Ecology
Epidemiology and Global Health

Landscape drives zoonotic malaria prevalence in non-human primates

Emilia Johnson, Reuben Sunil Kumar Sharma, Pablo Ruiz Cuenca, Isabel Byrne, Milena Salgado-Lynn, Zarith Suraya Shahar, Lee Col Lin, Norhadila Zulkifli, Nor Dilaila Mohd Saidi, Chris Drakeley, Jason Matthiopoulos, Luca Nelli, Kimberly Fornace

School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland • Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK • Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK • Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia • Lancaster University, Bailrigg, Lancaster LA1 4YW • Liverpool School of Tropical Medicine, Pembroke Place Liverpool, L3 5QA, UK • School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK • Wildlife Health, Genetic and Forensic Laboratory, Sabah Wildlife Department, Wisma Muis, 88100, Kota Kinabalu, Malaysia • Danau Girang Field Centre, Sabah Wildlife Department, Wisma MUIS 88100, Kota Kinabalu Sabah, Malaysia • Department of Infection Biology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK • Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
October 23, 2023 (this version)

Reviewed preprint version 1
July 13, 2023

Sent for peer review
April 24, 2023

Posted to bioRxiv
April 13, 2023

Abstract

Zoonotic disease dynamics in wildlife hosts are rarely quantified at macroecological scales due to the lack of systematic surveys. Non-human primates (NHPs) host *Plasmodium knowlesi*, a zoonotic malaria of public health concern and the main barrier to malaria elimination in Southeast Asia. Understanding of regional *P. knowlesi* infection dynamics in wildlife is limited. Here, we systematically assemble reports of NHP *P. knowlesi* and investigate geographic determinants of prevalence in reservoir species. Meta-analysis of 6322 NHPs from 148 sites reveals that prevalence is heterogeneous across Southeast Asia, with low overall prevalence and high estimates for Malaysian Borneo. We find that regions exhibiting higher prevalence in NHPs overlap with human infection hotspots. In wildlife and humans, parasite transmission is linked to land conversion and fragmentation. By assembling remote sensing data and fitting statistical models to prevalence at multiple spatial scales, we identify novel relationships between *P. knowlesi* in NHPs and forest fragmentation. This suggests that higher prevalence may be contingent on habitat complexity, which would begin to explain observed geographic variation in parasite burden. These findings address critical gaps in understanding regional *P. knowlesi* epidemiology and indicate that prevalence in simian reservoirs may be a key spatial driver of human spillover risk.

eLife assessment

This study presents **useful** findings regarding the impact of forest cover and fragmentation on the prevalence of malaria in non-human primates. The evidence supporting the claims of the authors is, however, **incomplete**, as the sampling design cannot adequately address the geospatial issues that this study focuses on.

Introduction

Zoonotic infectious diseases arise from the spillover of pathogens into human populations, typically from a reservoir in wildlife hosts. Anthropogenic land use and land cover change have now been widely linked to infectious disease outbreaks (Brock et al., 2019; Davidson et al., 2019a; Loh et al., 2016). Such practices, including deforestation, logging, clearing for cash-crop plantations or conversion of intact forest into arable land, are accelerating across tropical forests of Southeast Asia (Fornace et al., 2021; Imai et al., 2018)(Fornace et al., 2021; Imai et al., 2018). Mechanisms that underly the association between habitat disturbance and spillover risk from wildlife hosts are complex and occur over multiple spatial scales (Brock et al., 2019). In Brazil, re-emergence of Yellow Fever Virus in both NHPs and humans has been linked to areas with highly fragmented forest (Ilacqua et al., 2021). In part, an increase in 'edge' habitat in fragmented or mosaic landscapes can facilitate spatial overlap and altered contact patterns between wildlife, vectors and humans (Lehman et al., 2006). Such ecological interfaces are also thought to contribute to parasite spillover in other vector-borne diseases including Zika (J. Li et al., 2021), Babesiosis and Lyme disease (Simon et al., 2014), *Trypanosoma cruzi* (Vaz et al., 2007) and zoonotic malaria (Brock et al., 2019; Grigg et al., 2017). At the same time, habitat fragmentation can have detrimental impact on wildlife population viability, with reduced host species occupancy and reduced disease burden in highly disturbed habitats (Hanski and Ovaskainen, 2000). Disentangling this interplay is essential to inform ecological strategies for surveillance and mitigation of diseases in regions undergoing landscape change (Fornace et al., 2021).

Zoonotic *P. knowlesi* is a public health threat of increasing importance across Southeast Asia, following the identification of a prominent infection foci in Borneo in 2004 (Singh et al., 2004). *P. knowlesi* is a zoonosis, with a sylvatic cycle circulating in non-human primates (NHPs). Human cases currently occur only from spillover events (Cuenca et al., 2022; Fornace, 2022; Fornace et al., 2023; Lee et al., 2011). Human transmission requires bites from infective mosquitos, primarily anopheline mosquitos of the Leucosphyrus Complex (*Anopheles balabacensis, An. latens, An. introlactus)* and Dirus Complex (*An. dirus, An. cracens)* (Moyes et al., 2016; Vythilingam et al., 2006; Wong et al., 2015). Natural hosts for *P. knowlesi* are typically Long-tailed macaques (*Macaca fascicularis*) and Southern Pig-tailed macaques (*M. nemestrina)* (Moyes et al., 2016), both occurring widely across Southeast Asia. Currently, distribution of *P. knowlesi* cases is thought to be restricted to the predicted ranges of known vector and host species (Davidson et al., 2019b), though recent studies have also identified other NHPs found to be harbouring *P. knowlesi.* This includes Stump-tailed macaques *(M. arctoides),* which are now considered to be another natural reservoir (Fungfuang et al., 2020).

Progress towards malaria elimination in Malaysia has been stymied by a recent rise in human incidence of *P. knowlesi* malaria. Even after accounting for increases in surveillance and diagnostic improvements it is now recognised as the most common cause of clinical malaria in Malaysia (Cooper et al., 2020). Indeed, Malaysia was the first country not to qualify for malaria elimination due to ongoing presence of zoonotic malaria and the World Health Organisation (WHO) updated the guidelines to reflect zoonotic malaria as a public health threat ("Global Malaria Programme," n.d.). Emergence of *Plasmodium knowlesi* infections has been linked to changes in land cover and land use (Fornace et al., 2021). While sporadic cases have been reported across Southeast Asia, including in Indonesia (Setiadi et al., 2016), the Philippines (Fornace et al., 2018), Vietnam (Maeno et al., 2015), Brunei (Koh et al., 2019) and Myanmar (Ghinai et al., 2017), the majority of *P. knowlesi* cases are found in East Malaysia (Borneo) with hotspots in the states of Sabah and Sarawak (Jeyaprakasam et al., 2020), areas that have seen extensive deforestation and landscape modification. In Sabah, human prevalence of *P. knowlesi* infection

has recently been shown to be specifically associated with recent loss of intact forest, agricultural activities, and fragmentation across multiple localised spatial scales (Brock et al., 2019; Fornace et al., 2019b, 2016).

Prevalence of the pathogen in reservoir hosts is one of three crucial factors determining the force of infection in zoonotic spillover events (Murray and Daszak, 2013). Despite this, very little is known of the impact of rapid landscape change on the distribution of *P. knowlesi* in NHPs. Literature on the impacts of fragmentation on primates tends to focus on primate density and abundance (Link et al., 2010; ZUNINO et al., 2007). What is known is that effects of land cover changes on primate-pathogen dynamics are highly variable and context-specific. Although the vector species responsible for sylvatic transmission remain unknown, the *Anopheles leucospryphus* group, the only vector group implicated in *P. knowlesi* transmission, is widely associated with secondary, disturbed forest (Brant, 2011; Hawkes et al., 2019; Wong et al., 2015). Macaques have been known to preferentially rely on fringe habitat, a behaviour that may be exaggerated in response to habitat fragmentation and facilitate exposure to vectors (Lehman et al., 2006; Stark et al., 2019). Changes to land composition can also create the biosocial conditions for higher rates of parasitism in primates. Under conditions of limited resources and reduction in viable habitat, conspecific primate density may increase as troops compete for available space. In turn, this can favour transmission via intra-species contact or allow the exchange of pathogens between troops dwelling in interior forest versus edge habitat (Faust et al., 2018; Stark et al., 2019). Habitat use may also become more intensive, preventing parasite avoidance behaviours (Nunn and Dokey, 2006). Land cover change is also known to favour more adaptable, synanthropic species such as *M. fascicularis* (McFarlane et al., 2012). Considering the spillover risk posed by wildlife reservoirs of *P. knowlesi,* clarifying any relationships between environmental factors and parasitaemia in key host species may contribute to a more comprehensive understanding of *P. knowlesi* transmission patterns.

Earth Observation (EO) data provides novel opportunities to investigate epidemiological patterns of diseases which are linked to environmental drivers (Kalluri et al., 2007). In relation to *P. knowlesi,* utility of fine-scale remote-sensing data has been demonstrated: examples include satellite-derived data used to examine household–level exposure risk in relation to proximate land configuration (Fornace et al., 2019b), UAV-imagery used to link real–time deforestation to macaque host behavioural change (Stark et al., 2019), and remote-sensing data used to interrogate risk factors for vector breeding sites (Byrne et al., 2021). Though macroecological studies that utilise geospatial data are often confounded by issues of matching temporal and spatial scales, as well as by the quality and accuracy of available georeferencing, measures can be taken to account for this when examining the role of environmental factors in modulating disease outcomes. Furthermore, ecological processes occur and interact over a range of distances, or 'spatial scales' (Brock et al., 2019; Fornace et al., 2016; Loh et al., 2016). This applies to determinants of vector-borne disease ecology, from larval breeding microclimate to wildlife host foraging behaviour. As multiple influential variables are rarely captured by a single scale (Cohen et al., 2016), data-driven methods can be applied to examine risk factors over multiple scales and identify covariates at their most influential extent (Byrne et al., 2021).

We hypothesise that prevalence of *P. knowlesi* in primate host species is spatially heterogeneous and that higher prevalence is partially driven by forest loss and fragmentation, contributing to the strong associations described between land use, land cover and human *P. knowlesi* risk. This study is the first to systematically assess *P. knowlesi* prevalence in NHPs at a regional scale, and across a wide range of habitats. In conceptual frameworks and transmission models, it is often assumed that *P. knowlesi* infections in NHPs are chronic (low level, persistent infection) and ubiquitous (uniformly distributed across populations) (Brock et al., 2016; Jeyaprakasam et al., 2020). No studies have systematically assessed the extent and quality of all available data on *P. knowlesi* in NHPs. Independent studies investigating *P. knowlesi* in primates are typically constrained by small sample sizes and confined geographic areas, limiting inference that can be made about

relationships between infection dynamics and landscape characteristics. Systematic tools developed for epidemiological studies of disease prevalence in human populations are rarely applied to the study of wildlife disease prevalence; however, such tools can be used to capture the scale and contrast required in macroecological studies to quantify disease burdens regionally. Furthermore, while recent research has shown the impact of deforestation on the distribution of macaques in the context of *P. knowlesi* (Moyes et al., 2016; Stark et al., 2019), associations between landscape and variation in the prevalence of simian *Plasmodium* spp. in primates have not been explored. We aimed to 1) assemble a georeferenced dataset of *P. knowlesi* in NHPs; 2) evaluate variation in NHP *P. knowlesi* prevalence by geographic region; and 3) assess environmental and spatial risk factors for *P. knowlesi* prevalence in NHPs across Southeast Asia.

Results

A systematic literature review was conducted in Medline, Embase and Web of Science to identify articles reporting prevalence of naturally acquired *Plasmodium knowlesi* in NHPs. 23 research articles were identified (Akter et al., 2015; Amir et al., 2020; Chang et al., 2011; Fungfuang et al., 2020; Gamalo et al., 2019; Ho et al., 2010; Jeslyn et al., 2011; Lee et al., 2011; M. I. Li et al., 2021; Muehlenbein et al., 2015; Putaporntip et al., 2010; Saleh Huddin et al., 2019; Seethamchai et al., 2008; Unpublished, 2015, 2013; Vythilingam et al., 2008; Zarith et al, 2021; Zhang et al., 2016), containing 148 unique primate survey records to form the dataset for analyses (see SI for details of JBI Critical Assessment, Table S5) (Munn et al., 2015). Year of sampling ranges from 2004–2019. No primatological studies were identified from Vietnam, Brunei or Timor-Leste. Full characteristics of the articles and individual study methodologies are reported in Supplementary Information (Table S2). Spatial resolution of the survey sites varied from GPS point coordinates to country-level administrative boundaries (Supplementary Table S7). Geographic distribution of sampling is illustrated in **Figure 1**.

Overall, records report on a total of 6322 primates, with the largest proportion sampled from Peninsular Malaysia (48.5%, n=3069/6322). Primate surveys were primarily conducted on Long-tailed macaques (*Macaca fascicularis*) (90.5%, n=5720/6322) followed by Pig-tailed macaques (*M. nemestrina)* (n=532/6322) (Amir et al., 2020; Lee et al., 2011; Muehlenbein et al., 2015; Putaporntip et al., 2010) (Table S3). Reported prevalence of *Plasmodium knowlesi* in NHPs ranged from 0%–100%. Only 87 of the surveys (58.8%, n=87/148) reported a positive diagnosis, with the remaining 61 sites finding no molecular evidence of *P. knowlesi* infection (41.2%) in any primates tested. A full breakdown of *P. knowlesi* infection rates according to reported primate characteristics can be found in SI, Table S4.

Meta-analysis of *P. knowlesi* prevalence

To quantify regional heterogeneity in simian cases of *P. knowlesi,* a one-stage meta-analysis of prevalence (number positive out of the number sampled) was conducted on primate malaria survey data. Overall pooled estimate for *P. knowlesi* prevalence was 11.99% (CI95% 9.35–15.26). Overall heterogeneity was assessed using the I^2 statistic. Substantial between-study heterogeneity (I^2 ≥75%) was found across all prevalence records (I^2=80.5%; CI95% 77.3–83.1). In the sub-group analysis by region, pooled prevalence estimates are consistently low for Thailand (2.0%, CI95% 1.1–3.5%), moderate in Peninsular Malaysia (14.3%, CI95% 11.1–18.2) and elevated in Singapore (23.3%, CI95% 11.0–42.8) and Malaysian Borneo (41.1%, CI95% 20.8–64.9) (**Figure 2**). Sub-group heterogeneity was assessed using prediction intervals, derived from τ^2 statistic used to describe between-study variability. Prediction intervals indicate high heterogeneity of estimates within regions, consistent with expectations of high variability of prevalence across individual study sites. Detailed forest plots for individual prevalence estimates can be found in Supplementary Figures S6.

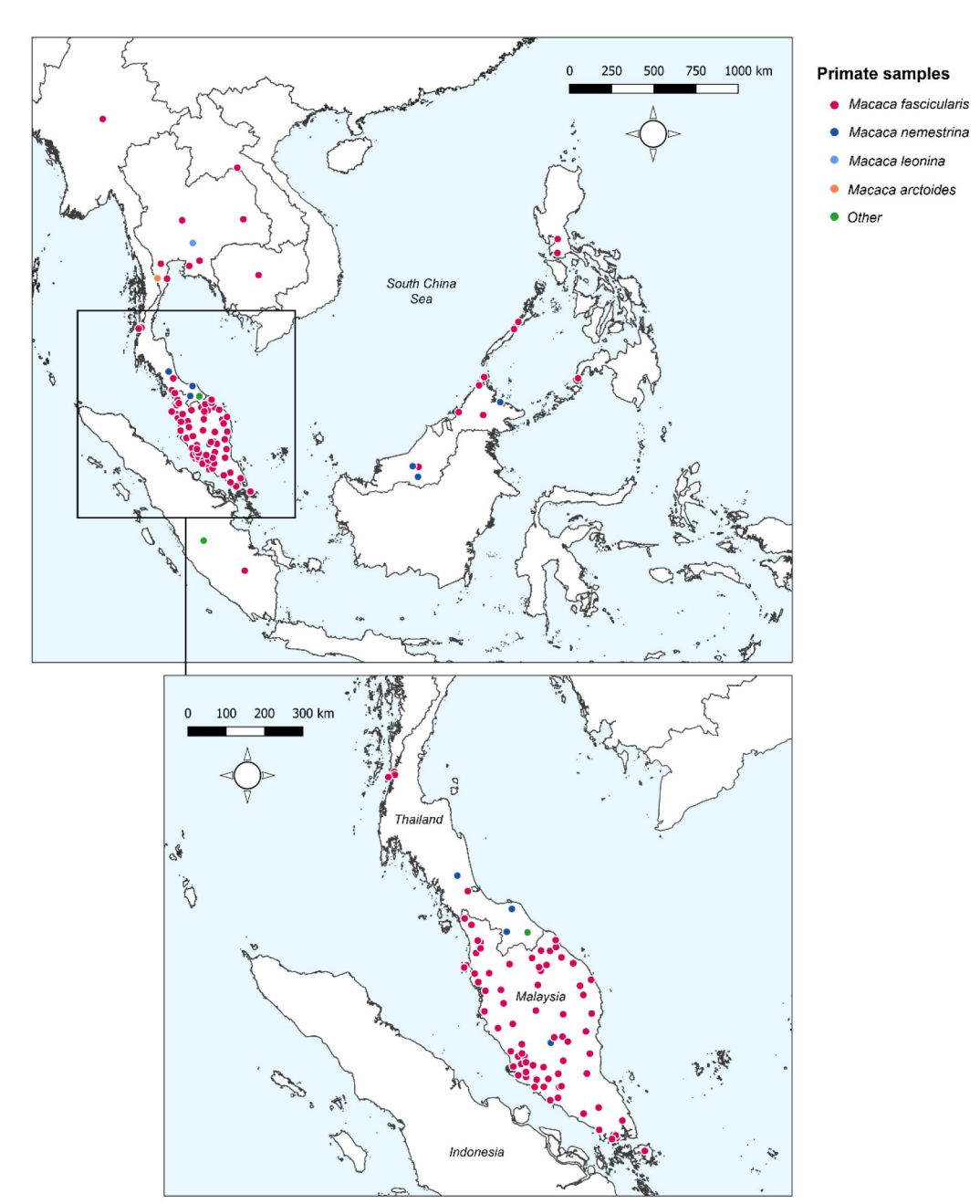

Figure 1.

Sampling sites and primate species sampled across Southeast Asia. 'Other' includes Trachypithecus obscurus *and undefined species from the genus* Presbytis. *Total surveys = 148*.

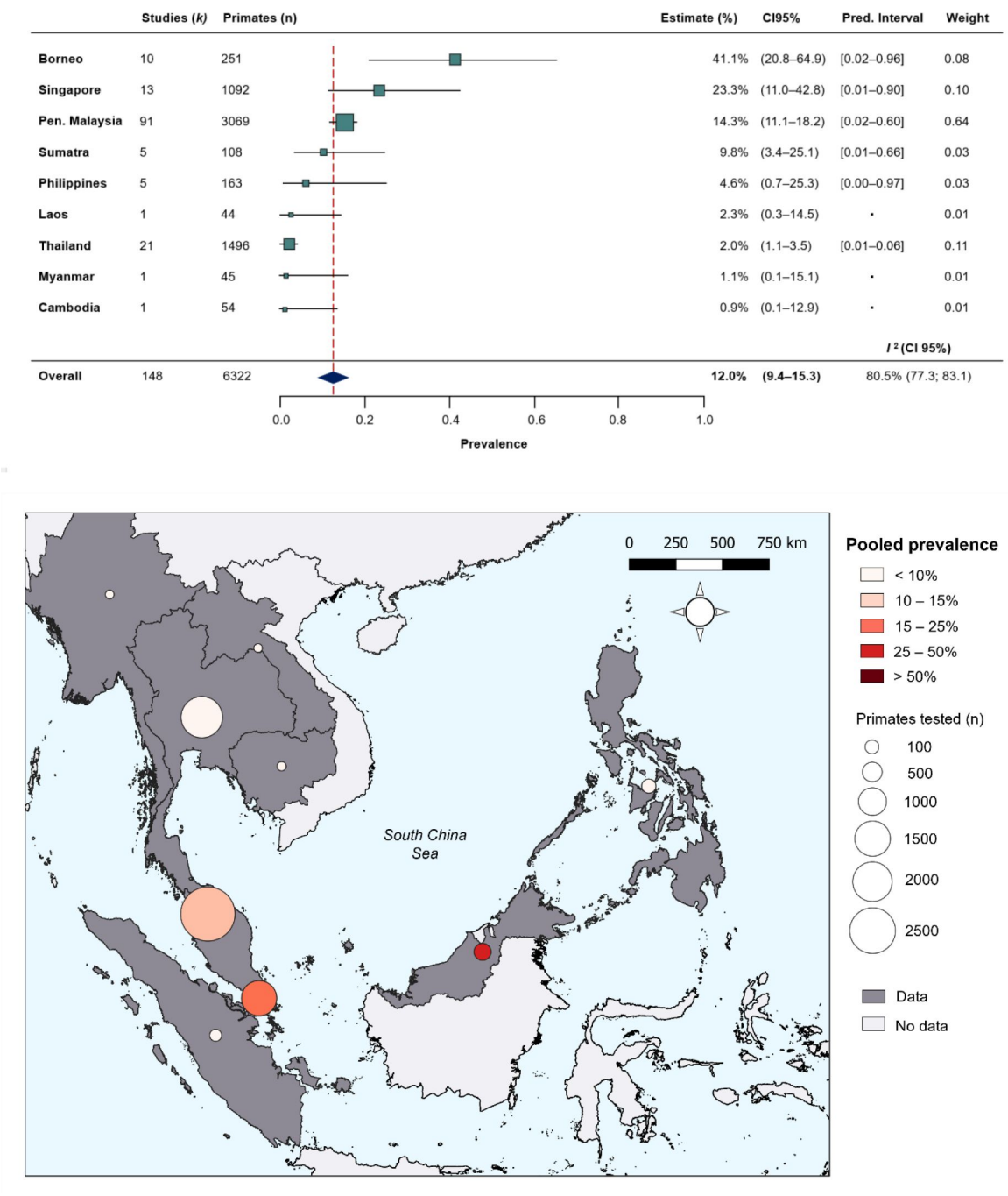

Figure 2

(A) *Forest plot of pooled estimates for* P. knowlesi *prevalence (%) in all non-human primates tested (n=6322) across Southeast Asia, disaggregated by species and sampling site (k=148). Random-effects meta-analysis, sub-grouped by region.* **(B)** *Map of regional prevalence estimates for P. knowlesi prevalence in NHP in Southeast Asia from meta-analysis. Point colour denotes pooled estimate (%). Size denotes total primates tested per region (n). Shading indicates data availability.*

Risk factor analysis

Covariate data and *P. knowlesi* prevalence data were used to fit additional models to explore the relationships between localised landscape configuration and NHP malaria prevalence. Environmental covariates were extracted from satellite-derived remote sensing datasets (**Table 1**) at either true sampling sites (GPS coordinates) or 10 random pseudo-sampling sites to account for geographic uncertainty in prevalence data. Host species was grouped as '*Macaca fascicularis*' or 'Other' due to sample counts of <10 for certain primate species. Only 57.4% (n=85/148 records) of data included year of sampling, deemed to be insufficient to assess temporal patterns in prevalence. Tree canopy cover ranged from negligible to near total cover (100%) within buffer radii (Supplementary Table S12). Details of covariate data processing is illustrated in Supplementary Information (Figure S7–8).

Following a two-stage approach for selection of explanatory variables, tree cover and fragmentation (measured by perimeter: area ratio, PARA) were retained at 5km as linear terms, human population density was retained at both 5km and 20km and primate species was retained as a categorical variable. Spearman's rank tests for residual correlation between final variables at selected scales indicates a strong negative correlation between tree cover and fragmentation index (PARA) (ρ= –0.75) (SI, Figure S14).

Adjusting for all other covariates in the model, we identified strong evidence of an effect between increasing tree canopy cover and higher prevalence of *P. knowlesi* in NHPs within a 5km radius (aOR=1.38, CI95% 1.19–1.60; p<0.0001). Evidence was also found for an association between likelihood of *P. knowlesi* and higher degrees of habitat fragmentation (PARA) within 5km (aOR=1.17, CI95% 1.02–1.34, p<0.0281). Evidence suggests that human population density within a 5km radius is associated with risk of *P. knowlesi* in NHP (aOR=1.36, CI95% 1.16–1.58, p=0.0001) whilst human density within 20km has an inverse effect on likelihood of *P. knowlesi* (aOR=0.56, CI95% 0.46–0.67, p<0.0001). *M. fascicularis* is also associated with higher prevalence relative to all other non-human primate species (aOR=2.50, CI95% 1.31–4.85; p=0.0051). Additional complexity did not improve optimal model fit and effect modification was not pursued. In sensitivity analyses removing datapoints with excessive spatial uncertainty, evidence was consistently found that tree canopy cover (5km) and host species exhibit a strong positive association with prevalence of *P. knowlesi* in NHP (Table S15-16). Final adjusted OR for the multivariable model can be visualised in **Figure 4** (Table S14).

Discussion

Land use and land cover change is widely linked to spillover of zoonotic pathogens from sylvatic reservoirs into human populations, and pathogen prevalence in wildlife host species is key in driving the force of infection in spillover events. Our initial analyses found that for *Plasmodium knowlesi*, there is substantial spatial heterogeneity and prevalence in non-human primates varies markedly between regions of Southeast Asia (Zhang et al., 2016). Consistent with our hypothesis that parasite density in primate hosts would be higher in areas experiencing habitat disturbance, we identified strong links between *P. knowlesi* in NHPs and measures of contemporaneous tree cover and habitat fragmentation. To our knowledge, this is the first systematic study to find evidence of landscape influencing the distribution of *P. knowlesi* prevalence in NHPs. Results offer evidence that *P. knowlesi* infection rates in NHPs are linked to changes in landscape across broad spatial scales, and that prevalence of *P. knowlesi* in reservoir species may be driving spillover risk across Southeast Asia. These findings could provide insight to improving surveillance of *P. knowlesi* and to the development of ecologically targeted interventions.

Covariate	Spatial res.	Temporal res.	Source
Human density (p/km^2)	1km	2012	WorldPop (WorldPop, 2018)
Elevation (m)	1km	2003	SRTM 90m Digital Elevation v4.1 (Jarvis, A., H.I. Reuter, A. Nelson, 2008)
Tree cover (1/0)	30m	Annual	Hansen's Global Forest Watch (Hansen et al., 2013)
Derivatives	Proportion canopy cover (%) Perimeter: area ratio (PARA>0)		

Table 1.

Spatial and temporal resolution and sources for environmental covariates. Summary metrics extracted within 5, 10 and 20km circular buffers.

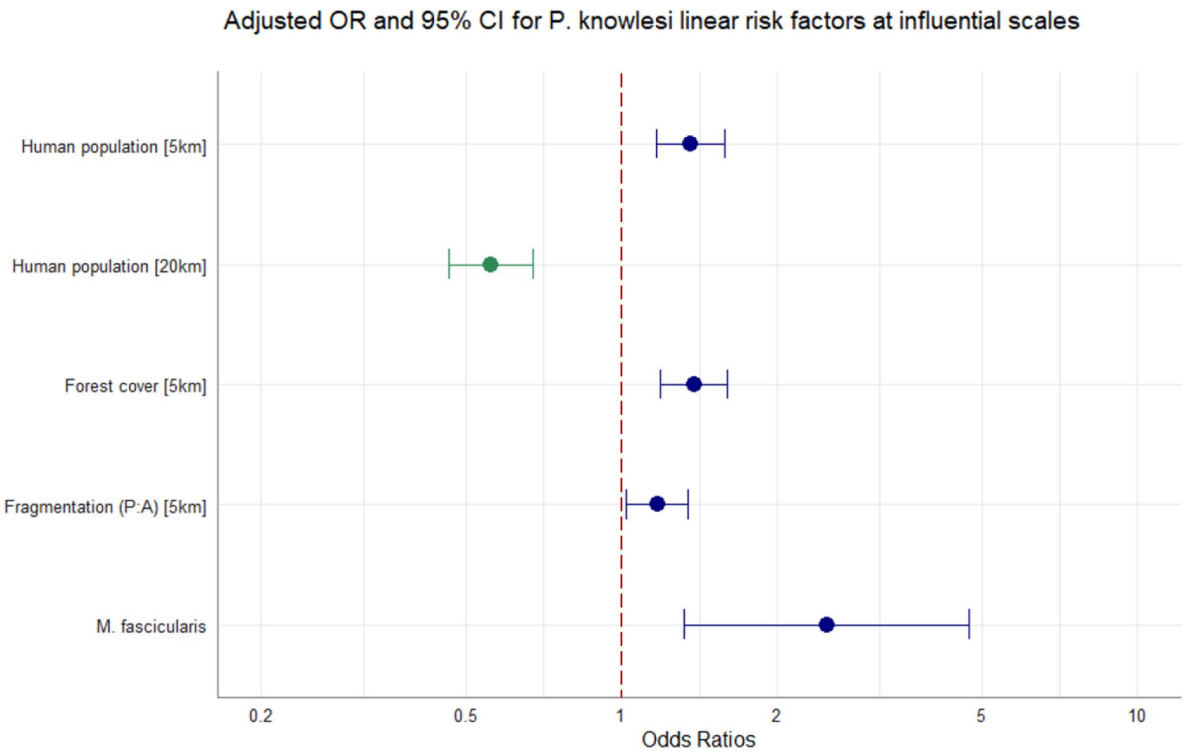

Figure 4.

Multivariable regression results. Spatial scale denoted in square bracket. Canopy cover = %. Adjusted OR (dots) and CI95% (whiskers) for factors associated with P. knowlesi *in NHPs at significant spatial scales. N=1354, accounting for replicate pseudo-sampling*.

While previous studies have estimated that *P. knowlesi* infection would be chronic in all macaques, or as high as 50–90% for modelling *P. knowlesi* transmission in Malaysia (Brock et al., 2016), this data strongly suggests that this is not the case. Overall prevalence of *P. knowlesi* infection in all NHPs is markedly lower than usual estimates, emphasising the importance of accounting for absence data in estimations of prevalence. Considerable heterogeneity was identified between and within regional estimates for *P. knowlesi* across Southeast Asia, which likely reflects genuine differences according to distinct climates and habitats (Shearer et al., 2016). Malaysian Borneo was found to have an estimated prevalence over five-fold higher than West Malaysia. Crucially, such extreme prevalence estimates for NHPs in Borneo align with the known hotspot for human incidence of *P. knowlesi* (Cooper et al., 2020). By comparison, for Peninsular Malaysia, estimated prevalence is far lower than anticipated. Cases of human *P. knowlesi* do occur in West Malaysia, though transmission has been found to exhibit spatial clustering (Phang et al., 2020) which may correspond to pockets of high risk within the wider context of low prevalence of *P. knowlesi* in macaque populations. Regional trends in *P. knowlesi* also mask differences in infection rates between sample locations, driven by more localised factors. Multiple studies reported finding *P. knowlesi* infections in wild macaques to be low or absent in peri-domestic or urbanised areas, attributed to the absence of vector species typically found in forest fringes (Brant et al., 2016; Chua et al., 2019; Manin et al., 2016). This pattern is seen in reports from Peninsular Malaysia (Saleh Huddin et al., 2019; Vythilingam et al., 2008), Singapore (Jeslyn et al., 2011; M. I. Li et al., 2021) and Thailand (Fungfuang et al., 2020; Putaporntip et al., 2010). The high heterogeneity of reports here suggests that the picture is even more complex. *P. knowlesi* infections may even vary between troops within a single study site, as was seen in the Philippines (Gamalo et al., 2019). Fine-scale interactions are unlikely to be captured by the scale of this study.

Ecological processes determining *P. knowlesi* infection are influenced by dynamic variables over multiple spatial scales (Cohen et al., 2016). We utilised a data-driven methodology to select variables at distances that capture maximum impact on *P. knowlesi* prevalence (Byrne et al., 2021; Fornace et al., 2019b), with tree cover and fragmentation influential at localised scales and human population density also exerting influence within wider radii. Contrary to previous studies on risk factors for human incidence of *P. knowlesi* (Fornace et al., 2019b, 2016), elevation was not found to be associated with *P. knowlesi* in NHPs at any scale. Vector and host species composition vary substantially across tropical ecotones, and it is likely that the study extent encompasses a range of putative vectors across different landscapes, such as those of the Minimus Complex in northern regions (Parker et al., 2015) or the recently incriminated *An.-collessi* and *An.-roperi* from the Umbrosus Group (de Ang et al., 2021). Given that the vector species driving sylvatic transmission remain elusive, it is conceivable that the elevation range covers multiple vector and host species niches and explains the lack of observed relationship between elevation and *P. knowlesi* in NHPs. Human population density was found to be significant at multiple distances, with contrasting effects on parasite prevalence in NHP. Previous studies have found a negative association between human density and vector density and biting rates in forested landscapes (Fornace et al., 2019a). Across wide spatial scales, increased vector density in less populated, more forested areas could generate higher parasite prevalence in NHPs. At the same time Long-tailed macaques, a species shown here to have higher prevalence rates, are notorious as nuisance animals and many of the available samples were collected opportunistically in urban areas, which might underly the observed positive association between localised high human density and higher prevalence in NHP. Whilst more data would be needed to understand this interaction, this further demonstrates the importance of using approaches to identify disease dynamics across multiple spatial scales (Brock et al., 2019).

A key finding is the link between high prevalence of *P. knowlesi* in primate host species with high degrees of habitat fragmentation. Habitat fragmentation is a key aspect of landscape modification, where large contiguous areas of habitat (for example, forests) are broken into a mosaic of smaller patches. This disturbs the ecological structure by increasing the density of fringes or ‘edges’,

dynamic habitat often at the boundaries between natural ecosystems and human-modified landscapes (Borremans et al., 2019). Other studies have linked habitat fragmentation to increased generalist parasite density in primates. In Uganda, a higher prevalence and infection risk of protozoal parasites was observed in wild populations of red colobus primates (*Procolobus rufomitratus*) inhabiting fragmented forests compared to those in undisturbed habitat (Gillespie and Chapman, 2008). For *P. knowlesi,* creation of edge habitat is thought to favour vectors of the Leucosphyrus Complex (Davidson et al., 2019a; Hawkes et al., 2019). *Anopheles* spp. presence can be predicted by indices of fragmentation in Sabah, Borneo, with land cover changes creating more suitable micro-climate for larval habitats (Byrne et al., 2021), and an increased abundance of *An. balabacensis* found in forest fringes (Hawkes et al., 2019; Wong et al., 2015). Increasing landscape complexity results in increased density of edge habitat, with conceivably higher density of vectors in forest fringes. Therefore, preferential use of fringe habitat and high exposure to vectors in forest fringes may contribute to higher conspecific transmission of *P. knowlesi* between primates in increasingly fragmented habitats. This finding also lends clarity to landscape fragmentation as a risk factor for human exposure to *P. knowlesi* in Malaysian Borneo (Brock et al., 2019; Fornace et al., 2019b), with changes in relative host density, vector density and wildlife parasite prevalence in nascent forest fringes potentially enhancing the spillover of this disease system into human populations in fragmented habitats.

Conversely, we saw a strong association between high parasite prevalence and high tree canopy coverage. Given that a strong inverse relationship with fragmentation was observed, with high tree density correlating to low fragmentation indices and vice versa, this speaks to a trade-off between dense canopy cover and high habitat complexity and suggests an 'ideal' amount of habitat fragmentation that facilitates prevalence in primate hosts. For animals with larger home ranges, individual-based disease models combined with movement ecology approaches have shown that the most highly fragmented areas are less favourable for maintaining parasite transmission (White et al., 2018). In Sabah, individual macaques were shown to increase ranging behaviour in response to deforestation (Stark et al., 2019). Forest edge density also peaks at intermediate levels of land conversion (Borremans et al., 2019). With smaller habitat patches in maximally fragmented landscapes potentially insufficient to support macaque troops, this interplay between disease ecology and metapopulation theory may explain why both tree density and habitat fragmentation appear to pose a greater risk for simian *P. knowlesi.* Likewise, this may relate to the finding that in Borneo, larger forest patches (lower fragmentation indices) were associated with *P. knowlesi* spillover in Borneo (Fornace et al., 2019b). Overall, this finding offers an insight to mechanisms that underpin the increased force of infection of *P. knowlesi* that is associated with landscape change.

There are limitations to consider in the available data and interpretation of these findings. 'Small-study effects' were observed in the dataset, suggestive of a bias toward positive effect estimates (Stewart et al., 2012). This may be a result of data disaggregation and small studies creating artefactually higher estimates or may reflect true bias in data collection toward areas known to be endemic for *P. knowlesi* and convenience sampling of macaques. Assumptions have also been made that sample site equates to habitat, which may not reflect actual habitat use, and even accurate georeferenced datapoints are unlikely to entirely reflect surrounding habitat within the macaque home range. Variability in study designs and data reporting also impacted geospatial accuracy. Steps were taken to account for spatial bias by extracting covariates at randomly generated pseudo-sampling points. Whilst uncertainty cannot be eliminated, we demonstrate a robust methodology to accommodate for geographical uncertainty in ecological studies. Future investigations should prioritise systematic, georeferenced sampling across a range of landscape scenarios.

Results show important regional ecological trends, but broad geographic patterns may not be generalisable at individual levels, or to all putative host species in all geographic contexts (Zhang et al., 2016). Follow up studies should be conducted at higher spatial and temporal resolution to

characterise the effect of local landscape configuration on wildlife *P. knowlesi* prevalence. Effects of fragmentation are likely to be dependent on land conversion type, species composition and surrounding matrix habitat (Fornace et al., 2019b). Use of perimeter: area ratio (PARA) as a fragmentation index was justified given high canopy coverage in study sites (Wang et al., 2014), though Edge Density (ED) or normalised Landscape Shape Index (nLSI) might be more appropriate in future analyses to account for variation in forest abundance. Specific land configurations have previously been linked to *P. knowlesi* exposure in Borneo (Fornace et al., 2019b), notably in areas where palm oil plantation is a dominant industry. Given this, broad forest classifications used here may mask important differences in *P. knowlesi* prevalence between land classes. As it was not possible to include contemporary land cover classifications in this analysis, future studies would also benefit from looking at specific habitat type (e.g., primary forest, agroforest, plantation).

Concluding remarks

Strong links have been identified between land use and land cover change and ecosystem perturbation that favours the transmission of vector-borne diseases (Loh et al., 2016). Prevalence of *P. knowlesi* in macaques is likely to be a crucial determinant of human infection risk, and more representative estimates of *P. knowlesi* prevalence derived here can better inform regional transmission risk models. This study also characterises landscape risk factors for heightened prevalence of *P. knowlesi* in NHPs. Findings provide evidence that *P. knowlesi* in primate hosts is partly driven by landscape modification across Southeast Asia. While the full complexity is not captured by the covariates used, it is clear that *P. knowlesi* infection in NHPs is not restricted to densely forested areas. This study also demonstrates the utility of systematic meta-analysis tools and remote-sensing datasets in the investigation of macroecological disease trends, in conjunction with methods to standardise a spatially heterogeneous dataset and data-driven selection of spatial scales. Gaps identified in data reporting should inform more systematic and localised primatological surveys to disentangle precise mechanisms. Notwithstanding limitations, this study highlights the marked spatial heterogeneity and role of landscape complexity in driving *P. knowlesi* infection rates in NHPs. Given the clear intersection between human epidemiology and wildlife ecology, it is essential that infection dynamics within wildlife reservoirs are considered in future public health interventions.

Methods

Study site

This study focused on the simian malaria *Plasmodium knowlesi* across Southeast Asia, within 28°30’00.0"N, 92°12’00.0"E and 11°00’00.0"S, 141°00’00.0"E. Climate mainly corresponds to the equatorial tropical zone, with high temperatures and high humidity.

Data assembly

A systematic literature review was conducted under the CoCoPop framework (Condition, Context, Population) (Cuenca et al., 2022; Munn et al., 2015). All studies identified in the literature review were screened for data on NHPs with a confirmed *P. knowlesi* diagnosis or absence data (zero counts of *P. knowlesi* with appropriate diagnostic methods). Exclusion criteria included (a) studies exclusively relying on microscopy (Antinori et al., 2013) (b) laboratory, animal model or experimental infection studies (c) data from outside of Southeast Asia. No limit was set on the temporal range for primate survey records. Duplicate records reporting results from the same surveys were removed, with one record per survey retained. Critical appraisal of the studies was

conducted using the Joanna Briggs Institute (JBI) checklist for prevalence studies (Munn et al., 2015) (see Supplementary Information (SI) for details and criteria). A flowchart of the selection process is illustrated in Figure S3, with a full list of articles included provided in Table S2.

Primary outcome was defined as *P. knowlesi* prevalence (*p*, proportion positive for *P. knowlesi* infection from *n* sampled NHPs). For each independent primate study, the following variables were extracted: year of data collection, primate species sampled, primate status (wild/captive), diagnostic test (PCR/sequencing) and target gene(s), sampling method (routine/purposive), number of *P. knowlesi* positive samples, number of *Plasmodium* spp. positive samples, total number of primates tested and geographical information.

In most studies identified, study site was only geolocated to a geographic area or descriptive location. Geolocation was assigned at the lowest available level of administrative polygon (i.e., district/state/country) by cross-referencing reported sampling location with GADM (v3.6) administrative boundaries. If specific location was given, GPS coordinates were assigned via Google Maps. For data visualisation, point coordinates were plotted in QGIS (3.10.14) and R (4.1.0) software.

Meta-analysis of *P. knowlesi* prevalence

Meta-analysis was conducted using methods that are standard in the analysis of human disease prevalence for individual participant datasets (IDP) (Liberati et al., 2009; Stewart et al., 2012). Data were disaggregated by geographic location (site) and primate species, to illustrate variance in prevalence by survey unit (Stewart et al., 2012). One-stage meta-analysis is considered appropriate for studies where the outcome may be infrequent, so data was included in a single model under the 'DerSimonian and Laird' variance estimator (Munn et al., 2015). Sensitivity analyses were conducted to compare methods for the back-transformation of prevalence estimates. For studies where prevalence estimates tend towards 0% or 100%, variance tends towards 0. To stabilise the variance and enable back-transformation of zero prevalence records, logit method was selected for the transformation of prevalence, with the inverse variance method used for individual study weights (see SI for details).

Overall heterogeneity of prevalence records was assessed using the I^2 statistic (Hippel, 2015), a relative estimate of true between-study variance. Sub-group analysis was conducted according to geographic region, with the heterogeneity of reported prevalence within regional sub-groups assessed using prediction intervals derived from the τ^2 statistic. Small-study effects, including selection and publication biases, were assessed by examining funnel plots and imputing 'missing' estimates using the trim-and-fill method (Lin and Chu, 2018). Full rationale and details of small-study effect assessments can be found in Supplementary Information.

Remote sensing data

Satellite-derived remote sensing datasets were used to assemble local environmental and anthropogenic covariates. Gridded UN-adjusted human population estimates were assembled at 1km resolution from WorldPop (WorldPop, 2018). Elevation data was obtained from NASA SRTM 90m Digital Elevation Database v4.1 (CGIAR-CSI) (Jarvis, A., H.I. Reuter, A. Nelson, 2008) with a spatial resolution of 1km. Contemporaneous tree cover was derived from Hanson's Global Forest Watch (30m) (Hansen et al., 2013), extracted for every year between 2006–2020.Tree cover was classified as ≥50% crown density, and then matched to primate data by sample site geolocation and by year of sample collection to account for rapid forest loss (SI, Figure S7). Where a broad timeframe of sampling was provided (≥3 years), median year was used. Full details for variable selection and processing can be found in Supplementary Information (Table S11–12, Figure S9).

Perimeter: area ratio (PARA, ratio of patch perimeter length to patch surface area) of given land class is a key metric for habitat conversion, where a higher PARA provides a measure of boundary complexity and indicates a more fragmented landscape (McGarigal and Cushman, 2012). Mean PARA was extracted from canopy cover within circular buffers. Habitat fragmentation has been shown to correlate with disease transmission parameters (Borremans et al., 2019; Faust et al., 2018), but definitions often lack precision and can be considered with respect to 'separation effects' (division and isolation of patches) and 'geometric effects' (changes to ratios of perimeter and core habitat) (Wilkinson et al., 2018). PARA provides a measure of edge density within the buffer area (PARA>0) and has been shown to provide a good index of fragmentation and good discrimination of spatial aggregation across areas where habitat abundance (tree canopy cover) is high. (Wang et al., 2014) (SI, Table S12, Figure S10).

Covariate assembly

For studies with exact GPS coordinates, precise environmental data at a single site could be obtained. For surveys published without GPS coordinates, there is considerable geographic uncertainty in the exact sampling location (SI, Table S7). Uncertainty in the spatial and environmental determinants of prevalence generates a sampling bias, with the precision of covariates correlated to certain studies. Use of a single centroid proxy site is standard procedure, but often generates erroneous estimates in large or heterogenous sampling units(Cheng et al., 2021). Alternative strategies were employed to account for and mitigate the effect of spatial uncertainty and spatial bias. Each prevalence observation was replicated and assigned a random sample of environmental realisations. 10 random sampling points were generated within the sampling area provided by the study, and covariates were extracted at each proxy sampling site (SI, Table S8). Selection of random points was validated by visual inspection of the stability of model coefficients with the inclusion of an increasing number of points. Number of points was selected conservatively at the point where coefficients stabilised (n=10).

For every georeferenced sampling point, mean values for all selected covariates were extracted within buffer radii at 5km, 10km and 20km (SI, Figure S11). Buffer area sizes were selected to investigate multiple spatial scales over which associations between risk factors and *P. knowlesi* prevalence might occur. A minimum radius of 5km was chosen to approximate the maximum ranging distance for *M. fascicularis* (Waxman et al., 2014), with wider radii (10–20km) included to account for the geographic uncertainties in areal data. Flowchart of data processing chain can be found in Supplementary Information (Figure S8).

Analysis of environmental risk factors

Generalised linear mixed-effect regression models (GLMM) were fitted to NHP prevalence data using a binomial distribution with a logit link. To account for within-study correlation in reported average prevalence, a unique identifier combining author and study was included as a random intercept in all models. Artificial inflation of sample size in the replicated data (10 pseudo-sampling sites for data geolocated to administrative areas) was accommodated by reducing individual observation weights to $1/10^{th}$ within the model.

Each covariate at each spatial scale was assessed for inclusion in the multivariable model based on bivariable analysis and a criterion of p >0.2 under likelihood ratio tests (LRT) (Table S13). A quadratic term for the fragmentation index 'PARA' was included to account for possible nonlinearity. Multicollinearity among independent predictors at multiple scales was examined via variance inflation factors (VIF). The VIF of each predictor variable was examined following a stepwise procedure, starting with a saturated model and sequentially excluding the variable with the highest VIF score from the model. Stepwise selection continued in this manner until the entire subset of explanatory variables in the global model satisfied a moderately conservative threshold of VIF ≤6 (Rogerson, 2001). Qualifying variables obtained were then assessed for model

inclusion using a backward stepwise strategy, removing variables with the highest *p* value (LRT) until a pre-defined threshold of $\alpha < 0.05$. Spearman's rank tests were conducted on the selected variables to observe residual correlation, plotted as a correlation matrix (Figure S14).

Fully adjusted OR for associations between environmental covariates and *P. knowlesi* prevalence were derived from the final multivariable GLMM with *p* values derived from LRT. Sensitivity analyses were conducted by excluding data points from administrative boundaries outside a reasonable size or above a reasonable threshold of environmental certainty, according to the standard deviation (SD) of the covariate values within each set of 10 environmental realisations (Table S15 and Table S16).

Ethics

Ethics approval was not required for this research, following assessment by the LSHTM Research Governance & Integrity Office (LSHTM MSc Ethics Ref: 25429).

Acknowledgements

Research was supported by the Sir Henry Dale Fellowship, jointly funded by the Wellcome Trust and the Royal Society (Grant Number 221963/Z/20/Z). Data collected in Peninsular Malaysia by the Faculty of Veterinary Medicine, Universiti Putra Malaysia, was funded by the Ministry of Higher Education Malaysia (Grant LRGS/1/2018/UM/01/1/). Additional funding was supported by the World Health Organization.

References

Akter R, Vythilingam I, Khaw LT, Qvist R, Lim YA-LL, Sitam FT, Venugopalan B, Sekaran SD, Vythilingam I, Khaw LT, Lim YA-LL, Sitam FT, Venugopalan B, Sekaran SD, Qvist R, Lim YA-LL, Sitam FT, Venugopalan B, Sekaran SD, Vythilingam I, Khaw LT, Lim YA- LL, Sitam FT, Venugopalan B, Sekaran SD (2015) **Simian malaria in wild macaques: first report from Hulu Selangor district, Selangor, Malaysia** *Malar J* **14** https://doi.org/10.1186/s12936-015-0856-3

Amir A, Shahari S, Liew JWK, de Silva JR, Khan MB, Lai MY, Snounou G, Abdullah ML, Gani M, Rovie-Ryan JJ, Lau YL. (2020) **Natural Plasmodium infection in wild macaques of three states in peninsular Malaysia** *Acta Trop* **211**:1–6 https://doi.org/10.1016/j.actatropica.2020.105596

Antinori S, Galimberti L, Milazzo L, Corbellino M (2013) **Plasmodium knowlesi: The emerging zoonotic malaria parasite** *Acta Trop* **125**:191–201 https://doi.org/10.1016/j.actatropica.2012.10.008

Borremans B, Faust C, Manlove KR, Sokolow SH, Lloyd-Smith JO (2019) **Cross-species pathogen spillover across ecosystem boundaries: Mechanisms and theory** *Philosophical Transactions of the Royal Society B: Biological Sciences* **374** https://doi.org/10.1098/rstb.2018.0344

Brant HL (2011) **Changes in abundance, diversity and community composition of mosquitoes based on different land use in Sabah** *Malaysia. Imperial College London*

Brant HL, Ewers RM, Vythilingam I, Drakeley C, Benedick S, Mumford JD (2016) **Vertical stratification of adult mosquitoes (Diptera: Culicidae) within a tropical rainforest in Sabah, Malaysia** *Malar J* **15**:1–9 https://doi.org/10.1186/s12936-016-1416-1

Brock PM, Fornace KM, Grigg MJ, Anstey NM, William T, Cox J, Drakeley CJ, Ferguson HM, Kao RR, Fornace KM, Cox J, Drakeley CJ, Grigg MJ, Anstey NM, William T, Kao RR (2019) **Predictive analysis across spatial scales links zoonotic malaria to deforestation** *Proc Biol Sci* **286** https://doi.org/10.1098/rspb.2018.2351

Brock PM, Fornace KM, Parmiter M, Cox J, Drakeley CJ, Ferguson HM, Kao RR (2016) **Plasmodium knowlesi transmission: integrating quantitative approaches from epidemiology and ecology to understand malaria as a zoonosis** *Parasitology* **143**:389–400 https://doi.org/10.1017/S0031182015001821

Byrne I, Drakeley CJ, Fornace KM, Manin BO, Chua TH, Aure W, Vythilingam I, Ferguson HM (2021) **Environmental and spatial risk factors for the larval habitats of Plasmodium knowlesi vectors in Sabah, Malaysian Borneo** *Sci Rep* **11** https://doi.org/10.1038/s41598-021-90893-1

Chang Q, Sun X, Wang J, Yin J, Song J, Peng S, Lu H, Zhou H, Jiang N, Chen Q (2011) **Identification of Hepatocystis species in a macaque monkey in northern Myanmar** *Res Rep Trop Med* **2**:141–146

Cheng Y, Tjaden NB, Jaeschke A, Thomas SM, Beierkuhnlein C (2021) **Using centroids of spatial units in ecological niche modelling: Effects on model performance in the context of environmental data grain size** *Global Ecology and Biogeography* **30**:611–621 https://doi.org/10.1111/geb.13240

Chua TH, Manin BO, Vythilingam I, Fornace K, Drakeley CJ (2019) **Effect of different habitat types on abundance and biting times of Anopheles balabacensis Baisas (Diptera: Culicidae) in Kudat district of Sabah, Malaysia** *Parasit Vectors* **12** https://doi.org/10.1186/s13071-019-3627-0

Cohen JM, Civitello DJ, Brace AJ, Feichtinger EM, Ortega CN, Richardson JC, Sauer EL, Liu X, Rohr JR (2016) **Spatial scale modulates the strength of ecological processes driving disease distributions** *Proceedings of the National Academy of Sciences* **113**:E3359–E3364 https://doi.org/10.1073/pnas.1521657113

Cooper DJ, Rajahram GS, William T, Jelip J, Mohammad R, Benedict J, Alaza DA, Malacova E, Yeo TW, Grigg MJ, Anstey NM, Barber BE (2020) **Plasmodium knowlesi Malaria in Sabah, Malaysia, 2015-2017: Ongoing increase in incidence despite nearelimination of the human-only plasmodium species** *Clinical Infectious Diseases* **70**:361–367 https://doi.org/10.1093/cid/ciz237

Cuenca R, Journal M, Cuenca PR, Key S, Lindblade KA, Vythilingam I, Drakeley C, Fornace K (2022) **Is there evidence of sustained human - mosquito - human transmission of the zoonotic malaria Plasmodium knowlesiA systematic literature review** *Malar J* :1–19 https://doi.org/10.1186/s12936-022-04110-z

Davidson G, Chua TH, Cook A, Speldewinde P, Weinstein Tock H., Weinstein P, Weinstein Tock H., Weinstein P. (2019) **The Role of Ecological Linkage Mechanisms in Plasmodium knowlesi Transmission and Spread** *Ecohealth* **16**:594–610 https://doi.org/10.1007/s10393-019-01395-6

Davidson G, Chua TH, Cook A, Speldewinde P, Weinstein P (2019) **Defining the ecological and evolutionary drivers of Plasmodium knowlesi transmission within a multi-scale framework** *Malar J* **18** https://doi.org/10.1186/s12936-019-2693-2

de Ang JX, Yaman K, Kadir KA, Matusop A, Singh B. (2021) **New vectors that are early feeders for Plasmodium knowlesi and other simian malaria parasites in Sarawak, Malaysian Borneo** *Sci Rep* **11**:1–12 https://doi.org/10.1038/s41598-021-86107-3

Faust CL, McCallum HI, Bloomfield LSP, Gottdenker NL, Gillespie TR, Torney CJ, Dobson AP, Plowright RK (2018) **Pathogen spillover during land conversion** *Ecol Lett* **21**:471–483 https://doi.org/10.1111/ele.12904

Fornace K (2022) **No evidence of sustained nonzoonotic Plasmodium knowlesi transmission: estimating species-specic spatiotemporally varying reproduction numbers for malaria in Malaysia** *Nature Portfolio* https://doi.org/10.21203/rs.3.rs-2278329/v1

Fornace KM, Abidin TR, Alexander N, Brock P, Grigg MJ, Murphy A, William T, Menon J, Drakeley CJ, Cox J, Abidin TR, Brock P, Grigg MJ, Murphy A, William T, Menon J (2016) **Association between landscape factors and spatial patterns of Plasmodium knowlesi infections in Sabah, Malaysia** *Emerg Infect Dis* **22**:201–208 https://doi.org/10.3201/eid2202.150656

Fornace KM, Alexander N, Abidin TR, Brock PM, Chua TH, Vythilingam I, Ferguson HM, Manin BO, Wong ML, Ng SH, Cox J, Drakeley C (2019) **Local human movement patterns and land use impact exposure to zoonotic malaria in Malaysian borneo** *Elife* **8** https://doi.org/10.7554/ELIFE.47602

Fornace KM, Brock PM, Abidin TR, Grignard L, Herman LS, Chua TH, Daim S, William T, Patterson CLEBEB, Hall T, Grigg MJ, Anstey NM, Tetteh KKAA, Cox J, Drakeley CJ (2019) **Environmental risk factors and exposure to the zoonotic malaria parasite Plasmodium knowlesi across northern Sabah** *Malaysia: a population-based cross-sectional survey. Lancet Planet Health* **3**:e179–e186 https://doi.org/10.1016/S2542-5196(19)30045-2

Fornace KM, Herman LS, Grignard L, Cox J, Tetteh KKA, Drakeley CJ, Abidin TR, Chua TH, Daim S, Lorenzo PJ, Espino F, Nuin NA, Ying LT, Grigg MJ, William Kimberly M., Lorenzo AO-, Pauline J. (2018) **Exposure and infection to Plasmodium knowlesi in case study communities in Northern Sabah, Malaysia and Palawan, The Philippines** *PLoS Negl Trop Dis* **12** https://doi.org/10.1371/journal.pntd.0006432

Fornace KM, Lines J, Drakeley CJ, Diaz Kimberly M., Diaz A v., Lines J, Drakeley CJ. (2021) **Achieving global malaria eradication in changing landscapes** *Malar J* **20** https://doi.org/10.1186/s12936-021-03599-0

Fornace KM, Topazian HM, Routledge I, Asyraf S, Jelip J, Lindblade KA, Jeffree MS, Ruiz Cuenca P, Bhatt S, Ahmed K, Ghani AC, Drakeley C (2023) **No evidence of sustained nonzoonotic Plasmodium knowlesi transmission in Malaysia from modelling malaria case data** *Nat Commun* **14** https://doi.org/10.1038/s41467-023-38476-8

Fungfuang W, Udom C, Tongthainan D, Kadir KA, Singh B (2020) **Malaria parasites in macaques in Thailand: Stump-tailed macaques (Macaca arctoides) are new natural hosts for Plasmodium knowlesi, Plasmodium inui, Plasmodium coatneyi and Plasmodium fieldi** *Malar J* **19**:1–7 https://doi.org/10.1186/s12936-020-03424-0

Gamalo LE, Dimalibot J, Kadir KA, Singh B, Paller VG (2019) **Plasmodium knowlesi and other malaria parasites in long-tailed macaques from the Philippines** *Malar J* **18**:1–7 https://doi.org/10.1186/s12936-019-2780-4

Ghinai I, Cook J, Hla TTW, Htet HMT, Hall T, Lubis IND, Ghinai R, Hesketh T, Naung Y, Lwin MM, Latt TS, Heymann DL, Sutherland CJ, Drakeley C, Field N (2017) **Malaria epidemiology in central Myanmar: identification of a multi-species asymptomatic reservoir of infection** *Malar J* **16** https://doi.org/10.1186/s12936-016-1651-5

Gillespie TR, Chapman CA (2008) **Forest fragmentation, the decline of an endangered primate, and changes in host-parasite interactions relative to an unfragmented forest** *Am J Primatol* **70**:222–230 https://doi.org/10.1002/ajp.20475

Programme Global Malaria **Global Malaria Programme. n.d. https://www.who.int/teams/global-malaria-programme/elimination**

Grigg MJ, Barber BE, Anstey NM, Cox J, Fornace KM, Drakeley CJ, William T, Jelip J, Brock PM, von Seidlein L, Yeo TW, Fornace KM, Brock PM, von Seidlein L, Barber BE, Anstey NM, Yeo TW, Drakeley CJ, Cox J, Fornace KM, Drakeley CJ, William T, Jelip J, Brock PM, von Seidlein L, Yeo TW, Fornace KM, Brock PM, von Seidlein L, Barber BE, Anstey NM, Yeo TW, Drakeley CJ. (2017) **Individual-level factors associated with the risk of acquiring human Plasmodium knowlesi malaria in Malaysia: a case-control study** *Lancet Planet Health* **1**:e97–e104 https://doi.org/10.1016/S2542-5196%2817%2930031-1

Hansen MC, Potapov P v., Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D (2013) **High-Resolution Global Maps of 21st-Century Forest Cover Change** *Science* **342**:850–53

Hanski I, Ovaskainen O (2000) **The metapopulation capacity of a fragmented landscape** *Nature* **404**:755–758 https://doi.org/10.1038/35008063

Hawkes FM, Manin BO, Cooper A, Daim S R H, Jelip J, Husin T, Chua TH, Cooper A, Jelip J, Husin Benny O., Homathevi R, Jelip J, Husin T, Chua TH. (2019) **Vector compositions change across forested to deforested ecotones in emerging areas of zoonotic malaria transmission in Malaysia** *Sci Rep* **9** https://doi.org/10.1038/s41598-019-49842-2

Hippel PT von (2015) **The heterogeneity statistic I2 can be biased in small meta-analyses** *BMC Med Res Methodol* **15** https://doi.org/10.1186/S12874-015-0024-Z

Ho GC, Lee CL, Sharma RSK (2010) **Association of Institutions for Tropical Veterinary Medicine (AITVM) Conference 23-26 August 2010 Bangkok** *Thailand*

Ilacqua RC, Medeiros-Sousa AR, Ramos DG, Obara MT, Ceretti-Junior W, Mucci LF, Marrelli MT, Laporta GZ (2021) **Reemergence of Yellow Fever in Brazil: The Role of Distinct Landscape Fragmentation Thresholds** *J Environ Public Health* **2021** https://doi.org/10.1155/2021/8230789

Imai N, Furukawa T, Tsujino R, Kitamura S, Yumoto T (2018) **Factors affecting forest area change in southeast Asia during 1980-2010** *PLoS One* **13**:1–14 https://doi.org/10.1371/journal.pone.0197391

Jarvis A., Reuter H.I., Nelson EG A. (2008) **Jarvis, A., H.I. Reuter, A. Nelson EG. 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).**

Jeslyn WPS, Huat TC, Vernon L, Irene LMZ, Sung LK, Jarrod LP, Singh B, Ching NL, Huat TC, Singh B, Vernon L, Jarrod LP, Irene LMZ, Sung LK, Jarrod LP, Singh B, Ching NL, Huat TC, Singh B, Vernon L, Jarrod LP (2011) **Molecular epidemiological investigation of Plasmodium knowlesi in humans and macaques in Singapore** *Vector-Borne and Zoonotic Diseases* **11**:131–135 https://doi.org/10.1089/vbz.2010.0024

Jeyaprakasam NK, Liew JWK, Low VL, Wan-Sulaiman WY, Vythilingam I (2020) **Plasmodium knowlesi infecting humans in southeast asia: What's next?** *PLoS Negl Trop Dis* **14**:1–16 https://doi.org/10.1371/journal.pntd.0008900

Kalluri S, Gilruth P, Rogers D, Szczur M (2007) **Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: A review** *PLoS Pathog* **3**:1361–1371 https://doi.org/10.1371/journal.ppat.0030116

Koh GJN, Ismail PK, Koh D (2019) **Occupationally Acquired Plasmodium knowlesi Malaria in Brunei Darussalam** *Saf Health Work* **10**:122–124 https://doi.org/10.1016/j.shaw.2018.09.002

Lee K-SS, Divis PCSS, Zakaria SK, Julin RA, Singh B, Matusop A, Conway DJ, Cox-Singh J, Julin RA, Conway DJ, Cox-Singh J, Singh B, Matusop A, Conway DJ, Cox-Singh J (2011) **Plasmodium knowlesi: Reservoir hosts and tracking the emergence in humans and macaques** *PLoS Pathog* **7** https://doi.org/10.1371/journal.ppat.1002015

Lehman SM, Rajaonson A, Day S (2006) **Edge effects and their influence on lemur density and distribution in southeast Madagascar** *Am J Phys Anthropol* **129**:232–241 https://doi.org/10.1002/ajpa.20241

Li J, Jia K, Liu Y, Yuan B, Xia M, Zhao W (2021) **Spatiotemporal distribution of zika virus and its spatially heterogeneous relationship with the environment** *Int J Environ Res Public Health* **18**:1–14 https://doi.org/10.3390/ijerph18010290

Li MI, Mailepessov D, Vythilingam I, Lee V, Lam P, Ng LC, Tan CH (2021) **Prevalence of simian malaria parasites in macaques of Singapore** *PLoS Negl Trop Dis* **15**:1–12 https://doi.org/10.1371/journal.pntd.0009110

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) **The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration** *PLoS Med* **6** https://doi.org/10.1371/JOURNAL.PMED.1000100

Lin L, Chu H (2018) **Quantifying Publication Bias in Meta-Analysis** *Biometrics* **74** https://doi.org/10.1111/BIOM.12817

Link A, de Luna A, Alfonso F, Giraldo-Beltran P, Ramirez F. (2010) **Initial effects of fragmentation on the density of three neotropical primate species in two lowland forests of Colombia** *Endanger Species Res* **13**:41–50 https://doi.org/10.3354/esr00312

Loh EH, Murray KA, Nava A, Daszak P (2016) **Evaluating the Links between Biodiversity** *Land-Use Change, and Infectious Disease Emergence. Tropical Conservation* :79–88

Maeno Y, Quang NT, Marchand RP, Culleton R, Kawai S, Masuda G, Nakazawa S (2015) **Humans frequently exposed to a range of non-human primate malaria parasite species through the bites of Anopheles dirus mosquitoes in South-central Vietnam** *Parasit Vectors* **8** https://doi.org/10.1186/s13071-015-0995-y

Manin BO, Ferguson HM, Vythilingam I, Fornace K, William T, Torr SJ, Drakeley C, Chua TH, Ferguson HM, Vythilingam I, Fornace K, Drakeley C, William T, Torr SJ, Drakeley C, Chua TH, Ferguson HM, Vythilingam I, Fornace K, Drakeley C, William T, Torr SJ, Drakeley C, Chua TH (2016) **Investigating the Contribution of Peri-domestic Transmission to Risk of Zoonotic Malaria Infection in Humans** *PLoS Negl Trop Dis* **10**:1–14 https://doi.org/10.1371/journal.pntd.0005064

McFarlane R, Sleigh A, McMichael T (2012) **Synanthropy of wild mammals as a determinant of emerging infectious diseases in the Asian-Australasian region** *Ecohealth* **9**:24–35 https://doi.org/10.1007/s10393-012-0763-9

McGarigal K, Cushman SA. (2012) **FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps** *Computer software program produced by the authors at the University of Massachusetts, Amherst*

Moyes CL, Shearer FM, Huang Z, Wiebe A, Gibson HS, Nijman V, Mohd-Azlan J, Brodie JF, Malaivijitnond S, Linkie M, Samejima H, O'Brien TG, Trainor CR, Hamada Y, Giordano AJ, Kinnaird MF, Elyazar IRFF, Sinka ME, Vythilingam I, Bangs MJ, Pigott DM, Weiss DJ, Golding N, Hay SI, Huang Z, Gibson HS, Sinka ME, Weiss DJ, Nijman V, Mohd-Azlan J, Brodie JF, Malaivijitnond S, Linkie M, Samejima H, O'Brien TG, Kinnaird MF, Trainor CR, Hamada Y, Giordano AJ, Elyazar IRFF, Vythilingam I, Bangs MJ, Pigott DM, Hay SI (2016) **Predicting the geographical distributions of the macaque hosts and mosquito vectors of Plasmodium knowlesi malaria in forested and non-forested areas** *Parasit Vectors* **9** https://doi.org/10.1186/s13071-016-1527-0

Muehlenbein MP, Pacheco MA, Taylor JE, Prall SP, Ambu L, Nathan S, Alsisto S, Ramirez D, Escalante AA, Pacheco MA, Taylor JE, Ambu L, Nathan S, Alsisto S, Ramirez D, Escalante AA (2015) **Accelerated diversification of nonhuman primate malarias in Southeast Asia: Adaptive radiation or geographic speciation?** *Mol Biol Evol* **32**:422–439 https://doi.org/10.1093/molbev/msu310

Munn Z, MClinSc SM, Lisy K, Riitano D, Tufanaru C. (2015) **Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data** *Int J Evid Based Healthc* **13**:147–153 https://doi.org/10.1097/XEB.0000000000000054

Murray KA, Daszak P (2013) **Human ecology in pathogenic landscapes: two hypotheses on how land use change drives viral emergence** *Curr Opin Virol* **3**:79–83 https://doi.org/10.1016/j.coviro.2013.01.006

Nunn CL, Dokey ATW (2006) **Ranging patterns and parasitism in primates** *Biol Lett* **2**:351–354 https://doi.org/10.1098/rsbl.2006.0485

Parker DM, Carrara VI, Pukrittayakamee S, McGready R, Nosten FH (2015) **Malaria ecology along the Thailand-Myanmar border** *Malar J* **14** https://doi.org/10.1186/s12936-015-0921-y

Phang WK, Hamid MHA, Jelip J, Mudin RN, Chuang T-W, Lau YL, Fong MY (2020) **Spatial and Temporal Analysis of Plasmodium knowlesi Infection in Peninsular Malaysia, 2011 to 2018** *Int J Environ Res Public Health* **17** https://doi.org/10.3390/ijerph17249271

Putaporntip C, Jongwutiwes S, Thongaree S, Seethamchai S, Grynberg P, Hughes AL (2010) **Ecology of malaria parasites infecting Southeast Asian macaques: evidence from cytochrome b sequences Chaturong** *Mol Ecol* **19**:3466–3476 https://doi.org/10.1111/j.1365-294X.2010.04756.x.Ecology

Rogerson PA (2001) **Statistical Methods for Geography**

Huddin A Saleh, Yusuf NA Md, Razak MRMA, Ogu Salim N, Hisam S (2019) **Genetic diversity of Plasmodium knowlesi among human and long-tailed macaque populations in Peninsular Malaysia: The utility of microsatellite markers** *Infection, Genetics and Evolution* **75** https://doi.org/10.1016/j.meegid.2019.103952

Seethamchai S, Putaporntip C, Malaivijitnond S, Cui L, Jongwutiwes S, Malaivijitnond S, Cui L (2008) **Malaria and Hepatocystis species in wild macaques, southern Thailand** *American Journal of Tropical Medicine & Hygiene* **78**:646–653 https://doi.org/10.4269/ajtmh.2008.78.646

Setiadi W, Sudoyo H, Trimarsanto H, Wangsamuda S, Asih PBS, Syafruddin D, Sihite BA, Saragih RJ, Juliawaty R (2016) **A zoonotic human infection with simian malaria, Plasmodium knowlesi, in Central Kalimantan, Indonesia** *Malar J* **15** https://doi.org/10.1186/s12936-016-1272-z

Shearer FM, Huang Z, Weiss DJ, Wiebe A, Gibson HS, Battle KE, Pigott DM, Brady OJ, Putaporntip C, Jongwutiwes S, Lau YL, Manske M, Amato R, Elyazar IRF, Vythilingam I, Bhatt S, Gething PW, Singh B, Golding N, Hay SI, Moyes CL (2016) **Estimating Geographical Variation in the Risk of Zoonotic Plasmodium knowlesi Infection in Countries Eliminating Malaria** *PLoS Negl Trop Dis* **10**:1–17 https://doi.org/10.1371/journal.pntd.0004915

Simon JA, Marrotte RR, Desrosiers N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe FJ, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V (2014) **Climate change and habitat fragmentation drive the occurrence of Borrelia burgdorferi, the agent of Lyme disease, at the northeastern limit of its distribution** *Evol Appl* **7**:750–764 https://doi.org/10.1111/eva.12165

Singh B, Sung LK, Matusop A, Radhakrishnan A, Shamsul SSG, Cox-Singh J, Thomas A, Conway DJ (2004) **A large focus of naturally acquired Plasmodium knowlesi infections in human beings** *The Lancet* **363**:1017–1024 https://doi.org/10.1016/S0140-6736(04)15836-4

Stark DJ, Fornace KM, Brock PM, Abidin TR, Gilhooly L, Jalius C, Goossens B, Drakeley CJ, Salgado-Lynn M (2019) **Long-Tailed Macaque Response to Deforestation in a Plasmodium knowlesi-Endemic Area** *Ecohealth* **16**:638–646 https://doi.org/10.1007/s10393-019-01403-9

Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA (2012) **Statistical Analysis of Individual Participant Data Meta-Analyses: A Comparison of Methods and Recommendations for Practice** *PLoS One* **7** https://doi.org/10.1371/JOURNAL.PONE.0046042

Unpublished. (2015) **Unpublished. 2015. Personal Correspondance.** *Personal Correspondance*

Unpublished. (2013) **Unpublished. 2013. Personal Correspondance.** *Personal Correspondance*

Vaz VC D'Andrea, Jansen AM PS (2007) **Effects of habitat fragmentation on wild mammal infection by Trypanosoma cruzi** *Parasitology* **134**:1785–93

Vythilingam I, Noorazian YM, Huat TC, Jiram AI, Yusri YM, Azahari AH, Norparina I, Noorrain A, Lokmanhakim S, Huat TC, Azahari AH, Noorrain A, Lokmanhakim S, Jiram AI, Yusri YM, Azahari AH, Norparina I, Noorrain A, Lokmanhakim S, Huat TC, Azahari AH, Noorrain A, Lokmanhakim S (2008) **Plasmodium knowlesi in humans, macaques and mosquitoes in peninsular Malaysia** *Parasit Vectors* **1** https://doi.org/10.1186/1756-3305-1-26

Vythilingam I, Tan CH, Asmad M, Chan ST, Lee KS, Singh B (2006) **Natural transmission of Plasmodium knowlesi to humans by Anopheles latens in Sarawak, Malaysia** *Trans R Soc Trop Med Hyg* **100**:1087–1088 https://doi.org/10.1016/j.trstmh.2006.02.006

Wang X, Blanchet FG, Koper N (2014) **Wang X, Blanchet FG, Koper N. 2014. Measuring habitat fragmentation : An evaluation of landscape pattern metrics 634–646. doi:10.1111/2041-210X.12198** *Measuring habitat fragmentation : An evaluation of landscape pattern metrics* :634–646 https://doi.org/10.1111/2041-210X.12198

Waxman D, Weinert LA, Welch JJ (2014) **Inferring host range dynamics from comparative data: the protozoan parasites of new world monkeys** *Am Nat* **184**:65–74 https://doi.org/10.1086/676589

White LA, Forester JD, Craft ME (2018) **Disease outbreak thresholds emerge from interactions between movement behavior, landscape structure, and epidemiology** *Proc Natl Acad Sci U S A* **115**:7374–7379 https://doi.org/10.1073/pnas.1801383115

Wilkinson DA, Marshall JC, French NP, Hayman DTS (2018) **Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence** *J R Soc Interface* **15** https://doi.org/10.1098/rsif.2018.0403

Wong ML, Chua TH, Leong CS, Khaw LT, Fornace K, Wan-Sulaiman W-Y, William T, Drakeley C, Ferguson HM, Vythilingam I, Chua TH, Fornace K, Drakeley C, William T, Ferguson HM (2015) **Seasonal and Spatial Dynamics of the Primary Vector of Plasmodium knowlesi within a Major Transmission Focus in Sabah, Malaysia** *PLoS Negl Trop Dis* **9** https://doi.org/10.1371/journal.pntd.0004135

WorldPop (2018) **WorldPop. 2018. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). 10.5258/SOTON/WP00671** *Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076)* https://doi.org/10.5258/SOTON/WP00671

Zarith, et al. (2021) **Molecular Prevalence, Spatial Distribution and Epidemiological Risk Factors of Plasmodium knowlesi infection among Macaca fascicularis in Peninsular Malaysia (pre-publication)**

Zhang X, Kadir KA, Quintanilla-Zariñan LF, Villano J, Houghton P, Du H, Singh B, Smith DG. **Distribution and prevalence of malaria parasites among long-tailed macaques (Macaca fascicularis) in regional populations across Southeast Asia** *Malar J* **15**:1–8 https://doi.org/10.1186/s12936-016-1494-0

Gabriel Zunino, Kowalewskii M, Oklander L. (2007) **Habitat Fragmentation and Population Size of the Black and Gold Howler Monkey (Alouatta caraya) in a Semideciduous Forest in Northern Argentina** *Am J Primatol* **69**:966–975 https://doi.org/10.1002/ajp

Article and author information

Emilia Johnson
School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland, Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK, Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK
For correspondence: emilia.johnson@glasgow.ac.uk
ORCID iD: 0000-0001-5460-1715

Reuben Sunil Kumar Sharma
Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

Pablo Ruiz Cuenca
Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK, Lancaster University, Bailrigg, Lancaster LA1 4YW, Liverpool School of Tropical Medicine, Pembroke Place Liverpool, L3 5QA, UK
ORCID iD: 0000-0002-2180-9509

Isabel Byrne
Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK
ORCID iD: 0000-0002-7800-3733

Milena Salgado-Lynn
School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK, Wildlife Health, Genetic and Forensic Laboratory, Sabah Wildlife Department, Wisma Muis, 88100, Kota Kinabalu, Malaysia, Danau Girang Field Centre, Sabah Wildlife Department, Wisma MUIS 88100, Kota Kinabalu Sabah, Malaysia

Zarith Suraya Shahar
Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

Lee Col Lin
Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

Norhadila Zulkifli
Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

Nor Dilaila Mohd Saidi
Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

Chris Drakeley
Department of Infection Biology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK
ORCID iD: 0000-0003-4863-075X

Jason Matthiopoulos
School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland
ORCID iD: 0000-0003-3639-8172

Luca Nelli
School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland
ORCID iD: 0000-0001-6091-4072

Kimberly Fornace
School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland, Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK, Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK, Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549
ORCID iD: 0000-0002-5484-241X

Copyright

Editors

Joint Public Review:

The study as a concept is well designed, although there is still one issue I see in the methodology.

I still have concerns with their attempts to combine the different scales of data. While the use of point data is great, it limits the sample size, and they have included the district to country level data to try and increase the sample size. The problem is that although they try to get an overall estimate at the district/state/country by taking 10 random sample points, which could be a method to get an estimate for the district/state/country. It would be a suitable method if the primates were evenly distributed across the district/state/country. The reality is that the primates are not evenly distributed across the district/state/country therefore the random point sampling is not a reasonable method to get an estimate of the environmental variables in relation to the macaques. For example if you had a mountainous country and you took 10 random points to estimate altitude, you would end up with a large number, but if all the animals of interest lived on the coast, your average altitude is meaningless in relation to the animals of interest as they are all living at low altitude. The fact that the model relies less on highly variable components and places more reliance on less variable components, is really not relevant as the district/state/country measurements have no real meaning in relation to the distribution of masques.

A simple possible way forward could be to run the model without the district/state/country samples and see what the outcome is. If the outcome is similar then the random point method may be viable (but if it gives the same outcome as ignoring those samples then you don't need the district/state/country samples). If you get a totally different outcome then it should raise concerns about using the district/state/country samples.

This paper is a really nice piece of work and is a valuable contribution but the district/state/country sample issue really needs to be addressed.

Author Response

The following is the authors' response to the original reviews.

> ### *Reviewer #1 (Public Review):*
>
> *The study as a concept is well designed, although there are two issues I see in the methodology (these may be just needing further explanation or if I am correct in my interpretation of what was done, may need reanalysis to take into account). Both issues relate to the data that was extracted from the published literature on zoonotic malaria prevalence in the study area.*

> *1. No limit was set on the temporal range*
>
> *With no temporal limit on the range of studies, the landscape in many cases will have changes between the study being conducted and the spatial data. This will be particularly marked in areas where there has been clearing since the zoonotic malaria prevalence study. Also, population changes (either through population growth, decline or movement) will have occurred. All research is limited in what it can do with the available data, so I realise that there may not be much the authors can do to correct this. One possible solution would be to look at the land use change at each site between the prevalence study and the remote sensing data. I'm not sure if this is feasible, but if it is I would recommend the authors attempt this as it will make their results stronger.*

Thank you for the comments. We agree that matching the date of remote sensing data to samples is particularly important for environmental variables that change rapidly (such as forest loss). To clarify, no limit was set on the date range of the studies identified from the literature to ensure no articles were excluded due to arbitrary date restrictions. We have edited the manuscript to clarify this (line 422). Regarding landscape and environmental features, remote sensing data was extracted annually for every year for the full date range of the data (see Table 1 and S11, annual temporal resolution from 2006 to 2020). Forest was then matched contemporaneously (see lines 467–473) meaning that, insofar as it was possible, forest data was extracted for the same year as the data was collected. Where a date range was given for the primate data, the mean year was used. For human population density, covariate data were extracted for multiple years but were found to be relatively stable over the time period for the sites covered, so median year was used (see Supplementary Information, Appendix E and Table S11). Elevation is stable and typically only one time point is used as reference (in this instance the SRTM 90m Digital Elevation model, 2003).

> *1. Most studies only gave a geographic area or descriptive location.*
>
> *The spatial analysis was based on a 5km and 20km radius of the 'study site' location, but for many of the studies the exact site is not known. Therefore the 'study site' was artificially generated using a polygon centroid. Considering that the polygon could be an administrative boundary (i.e., district/state/country), this is an extremely large area for which a 5km radius circle in the middle of the polygon is being taken as representative of the 'study site'. This doesn't make sense as it assumes that the landscape is uniform across the district, which in most cases it will not be (in rural areas it is going to be a mixture of villages, forest, plantation, crops etc which will vary across the landscape). This might just be a case of misunderstanding what was done (in which case the text needs rewording to make it clearer) or if I have interpreted it correctly the selection of the centroid to represent the study area does not make sense. I am not sure how to overcome this as it probably not possible to get exact locations for the study sites. One possibility could be to make the remote sensing data the same scale as the prevalence data ie if the study site is only identifiable at the polygon level, then the remote sensing data (fragmentation, cover and population) is used at the polygon level.*
>
> *Both these issues could have an impact on the study's findings. I would think that in both cases it might make the relationship between the environmental variables and prevalence even clearer.*

We would like to thank the reviewer for their concerns and provide some clarification on the methods used to extract environmental variables:

• Centroid was initially explored, but not pursued for the same concerns raised by the reviewer. Taking the centroid would be arbitrary and the central point of a large polygon is

not likely to be representative of habitat across the entire sampling area and introduces error so this was not pursued(Cheng et al., 2021). We have clarified the wording in the manuscript with reference to centroids to avoid confusion on this point (line 491).

• We demonstrate a method to account for the lack of precise geolocation by taking 10 'pseudo-sampling' points instead of a single random location, with environmental variables extracted at 5, 10 and 20km for each site (lines 487-500). By including 10 environmental realisations, surveys conducted in smaller or more uniform landscapes will have more consistent covariates and this will lend more weight to the model. Conversely, samples taken from large administrative polygons are likely to be highly variable, and these associations will have less representation in the final model. This approach was used to demonstrate an alternative to using a single arbitrary site to represent the area.

To further support the validity of this technique:

• Figures illustrating the variance of the environmental variables across the 10 sampling sites at 5, 10 and 15km for GADM administrative classifications at country level (GID0), state (GID1), district (GID2) and exact coordinates (GPS) are now included in the SI (Figure S12).

• Sensitivity analyses were conducted, in which final GLMM models were fit again but using only acceptable levels of variance in environmental variables and/or acceptable size of administrative boundary (Table S15 and S16). In sensitivity analyses, forest cover and fragmentation retained a significant effect on prevalence of P. knowlesi in macaques, suggesting this effect is robust to spatial uncertainty.

We would also like to highlight that the main finding of this research is the novel synthesis of regional prevalence of P. knowlesi in simian reservoirs across Southeast Asia, which was formerly assumed to be ubiquitous high prevalence, and which can now be used to inform regionally specific transmission modelling, better estimate spatial risk and parameterise early warning systems for P. knowlesi malaria in countries approaching elimination of human malarias. The risk factor analysis here is provided to begin to understand what may be driving this geographic heterogeneity in P. knowlesi prevalence at finer scales and demonstrate methods that could be used to accommodate spatial uncertainty in secondary data. We appreciate that this may not have been clear and have edited the manuscript accordingly.

Reviewer #2 (Public Review):

This is the first comprehensive study aimed at assessing the impact of landscape modification on the prevalence of P. knowlesi malaria in non-human primates in Southeast Asia. This is a very important and timely topic both in terms of developing a better understanding of zoonotic disease spillover and the impact of human modification of landscape on disease prevalence.

This study uses the meta-analysis approach to incorporate the existing data sources into a new and completely independent study that answers novel research questions linked to geospatial data analysis. The challenge, however, is that neither the sampling design of previous studies nor their geospatial accuracy are intended for spatially-explicit assessments of landscape impact. On the one hand, the data collection scheme in existing studies was intentionally opportunistic and does not represent a full range of landscape conditions that would allow for inferring the linkages between landscape parameters and P. knowlesi prevalence in NHP across the region as a whole. On the other hand, the absolute majority of existing studies did not have locational precision in reporting results and thus sweeping assumptions about the landscape representation had to be made for the modeling experiment. Finally, the landscape characterization was oversimplified in this study, making it difficult to extract meaningful relationships

between the NHP/human intersection on the landscape and the consequences for P. knowlesi malaria transmission and prevalence.

Thank you for the feedback on the manuscript. We agree that the data was not originally intended for spatial assessment of landscape impact nor represents a full range of landscape conditions across the region. However, we would like to highlight the first set of results from the meta-analysis. Here, the synthesis of all available data allows for the detection of regional disparities and geographic heterogeneity of prevalence in host species, which individual small-scale opportunistic studies are not powered to do, and which had not been identified before this investigation.

In this context, the risk factor analysis is an exploratory analysis to understand what may be driving the observed geographic variation at broad scales as well as provide a framework for dealing with spatial uncertainty. Landscape data was extracted at a level deemed appropriate given the limitations of the data. The majority were geolocated to district level and sensitivity analysis showed a reasonable consistency of landscape features at our chosen scales (Table S8, Figure S12A). To address some of these concerns, we conducted further analysis to explore the deviation of environmental covariates in each sampling area and ran sensitivity analysis by removing extremely variable datapoints (Table S15 and Table S16). When removing highly uncertain data and/or countrylevel data, effects of canopy cover on non-human primate malaria prevalence is retained, supporting the original findings.

Despite many study limitations, the authors point to the critical importance of understanding vector dynamics in fragmented forested landscapes as the likely primary driver in enhanced malaria transmission. This is an important conclusion particularly when taken together with the emerging evidence of substantially different mosquito biting behaviors than previously reported across various geographic regions.

Another important component of this study is its recognition and focus on the value of geospatial analysis and the availability of geospatial data for understanding complex human/environment interactions to enable monitoring and forecasting potential for zoonotic disease spillover into human populations. More multi-disciplinary focus on disease modeling is of crucial importance for current and future goals of eliminating existing and preventing novel disease outbreaks.

Reviewer #1 (Recommendations For The Authors):

A couple of minor points

1. Was the human density and forest cover correlated? If so was this taken into account

Human density and forest cover at selected scales were not found to be strongly correlated (Spearman's rank values -0.38 and -0.45 within 5km and 20km buffer radii for human population density respectively).

In selecting variables for inclusion in the final model, we examined variance inflation factors (VIF) to detect and minimise multicollinearity in the model. VIF measures the correlation and strength of correlation between independent predictors. VIF of each predictor variable was examined starting with a saturated model and sequentially excluding the variable with the highest VIF score from the model. Stepwise selection continued until the entire subset of explanatory variables in the global model satisfied a conservative threshold of VIF ≤6 (Rogerson, 2001), which ensures that the remaining variables included in the final model have minimal correlation. Spearman's correlation matrices for all variables at all scales and

final selected variables (below VIF threshold) are included in the Supplementary Information (Figure S13 and Figure S14).

> *1. Reference (Speldewinde et al., 2019) is down as Davidson et al. in the reference list*

Thank you for the thoroughness in this review. There are two similar but separate references, both published in 2019 with the same co-authors, and the (Speldewinde et al, 2019) was incorrectly referenced. They should be (Davidson et al., 2019a) and Davidson et al., 2019b) respectively. This has now been corrected in the manuscript.

Davidson, G., Chua, T.H., Cook, A. et al. Defining the ecological and evolutionary drivers of Plasmodium knowlesi transmission within a multi-scale framework. Malar J 18, 66 (2019). https://doi.org/10.1186/s12936-019-2693-2

Davidson G, Chua TH, Cook A, Speldewinde P, Weinstein P. The Role of Ecological Linkage Mechanisms in Plasmodium knowlesi Transmission and Spread. Ecohealth. 2019;16(4):594-610. https://doi:10.1007/s10393-019-01395-6

> ## *Reviewer #2 (Recommendations For The Authors):*
>
> *Line 143: "We hypothesise that higher prevalence of P. knowlesi in primate host species is driven by landscape change..." without specifying here the kind of landscape change (e.g. "forest degradation and fragmentation") it is virtually impossible to confirm or reject this hypothesis.*

We agree that the wording of the hypotheses needed to be more specific. We have edited lines 142 – 145 to specify forest fragmentation as our landscape variable of interest, and to more explicitly include the regional meta-analysis of P. knowlesi prevalence.

> *Table 1 vs Table S11 discrepancy regarding spatial resolution of Forest cover and fragmentation variables. The original dataset resolution is 30m but I don't think one can compute a PARA index at 30 m since it really requires a polygon that is larger than the single value pixel. Table S11 indicates a 30 km gridcell with some postprocessing of the original datasets.*

We appreciate this being identified. The resolution refers to the input layer (tree canopy cover, 30m). PARA was calculated from the binary forest cover layer (30m resolution) within each buffer radii 5, 10 and 20km. We have edited both Table 1 and Table S11 to help clarify this.

> *It would be very helpful if you provided justification for selecting specific metrics to represent the key landscape variables. How are these particular landscape variables relevant? Why not other land cover/land use components?*

We have now included a paragraph in the Supplementary Information (Appendix D) to explain the choice of environmental covariates. Elevation was chosen as an important proxy for vector distribution (but was not retained in model selection). Human population density was chosen as a measure of proximity to human settlement, rather than relying on qualitative assessment of rural/peri-urban/urban. Tree canopy cover and fragmentation indices are key determinants of primate habitat selection and of vector breeding habitat, and justification for the use of perimeter: area ratio is included in the methods section (section beginning line 462).

> *I think the other issues present substantial weaknesses that you cannot address without redoing the study. I will list those below just for reference.*
>
> 1. *If the forest is so dominant (which I would agree with based on my understanding of macaque ecology), how does it make sense to select completely random points (especially at the country or even state level) to represent landscape covariates? At a minimum, I would suggest getting random points within the forest or better yet forest edge habitat. But even then, I doubt that these points would be at all representative of the conditions of a specific study. The geospatial uncertainty is just too large. The dataset simply doesn't support the analysis that is attempted here.*

On the point of selecting from only within forest: forest is a dominant habitat, but Long-tailed macaques are anthropophilic and not exclusively found in forest (Stark et al., 2019), and a proportion of the more opportunistic and nuisance samples caught were found in areas more associated with human activity (Li et al., 2021). As such, random points only within forested areas is also unlikely to capture the true habitat of the primates sampled and selecting only from forested areas would bias the results.

Whilst fully georeferenced samples would be the ideal scenario, the idea behind selecting random points from the sampling polygon is that for smaller areas (with higher spatial certainty), habitat would be more consistent between random points and lend more weight to the final model, whereas large polygons with high uncertainty are likely to vary and lend less weight to the final model. In response to these comments, we have further supported this by running regression models only on samples within a reasonable administrative boundary size and on samples within reasonable threshold of uncertainty (i.e., data points are removed if the deviation of environmental covariates across the 10 random points is so high that the sample is uninformative, or if datapoints can only be geolocated to country-level). In these sensitivity analyses, forest cover and species are retained as factors associated with higher malarial prevalence in non-human primates (Table S15S16).

> 1. *Hansen et al. dataset reflects "tree cover" - which is not the same as "forest cover" since it would also include plantations that are very widely distributed across Southeast Asia. If the animal use of plantations differs from that of natural forests, it will present a large issue for the study.*

In this analysis the feature of interest was habitat configuration (fragmentation) and deforestation (forest loss) rather than specific land class. We have defined forest as >50% canopy cover, which considers canopy density given historical forest loss and has precedence in other work (Fornace et al.,, 2016). In addition to importance to macaque ecology, forest (canopy) cover, forest loss and forest edge are noted to be key determinants of vector breeding and vector habitat (Byrne et al., 2021, Chua et al., 2019). For this reason, these are important variables to include in analyses. More specific landscape variables were explored, but the temporal and spatial range of the data precluded fine-scale land classification data. To investigate preliminary links to landscape configuration and habitat fragmentation at broad scales this is felt to be sufficient. We have also amended the manuscript to be more discerning with the use of ‘forest’ to avoid confusion throughout.

> 1. *Tree regrowth in the ecosystems of monsoonal Asia is very rapid. Based on the study description, tree regrowth was not accounted for in the study which could potentially lead to a very large underestimation of tree cover if only tree loss since 2000 was monitored. Again unless there is a reason to assume that macaques do not use young successional forests or use it at a highly reduced rate. Both of these points are acknowledged as limitations at the end of the discussion section but in my opinion they have a very strong impact on the study, making the results non-significant.*

This is an interesting suggestion. Macaques do forage in plantations and cultivated landscapes to supplement food, but preferentially roost and range in forest edges and interior forest, though ranging behaviour will be complex and vary across Southeast Asia. In this study the primary interest was in deforestation (forest loss) and fragmentation of old growth forested landscapes, which are key variables both for macaque ecology and for vector breeding sites. Therefore, it was felt that forest loss (transition from >50% canopy cover to <50% canopy cover since 2000) was sufficient to capture this. Ranging behaviour of individual animals and macaque troops would not be captured at this scale, and higher spatial and temporal resolution would be required to characterise relationships with tree regrowth and young plantations which is outside the scope of this study. In all regions, purposeful fine scale follow-up studies would be required to unpick fine scale relationships across a habitat gradient.

> *I am not 100% sure I understand the geospatial design fully. The pieces are distributed between different subsections and it was challenging to string together the processing chain between subsections of the manuscript and the supplemental information. I would help to add a figure (a flowchart, perhaps?) to the supplemental section that walks through the entire geospatial covariates assembly. E.g.*
>
> - *GPS location create 5, 10, and 20 km buffers mean elevation, mean population, % (?) Forest, PARA(?) for each buffer - I still don't understand the 30m or 30 km spatial resolution reference for forest and PARA in this context.*

This was an error in the table in the Supplementary Information and has been corrected – the forest cover raster has a resolution of 30m, and the perimeter: area ratio is calculated within 5, 10 and 20km buffers.

> - *landscape covariates receive the full weight (1) in the model. - This is defensible even though not ideal*

This is equivalent, but we felt more intuitive, to sampling GPS points x10 and inputting with equal weights to the areal data.

> - *No GPS location assign to the best identifiable administrative unit (country, state, or district) generate 10 random points within the administrative unit create 5, 10, and 20 km buffers mean elevation, mean population, %(?) Forest, PARA(?) for each buffer landscape covariates from each point receive the proportional weight (0.1) in the model. I do not believe that this approach is representative of macaque habitat/macaque human interaction characterization.*

In other examples dealing with spatial uncertainty, the centroid is taken to be representative of an area. This method generates considerable bias and uncertainty – particularly if the uncertainty is not then accounted for by weighting subsequent models (Cheng, 2021). In this exploratory analysis, pseudo-sampling from 10 random sites generates a more realistic

generalised environmental realisation than taking a centroid/random point. This was used as an exploratory analysis to explain broad regional trends in prevalence between, which can be used to guide further investigation on fine scale studies which are required to completely describe disease dynamics in specific macaque habitats.

Thank you for this useful suggestion – we have taken this advise and added a flowchart of data processing to the Supplementary Information (Appendix D, Figure S8).

> *Discussion:*
>
> *Based on information in Table S4, sampled NHPs were predominantly from human-dominated (peridomestic, agricultural, and urban) landscapes. In forested landscapes, only macaques that live in forest edge habitats were likely sampled in the first place just simply due to extreme challenges in getting to macaques in remote inaccessible areas. There is a very substantial spatial bias in sampling will undoubtedly reflect that fragmented habitat is a key landscape component impacting the prevalence of Pk in NHP, especially as the authors point out in the later part of the discussion, the critical vectors for transmission are also associated with forest edge habitats. High forest fragmentation is also linked to the presence/ increase in migrant human workers (logging or plantation activities) - a population also strongly associated with higher malaria prevalence for a variety of P spp (although I am not aware of studies that are specific to Pk malaria). However, the living conditions for migrant workers have frequently been implicated in higher rates of malaria transmission and thus those could, hypothetically, also contribute to Pk infection rates in NHP. Ultimately, the discussion appears to suggest that the biggest gap in our understanding is within vector ecology and understanding the NHP-vector-human dynamics within local landscape settings. It is an interesting finding. However, my overall conclusion would be that the sampling strategy (both for NHP and geospatial covariates) renders this study as "exploratory" at maximum and that all findings would need to be tested and verified through independent and more rigorously designed studies.*

Thank you to the reviewer for a comprehensive assessment. We would first like to highlight the regional meta-analysis, which was one of the main findings. This is a novel result for P. knowlesi literature; being the first demonstration of regional differences in prevalence that correlate to regional hotspots of human incidence, the force of infection from NHP may drive hotspots of P. knowlesi in human populations.

We include a risk factor analysis that suggests a method for dealing with high spatial uncertainty, and an exploratory analysis that finds landscape complexity may be a contributory factor to broad regional heterogeneity. These associations are robust to sensitivity analysis where data with extreme variability in environmental variables is removed (Table S15-S16).

Habitat descriptions in original studies are qualitative, likely subjective, and whilst there is likely to be an important sampling bias there was also evident differences in prevalence between the NHP sampled in different environments from the available data that we have further characterised. Risk factors for human P. knowlesi do include forest loss (reduction in canopy cover) within 5 years and within 2km, as well as contact with macaques and occupations in plantations (Fornace et al., 2014; Fornace et al., 2016). Reverse spillover from humans to NHP is an interesting suggestion, but outside the scope and scale of the study. Given known links of deforestation (forest loss) with human incidence of P. knowlesi and also with increased vector breeding sites (Byrne et al., 2021), this analysis explores whether deforestation is linked to prevalence in reservoir species thus contributing to the force of infection at broad scales.