

REVIEW

Biases, gaps, and opportunities in mammalian extinction risk research

Luis D. VERDE ARREGOITIA* *School of Biological Sciences, University of Queensland, Brisbane, 4072, Australia. E-mail: luisd@ciencias.unam.mx*

Keywords

bibliometric, hotspots, macroecology, modelling, threat status

*Present address and correspondence: Museo de Zoología "Alfonso L. Herrera", Facultad de Ciencias, Universidad Nacional Autónoma de México, A.P. 70-399, Ciudad Universitaria, México, DF 04510, México.

Submitted: 3 February 2015
Returned for revision: 6 July 2015
Revision accepted: 13 July 2015
Editor: KH

doi:10.1111/mam.12049

ABSTRACT

1. A subdiscipline of conservation science involves the study of differences between threatened and non-threatened species. The relevance of multispecies analyses of extinction risk to conservation practice has been questioned, but there has been no synthesis of the allocation of research effort to different regions and groups of mammals to assess whether or not sufficient knowledge is available to support conservation science where it is most needed.

2. I reviewed 68 comparative studies of mammalian extinction risk to test whether existing research reflects our concern for threatened species. Additionally, I examined the variables used and various methodological issues that can lead to uninformative results.

3. Known spatial and taxonomic biases in conservation science persisted in extinction risk research, leaving large proportions of globally threatened taxa unstudied. Primates and carnivores had more dedicated studies, whereas small mammals such as rodents and the Eulipotyphla (true shrews, talpids, solenodons, gymnures, and hedgehogs) lack research effort despite their high diversity, threat, and extinction record. Except for the Australian mammal fauna (a clear priority given the number of threatened, extinct, and endemic taxa), most areas of conservation importance remain underrepresented in these types of studies.

4. Detailed country-level analyses can provide applicable results for understudied regions. I propose Southeast Asia and the Caribbean for further research, given their high levels of extinction, threat and endemism, and their unique biogeographic histories. Finally, I offer suggestions for general methodological improvements to avoid problems with missing data and statistical circularity in order to maximise conservation relevance.

INTRODUCTION

The rate of biodiversity loss in recent centuries points to an ongoing mass extinction (Barnosky et al. 2011, De Vos et al. 2015). Human activity is causing widespread wildlife declines (Dirzo et al. 2014), but not all species are affected equally (Bennett & Owens 1997). Analyses of past and projected extinctions often find non-random patterns with respect to geography and taxonomy (McKinney 1997, Russell et al. 1998, Purvis et al. 2000a, Davies et al. 2008). The selectivity of extinction offers clues about the importance of different threatening processes, and to how species' traits mediate their responses to these threats (Hanna & Cardillo 2014b).

By synthesising information across large numbers of species, comparative analyses can be used to identify the major drivers behind large-scale extinction processes and to determine which species are most vulnerable (Fisher & Owens 2004, Murray et al. 2014). Comparative extinction risk analyses test for associations between a species' level of endangerment and information on its biology, distribution, and threats (Fig. 1). Common findings for vertebrates are that large body size, restricted distributions, and human pressures all increase a species' likelihood of becoming extinct in the foreseeable future (Fisher & Owens 2004).

Despite their potential for guiding conservation efforts, comparative extinction risk analyses have had little influence on conservation policy (Cardillo & Meijaard 2012).

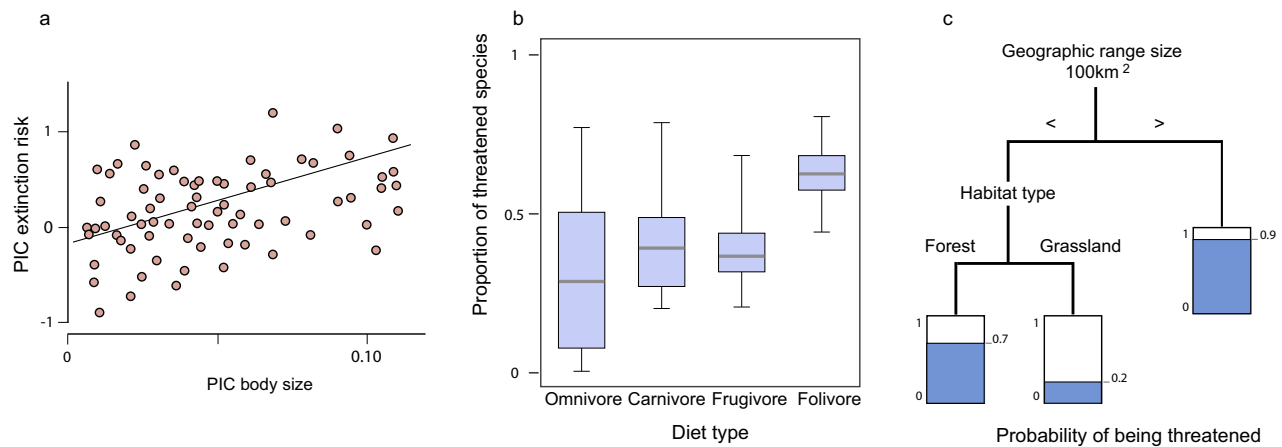


Fig. 1. Hypothetical schematic representations of common results in extinction risk studies for three frequently used methods: a) linear regression on Phylogenetically Independent Contrasts (PIC); b) effect of a categorical variable on the threat status of the species assigned to each category; c) tree-based method for subsetting species into groups with different probabilities of being threatened based on continuous and categorical data.

The results from these studies may be too general for implementation, or too specific to be transferred between taxa, regions, or time periods. Murray et al. (2014) suggest that studies lack detailed information on the threats affecting species (e.g. habitat loss, exploitation). Researchers are focusing on identifying the biological factors that predispose species to extinction without investigating why they are already threatened.

The trade-off between predictive power and generality is obvious in comparative studies. Powerful, informative models tend to have narrow scopes (Cardillo et al. 2008). Fine-scale studies can reveal patterns that are not otherwise apparent because they focus on taxa that share the same environmental context. Nonetheless, global studies are still useful because they have large sample sizes, and can identify the processes that threaten species indiscriminately.

Although mammals receive a disproportionate amount of conservation research relative to other animal groups, research effort is not evenly allocated (Clark & May 2002, Hoffmann et al. 2011). Gaps and biases in the evidence base can mislead or hinder conservation efforts. Exhaustive reviews of the conservation literature have shown that the distribution of research output among species and regions is uneven and that biases vary within different branches of conservation science (Lawler et al. 2006, Stroud et al. 2014). Trimble and van Aarde (2012) argue that threat status, economic importance, or ecological impact drive biases, but to a lesser extent than the personal affinities of scientists, funders, and reviewers towards certain species' characteristics that may be unrelated to research needs (Bonn et al. 2002, Lawler et al. 2006).

Recent publications have addressed some of the issues surrounding the data and methods used in comparative

extinction risk studies (Cardillo & Meijaard 2012, González-Suárez et al. 2012, Murray et al. 2014), but a targeted review of the existing literature on mammals can summarise what has been done and how we can move forward. Gaps in the available data or difficulties with statistical analyses are both important but are separate issues from biases in the overall research effort allocation in mammals. If research interests are misaligned with research needs, then scientists, publishers, and funding agencies should make efforts to ensure sufficient scientific knowledge is available to support conservation where it is most urgent. Specifically, research should be prioritised in areas or for species groups where the topic is understudied despite high threat (Trimble & van Aarde 2012).

It is also necessary to evaluate whether extinction risk research is being done at a scale that considers geopolitical realities, and whether it is focused on areas that would benefit from these studies. Species do not recognise borders, but conservation decisions are usually made and implemented by political units, so nationally endemic taxa represent important conservation targets (Bonn et al. 2002).

In this survey of the literature, I review 68 comparative studies of mammalian extinction risk, aiming to: 1) identify biases and gaps in research effort by comparing the taxonomic and geographic distribution of research output with metrics that are likely to drive conservation research; 2) evaluate the methods and variables used to identify any statistical issues and how they can be addressed; 3) examine the variation in analysis scale; and 4) summarise the state of the field and provide suggestions for effective research on understudied groups or regions.

METHODS

I searched the peer-reviewed literature published up to July 2015 using the Google Scholar search engine (<http://scholar.google.com>). I examined titles, abstracts, and keywords using combinations of 'extinction, decline, persistence' + 'risk, proneness, vulnerability, susceptibility, probability, selectivity' + 'mammal*'. The final results include only peer-reviewed, English language, multi-species studies. Similar searches in Spanish did not yield any peer-reviewed papers.

For each study I recorded the response and predictor variables, taxonomic group, geographic focus, and various details about the analytical techniques employed. I documented the data sources used and determined whether or not authors incorporated primary data. All data were stored in a relational database (see Appendices S1, S6, and S7), manipulated as flat tables in R version 3.2.1 using the 'dplyr' package (Wickham & Francois 2014).

Gaps and biases

In conservation, it is best to measure biases in terms of whether there is disproportionate focus on non-threatened taxa or a limited sample of diversity (Griffiths & Dos Santos 2012). Rather than attempting to explain what drives research output, I monitor the distribution of published research in comparison with emerging research requirements.

The criteria that I consider relevant to this relatively narrow field within conservation include: data richness, extinctions, threat status, and endemism. Data-rich groups could be better represented in the literature, and previous studies have documented taxonomic bias in the availability of biodiversity data (e.g. Boakes et al. 2010). Similarly, taxa or regions with more recorded extinctions or a higher proportion of threatened species are likely to receive more research aimed at understanding the causes and circumstances of these irreversible losses. Lastly, places with more endemic taxa may receive more research interest. The proportion of threatened species in an order is also a relevant measure to indicate vulnerable taxa. Because this proportion is sensitive to taxon size (the number of species in a taxon) and to how many species in an order have been assessed, and because the diversity-adjusted proportion of threatened species correlates tightly with the number of extinctions in an order ($r = 0.87$, d.f. = 24, $P < 0.001$), I did not consider it explicitly in analyses of research bias.

To examine the relative degree of bias among taxonomic groups and geographic regions, I followed Seddon et al. (2005) for calculating the number of studies that would be expected if studies were distributed in proportion to species richness, threat, extinction, and data richness. The expected

number of studies for a given criteria and category is calculated as: (observed value / sum of observed values across all categories) * total number of studies. For example: four out of 26 regional studies are focussed on Africa, where seven mammal extinctions have been recorded out of the 83 global total. Approximately 8% (7/83) of all observed extinctions happened in Africa. If study effort was proportional to observed extinctions, two studies (8% of the observed 26) would relate to Africa.

To quantify the differences between observed and expected number of studies, I examined the pattern of the Pearson residuals (difference between observed and expected values / square root of expected values). Pearson residuals (r_{ij}) are directly comparable irrespective of absolute frequencies, and because they are approximately standard normal (Zeileis et al. 2007), they can be binned into small ($|r_{ij}| < 2$), medium ($2 \leq |r_{ij}| < 4$), and large residuals ($|r_{ij}| \geq 4$). Medium and large residuals can be considered to be individually significant at approximately the $\alpha = 0.05$ and $\alpha = 0.0001$ levels, so they reflect the pattern of deviation from expected values and highlight regions or taxa with biases in research effort. In the case of mammal extinctions in Africa: observed studies [4] – expected studies [2] / square root (expected studies [2]) = 1.414, which indicates positive bias, but not at a significant level.

I derived spatial and taxonomic values of richness, extinction, and threat from the International Union for Conservation of Nature and Natural Resources (IUCN) Red List database (<http://www.iucnredlist.org>). Taxonomic patterns are reported for orders, and spatial patterns for a modification to the United Nations geoscheme (see Table 2 and Appendices S2 and S3). To compare research output with taxonomic patterns in data availability, I calculated the proportion of species with complete data for three important extinction-biasing traits: body mass, gestation length, and geographic range size, from the most comprehensive and widely used database of mammal species' traits (PanTHERIA; Jones et al. 2009). This set of traits follows González-Suárez et al. (2012), representing an indirect measure of how well a species has been studied, and also functions as reliable set of predictors for fitting multivariate models of extinction risk.

Nationally endemic species and threat status

For conservation pragmatism, I consider endemism (an intrinsically relative measure) at the country level. I examined the spatial focus of different studies in order to provide a descriptive evaluation of whether the known geopolitical hotspots of mammalian endemism, extinction, and threat are covered by existing research.

To identify relevant countries in mammalian conservation for comparison with existing research, I used IUCN

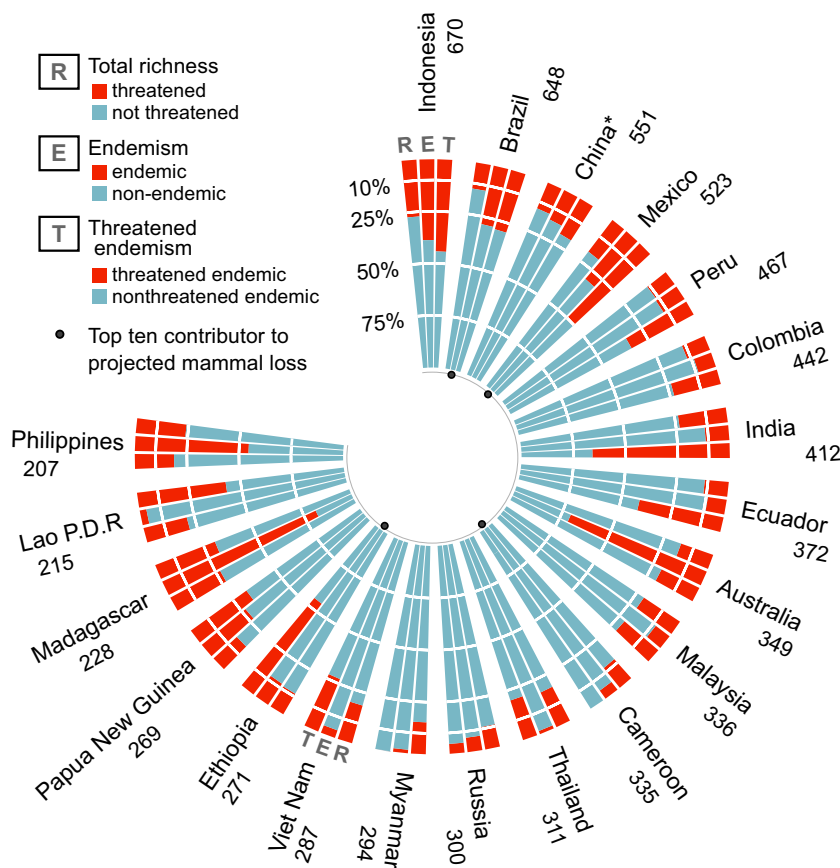


Fig. 2. Patterns of species richness, threat, and endemism for the top 20 countries listed on the basis of having over 200 mammal species of which more than 10% are threatened. Total species richness is shown under each country name. Some country labels (seven o'clock to nine o'clock positions) are flipped for legibility, but the clockwise order of the bars remains unchanged. Sources: IUCN Global Mammal Assessment and Visconti et al. (2011). *Includes the provinces of Hong Kong and Macau, but not Taiwan.

Red List data (Appendix S2) to list and visualise patterns of richness, threat, extinction, and endemism for the 20 countries that have more than 200 species of mammals of which more than 10% are threatened (Fig. 2). To consider future losses, I compared these countries with the top ten national hotspots of projected terrestrial mammal loss identified by Visconti et al. (2011) using global scenarios of socio-economic development.

Variable use

I quantified and annotated the predictor variables used in mammalian extinction risk studies. After standardising the variable names for those that refer to the same attribute under different names (e.g. 'mean body weight' and 'average adult body mass'), I grouped variables used into 11 categories: Behaviour, Conservation response, Distribution, Ecology, Environment, Human pressure, Morphology, Other extrinsic pressure, Phylogeny, Physiology, and Reproduction. Categorising variable types is not straightforward; this classification is based on the definitions in Appendix S4, previously used schemes (Redding et al. 2010, Di Marco et al. 2014), and on how the authors themselves would refer to different types of traits. Eleven categories allowed for

unambiguous classifications, and the risk of incorrectly lumping different metrics was low. I kept measures of central tendency and measures of variability separate. All the information on the units, sample sizes, categorisation schemes for discrete variables, data sources, methodological treatments, and relevance in predicting extinction risk is available in Appendices S1–S7.

Pathways to data loss, scale, and other statistical issues

Comparative analyses perform best if all taxa in the focal clade or region are accounted for, but missing data and other issues can complicate the statistical methods and lead to reduced or biased datasets. Data loss can affect the analysis and interpretation of any results, with consequences on their uptake in policy and management. I checked each paper's methods for statements relating to missing data, or for phrasing that suggested the exclusive use of taxa with complete information. For example, Hanna and Cardillo (2014a) exclude species that lack data on torpor from their analysis, and Fisher et al. (2003) use multiple imputation to complete their species' data. In other cases, the statistical design or study focus involves only taxa with complete data.

I assigned each paper a 'missing data action', selecting from: exclude, impute, assign, accommodate missing data statistically, or not applicable (i.e. no missing data).

Given the importance of scale in extinction risk patterns, I assigned each spatially defined study into a global, continental, or local scale and quantified the mean sample sizes used. I recorded the ways in which data were lost before analysis, noting the treatment of missing data and uncertain phylogenetic relationships. Another reason for the casewise removal of taxa includes the common issue of circularity arising when threat status is not independent from a predictor of extinction risk (e.g. geographic range size).

RESULTS

I analysed a total of 68 studies representing (to my best knowledge) all the published literature for mammals (Appendix S5). More studies were focused on taxonomic groups (28) than on spatially defined (21) or global datasets (19). Taxonomically defined studies were focused mainly on individual orders or families; spatially defined studies varied in their study scale: from single localities to countries, continents, and entire biogeographic realms.

Primates and carnivores outnumbered all other groups in taxonomically defined studies, and there were more studies for these two orders than would be expected given their species richness, extinctions, and threat levels (medium and large positive Pearson residuals, Table 1). Rodents are the most diverse mammal group and they also have the most threatened and extinct species, but they were underrepresented in extinction risk research (negative residuals), and no attempts have been made at a comprehensive cladewise analysis. Bats had more studies than expected for their data availability, and the Eulipotyphla (true shrews and moles, solenodons, gymnures, and hedgehogs) were absent from taxon-specific literature despite their diversity and extinction record. Rodents, primates, bats, Cetartiodactyla, and Eulipotyphla have the highest diversity-adjusted proportion of threatened species, and of these five groups only the primates were adequately represented in the literature.

Geographically, Oceania was particularly well represented in mammalian extinction risk research, followed by Africa and South America. The research output expected from extinction, richness, and threat patterns was exceeded for Oceania (Australia in particular) but lacking for Asia and the Caribbean region. The Caribbean, Europe, and the Middle East had no dedicated studies at all (Table 2). Only three countries had dedicated comprehensive studies: Australia, India, and Bolivia, showing little correspondence with the 20 countries with high threat and endemism in Fig. 2.

Australia's mammal extinction record (globally the highest) was well addressed by the literature, and these

papers include the Diprotodontia and Dasyuromorphia, two groups of mostly Australian marsupials with approximately a third of their species threatened and eight already extinct. Other mammal extinction hotspots such as Mexico, Madagascar, and the Caribbean lacked dedicated studies. Only the African hotspots (Angola, Cameroon, the Democratic Republic of the Congo, Ethiopia, Nigeria, South Africa, and Tanzania) for projected mammal losses from Visconti et al. (2011) were covered by a small number of studies, all remaining areas predicted to lose significant mammalian diversity and endemism remained understudied.

Variable selection

A total of 617 predictor variables appeared in the 68 studies, of which 187 refer to unique metrics and appear in multiple studies. The studies reviewed incorporated between one and 29 predictor variables, with an average (arithmetic mean) of 9.07 and a median of nine. The most widely used variable types describe species' ecology and life history, followed by morphology, and spatial distribution (Fig. 3a). Body size was the most prevalent variable, appearing in 96% (65/68) of studies, followed by geographic range size, used in 45% (31/68) of studies. These two variables were found to be statistically significant predictors of extinction risk more often than any other predictors (Fig. 3b). However, a higher total number of significant results could simply reflect the fact that they were tested more often, and other traits could be more relevant. Proportionally: weaning age, geographic range, and precipitation were the most consistent predictors of extinction risk, showing significant associations with extinction risk in 42%, 41%, and 36% (respectively) of the studies that included them. The high proportional significance for weaning age is not a result of a small number of appearances in extinction risk studies. The prevalent body size variable was significantly related to extinction risk in only a third (32%) of the studies in which it was used.

Kamilar and Paciulli (2008) summarised that variable selection in comparative studies depends on: 1) relevance to extinction vulnerability; 2) inclusion in previous publications; and 3) data availability. The recurring use of certain variables such as body size and geographic range indicates that they probably meet all three criteria, whereas other predictors are introduced when testing new hypotheses about what puts species at risk. The frequent use of some variables (e.g. home range size) did not coincide with how often they were found to be significantly linked with extinction risk. This discrepancy may reflect the difficulty in measuring these traits in wild populations, interspecific variation, or a large number of data that are missing non-randomly.

Table 1. Taxonomically defined mammalian extinction risk studies. Medium and large residuals appear in bold to show biases in research effort. Complete data refers to the percentage of species in the order with no missing data for body mass, gestation length, and geographic range size in the PanTHERIA database. Orders with fewer than 15 species are not shown

Order	Studies	Total species (tot sp)	Extinct species (ext sp)	Threatened species (thr sp)	Complete data % (dat)	Expected number of studies & standardised residuals			
						tot sp	ext sp	thr sp	dat
Rodentia	3	2255	37	355	17	17	20	12	1
						-3.40	-3.80	-2.60	2.92
Primates	9	414	2	257	38	3	1	9	1
						3.28	7.70	0	6.27
Chiroptera	5	1150	5	172	13	9	3	6	0
						-1.33	1.15	-0.41	6.39
Cetartiodactyla	3	329	9	111	50	3	5	3	0
						0	-0.89	-0.50	0.71
Eulipotyphla	0	450	7	84	8	3	4	3	0
						-1.73	-2	-1.73	-0.55
Carnivora	8	285	5	70	46	2	3	2	2
						4.24	2.89	4.24	4.24
Diprotodontia *	4	146	7	45	31	1	4	2	1
						2.73	0	1.41	2.59
Lagomorpha	1	93	1	18	37	1	1	1	1
						0.34	0.64	0.47	-0.35
Afrosoricida	0	54	0	17	18	0	0	1	1
						-0.64	-	-0.77	-0.83
Perissodactyla	0	16	0	13	82	0	0	0	3
						-0.35	-	-0.67	-1.73
Dasyuromorphia*	4	74	1	12	49	1	1	0	2
						4.56	4.76	5.52	1.41
Didelphimorphia	1	95	1	9	9	1	1	0	0
						0.32	0.64	1.22	1.12
Peramelemorphia*	4	22	3	6	43	0	2	0	2
						0.32	1.41	1.22	1.41
Cingulata	0	21	0	3	43	0	0	0	2
						-0.40	-	0.32	-1.41
Macroscelidea	0	16	0	3	47	0	0	0	2
						-0.35	-	-0.32	-1.41
Scandentia	0	20	0	2	15	0	0	0	1
						0.57	-	-0.26	-0.76

*The number of studies shown for these orders is derived from a pattern of research on 'Australian marsupials', none of these has single-order studies.

Missing values, primary data, uncertainty, and information loss

When the studies I reviewed acknowledged taxa with undefined extinction risk values, the most common treatment was to exclude them. In a limited number of cases, the status of these taxa was predicted, or a value was assigned before model fitting (Table 3). Authors who considered taxa and variables with missing predictor values tended to use heuristic model-fitting approaches that accommodate missing values or machine learning methods with no assumptions of completeness. Few studies used imputation techniques or assigned values from related

taxa (Table 3). Only about one-third (25/68) of the studies reviewed used primary data from field work, museum specimens, interviews, or derived new datasets from a range of previously unused sources using novel methods.

Comparing the sample sizes used with the actual total diversity of the study group and/or region defined in the aims of each study shows that smaller scale studies have more representative datasets, often including 75–100% of the taxa outlined in the aims (Fig. 4). More ambitious global studies that involve thousands of species tend to include a smaller proportion (~50%) of the total diversity of the study group defined in the aims.

Table 2. Spatially defined mammalian extinction risk studies. Regionalisation follows the United Nations geoscheme, modified here so that Middle East refers to Western Asia. Medium and large residuals appear in bold to show biases in research effort

Region	Extinct species	Threatened species	Species richness*	Observed studies	Expected number of studies and standardised residuals		
					Extinct species	Threatened species	Species richness
Africa	7	315	80.61	4	2	6	5
Asia	5	411	99.52	1	1.41	-0.82	-0.45
Central America	7	113	43.92	1	2	8	6
Europe	2	30	14.03	0	-0.71	-2.27	-2.04
Middle East	2	56	30.61	0	2	2	3
North America	3	36	26.84	1	-0.71	-0.71	-1.15
Oceania	23	129	37.35	14	1	1	1
South America	11	201	76.43	5	-1.00	-1.00	-1
Caribbean	23	32	16.71	0	1	1	2
					-1.00	-1.00	-1.41
					1	1	2
					0	0	-0.71
					7	3	2
					2.65	6.35	8.49
					3	4	5
					1.15	0.50	0
					7	1	1
					-2.65	-1.00	-1.00

*Area-adjusted species richness (richness/log(area)).

DISCUSSION

Spatial and taxonomic biases reported in the overall conservation literature persist in mammalian extinction risk research, leaving large proportions of threatened taxa underrepresented. With a few exceptions, global and local patterns of threat, diversity, endemism, and data richness do

not correspond with research effort allocation. The overrepresentation of some groups and regions relative to their threat status and diversity points to other factors driving research interest. Identifying the gaps and biases can help outline opportunities for future research on understudied groups or regions, as well as potentially useful variables. Examining the methodological details of these studies

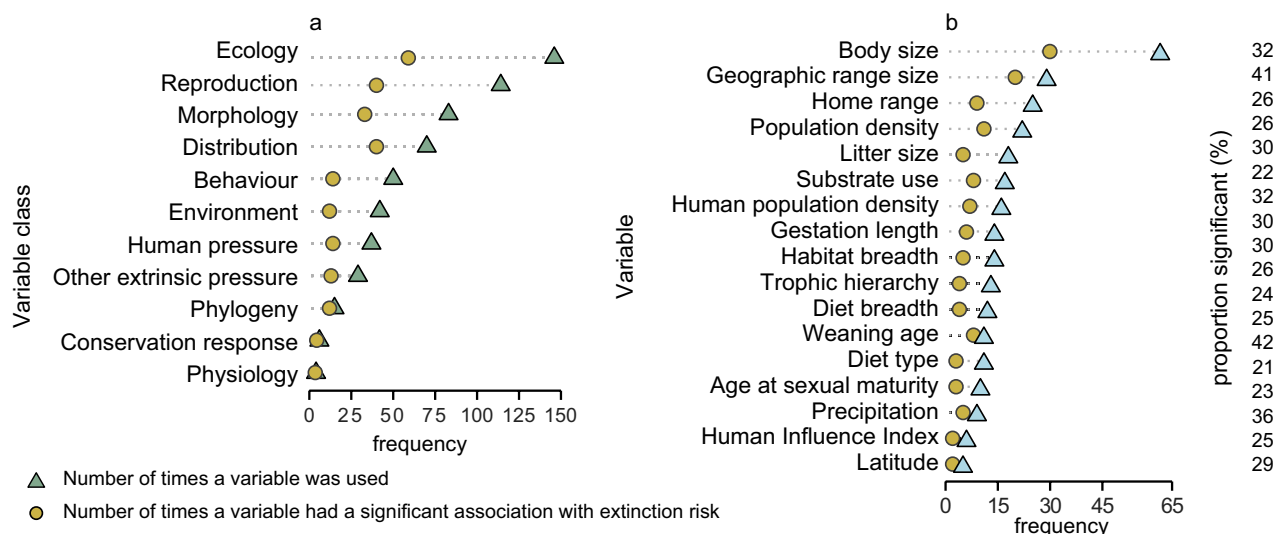
**Fig. 3.** Summary of variable use in 68 comparative extinction risk studies for mammals: a) all 187 unique variables grouped into 11 classes; b) variables that appear in at least five studies.

Table 3. Treatment of missing data in 68 mammalian extinction risk studies, separated into approaches for addressing missing predictor and missing response data (undefined extinction risk)

Treatment of missing data:	
Approach	Frequency
Predictor variables	
Not applicable	38
Accommodate incomplete cases	14
Exclude	9
Impute	6
Assign from related taxa	1
Response variables	
Not applicable	34
Exclude	26
Exclude but predict using fitted models	5
Assign value	3

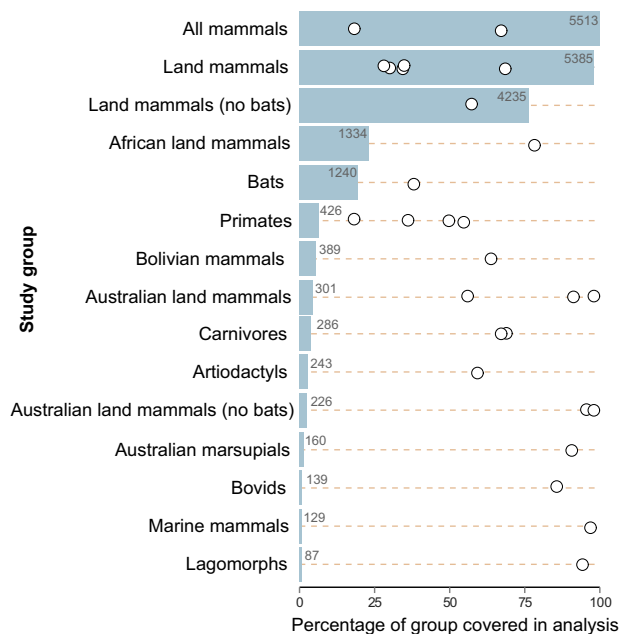
also allowed me to identify potential improvements for generating accurate, informative, and conservation-relevant models in future investigations.

Gaps and biases

Taxonomic biases are directed at primates and carnivores. These two orders are often considered flagship taxa for aesthetic reasons (Smith et al. 2012), and the ecological reasons for the extensive research attention focused on them include

their contributions to processes such as predation, herbivory and seed dispersal, and their role as apex consumers. Primates and carnivores are relatively data-rich orders, and they include a number of large-bodied, wide-ranging species. They also represent well-resolved clades within the mammal tree of life (Davies & Yessoufou 2013), with phylogenetic hypotheses that became available earlier than those of other orders (Purvis 1995, Bininda-Emonds et al. 1999). Most comparative studies rely on phylogenetic regression techniques, so primates and carnivores provided a testing ground for comparative methods research.

Spatially, extinction risk research is clearly biased towards Oceania, and Australia in particular. This research may represent a genuine effort to understand why so many species declined and identify vulnerable species for conservation planning. Alternatively, extinction risk research in Australia may be driven by a notable exception to the positive body size – extinction risk association that is well known for mammals: the ‘Critical Weight Range’. The validity and biological interpretation of this pattern is an ongoing debate (e.g. Woinarski 2014) and may account for the disproportionate research effort. Australia also represents an exception to the latitudinal pattern of disparity in biodiversity data. Tropical regions and the Southern Hemisphere have fewer biodiversity data than temperate, boreal biomes, but more emphasis is placed on biodiversity data collection and conservation attention in Australia than in other Southern Hemisphere countries (Collen et al. 2008).

**Fig. 4.** Coverage in 31 taxonomically defined comparative extinction risk studies. Labelled bars show current total numbers of species in each group following the IUCN Red List taxonomy, open circles show the percentage of each group included by each study.

Scale

Although it is difficult to quantify, I speculate that small-scale studies are likely to incorporate more detailed data and cover a higher proportion of the study group or region. At fine scales, researchers are more apt to apply expert knowledge and consult specialised or grey literature, leading to more complete data sets and more informed choices of predictor variables. For example, even within the colobine monkey clade (known dietary specialists), subtle variations in the degree of folivory play a significant role in predicting persistence (Kamilar & Paciulli 2008). These findings emphasise the importance of examining correlates of extinction risk at various levels of analysis. Categorising species in broad, artificially discrete categories (a common practice in large-scale studies with large sample sizes) may obscure details that could be critical for assigning levels of extinction risk.

Despite the conservation relevance of country-level studies, these represent a small minority of the overall literature, making any efforts to compare research patterns with socio-economic data unfeasible. Other than Australia, none of the nations that meet multiple conservation criteria has dedicated studies. At a national scale, threatened species

lists are relatively easier to update and reflect the magnitude of the extinction threat realistically (Ceballos & Brown 1995). National taxonomic checklists tend to be more regularly updated than global taxonomies and can help establish the identity of species before introducing uncertainty in further analyses. Finally, the findings from country-oriented studies should be easier to incorporate into effective conservation strategies if the results are not confounded by processes that do not operate within the study area and there are no conflicts in conservation policy.

Suggesting multiple studies for every region or taxonomic group with threatened or endemic species is unreasonable. Although taxon-specific studies are interesting for macroecology, thorough analyses for a few key regions have the potential to produce novel findings and support realistic conservation measures. The geographic overlap of multiple criteria in political units has practical advantages, so most of the study themes suggested below are at the country level. As a clear example of the relevance of political units in conservation, eight countries – Australia, China, Colombia, Ecuador, Indonesia, Malaysia, Mexico, and the USA – cover less than 1% of the world's surface but were identified as responsible for more than half of the global deterioration in the conservation status of vertebrates (Rodrigues et al. 2014).

Research opportunities

Setting conservation priorities is best left to informed models that consider socio-economic (e.g. governance, policy, welfare), biological, and human pressure data in a structured decision-making framework. However, the following regions are absent from extinction risk research, and they represent important opportunities based on their unique combinations of conservation issues, mammalian faunas, and biogeography. Formal assessments consistently identify these areas as 'hotspots' for conservation (Myers et al. 2000, Mittermeier et al. 2011). Below, I suggest regions that represent important understudied opportunities for dedicated analyses of mammalian extinction risk, which were underrepresented in the literature, based on the criteria of threatened species, richness, and extinctions (Table 2).

Asia, and Southeast Asia in particular, harbours an exceptionally high number of endemic mammals that are threatened by the loss of >70% of original habitat (Myers et al. 2000). Various socio-economic factors present formidable challenges for conservation in the region (Sodhi et al. 2010). Aside from high endemism and unique biogeographic histories, two countries in this region contain vast areas that are irreplaceable in the context of cost-effective conservation: Indonesia and the Philippines (Carwardine et al. 2008).

Indonesia tops the lists of mammalian endemism, richness, and threat, and it lies in an area of interchange between historically isolated biogeographic provinces (Ceballos & Brown 1995). Quantifying the relative importance of different threats at a national level can direct enforcement actions, and any discrepancies between species' actual and predicted extinction risk may prompt formal re-evaluations. Mammals in the Philippines represent the foremost example of diversification within an ancient oceanic archipelago (Rickart et al. 2011). For its area, the Philippines has the greatest concentration of endemic mammals in the world, and these species are distributed heterogeneously across islands that represent unique biogeographic units (Esselstyn et al. 2004). The drivers of threat for these species can only be investigated adequately with an in-depth analysis that captures the idiosyncratic nature of the threats to mammals throughout the archipelago, considering the increased sensitivity of island endemics to anthropogenic habitat changes (Fordham & Brook, 2010).

The Caribbean region includes 28 nations. This may complicate conservation management, but an investigation of extinction risk for this unique mammalian fauna is warranted. Human settlement and biological invasions in the Lesser and Greater Antilles have caused 23 confirmed mammal extinctions (the same as Australia in only a fraction of the land area) and the loss of up to 90% of the original vegetation (Myers et al. 2000).

Methodological suggestions

Reviewing the methods in mammalian extinction risk studies shows that two common issues have the potential to alter the results and interpretation: 1) reduced and unrepresentative datasets; and 2) lacking primary data. These two issues are not a novel finding, but they had not been quantified across the extinction risk literature for any taxonomic group. Any future study can benefit from avoiding data loss and incorporating new primary data.

Unnecessarily degraded datasets

There were multiple pathways to reduced datasets, mainly through the casewise removal of species with missing values, statistical circularity, or unclear phylogenetic relationships (Table 3). Various approaches were used to address missing data and avoid case-deletion. Purvis et al. (2000b) described a heuristic approach for searching model space, using multiple regression with model simplification to find minimum adequate models. This approach, and recent modifications to it (Hanna & Cardillo 2013a), can accommodate missing data and was the most commonly used. Alternatively, imputation methods provide valuable

alternatives to removing missing observations. They produce low errors and retain relationships among traits (Penone et al. 2014). Current modelling approaches can account for imputation error, and modern studies have successfully employed imputation to avoid biases (e.g. Di Marco et al. 2012).

Indiscriminately excluding species with no extinction risk data can leave entire political units or taxonomic groups out of the analysis, but undefined extinction risk values should not hinder comparative studies. These taxa can be included at different points in the modelling process. They can be omitted from model fitting but then the models can be used to predict their threat status. Alternatively, a status can be assigned using either a precautionary approach that classifies these species as threatened, or a relaxed approach that considers them as non-threatened. (e.g. González-Suárez et al. 2013). Several data-deficient taxa have listed threats and available trait data, and recent papers provide methods to estimate the conservation status of data-deficient mammals so that these can be considered in conservation research (Bland et al. 2014, Jetz & Freckleton 2015).

Some species are considered to be at risk simply because of limited geographic occurrence regardless of population status, creating circularity when using geographic range size as a predictor (Purvis et al. 2000b, Grandcolas et al. 2010, Murray et al. 2011). Circularity is usually avoided by excluding all species that fit this criterion. If species with restricted ranges are not randomly distributed in space or across the mammal phylogeny, excluding them from analysis can lead to misleading results and interpretation. The number of mammal species listed this way varies from 370 to 380 throughout different iterations of the IUCN Red List. This apparent circularity can be overcome in various ways. First, alternative measures of extinction risk may be used if they do not include range size as a basis (e.g. population trends). Alternatively, preliminary analyses to determine the effect of including or excluding range size as a predictor, and including or excluding geographically restricted species, can guide the final model fitting (González-Suárez & Revilla 2013, Di Marco et al. 2014). Finally, geographic range size can simply be excluded as a predictor of extinction risk, because other variables can take its place. Extrinsic pressures not only increase model accuracy (Murray et al. 2014); they were found (in combination with climate) to predict species' geographic range sizes accurately (Di Marco & Santini 2014). If extrinsic and environmental variables are measured adequately and incorporated in analyses, they can ultimately replace geographic range size and improve model interpretation.

Phylogeny is rarely used as a predictor variable (Verde Arregoitia et al. 2013), but it is a critical component of the comparative method. Sometimes taxa with complete data are excluded from analyses because of unclear phylogenetic

affinities (e.g. Liow et al. 2009). The growing number of phylogenetic data sources and methods that allow for phylogenetic uncertainty in analyses should prevent species from being excluded on the basis of unclear affinities. When analytical methods do not explicitly incorporate phylogenetic data, or when they are non-existent for a group or region, taxonomic random effects and exploratory sensitivity analyses can solve this problem (e.g. Matthews et al. 2011, Davidson et al. 2012, Hanna & Cardillo 2013b).

Despite analytical advances to overcome missing values and uncertainty, the need for ongoing data collection on species' traits is critical. In a promising development, Kissling et al. (2014) present a workflow for digitising, extrapolating, and validating global trait datasets, and use it to compile a comprehensive dataset of mammal diets. This workflow can be applied to other types of biological data, such as those identified here as commonly used and significant (e.g. weaning age).

CONCLUSION

Comparative extinction risk studies are mostly desk studies that are not resource-intensive. The suggestions I propose here represent realistic evidence-based study topics with relevance to conservation science. Biodiversity is not evenly distributed through space, and neither are the drivers of its decline, so comparative studies can still play an important role when investigating how to conserve species and areas in peril and reduce the threats with the highest impacts (Rodrigues et al. 2014).

Conservation scientists need to tackle any perceived difficulties associated with studying 'uncharismatic' or data-poor mammals and to prioritise underrepresented geographic regions and taxonomic groups. The methodological recommendations provided here should help researchers meet the level of scientific rigor required by conservation journals, and the dataset provided with this paper (Appendices S6 and S7, a comprehensive metadata compilation of all the reviewed studies) can be used to conduct further analyses that will lead to new insights. Likewise, the scientific publishing environment should put aside any issues with novelty or general interest for taxonomically narrow research (e.g. single-order studies) that accentuate biases and remove incentives for the basic research or capacity building that is urgently needed in conservation.

ACKNOWLEDGEMENTS

I was supported by funding from Consejo Nacional de Ciencia y Tecnología (CONACYT) scholarship 308685 and by a University of Queensland UQ International stipend. Special thanks to Emily Hanna for comments on the initial outline of this manuscript.

REFERENCES

- Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB et al. (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.
- Bennett PM, Owens IPF (1997) Variation in extinction risk among birds: chance or evolutionary predisposition? *Proceedings of the Royal Society B: Biological Sciences* 264: 401–408.
- Bininda-Emonds ORP, Gittleman JL, Purvis A (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews* 74: 143–175.
- Bland LM, Collen B, Orme CDL, Bielby J (2014) Predicting the conservation status of data-deficient species. *Conservation Biology* 29: 250–259.
- Boakes EH, McGowan PJ, Fuller RA, Chang-qing D, Clark NE, O'Connor K, Mace GM (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology* 8: e1000385.
- Bonn A, Rodrigues ASL, Gaston KJ (2002) Threatened and endemic species: are they good indicators of patterns of biodiversity on a national scale? *Ecology Letters* 5: 733–741.
- Cardillo M, Meijaard E (2012) Are comparative studies of extinction risk useful for conservation? *Trends in Ecology and Evolution* 27: 167–171.
- Cardillo M, Mace GM, Gittleman JL, Jones KE, Bielby J, Purvis A (2008) The predictability of extinction: biological and external correlates of decline in mammals. *Proceedings of the Royal Society B: Biological Sciences* 275: 1441–1448.
- Carwardine J, Wilson KA, Ceballos G, Ehrlich PR, Naidoo R, Iwamura T, Hajkiewicz SA, Possingham HP (2008) Cost-effective priorities for global mammal conservation. *Proceedings of the National Academy of Sciences* 105: 11446–11450.
- Ceballos G, Brown JH (1995) Global patterns of mammalian diversity, endemism, and endangerment. *Conservation Biology* 9: 559–568.
- Clark JA, May RM (2002) Taxonomic bias in conservation research. *Science* 297: 191–192.
- Collen B, Ram M, Zamin T, McRae L (2008) The tropical biodiversity data gap: addressing disparity in global monitoring. *Tropical Conservation Science* 1: 75–88.
- Davidson AD, Boyer AG, Kim H, Pompa-Mansilla S, Hamilton MJ, Costa DP, Ceballos G, Brown JH (2012) Drivers and hotspots of extinction risk in marine mammals. *Proceedings of the National Academy of Sciences* 109: 3395–3400.
- Davies TJ, Yessoufou K (2013) Revisiting the impacts of non-random extinction on the tree-of-life. *Biology Letters* 9: 1–4.
- Davies TJ, Fritz SA, Grenyer R, Orme CDL, Bielby J, Bininda-Emonds ORP et al. (2008) Phylogenetic trees and the future of mammalian biodiversity. *Proceedings of the National Academy of Sciences* 105: 11556–11563.
- De Vos JM, Joppa LN, Gittleman JL, Stephens PR, Pimm SL (2015) Estimating the normal background rate of species extinction. *Conservation Biology* 29: 452–462.
- Di Marco M, Santini L (2014) Human pressures predict species' geographic range size better than biological traits. *Global Change Biology* 21: 2169–2178.
- Di Marco M, Cardillo M, Possingham HP, Wilson KA, Blomberg SP, Boitani L, Rondinini C (2012) A novel approach for global mammal extinction risk reduction. *Conservation Letters* 5: 134–141.
- Di Marco M, Buchanan GM, Szantoi Z, Holmgren M, Grottolo Marasini G, Gross D, Tranquilli S, Boitani L, Rondinini C (2014) Drivers of extinction risk in African mammals: the interplay of distribution state, human pressure, conservation response and species biology. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 1–12.
- Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the anthropocene. *Science* 345: 401–406.
- Esselstyn JA, Widmann P, Heaney LR (2004) The mammals of Palawan island, Philippines. *Proceedings of the Biological Society of Washington* 117: 271–302.
- Fisher DO, Owens IPF (2004) The comparative method in conservation biology. *Trends in Ecology & Evolution* 19: 391–398.
- Fisher DO, Blomberg SP, Owens IPF (2003) Extrinsic versus intrinsic factors in the decline and extinction of Australian marsupials. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: 1801–1808.
- Fordham DA, Brook BW (2010) Why tropical island endemics are acutely susceptible to global change. *Biodiversity and Conservation* 19: 329–342.
- González-Suárez M, Revilla E (2013) Variability in life-history and ecological traits is a buffer against extinction in mammals. *Ecology Letters* 16: 242–251.
- González-Suárez M, Lucas PM, Revilla E (2012) Biases in comparative analyses of extinction risk: mind the gap. *Journal of Animal Ecology* 81: 1211–1222.
- González-Suárez M, Gómez A, Revilla E (2013) Which intrinsic traits predict vulnerability to extinction depends on the actual threatening processes. *Ecosphere* 4: art76.
- Grandcolas P, Nattier R, Legendre F, Pellens R (2010) Mapping extrinsic traits such as extinction risks or modelled bioclimatic niches on phylogenies: does it make sense at all? *Cladistics: The International Journal of the Willi Hennig Society* 27: 181–185.
- Griffiths RA, Dos Santos M (2012) Trends in conservation biology: progress or procrastination in a new millennium? *Biological Conservation* 153: 153–158.
- Hanna E, Cardillo M (2013a) A comparison of current and reconstructed historic geographic range sizes as predictors of extinction risk in Australian mammals. *Biological Conservation* 158: 196–204.

- Hanna E, Cardillo M (2013b) Island mammal extinctions are determined by interactive effects of life history, island biogeography and mesopredator suppression. *Global Ecology and Biogeography* 23: 395–404.
- Hanna E, Cardillo M (2014a) Clarifying the relationship between torpor and anthropogenic extinction risk in mammals. *Journal of Zoology* 293: 211–217.
- Hanna E, Cardillo M (2014b) Predation selectively culls medium-sized species from island mammal faunas. *Biology Letters* 10: 1–4.
- Hoffmann M, Belant JL, Chanson JS, Cox NA, Lamoreux J, Rodrigues ASL, Schipper J, Stuart SN (2011) The changing fates of the world's mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 2598–2610.
- Jetz W, Freckleton RP (2015) Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370: 1–10.
- Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90: 2648–2648.
- Kamilar JM, Paciulli LM (2008) Examining the extinction risk of specialized folivores: a comparative study of Colobine monkeys. *American Journal of Primatology* 70: 816–827.
- Kissling WD, Dalby L, Fløjgaard C, Lenoir J, Sandel B, Sandom C, Trøjelsgaard K, Svenning J-C (2014) Establishing macroecological trait datasets: digitalization, extrapolation, and validation of diet preferences in terrestrial mammals worldwide. *Ecology and Evolution* 4: 2913–2930.
- Lawler JJ, Aukema JE, Grant JB, Halpern BS, Kareiva P, Nelson CR et al. (2006) Conservation science: a 20-year report card. *Frontiers in Ecology and the Environment* 4: 473–480.
- Liow LH, Fortelius M, Lintulaakso K, Mannila H, Stenseth NC (2009) Lower extinction risk in sleep-or-hide mammals. *The American Naturalist* 173: 264–272.
- McKinney ML (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annual Review of Ecology and Systematics* 28: 495–516.
- Matthews LJ, Arnold C, Machanda Z, Nunn CL (2011) Primate extinction risk and historical patterns of speciation and extinction in relation to body mass. *Proceedings of the Royal Society B: Biological Sciences* 278: 1256–1263.
- Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C (2011) Global Biodiversity Conservation: The Critical Role of Hotspots. In: Zachos FE, Habel JC (eds) *Biodiversity Hotspots*, 3–22. Springer, Berlin Heidelberg.
- Murray KA, Rosauer D, McCallum H, Skerratt LF (2011) Integrating species traits with extrinsic threats: closing the gap between predicting and preventing species declines. *Proceedings of the Royal Society B: Biological Sciences* 278: 1515–1523.
- Murray KA, Verde Arregoitia LD, Davidson A, Di Marco M, Di Fonzo MMI (2014) Threat to the point: improving the value of comparative extinction risk analysis for conservation action. *Global Change Biology* 20: 483–494.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Penone C, Davidson AD, Shoemaker KT, Marco MD, Rondinini C, Brooks TM, Young BE, Graham CH, Costa GC (2014) Imputation of missing data in life-history traits datasets: which approach performs the best? *Methods in Ecology and Evolution* 5: 961–970.
- Purvis A (1995) A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 348: 405–421.
- Purvis A, Agapow P-M, Gittleman JL, Mace GM (2000a) Nonrandom extinction and the loss of evolutionary history. *Science* 288: 328–330.
- Purvis A, Gittleman JL, Cowlshaw G, Mace GM (2000b) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267: 1947–1952.
- Redding DW, DeWolff CV, Mooers AO (2010) Evolutionary distinctiveness, threat status, and ecological oddity in Primates. *Conservation Biology* 24: 1052–1058.
- Rickart EA, Balet DS, Rowe RJ, Heaney LR (2011) Mammals of the northern Philippines: tolerance for habitat disturbance and resistance to invasive species in an endemic insular fauna. *Diversity and Distributions* 17: 530–541.
- Rodrigues ASL, Brooks TM, Butchart SHM, Chanson J, Cox N, Hoffmann M, Stuart SN (2014) Spatially explicit trends in the global conservation status of vertebrates. *PLoS ONE* 9: e113934.
- Russell GJ, Brooks TM, McKinney MM, Anderson CG (1998) Present and future taxonomic selectivity in bird and mammal extinctions. *Conservation Biology* 12: 1365–1376.
- Seddon PJ, Soorae PS, Launay F (2005) Taxonomic bias in reintroduction projects. *Animal Conservation* 8: 51–58.
- Smith RJ, Veríssimo D, Isaac NJB, Jones KE (2012) Identifying Cinderella species: uncovering mammals with conservation flagship appeal. *Conservation Letters* 5: 205–212.
- Sodhi NS, Posa MRC, Lee TM, Bickford D, Koh LP, Brook BW (2010) The state and conservation of Southeast Asian biodiversity. *Biodiversity and Conservation* 19: 317–328.
- Stroud JT, Rehm E, Ladd M, Olivas P, Feeley KJ (2014) Is conservation research money being spent wisely? Changing trends in conservation research priorities. *Journal for Nature Conservation* 22: 471–473.
- Trimble MJ, van Aarde RJ (2012) Geographical and taxonomic biases in research on biodiversity in human-modified landscapes. *Ecosphere* 3: art119.
- Verde Arregoitia LD, Blomberg SP, Fisher DO (2013) Phylogenetic correlates of extinction risk in mammals: species in older lineages are not at greater risk. *Proceedings of the Royal Society B: Biological Sciences* 280: 1–7.
- Visconti P, Pressey RL, Giorgini D, Maiorano L, Bakkenes M, Boitani L et al. (2011) Future hotspots of terrestrial mammal

- loss. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 2693–2702.
- Wickham H, Francois R (2014) dplyr: a grammar of data manipulation. *R package version 0.3.0.2*.
- Woinarski JCZ (2014) Critical-weight-range marsupials in northern Australia are declining: a commentary on Fisher et al. (2014) ‘The current decline of tropical marsupials in Australia: is history repeating?’. *Global Ecology and Biogeography* 24: 118–122.
- Zeileis A, Meyer D, Hornik K (2007) Residual-based shadings for visualizing (conditional) independence. *Journal of Computational and Graphical Statistics* 16: 507–525.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s web-site.

Appendix S1. Data definitions for the flat data tables recording different aspects of 68 comparative extinction risk

studies for mammals. The *AuthorYear* field can be used to establish the relationship between both tables for database manipulations.

Appendix S2. IUCN Global Mammal Assessment data for the 20 countries depicted in Fig. 2.

Appendix S3. Threatened and extinct species for extant mammal orders. Dermoptera, Hyracoidea, Microbiotheria, Notoryctemorphia, and Tubulidentata are not shown because they have no threatened or extinct species. Source: IUCN Global Mammal Assessment summary statistics.

Appendix S4. Definitions used to categorise predictor variables into 11 types when authors of the studies reviewed did not classify or define the trait types used.

Appendix S5. Full list of studies reviewed.

Appendix S6. Table of studies used.

Appendix S7. Table of variables used.