species that meet for the first time. However, it is possible that the *rate* of this process has peaked (i.e., that much of the possible climate-linked cross-species transmission has happened, but not all of it).

This is the hypothesis we’ve set out to test as encouraged by the reviewer and editor; in our appeal we noted that “our guess is that this is unlikely: this would mean that species have entered the majority of new habitats they ever will by this point in the century.” **To our surprise, our guess was incorrect, and we present that fully transparently in our revision.**

To replace the smaller ERA5 analysis we produced previously, we executed the new analysis by using the CHELSA v2.1 dataset, which has also only just released CMIP6 predictions in the last year. This was suggested to us by other climate scientist collaborators, who noted that if we were to work with ERA5 outputs, it would be best for us to downscale and bias correct the future CMIP6 projections ourselves, adding several more months to an analysis that has already added half a year. CHELSA works just as well for our purposes, and has allowed us to once again compare a recent past baseline (1981-2010) and three futures including a “present” centered on 2025 (2011-2040), as well as 2041-2070, and 2071-2100. As requested, we’ve executed the entire analysis for all five GCMs and three SSP-RCPs that were available (15 climate futures). We decided to treat this as an addition to our already extensive results, rather than a replacement, because (1) our existing climate methodology had already undergone several years of review, and (2) allowed us to present a substantially greater total range of future climate uncertainty (9 GCMs x 4 SSP-RCPs = 36 climate futures, including the RCP4.5 scenario for intermediate levels of greenhouse gas emissions). By showing that the entire phenomenon can be reproduced in two distinct climate datasets, we also implicitly show a third dimension of how our results are robust to climate model uncertainty. In total, this makes our study one of the largest species distribution modeling ensemble studies ever attempted.

Major points of the new analysis are that

- The majority of first encounters among species are indeed projected to occur by the “present”/near term (2025), counter to our expectations. (This makes sense, in retrospect: species’ ranges will achieve a few pixels’ worth of overlap much earlier than when the majority of range change is projected to occur.)
- This confirms that the majority of opportunities for novel viral sharing may already exist in the world today (or in the near-term future); however, species will continue to meet for the first time up to 2100. We show this in a new Figure 6 as part of a new section of the main text (elevating this analysis from the Methods, where the ERA5 analysis had been added in a previous revision).
- Differences in the number of first encounters between time points greatly exceeded differences in first encounters due to climate model uncertainty and climate scenario differences (i.e., differences among climate models and differences among SSP-RCPs); but they were far less than the difference in number of first encounters between the scenarios with and without dispersal limitation. This suggests, perhaps unsurprisingly, that how fast species move will matter more to the timing of first encounters than how fast suitable habitat moves.
- In the dispersal constrained scenario, first encounters also plateau over time *much less* steeply, indicating that a much smaller percentage of species might already have met for the first time when dispersal constraints are considered compared to when only climate change is considered. Given

that our dispersal velocity estimates are a *maximum* (a biological upper limit), we therefore talk about these results as describing the potential timing of opportunity, noting that if species are moving even slower, even more opportunities may be unrealized in the present day.

- The geographic distribution of first encounters was essentially constant over time, reinforcing our existing analyses as robust (and we added two extended data figures to show this).
- The implications of our results for the timing of viral sharing are less straightforward, and we have approached them with caution. The Albery model of viral sharing we use here is essentially meant to describe the host-virus network at equilibrium, and so when we predict viral sharing associated with first encounters, the first encounters are anchored to a particular time point, but the viral sharing is an equilibrium projection that eventually results from that change. Over a century, much or all of the possible viral sharing driven by climate change could easily happen; over a 30-year interval, we feel much less comfortable making the assertion that it will (which we would have to if we were comparing viral sharing between two time points). Predicting the timing of specific jumps is beyond the limits of disease ecology as a science today, though it might be possible in the future to predict the timing of host jumps in aggregate as a coarse average, if biological invasions can be used to train such a model (some work pointing towards this possibility: Schatz, Annakate M., and Andrew W. Park. “Host and parasite traits predict cross-species parasite acquisition by introduced mammals.” *Proceedings of the Royal Society B* 288.1950 (2021): 20210341.). However, our model cannot make that prediction.
- Given concerns that predicting the timing of viral sharing this precisely might be misleading or simply incorrect (if the lag between the first contact and viral establishment is longer than, on average, 15 years), we felt it would be inappropriate to plot the estimated novel viral sharing that accompanies each time point (and these plots would contain largely identical information about spatiotemporal patterns to the ones we include in Figure 6 and ED Figures 8 and 9). However, we interpret in the text that species have indeed probably been given the majority of climate-related *opportunity* to share viruses by the present / near-term (conditional on dispersal rates). The reviewer even emphasizes this is the main point of our study: [assessing the impact of climate change on the opportunity for novel cross-species virus sharing events](#).
- Finally, we have also, returning to the second revision, adopted the reviewer’s suggested new title (“Climate change *drives* cross-species viral transmission”), given our new results and the phocine distemper study we discuss. We still feel slightly hesitant that this might mislead readers to believe we are documenting evidence those changes have happened with viral surveillance data, but we have written the results of the timing analysis in such a way as to try to avoid this. One alternate title that may be more indisputable which we could use is “Climate change increases cross-species viral transmission risk,” and we will defer to the reviewer and editor if this might be more appropriate.

In summary, this revision contains a new analysis that entirely reproduces our study’s results from scratch, includes a new main text section that shows the point the reviewer asked us to verify, and includes one new main text figure and two extended data figures. We appreciate the reviewer’s push for us to engage fully in this point, and we believe it makes the paper much stronger so that we can, confidently, say that the most most of the climate-driven opportunities for virus sharing due to climate change may already have happened by the present / near-term. We are excited to resubmit and hope that the reviewer will feel that with these new results, as indicated in the second revision, our manuscript is finally ready to be recommended for publication.

Reviewer Reports on the Fourth Revision:

Referees' comments:

Referee #3 (Remarks to the Author):

I applaud the authors for getting to this point. I think the study has benefitted substantially from the new analysis in this latest revision.

I am now happy to recommend publication, with only a few minor comments to address:

1) Abstract -- I think the results from the "timing of ecological opportunity" (i.e., the new work) should be mentioned here. For example, you could say something like:

"Counter to expectations, holding warming under 2°C within the century does not reduce new viral sharing, due to greater potential range expansions. Our results even suggest that novel viral sharing may be close to its peak in the current climate (already ~1C above preindustrial)..."

I leave the wording up to the authors, but I think their results here are about more than just the 2C limit. Including this point would allow readers to interpret the findings in context of the 1.5C limit, and would also strengthen the case for needing to urgently strengthen surveillance.

2) With reference to the title, I think either:

(a) "Climate change drives novel cross-species viral transmission"

or:

(b) "Climate change increases cross-species viral transmission risk"

is acceptable. I hope the editor will comment if they think the current title (a -- which I suggested in the last round of review) should be replaced by b).

Congratulations to the authors.

Signed: Tom Matthews (King's College London).

Author Rebuttals to Fourth Revision:

Referees' comments:

Referee #3 (Remarks to the Author):

I applaud the authors for getting to this point. I think the study has benefitted substantially from the new analysis in this latest revision.

I am now happy to recommend publication, with only a few minor comments to address:

1) Abstract -- I think the results from the "timing of ecological opportunity" (i.e., the new work) should be mentioned here. For example, you could say something like:

"Counter to expectations, holding warming under 2°C within the century does not reduce new viral sharing, due to greater potential range expansions. Our results even suggest that novel viral sharing may be close to its peak in the current climate (already ~1C above preindustrial)..."

I leave the wording up to the authors, but I think their results here are about more than just the 2C limit. Including this point would allow readers to interpret the findings in context of the 1.5C limit, and would also strengthen the case for needing to urgently strengthen surveillance.

We appreciate this point. Given the requested substantial reduction in abstract length, we've balanced this with a fairly short addition:

"Surprisingly, we find that this ecological transition may already be underway, and holding warming under 2°C within the century will not reduce future viral sharing."

2) With reference to the title, I think either:

(a) "Climate change drives novel cross-species viral transmission"

or:

(b) "Climate change increases cross-species viral transmission risk"

is acceptable. I hope the editor will comment if they think the current title (a -- which I suggested in the last round of review) should be replaced by b).

As requested by the editor, we have used “Climate change increases cross-species viral transmission risk.”

Congratulations to the authors.

Signed: Tom Matthews (King's College London).