# Early-stage loss of ecological integrity drives the risk of zoonotic disease emergence

**Authors:**

**Lara Marcolin** Department of Biology and Biotechnologies "Charles Darwin", Sapienza Università di Roma, Rome, Italy // ORCID: 0009-0009-3324-8161

**Andrea Tonelli** Department of Biology and Biotechnologies "Charles Darwin", Sapienza Università di Roma, Rome, Italy // ORCID: 0009-0009-1435-8381

**Moreno Di Marco** Department of Biology and Biotechnologies "Charles Darwin", Sapienza Università di Roma, Rome, Italy // ORCID: 0000-0002-8902-4193

## Abstract

Anthropogenic pressures have increasingly disrupted integrity of ecosystems worldwide, jeopardizing their capacity to provide essential contributions to human well-being. Recently, the role of natural ecosystems in reducing disease emergence risk has gained prominence in decision-making processes, as scientific evidence indicates that human-driven pressure, such as habitat destruction and deforestation, can trigger the emergence of zoonotic infectious diseases. However, the intricate relationship between biodiversity and emerging infectious diseases (EIDs) remains only partially understood. Here, we updated the most comprehensive zoonotic EID event database with the latest reported events to analyse the relationship between EIDs of wildlife origin (zoonoses) and various facets of ecological integrity. We found EID risk was strongly predicted by structural integrity metrics such as human footprint and ecoregion intactness, in addition to environmental variables such as tropical rainforest density and mammal species richness. EID events were more likely to occur in areas with intermediate levels of compositional and structural integrity, underscoring the risk posed

by human encroachment into pristine, undisturbed lands. Our study highlights the need to identify novel indicators and targets that can effectively address EID risk alongside other pressing global challenges in sustainable development, ultimately informing strategies for preserving both human and environmental health.

## 1. Introduction

Emerging infectious diseases (EIDs) of zoonotic origin pose a major threat to public health and socio-economic stability [1, 2]. Most EIDs in recent decades derive from pathogens originating in wildlife [3], but the link between biodiversity and EID risk remains only partially explored. As zoonotic spillover cannot occur in the absence of reservoir hosts that maintain the pathogen, wildlife can be considered the necessary but not sufficient condition for the occurrence of zoonotic spillover [4]. Global-scale analyses [3, 5] identified mammal species richness as a key predictor of zoonotic EIDs. The relationship between mammal species richness and EID risk generally shows idiosyncratic trends, where both low and high values of species richness correspond to high risk [5]. Such trends suggest that the biodiversity-risk relationship is linked to context-dependent dynamics of pathogen, host, and human interaction [6], and is likely influenced by anthropogenic drivers that alter these dynamics. A recent study [7] showed that the proportion of zoonotic hosts in wildlife communities increases in human-dominated systems, compared to undisturbed areas. Both richness and abundance of host species in sites under substantial human use (such as agricultural and urban ecosystems) were higher compared to those found in assemblages sited in undisturbed habitats. This finding highlights the role of anthropogenic disturbance in increasing EID risk. Moreover, it emphasises that efforts to anticipate EID risk need to take into consideration ecosystem-level metrics that capture the effects of human activities and related environmental changes on biodiversity. In this perspective, EID risk is expected to respond positively to high levels of human pressures that cause habitat degradation, fragmentation, and biodiversity loss, thereby altering host communities' composition and abundance in a way that facilitates pathogen circulation and their consequent transmission to human hosts [7,8]. However, the risk of EIDs can be expected to decrease beyond a certain level of habitat modification, because highly modified areas support fewer species of wildlife, leading to a reduced diversity of available pathogens.

Habitat degradation caused by land-use changes has a direct influence on human exposure to wildlife pathogens, increasing contact rates at the human-wildlife interface and leading to higher risk of zoonotic spillover [9,10]. Also, habitat degradation causes substantial changes in species assemblages, which may alter disease dynamics within wildlife communities. In multi-host disease systems, higher species diversity may reduce the transmission of a pathogen through a variety of mechanisms collectively referred to as dilution effects [11]. As different species have different levels of host competence for a given pathogen [12] (i.e., the ability to harbour the pathogen and transmit it to new hosts or vectors), the dilution effect hypothesis assumes that in more diverse wildlife communities, the presence of less competent hosts reduces contact rates between highly competent

hosts (known as the "encounter reduction" mechanism), limiting the opportunities for pathogens to spread. Another dilution mechanism is "host regulation", which occurs when predators or competitors of competent hosts are abundant. Alterations that decrease species diversity may therefore increase pathogen transmission in wildlife communities, resulting in a higher risk of zoonotic spillover once dilution effects are suppressed. On the other hand, decline in biodiversity itself doesn't necessarily result in increasing disease risk. Indeed, high levels of biodiversity loss reduce disease risk if competent host species disappear from highly modified environments [11].

Following the same logic, high levels of environmental and biodiversity integrity are supposed to reduce EID risk [13,14]. While anthropogenic pressures alter species diversity and abundance, intact lands support natural ecological and evolutionary processes operating with minimal human disturbance. Intact forests provide essential functions on climate mitigation, conservation of biodiversity and, importantly in this context, human health [14]. When compared to degraded ones, intact forests host a higher richness of forest-dependent species, a higher functional and intraspecific genetic diversity, as well as a higher connectivity important for gene flow and genetic adaptation. All these features are supposed to give intact forests the potential to reduce infectious disease risk. However empirical evidence of how various components of ecological integrity contribute to reducing EID risk remain unknown.

The recently approved Kunming-Montreal Global Biodiversity Framework [15] formalises the importance of conserving ecological integrity, in Target 1: "*... bring the loss of areas of high biodiversity importance, including ecosystems of high ecological integrity, close to zero by 2030…*". The Framework also recognises the value of nature in reducing disease risk, in Target 11 "*Restore, maintain and enhance nature's contributions to people, including ecosystem functions and services, such as […] reduction of disease risk…*". We argue that these two target elements have the potential to generate important synergies in the implementation of the Framework [16] and its relationship with the broader 2030 Sustainable Development Agenda [17], but the relationship between ecological integrity and disease risk reduction needs to be tested empirically. A comprehensive large-scale test of these effects has not been performed yet, as large-scale modelling exercises have mostly focused on species richness and ignored ecological integrity and other aspects of biodiversity [3, 5]. In this work, we explore how global zoonotic EID risk responds to ecological integrity, analysing the relationship between zoonotic EID events of wildlife origin and several biodiversity metrics that reflect anthropogenic alteration.

## 2. Methods

### 2.1. Mapping zoonotic EID events

We followed the zoonotic EID event definition given by Jones et al. [3], considering an EID event as "*the first temporal emergence of a pathogen in a human population which was related to the increase in distribution, increase in incidence or increase in virulence or other factor which led to that pathogen being classed as an emerging disease*". The zoonotic EID event database used in our work is an updated version of the database used in Jones et al. [3] and modified by Allen et al. [5], where events range in time from 1940 to 2008 (n=224). Based on the Emerging Infectious Disease Repository database [18], we filtered and updated 7 out of 33 reported events until 2013 (n=231) (Table S1) following the same criteria used by Allen et al. [5], so as to include zoonotic diseases of wildlife origin reported in the peer-reviewed literature, keeping only the first emergence of a new disease-causing agent. Single case reports and events lacking in evidence for their geographic and temporal origin were excluded from the analysis. For each new event we created a spatial polygon corresponding to the most precise municipal region where the event occurred. Since there are multiple temporal layers among covariates, such as land use and population variables, events were selected keeping only those which occurred since 1970 (n=154) in order to temporally match them with predictors.

### 2.2. Predictors of zoonotic EID risk

Following Allen et al. [5], we first collected spatial data layers for 19 environmental and anthropogenic drivers already tested for their contribution to zoonotic EID risk prediction. Then, we collected six biodiversity metrics representing different aspects of ecological integrity, to test our hypothesis about the role of integrity in mitigating EID risk. These data were downscaled to the lowest common spatial resolution of 1° (WGS84, approximately 110 km at the equator). Full list of predictor layers and details of sources, original resolutions and rescaling are shown in Table 1.

132 **Table 1.** List of predictors included in the models.

| Variable | Type | Source data set | Processing | Temporal resolution and extent |
|---|---|---|---|---|
| Human population | Human activity | GPCGv1 [25] | Rescaled | Decadal (1970-2000) |
| Population change | Human activity | GPCGv1 [25] (calculated) | Calculated from rescaled layers | Inter-decadal (1970-2000) |
| Cropland | Human activity | LUH2 [29] | Rescaled | Decadal (1970-2000) |
| Cropland change | Human activity | LUH2 [29] (calculated) | Calculated from rescaled layers | Inter-decadal (1970-2000) |
| Pasture | Human activity | LUH2 [29] | Rescaled | Decadal (1970-2000) |
| Pasture change | Human activity | LUH2 [29] (calculated) | Calculated from rescaled layers | Inter-decadal (1970-2000) |
| Urban land | Human activity | EarthEnv [20] | Rescaled | N/A |
| Managed/cultivated vegetation | Human activity | EarthEnv [20] | Rescaled | N/A |
| Global environmental stratification | Environment | GEnS [19] | Rescaled | N/A |
| Evergreen/deciduous needleleaf trees | Environment | EarthEnv [20] | Rescaled | N/A |
| Evergreen broadleaf trees | Environment | EarthEnv [20] | Rescaled | N/A |
| Deciduous broadleaf trees | Environment | EarthEnv [20] | Rescaled | N/A |
| Mixed/other trees | Environment | EarthEnv [20] | Rescaled | N/A |
| Shrubs | Environment | EarthEnv [20] | Rescaled | N/A |
| Herbaceous vegetation | Environment | EarthEnv [20] | Rescaled | N/A |
| Regularly flooded vegetation | Environment | EarthEnv [20] | Rescaled | N/A |
| Mammal species richness | Animal/Host | IUCN [32] | (See supplementary methods) | Static (2022) |
| Livestock mammal headcount | Animal/Host | GLW [30] | Rescaled | Static (2016) |
| Poultry | Animal/Host | GLW [30] | Rescaled | Static (2016) |
| Biodiversity habitat index | Biodiversity | Harwood et al. 2022 | Reprojected and rescaled | Decadal (2000-2020) |
| Biodiversity intactness index | Biodiversity | Newbold et al. 2016 | Reprojected and rescaled | Static (2016) |
| Contextual intactness | Biodiversity | Mokany et al. 2020 | Reprojected and rescaled | Static (2013) |
| Ecoregion intactness | Biodiversity | Beyer et al. 2020 | Reprojected and rescaled | Decadal (1993-2009) |
| Human footprint | Biodiversity | Venter et al. 2016 | Reprojected and rescaled | Decadal (1993-2009) |
| Wilderness | Biodiversity | Allan et al. 2017 | (See supplementary methods) | Decadal (1993-2009) |

133

134

### 2.2.1. Environmental and anthropogenic variables

We represented environmental conditions using eight climatic and land cover features, respectively from the Global Environmental Stratification dataset [19] and the EarthEnv database [20]. Such environmental variables are known to greatly influence the distribution of terrestrial mammals and the pathogens they host [21–23].

We then included data on human population density, population change, and land-use changes to control for the human-driven pressure that increases exposure to EIDs. Population density has a strong correlation with the distribution of emerging infectious diseases, affecting disease transmission dynamics and increasing outbreak detection probability in densely populated areas [3,5,23]. Population change serves as an indicator of changing demands on ecosystems, leading to environmental disruptions and disease emergence [13]. These variables were derived from the Global Population Count Grid Time Series Estimates, v1 database [25]. Land-use changes, such as agriculture and deforestation, impact terrestrial species richness and abundance, altering host and vector communities, and facilitating human-wildlife contact [13,26–28]. Land-use data were obtained from the Land Use Harmonization 2 database (LUH2) [29]. Domestic animal density was also included due to the role of livestock as intermediate or amplifying hosts in disease outbreaks in humans [23, 29]. Data on livestock density were retrieved from the Gridded Livestock of the World dataset [30].

### 2.2.2. Biodiversity and Ecological Integrity variables

We included mammal species richness, a well-known correlate of EID risk [5], as a proxy of pathogen species richness. Here, for simplicity, we assumed that the pool of pathogens with zoonotic potential increases with an increased number of mammal species richness [31], although we acknowledge the question on whether zoonotic pathogen richness is homogeneously distributed across mammalian taxonomy is still debated. Terrestrial mammals' distribution data were obtained from the IUCN Red List [32] which includes spatial maps for 5624 terrestrial mammal species' known range. We filtered data excluding ranges where species were declared extinct prior to 1970 and we then assessed mammal species richness index on a global scale by overlaying each species' spatial polygon which overlaps with each 1-degree resolution grid cell and counting the number of species present in each grid cell.

To explore the relationship between ecological integrity and EID risk, we selected six biodiversity metrics, each capturing unique dimensions of ecological integrity. As defined by the Convention on Biological Diversity, ecological integrity is 'an ecosystem's capacity to maintain its composition,

structure, and functioning within a natural range of variability over time'. Structure refers to the three-dimensional component of ecosystems, encompassing the biotic and abiotic elements that shape the heterogeneous matrix which supports the composition and functioning of the ecosystem; composition pertains to the diversity and range of organisms existing within the ecosystem; function relates to the ecological processes and ecosystem services provided by the ecosystem [33]. Changes in the structure of ecosystems due to human activities can have significant consequences for EID risk. Alterations of ecosystem structure can affect the spatial distribution of host species, vectors, and potential intermediate hosts involved in the transmission of diseases. Anthropogenic pressures, such as deforestation and habitat conversion, can lead to changes in ecosystem structure, resulting in increased contact between humans, livestock, and pathogens. This can facilitate the transmission of EIDs [9] by increasing pathogen sharing from wildlife to humans and creating novel opportunities for pathogen evolution and adaptation to new hosts [33,34]. Changes in ecosystems' composition due to human activities, such as deforestation, habitat destruction, and land-use change, can cause shifts in pathogens' ecology by altering zoonotic host communities' composition through loss, turnover and homogenization of biodiversity [7,8]. Biodiversity loss may increase disease transmission and incidence when the lost species are less competent hosts, by intensifying encounter rates between pathogens and competent hosts [26,36]. More diverse host communities instead may inhibit the spread of pathogens according to the dilution effect hypothesis [37].

We chose four ecosystem-level metrics to account for the overall impact of anthropogenic pressures on ecosystem's structure.

(i) The human footprint index (HFP) [38] is an indicator of cumulative human pressures on natural ecosystems. HFP is a global index with a resolution of 1 $km^2$ based on eight key layers of human pressures such as infrastructures, land cover and human access to natural areas, which are a suite of anthropogenic stressors that are known to impact ecological systems.

(ii) Wilderness areas [39,40] defined as "ecologically intact areas free of industrial scale activities and other human pressures which result in significant biophysical disturbance" [41]. Hosting original species assemblages in the absence of large-scale anthropogenic disturbances, fully functional ecosystems may mitigate pathogen spillover risk [14]. We derived the wilderness global index using wilderness areas maps from Allan et al. [40], representing 'pressure free' lands with a contiguous area >10,000 $km^2$. It's important to notice that this definition of wilderness does not preclude human presence (e.g., indigenous populations), rather industrial-scale activities. The proportion covered by wilderness areas for each 1-degree cell was binarized using a threshold of 25% (i.e., cells with >25% wilderness coverage were considered to include wilderness).

(iii) Ecoregion intactness [42] quantifies the extent to which natural ecosystems within an ecoregion remain undisturbed and retain their original ecological characteristics. It is a measure of human alteration of terrestrial ecosystems which captures habitat loss, quality and fragmentation effects arising from anthropogenic disturbance at ecoregional scale.

(iv) Contextual Intactness [43] is obtained by combining local habitat condition with estimates of spatial turnover in species composition, via generalized dissimilarity modelling [44,45] with a resolution of 30 arcsec. The metric identifies valuable local habitat conditions, in the context of other locations where similar species assemblages are found. It measures the proportion of all locations expected to have once supported a similar assemblage of species to the focal grid cell, but which have suffered higher impact from human activities than the focal cell. Contextual intactness values range between 0 and 1, where higher values mean the focal cell has a higher level of intactness within the context of biologically similar cells.

We also accounted for intactness in biological communities, as determined by levels of anthropogenic alterations on species richness and abundance. We chose two metrics:

(i) The biodiversity habitat index [46] is a global index with a resolution of 30 arcsec designed to quantify the level of species diversity expected to be retained within a given geographic area, compared to its original (pristine) condition. This estimate is based on factors such as the unit's size, and the connectivity and integrity of natural ecosystems across it.

(ii) The biodiversity intactness index [47] expresses the average abundance of the native terrestrial species, relative to their abundance in an undisturbed location. It is a global spatial raster with a resolution of 30 arcsec which integrates a measure of abundance-based compositional similarity with pressure variables of land use, land use intensity, human population density, and proximity to the nearest road.

### 2.3. Predicting zoonotic EID risk

We explored the relationship between EID risk and biodiversity variables with a Random Forest classifier algorithm [48] in the R package 'ranger' [49]. We run 6 distinct models, one for each metric of ecological integrity. Every model included the same set of 19 environmental and anthropic predictors used in the original Allen et al. [5] paper, including mammal species richness, plus a new variable referring to ecological integrity (i.e., human footprint, biodiversity habitat index, biodiversity intactness index, ecoregion intactness, contextual intactness, wilderness). We decided not to run a full

model with all integrity metrics together in order to avoid effects of collinearity between these predictors in our analysis (Table S2). We selected Random Forest models over the Boosted Regression Trees models used by Allen et al. [5] after conducting a preliminary random 10-fold cross-validation, wherein Random Forest models exhibited a higher predictive performance, as measured by the true skill statistics metric (TSS) [50] (Table S3).

The fitted models were used to generate relative influence box plots and partial dependence plots with empirical 90% confidence intervals. The importance for each variable was measured as the mean decrease accuracy (MDA) which measures the decrement of the accuracy of model predictions caused by the variable's permutation, averaged across all trees. The MDA scores for each model were then normalized by their largest value, so that the most influential variable in every model took a value of 1 and every other variable was rescaled accordingly, in order to compare variables' importance among models.

We also evaluated how EID risk responds to the interaction between each integrity metric and levels of forest cover, a well-known predictor of EID risk, by implementing bivariate partial dependence plots. We used the R package 'pdp' [51] to create both partial dependence plots and bivariate partial dependence plots.

### 2.4. Dealing with uncertainty in EID events

To account for spatial and sampling uncertainty in EID events, we used a bootstrap resampling regime: each model was fitted on 100 bootstrap samples consisting of 154 zoonotic EID "events", and the resultant predictions from the multiple models were then combined to create a single score and to generate empirical confidence intervals. For each event, we selected a presence point, which is one of the grid cells intersecting the spatial polygon where that event is known to have occurred (a municipality, a region, a country). Each presence point was associated to a background point, that is a grid cell chosen within all other cells which don't intersect that polygon. Both presence and background points were sampled randomly, but the probability of selecting them was weighted based on event coverage and reporting effort. Event coverage is the proportion of a grid cell that overlaps the polygon within the event has occurred. We assigned higher weights to grid cells with larger overlap (i.e., a higher probability of being selected as presences). Reporting effort is the extent to which disease events are reported to become part of the scientific literature, which is a proxy for the power of disease detection and reporting (for details see Allen et al. [5]). Cells with a high value of reporting effort have a higher probability to be chosen both as presence and background points.

### 2.5. Model fitting and validation

For each model, we selected the optimal number of variables randomly sampled as candidates for each split (mtry) and node size using a 5-fold cross validation. We defined a grid of hyperparameters values (n=200) and repeated the cross-validation process for each combination. We then calculated the average performance of the model across all folds for each hyperparameter set and chose the combination with the highest average performance, quantified by the true skill statistic metric.

To estimate models' predictive performance, we ran two validations: a classical random cross-validation, and a stricter spatial block cross-validation. In the random cross-validation data were randomly partitioned into 10 folds, data from 9 folds were used to fit the model whereas data from the remaining fold were used to test predictive ability. We fitted the models iteratively on all folds but one and validated on the left-out fold. In the spatial block cross validation, we created 6 different spatial blocks, one for each continent. We then fitted the models iteratively on all blocks but one and validated on the left-out block. This latter strategy allowed us to separate training and testing sets by using spatial blocks [52] which take into account the spatial autocorrelation of ecological data. Models' predictive performance was estimated with the true skill statistic.

## 3. Results

### 3.1. Model accuracy

For each integrity model, the ability to recognise both presence and background points was always better than random with specificity and sensitivity both higher than 0.5 (Table S4). The random 10-fold cross-validation showed a high predictive power across all models per sample combinations (averaged TSS=0.46, range:0.33–0.60), whereas the spatial cross validation showed a lower performance (averaged TSS=0.27, range: 0.12–0.52) when extrapolating into novel geographical areas [52].

### 3.2. Variables importance in the Random Forest models

Evergreen broadleaf trees variable was the strongest predictor of EID events across all models (median MDA = 0.76, range: 0.74 - 0.80; Figure 1), followed by global environmental stratification (median MDA = 0.60, range: 0.54 - 0.61), and evergreen/deciduous needleleaf trees (median MDA = 0.54, range: 0.45 - 0.57). Additionally, several variables related to human activity showed a high level of relative importance, including human population (median MDA = 0.44, range: 0.41 - 0.48),

cropland coverage (median MDA = 0.37, range: 0.34 - 0.38), and cropland change (median MDA = 0.35, range: 0.34 - 0.37). Notably, while the TSS scores were similar across all models, and also similar to a baseline model without any integrity metric, there was high variation in the relative predictive importance of integrity metrics. Specifically, human footprint showed the largest relative importance (median MDA = 0.55) among the biodiversity and ecological integrity metrics, immediately followed by mammal species richness (median MDA = 0.53, range: 0.49 - 0.57). Ecoregion intactness (median MDA = 0.44) and biodiversity habitat index (median MDA = 0.38) also showed a good predictive power, while wilderness had the lowest normalized MDA (median MDA = 0).

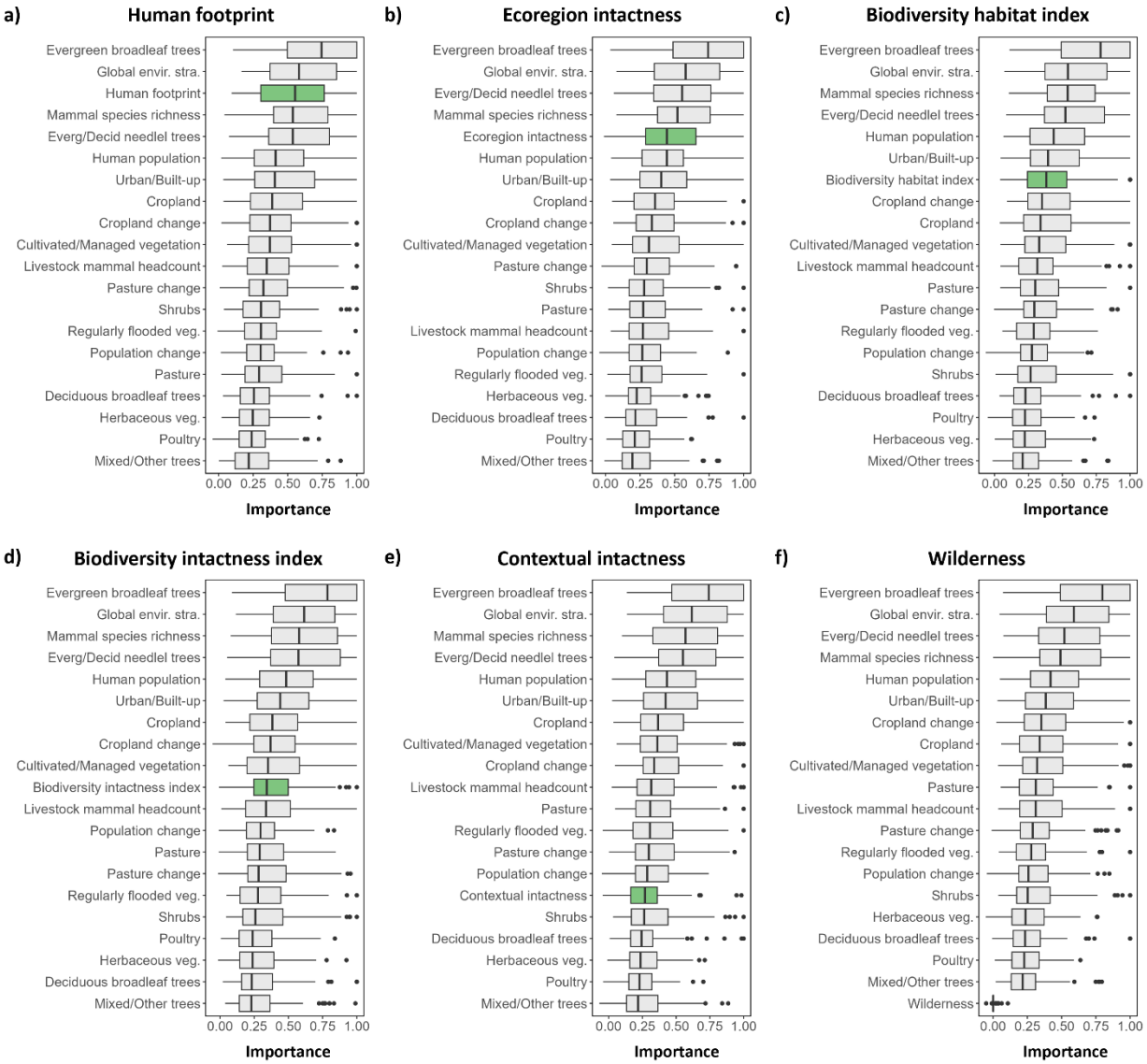


**Figure 1.** Variable's importance plots for each separate model run with the same set of environmental and anthropogenic predictors plus a single integrity metric (in green): (a) human footprint; (b) ecoregion intactness; (c) biodiversity habitat index; (d) biodiversity intactness index; (e) contextual

intactness; (f) wilderness. The importance of each variable is quantified as the MDA of models' predictions caused by variable permutation. Accuracy measures the proportion of observations correctly classified out of the total observations in the dataset. Variables' MDA scores within each model were normalized, so that the most influential variable took a value of 1, and the others were rescaled accordingly. Here, the boxplots show the minimum, first quartile, median, third quartile and maximum MDA scores across 100 replicate models.

We employed partial dependence plots to display the relationship between EID risk and explanatory variables (Figure 2; Figure S1). The partial dependence plot for human footprint showed the highest probability of EID events for values which correspond to rural areas subject to intermediate modification (HFP ~ 5), while the risk is low both for values indicating highly modified areas and for values corresponding to intact and near-intact lands (HFP <3). EID risk correlated with ecoregion intactness and biodiversity habitat index in a similar way, with the highest risk for intermediate values of integrity. For the remaining integrity metrics, the predicted probability showed little variation, indicating a weak influence on EID risk.

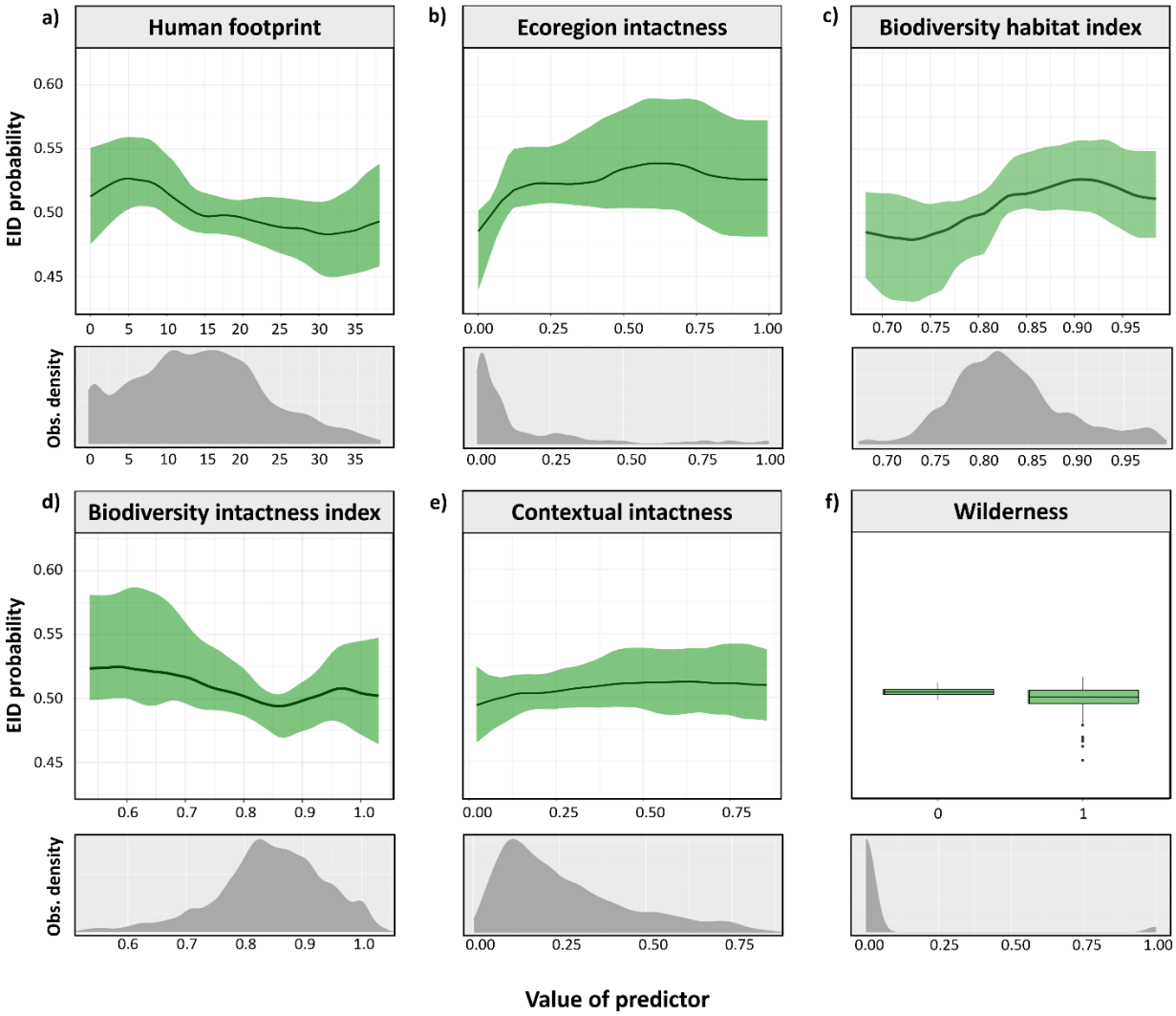


**Figure 2.** Partial dependence plots showing the effect of integrity metrics on EID risk. Each plot represents how the risk (on the y-axis) changes given different values of each variable (on the x-axis): (a) human footprint; (b) ecoregion intactness; (c) biodiversity habitat index; (d) biodiversity intactness index; (e) contextual intactness;(f) wilderness. Black lines show the median and coloured areas show the 90% confidence intervals, computed using a bootstrap resampling regime incorporating uncertainty in EID event locations. Below the plots, we report the density distribution of each biodiversity metric upon which the model was built.

We implemented bivariate partial dependence plots to display how the interaction between integrity metrics and forest cover affects EID risk (Figure 3). For all integrity metrics, the highest risk of EID was observed at intermediate values of forest cover and integrity values, namely: human footprint values of approximately 5, ecoregion intactness around 0.6, and biodiversity habitat index around 0.9. While the predicted probability for wilderness areas exhibited minimal variation between wilderness

(indicated as 1) and non-wilderness areas (indicated as 0), it is noteworthy that EID risk consistently remained lower in wilderness areas.

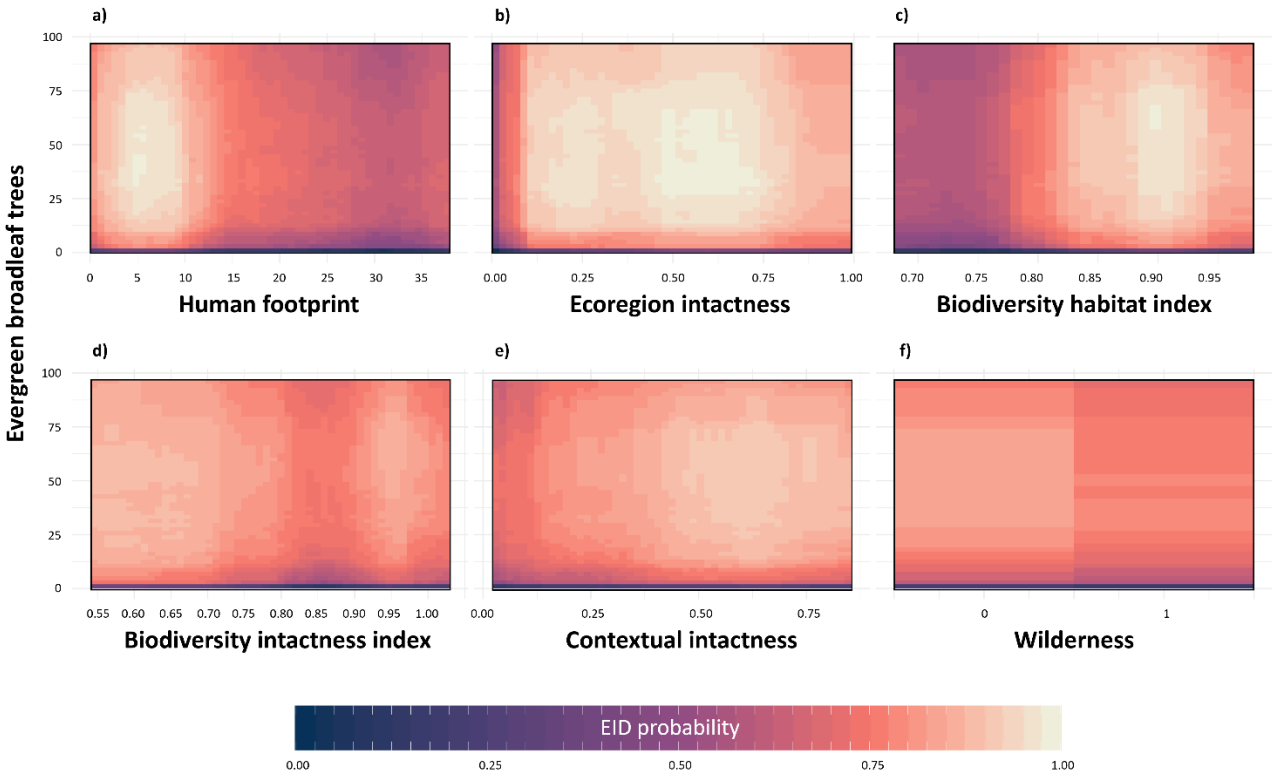


**Figure 3.** Bivariate partial dependence plots showing the effect on EID risk of the interaction between tropical rainforest cover and integrity metrics: (a) human footprint; (b) ecoregion intactness; (c) biodiversity habitat index; (d) biodiversity intactness index; (e) contextual intactness; (f) wilderness. EID probability has been rescaled by its minimum and maximum to allow comparison across integrity metrics.

## 4. Discussion

We developed spatial models to examine how the risk of zoonotic EID changes in response to different metrics of ecological integrity. Our approach allowed us to evaluate the predictive power of these metrics in assessing disease risk and to investigate the intensity and direction of their relationship with EID event probability. This research explores various ecosystem-level metrics, extending beyond the traditional focus on mammal species richness alone. We showed that certain aspects of ecological integrity are important for predicting EID risk, both individually and in interaction with levels of forest cover.

Models consistently performed well in recognizing presence and background points, with high specificity and sensitivity, especially in random 10-fold cross-validation, but showed lower

performance in spatial cross-validation. The differences in model accuracy resulting from the two validation methods are in line with those expected [52] because random cross validation often suffers for the dependence structure of the ecological data, since when validation data are randomly selected from the entire spatial domain, training and validation data from nearby locations will be dependent, and this spatial autocorrelation results in an overoptimistic assessment of models' performance.

Environmental variables were the features that better characterized EID risk, with evergreen broadleaf trees as the strongest predictor of the distribution of EID events, followed by global environmental stratification and mammal species richness. This pattern was already identified in the original Allen et al. model [5], and we confirmed it here with models which include integrity metrics and more recent disease events. Given the high biodiversity levels of tropical forests [53], it is possible that these variables represent a similar process. This trend is consistent with existing theories which suggest that greater host biodiversity expands the pool of available pathogens, thus increasing the probability for novel zoonotic pathogens to emerge [13].

Among integrity metrics, human footprint and ecoregion intactness – respectively representing human pressure and its effects on ecosystems' structure – were especially important in shaping EID risk. The biodiversity habitat index, which measures the effects of anthropogenic disturbances on ecosystems' biological composition, also played an important role in predicting risk. EID risk was high for values of ecological integrity that indicate moderate levels of anthropogenic alteration. These patterns remain consistent regardless of the extent of forest cover, suggesting that human activities altering ecologically intact and nearly intact areas pose a more significant risk of zoonotic disease emergence compared to intensification of activities within areas that have already undergone human modification [13].

The human footprint index was the most important metric in predicting zoonotic EID risk. A low risk of EID was associated both to very high values (>30), and low values of HFP (< 3) which correspond to heavily impacted areas and intact or near-intact lands, respectively [38]. A HFP threshold of 3 has already proved to be an important predictor of transitions in extinction risk for terrestrial mammals [54,55] suggesting that the same drivers of extinction risk, such as habitat loss, degradation, and fragmentation, are also associated with increased zoonotic risk. In fact, the probability of EID events is lower in wilderness areas compared to non-wilderness areas, and this pattern occurs for each level of forest cover. At the same time, highly modified areas are less susceptible to EID risk from wildlife, as very few species are able to live in these areas, leading to a much lower diversity of pathogens. The highest risk was associated with HFP values around 5. Values of HFP between 3 and 5 indicate moderately modified rural areas subject to intermediate levels of human pressures and land

conversion. HFP greater than 3 corresponds to a level of human pressure comparable to that found in pasture lands, while an HFP of 7 or greater is considered equivalent to intensive agriculture [38]. The relationship between human footprint and EID risk emerging from this analysis is consistent with previous findings about the role of human presence and anthropogenic land-use changes in increasing EID risk. Importantly, human footprint was identified as a key predictor of epidemic cases of major vector-borne diseases such as dengue, chikungunya, and Zika, due to human-driven ecological changes that affect vector species distribution and disease incidence [56]. Rural areas are associated with agricultural expansion and intensification that promote EID risk in close proximity with natural highly-diverse habitats [28,57]. Anthropogenic pressures have been shown to alter ecological systems and continue to expand into new areas, with land conversion mostly occurring in tropical rainforests [58]. Changes in land-use increase contact rates between humans and wildlife, disrupting natural disease dynamics [26,28]. These changes often alter species diversity and abundance within ecological communities, affecting pathogens' transmission rates and exposing novel hosts to infection [13,27]. Moreover, the expansion and intensification of livestock production bring domestic animals in close proximity to wildlife habitats, creating a pathway for transmission of zoonotic pathogens from wildlife species to farming communities. High livestock population density, coupled with poor genetic diversity and health conditions, can increase the probability of zoonotic spillover to humans by facilitating pathogen shedding and transmission within livestock, leading to an "amplification effect" [24]. Land conversion is also associated with predictable changes in the local diversity and taxonomic composition of known wildlife hosts of zoonotic pathogens, probably mediated by covariance between traits that influence both host status and tolerance to humans [7]. This could be explained by some life-history features such as small body size, short lifespans, and fast reproduction [59]. These features enable some wildlife species to cope with human disturbance, but at the same time make them more prone to acquire and transmit pathogens due to low energetic investment in adaptive immunity [60,61]. Furthermore, when compared to natural habitats, assemblages in areas under land-use change have more widespread species on average, especially in tropical latitudes [8]. In fact, land-use changes can increase the abundance of selected species that have habitat and dietary niches which overlap with humans enabling direct and indirect contact with similarly adapted sympatric species, domesticated species, and humans [62]. Johnson et al. [35] suggest that spillover risk is higher from animal species adapted to human-dominated landscapes that have increased in abundance and expanded their range following land-use change. Additionally, their research found that less common wildlife species, categorized with increasingly threatened status by the IUCN Red List, host significantly fewer viruses shared with humans, when compared to widespread and abundant wild mammalian species. These trends of biological homogenization occurring in

ecosystems under anthropogenic transformation may amplify disease risk in human-modified lands when widespread, generalist, and synanthropic species are also more competent hosts for pathogens [7].

Along with human footprint, other integrity metrics resulted to be important in predicting EID risk: ecoregion intactness, which accounts for the extent of habitat degradation, fragmentation and the loss of habitat quality caused by human pressures, and biodiversity habitat index, which quantifies the degree of biodiversity intactness in terms of persistence of species diversity. Once again, values representing intermediate levels of human pressures – ~0.67 for ecoregion intactness and ~0.9 for biodiversity habitat index, according to the original sources [42,46] – are associated with an increase in risk. A consequence of the processes of land conversion is the transformation of contiguous natural habitats into smaller, discrete remnant patches embedded in a matrix of human-modified land. The resulting changes in edge density may have cascading ecological effects that influence resource availability, population carrying capacities, species persistence, and the community composition, as well as increase interspecies contact rates influencing how pathogens are transmitted within and between species. As demonstrated by Faust et al. [57], intermediate levels of habitat loss correspond to the maximum edge density, while at high levels of habitat conversion edge density declines, reducing spillover risk. Intermediate values of the biodiversity habitat index indicate an environment where moderate habitat changes can lead to altered population dynamics among wildlife species, creating ecological conditions that can facilitate pathogen transmission among species. The loss of diversity and changes in species composition can increase the risk of emerging infectious diseases, especially when it involves animal species that can play a key role in pathogen transmission [63]. Biodiversity loss due to habitat loss, fragmentation and degradation is size-selective, especially in tropical forests, and the species most likely to disappear are large-bodied species [64], while smaller-bodied species tend to increase in abundance [63]. Smaller-bodied species, such as rodent, are more likely to be competent hosts for many pathogens [60] and several disease systems, such as Lyme disease [36], showed a correlation between large wildlife loss, fluctuations in the abundance of susceptible hosts and increased disease risk [65]. The positive but weak correlation between wilderness and EID risk could derive from an oversimplification of habitat conditions, as both near-intact and impact areas are classified as “non wilderness” in this case. Contextual intactness also revealed a weak influence on zoonotic risk. This could be associated with the complexity of that metric, which accounts for both habitat condition and patterns of compositional dissimilarity [43]. In fact, locations in very different environmental conditions might exhibit similar contextual intactness values, if both are found at the edge of more modified areas, while determining different EID risk levels.

This study acknowledges significant limitations. Firstly, data on disease emergence events are constrained because the primary source, the Emerging Infectious Disease Repository (EIDR) database [18], was last updated in 2013, omitting events since then. Additionally, many spillover events, particularly those associated with mild or non-specific symptoms and no human-to-human transmission, often go undetected or unreported, and the specificity of geographic information for these events varies widely, ranging from precise coordinates to broader regions. Furthermore, we acknowledge the presence of representation bias in our analysis, as reports of emerging diseases are mostly from developed countries [3], while surveillance and spillover detection are lacking in rural and low-income regions. To address this bias, we incorporated a reporting effort index derived from Allen et al. [5] which accounts for the varying likelihood of disease events being reported in the scientific literature. The response variable includes all zoonotic EID events originating from wildlife reported since 1970, regardless of pathogen type or transmission route (Figure S2). This introduces a potential limitation as the effects associated with biodiversity metrics may indirectly reflect variations in pathogen biology and ecology, assuming that not all diseases respond to biodiversity in the same way [56]. In fact, the degree to which biodiversity affects disease risk may depend on several factors related to pathogens: host-specialist diseases that are directly transmitted, without free-living stages, intermediate hosts or vectors are less likely to respond to changes in biodiversity, compared to vector-borne and multi-host pathogens with complex life cycles or free-living infectious stages. Additionally, the influence of biodiversity on disease risk is contingent on various pathogen-related factors, such as transmission type, with density-dependent pathogens responding differently from frequency-dependent ones [57,66]. Pathogens may have multiple transmission routes, and their primary pathway can vary depending on geographic region and host population, increasing uncertainty in large-scale global analyses [67]. It is also possible that at larger spatial scales, models might be unable to capture the mechanisms underlying the relationship between biodiversity metrics and EID risk. Johnson et al. [68] suggested that this relationship varies with scale, being most pronounced locally where species interactions influence it, and weaker at larger scales where climate and abiotic factors take over as dominant drivers. The spatial mismatch between the data of zoonotic spillover events and the explanatory variables also presents a challenge. For instance, the biodiversity intactness index and the contextual intactness variables have a fine resolution reflecting the spatial compositional turnover at the local ecological scale of 1 km. However, when aggregated to a coarse one-degree resolution, these variables might lose valuable information, potentially reducing their importance within risk models. Hence, interpreting these results in absolute terms requires caution, as the predictive power and effect of these variables may vary when assessed on a different, finer scale.

## 5. Conclusion

We showed how metrics that represent different dimensions of ecological integrity affect EID risk, providing evidence that goes beyond simple biodiversity metrics (such as species richness) typically used before. By considering these factors, the study offers a broader picture of how anthropogenic pressures influence ecosystem structure and composition, subsequently affecting zoonotic risk. Our findings reveal that the highest EID risk occurs at the initial stages of ecosystem degradation, as demonstrated by human footprint, ecoregion integrity, and biodiversity habitat index. This suggests that altering areas with high structural and compositional integrity may result in an immediate increase in EID risk. Another important finding emerging from this study is the complex relationship between tropical forest cover and disease risk. We found the assumption that high levels of forest density directly result in high risk is an oversimplification. Our results clearly illustrate this complexity, showing that risk levels increase when, under conditions of consistent forest density, intermediate levels of compositional or structural integrity are present. This study underscores the need to expand our perspective on biodiversity metrics beyond species richness alone. While species diversity is important, it is not the only factor influencing EID risk. Among the integrity variables we examined, the human footprint emerged as a highly predictive factor indicating the importance of the cumulative environmental impact of anthropogenic pressures on zoonotic risk predictions [56].

Empirical findings from our study underscore the importance of preserving ecosystems with high ecological integrity. These areas not only play a crucial role in conserving biodiversity but also contribute to the reduction of zoonotic disease risk, aligning with the targets outlined in the Kunming-Montreal Global Biodiversity Framework [14]. The relationship found between human footprint and EID risk supports the Framework's objectives and emphasizes the need for integrated efforts to safeguard both ecosystem integrity and public health in the face of global environmental changes. Future research should continue to explore the complex connections between biodiversity conservation, ecological integrity, and disease risk, providing a robust empirical foundation for the effective implementation of the Framework's goals [17].

## Ethics

This work did not require ethical approval from a human subject or animal welfare committee.

## Declaration of AI use

We have not used AI-assisted technologies in creating this article.

## Data accessibility

The datasets generated and/or analysed during the current study along with the essential code employed for statistical modelling, including the final modelled dataset, are available from the Zenodo Repository, https://doi.org/10.5281/zenodo.10987602 [69].

## Authors' contributions

L.M.: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; M.D.M.: conceptualization, methodology, supervision, funding acquisition, writing—original draft, writing—review and editing; A.T.: conceptualization, methodology, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

## Conflict of interest declaration

We declare we have no competing interests.

## Funding

This research was supported by EU funding within the NextGeneration EU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT Spoke 4).

## References


1. Morens DM, Fauci AS. 2020 Emerging Pandemic Diseases: How We Got to COVID-19. *Cell*. **182**, 1077–1092. (doi:10.1016/j.cell.2020.08.021)

2. Bernstein AS, Ando AW, Loch-Temzelides T, Vale MM, Li B V., Li H, Busch J, Chapman CA, Kinnaird M, Nowak K, Castro MC, Zambrana-Torrelio C, Ahumada JA, Xiao L, Roehrdanz P, Kaufman L, Hannah L, Daszak P, Pimm SL, Dobson AP. 2022 The costs and benefits of primary prevention of zoonotic pandemics. *Sci. Adv.* **8**. (doi:10.1126/sciadv.abl4183)

3. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2017 Global trends in emerging infectious diseases. *Nature.* **451**, 990–993 (doi: 10.1038/nature06536)

4. Hosseini PR, Mills JN, Prieur-Richard AH, Ezenwa VO, Bailly X, Rizzoli A, Suzán G, Vittecoq M, García-Penã GE, Daszak P, Guégan JF, Roche B. 2017 Does the impact of biodiversity differ between emerging and endemic pathogens? The need to separate the concepts of hazard and risk. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **372**. (doi:10.1098/rstb.2016.0129)

5. Allen T, Murray KA, Zambrana-Torrelio C, Morse SS, Rondinini C, Di Marco M, Breit N, Olival KJ, Daszak P. 2017 Global hotspots and correlates of emerging zoonotic diseases. *Nat. Commun.* **8**. (doi: 10.1038/s41467-017-00923-8)

6. Salkeld DJ, Padgett KA, Jones JH. 2013 A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. *Ecol. Lett.* **16**, 679–686. (doi: 10.1111/ele.12101)

7. Gibb R, Redding DW, Chin KQ, Donnelly CA, Blackburn TM, Newbold T, Jones KE. 2020 Zoonotic host diversity increases in human-dominated ecosystems. *Nature.* **584**,398–402. (doi:10.1038/s41586-020-2562-8)

8. Newbold T, Hudson LN, Contu S, Hill SLL, Beck J, Liu Y, Meyer C, Phillips HRP, Scharlemann JPW, Purvis A. 2018 Widespread winners and narrow-ranged losers: Land use homogenizes biodiversity in local assemblages worldwide. *PLoS Biol.* **16**. (doi:10.1371/journal.pbio.2006841)

9. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, Lloyd-Smith JO. 2017 Pathways to zoonotic spillover. Vol. 15, *Nat. Rev. Microbiol.* **15**, 502–510. (doi: 10.1038/nrmicro.2017.45)

10. Plowright RK, Reaser JK, Locke H, Woodley SJ, Patz JA, Becker DJ, Oppler G, Hudson PJ, Tabor GM. 2021 Land use-induced spillover: a call to action to safeguard environmental, animal, and human health. *Lancet Planet Health*. **5**, 237-245. (doi: 10.1016/S2542-5196(21)00031-0)

11. Keesing F, Ostfeld RS. 2021 Dilution effects in disease ecology. *Ecol. Lett.* **24**,2490–2505. (doi:10.1111/ele.13875)

12. Stewart Merrill TE, Johnson PTJ. 2020 Towards a mechanistic understanding of competence: A missing link in diversity-disease research. *Parasitology.* **147**, 1159–1170. (doi:10.1017/S0031182020000943)

13. 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. (doi:10.1016/j.coviro.2013.01.006)

14. Watson JEM, Evans T, Venter O, Williams B, Tulloch A, Stewart C, Thompson I, Ray JC, Murray K, Salazar A, McAlpine C, Potapov P, Walston J, Robinson JG, Painter M, Wilkie D, Filardi C, Laurance WF, Houghton RA, Maxwell S, Grantham H, Samper C, Wang S, Laestadius L, Runting RK, Silva-Chávez GA, Ervin J, Lindenmayer D. 2018 The exceptional value of intact forest ecosystems. Nat. Ecol. Evol. **2**, 599–610. (doi:10.1038/s41559-018-0490-x)

15. Convention on Biological Diversity. 2022 Monitoring framework for the Kunming-Montreal Global Biodiversity Framework. Montreal, Canada.

16. Di Marco M, Butchart SHM, Visconti P, Buchanan GM, Ficetola GF, Rondinini C. 2016 Synergies and trade-offs in achieving global biodiversity targets. *Conserv. Biol.* **30**,189–195. (doi:10.1111/cobi.12559)

17. Caceres-Escobar H, Maiorano L, Rondinini C, Cimatti M, Morand S, Zambrana-Torrelio C, Peyre M, Roche B, Di Marco M. 2023 Operationalizing One Health: Environmental Solutions for Pandemic Prevention. Ecohealth **20**,156–164. (doi:10.1007/s10393-023-01644-9)

18. EcoHealth Alliance. Emerging Infectious Diseases Repository [Internet] 2016 [cited 2022 Apr 21]. Available from: https://eidr.ecohealthalliance.org/

19. Metzger MJ, Bunce RGH, Jongman RHG, Sayre R, Trabucco A, Zomer R. 2013 A high-resolution bioclimate map of the world: A unifying framework for global biodiversity research and monitoring. *Glob. Ecol. Biogeogr.* **22**,630–638. (doi:10.1111/geb.12022)

20. Tuanmu MN, Jetz W. 2014 A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **23**,1031–1145. (doi:10.1111/geb.12182)

21. Rondinini C, di Marco M, Chiozza F, Santulli G, Baisero D, Visconti P, Hoffmann M, Schipper J, Stuart SN, Tognelli MF, Amori G, Falcucci A, Maiorano L, Boitani L. 2011 Global habitat suitability models of terrestrial mammals. *Philos. Trans. R. Soc. B: Biol. Sci.* **366**, 2633–2641. (doi:10.1098/rstb.2011.0113)

22. Murray KA, Preston N, Allen T, Zambrana-Torrelio C, Hosseini PR, Daszak P. 2015 Global biogeography of human infectious diseases. *PNAS.* **112**,12746–12751. (doi:10.1073/pnas.1507442112)

23. Carlson CJ, Albery GF, Merow C, Trisos CH, Zipfel CM, Eskew EA, Olival KJ, Ross N, Bansal S. 2022 Climate change increases cross-species viral transmission risk. *Nature.* **607**,555–562. (doi:10.1038/s41586-022-04788-w)

24. Karesh WB, Dobson A, Lloyd-Smith JO, Lubroth J, Dixon MA, Bennett M, Aldrich S, Harrington T, Formenty P, Loh EH, Machalaba CC, Thomas MJ, Heymann DL. 2012 Ecology of zoonoses: natural and unnatural histories. *The Lancet.* **380**,1936–1945. (doi:10.1016/S0140-6736(12)61678-X)

25. Center for International Earth Science Information Network. Global Population Count Grid Time Series Estimates, v1 [Internet] 2017[cited 2023 Jan 12]. Available from: https://doi.org/10.7927/H4CC0XNV

26. Keesing F, Belden LK, Daszak P, Dobson A, Harvell CD, Holt RD, Hudson P, Jolles A, Jones KE, Mitchell CE, Myers SS, Bogich T, Ostfeld RS. 2010 Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature.* **468**, 647–52. (doi:10.1038/nature09575)

27. Rohr JR, Barrett CB, Civitello DJ, Craft ME, Delius B, DeLeo GA, Hudson PJ, Jouanard N, Nguyen KH, Ostfeld RS, Remais J V., Riveau G, Sokolow SH, Tilman D. 2019 Emerging human infectious diseases and the links to global food production. *Nat. Sust.* **2**,445–456. (doi:10.1038/s41893-019-0293-3)

28. Gottdenker NL, Streicker DG, Faust CL, Carroll CR. 2014 Anthropogenic Land Use Change and Infectious Diseases: A Review of the Evidence. *Ecohealth.* **11**,619–632. (doi: 10.1007/s10393-014-0941-z)

29. Hurtt GC, Chini L, Sahajpal R, Frolking S, Bodirsky BL, Calvin K, Doelman JC, Fisk J, Fujimori S, Klein Goldewijk K, Hasegawa T, Havlik P, Heinimann A, Humpenöder F, Jungclaus J, Kaplan J, Kennedy J, Krisztin T, Lawrence D, Lawrence P, Ma L, Mertz O, Pongratz J, Popp A, Poulter B, Riahi K, Shevliakova E, Stehfest E, Thornton P, Tubiello FN, Van Vuuren DP, Zhang X. 2020 Harmonization of Global Land-Use Change and Management for the Period 850-2100 (LUH2) for CMIP6. *Geosci. Model Dev.* **13**,5425–5464 (doi:10.5194/gmd-13-5425-2020)

30. Robinson TP, William Wint GR, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, Cinardi G, D'Aietti L, Hay SI, Gilbert M. 2014 Mapping the global distribution of livestock. *PLoS One.* **9**. (doi:10.1371/journal.pone.0096084)

31. Mollentze N, Streicker DG. 2020 Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. *PNAS.* **117**,9423–9430. (doi:10.1073/pnas.1919176117)

32. IUCN. The IUCN Red List of Threatened Species. [Internet]. 2019 [cited 2022 Oct 10]. Available from: https://www.iucnredlist.org

33. Hill SLL, Fajardo J, Maney C, Harfoot M, Harrison M, Guaras D, Jones M, Oliva MJ, Danks F, Hughes J, Burgess ND. The Ecosystem Integrity Index: a novel measure of terrestrial ecosystem integrity with global coverage [internet]. *bioRxiv* [Preprint]. 2022 [cited 2022 Oct 10] Available from: https://www.biorxiv.org/content/10.1101/2022.08.21.504707v1

34. Pulliam JRC, Epstein JH, Dushoff J, Rahman SA, Bunning M, Jamaluddin AA, Hyatt AD, Field HE, Dobson AP, Daszak P. 2012 Agricultural intensification, priming for persistence and the emergence of Nipah virus: A lethal bat-borne zoonosis. *J. R. Soc. Interface.* **9**,89–101. (doi:10.1098/rsif.2011.0223)

35. Johnson CK, Hitchens PL, Pandit PS, Rushmore J, Evans TS, Young CCW, Doyle MM. 2020 Global shifts in mammalian population trends reveal key predictors of virus spillover risk. *Proc. R. Soc. B: Biol.* **287**. (doi:10.1098/rspb.2019.2736)

36. Keesing F, Brunner J, Duerr S, Killilea M, LoGiudice K, Schmidt K, Vuong H, Ostfeld RS. 2009 Hosts as ecological traps for the vector of Lyme disease. *Proc. R. Soc. B: Biol.* **276**,3911–3919. (doi:10.1098/rspb.2009.1159)

37. Keesing F, Ostfeld RS. 2021 Impacts of biodiversity and biodiversity loss on zoonotic diseases. *PNAS.* **118**. (doi:10.1073/pnas.2023540118)

38. Venter O, Sanderson EW, Magrach A, Allan JR, Beher J, Jones KR, Possingham HP, Laurance WF, Wood P, Fekete BM, Levy MA, Watson JEM. 2016 Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* **23**. (doi:10.1038/ncomms12558)

39. Watson JEM, Shanahan DF, Di Marco M, Allan J, Laurance WF, Sanderson EW, Mackey B, Venter O. 2016 Catastrophic Declines in Wilderness Areas Undermine Global Environment Targets. *Current Biology.* **26**,2929–2934. (doi:10.1016/j.cub.2016.08.049)

40. Allan JR, Venter O, Watson JEM. 2017 Temporally inter-comparable maps of terrestrial wilderness and the Last of the Wild. *Sci. Data.* **12**. (doi:10.1038/sdata.2017.187)

690 41. Lesslie RG, Mackey BG, Preece KM. 1988 A Computer-based Method of Wilderness Evaluation. *Environ. Conserv.* **15**,225–32. (doi:10.1017/S0376892900029362)
691
692
693 42. Beyer HL, Venter O, Grantham HS, Watson JEM. 2020 Substantial losses in ecoregion intactness highlight urgency of globally coordinated action. *Conserv. Lett.* **13**. (doi:10.1111/conl.12692)
694
695
696
697 43. Mokany K, Ferrier S, Harwood TD, Ware C, Di Marco M, Grantham HS, Venter O, Hoskins AJ, Watson JEM. 2020 Reconciling global priorities for conserving biodiversity habitat. *PNAS.* **117**,9906–9911. (doi:10.1073/pnas.1918373117)
698
699
700
701 44. Ferrier S, Manion G, Elith J, Richardson K. 2007 Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* **13**,252–264. (doi: 10.1111/j.1472-4642.2007.00341.x)
702
703
704
705 45. Hoskins AJ, Harwood TD, Ware C, Williams KJ, Perry JJ, Ota N, Croft JR, Yeates DK, Jetz W, Golebiewski M, Purvis A, Robertson T, Ferrier S. Supporting global biodiversity assessment through high-resolution macroecological modelling: Methodological underpinnings of the BILBI framework. *bioRxiv* [Preprint]. 2019 [cited 2022 Oct 10] Available from: https://www.biorxiv.org/content/10.1101/309377v3
706
707
708
709
710 46. Harwood T, Ware C, Hoskins A, Ferrier S, Bush A, Golebiewski M, Hill S, Ota N, Perry J, Purvis A, Williams K. 2022 BHI v2: Biodiversity Habitat Index: 30s global time series. v1. CSIRO. Data collection. (doi:10.25919/3j75-f539)
711
712
713
714 47. Newbold T, Hudson LN, Hill SLL, Contu S, Gray CL, Scharlemann JPW, Börger L, Phillips HRP, Sheil D, Lysenko I, Purvis A. 2016 Global patterns of terrestrial assemblage turnover within and among land uses. *Ecography.* **39**,1151–1163. (doi:10.1111/ecog.01932)
715
716
717
718 48. Breiman L. 2001 Random Forests. *Mach. Learn*. **45**,5–32. (doi: 10.1023/A:1010950718922)
719
720 49. Wright MN, Ziegler A. 2017 Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**. (doi:10.18637/jss.v077.i01)
721
722
723 50. Allouche O, Tsoar A, Kadmon R. 2006 Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**,1223–1232. (doi:10.1111/j.1365-2664.2006.01214.x)
724
725
726
727 51. Greenwell BM. pdp: An R Package for Constructing Partial Dependence Plots [Internet]. 2017. Available from: https://CRAN.R-project.org/package=pdp
728
729
730 52. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF. 2017 Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography.* **40**,913–29. (doi:10.1111/ecog.02881)
731
732
733
734
735 53. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000 Biodiversity hotspots for conservation priorities. *Nature*. **403**. (doi:10.1038/35002501)
736
737
738 54. Di Marco M, Venter O, Possingham HP, Watson JEM. 2018 Changes in human footprint drive changes in species extinction risk. *Nat. Commun.* **9**. (doi:10.1038/s41467-018-07049-5)
739

740
741 55. Ramírez-Delgado JP, Di Marco M, Watson JEM, Johnson CJ, Rondinini C, Corredor Llano X, Arias M,
742 Venter O. 2022 Matrix condition mediates the effects of habitat fragmentation on species extinction risk.
743 *Nat. Commun.* **13**. (doi:10.1038/s41467-022-28270-3)
744
745 56. Skinner EB, Glidden CK, MacDonald AJ, Mordecai EA. 2023 Human footprint is associated with shifts in
746 the assemblages of major vector-borne diseases. *Nat. Sustain.* **6**,652–661. (doi: 10.1038/s41893-023-01080-
747 1)
748
749 57. Faust CL, McCallum HI, Bloomfield LSP, Gottdenker NL, Gillespie TR, Torney CJ, Dobson AP, Plowright
750 RK. 2018 Pathogen spillover during land conversion. *Ecol. Lett.* **2**,471–483. (doi: 10.1111/ele.12904)
751
752 58. Pendrill F, Persson UM, Godar J, Kastner T. 2019 Deforestation displaced: Trade in forest-risk commodities
753 and the prospects for a global forest transition. *Environ. Res. Lett.* **14**. (doi:10.1088/1748-9326/ab0d41)
754
755 59. Tonelli A, Caceres-Escobar H, Blagrove M, Wardeh M, Di Marco M. Identifying life-history patterns along
756 the fast-slow continuum of mammalian viral carriers. *Research Square* [Preprint]. 2023 [cited 2024 Mar 18]
757 Available from: https://www.researchsquare.com/article/rs-2722217/v1
758
759 60. Plourde BT, Burgess TL, Eskew EA, Roth TM, Stephenson N, Foley JE. 2017 Are disease reservoirs
760 special? Taxonomic and life history characteristics. *PLoS One.* **12**. (doi:10.1371/journal.pone.0180716)
761
762 61. Lee KA. 2006 Linking immune defenses and life history at the levels of the individual and the species.
763 *Integr. Comp. Biol.* **46**,1000–1015. (doi:10.1093/icb/icl049)
764
765 62. Wilkinson DA, Marshall JC, French NP, Hayman DTS. 2018 Habitat fragmentation, biodiversity loss and
766 the risk of novel infectious disease emergence. *J. R. Soc. Interface.* **15**. (doi:10.1098/rsif.2018.0403)
767
768 63. Keesing F, Young TP. 2014 Cascading consequences of the loss of large mammals in an African Savanna.
769 *BioScience.* **64**, 487–495. (doi:10.1093/biosci/biu059)
770
771 64. Fritz SA, Bininda-Emonds ORP, Purvis A. 2009 Geographical variation in predictors of mammalian
772 extinction risk: Big is bad, but only in the tropics. *Ecol. Lett.* **12**, 538–549. (doi:10.1111/j.1461-
773 0248.2009.01307.x)
774
775 65. Young HS, Dirzo R, Helgen KM, McCauley DJ, Billeter SA, Kosoy MY, Osikowicz LM, Salkeld DJ,
776 Young TP, Dittmar K. 2014 Declines in large wildlife increase landscape-level prevalence of rodent-borne
777 disease in Africa. *PNAS.* **111**,7036–7041. (doi:10.1073/pnas.1404958111)
778
779 66. Keesing F, Holt RD, Ostfeld RS. 2006 Effects of species diversity on disease risk. *Ecol. Lett.* **9**,485–498.
780 (doi:10.1111/j.1461-0248.2006.00885.x)
781
782 67. Cwynar P, Stojkov J, Wlazlak K. 2019 African swine fever status in Europe. *Viruses.* **11**.
783 (doi:10.3390/v11040310)
784
785 68. Johnson PTJ, Ostfeld RS, Keesing F. 2015 Frontiers in research on biodiversity and disease. *Ecol. Lett.*
786 **18**,1119–1133. (doi:10.1111/ele.12479)
787
788 69. Marcolin L, Tonelli A, Di Marco M. 2024 Early-stage loss of ecological integrity drives the risk of zoonotic
789 disease emergence. (https://doi.org/10.5281/zenodo.10987602)
790

70. Marcolin L, Tonelli A, Di Marco M. 2024 Supplementary Materials from: Early-stage loss of ecological integrity drives the risk of zoonotic disease emergence. FigShare (doi:10.6084/m9.figshare.25737704)