

Correlates of rediscovery and the detectability of extinction in mammals

Diana O. Fisher* and Simon P. Blomberg

School of Biological Sciences, Goddard Building (8), The University of Queensland, St Lucia, Queensland 4072, Australia

Extinction is difficult to detect, even in well-known taxa such as mammals. Species with long gaps in their sighting records, which might be considered possibly extinct, are often rediscovered. We used data on rediscovery rates of missing mammals to test whether extinction from different causes is equally detectable and to find which traits affect the probability of rediscovery. We find that species affected by habitat loss were much more likely to be misclassified as extinct or to remain missing than those affected by introduced predators and diseases, or overkill, unless they had very restricted distributions. We conclude that extinctions owing to habitat loss are most difficult to detect; hence, impacts of habitat loss on extinction have probably been overestimated, especially relative to introduced species. It is most likely that the highest rates of rediscovery will come from searching for species that have gone missing during the 20th century and have relatively large ranges threatened by habitat loss, rather than from additional effort focused on charismatic missing species.

Keywords: rediscovery; mammal extinction; habitat loss; introduced predators; overkill; extinction rate

1. INTRODUCTION

Species presumed to be extinct are often rediscovered. For example, 89 Australian vascular plants were rediscovered between 1981 and 2001, and rediscovery was the reason for disqualification of 13 mammals from a list of 144 candidate extinct species analysed in 1999 [1,2]. Conservation resources are wasted searching for species that have no chance of rediscovery, while most missing species receive no attention. Public understanding of the extent of the current extinction crisis may be compromised if lists of extinct species are too conservative [3]. On the other hand, premature designation of species extinction will lead to the withdrawal of conservation effort [4]. Quantitative predictions of which missing species are most likely to be extinct and which might be rediscovered are therefore needed.

Modern-era species extinctions (since 1500) have been assessed based on two criteria. Provided that a species is taxonomically accepted [2], the species is considered extinct if it remains missing either for a prescribed waiting period (the pre-1995 IUCN threshold was 50 years), or after a specified search effort [5]. The current IUCN criteria are phrased in terms of search effort: 'exhaustive surveys in known and/or expected habitat, at appropriate times, throughout its historic range have failed to record an individual... over a time frame appropriate to the taxon's life cycle and life form' [5]. The emphasis is now on demonstrating adequate search effort, because extinction is difficult to detect, even in the best-known taxa. The current IUCN criteria acknowledge this by including a status category of 'Critically Endangered' with a tag of '(Possibly Extinct)'. A decade ago, a comprehensive review found only 36 per cent of

purported mammal extinctions to be resolved. The rest had insufficient evidence of extirpation or taxonomic validity, or had been rediscovered [2]. MacPhee & Flemming [2] argued that although the 50-year threshold was not inherently meaningful, some substantial, unspecified waiting period is necessary because of frequent rediscoveries of supposedly extinct mammals. Lack of detection for 50 years might not constitute good evidence of extinction, depending on the distribution of previous sightings [6]. For species with five or more dated sightings, quantitative techniques have recently been developed to estimate the probability of extinction, based on the distribution of sighting records [6,7]. However, 70 per cent of purportedly extinct mammal species are known from fewer than five sightings (D. Fisher & S. Blomberg 2010, unpublished data). We do not know which traits are correlated with longer gaps in species detection records, or the frequency of species rediscovery.

We might expect the rediscovery rate of missing species to depend on factors including search effort [8] and the size of the area that investigators need to search. We predicted that the probability of species rediscovery would also depend on species traits that predict extinction risk in mammals, such as small geographical range size, ecological specialization, large body size and slow reproduction [9,10]. We might further expect to find interactions between the cause of decline in missing species and traits that predict elevated extinction risk, because small range size and specialization are expected to disproportionately increase extinction risk in species affected by habitat loss [11,12], and large body size to increase extinction risk particularly in harvested and persecuted species, and those subject to overkill by introduced predators [13,14]. The aims of this study were to quantify the frequency of species rediscovery, and to identify traits that predict the probability of rediscovery in missing mammals.

* Author for correspondence (d.fisher@uq.edu.au).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsob.2010.1579> or via <http://rsob.royalsocietypublishing.org>.

2. MATERIAL AND METHODS

(a) *Data and definitions*

In order to compare traits and detection rates of currently detected (extant) and undetected (missing) species, we first needed to define an appropriate dataset. We compiled a global database of species with long gaps in their sighting records: (i) species that have been reported extinct (including 'Extinct' (EX), 'Critically Endangered (Possibly Extinct)' (CR (PE)) and 'Extinct in the Wild' (EW) species in the Red List) or flagged as missing in the literature (so have the most recent detection date reported); and (ii) species that have published accounts of rediscovery, including records of detection history (see electronic supplementary material, dataset S1). To minimize the possibility that our choice of definition of 'missing' influenced our results, we analysed both the full dataset and subsets of the data with a stricter definition of assumed extinction, omitting species classed as data-deficient (but with species account information indicating that they are very likely extinct), now extinct in the wild, or subject to taxonomic disagreement, and including only missing species currently designated as EX or CR (PE). We also analysed a subset of the data including only the most recent two centuries of records (see below). We established the status and detection history of these species after they were reported missing using past and present IUCN Red Lists and related publications, primary literature, books and the Committee on Recently Extinct Organisms mammal database [2]. We classed each species as currently missing/extinct or rediscovered, and recorded the date of last sighting and rediscovery.

In order to calculate rediscovery rates, we included the dates of all rediscoveries (i.e. species that were rediscovered but are now again missing were classified as rediscovered). Only one rediscovered species is classified as Extinct by the IUCN: the desert rat kangaroo (*Caloprymnus campestris*), which was found in 1931 after having been missing for 90 years [15], and disappeared again in 1935, shortly after the red fox (*Vulpes vulpes*) reached the lake Eyre basin of South Australia, where it occurred [16]. Two other rediscovered Australian species (the central rock rat, *Zyzomys pedunculatus* and Christmas island shrew, *Crociodura trichura*), a Cuban species (Garrido's hutia, *Mysateles garridoi*) and a Solomon Islands species (the Vanikoro flying fox, *Pteropus tuberculatus*) are classified as Critically Endangered (Possibly Extinct). We omitted species that have been 'rediscovered' through taxonomic revision (e.g. 'splitting'), and only included mammals that are named as taxonomically accepted full species (not subspecies) in the IUCN Red List, which lists assessments of most current taxonomically accepted species of mammals [17].

For each species, we recorded available data on the cause of extinction or the threat associated most with decline. A main threat was assigned if one threat was reported to be the major one, whether it be habitat loss (deforestation, agricultural clearing, fragmentation, degradation, overgrazing), overkill (harvesting, hunting, exploitation, persecution or bycatch) or introduced species (invasive predators or diseases, predominantly comprising: black rat, *Rattus rattus*; domestic cat, *Felis catus*; red fox, *Vulpes vulpes*; Indian mongoose, *Herpestes javanicus*; domestic dog, *Canis lupus familiaris*; pig, *Sus scrofa*; brown rat, *Rattus norvegicus*; kiore, *Rattus exulans*; and a trypanosome contracted from black rats). We also recorded additional threats interacting

with the major extinction driver, such as loss of vegetation cover exacerbating predation. If no threat was reported to be the major one (two or three of the three main classes of threat were responsible, or the likely cause was not known), we assigned no threat to that species (i.e. 'threat' was treated as missing data). We recorded geographic range rank (range with a precision of one order of magnitude: 1 = up to 1 km²; 2 = 1–10 km²; 7 = 100 000–1 000 000 km²). We used a rank because range was often not known with sufficient precision to treat it as a continuous variable. Range estimates (e.g. [18]) are from different sources, so we assume that estimates of range are unlikely to vary by more than an order of magnitude between authors. We defined search effort as the number of reported search expeditions targeted to the species after it was reported missing and before it was rediscovered (if rediscovered). We divided the number of reported searches into three ranked intervals: low (0–2), medium (3–6) and high (>10). Reported searches are likely to be underestimates, but we assume that the three broad categories reflect meaningful relative ranks of recent search effort, because except in the case of undisputed pre-20th century extinctions (which we ranked as low effort; e.g. Steller's sea cow, *Hydrodamalis gigas*), authors of the IUCN Red List always noted search effort in their accounts of extinct and possibly extinct species, and publications that report rediscoveries invariably discussed the frequency of previous sightings and unsuccessful search expeditions.

We assessed the original reported density of each species under three categories (sparse, 1; locally common, 0.5; dense, 0), if we found unambiguous published statements about abundance. For example, the Christmas Island shrew, *C. trichura*, was assessed as dense because it was 'once extremely common all over the island and its distinctive shrill squeaks could be heard all around as one stood quietly in the rainforest' [19]; the desert bandicoot, *Perameles eremiana*, was assessed as locally common because 'according to Finlayson, it was common in northwestern South Australia, the southwest of the Northern Territory and adjacent parts of western Australia in the 1930s. Its range extended as far north as the Tanami desert' [20]; and the Jamaican rice rat, *Oryzomys antillarum*, was assessed as sparse because 'doubtless this is the field mouse described by P. H. Gosse in his *Naturalist's sojourn in Jamaica* as... far from numerous' [21]. We assessed research effort as the mean number of Web of Science citations in 2009 that had the topic keywords 'taxonomy' or 'conservation', and an address keyword as the countries in which the species occurred. We also recorded discovery date (the date that the type specimen was collected, not necessarily the date of description); elevation (coastal, up to 50 m.a.s.l.; mid, 50–1000 m.a.s.l.; high, over 1000 m.a.s.l.); habitat openness (closed: forest or swamp; open: grassland, desert, rocky coast or shrubland); island status (island versus continent); body mass (g); last sighting date; century (19th, 20th or pre-19th); colour (cryptically coloured, black, grey, brown or white versus conspicuously coloured, spotted or striped); arboreality (arboreal or terrestrial); diurnality (diurnal versus nocturnal or crepuscular); gregariousness (group-living versus solitary); and current human density rank (mean human density in the geographical range in categories of 50 km⁻²: from 0 = up to 1 person km⁻² to 15 = more than 651 km⁻²). Human density rank was obtained from the most recent data in [18] for most species, and from the United Nations Population Division

Table 1. The probability of rediscovery versus predictive variables in the final model of a Cox proportional hazards regression with multiple imputation, using data from all centuries, excluding the effect of century and including search effort (s.e. is standard error, p is the p -value for the final imputed dataset). Results from the broad dataset are non-italicized. If different, results from the restricted dataset (with a narrower definition of ‘missing’) are italicized.

variable	coefficient	exp(coef)	s.e.	p
range rank	0.457 <i>0.395</i>	1.58 <i>1.48</i>	0.009 <i>0.101</i>	<0.0001
threat (habitat loss)	1.217 <i>1.379</i>	3.38 <i>3.97</i>	0.034 <i>0.363</i>	<0.001
threat (overkill)	0.612 <i>0.841</i>	1.84 <i>2.32</i>	0.390 <i>0.402</i>	0.04
density category	0.762 <i>0.898</i>	2.14 <i>2.46</i>	0.401 <i>0.433</i>	<0.01
search category (medium)	1.068 <i>0.812</i>	2.91 <i>2.25</i>	0.302 <i>0.326</i>	<0.001 <i>0.01</i>

database (<http://esa.un.org/unpp/>) for some island species, assuming that human density on the island was representative of the species range, in mammals restricted to small islands.

Clavero & Garcia-Berthou [22] noted that the searchable classification system of threats in the IUCN Red List does not cover all species, so researchers need to assess detailed information about the causes of extinction provided in other fields of the database. Therefore, we assessed all extended species accounts, in order to extract information on threatening processes and detection history. We include our data on missing and rediscovered mammals and sources as electronic supplementary material (dataset S1 and text S1).

(b) Statistical analysis

We tested for associations between species characteristics and probability of rediscovery using Cox proportional hazards regression [23]. We modelled the annual ‘survival’ of missing/extinct status with respect to potential covariates (i.e. rediscovery was analogous to death in a survival model, in which mean survival is compared between groups with different traits, and the result expressed as a mean survival curve for each group). In species missing before the 19th century, there was no variation in search effort, and no species in which extinction was attributed mainly to habitat loss. Therefore, we did two tests: (i) using a dataset including species from all centuries and search effort, but excluding the ‘century’ variable; and (ii) with a reduced dataset including century (19th and 20th), but excluding search effort and species missing before the 19th century.

Case-wise deletion of species without complete data is likely to give misleading results in a multiple regression [24], so we dealt with missing data using multiple imputation (MI [25]). MI provides unbiased estimates of parameters from the data, together with standard errors that take account of the imputation procedure [26]. Five datasets were generated after 10 000 iterations of a Gibbs sampler using the ‘mice’ package for R [27,28]. Convergence was assessed by visual examination of the trace plots for each variable. Missing data were never greater than 6 per cent for any of the variables in our dataset. Data were imputed for the continuous variables range rank, log mass, human rank and inverse density using a Bayesian multiple regression, with the conventional improper flat prior [29]. The three binary variables colour, diurnality and gregariousness were imputed using logistic regression. Threat, an ordinal variable, was imputed using polytomous regression. Combined inference for the multiply imputed datasets was conducted using the ‘mitools’ package for R [30].

For analyses of the ‘all centuries’ and the ‘19th and 20th century’ datasets, we tested the assumptions of Cox regression (the hazard for any species is a fixed proportion

of the hazard for any other species, risk is multiplicative) by calculating weighted residuals using the `cox.zph` function in R [31]. We tested for multi-collinearity using variance inflation factors, calculated using the ‘car’ package in R [32]; no values were >2 (indicating no problematic multi-collinearity for either dataset [33]).

We did not test for correlates of rediscovery rate using a method to control for phylogenetic non-independence, because rate data are strongly skewed (non-normally distributed), and no method currently exists to account for phylogenetically correlated data in a Cox regression. However, rediscovered species were dispersed evenly across the mammal phylogeny, and showed no apparent phylogenetic signal (see electronic supplementary material, figure S1).

Using the ‘all centuries’ dataset, we tested for correlations between species characteristics and discovery date, using both generalized least-squares models to account for phylogenetic non-independence and linear models of raw species data. We report the results of all models using raw species data, because there was no significant difference between any of the phylogenetic and non-phylogenetic linear model results ($p > 0.99$ in each case). Rediscoveries are scattered across families and genera of mammals, rather than being clumped in certain phylogenetic lineages (for details of the phylogeny and tree-building methods, see electronic supplementary material, figure S1).

3. RESULTS

(a) Detectability of extinction in all centuries

We identified 187 mammal species that have been missing (claimed or suspected to be extinct) since 1500. This number includes all such mammals for which we were able to find key variables for analysis. In the complete dataset, 67 species that were once missing have been rediscovered. When species from all centuries were included, mammals that declined mainly due to habitat loss were much more likely to have been wrongly suspected to be extinct than species affected by overkill or introduced species. The rate of rediscovery in species affected by loss of habitat was 3.4 times, and in species affected by persecution 1.8 times as high as in species affected by introduced predators and diseases (table 1 and figure 1). Mammals with larger geographical ranges and lower original population densities were also more likely to be wrongly suspected to be extinct (table 1). To check that our results did not depend on interpretation of the major threat and multiple threat interactions, we also repeated this analysis using only species listed as affected by only one threat. The results were qualitatively unchanged, with the same significant variables and

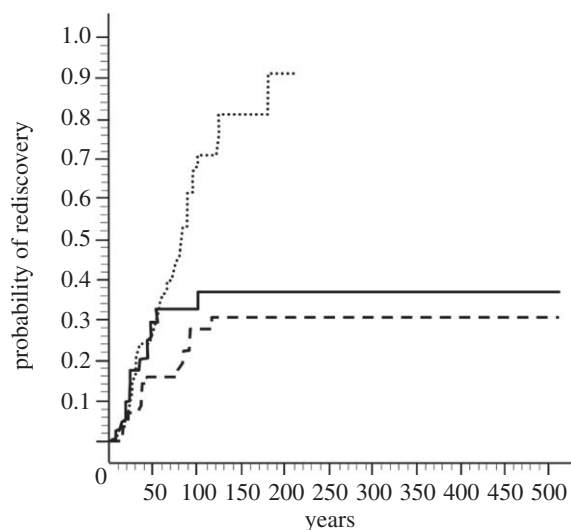


Figure 1. Cox proportional hazards regression showing the mean difference in the probability of rediscovery between species affected by different threats, in all centuries. Habitat loss (including fragmentation and degradation), dotted line; introduced species (including invasive diseases and predators), dashed line; overkill (including harvesting, persecution and exploitation), solid line. The y -axis shows the probability of rediscovery for each category per year. Time zero is when all species in the category are missing (0 = zero rate of rediscovery, 1 = 100% rediscovered). Cox proportional hazards regression does not extrapolate beyond the data: the curve for species affected by habitat loss ends at 180 years because the earliest suspected extinction attributed to habitat loss was 180 years ago (the Bahian tree rat, *Phyllomys unicolor*). Several species affected by introduced species or overkill disappeared in the 16th century ($n = 187$).

relative effect sizes (electronic supplementary material, table S1). In the reduced dataset with a narrower definition of ‘missing’, 56 species were rediscovered and 98 were missing. The conclusions based on this reduced dataset were also unchanged, with the same significant variables, and even larger effect sizes for the main correlates (table 1).

The distribution of search effort (the number of reported search expeditions targeting each missing species) was highly skewed. Most species were subject to two or fewer expeditions, but six species in the broad dataset that were exterminated in the 20th century were the targets of more than 11 reported searches each (figure 2). Search effort affected the detection rate of missing species. Most species with up to two targeted searches (low effort) remain missing, but most species with three to six searches (intermediate effort) have been rediscovered (figure 2). Species with intermediate effort were rediscovered at a rate 2.9 times higher than species with low effort (table 1). However, the association between effort and detectability was nonlinear. None of the missing species subject to more than 11 rediscovery attempts have been found. No other modelled variables (including body size, life history, habitat, appearance, cryptic habits or density of overlapping human populations) were significantly associated with the probability of rediscovery (table 1).

Harvested, persecuted or exploited mammals elicited more attention than species affected by other threats.

They were targeted by more than twice as many reported searches as species that declined from habitat loss or introduced species ($F_{2,174} = 4.6$, $p = 0.01$; 3.4 ± 0.93 (s.e.) searches versus 1.5 ± 0.18 for habitat loss and 1.7 ± 0.28 for introduced species).

(b) Detectability of extinction in the 19th and 20th centuries

Habitat loss was not considered a main cause of extinction until after 1800 (electronic supplementary material, dataset S1). In the model including only 19th and 20th century extinctions and effects of century, an interaction between threat and range size was the only significant effect. In mammals that declined from habitat loss, species with larger geographical ranges were much more likely to be wrongly thought of as extinct, and, conversely, claims of extinction in species with very small ranges have nearly always been confirmed (table 2 and figure 3). In these species, each order of magnitude increase in range increased the odds of rediscovery by a factor of 1.6. Species with the largest ranges (100 000–1 000 000 km², rank of 7) therefore had odds of rediscovery 26.84 times higher than species with the smallest ranges that declined from habitat loss (up to 1 km², rank of 1; table 2 and figure 3). There was no interaction effect between range size and overkill, or range size and introduced species. The results based on the reduced dataset were unchanged, with the same significant variables and even larger effect sizes (table 2).

Mammals missing in the 20th century were nearly three times as likely to be rediscovered as those that disappeared in the 19th century. There was also an interaction between century and threat; mammals affected by overkill were 6.62 times as likely to be rediscovered in the 20th century as in the 19th century. However, these effects of century were marginally non-significant (table 2). No other modelled variables (including body size, life history, habitat, appearance, cryptic habits, or density of overlapping human populations) were associated with probability of rediscovery (table 2).

4. DISCUSSION

A substantial proportion (more than a third) of mammal species that have been classified as extinct or possibly extinct, or flagged as missing, have been rediscovered. Searching for missing species takes substantial effort and funding, and many missing species have a high scientific or public profile and high potential conservation importance if found [3,7,8]. It is therefore important that investigators prioritize their effort to missing species that are most likely to be detected [7].

The missing species most likely to be rediscovered are those with large ranges that declined from habitat loss. Mammal extinctions have been attributed to habitat loss only in the last two centuries, and our analysis of this time period showed that larger range size predicted higher probability of rediscovery only in species affected by habitat loss. This is consistent with most evidence in birds, showing that habitat loss causes disproportionate global and local extinction of restricted range endemics in comparison with other threats [13], and with models showing that endemics–area relationships predict

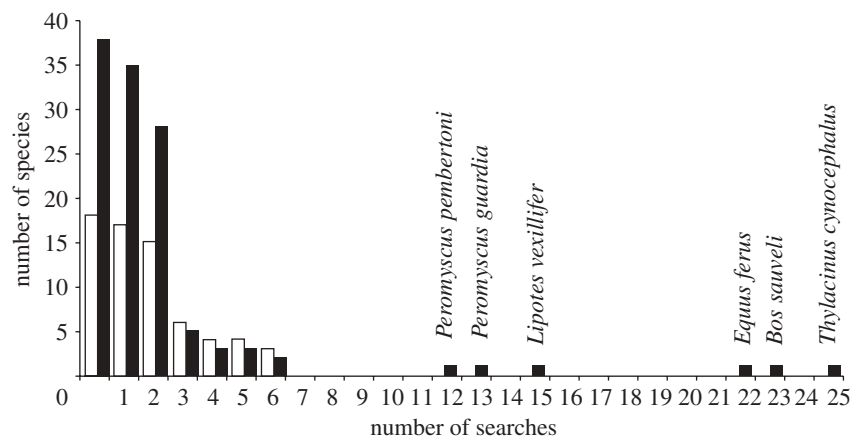


Figure 2. Frequency distribution of the number of reported searches targeting particular missing species, in all centuries. Rediscovered species are white bars, missing species are black bars. Species with high numbers of searches (≥ 12 , all unsuccessful) are Pemberton's deer mouse (*Peromyscus pembertonii*) and the Angel Island deer mouse (*Peromyscus guardia*) from Mexico; the baiji (*Lipotes vexillifer*) from China; wild horse (*Equus ferus*) from China, Mongolia, Belarus, Germany, Kazakhstan, Lithuania, Poland, the Russian Federation and Ukraine; kouprey (*Bos sauveli*) from Thailand, Laos and Cambodia; and the thylacine (*Thylacinus cynocephalus*) from Australia (for an explanation of sources for search effort, see electronic supplementary material, text S1; $n = 178$).

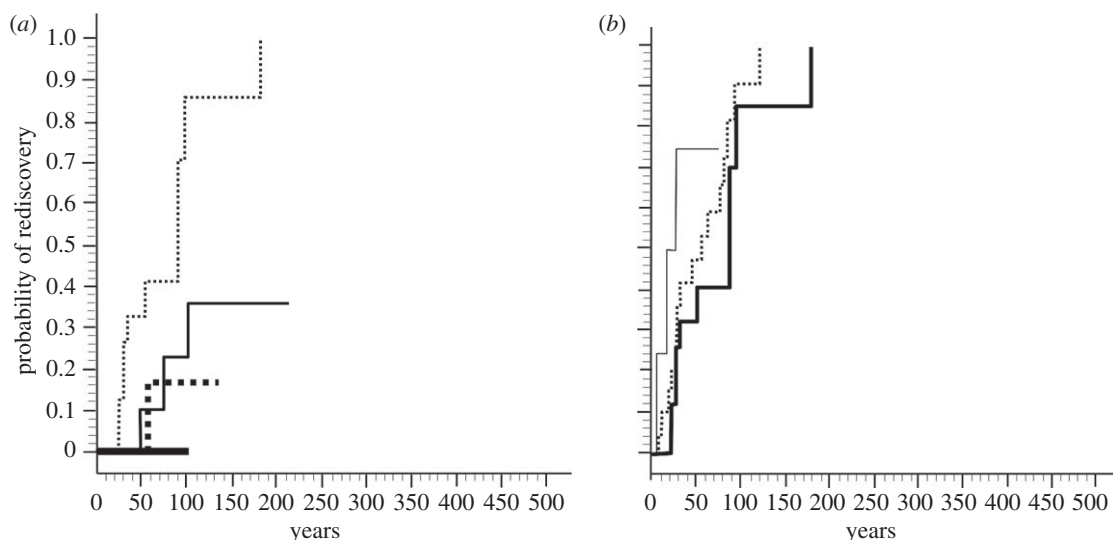


Figure 3. Cox proportional hazards regression showing mean probability of rediscovery for species affected by habitat loss in seven categories of range size. All species were last recorded in the 19th or 20th century. (a) Small geographical ranges (thick solid line, <1 km²; thick dotted line, 1–10 km²; thin solid line, 11–100 km²; thin dotted line, 101–1000 km²); (b) larger geographical ranges (thick solid line, 1001–10 000 km²; dotted line, 10 001–100 000 km²; thin solid line, 100 001–1 000 000 km²). Time zero is when all species in the category are missing (0 = zero rate of rediscovery, 1 = 100% rediscovered; $n = 164$).

Table 2. The probability of rediscovery versus predictive variables in the final model of a Cox proportional hazards regression with multiple imputation, using data from the 19th and 20th centuries, including the effect of century and excluding search effort (s.e. is standard error, p is the p -value for the final imputed dataset). Results from the broad dataset are non-italicized. If different, results from the restricted dataset (with a narrower definition of 'missing') are italicized.

variable	coefficient	exp(coef)	s.e.	p
century	1.07 0.93	2.92 2.53	0.57 0.59	0.061 0.12
threat (overkill) : century (20th)	1.89 2.45	6.62 11.59	0.99 1.23	0.058 0.046
threat (habitat loss) : range rank	0.47 0.51	1.59 1.67	0.19 0.21	0.016

extinction rates from habitat loss better than species–area relationships [34].

Our finding that mammals affected by habitat loss are most likely to be rediscovered suggests that the current number of species considered extinct owing to habitat

loss is likely to be overestimated. Because small range is a variable used to both ascertain extinction risk and assign the cause (habitat loss), circularity might lead to overestimation of the proportion of extinctions that are due to habitat loss [11]. Severely declined mammals are

likely to be considered as specialists on their last detected habitat. Now being restricted to a small range, they will be categorized as threatened by habitat loss, even if the cause of previous decline was different. This will inflate perceived extinction risk owing to habitat loss if some of these species actually persist undetected in other habitats or distant sites. Fisher [35] found that species affected by habitat loss are more likely to be rediscovered at the periphery than the centre of their former range, suggesting that spreading habitat change has pushed them to the range edge, and that high human population pressure was associated with rediscovered species changing habitat from previous records in primary forest, to rediscovery in marginal habitat such as regrowth, cropland and plantations. Both of these effects are likely to make mammals that have declined from habitat loss particularly hard to detect. We found no significant effect of human population density on the probability of rediscovery, although increased frequency of extinction from habitat loss and overkill might be expected in more populated regions. It is possible that this effect was cancelled out because there were also more opportunities for rediscovery in populated areas, because of increased encounter rates and number of people with identification skills.

Across all centuries, range size was strongly correlated with the probability of rediscovery of missing mammals, and species with very small ranges were unlikely to be rediscovered. This might simply be due to the elevated extinction risk associated with small ranges. All recent analyses have concluded that small range size and the closely correlated trait of small population size are the most important indicators of extinction risk and declines of threatened mammals [10,11,36]. The rediscovery rate of species with large ranges might also be higher because scattered remnant populations are more likely to escape detection, an explanation reinforced by the finding that species originally occurring at lower population density were rediscovered at higher rates, despite the fact that low population density predicts extinction risk in mammals generally [10]. This interpretation also seems inconsistent with previous assertions that large geographical range is the best predictor of early species description, because it increases the encounter rate with collectors [37]. However, the high detectability of initially widespread species before decline, and their low detectability after decline, makes sense if they contracted to a very small range that was not at the site where they were last seen but one anywhere within the former wide distribution, or at a remote edge of it [35,38].

Our finding that, throughout historical time, species with small ranges are unlikely to be rediscovered is not an effect of island endemics being extirpated by introduced predators. Unlike birds, which are disproportionately exterminated by predators introduced to islands [39,40], invasive predators have had continental-scale impacts on mammal extinction rates [2,41,42]. Being restricted to islands was not correlated with the probability of rediscovery in our analyses (60% of rediscoveries were on continents). We found that, overall, mammals were unlikely to be rediscovered if the cause of extinction was an introduced predator or disease, but they were likely to be rediscovered if the cause was habitat loss. This conclusion parallels recent findings in birds.

Although more birds are classified as threatened to some degree by habitat loss than by biological invasion, bird families threatened mainly by invasive species are more extinction-prone, and families containing species primarily threatened by habitat loss are less extinction-prone [39].

Moderate search effort was associated with increased rediscoveries, in comparison with low search effort. We could not separate this from the effect of century, because all species missing in the 19th century and before were subject to low search effort (two or fewer expeditions), except for the Talaud flying fox (*Acerodon humilis*), which was missing in 1897 and found alive in 1999 after three searches. Most missing mammals have not been adequately searched for, but a few flagship species (charismatic large mammals) received disproportionately high numbers of searches. The highest search effort in our dataset was confined to a handful of species that remain missing, namely the thylacine (*Thylacinus cynocephalus*), wild horse (*Equus ferus*, extinct in the wild), kouprey (*Bos sauveli*) and Baiji (*Lipotes vexillifer*). We suggest that this is because it is possible to keep searching indefinitely without success if the species is actually extinct. These large-bodied mammals all declined mainly from overkill. However, body size did not independently predict rediscovery rate in any of our models, although persecuted and harvested species are predominantly large and conspicuous [11,13,43]. A species must be identifiable and detectable to be persecuted, exploited or harvested, so publicity about its supposed extinction is also more likely, which might result in more search effort. Our data suggest that mammals purportedly exterminated by overkill receive more attention, because they were targeted by more than twice as many reported searches on average as the more enigmatic species that declined from habitat loss or introduced predator impacts. Mammals that declined from human persecution were more likely to be rediscovered than those presumed to have been driven extinct by introduced species, particularly in the 20th century. Increased public attention and searching probably explain why species that declined in the 20th century tended to be rediscovered more frequently, especially if they declined from overkill.

Our major findings are robust to varying definitions and time scales, because the same conclusions were important whether we used the overall dataset with a broad definition of missing species, or subsections (19th and 20th centuries only, species with one reported threat only, or the restricted dataset of species with a narrower definition of 'missing').

5. CONCLUSIONS

Rediscovery in purportedly extinct and missing mammals is not a random process, but the chance of success depends on search effort, search area, time missing and traits known to be associated with extinction risk such as population density and range size, which interacts with the cause of extinction as predicted by theory. Past effort has focused on a handful of species. Rather than allocating even more effort to these charismatic mammals affected by overkill that are certainly extinct, such as the thylacine, we recommend particularly targeting neglected

species missing later than the 18th century, with relatively large ranges, threatened by habitat loss. It is most likely that some of these species survive, and locating them will enable us to protect their final habitats and avert extinction.

We thank Kate Jones, Jaime Jiminez, David & Meredith Happold, Andrew Cockburn, Hideki Endo, Rainer Hutterer, Carla Kishinami, Stefan Klose, Friederike Spitzenberger, Craig Hilton-Taylor and Richard Fuller for providing or helping us to locate data, and Kerrie Wilson, Peter Baxter, Hugh Possingham, James Brown, Anne Goldizen and anonymous reviewers for discussions and/or comments. This work was supported by funding from the Australian Research Council (ARF DP0773920).

REFERENCES

- Keith, D. A. & Burgman, M. A. 2004 The Lazarus effect: can the dynamics of extinct species lists tell us anything about the status of biodiversity? *Biol. Conserv.* **117**, 41–48. (doi:10.1016/S0006-3207(03)00261-1)
- MacPhee, R. D. E. & Flemming, C. 1999 Requiem aeternam: the last five hundred years of mammalian species extinctions (database: <http://www.creo.amnh.org>). In *Extinctions in near-time* (ed. R. D. E. MacPhee), pp. 333–371. New York, NY: Kluwer Academic/Plenum.
- Diamond, J. M. 1987 Extant unless proven extinct? Or, extinct unless proven extant? *Conserv. Biol.* **1**, 77–79. (doi:10.1111/j.1523-1739.1987.tb00012.x)
- Collar, N. J. 1998 Extinction by assumption; or, the Romeo Error on Cebu. *Oryx* **32**, 239–244. (doi:10.1046/j.1365-3008.1998.d01-51.x)
- Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akcakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J. & Stuart, S. N. 2008 Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.* **22**, 1424–1442. (doi:10.1111/j.1523-1739.2008.01044.x)
- Solow, A. R. 2005 Inferring extinction from a sighting record. *Math. Biosci.* **195**, 47–55. (doi:10.1016/j.mbs.2005.02.001)
- Roberts, D. L., Elphick, C. S. & Reed, J. M. 2010 Identifying anomalous reports of putatively extinct species and why it matters. *Conserv. Biol.* **24**, 189–196. (doi:10.1111/j.1523-1739.2009.01292.x)
- Scott, J. M., Ramsey, F. L., Lammertink, M., Rosenberg, K. V., Rohrbaugh, R., Wiens, J. A. & Reed, J. M. 2008 When is an 'extinct' species really extinct? Gauging the search efforts for Hawaiian forest birds and the Ivory-billed Woodpecker. *Avian Conserv. Ecol.* **3**, 3.
- Cardillo, M., Mace, G. M., Gittleman, J. L., Jones, K. E., Bielby, J. & Purvis, A. 2008 The predictability of extinction: biological and external correlates of decline in mammals. *Proc. R. Soc. B* **275**, 1441–1448. (doi:10.1098/rspb.2008.0179)
- Davidson, A. D., Hamilton, M. J., Boyer, A. G., Brown, J. H. & Ceballos, G. 2009 Multiple ecological pathways to extinction in mammals. *Proc. Natl Acad. Sci. USA* **106**, 10 702–10 705. (doi:10.1073/pnas.0901956106)
- Fisher, D. O. & Owens, I. P. F. 2004 The comparative method in conservation biology. *Trends Ecol. Evol.* **19**, 391–398. (doi:10.1016/j.tree.2004.05.004)
- Hockey, P. A. R. & Curtis, O. E. 2009 Use of basic biological information for rapid prediction of the response of species to habitat loss. *Conserv. Biol.* **23**, 64–71. (doi:10.1111/j.1523-1739.2008.01028.x)
- Owens, I. P. F. & Bennett, P. M. 2000 Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proc. Natl Acad. Sci. USA* **97**, 12 144–12 148. (doi:10.1073/pnas.200223397)
- Purvis, A., Gittleman, J. L., Cowlishaw, G. & Mace, G. M. 2000 Predicting extinction risk in declining species. *Proc. R. Soc. Lond. B* **267**, 1947–1952. (doi:10.1098/rspb.2000.1234)
- Finlayson, H. H. 1932 Rediscovery of *Caloprymnus campestris* (Marsupialia). *Nature* **129**, 871–871. (doi:10.1038/129871b0)
- Carr, S. G. & Robinson, A. C. 1997 The present status and distribution of the desert rat-kangaroo *Caloprymnus campestris* (Marsupialia: Potoroidae). *S. A. Naturalist* **72**, 4–27.
- IUCN. 2009 *IUCN red list of threatened species*. Version 2009.2. See <http://www.iucnredlist.org>.
- Jones, K. E. *et al.* 2009 PanTHERIA: a species-level database of life-history, ecology and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648. (doi:10.1890/08-1494.1)
- Schultz, M. 2002 National recovery plan for the Christmas Island shrew, *Crocidura attenuata trichura*. Canberra, Australia: Department of the Environment and Heritage.
- Gordon, G. 2008 Desert bandicoot. In *The Australian Museum complete book of Australian mammals* (eds S. V. Dyck & R. Strahan), pp. 185–186. Sydney, Australia: Reed Books.
- Allen, G. M. 1942 *Extinct and vanishing mammals of the Western Hemisphere with the marine species of all the oceans*. New York, NY: American Committee for International Wild Life Protection.
- Clavero, M. & Garcia-Berthou, E. 2005 Invasive species are a leading cause of animal extinctions. *Trends Ecol. Evol.* **20**, 110–110. (doi:10.1016/j.tree.2005.01.003)
- Hosmer, D. W. J. & Lemeshow, S. 2000 *Applied survival analysis: regression modeling of time to event data*. Malden, MA: Wiley-Interscience.
- Nakagawa, S. & Freckleton, R. P. 2008 Missing in action: the dangers of ignoring missing data. *Trends Ecol. Evol.* **23**, 592–596. (doi:10.1016/j.tree.2008.06.014)
- Little, R. J. A. & Rubin, D. B. 2002 *Statistical analysis with missing data*. Hoboken, NJ: John Wiley and Sons Inc.
- Fisher, D. O., Blomberg, S. P. & Owens, I. P. F. 2003 Extrinsic versus intrinsic factors in the decline and extinction of Australian marsupials. *Proc. R. Soc. Lond. B* **270**, 1801–1808. (doi:10.1098/rspb.2003.2447)
- R Development Core Team. 2009 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.R-project.org>.
- van Buuren, S. & Oudshoorn, C. G. M. 2009 MICE: multivariate imputation by chained equations. See <http://www.stefvanbuuren.nl/publications/MICEinR-Draft.pdf>.
- Rubin, D. B. 1987 *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley and Sons.
- Lumley, T. 2009 *mitools: tools for multiple imputation of missing data*. Vienna, Austria: R Foundation for Statistical Computing.
- Grambsch, P. & Therneau, T. 1994 Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 515–526. (doi:10.1093/biomet/81.3.515)
- Fox, J. *et al.* 2009 *Car: companion to applied regression. R package version 1.2-16*. Vienna, Austria: R project. See <http://www.CRAN.R-project.org/package=car>.
- Menard, S. 1995 *Applied logistic regression analysis*. Thousand Oaks, CA: Sage Publications.
- Kinzig, A. P. & Harte, J. 2000 Implications of endemics-area relationships for estimates of species extinctions. *Ecology* **81**, 3305–3311.

- 35 Fisher, D. O. In press. Trajectories from extinction: where are missing mammals rediscovered? *Glob. Ecol. Biogeogr.*
- 36 Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R. P., Sechrest, W., Orme, C. D. L. & Purvis, A. 2005 Multiple causes of high extinction risk in large mammal species. *Science* **309**, 1239–1241. (doi:10.1126/science.1116030)
- 37 Collen, B., Purvis, A. & Gittleman, J. L. 2004 Biological correlates of description date in carnivores and primates. *Glob. Ecol. Biogeogr.* **13**, 459–467. (doi:10.1111/j.1466-822X.2004.00121.x)
- 38 Channell, R. & Lomolino, M. V. 2000 Dynamic biogeography and conservation of endangered species. *Nature* **403**, 84–86. (doi:10.1038/47487)
- 39 Clavero, M., Brotons, L., Pons, P. & Sol, D. 2009 Prominent role of invasive species in avian biodiversity loss. *Biol. Conserv.* **142**, 2043–2049. (doi:10.1016/j.biocon.2009.03.034)
- 40 Duncan, R. P. & Blackburn, T. M. 2007 Causes of extinction in island birds. *Anim. Conserv.* **10**, 149–150. (doi:10.1111/j.1469-1795.2007.00110.x)
- 41 Johnson, C. N. 2006 *Australia's mammal extinctions: a 50 000 year history*. Cambridge, UK: Cambridge University Press.
- 42 Johnson, C. N., Isaac, J. L. & Fisher, D. O. 2007 Rarity of a top predator triggers continent-wide collapse of mammal prey: dingoes and marsupials in Australia. *Proc. R. Soc. B* **274**, 341–346. (doi:10.1098/rspb.2006.3711)
- 43 Isaac, N. J. B. & Cowlshaw, G. 2004 How species respond to multiple extinction threats. *Proc. R. Soc. Lond. B* **271**, 1135–1141. (doi:10.1098/rspb.2004.2724)