


# A strategy to assess spillover risk of bat SARS-related coronaviruses in Southeast Asia

Cecilia A. Sánchez <sup>1</sup>, Hongying Li <sup>1</sup>, Kendra L. Phelps <sup>1</sup>, Carlos Zambrana-Torrel <sup>2</sup>, Lin-Fa Wang <sup>3</sup>, Peng Zhou<sup>4</sup>, Zheng-Li Shi <sup>4</sup>, Kevin J. Olival <sup>1</sup> & Peter Daszak <sup>1</sup> 

Emerging diseases caused by coronaviruses of likely bat origin (e.g., SARS, MERS, SADS, COVID-19) have disrupted global health and economies for two decades. Evidence suggests that some bat SARS-related coronaviruses (SARSr-CoVs) could infect people directly, and that their spillover is more frequent than previously recognized. Each zoonotic spillover of a novel virus represents an opportunity for evolutionary adaptation and further spread; therefore, quantifying the extent of this spillover may help target prevention programs. We derive current range distributions for known bat SARSr-CoV hosts and quantify their overlap with human populations. We then use probabilistic risk assessment and data on human-bat contact, human viral seroprevalence, and antibody duration to estimate that a median of 66,280 people (95% CI: 65,351–67,131) are infected with SARSr-CoVs annually in Southeast Asia. These data on the geography and scale of spillover can be used to target surveillance and prevention programs for potential future bat-CoV emergence.

<sup>1</sup>EcoHealth Alliance, New York, NY, USA. <sup>2</sup>Department of Environmental Science and Policy, George Mason University, Fairfax, VA, USA. <sup>3</sup>Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore. <sup>4</sup>Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China. email: [daszak@ecohealthalliance.org](mailto:daszak@ecohealthalliance.org)

Emerging coronaviruses (CoVs) of wildlife origin have significantly disrupted global health security and economies during the last two decades<sup>1,2</sup>. Severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) CoVs caused significant human morbidity and mortality in 2002 and 2012 respectively<sup>3,4</sup>. Swine Acute Diarrheal Syndrome CoV caused substantial mortality in pigs in southern China during 2016 and 2019<sup>5,6</sup>. The emergence of SARS-CoV-2 in 2019 led to the current COVID-19 pandemic that has caused millions of cases and deaths, with economic loss likely to be in the tens of trillions of US dollars<sup>2,7</sup>. Efforts to increase preparedness and improve surveillance for emerging coronaviruses therefore represent a priority for global health programs<sup>8</sup>.

Phylogenetic analysis suggests that SARS-CoV, MERS-CoV, SADS-CoV, and SARS-CoV-2 originate within CoV lineages from bat reservoir hosts<sup>5,9–11</sup>. The initial spillovers of SARS and MERS into human populations are thought to have occurred via intermediate hosts (palm civets and dromedary camels, respectively<sup>12,13</sup>). However, the role of civets in the emergence of SARS is uncertain, and other bat SARSr-CoVs can directly infect human cells, including airway epithelial cells, and thus have potential to spill over directly from bats to humans<sup>14–16</sup>. In support of this idea, serological evidence of prior infection with SARSr-CoVs was found in communities living near bat populations in China prior to the emergence of COVID-19, including in people who reported no contact with SARSr-CoV intermediate hosts<sup>17,18</sup>. Direct bat-to-human spillover events may occur more frequently than has been reported, but go unrecognized because they cause mild symptoms, cause symptoms similar to those of other infections, result in small case numbers, or lack sustained chains of human-to-human transmission. However, every wildlife-to-human spillover event represents an opportunity for viral adaptation that could permit human-to-human spread<sup>19–22</sup>. Estimating the extent of these undetected spillovers could therefore be important for identifying the likelihood of future epidemics or pandemics.

Surveys of bats in China have revealed high diversity of SARSr-CoVs, and often high infection prevalence (5–10%) in rhinolophid and hipposiderid species that are widely distributed and abundant, with varying resilience to habitat perturbation and many synanthropic species (having contact and often interactions with human populations)<sup>23–25</sup>. Many of the bat species and genera known to harbor these  $\beta$ -CoVs occur in Southeast Asia, a hotspot of bat diversity with 441 species reported, 115 (around a quarter) of which are rhinolophids or hipposiderids<sup>26</sup>. Diversity of SARSr-CoVs is also likely high in this region<sup>23</sup>, but may be underestimated as CoV research effort in China appears to have been far more intense than in nearby Southeast Asian or South Asian countries<sup>27</sup> (see Methods and Supplementary Fig. 1). Furthermore, many of these less well-sampled countries are undergoing dynamic social and environmental changes correlated with zoonotic emergence (e.g., rapid human population growth, movement of rural residents to urban centers, extensive wildlife farming and trade, rapid land conversion from forested habitats to agricultural land), and thus might represent hitherto unreported hotspots for coronavirus spillover<sup>23,28–33</sup>.

In this study, we use host distribution modeling as well as human behavioral and epidemiological data to estimate the geographic distribution of SARSr-CoV bat hosts, and the likely rate of zoonotic spillover in China, South and Southeast Asia. Our results provide detailed estimates of the distribution of bat hosts of SARSr-CoVs and suggest that the magnitude of SARSr-CoV spillover from bats to humans may be substantially underestimated. Our approach provides proof of concept for assessing the rate of zoonotic spillover and identifies key geographic areas that can be prioritized for targeted surveillance of animals and humans. Given the challenges of identifying the origins of

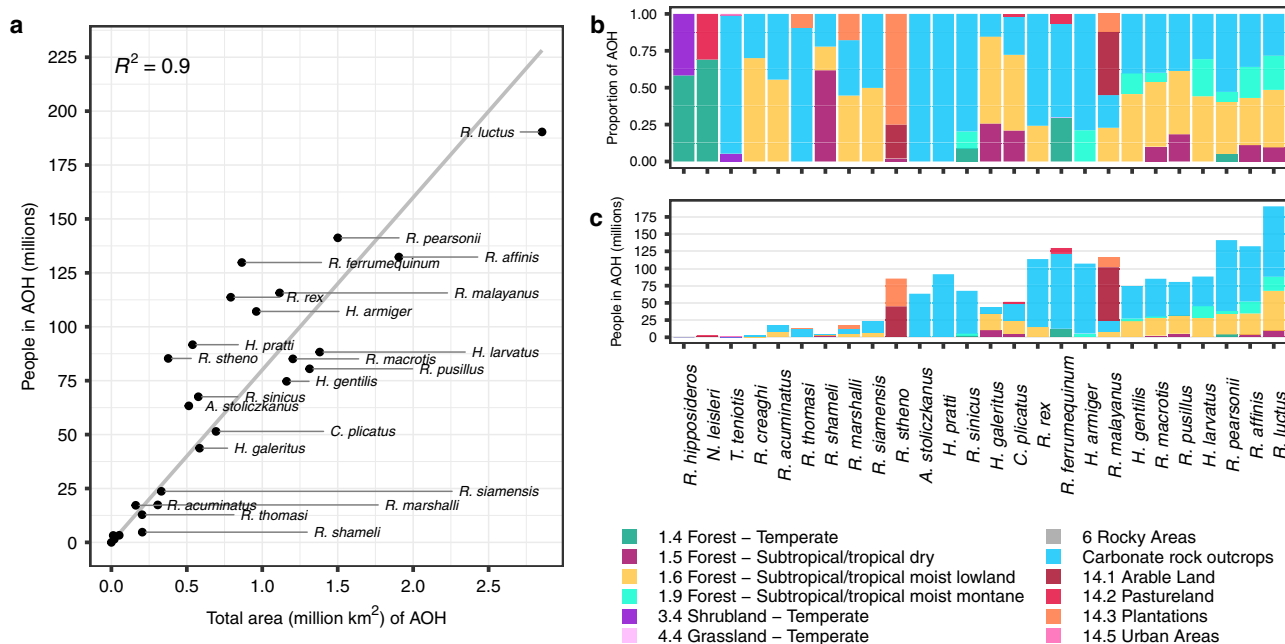
COVID-19 and pathways by which SARS-CoV-2 spilled over to people<sup>34,35</sup>, our results may also aid efforts to identify the geographic sites where spillover first occurred.

## Results

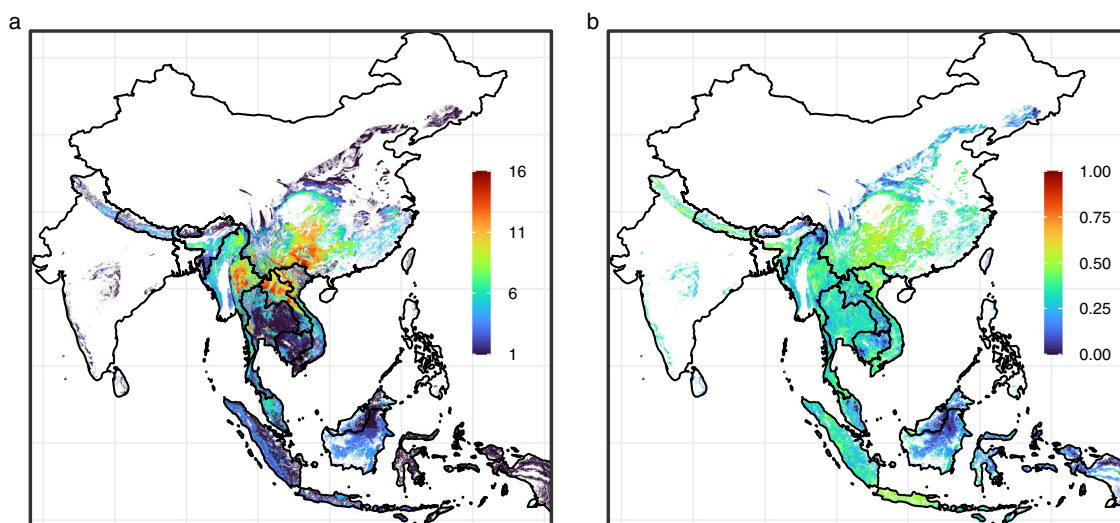
We assembled a list of 26 known SARSr-CoV bat host species that occur in our geographic region of interest (a broad region including parts of South Asia, China and Southeast Asian countries; see Methods for a list of countries and administrative regions included in the analysis). Host species were mainly members of *Rhinolophidae* and *Hipposideridae* families, but also included two members of *Molossidae* and one member of *Vespertilionidae* (Supplementary Table 1). We derived the area of habitat (AOH)<sup>36</sup> for each species by refining the IUCN geographic range of each species according to habitat suitability, elevation limits, and the boundaries of our region of interest. Removing unsuitable areas within the IUCN geographic ranges greatly reduced the size of species distributions. For example, the reduction in area from the original IUCN range to the more refined AOH ranged from 42% for *Rhinolophus malayanus* to almost 100% for *R. hipposideros*, with a median of 65% reduction across all species (Supplementary Fig. 2). The reduction in area was 55% for *R. rex*, the only species in our list assessed as endangered by the IUCN Red List.

We validated species AOHs using cleaned occurrence records downloaded from the Global Biodiversity Information Facility (GBIF; see Methods for details of data cleaning and validation). After data cleaning, no occurrence records remained for *R. hipposideros*, while 1–621 (median = 42) occurrence points remained for all other species (Supplementary Table 2). In validating each species' AOH with GBIF occurrence points, we found among species with  $\geq 1$  occurrence point, the median percent of points (buffered by 5 km) that overlapped a species' AOH was 62% (Supplementary Table 2). Among species with at least 40 occurrence points ( $n = 15$ ), the median overlap was 76%. Overlap was  $\geq 80\%$  for five species: *R. luctus* (63/69, 91%), *R. creaghi* (37/41, 90%), *R. malayanus* (45/52, 87%), *R. shameli* (35/42, 83%), and *Hipposideros galeritus* (64/80, 80%).

The size of individual species AOHs varied widely, but generally, the largest AOHs encompassed the most people (Fig. 1a). For example, the AOH of *R. luctus* was the largest of all species, covering  $\sim 2.9$  million km<sup>2</sup>, and encompassed  $\sim 190$  million people, the most of any species. Only three other species had AOHs encompassing more than 125 million people (*R. pearsonii*, *R. affinis*, and *R. ferrumequinum*), with the AOHs of *R. affinis* and *R. ferrumequinum* both encompassing  $\sim 130$  million people despite the AOH of *R. ferrumequinum* being less than half the size of that of *R. affinis*. Fewer than 5 million people live in the AOHs of *R. hipposideros*, *Nyctalus leisleri*, *Tadarida teniotis*, *R. creaghi*, and *R. shameli*. Two species had AOHs with more limited area but relatively high human population density: *N. leisleri* ( $\sim 262$  people/km<sup>2</sup>) and *R. stheno* ( $\sim 226$  people/km<sup>2</sup>; Fig. 1a, Supplementary Fig. 3). Within species AOHs, forest habitats and carbonate (limestone) rock outcrops—used as a proxy for cave distribution—typically comprised the largest proportion of suitable habitat (Fig. 1b). Four species were highly reliant on carbonate rock outcrops (i.e., their AOHs were comprised of  $>90\%$  of this habitat type): *T. teniotis*, *R. thomasi*, *Aselliscus stoliczkanus*, and *H. pratti*. Only *R. stheno* and *R. malayanus* had AOHs comprising  $>50\%$  artificial habitats, including plantations and arable land. Examining the number of people in the habitat type of each species revealed that people tended to be over-represented in carbonate rock outcrop areas compared to forest areas, relative to the size of these habitats (Fig. 1c). For example, carbonate rock outcrops comprised 28%, by size, of the AOH of *R. luctus*, but



**Fig. 1 Relationships among area of habitat (AOH) size, number of people, and habitat proportions.** **a** Scatterplot showing the total number of people living in each AOH versus the total area, for each SARSr-CoV bat host species. A best-fit line was fit through the origin, for which the  $R^2$  is displayed. The four unlabeled species at the bottom left corner are (left to right): *R. hipposideros*, *N. leisleri*, *T. teniotis*, and *R. creaghi*. **b** Proportion of each habitat type (by area) within species AOHs. **c** People living in the AOH of each species, separated by habitat type. For **b** and **c**, species are listed in order of increasing AOH size, and the species order is the same in both panels. Genus abbreviations in all panels: A. = *Aselliscus*, C. = *Chaerephon*, H. = *Hipposideros*, N. = *Nyctalus*, R. = *Rhinolophus*, T. = *Tadarida*.



**Fig. 2 Hotspots of SARSr-CoV bat host species richness and human overlap in Southeast Asia.** **a** Species richness of SARSr-CoV bat host species in Southeast Asia, created by overlaying area of habitat maps for all 26 SARSr-CoV bat host species known for this region. **b** Relative bat-human overlap: bat host species richness multiplied by human population count. Values were  $\ln(x + 1)$  transformed and then normalized to a 0–1 scale. For both panels, redder colors indicate larger values and bluer colors indicate smaller values.

represented 54% in terms of people living in this habitat type; similarly, these percentages were 53% and 73% for *R. pearsonii*. Arable land also appeared overrepresented in terms of the number of people living in this habitat versus its proportion by size, as observed for *R. stheno* and *R. malayanus*.

The consensus area of all SARSr-CoV bat host species, created by overlaying the 26 species AOHs, comprised ~5.1 million km<sup>2</sup>. We calculated that ~499 million people live within this consensus area, which covered most of Lao PDR, Cambodia, Thailand, Vietnam, Nepal, Bhutan, peninsular Malaysia, Myanmar,

southeast China, and the western islands of Indonesia (Fig. 2a). Bat species distribution was patchier in India, Sri Lanka, East Malaysia, and the Philippines. Species richness ranged from 1–16 species, with the highest richness of SARSr-CoV bat host species in southern China, eastern Myanmar, and northern Lao PDR (Fig. 2a). When we visualized areas with both high host richness and large human populations (bat-human overlap), southern China remained a hotspot, while other areas emerged as important because of their high human population sizes (e.g., Java, parts of northern India, parts of Myanmar; Fig. 2b).

After calculating the number of people living in the consensus area of SARSr-CoV bat host species, we incorporated data from the literature on human-bat contacts, viral seroprevalence among humans reporting bat contact, and human SARS antibody duration to estimate spillover risk of SARSr-CoVs in Southeast Asia (see Methods for details). For this analysis, we define spillover risk as the annual number of people infected by SARSr-CoVs via bat-to human transmission events. This definition draws from the work of Hosseini and colleagues<sup>37</sup>, who define risk as equal to hazard  $\times$  exposure  $\times$  vulnerability, where a hazard is a potential source of harm from a microbe, exposure is the likelihood of contact between humans and hazards, and vulnerability is the chance of a hazard causing harm, given exposure. Here, the hazard is a bat belonging to a species known to host SARSr-CoVs, exposure is the likelihood of contact between a person and a bat, and vulnerability is the chance that a human-bat contact event results in a human infection that produces detectable antibodies. Note that we do not consider whether a human infection will lead to sickness or an outbreak, because these outcomes depend on a range of other factors including viral genotype and phenotype, host susceptibility (e.g., age, predisposing conditions), and population-level factors that would affect the ability of a pathogen to spread; these are outside the scope of the current study. We estimated that within the consensus area of SARSr-CoV bat host species, a median of 66,280 people (95% confidence interval for the median: 65,351–67,131; overall range: 1–54,905,544; interquartile range: 3,250–519,557) are infected with SARSr-CoVs annually in Southeast Asia (Fig. 3a).

Sensitivity analyses (see Methods) indicated that two parameters—the probability of human contact with a bat ( $P_{contact}$ ), and the probability of antibody detection given contact with a bat ( $P_{detect}$ )—primarily contributed to variance in the outcome (Fig. 3b, c). Specifically,  $P_{detect}$  contributed most to variance in the outcome (first-order sensitivity index: 0.486; 95% CI: 0.475–0.496; total-order sensitivity index: 0.873; 95% CI: 0.864–0.881), while  $P_{contact}$  contributed moderately to variance in the outcome (first-order sensitivity index: 0.125; 95% CI: 0.117–0.134; total-order sensitivity index: 0.509; 95% CI: 0.499–0.519). We therefore explored how refitting the distributions for these parameters (Supplementary Fig. 4), after excluding the highest estimates of  $P_{contact}$  and  $P_{detect}$  found in our literature searches, changed our estimates of spillover (Fig. 3a). Refitting only the  $P_{contact}$  distribution lowered the median number of people infected to 53,290 (95% CI: 52,621–53,957; range: 1–33,119,404; interquartile range: 2,817–375,452), while refitting only the  $P_{detect}$  distribution lowered the median to 49,599 people (95% CI: 49,151–50,095; range: 1–20,694,204; interquartile range: 6,262–234,140). When both distributions were refit, the median number of people infected was 38,910 (95% CI: 38,551–39,268; range: 1–11,653,741; interquartile range: 5,538–163,053).

## Discussion

Our paper reports an analytical framework to assess SARSr-CoV spillover risk in a region that includes the site of the first detected spillover of SARS-CoV, and likely of SARS-CoV-2. We first provide detailed maps of SARSr-CoV bat host richness and bat-human overlap in South and Southeast Asia. Using information gathered from the literature on human-bat contacts, viral seroprevalence among humans with bat contact, and human SARS antibody duration, we then provide the first estimates, to our knowledge, of the number of people infected by SARSr-CoVs annually. The analytical framework, maps, and other results produced here may assist public health measures by identifying regions for targeted surveillance and early detection of ongoing spillover events, for viral discovery programs to identify novel

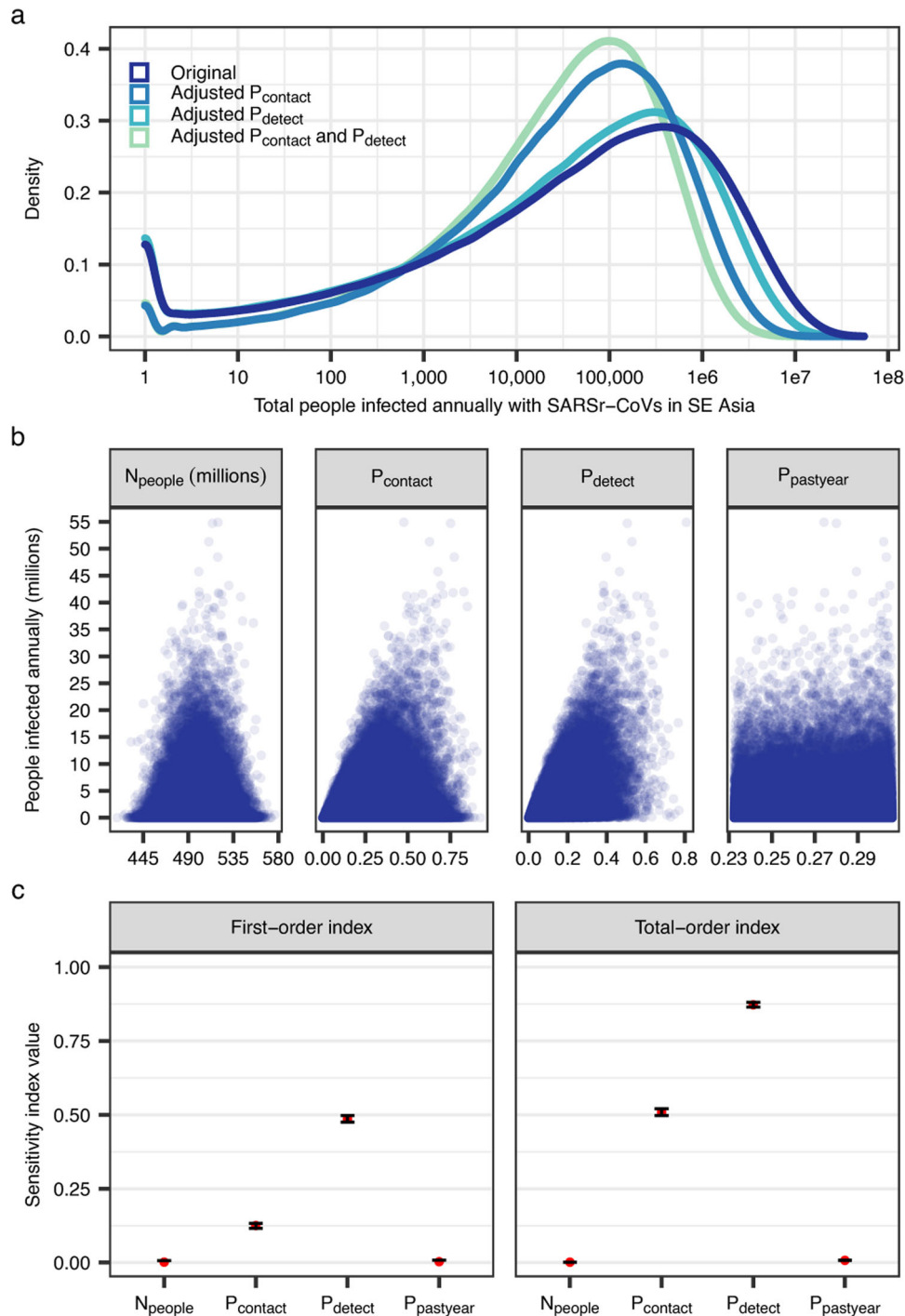
bat-CoVs, and for COVID-19 origin tracing. All of these are key goals for pandemic preparedness and prevention<sup>38–41</sup>, and if used to target future surveillance and disease control, may help to reduce the possibility of future COVID-like outbreaks.

Our analysis identifies regions in southern China, northeastern Myanmar, Lao PDR, and northern Vietnam as having the highest diversity of SARSr-CoV bat host species, in concordance with other recent efforts to map the distribution of SARSr-CoV bat host species<sup>42</sup>. These hotspots of SARSr-CoV bat reservoir host diversity may be particularly fruitful sites for viral discovery of novel SARSr-CoVs, assuming that viral diversity scales with host species diversity<sup>43</sup>. This finding supports conclusions from prior phylogenetic analyses that particularly diverse SARSr-CoV lineages are found in southern China<sup>23</sup>. Our results also help explain the recent identification of multiple strains of SARSr-CoVs in southern China<sup>33,44,45</sup> and southeast Asian countries<sup>30</sup>, despite small sample sizes. Our findings suggest that less intense sampling in countries bordering southern China (Supplementary Fig. 1) could have led to an underestimate of the diversity of these viruses there<sup>27,32</sup>. Given that the bat species known to host the closest relatives of SARS-CoV-2 are found in this region, our species richness map may also guide efforts to identify the viral clade from which a progenitor virus emerged<sup>23,33,35</sup>.

The map of bat-human overlap reveals hotspots in southern China and bordering countries, but also in the populous regions of Indonesia. This map may be useful in targeting surveillance to identify SARSr-CoV spillover events in people, including syndromic surveillance for SARS- or COVID-like respiratory disease in communities within these hotspots. This has been proposed previously as a tool for proactive surveillance for novel infections that could become standardized across emerging disease hotspots<sup>38,43</sup>. In this case, surveillance in clinics for severe/acute respiratory illness, influenza-like illness, and fevers of unknown origin would capture pneumonia or rapid-onset respiratory distress typical of SARS and COVID-19 patients, and likely to be expected from other novel bat-SARSr-CoV spillover infections<sup>46</sup>. More common infections such as influenza and bacterial or fungal pneumonia could be ruled out relatively easily<sup>47</sup>. The bat-human overlap map may also provide guidance for studies of where initial spillover of the progenitor of SARS-CoV-2 may have occurred, although epidemiological and other data suggest this would most likely have been in China or neighboring countries<sup>34,35</sup>.

Our estimate that a median of ~66,000 people are infected with SARSr-CoVs each year in Southeast Asia suggests that bat-to-human SARSr-CoV spillover is common in the region, and is undetected by surveillance programs and clinical studies in the majority of cases. While our results suggest significant levels of spillover, many of the diverse viral strains that infect people in the region each year may not be able to replicate well in people, cause illness, or be transmitted sufficiently among people to cause an outbreak. This has been shown in theoretical models of disease emergence<sup>21,48</sup>, and supports earlier evidence from studies of non-human primate virus spillover<sup>49</sup>, that cross-species transmission of novel animal-origin viruses is not the rate-limiting step in pandemic viral emergence. However, given the relatively large number of people likely to be infected each year with bat-CoVs, it is plausible that illnesses or clusters of cases due to novel bat-CoV infection occur regularly within the region, and are either not reported or are missed by clinical surveillance. Evidence of underreporting has been demonstrated for other bat-origin viral infections. For example, targeted syndromic surveillance of encephalitis patients in a small number of clinics in Bangladesh showed that Nipah virus causes outbreaks annually with an overall mortality rate of ~70%, despite it only recently being reported from the country<sup>50</sup>. Efforts to increase surveillance





**Fig. 3 Spillover simulations and sensitivity analyses.** **a** Density plots of the estimated total number of people in Southeast Asia infected with SARSr-CoVs by bats each year. Line colors represent scenarios exploring how estimated spillover changes if one or both of two parameter distributions ( $P_{\text{contact}}$  and  $P_{\text{detect}}$ ) are refit after excluding the highest estimates of these parameters gathered via literature searches (see Methods for details). Note that the x-axis is on a  $\log_{10}$  scale. **b** Estimated total number of people in Southeast Asia infected with SARSr-CoVs by bats each year, plotted as a function of four input variables. Values correspond to the original scenario (i.e., no adjusted parameters). **c** Sobol sensitivity indices, indicating the amount of variance in the outcome due to each input on its own (first-order index) and the amount of variance in the outcome due to each input including interactions with other inputs (total-order index). Sobol indices were calculated with two random samples of 200,000 points each. Red dots represent calculated index values and error bars represent 95% confidence intervals. Values correspond to the original scenario (i.e., no adjusted parameters).

for novel SARSr-CoVs (and other emerging viruses) in clinical cohorts, particularly using syndromic surveillance, may identify the rate of missed cases and pre-empt larger scale outbreaks. Estimating the true rate of spillover of previously unknown, potentially zoonotic animal-origin viruses is difficult without

serological or genomic surveillance data. For most viruses, the duration of infection in humans is relatively short (e.g., the infectious period for COVID-19 is 2–3 weeks<sup>51</sup>), and if spillover is rare, PCR surveillance is unlikely to give valuable data on spillover rates due to lack of positives. One exception is for

viruses with long infectious periods such as lentiviruses and some retroviruses; a previous genomic surveillance study of wildlife hunters in Africa was able to find 10/1099 (0.9%) prevalence of non-human primate origin simian foamy viruses<sup>52</sup>. Because detectable antibodies or T cell responses are more long-lived, serological surveys or activated T cell testing represent more valid strategies to estimate spillover rates.

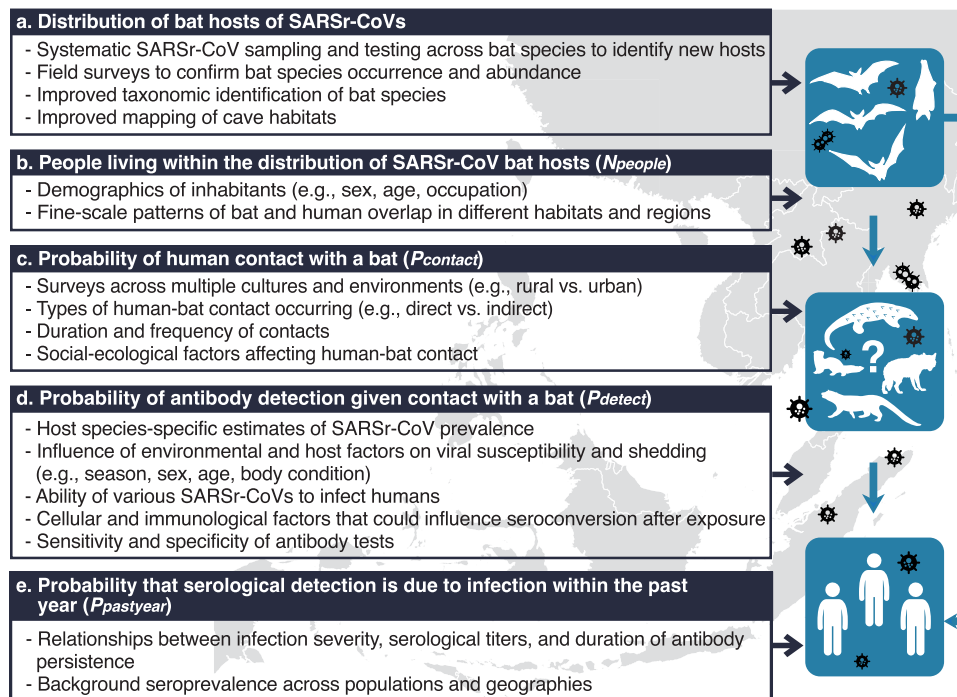
Our calculation of undetected spillover represents the first published attempt, to our knowledge, to identify the spillover risk of SARSr-CoVs from bats to people. It relies on a number of input variables, and we attempted to account for uncertainty (e.g., potential inaccuracy of serological tests) and variation (e.g., differences in human behavior due to gender, occupation, cultural norms) associated with these inputs by assigning a probability distribution to each variable, based on data gathered from the literature. We discuss these data, and some of the assumptions we make in using them, in more detail below.

Human-bat interactions likely vary widely across our study region, and are influenced by a variety of social, ecological, and economic factors at individual and community scales and beyond<sup>53–55</sup>. We therefore performed a systematic literature search to gather estimates of human-bat contact prevalence reported in previous surveys and ethnographic investigations. Our literature search revealed that limited data exist on human contact with bats in Southeast Asia, and that most studies of human-bat contact have targeted rural communities where bats are present and contact events could occur due to occupational or environmental exposure. As a result, estimates of human-bat contact reported in these studies might be higher than would be found in other human populations (i.e., more urban communities). However, we consider it plausible that urban residents face comparable levels of contact with bats as rural residents, though the interfaces may be different. For example, a recent study showed that rural and urban human populations in China displayed similar patterns of contact with wild animals, including bats<sup>56</sup>. Although some dense urban areas may provide less suitable roosting habitat for many bat species as compared to rural or peri-urban areas (meaning urban residents could be less likely to have bats in their home, visit caves, or hunt bats), several synanthropic bat species are able to thrive in human-developed landscapes<sup>57</sup>. Of the host species considered in this study, IUCN habitat type 14.5 (Urban areas) is considered suitable for two species, while other artificial terrestrial habitat types (14.1: Arable land, 14.2: Pastureland, 14.3: Plantations, 14.4: Rural gardens, 14.6: Subtropical/tropical highly degraded former forest) are suitable for 2–5 species each (Supplementary Table 1). Further, interfaces such as the wildlife trade—including physical markets and the supply chain—represent a major potential exposure pathway for people living in more densely populated urban settings. For instance, a study of the bat trade in Indonesia demonstrated that bats are harvested in rural areas but are directly transported to more urban and developed areas with more buying power<sup>58</sup>. Other work has found that wildlife (*Rattus* and *Bandicota* spp.) CoV prevalence increases along the distribution chain (i.e., from traders to large markets to restaurants), and suggested that end consumers face “maximal risk”<sup>59</sup>. Given this evidence, and in the absence of better data on contact with bats for urban residents, we believe it is reasonable to apply the bat-human contact rates generated from our literature search across our study region.

In determining the distribution for  $P_{detect}$  (the parameter describing the probability that a human-bat contact leads to a serologically detectable human infection), we incorporated literature-gathered seroprevalence estimates from people exposed to CoVs and other directly transmitted RNA viruses. We believe we are justified in using non-CoV data in our analyses for several

reasons. First, there are very few pre-COVID human serosurveys that specifically tested for SARSr-CoV exposure<sup>17,60</sup>, and we wanted to parameterize our spillover estimate with a wide range of possible values for infection and seroconversion using other RNA viruses that are known or likely able to be transmitted via direct exposure to bats or bat excreta. We acknowledge that there may be substantial variation in the ability of different SARSr-CoVs or other bat-harbored viruses to cause infection and seroconversion in humans; however, we know little about the large diversity of viral strains circulating in bat populations and the host cellular and immunological factors that could influence seroconversion after exposure for most viruses<sup>61</sup>. Indeed, many seroprevalence estimates reported in the literature may be an underestimate. For example, some individuals subclinically infected by SARS-CoV-2 may clear infection prior to production of antibodies<sup>62</sup> and some bat SARSr-CoVs may require limited evolutionary adaptation before antibody responses can be produced (e.g., data from RaTG13 binding<sup>63</sup>). A study of memory T-cells in people in Singapore detected SARS-CoV-2-specific interferon- $\gamma$  responses in approximately half (19/37) of people negative for SARS-CoV-1 or SARS-CoV-2 infection, and suggested unknown, potentially animal-origin CoVs could induce cross-reactive SARS-CoV-2 T cells<sup>64</sup>. Data from serological analysis also demonstrates high seroprevalence in Lao PDR. Specifically, Virachith and colleagues detected SARS-CoV-2-specific anti-N and anti-S antibodies in 5.3% and 1.1%, respectively, of human serum samples collected in 2018 (i.e., pre-COVID-19) in Lao PDR<sup>60</sup>. The authors also tested serum samples collected during the pandemic (August–September 2020) and detected antibodies in 5.2% (anti-N) and 2.1% (anti-S) of the general public, and in 20.3% (anti-N) and 6.8% (anti-S) of bat/wildlife contacts (bat guano collectors, wildlife animal vendors, and family members of bat contacts, catchers or guano collectors). The authors noted that high seroprevalence may also reflect the frequent trade in bush meat and low biosafety awareness reported in earlier surveys in Lao PDR<sup>60,65</sup>.

Previous authors have commented on the need to develop approaches to assessing zoonotic spillover risk that explicitly incorporate uncertainty<sup>66</sup>, and we recognize that further data would likely help refine estimates of SARSr-CoV spillover rates (Fig. 4). Perhaps most critical is a lack of information on the role of intermediate hosts in the emergence of SARSr-CoVs. It has been postulated that civets and other commonly farmed and traded mammalian species played a role in the emergence of SARS-CoV by acting as efficient amplifier hosts within wildlife farms and markets<sup>12,67–70</sup>. SARS-CoV-2 can infect civets, raccoon dogs, and other mammals commonly farmed and traded for food in the region<sup>71,72</sup>. SARS-CoV-2 has also caused significant outbreaks in animals bred for fur (e.g., mink, raccoon dogs) that have in some cases led to transmission to people<sup>73,74</sup>. These and other data led an international team to conclude that the most likely pathway for COVID emergence was from bats to people through a farmed mammalian intermediate host<sup>34</sup>. Accurate data are not available on the number of wildlife farms and potential intermediate hosts bred each year, or on the market systems that supply live animals into cities within China and Southeast Asia. It was estimated that 14 million people were employed in the wildlife farming system within China alone during 2016, in an industry worth \$77 billion annually<sup>75,76</sup>. Additionally, some of these potential intermediate hosts occur naturally in the region (e.g., pangolins, civets), and other livestock species that are common (e.g., pigs, cattle, rabbits) are susceptible to SARSr-CoVs either naturally or experimentally<sup>72</sup>. Spillover of SARSr-CoVs may therefore be substantially skewed to people who have high contact with these species, and this would likely have been missed in the serological surveys upon which our analyses are currently



**Fig. 4** Key data inputs and future research needs to improve estimates of SARSr-CoVs spillover from bats to humans. Additional data inputs are organized according to steps in our probabilistic risk assessment to improve our understanding of: **a** wildlife reservoir host (bat) distribution, **b** people overlapping with bat hosts, **c** probability of human-bat contact, **d** probability of detecting antibodies given contact, and **e** dynamics of antibody response in individuals.

based. Thus, better estimates of the role of farmed and traded intermediate hosts are likely to substantially increase the estimated spillover rates of SARSr-CoVs across the region.

Other data are needed to improve and help validate our estimates from limited knowledge. Useful research could include bat surveys to identify contemporary, accurate species presence records, and sampling a range of bat species for SARSr-CoVs to identify previously unknown SARSr-CoV hosts (Fig. 4a). Statistical models can inform prioritization of species to sample to optimize viral discovery<sup>77</sup>. Further data are also needed on the role of host biology and ecology in spillover: for instance, improved species-specific estimates of SARSr-CoV prevalence, and the effects of environmental factors (e.g., location, season, year) and host factors (e.g., sex, age, body condition, reproductive status) on viral shedding<sup>78–81</sup> (Fig. 4d). As described previously, there is a need to better understand how and why people and bats come into contact in regions of overlapping occurrence (Fig. 4b, c). Crucial immunological information includes expanded survey data on the background seroprevalence of bat SARSr-CoVs in people across different populations and geographies, improved understanding of factors that could influence seroconversion after exposure, and studies of how serological titers and duration of antibody persistence relate to severity of infection (Fig. 4d, e). Finally, it is also possible that taxonomic errors have occurred in the data we analyzed. Taxonomic standards in viral discovery and surveillance studies vary widely, and point to a need for better taxonomic training of field teams, standardized DNA barcoding of hosts, collection of voucher specimens, and closer collaboration among disease ecologists, virologists, field biologists and taxonomists<sup>66,82,83</sup>.

Our refinement of species ranges has produced a detailed picture of SARSr-CoV bat host species potential occurrence. The AOHs we produced may be useful for targeting surveillance to key species. For example, only a few species' AOHs contained a sizable proportion ( $\geq 25\%$ ) of one or more artificial habitats (Fig. 1b): *R. malayanus* (arable land, plantations), *R. stheno* (plantations, arable land), and

*N. leisleri* (pastureland). These species may be more important as spillover hosts considering that these modified landscapes have a higher opportunity for human-wildlife contact and are known to have heightened diversity of other zoonotic disease hosts<sup>84</sup>. Cave habitats were classified by the IUCN as suitable for nearly all species in our analyses, and carbonate rock outcrops (used here as a proxy for caves) made up a large proportion of species AOHs. Public presence in caves and livelihoods or occupational activities involved in bat guano collection are likely particularly hazardous given these findings<sup>85–88</sup>. Our validation process indicated good agreement between species AOHs and GBIF occurrence data, suggesting that the maps we generated accurately reflect species presence for occurrence records collected after 1990. We note that external forces could drive bats away from otherwise suitable habitats; for example, many cave-dwelling bats face disturbance in the forms of tourism, guano harvesting, vandalism, land use change, and more<sup>89–91</sup>. Limitations of AOH maps include the potential inaccuracy of the IUCN species ranges, habitat suitability assignments, and elevation limits<sup>36,92</sup>. All but one of the species in our dataset were last assessed by the IUCN in 2016 or more recently (*A. stoliczkanus* was last assessed in 2008); in future, revised AOHs can be developed based on new IUCN range data. The map of terrestrial habitat types used in our analyses<sup>93,94</sup> is not a perfect representation of habitat, but is the most accurate and detailed analysis of its kind to date. In particular, the map did not include caves, an important habitat type for many bat species. We used carbonate rock outcrop data as a proxy for cave distribution and this should be ground-truthed. Finally, though grid cells are assigned to one dominant habitat class, at a fine-scale level they are comprised of various microhabitats which may be more or less attractive to wildlife.

Our analytical framework provides a strategy that has potential for improving preparedness for emerging diseases and pandemics. However, it is important to recognize that our results specifically identify spillover risk, which we defined as the annual number of people infected by SARSr-CoVs via bat-to-human



transmission events. Our analyses do not assess the impact of spillover, the severity of disease that results (if any), the ability of infections to lead to human-to-human (community) transmission, or the likelihood of a regional outbreak or a pandemic<sup>37</sup>. Furthermore, we did not account for differences in vulnerability among populations, communities, or individuals. However, our analyses have public health value in that the maps can be used to conduct more cost-effective field surveys for viral discovery programs, guide human surveillance to identify clusters of cases of a new CoV infection earlier and help prevent spread, and guide targeted epidemiological studies of small-scale processes that increase spillover risk<sup>95–97</sup>. Our analysis pipeline and framework are based on open-source code and can therefore serve as a resource to update and modify spillover risk maps and estimates as new data become available. Finally, our framework can be rapidly adapted for spillover risk assessment of other viral groups, such as the HKU-2/SADS-CoV  $\alpha$ -CoVs that have recently been found able to infect primary human airway epithelial cells in vitro, and therefore pose a heightened spillover risk, or any of the other ~25 viral families that include known zoonoses<sup>43</sup>.

## Methods

All analyses were performed in the R statistical environment v. 4.0.3 (R Core Team 2020).

**Estimation of coronavirus research effort.** To assess CoV research effort in Southeast Asia, we used the R package `rentrez` v1.2.3<sup>98</sup> to query the PubMed database for the number of publications from each country in Southeast Asia (see below), as of November 17, 2021, using the search string: (“bat” OR “bats” OR “Chiroptera”) AND (“coronavirus” OR “coronaviruses”)<sup>93</sup>. We visualized a map of CoV research effort using the R package `rworldmap` v1.3–6<sup>99</sup>.

**Compilation of SARSr-CoV bat host data.** We identified all bat species from which molecular evidence of SARSr-CoV infection had been reported, and for which associated sequence confirmation data were available. We supplemented a recently compiled list<sup>100</sup> with hosts listed in other more recent publications<sup>23,30,32,44,101,102</sup>. We refrained from including bat species that have been predicted, but not confirmed, to be SARSr-CoV hosts. We considered *R. paradoxolophus* to be a subspecies of *R. rex*<sup>103</sup>. We excluded *R. monoceros* because, although recent work has retained it as a distinct species<sup>104</sup>, no recent assessment of the species has been published by the IUCN Red List of Threatened Species. *H. pomona* and *H. gentilis* were split in 2018, with *H. pomona* restricted to a small area in southern India while *H. gentilis* is broadly distributed across Southeast Asia<sup>105</sup>. Although Latinne and colleagues<sup>23</sup> listed *H. pomona* as a host species, because the field sampling was conducted in China, we substituted *H. gentilis* in our analyses.

We included only bat hosts with geographic distributions either partially or entirely within Southeast Asia (see Supplementary Table 1 for the finalized list of SARSr-CoV bat hosts). The area covered by the following countries and administrative regions is included in our analysis extent: Bangladesh, Bhutan, Brunei, Cambodia, China, Hong Kong SAR, Macao SAR, India, Indonesia, Lao People’s Democratic Republic, Malaysia, Myanmar, Nepal, the Philippines, Singapore, Sri Lanka, Thailand, Timor-Leste, and Vietnam. Note that while we use the term Southeast Asia, the region we considered has broader range than is used for political definitions of the region and includes China and parts of South Asia, due to the extensive range of some of the bat host species. All analyses described below were restricted by the regional boundaries of Southeast Asia as defined above.

**Calculation and validation of host area of habitat.** To assess SARSr-CoV bat host distribution and richness across our study area, we derived the area of habitat (AOH) for each species. AOH describes the habitat suitable for a species within its range according to the species’ habitat preferences and elevation limits<sup>36</sup>. Previous validation of AOH maps with species occurrence data has demonstrated their high accuracy; for example, ~92% (241/263) of terrestrial mammal species AOHs predicted species point localities better than their IUCN geographical range<sup>106</sup>. Determining a species AOH does not rely on presence data, making it useful when assessing species with few occurrence records (as was the case for many species in our analysis; see Results).

To derive the AOH for each species, we downloaded its geographic range from the IUCN Red List of Threatened Species<sup>26</sup> (last update: March 25, 2021), overlaid it onto a raster map of terrestrial habitat types<sup>93,94</sup>, and selected areas of the map that occurred within its range. We then selected areas of suitable habitat for each species. Specifically, using the IUCN Red List data, we assessed if each of 22 habitat types was suitable for each bat host species. According to the IUCN Habitats

Classification Scheme, a designation that a habitat is suitable means “the species occurs in the habitat regularly or frequently”<sup>107</sup>. The 22 habitat types we examined included various forest habitats, shrublands, rocky areas and caves, and artificial habitats, chosen from a larger list of ~100 habitat types<sup>107</sup> as being influential in determining the occurrence of one or more target host species (Supplementary Table 1). A global land cover map of these habitat types<sup>93,94</sup> included 18 of the 22 target habitats. Because two habitat types (7.1. Caves and Subterranean Habitats (non-aquatic)—Caves; 7.2. Caves and Subterranean Habitats (non-aquatic)—Other subterranean habitats) were not included in the global land cover map, but were suitable for many host species, we supplemented the land cover dataset with the World Map of Carbonate Rock Outcrops v3.0<sup>108</sup> as a proxy for karst landscape and cave distribution. This dataset comprises two layers, one with areas of continuous carbonate rocks and another with abundant but not continuous rocks; we used the former for our analyses. We rasterized the carbonate rock outcrop shapefile using the R package `fasterize` v1.0.3<sup>109</sup> and combined this with the land cover raster. When restricting species IUCN ranges by suitable habitat, we used the combined land cover/carbonate rock raster if a species was found in habitat type 7.1, and the land cover-only raster if a species was not found in this habitat type. The other two habitat types not represented on the land cover raster were only suitable for 1–2 host species (i.e., 14.6. Subtropical/Tropical Heavily Degraded Former Forest; 15.8. Seasonally Flooded Agricultural Land); therefore, we did not attempt to find other data to use as a proxy for these habitats.

Finally, we overlaid elevation data (Shuttle Radar Topography Mission data obtained with the `getData` function of the R package `raster` v3.4–5<sup>110</sup>) onto each species’ range and habitat map and selected areas that fell within a species’ elevation limits (Supplementary Table 1); these remaining areas represented the species’ AOH. Using the `area` function of the `raster` package<sup>110</sup>, we calculated the size (in km<sup>2</sup>) of each species’ AOH and compared it to the size of its original IUCN range.

We validated the AOH of each species using occurrence data downloaded from GBIF using the R package `rgbif` v3.6.0<sup>111–113</sup>. We cleaned GBIF data by removing records with inaccurate or imprecise coordinate data (i.e., no coordinates, identical longitude and latitude, coordinates of 0, coordinate uncertainty >35 km, coordinates in the ocean, country–coordinate mismatch), records from areas outside our geographically defined region, records of absence rather than presence, and records with a Basis of Record other than human observation, machine observation, observation, material sample or preserved specimen. We removed records within 5 km of country capitals and within 1 km of country/province centroids (as coordinates are likely to be assigned to capitals and centroids in the absence of detailed location information), and within 100 m of biodiversity institutions (as these coordinates would represent the physical location of a specimen, rather than its capture location). We removed records before 1991, as older records tend to have less precise location data<sup>114</sup> and could reflect species ranges that have since shifted. Within each species, we removed records with duplicate coordinates. Cleaning was facilitated with the R package `CoordinateCleaner` v2.0–18<sup>115</sup>. For each species, we buffered each of its occurrence points by a five-kilometer radius and calculated the percent of buffered points that overlapped the species’ AOH.

**Estimation of bat-human overlap, bat species richness, and habitat proportions.** To identify regions where human populations might be exposed directly or indirectly to SARSr-CoV bat hosts, we overlaid human population count data on each host species’ AOH. We used a 1-km resolution raster of 2020 population count data from `WorldPop`<sup>116</sup> and resampled it with bilinear interpolation using the R package `gdalUtils` v2.0.3.2<sup>117</sup> so that its extent and resolution matched that of the habitat raster. We calculated the number of people living within the AOH of each species separately, and divided this by the size of each AOH to obtain average human population density values. As living in areas with high diversity of SARSr-CoV bat hosts and large human populations may increase the likelihood of human-bat contact, viral spillover, and subsequent pathogen spread, we visualized bat species richness as well as bat species richness multiplied by human population size (which we term bat-human overlap) across the consensus map. To examine the relative importance of different habitat types to SARSr-CoV bat hosts, we calculated the proportion (by area) of each habitat type within the AOH of each species. To examine which habitat types might be most important for human exposure to bats, we calculated the number of people living in each habitat type of each species’ AOH.

**Estimating human infection with SARSr-CoVs.** We used a probabilistic risk assessment to estimate SARSr-CoV spillover risk in Southeast Asia. For this analysis, we define spillover risk as the annual number of people infected by SARSr-CoVs via bat-to-human transmission events. We draw on an established framework that defines risk as equal to hazard times exposure times vulnerability, where a hazard is a potential source of harm from a microbe, exposure is the likelihood of contact between humans and hazards, and vulnerability is the chance of a hazard causing harm, given exposure<sup>37</sup>. Here, the hazard is a bat belonging to a species known to host SARSr-CoVs, exposure is the likelihood of contact between a person and a bat, and vulnerability is the chance that a human-bat contact event results in a human infection that produces detectable antibodies.



We assumed that spillover risk can be approximated as the number of people who live in the consensus area of SARSr-CoV bat hosts ( $N_{\text{people}}$ ), multiplied by the probability that a human comes into contact with a bat ( $P_{\text{contact}}$ ), multiplied by the probability that a human-bat contact leads to a serologically detectable human infection ( $P_{\text{detect}}$ ), multiplied by the probability that serological detection is due to an infection within the previous year ( $P_{\text{pastyear}}$ ) (1).

$$\text{Spillover risk} = N_{\text{people}} * P_{\text{contact}} * P_{\text{detect}} * P_{\text{pastyear}} \quad (1)$$

We accounted for uncertainty (e.g., potential inaccuracy of serological tests) and variation (e.g., differences in human behavior due to gender, occupation, cultural norms) associated with the input variables by assigning a probability distribution (rather than a single fixed value) to each. We performed Latin hypercube sampling with the R package `lhs` v1.1.1<sup>118</sup> to generate 400,000 sets of input combinations and calculated the corresponding spillover risk for each set of inputs, creating an output distribution for spillover risk. We then calculated summary statistics (median, 95% confidence interval for the median, overall range, interquartile range) for this distribution. The 95% confidence interval for the median was calculated using the `MedianCI` function in the R package `DescTools` v0.99.41<sup>119</sup> with 1000 bootstrap replicates. Our choices of probability distributions for  $N_{\text{people}}$ ,  $P_{\text{contact}}$ ,  $P_{\text{detect}}$ , and  $P_{\text{pastyear}}$  are described below.

For  $N_{\text{people}}$ , we assigned the normal distribution  $N(499101844, 16636728^2)$ . The mean value was derived from our calculation of the number of people living in the consensus area of SARSr-CoV bat hosts within Southeast Asia, based on 2020 WorldPop population count data. The standard deviation was chosen so that the tails of the distribution (3 standard deviations) would extend 10% beyond the mean, to reflect potential uncertainty associated with the WorldPop dataset.

To inform our choice of distribution for  $P_{\text{contact}}$ , we gathered data on bat-human contacts in Southeast Asia using a systematic search of Google Scholar, PubMed, and Web of Science with the following keywords: (“bat contact” OR “human bat contact” OR “human bat interaction” OR “bat human contact” OR “bat human interaction”) AND (“Bangladesh” OR “Bhutan” OR “Brunei” OR “Cambodia” OR “China” OR “India” OR “Indonesia” OR “Lao PDR” OR “Malaysia” OR “Myanmar” OR “Nepal” OR “Philippines” OR “Singapore” OR “Sri Lanka” OR “Thailand” OR “Timor-Leste” OR “Vietnam”). English language papers published from January 1, 2000 to May 07, 2021 were included in the search. A total of 421 records were initially identified, of which 24 were retained following title and abstract screening that excluded duplicate articles, studies that only focused on bat-borne pathogens or bat ecology, and studies based outside of Southeast Asia. Full text review was performed on these 24 articles. Six articles provided data on the frequency of human contact with bats or their excreta within the last 12 months. These studies included specific data on behaviors and practices that might allow viral transmission (e.g., eating bats, receiving a bat scratch or bite, exposure to urine or guano, being in a bat cave, having bats in one’s house; summarized in Supplementary Table 3). Given that the probability of contact is bounded between 0 and 1, we assigned the beta distribution  $\beta(0.9366017, 5.2604551)$  to the 24 estimates of bat-human contact reported in the final set of six articles, using the R package `fitdistrplus` v1.1–3<sup>120</sup> to determine the shape parameters (Supplementary Fig. 4a).

In a similar manner, to inform our choice of distribution for  $P_{\text{detect}}$ , we gathered data on viral seroprevalence among people with self-reported bat contact using a systematic search of Google Scholar, PubMed, and Web of Science with the following keywords: (“bat contact” OR “human bat contact” OR “human bat interaction” OR “bat human contact” OR “bat human interaction”) AND (“serological prevalence” OR “seroprevalence” OR “serological evidence” OR “human infection” OR “spillover”). English language papers published from January 1, 2000 to June 21, 2021 were included in the search. A total of 339 records were initially identified, of which 12 were retained following title and abstract screening that excluded (i) duplicate articles, (ii) studies that reported seroprevalence in animal populations, (iii) studies that provided either seroprevalence or human-bat contact but not both, (iv) studies that detected pathogens that are unlikely to be transmitted from contact with bats, and (v) studies that analyzed human-bat contact only among identified seropositive human cases. Full text review was performed on these 12 articles. Seven articles provided data on human viral seroprevalence among populations who reported contact with bats. An additional relevant article was discovered during the revision process, making a total of eight final articles included and 16 reported estimates of seroprevalence (Supplementary Table 4). We assigned the beta distribution  $\beta(0.2405859, 6.57698)$  to these data, again using the `fitdistrplus` package<sup>120</sup> to determine the shape parameters (Supplementary Fig. 4c). We note that it is unlikely that all known and as-yet-undiscovered bat SARSr-CoVs are likely able to infect people directly, or at all. However, the use of serology data from human surveys provides a way to account for this by providing information on exposure that has led to prior infection, albeit that the severity of that infection remains unknown.

Studies of SARS patients have provided a range of estimates for the persistence of antibodies at diagnostically detectable levels, with a maximum of six years, suggesting that the proportion of those testing seropositive could include people infected more than one year prior to testing. We note that studies of memory T cells have shown that it is possible to identify antibodies as long as 17 years after SARS infection<sup>121</sup>. However, our analysis uses published data on serological tests, not memory T-cell activation studies, therefore we set the maximum at six years,

which is the longest time after exposure that published data for serological testing provided evidence of detectable IgG<sup>122</sup>. Given that our aim was to estimate the amount of bat-to-human spillover each year, we performed a calculation to estimate the likelihood that someone who tested seropositive was infected within the past year. We gathered data from papers that reported the duration of SARS-CoV immunoglobulin G (IgG) antibody detection among patients who recovered from SARS (Supplementary Table 5). Generally, the detectability of IgG rose rapidly following infection, peaked at 3–4 months following symptom onset and remained high for the first 16 months following symptom onset. Detectability then decreased over time, with ~40–70% of patients having detectable IgG after three years, and only ~10% having detectable IgG after six years. We fit a 2<sup>nd</sup> order polynomial to these data, then divided the integral from 0 to 12 months by the integral from 0 to 71.5 months (when antibody detection declines to ~0 on the curve) to calculate the probability that antibody detection is due to an infection in the past year. We then repeated this process excluding the only data point at 72 months, this time dividing the integral from 0 to 12 months by the integral from 0 to 49.5 months (when antibody detection goes to ~0 on the curve). We used the two values generated from these processes to assign the uniform distribution  $\beta(0.232444, 0.3059904)$  for  $P_{\text{pastyear}}$ .

Finally, we performed a global sensitivity analysis by calculating first-order and total-order Sobol sensitivity indices<sup>123</sup> to understand the relative contribution of each input variable to the outcome. Sobol sensitivity indices measure the amount of variance in an outcome (here, spillover risk) due to input variables, and can be used for non-linear relationships. The first-order index represents the amount of variance in the outcome due to each input on its own. The total-order index represents the amount of variance in the outcome due to each input including interactions with other inputs. First-order and total-order indices are always  $\leq 1$ , and first-order indices are  $\leq$  total-order indices. All sensitivity indices were calculated using the R package `sensitivity` v1.25.0<sup>124</sup> using the Janon-Monod method, and 95% confidence intervals were calculated using 100 bootstrap replicates using the same package<sup>125,126</sup>.

Sensitivity analyses indicated that  $P_{\text{contact}}$  and  $P_{\text{detect}}$  primarily contributed to variance in the outcome (see Results). Therefore, we repeated our calculation of spillover risk by excluding the three highest estimates of  $P_{\text{contact}}$  and the two highest estimates of  $P_{\text{detect}}$  found in our literature searches and refitting the distributions for these parameters. We explored three scenarios: (1) only the distribution for  $P_{\text{contact}}$  was refit, (2) only the distribution for  $P_{\text{detect}}$  was refit, and (3) both distributions were refit. See Supplementary Fig. 4 for an illustration of the original and adjusted distributions for  $P_{\text{contact}}$  and  $P_{\text{detect}}$ .

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Bat distribution shapefiles are available at <https://www.iucnredlist.org/resources/spatial-data-download>. Elevation and habitat suitability for individual species are available at <https://www.iucnredlist.org/>. Bat occurrence data are available at <https://doi.org/10.15468/dl.8w26d8>. Human population count data are available at <https://doi.org/10.5258/SOTON/WP00647>. Carbonate rock outcrop data are available at <https://crc806db.uni-koeln.de/layer/show/296/>. A global map of terrestrial habitat types (version 001) is available at <https://doi.org/10.5281/zenodo.3666246>. A global land shapefile is available at <https://www.naturalearthdata.com/downloads/50m-physical-vectors/>. Data from published sources are available in Supplementary Tables and also at <https://doi.org/10.5281/zenodo.5251725>.

## Code availability

All R scripts needed to reproduce the analysis and figures are available in a Github repository ([https://github.com/ecohealthalliance/sars\\_cov\\_risk](https://github.com/ecohealthalliance/sars_cov_risk)), also deposited in Zenodo at <https://doi.org/10.5281/zenodo.5251725>.

Received: 24 August 2021; Accepted: 15 June 2022;

Published online: 09 August 2022

## References

- Lee, J.-W. & McKibbin, W. J. Globalization and disease: the case of SARS. *Asian Economic Pap.* **3**, 113–131 (2004).
- Cutler, D. M. & Summers, L. H. The COVID-19 pandemic and the \$16 trillion virus. *JAMA* **324**, 1495–1496 (2020).
- Peiris, J. S. M., Guan, Y. & Yuen, K. Y. Severe acute respiratory syndrome. *Nat. Med.* **10**, S88–S97 (2004).
- Raj, V. S., Osterhaus, A. D. M. E., Fouchier, R. A. M. & Haagmans, B. L. MERS: emergence of a novel human coronavirus. *Curr. Opin. Virol.* **5**, 58–62 (2014).

5. Zhou, P. et al. Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. *Nature* **556**, 255–258 (2018).
6. Zhou, L. et al. The re-emerging of SARS-CoV infection in pig herds in Southern China. *Transbound. Emerg. Dis.* **66**, 2180–2183 (2019).
7. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
8. Daszak, P., Keusch, G. T., Phelan, A. L., Johnson, C. K. & Osterholm, M. T. Infectious disease threats: a rebound to resilience. *Health Aff.* **40**, 204–211 (2021).
9. Anthony, S. J. et al. Further evidence for bats as the evolutionary source of Middle East respiratory syndrome coronavirus. *mBio* **8**, e00373-17 (2017).
10. Li, W. D. et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676–679 (2005).
11. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
12. Wang, L. F. & Eaton, B. T. In *Wildlife and Emerging Zoonotic Diseases: The Biology, Circumstances and Consequences of Cross-Species Transmission* (eds J. E. Childs, J. S. Mackenzie, & J. A. Richt) 325–344 (Springer Berlin Heidelberg, 2007).
13. Dudas, G., Carvalho, L. M., Rambaut, A. & Bedford, T. MERS-CoV spillover at the camel-human interface. *eLife* **7**, e31257 (2018).
14. Ge, X. Y. et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535–538 (2013).
15. Menachery, V. D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* **21**, 1508–1513 (2015).
16. Menachery, V. D. et al. SARS-like WIV1-CoV poised for human emergence. *Proc. Natl Acad. Sci. USA* **113**, 3048–3053 (2016).
17. Li, H. et al. Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. *Biosaf. Health* **1**, 84–90 (2019).
18. Wang, N. et al. Serological evidence of bat SARS-related coronavirus infection in humans, China. *Virologica Sin.* **33**, 104–107 (2018).
19. Wasik, B. R. et al. Onward transmission of viruses: how do viruses emerge to cause epidemics after spillover? *Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **374**, 20190017 (2019).
20. Parrish, C. R. et al. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol. Mol. Biol. Rev.* **72**, 457–470 (2008).
21. Lloyd-Smith, J. O. et al. Epidemic dynamics at the human-animal interface. *Science* **326**, 1362–1367 (2009).
22. Gray, G. C., Robie, E. R., Studstill, C. J. & Nunn, C. L. Mitigating future respiratory virus pandemics: new threats and approaches to consider. *Viruses* **13**, 637 (2021).
23. Latinne, A. et al. Origin and cross-species transmission of bat coronaviruses in China. *Nat. Commun.* **11**, 4235 (2020).
24. McFarlane, R., Sleight, A. & McMichael, T. Synanthropy of wild mammals as a determinant of emerging infectious diseases in the Asian-Australasian region. *EcoHealth* **9**, 24–35 (2012).
25. Hu, B. et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLOS Pathog.* **13**, e1006698 (2017).
26. IUCN. *The IUCN Red List of Threatened Species. Version 2021-1*, <https://www.iucnredlist.org> (2021).
27. Ruiz-Aravena, M. et al. Ecology, evolution and spillover of coronaviruses from bats. *Nat. Rev. Microbiol.* **20**, 299–314 (2022).
28. Coker, R. J., Hunter, B. M., Rudge, J. W., Liverani, M. & Hanvoravongchai, P. Emerging infectious diseases in southeast Asia: regional challenges to control. *Lancet* **377**, 599–609 (2011).
29. Horby, P. W., Pfeiffer, D. & Oshitani, H. Prospects for emerging infections in East and Southeast Asia 10 years after severe acute respiratory syndrome. *Emerg. Infect. Dis.* **19**, 853–860 (2013).
30. Wacharapluesadee, S. et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat. Commun.* **12**, 972 (2021).
31. Rulli, M. C., D’Odorico, P., Galli, N. & Hayman, D. T. S. Land-use change and the livestock revolution increase the risk of zoonotic coronavirus transmission from rhinolophid bats. *Nat. Food* **2**, 409–416 (2021).
32. Delaune, D. et al. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat. Commun.* **12**, 6563 (2021).
33. Zhou, H. et al. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell* **184**, 4380–4391 (2021).
34. World Health Organization. WHO-convened global study of origins of SARS-CoV-2: China Part. (2021).
35. Holmes, E. C. et al. The origins of SARS-CoV-2: a critical review. *Cell* **184**, 4848–4856 (2021).
36. Brooks, T. M. et al. Measuring terrestrial Area of Habitat (AOH) and its utility for the IUCN Red List. *Trends Ecol. Evol.* **34**, 977–986 (2019).
37. Hosseini, P. R. et al. Does the impact of biodiversity differ between emerging and endemic pathogens? The need to separate the concepts of hazard and risk. *Philos. Trans. R. Soc. B: Biol. Sci.* **372**, 20160129 (2017).
38. Dobson, A. P. et al. Ecology and economics for pandemic prevention. *Science* **369**, 379–381 (2020).
39. Petrovan, S. O. et al. Post COVID-19: a solution scan of options for preventing future zoonotic epidemics. *Biol. Rev.* **96**, 2694–2715 (2021).
40. Roche, B. et al. Was the COVID-19 pandemic avoidable? A call for a “solution-oriented” approach in pathogen evolutionary ecology to prevent future outbreaks. *Ecol. Lett.* **23**, 1557–1560 (2020).
41. Naguib, M. M., Ellström, P., Järhult, J. D., Lundkvist, Å. & Olsen, B. Towards pandemic preparedness beyond COVID-19. *Lancet Microbe* **1**, e185–e186 (2020).
42. Muylaert, R. L. et al. Present and future distribution of bat hosts of sarbecoviruses: implications for conservation and public health. *Proc. Roy. Soc. B.* **289**, 20220397 (2022).
43. Carroll, D. et al. The global virome project. *Science* **359**, 872–874 (2018).
44. Zhou, H. et al. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* **30**, 2196–2203.e2193 (2020).
45. Li, L.-L. et al. A novel SARS-CoV-2 related coronavirus with complex recombination isolated from bats in Yunnan province, China. *Emerg. Microbes Infect.* **10**, 1683–1690 (2021).
46. Pormohammad, A. et al. Comparison of confirmed COVID-19 with SARS and MERS cases - Clinical characteristics, laboratory findings, radiographic signs and outcomes: A systematic review and meta-analysis. *Rev. Med. Virol.* **30**, e2112 (2020).
47. Brehm, T. T. et al. Comparison of clinical characteristics and disease outcome of COVID-19 and seasonal influenza. *Sci. Rep.* **11**, 5803 (2021).
48. Wolfe, N. D., Dunavan, C. P. & Diamond, J. Origins of major human infectious diseases. *Nature* **447**, 279–283 (2007).
49. Wolfe, N. D. et al. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc. Natl Acad. Sci. USA* **102**, 7994–7999 (2005).
50. Nikolay, B. et al. Transmission of Nipah virus—14 Years of investigations in Bangladesh. *N. Engl. J. Med.* **380**, 1804–1814 (2019).
51. Byrne, A. W. et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open* **10**, e039856 (2020).
52. Wolfe, N. D. et al. Naturally acquired simian retrovirus infections in central African hunters. *Lancet* **363**, 932–937 (2004).
53. Mildenstein, T., Tanshi, I. & Racey, P. A. Exploitation of bats for bushmeat and medicine. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (eds Voigt, C. C. & Kingston, T.) Ch. 12, 325–375 (Springer International Publishing, 2016).
54. Low, M.-R. et al. Bane or blessing? Reviewing cultural values of bats across the Asia-Pacific region. *J. Ethnobiol.* **41**, 18–34 (2021).
55. Kingston, T. Cute, creepy, or crispy—How values, attitudes, and norms shape human behavior toward bats. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (eds Voigt, C. C. & Kingston, T.) 571–595 (Springer International Publishing, 2016).
56. Li, H. et al. Knowledge, attitude, and practice regarding zoonotic risk in wildlife trade, Southern China. *EcoHealth* **18**, 95–106 (2021).
57. Jung, K. & Threlfall, C. G. Urbanisation and its effects on bats—A global meta-analysis. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (eds Voigt, C. C. & Kingston, T.) Ch. 2, 13–33 (Springer International Publishing, 2016).
58. Latinne, A. et al. Characterizing and quantifying the wildlife trade network in Sulawesi, Indonesia. *Glob. Ecol. Conserv.* **21**, e00887 (2020).
59. Huong, N. Q. et al. Coronavirus testing indicates transmission risk increases along wildlife supply chains for human consumption in Viet Nam, 2013–2014. *PLOS ONE* **15**, e0237129 (2020).
60. Virachith, S. et al. Low seroprevalence of COVID-19 in Lao PDR, late 2020. *Lancet Regional Health – West. Pac.* **13**, 100197 (2021).
61. Letko, M., Seifert, S. N., Olival, K. J., Plowright, R. K. & Munster, V. J. Bat-borne virus diversity, spillover and emergence. *Nat. Rev. Microbiol.* **18**, 461–471 (2020).
62. Swadling, L. et al. Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2. *Nature* **601**, 110–117 (2022).
63. Liu, K. et al. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other species. *Cell* **184**, 3438–3451.e3410 (2021).
64. Le Bert, N. et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).
65. Philavong, C. et al. Perception of health risks in Lao market vendors. *Zoonoses Public Health* **67**, 796–804 (2020).
66. Carlson, C. J. et al. The future of zoonotic risk prediction. *Philos. Trans. R. Soc. B: Biol. Sci.* **376**, 20200358 (2021).

67. Bell, D., Robertson, S. & Hunter, P. R. Animal origins of SARS coronavirus: possible links with the international trade in small carnivores. *Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci.* **359**, 1107–1114 (2004).
68. He, J. F. et al. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. *Science* **303**, 1666–1669 (2004).
69. Tu, C. et al. Antibodies to SARS-Coronavirus in Civets. *Emerg. Infect. Dis.* **10**, 2244–2248 (2004).
70. Guan, Y. et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science* **302**, 276–278 (2003).
71. Freuling, C. et al. Susceptibility of raccoon dogs for experimental SARS-CoV-2 infection. *Emerg. Infect. Dis.* **26**, 2982–2985 (2020).
72. OIE-World Organisation for Animal Health. Infection with SARS-CoV-2 in animals. <https://www.oie.int/app/uploads/2021/11/en-factsheet-sars-cov-2-20211025.pdf> (2021).
73. Oreshkova, N. et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Eurosurveillance* **25**, 2001005 (2020).
74. Oude Munnink, B. B. et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172–177 (2021).
75. Daszak, P. et al. Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services. (Bonn, Germany, 2020).
76. Chinese Academy of Engineering. Report on sustainable development strategy of China's wildlife farming industry. (2017).
77. Becker, D. J. et al. Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. *The Lancet Microbe*, [https://doi.org/10.1016/S2666-5247\(21\)00245-7](https://doi.org/10.1016/S2666-5247(21)00245-7) (2022).
78. Wacharapluesadee, S. et al. Longitudinal study of age-specific pattern of coronavirus infection in Lyle's flying fox (*Pteropus lylei*) in Thailand. *Viol. J.* **15**, 38 (2018).
79. Luo, Y. et al. Longitudinal surveillance of Betacoronaviruses in fruit bats in Yunnan Province, China during 2009–2016. *Virologica Sin.* **33**, 87–95 (2018).
80. Maganga, G. D. et al. Genetic diversity and ecology of coronaviruses hosted by cave-dwelling bats in Gabon. *Sci. Rep.* **10**, 7314 (2020).
81. Epstein, J. H. et al. Nipah virus dynamics in bats and implications for spillover to humans. *Proc. Natl Acad. Sci. USA* **117**, 29190 (2020).
82. Thompson, C. W. et al. Preserve a voucher specimen! The critical need for integrating natural history collections in infectious disease studies. *mBio* **12**, e02698–02620 (2021).
83. Phelps, K. L. et al. Bat research networks and viral surveillance: gaps and opportunities in Western Asia. *Viruses* **11**, 240 (2019).
84. Gibb, R. et al. Zoonotic host diversity increases in human-dominated ecosystems. *Nature* **584**, 398–402 (2020).
85. Robertson, K. et al. Rabies-related knowledge and practices among persons at risk of bat exposures in Thailand. *Plos Negl. Trop. Dis.* **5**, e1054 (2011).
86. Wacharapluesadee, S. et al. Group C Betacoronavirus in bat guano fertilizer, Thailand. *Emerg. Infect. Dis.* **19**, 1349–1352 (2013).
87. Suwannarong, K. et al. Risk factors for bat contact and consumption behaviors in Thailand; a quantitative study. *BMC Public Health* **20**, 841 (2020).
88. Valitutto, M. T. et al. Detection of novel coronaviruses in bats in Myanmar. *PLoS ONE* **15**, e0230802 (2020).
89. Phelps, K., Jose, R., Labonite, M. & Kingston, T. Assemblage and species threshold responses to environmental and disturbance gradients shape bat diversity in disturbed cave landscapes. *Diversity* **10**, 55 (2018).
90. Quibod, M. N. R. M. et al. Diversity and threats to cave-dwelling bats in a small island in the southern Philippines. *J. Asia-Pac. Biodivers.* **12**, 481–487 (2019).
91. Furey, N. M. & Racey, P. A. Conservation ecology of cave bats. In *Bats in the Anthropocene: Conservation of Bats in a Changing World* (eds C. C. Voigt & T. Kingston) 463–500 (Springer International Publishing, 2016).
92. Herkt, K. M. B., Skidmore, A. K. & Fahr, J. Macroecological conclusions based on IUCN expert maps: a call for caution. *Glob. Ecol. Biogeogr.* **26**, 930–941 (2017).
93. Jung, M. et al. A global map of terrestrial habitat types. *Sci. Data* **7**, 256 (2020).
94. Jung, M. et al. A global map of terrestrial habitat types (Version 001), <https://doi.org/10.5281/zenodo.3666246> (2020).
95. Faust, C. L. et al. Null expectations for disease dynamics in shrinking habitat: dilution or amplification. *Philos. Trans. R. Soc. B: Biol. Sci.* **372**, 20160173 (2017).
96. Redding, D. W., Moses, L. M., Cunningham, A. A., Wood, J. & Jones, K. E. Environmental-mechanistic modelling of the impact of global change on human zoonotic disease emergence: a case study of Lassa fever. *Methods Ecol. Evol.* **7**, 646–655 (2016).
97. Hassell, J. M. et al. Towards an ecosystem model of infectious disease. *Nat. Ecol. Evol.* **5**, 907–918 (2021).
98. Winter, D. J. rentrez: An R package for the NCBI eUtils API. *R. J.* **9**, 520–526 (2017).
99. South, A. rworldmap: A New R package for Mapping Global Data. *R. J.* **3**, 35–43 (2011).
100. Olival, K. J. et al. Possibility for reverse zoonotic transmission of SARS-CoV-2 to free-ranging wildlife: a case study of bats. *PLOS Pathog.* **16**, e1008758 (2020).
101. Anthony, S. J. et al. Global patterns in coronavirus diversity. *Virus Evolution* **3**, vex012 (2017).
102. Murakami, S. et al. Detection and characterization of Bat Sarbecovirus phylogenetically related to SARS-CoV-2, Japan. *Emerg. Infect. Dis.* **26**, 3025–3029 (2020).
103. Zhang, L. et al. Multilocus phylogeny and species delimitation within the philippinensis group (Chiroptera: Rhinolophidae). *Zoologica Scr.* **47**, 655–672 (2018).
104. Wilson, D. E. & Mittermeier, R. A. *Handbook of the Mammals of the World. Vol. 9. Bats.* (Lynx Edicions, 2019).
105. Srinivasulu, B. & Srinivasulu, C. In plain sight: Bacular and noseleaf morphology supports distinct specific status of Roundleaf Bats *Hipposideros pomona* Andersen, 1918 and *Hipposideros gentilis* Andersen, 1918 (Chiroptera: Hipposideridae). *J. Threatened Taxa* **10**, 12018–12026 (2018).
106. Rondinini, C. et al. Global habitat suitability models of terrestrial mammals. *Philos. Trans. R. Soc. B: Biol. Sci.* **366**, 2633–2641 (2011).
107. IUCN. *Habitats Classification Scheme (Version 3.1)*, <https://www.iucnredlist.org/resources/habitat-classification-scheme> (2021).
108. Williams, P. & Fong, Y. T. World Map of Carbonate Rock Outcrops v3.0 (ed The University of Auckland) (2010).
109. Ross, N. fasterize: Fast Polygon to Raster Conversion. R package version 1.0.3 (2020).
110. Hijmans, R. J. raster: Geographic Data Analysis and Modeling. R package version 3.4-5. (2020).
111. Chamberlain, S. & Boettiger, C. R Python, and Ruby clients for GBIF species occurrence data. *PeerJ Preprints* **5**, <https://doi.org/10.7287/peerj.preprints.3304v1> (2017).
112. Chamberlain, S. et al. rgif: Interface to the Global Biodiversity Information Facility API. R package version 3.6.0 (2022).
113. GBIF.org. *GBIF Occurrence Download*, <https://doi.org/10.15468/dl.8w26d8> (2021).
114. Feng, X. et al. A checklist for maximizing reproducibility of ecological niche models. *Nat. Ecol. Evol.* **3**, 1382–1395 (2019).
115. Zizka, A. et al. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).
116. WorldPop. *Unconstrained global mosaic 2020 (1km resolution)*, <https://doi.org/10.5258/SOTON/WP00647> (2018).
117. Greenberg, J. A. & Mattiuzzi, M. gdalUtils: Wrappers for the Geospatial Data Abstraction Library (GDAL) Utilities. R package version 2.0.3.2. (2020).
118. Carnell, R. lhs: Latin Hypercube Samples. R package version 1.1.1. (2020).
119. Signorell, A. et al. DescTools: Tools for Descriptive Statistics v. 0.99.41 (2021).
120. Delignette-Muller, M. L. & Dutang, C. fitdistrplus: an R Package for fitting distributions. *J. Stat. Softw.* **64**, 1–34 (2015).
121. Tan, C. W. et al. A SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage of ACE2–spike protein–protein interaction. *Nat. Biotechnol.* **38**, 1073–1078 (2020).
122. Tang, F. et al. Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J. Immunol.* **186**, 7264 (2011).
123. Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Computers Simul.* **55**, 271–280 (2001).
124. Iooss, B., Da Veiga, S., Janon, A. & Pujol, G. sensitivity: Global Sensitivity Analysis of Model Outputs. R package version 1.25.0. (2021).
125. Monod, H., Naud, C. & Makowski, D. Uncertainty and sensitivity analysis for crop models. In *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications* (eds Wallach, D., Makowski, D. & Jones, J.) (Elsevier Science, 2006).
126. Janon, A., Klein, T., Lagnoux, A., Nodet, M. & Prieur, C. Asymptotic normality and efficiency of two Sobol index estimators. *ESAIM: Probab. Stat.* **18**, 342–364 (2014).

## Acknowledgements

We acknowledge funding from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (R01AI110964 & NIAID-CREID U01AI151797 “EID-SEARCH”), the Samuel Freeman Charitable Trust, Pamela Thye, The Wallace Fund, and an Anonymous Donor c/o Schwab Charitable. L.F.W. is supported by grants from NRF (NRF2012NRF-CRP001-056; NRF2016NRF-NSFC002-013), NMRC (MOH-OFIRG19MAY-0011; COVID19RF-003) and MOE (MOE2019-T2-2-130) of Singapore. Z.L.S. was funded by the National Natural Science Foundation of China (31727901). We thank Cadhla Firth, Noam Ross, and Emma Mendelsohn for useful discussions.

## Author contributions

P.D. conceived the study, C.A.S., H.L., K.L.P., C.Z.T., and K.J.O. designed the methods, C.A.S. conducted the analyses, H.L. performed the literature searches, C.A.S. and H.L.

designed figures and tables, C.A.S. wrote the first draft of the manuscript, and H.L., K.L.P., C.Z.T., L.F.W., P.Z., Z.L.S., K.J.O., and P.D. provided comments and edits to the manuscript throughout the writing process. All authors approved the submitted manuscript and agree to be held accountable for their own contributions and ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated, resolved, and the resolution documented in the literature.

### Competing interests

P.D. served as a member of the WHO-China joint study on COVID-19 origins, is a current member of the Taskforce on the Origins and Early Spread of COVID-19 and One Health Solutions to Future Pandemics, and was Chair of the IPBES Pandemics and Biodiversity Workshop. P.D. has made numerous public statements both independently, and as part of these groups, on the likely origins of COVID-19 and the need to assess risk and prioritize targeted surveillance for future disease emergence. L.F.W. serves on multiple committees for WHO, FAO, and OIE on COVID-19 including assay and vaccine development and animal models; has ongoing research investigating the origin of COVID-19; and has made statements on this issue to the media. P.Z. serves on multiple committees for MOST (China) on COVID-19 including pathogen identification, monitoring viral mutations, and the likely origin of COVID-19. Z.L.S. has made statements to the media about the likely origins of COVID-19. The remaining authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-31860-w>.

**Correspondence** and requests for materials should be addressed to Peter Daszak.

**Peer review information** *Nature Communications* thanks Emma Glennon, James Hassell, Renata Muylaert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022