

Extinct or still out there? Disentangling influences on extinction and rediscovery helps to clarify the fate of species on the edge

Running Header: Disentangling influences on extinction and rediscovery

T. E. Lee

Mathematical Institute, Andrew Wiles Building, University of Oxford, UK

Tamsin.lee@maths.ox.ac.uk

D.O. Fisher

School of Biological Sciences, University of Queensland, Queensland, 4072, Australia

d.fisher1@uq.edu.au

S. P. Blomberg

School of Biological Sciences, University of Queensland, Queensland, 4072, Australia

s.blomberg1@uq.edu.au

B. A. Wintle

School of Biosciences, University of Melbourne, VIC, 3010, Australia.

b.wintle@unimelb.edu.au

Keywords, Lazarus, detectability, cure rate model, search effort, Mammalia

Abstract

Each year, two or three species that had been considered to be extinct are rediscovered. Uncertainty about whether or not a species is extinct is common, because rare and highly threatened species are difficult to detect. Biological traits such as body size and range size are expected to be associated with extinction. However, these traits, together with the intensity of search effort, might influence the probability of detection and extinction differently. This makes statistical analysis of extinction and rediscovery challenging.

Here we use a variant of survival analysis known as cure rate modelling to differentiate factors that influence rediscovery from those that influence extinction. We analyse a global dataset of 99 mammals that have been categorised as extinct or possibly extinct. We estimate the probability that each of these mammals is still extant, and thus estimate the

proportion of missing (presumed extinct) mammals that are incorrectly assigned extinction.

We find that body mass and population density are predictors of extinction, and body mass and search effort predict rediscovery. In mammals, extinction rate increases with body mass and population density, and these traits act synergistically to greatly elevate extinction rate in large species that also occurred in formerly dense populations. However, when they remain extant, larger-bodied missing species are rediscovered sooner than smaller species. Greater search effort increases the probability of rediscovery in larger species of missing mammals, but has a minimal effect on small species, which take longer to be rediscovered, if extant.

By separating the effects of species characteristics on extinction and detection, and using models with the assumption that a proportion of missing species will never be rediscovered, our new approach provides estimates of extinction probability in species with few observation records and scant ecological information.

Introduction

Incorrectly classifying a species as extinct can lead to failure to conserve a threatened species (Chades et al., 2008), and public confidence in conservation listings can erode when a species reappears after being classified as extinct (Keith & Burgman, 2004). Extinction rates are used to assess the effectiveness of biodiversity conservation efforts (May et al., 1995; Flannery & Schouten, 2001; IUCN 2015). The high turnover in species classified as extinct highlights the difficulty of detecting real changes in the status of biodiversity using the IUCN Red List (Burgman, 2002; Possingham et al., 2002). Before 1995, the Red List criterion was 'not definitely recorded for 50 years'. It now classifies species as extinct when there is 'no reasonable doubt that the last individual has died', based on appropriate case-by-case surveys. Scheffers et al., (2011) concluded that at least 351 species have been rediscovered in the past 122 years. Since the 16th century around a third of mammals suspected to be extinct have been rediscovered (Fisher & Blomberg, 2011).

Fisher and Blomberg (2011) found that low search effort and large species range size were associated with low rediscovery rate in mammals. Search effort is generally low for small mammals. For example, the Australian central rock rat *Zyzomys pedunculatus* was feared extinct between 1960 and 1996, but it was probably difficult to detect because it is small-bodied, inconspicuous, and had a formerly large range in remote arid Australia.

Species traits might differ in their influence on the probability of extinction and detectability. For example, large body size is strongly associated with extinction risk in mammals (Cardillo et al., 2005; Davidson et al., 2009; Di Marco et al., 2015) because large species are targeted by hunters, and have low reproductive rates, and therefore low capacity to compensate for over-harvesting and other threats (Fisher & Owens, 2004).

Fisher & Blomberg (2011) proposed that large body size may increase extinction rate, but also increase the chance that species are detected, because large mammals are easier to find and recognise. This might mean that the effects of large body size on extinction and detection cancel one another out, so that body size is not associated with mammal rediscovery rate overall.

Low local population density may promote extinction because it reflects small population size (Johnson, 1998), and because mammals with smaller geographic ranges (and therefore higher extinction risk) are likely to have low population densities in relation to body mass, and more specialised habitat preferences (Johnson, 1998; Slatyer et al., 2013). Species with restricted distributions and narrow niches are vulnerable to threats such as habitat loss that may affect their entire range, and some threats such as invasive predators on small islands are particularly dangerous to endemic species (Davidson et al., 2009). Species in sparse populations ought to be more difficult to detect. However, missing mammals with formerly sparse populations were more likely to be rediscovered than those in dense populations (Fisher & Blomberg, 2011). This may be because extinction is easier to confirm in species that were formerly easy to find due to their high local densities. Often, these were species on islands, and have been eliminated by invasive predators or disease. Higher search effort increases the chance that elusive species are detected, and prolonged high and unsuccessful search effort indicates likely extinction (Fisher & Blomberg, 2011).

Mathematical models have been developed to assess the probability of species extinction based on dated sighting records (Solow, 1993a; Solow & Roberts, 2003; Rout et al., 2009; Solow et al., 2011; Thompson et al., 2013; Lee et al., 2014). By assessing the frequency of sightings when the species is known to exist, the significance of an extended period of no sightings can be quantified. However, often sighting records are often not available. For example, 45 of 116 extinct and potentially extinct mammals studied by Fisher and Blomberg (2012) are known only from a single dated specimen, and 82 have insufficient data to use Solow's and Roberts' methods. For most of these species, only broad categories of threats and biological traits such as body mass and approximate former range size are known.

Cox proportional hazards models have been used to estimate rediscovery times for missing mammals, to find traits influencing rediscovery, and to predict which species are most likely to be rediscovered (Fisher & Blomberg, 2011; 2012). The Cox proportional hazards model is a survival model that is also used commonly in medicine and manufacturing (Cox, 1972). It models the time until an event occurs, such as death of a patient or failure of a

component, as a function of particular traits of the participants. For example, in a longitudinal study of correlates of a disease, traits such as weight, smoking status or age may vary with the participants' time until death from the disease, and this may support or refute hypotheses about causes. In Fisher's and Blomberg's (2011; 2012) analyses, the participant is a species and the event is rediscovery. At any time, given that the species has remained undetected to that point, the species may be detected (rediscovered). This potential is called the 'hazard' function for the species, and estimates the time to rediscovery. An assumption of this survival model is that the event (rediscovery) will eventually occur for all participants. However, in the case of modelling the rediscovery of species, this assumption does not hold because some species are truly extinct.

The cure rate model (Boag, 1949; Berkson & Gage, 1952) is a form of survival model that allows for some participants not to experience the event of 'death'. The model estimates the proportion of participants that are 'cured' of the threat: this proportion is the 'cure rate'. For example, when testing a new medicine, some patients will not incur death from the disease – and perhaps traits such as weight and age affect the chances of being cured. In our context, this is the extinction rate; the proportion of species that are truly extinct and will never be rediscovered. The cure rate model also estimates a maximum time by which all subjects that will die from the disease, have died. In the example above, the model would estimate the maximum time of death from the disease for uncured patients with different weight, smoking status and age. In our model of missing and extinct mammals, the model estimates the maximum time by which a species with particular traits will be rediscovered, if it is going to be rediscovered at all (that is, it has not gone extinct). Using cure rate models and data on rediscovered species, we aim to clarify which traits and processes influence extinction, which influence rediscovery time, and which influence both. We use the cure rate model to predict which species currently classified as 'Extinct' or 'Critically Endangered' are likely to be extant, but undetected.

Materials and Methods

In the literature, there are two main mixture cure models (Cai et al., 2012). One assumes proportional hazards, as in the Cox Proportional hazards model. This means that if species

A is twice the size of species B (and small species take longer to rediscover), species B would take twice as long to rediscover than species A. An alternative cure rate model, the accelerated failure time model, relaxes this assumption. Covariate effects on failure time are modelled directly, and interpretation of covariate effects may be simpler than in the proportional hazards model (Zhang & Peng, 2012). To allow for the possibility that hazards are not proportional, we use the accelerated failure time mixture cure model.

Cure rate models

We use the cure rate approach (Boag, 1949; Berkson & Gage, 1952) to model the influence of species traits, threats, and survey effort on the processes of extinction and rediscovery. Rediscovery is directly observed in the data, whereas extinction is a latent, unobservable process. Observations of rediscovery (or not) at a given time are assumed to be drawn from a mixture of two survival models; extant species that are either rediscovered or not, and an extinct species that can never be rediscovered. In our mixture cure model, the proportion of extinct species and the non-detection distribution of extant species are modelled separately against the covariates.

Assume that t is the non-detection time, a random variable. First we denote the proportion of truly extinct species by $E(z)$, which depends upon observed values of covariate vector $z=(z_1, z_2, \dots, z_i, \dots, z_N)$ for N species. Second, we consider the case where extinction does not occur ($1-E(z)$), and rediscovery happens (or will happen). For this fraction, we denote the non-detection function as $U(t|x)$, which depends upon observed values of covariate vector x for N species. The other fraction, the truly extinct species, does not have a non-detection function. Therefore, the non-detection function is a mixture of both of these populations

$$U(t|x, z) = E(z) + (1 - E(z))U(t|x) \quad (1)$$

A logistic regression formulation is usually assumed for $E(z)$ so that

$$E(z) = \frac{e^{bz}}{1+e^{bz}} \quad (2)$$

where the vector of unknown parameters b is used to model the effects of z .

The non-detection function for the extant population $U(t|x)$ depends on the hazard function $\lambda(t)$. The hazard function defines the potential for the species to be detected at a particular time, given that it has remained undetected until that time: $\lambda = \lambda(t_i | \text{extant}, x_i)$ where t_i is the non-detection time for a species with covariate vector x_i . We do not assume proportional hazards, thus $U(t|x) = U_0(te^{\lambda t})$ (Zhang & Peng 2007), where U_0 is the baseline non-detection function of extant species when $x=0$. In the case where proportional hazards are assumed, then $U(t|x) = U_0(t)e^{\lambda t}$. This proportional hazards model would correspond to equation (1) in Garrad et al., (2008) who used it to model time to first detection of rare plant species in biological surveys.

Parametric approaches exist for mixture cure models (Farewell, 1982; Peng et al., 1998). However, the parametric assumptions are limiting and difficult to verify, so we use semiparametric methods (Li & Taylor, 2002; Zhang & Peng, 2007; Sy & Taylor, 2000). We used the `smcure` package in R (Cai et al., 2012) to estimate the cure rate model. We test the performance of the cure rate model against simulated extinction and rediscovery data where species extinction risk varies as a function of traits (Appendix A).

Applying the cure rate model to rediscovery data

We fit the cure rate model to an updated version of the dataset in Fisher and Blomberg (2011) to infer the influence of traits on extinction and rediscovery probabilities, and to estimate the probability that listed extinct and long unseen endangered animals are truly extinct. The few mammals that have been missing before the 17th century were excluded, leaving 167 species, 68 of which have been rediscovered. The data set is available as Supporting Information online.

We obtained the probability of being extinct and the maximum time to rediscovery for each species in the dataset, using the non-detection function. An example showing how we derived these parameters from the non-detection function is provided in Appendix B. Rediscovery may occur before the maximum time, but if the species is not detected after a maximum time we classify it as extinct. We also provide the probability that the mammal is extinct in 2014.

199
200 We considered covariates that were found to be important in the models of rediscovery and
201 extinction of mammals in Fisher & Blomberg (2011; 2012) (Table 1). Previous research also
202 indicates that these traits are likely to have a causal link with either extinction or detection
203 probability, or both, and are not highly correlated with one another. Body size is one of the
204 most important predictors of indices of extinction risk in mammals (Cardillo et al., 2005;
205 Fritz et al., 2009) and increase in extinction risk over time (Di Marco et al., 2015).
206 Population density reflects local population size, which is one of the most important
207 predictors of extinction in all organisms (Johnson., 1998; Gaston et al., 1997), and is used in
208 IUCN Red List criteria to assign species to categories of extinction risk (IUCN, 2015).
209 Previous studies have identified other variables associated with extinction risk in
210 mammals. For example, Arregoitia et al. (2016) reviewed ~ 60 traits reported in the
211 literature. In preliminary exploratory analyses, we tested a range of variables (Appendix C)
212 that influence mammal extinction rates, including geographic range, mean human
213 population density in the range and major threats. Larger models did not converge to
214 produce stable estimates of regression coefficients. A limitation of our approach at present
215 is the restriction on the number of covariates and interactions that can be included in the
216 cure rate model. However, body mass is strongly correlated with other life history variables
217 such as reproductive rate (Promislow & Harvey, 1990), and geographic range (Johnson,
218 1998; Gaston et al., 1997). We omitted some additional variables that may be linked to
219 detectability, such as technological change and accessibility, due to the lack of reliable
220 information on a global scale. Acknowledging these limitations, we test how population
221 density, body mass and search effort (Table 1) are associated with rediscovery through
222 their influences on both detection and extinction . We used an ad-hoc backward selection
223 approach to test the effects of these covariates.

224
225 We dealt with missing covariate data using multiple imputation (Rubin, 1987). There were
226 more missing data for presumed extinct species than for rediscovered species. To minimise
227 bias due to deletion of species with missing data, we imputed five replicate datasets using
228 multiple imputation (R package mi, Su et al., 2011). We ran the cure rate model on all five
229 versions of the covariate data, and report the average of the five extinction probabilities

and five maximum rediscovery times. Additionally, we calculated the extinction probability using leave one out analysis. We compare the maximum predicted rediscovery time with the actual rediscovery time for the 68 rediscovered mammals.

Results

When comparing the leave one out analysis with results from using the full data set, we found the results are very similar, with very little variation (mean absolute difference of 0.022 with standard deviation 0.034).

Extinction probabilities and effects of density and body mass

Body mass, population density, and the interaction between population density and body mass were associated with extinction risk in mammals. The probability of extinction increased as body mass increased (Fig. 1a). Missing species with a body mass of ~10g had a median probability of being extinct of ten percent (0.101), but missing species weighing approximately 100kg had a median probability of being extinct of nearly 80% (0.780). The large variance indicates that other factors are also important, especially for the smaller species (Fig. 1a). The two largest species in the dataset are the extinct kouprey *Bos sauveli* (~800kg) and the extinct Steller's sea cow *Hydrodamalis gigas* (~10,000kg), and there are no species with a mass intermediate between these two species. This may explain the highly variable results for the largest species (Fig 1a).

In agreement with Fisher and Blomberg (2011, 2012), we found that presumed-extinct species in the highest category of former population density were eight times as likely to be extinct than the sparsest species (0.723 median probability of being extinct versus 0.090 for the lowest population density category). The small interquartile range of species in the highest density category indicates that these missing species have a high probability of being truly extinct (Fig. 1b). The interaction of population density and body mass shows that there is an elevated risk of extinction in species that are both large and occurred in dense populations (Fig. 2a). Large interquartile ranges (Fig. 1b) are due to the interaction of body mass with population density. For example, missing mammals that occurred in sparse populations and weigh ~10g have a 1% (0.01) probability of being extinct (e.g. Wimmer's

shrew, *Crocidura wimmeri*) while species of the same body mass that occurred in dense populations have a 70% chance of being extinct (e.g. Angel island mouse, *Peromyscus guardia*). Sparsely distributed missing species had a variable probability of being extinct that ranged between 0.01 and 0.88, whereas species that had occurred in dense populations had a narrower range of probabilities of being extinct, 0.70 to 0.82 (Fig. 2a).

Rediscovery times and effects of body mass and search effort

Body size and search effort were both associated with the chance of rediscovery. Larger missing species were more likely to be rediscovered sooner, with little variation (Fig. 1c). Species that weighed ~10g (e.g. island-shrews *Nesophontes* spp and bats) had a median maximum rediscovery time of 292 years, regardless of search effort (Fig. 2d). Increasing search effort had a varied effect: increasing effort from low to high reduced the median maximum time of rediscovery from 194 to 39 years (Fig. 1d), but the interquartile range around the median of 39 years ranges from 28 to 292 years.

The maximum time to rediscovery was less in species that were both large and subject to high search effort: search effort was more beneficial for larger mammals (Fig. 2b). Search effort had no effect on rediscovery time of very small mammals, but high search effort halved the mean rediscovery time of mammals ~10kg, and the difference was even greater for mammals ~100kg. For example, the Angel Island mouse, *Peromyscus guardia* received a high search effort and the Somali golden mole, *Calcochloris tytonis* received low search effort, yet they are both predicted to be rediscovered (if at all) within 292 years. Two large species, the wild horse, *Equus ferus* and red gazelle, *Eudorcas rufina* are also of similar size. The horse received high search effort, and had a maximum time to rediscovery of 28 years, whereas the gazelle received low search effort and had a maximum time to rediscovery more than double that of the horse, 73 years.

The expected time of rediscovery is related to the maximum time to rediscovery by

$$\text{Rediscovery_time} = 20.8 + 0.19\text{Max_time} \quad (3)$$

(Fig. 3). However, the actual time to rediscovery is highly variable, so this equation cannot be relied upon to provide precise and accurate predictions of rediscovery time. Forty-three percent of rediscovered mammals were found within a quarter of their maximum

rediscovery time; and after half of their maximum rediscovery time had passed, 79% of rediscovered mammals had been rediscovered. We compared the maximum predicted rediscovery times with the times that species presumed extinct were missing, and found that ~14% of mammals have been missing for longer than the corresponding maximum predicted rediscovery time. We infer that these mammals are highly likely to be extinct (Table 3).

Which species are extinct?

Of 99 missing mammals, the cure rate model indicated that 44 are over 70% likely to be extinct, 36 are over 80% likely to be extinct, 24 are over 90% likely to be extinct, and 14 of the 99 have remained undetected for longer than their maximum rediscovery times and are nearly certain to be extinct based on this model (Fig. 4). This includes the only two mammals in our data set that were last seen in the 18th century (Steller's sea cow *Hydrodamalis gigas* last seen 1768 and the bluebuck *Hippotragus leucophaeus* last seen 1799), both large mammals. The IUCN Red List also states that these two mammals are extinct (Table 2). Of the 23 mammals last seen in the 19th century (missing for >113 years), only five were classified as certainly extinct by the cure rate model (Table 3): the Falkland Islands wolf *Dusicyon australis*, the Arabian gazelle *Gazella arabica*, the sea mink (from the eastern USA and Canada) *Neovison macrodon*, the Oriente cave rat *Boromys offella*, and the Aru Island flying fox (from Indonesia) *Pteropus aruensis*.

Australia, The Caribbean Islands and Mexico have the world's highest rates of mammal extinction. Most Australian missing species are listed as unequivocally extinct by the IUCN, but were not classed as certain to be extinct based on maximum rediscovery time according to the cure rate model. The one exception is the thylacine *Thylacinus cynocephalus*. This species is classed as extinct with high certainty based on its large body size and high search effort, giving a maximum time to rediscovery that expired in 1983. Of the 40 missing Caribbean Island and Mexican species, most are considered unequivocally extinct by the IUCN, but only two are classed as definitely extinct based the cure rate model prediction of maximum rediscovery times: the Caribbean monk seal *Monachus tropicalis* and the Oriente cave rat *Boromys offella*. The maximum rediscovery time will expire for a further 11

mammals before 2050, including three more Australian species *Lagorchestes leporides* (eastern hare-wallaby), *Macropus greyi* (toolache wallaby), and *Potorous platyops* (broad-faced potoroo), and one Caribbean species *Megalomys desmarestii* (pilorie). Eight of these eleven species with very high probabilities of extinction, according to our cure rate model, are classed as extinct by the IUCN, and the other three as critically endangered (Table 2). Two Critically Endangered species (the south east Asian kouprey *Bos sauveli* and Indonesian Aru flying fox *Pteropus aruensis*) are included amongst the 14 which we predict are definitely extinct, two others have a very high chance of being extinct within 50 years since last reporting (the wild horse *Equus ferus* and baiji *Lipotes vexillifer*, both in China. Two further CR species in Melanesia have a very high probability of being extinct (>75%) (Wondiwoi tree kangaroo *Dendrolagus mayri* and Emperor rat *Uromys imperator*) (Table 2).

Discussion

Birdlife International uses the category 'Possibly Extinct' (Birdlife.org 2015), and the IUCN flags some 'Critically Endangered' species as 'Possibly Extinct' in the text of species accounts. Our method could help to assign mammals to a formal 'Possibly Extinct' category. The 'Possibly Extinct' category could include species that are difficult to detect, based on a quantitative threshold. For example, this might include small mammals that can remain undetected for over 50 years although they have a <10% chance of being extinct under our model (Table 3). In an opinion article, Ladle et al. (2011) proposed that rather than being tagged 'possibly extinct', missing CR and DD species could be categorized according to the time since the last record (e.g. in 25 year blocks) to reduce misinterpretation of rediscoveries. Our method provides a time frame associated with the probability that each species is extinct. This diminishing probability could also potentially be updated during IUCN Red List revisions until either rediscovery, or the probability of rediscovery reaches zero. Additionally, the model could enhance decisions about search effort. For example, increasing the search effort from low to medium for a species with an approximate body mass of 10kg reduced the period of uncertainty from 103 to 71 years. However, increasing the search effort in similar manner for a small species makes little difference. This indicates that investment in search effort for relatively large mammals is likely to be more efficient. Criterion E of the IUCN Red List recommends that a species should move from 'Endangered'

to 'Critically Endangered' when quantitative analysis shows that the probability of extinction in the wild is at least 50% within 10 years or three generations of the last sighting, whichever is longer, up to a maximum of 100 years (IUCN 2015). Our model is one way to do such an analysis, based on a combination of species traits. Assessors could also potentially use this method to modify the maximum threshold of 100 years to be specific to the traits of a particular species. Once the maximum time to rediscovery has expired, the species could be moved from 'Critically Endangered' to 'Extinct'.

Our new approach using cure rate models to clarify which traits influence extinction, and which influence rediscovery times, found that of the 99 mammals that are missing, and feared extinct in our dataset, 44 currently have a less than 50% chance of being truly extinct, based on their body mass and density in a cure rate model. This indicates a relatively high chance of persistence in 45% of missing mammals in the data set. This is consistent with the observation of Fisher & Blomberg (2011), that 36% of mammal species ever considered extinct have been rediscovered. Based on an independent dataset, Scheffers et al. (2011) found that 42% of mammals that have been declared extinct have been rediscovered.

We found that the probability of rediscovery of missing mammal species is most strongly correlated with small body size. In larger mammals, higher search effort also increased the probability of rediscovery. Based on our model, we expect that small-bodied missing species last seen in the past 180 years would be most likely to be rediscovered. Examples include the 60g Puebla deer mouse *Peromyscus mekisturus* from Mexico, and 4g Santa Cruz tube-nosed bat *Nyctimene santacrucis*, which we estimate have a 8.2% and 6.5% chance of being extinct despite not having been seen since 1948 and 1907, respectively (Table 3). Large body size is a key biological trait associated with indices of extinction threat in mammals (e.g. Red List status) (Cardillo et al., 2005; Davidson et al., 2009; Fisher & Owens, 2004; Fritz et al., 2009; Fritz & Purvis, 2010; Gonzalez-Suarez & Revilla, 2012; Arregoitia et al., 2013; Di Marco et al., 2015). Large mammals are vulnerable to overharvesting, are targeted by hunters, and require large areas of habitat. They are also more likely to be threatened because large size is correlated with slow life histories and low capacity to

recover from declines (Cardillo et al., 2005; Fisher & Owens 2004, Fritz et al., 2009). By separating extinction from rediscovery probability, our results confirm that larger species are not only currently more threatened, but have also been more prone to becoming extinct in the past (Fritz et al., 2009; Di Marco et al., 2015).

In contrast, Fisher and Blomberg (2011) found no association between mammal body mass and probability of rediscovery. They suggested that this was because large body size increases both extinction rate and detectability. Larger-bodied mammals are more conspicuous, more likely to be identifiable by non-experts, and easier to observe without specialist trapping equipment. Small body size decreases detectability but also decreases extinction rate, so there may be no net effect of body size on rediscovery rate when effects of extinction and detection are not separated by the modelling approach, as in their (non cure-rate) Cox proportional hazards model. Our approach separates effects of extinction and detection on rediscovery and confirms that missing but extant species are more likely to be rediscovered if they are large. This implies that the longer a large species remains missing, the more likely it is to be extinct, both because large mammals are extinction-prone, and because if extant, it would be more likely to have been detected.

Our finding that missing species that once existed in dense populations were more likely to be truly extinct agrees with Fisher's and Blomberg's (2011) results. This finding is counter-intuitive, given that a central conclusion of macroecology is that high local abundance (i.e. population density) is associated with large geographic range size and species persistence, because species with flexible habitat requirements and large populations are most likely to avoid extinction (Gaston et al., 1997; Slayter et al., 2013). However, many of these locally abundant missing species in our dataset occurred in small ranges. We found that missing species with both formerly high local population density and large body size were especially likely to be extinct. Such species included social ungulates such as the kouprey (*Bos sauveli*) and gazelles, and aquatic mammals that concentrated at particular sites (a single river or limited island group) such as the baiji (*Lipotes vexillifer*, the Yangtze River dolphin), Caribbean Monk Seal *Monachus tropicalis*, and Steller's sea cow *Hydrodamalis gigas*. Concentrations of large mammals are attractive to hunters, and these species were

affected by overkill. Higher detectability might also explain why large-bodied missing species that had denser populations are more likely to be truly extinct. Fisher and Blomberg (2011) argued that we are less likely to miss abundant species in small ranges if they are in fact extant, and this also applies to large species.

Our models suggest that high search effort for large-bodied mammals can reduce the maximum time to rediscovery considerably. However, greater search effort did not improve the chance of rediscovery of small mammals. This may be because it is rare that small mammals received the high levels of effort that more charismatic large missing species such as the thylacine and baiji did. Such species are often termed ‘flagship species’ for conservation (Verissimo et al., 2009), especially mammals (Leader-Williams & Dublin, 2000). Large species in Table 2 with relatively high probabilities of being extant are likely to be, or become flagship species. However, threatened small mammals are disadvantaged by lack of conservation attention. Conservationists and researchers spend less money and effort seeking out remnant, remote or refuge populations of smaller mammals (Fisher 2011). Trapping cryptic species requires specialist equipment and extra effort. Compounding the issue of scant conservation resources, small mammals often cannot be identified unless they are trapped by expert taxonomists or sent as specimens to museums (Verde Arregoitia 2015).

Because of model convergence failures due to low power, we have not been able to include some variables that were important in previous analyses, such as threat type (Appendix C). Nonetheless, we suggest that future work on this novel approach can overcome this problem. For example, model selection could be established by calculating an AIC value based on optimising the estimating equation for the AFT model.

Many quantitative methods that estimate extinction probabilities are based on the rate of sightings (reviewed by Boakes et al., 2015), and require multiple reliably dated records as the species approaches extinction. Our alternative approach is suitable for species with very few dated records, and can be used for any species with known body mass, and known search effort and past population density (local population size) in broad categories. This

method can provide a basis for prioritising efforts to find populations of species on the edge of extinction.

References

- [1] Arregoitia, L. D. V., Blomberg, S. P., & Fisher, D. O. (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**(1765), 20131092.
- [2] Berkson, J. & Gage, R. P. (1952) Survival curve for cancer patients following Treatment. *J. Am. Statist. Assoc.*, **47**, 501–515.
- [3] Boag, J. W. (1949) Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **11**, 15–53.
- [4] Burgman, M. A. (2002) Turner Review No. 5. Are listed threatened plant species actually at risk? *Aust. J. Bot.* **50**, 1–13.
- [5] Cai, C., Zou, Y., Peng, Y. & J. Zhang (2012) smcure: An R-package for estimating semiparametric mixture cure models. *Comput. Meth. Prog. Bio.* **108**, 1255–1260.
- [6] Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R., Sechrest, W., Orme, C. D., & Purvis, A. (2005). Multiple causes of high extinction risk in large mammal species. *Science*, 309(5738), 1239-1241.
- [7] Chades, I., McDonald-Madden, E., McCarthy, M., Wintle, B., & Possingham, H. (2008). Save, survey or surrender: optimal management of a cryptic threatened species. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13936–13940.
- [8] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 187–220.
- [9] Davidson, A. D., Hamilton, M. J., Boyer, A. G., Brown, J. H., & Ceballos, G. (2009). Multiple ecological pathways to extinction in mammals. *Proceedings of the National Academy of Sciences*, **106**(26), 10702-10705.
- [10] Di Marco, M., Collen, B., Rondinini, C., & Mace, G. M. (2015). Historical drivers of extinction risk: using past evidence to direct future monitoring. *Proc. R. Soc. B* **282**(1813) 20150928.

- [11] Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* 1041–1046.
- [12] Fisher, D. O. (2011) Cost, effort and outcome in mammal rediscovery: neglect of small species. *Biol. Cons.* **144**, 1712–1718.
- [13] Fisher, D. O. & Blomberg, S. P. (2011) Correlates of rediscovery and the detectability of extinction in mammals. *Proc. R. Soc. B.* **278**, 1090–1097.
- [14] Fisher, D. O. & Blomberg, S. P. (2012) Inferring extinction of mammals from sighting records, threats, and biological traits. *Conserv. Biol.* **26**, 57–67.
- [15] Fisher, D. O., & Owens, I. P. (2004). The comparative method in conservation biology. *Trends in Ecol& Evolut*, **19(7)**, 391-398.
- [16] Flannery, T. & Schouten, P. (2001). *A Gap in Nature: Discovering the World's Extinct Animals*. Text Publishing, Melbourne.
- [17] Fritz, S. A., Bininda-Emonds, O. R., & Purvis, A. (2009). Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters*, **12(6)**, 538-549.
- [18] Fritz, S. A., & Purvis, A. (2010). Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, **24(4)**, 1042-1051.
- [19] Garrad, G. E., Bekessy, S. A., McCarthy, M. A. & Wintle, B. A. (2008) When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral. Ecol.* **33**, 986–998.
- [20] Gaston, K. J., Blackburn, T. M., & Lawton, J. H. (1997). Interspecific abundance-range size relationships: an appraisal of mechanisms. *Journal of Animal Ecology*, 579-601.
- [21] González-Suárez, M., Lucas, P. M., & Revilla, E. (2012). Biases in comparative analyses of extinction risk: mind the gap. *Journal of Animal Ecology*, **81(6)**, 1211-1222.
- [22] IUCN (2015). *The IUCN Red List of threatened species*. (<http://www.redlist.org>).
- [23] Johnson, C. N. (1998). Species extinction and the relationship between distribution and abundance. *Nature*, **394(6690)**, 272-274.
- [24] 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. Cons.* **117**, 41–48.

- [25] Ladle, R., Jepson, P., Malhado, A., Jennings, S., & Barua, M. (2011). Perspective: The causes and biogeographical significance of species' rediscovery. *Frontiers of Biogeography*, **3**(3).
- [26] Li, C., Taylor, J. (2002). A semi-parametric accelerated failure time cure model. *Statistics in Medicine* **21**, 3235–3247.
- [27] Leader-Williams, N., & H. T. Dublin. (2000). Charismatic megafauna as 'flagship species'. Pages 53–81 in A. Entwistle, and N. Dunstone, editors. Priorities for the Conservation of Mammalian Diversity: Has the Panda Had Its Day. *Cambridge University Press*.
- [28] Lee, T. E., McCarthy, M. A., Wintle, B., Bode, M., Roberts, D. L., & Burgman, M. A. (2013). Inferring extinctions from sighting records of variable reliability. *Journal of Applied Ecology*, DOI
- [29] Maller, R. A. & Zhou, S. (1996). Survival Analysis With Long-Term Survivors, Wiley: Chichester, U.K.
- [30] May, R. M., Lawton, J. H. & Stork, N. E. (1995). *Assessing extinction rates*. Oxford University Press, Oxford, 1–24.
- [31] Peng, Y., Dear, K. B., & Denham, J. W. (1998). A generalized F mixture model for cure rate estimation. *Statistics in medicine* **17**(8) 813–830.
- [32] Pimm, S. L. & Raven, P. (2000). Extinction by numbers. *Nature* **403**, 843–845.
- [33] Possingham, H. P, Andelman, S. J., Burgman, M. A., Medellin, R. A., Master, L. L. & Keith, D. A. (2002) Limits to the use of threatened species lists. *Trends Ecol. Evol.* **17**, 503–507.
- [34] Promislow, D. E., & Harvey, P. H. (1990). Living fast and dying young: A comparative analysis of life-history variation among mammals. *Journal of Zoology*, **220**(3), 417–437.
- [35] Rout, T. M., Salomon, Y. & McCarthy, M. A. (2009). Using sighting records to declare eradication of an invasive species. *Journal of Applied Ecology*, **46**, 110–117.
- [36] Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. *Journal of the American Statistical Association*, 543–546.

- [37] Scheffers, B. R., Yong, D. L., Harris, J. B. C., Giam, X., & Sodhi, N. S. (2011). The world's rediscovered species: back from the brink? *-PloS one*, **6(7)**, e22531.
- [38] Slatyer, R. A., Hirst, M., & Sexton, J. P. (2013). Niche breadth predicts geographical range size: a general ecological pattern. *Ecology letters*, **16(8)**, 1104-1114.
- [39] Solow, A. & Roberts, D. L. (2003). A nonparametric test for extinction based on a sighting record. *Ecology*, **84**, 1329–1332.
- [40] Solow, A. (1993a). Inferring extinction from sighting data. *Ecology*, **74**, 962–964.
- [41] Solow, A., Smith, W., Burgman, M. A., Rout, T., Wintle, B. & Roberts, D. (2011). Uncertain sightings and the extinction of the ivory-billed woodpecker. *Conservation Biology*, **26**, 1–5.
- [42] Su, Y. S., Yajima, M., Gelman, A. E., & Hill, J. (2011). Multiple imputation with diagnostics (mi) in R: opening windows into the black box. *Journal of Statistical Software*, **45(2)**, 1–31.
- [43] Sy, J. P., & Taylor, J. M. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics* **56(1)** 227-236.
- [44] Thompson, C. J., Lee, T. E., Stone, L., McCarthy, M. A., & Burgman, M. A. (2013).
- [45] Veríssimo, D., I. Fraser, R. Bristol., J. Groombridge, & D. MacMillan. (2009). Birds as tourism flagship species: A Case Study on Tropical Islands. *Animal Conservation* **12** 549-558.
- [46] Zhang, J., & Peng, Y. (2007). A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in medicine* **26(16)** 3157–3171.
- [47] Endangered Species Act 5-Year Review: Caribbean Monk Seal (*Monachus tropicalis*). (2008) *National Oceanic and Atmospheric Administration National Marine Fisheries Service*.

A Simulated data example

To test the performance of the model we used simulated data for 300 hypothetical species, 150 of which are truly extinct. The times since the truly extinct fraction became missing were randomly generated from a log-normal distribution, with geometric mean of 97 and geometric standard deviation of 2.2. This distribution provided times which are similar to the empirical observations of non-detection times (Fisher & Blomberg, 2012). Similarly, a log-normal distribution with geometric mean 35 and geometric standard deviation of 1.8 resembles observed rediscovery data. Hence, this distribution was used to generate the time to rediscovery for the extant fraction. Using a Bernoulli distribution, we randomly allocated 90% of the rediscovered species with a trait, and 5% of the truly extinct species with the same trait. Therefore, if the species did not have the trait, then it was likely to be extinct. This was the 'true' record. Using this record, we would expect the model to predict high extinction for species without the trait, and low extinction for species with the trait. However, to test the performance of the model we randomly reset 50 of the rediscovered species to 'missing'. If the model could still predict high extinction for species without the trait, and low extinction for species with the trait, then the model was deemed to be performing adequately.

When using the simulated data, the model predicted an extinction rate of 0.9128 for species without the trait, and 0.0865 for species with the trait. From this simulation we concluded that the model performed as expected.

B Simple example of a cure rate model

Consider the simplest example where we examine how extinction and rediscovery is influenced by body mass only. Even simpler still, we consider only two categories: small mammals and large mammals. Since the median log body mass for the data set is 2.43g (268g), we consider mammals with log body mass less than 2.5g (315g) as small mammals, and all others as large mammals. Since there are two subgroups, the cure rate model provides two non-detection functions, see Fig. 5. Each non-detection function gives the proportion of species in that subgroup (y -axis) which remain undetected at a given time since they were last observed (x -axis).

In cure-rate analysis, the proportion of species that go extinct, and the maximum time it would take to rediscover all extant species, is given by the coordinates at which the non-detection levels out. In this analysis, the extinction rate for mammals under 2.5g log body mass is 16.8%, and all survivors in this group are rediscovered within 318 years of the last sightings; whereas the extinction rate for those over 2.5g is 40.2%, and all survivors in this group are rediscovered within 125 years of the last sighting. This simple example demonstrates that unseen large mammals are more than twice as likely to be truly extinct than unseen small mammals.

The non-detection curves provides further information. In this example, the re-discovery profile of the larger body mass species tends to be more steep, indicating that they are re-discovered at a higher rate than the smaller species.

The probability that a mammal is extinct in a particular year is given by the non-detected proportion as a ratio of the total proportion which will be detected. For example, suppose one large and one small which were both last seen 50 years ago, then the probability that the small mammal would be extinct in 2014 is $16.8/86.3=19.5\%$, and probability that the large mammal would be extinct in 2014 is $40.2/64.1=62.7\%$.

C: Testing variables

We list the models considered before settling on the chosen model:

Discovery: searcheffort * bodymass

Extinction: density * bodymass

The asterisk indicates interaction. The models below did not converge due to the dataset not being large enough. Italics indicate the changes made from one model to the next.

1. Discovery: density + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range * bodymass + searcheffort * bodymass + searcheffort * overkill + searcheffort * invasion * searcheffort * habitatloss + colouring + arboreal + diurnal + *elevation * range* + social + openhabitat + *humandensity * elevation* + humandensity * range

Extinction: openhabitat + density * bodymass + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range + humandensity * elevation + humandensity * overkill + humandensity * invasion + humandensity * habitatloss

2. Discovery: density + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range * bodymass + searcheffort * bodymass + *searcheffort * overkill + searcheffort * invasion * searcheffort * habitatloss* + colouring + arboreal + diurnal + social + openhabitat + humandensity * range

Extinction: openhabitat + density * bodymass + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range + humandensity * elevation + humandensity * overkill + humandensity * invasion + humandensity * habitatloss

3. Discovery: density + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range * bodymass + searcheffort * bodymass + colouring + arboreal + diurnal + social + openhabitat + humandensity * range

Extinction: openhabitat + density * bodymass + overkill * bodymass + invasion * bodymass + habitatloss * bodymass + range + *humandensity * elevation* + humandensity * overkill + humandensity * invasion + humandensity * habitatloss

- 643 **4. Discovery:** density + overkill * bodymass + invasion * bodymass + habitatloss *
644 bodymass + range * bodymass + searcheffort * bodymass + colouring + arboreal +
645 diurnal + social + openhabitat + humandensity * range
646 **Extinction:** openhabitat + density * bodymass + overkill * bodymass + invasion *
647 bodymass + habitatloss * bodymass + range + *humandensity * overkill +*
648 *humandensity * invasion + humandensity * habitatloss*
- 649 **5. Discovery:** density + overkill * bodymass + invasion * bodymass + habitatloss *
650 bodymass + range * bodymass + searcheffort * bodymass + colouring + arboreal +
651 diurnal + social + openhabitat + *humandensity * range*
652 **Extinction:** openhabitat + density * bodymass + overkill * bodymass + invasion
653 * bodymass + habitatloss * bodymass + range
- 654 **6. Discovery:** density + *overkill * bodymass + invasion * bodymass + habitatloss **
655 *bodymass + range * bodymass + searcheffort * bodymass + colouring + arboreal +*
656 *diurnal + social + openhabitat*
657 **Extinction:** openhabitat + density * bodymass + overkill * bodymass + invasion *
658 bodymass + habitatloss * bodymass + range
- 659 **7. Discovery:** density + range * bodymass + searcheffort * bodymass + *colouring +*
660 *arboreal + diurnal + social + openhabitat*
661 **Extinction:** openhabitat + density * bodymass + *overkill * bodymass + invasion **
662 *bodymass + habitatloss * bodymass + range*
- 663 **8. Discovery:** density + *range * bodymass + searcheffort * bodymass + openhabitat*
664 **Extinction:** density * bodymass + range
- 665 **9. Discovery:** density + searcheffort * bodymass + openhabitat
666 **Extinction:** density * bodymass + range
- 667 **10. Discovery:** searcheffort * bodymass + openhabitat
668 **Extinction:** density * bodymass + range
- 669
670
671
672
673

Covariate	Categories/Units	Influences
Species density	Sparse (1); Common (0.5); Dense (0)	Extinction
Species body mass	Log-transformed, $\log_{10}(\text{mass(g)}+1)$, and rounded to the nearest integer	Both
Search effort	Low: 0-2 searches (1); Medium: 3-6 searches (2); High: 10+ searches (3).	Rediscovery

Table 1: Covariates of extinction and rediscovery used in the model. Values of these variables for each species were extracted from Fisher and Blomberg (2011).

678

679

Species (yr last seen)	Yrs Missing	Redlist	Density	Search	Log body mass	E. 2014	E. rate	M. time (year)
Megalomysdesmarest(1902)	112	EX	0.5	1	3	0.7577	0.3494	146 (2048)
Gazella saudiya (1936)	78	EX	0	1	4	0.9463	0.7684	103 (2039)
Lagorchestes leporid (1889)	82	EX	0	1	3	0.9675	0.7466	146 (2035)
Lipotes vexillifer (2006)*	8	CR	1	3	5	0.7245	0.6345	28 (2034)
Macropus greyi (1927)	87	EX	0	1	4	0.9706	0.7684	103 (2030)
Dendrolagus mayri (1928)	86	CR	1	1	4	0.8767	0.4473	103 (2031)
Zalophus japonicus (1951)	63	EX	0.5	1	NA	0.9412	0.7103	74 (2025)
Gazella bilkis (1951)	63	EX	0.5	1	5	0.9635	0.7197	73 (2024)
Uromys imperator (1887)	127	CR	0	1	NA	0.7523	0.3335	137 (2024)
Potorous platyops (1875)	139	EX	0.5	1	3	0.8929	0.3494	146 (2021)
Pteropus pilosus (1873)	141	EX	0.5	1	3	0.9217	0.427	146 (2019)
Pteropus aruensis (1867)	147	CR	1	1	3	1	0.1443	146 (2013)
Elaphurus	75	EW	0.5	1	5	1	0.7197	73

Davidianus (1939)								(2012)
Eudorcas rufina (1936)	78	DD	1	1	5	1	0.6345	73
								(2009)
Bos sauveli (1986)*	28	CR	1	3	6	1	0.878	16
								(2002)
Equus ferus (1969)*	45	EX	0.5	3	5	1	0.7197	28
								(1997)
Monachus tropicalis	62	EX	0	2	5	1	0.7879	45
(1952)								(1997)
Boromys offella (1850)	164	EX	0.5	1	3	1	0.3494	146
								(196)
Neovison macrodon	134	EX	1	1	4	1	0.2925	103
(1880)								(1983)
Rucervus	76	EX	0.5	2	5	1	0.7197	45
Schomburg (1938)								(1983)
Thylacinus cynoceph	81	EX	1	3	4	1	0.2925	50
(1933)								(1983)
Dusicyon australis (1876)	138	EX	0	1	4	1	0.7684	103
								(1979)
Gazella arabica (1825)	189	DD	0.5	1	4	1	0.3966	103
								(1928)
Hippotragus leucoph	215	EX	1	1	5	1	0.6345	73
(1799)								(1872)
Hydrodamalis gigas	246	EX	0	1	7	1	0.8205	36
(1768)								(1804)

680

681 **Table 2:** The 25 missing or extinct mammals whose maximum time to rediscovery expires
682 before 2050 under our model (including mammals whose maximum time to rediscovery
683 expired before 2014), and will therefore be deemed definitely extinct by 2050. The columns
684 show the species name (with the year it was last seen); the number of years since it was last
685 seen, the IUCN Red List classification (E extinct, CR critically endangered, EW extinct in the

686 wild, DD data deficient); population density category (0 dense, 0.5 locally common, 1
687 sparse); search effort category (1 low, 2 medium, 3 high); log body mass to the nearest
688 integer (3~1kg, 4~10kg, 5~100kg, etc.); the extinction probability in 2014 (E. 2014); the
689 extinction rate (E. rate); the maximum time to rediscovery (M. time); and the year that time
690 expired. Species with a very high probability of being extinct within 50 years since last
691 reporting are indicated with an asterisk.

692

Species (yr last seen)	E. 2014	E. rate	M. time (yr)
<i>Hydrodamalis gigas</i> (1768)	1.0000	0.8205	36 (1804)
<i>Hippotragus leucophaeus</i> (1799)	1.0000	0.6345	73 (1872)
<i>Gazella arabica</i> (1825)	1.0000	0.3966	103 (1928)
<i>Dusicyon australis</i> (1876)	1.0000	0.7684	103 (1979)
<i>Thylacinus cynocephalus</i> (1933)	1.0000	0.2925	50 (1983)
<i>Rucervus schomburgki</i> (1938)	1.0000	0.7197	45 (1983)
<i>Neovison macrodon</i> (1880)	1.0000	0.2925	103 (1983)
<i>Boromys offella</i> (1850)	1.0000	0.3494	146 (1996)
<i>Monachus tropicalis</i> (1952)	1.0000	0.7879	45 (1997)
<i>Equus ferus</i> (1969)	1.0000	0.7197	28 (1997)
<i>Bos sauveli</i> (1986)	1.0000	0.8780	16 (2002)
<i>Eudorcas rufina</i> (1936)	1.0000	0.6345	73 (2009)
<i>Elaphurus davidianus</i> (1939)	1.0000	0.7197	73 (2012)
<i>Pteropus aruensis</i> (1867)	1.0000	0.1443	146 (2013)
<i>Macropus greyi</i> (1927)	0.9706	0.7684	103 (2030)
<i>Lagorchestes leporides</i> (1889)	0.9675	0.7466	146 (2035)
<i>Gazella bilkis</i> (1951)	0.9635	0.7197	73 (2024)
<i>Boromys torrei</i> (1850)	0.9522	0.7226	206 (2056)
<i>Gazella saudiya</i> (1936)	0.9463	0.7684	103 (2039)
<i>Zalophus japonicus</i> (1951)	0.9412	0.7103	74 (2025)
<i>Pseudomys gouldii</i> (1857)	0.9384	0.7226	206 (2063)
<i>Pteropus subniger</i> (1864)	0.9319	0.7226	206 (2070)
<i>Pteropus pilosus</i> (1873)	0.9217	0.4270	146 (2019)
<i>Lagorchestes asomatus</i> (1932)	0.8964	0.7466	146 (2078)
<i>Neotoma bunkerii</i> (1931)	0.8964	0.7466	146 (2077)
<i>Potorous platyops</i> (1875)	0.8929	0.3494	146 (2021)
<i>Oligoryzomys victus</i> (1892)	0.8925	0.7226	206 (2098)
<i>Dendrolagus mayri</i> (1928)	0.8767	0.4473	103 (2031)

<i>Neotoma martinensis</i> (1925)	0.8668	0.7226	182 (2107)
<i>Neotoma anthonyi</i> (1926)	0.8659	0.7226	182 (2108)
<i>Oryx dammah</i> (1988)	0.8602	0.7879	73 (2061)
<i>Rattus macleari</i> (1903)	0.8580	0.7015	221 (2124)
<i>Geocapromys thoracatus</i> (1955)	0.8463	0.7466	146 (2101)
<i>Nesoryzomys indefessus</i> (1934)	0.8242	0.7226	206 (2140)
<i>Mesocapromys sanfelipensis</i> (1978)	0.8237	0.7466	114 (2092)
<i>Mesocapromys nanus</i> (1937)	0.7816	0.7074	277 (2214)
<i>Notomys mordax</i> (1845)	0.7802	0.4054	206 (2051)
<i>Onychogalea lunata</i> (1956)	0.7758	0.5400	103 (2059)
<i>Peromyscus pembertoni</i> (1931)	0.7738	0.6965	292 (2223)
<i>Megalomys desmarestii</i> (1902)	0.7577	0.3494	146 (2048)
<i>Uromys imperator</i> (1887)	0.7523	0.3335	137 (2024)
<i>Pappogeomys alcorni</i> (1998)	0.7300	0.7226	206 (2204)
<i>Lipotes vexillifer</i> (2006)	0.7245	0.6345	28 (2034)
<i>Peromyscus guardia</i> (1991)	0.7054	0.6965	292 (2283)
<i>Rattus nativitatis</i> (1903)	0.6886	0.3080	156 (2059)
<i>Notomys macrotis</i> (1843)	0.6574	0.4428	206 (2049)
<i>Canis rufus</i> (1980)	0.6558	0.5400	103 (2083)
<i>Macrotis leucura</i> (1931)	0.6130	0.3494	146 (2077)
<i>Megalomys luciae</i> (1852)	0.6109	0.1974	206 (2058)
<i>Lagostomus crassus</i> (1910)	0.6009	0.3824	205 (2115)
<i>Calcochloris tytonis</i> (1964)	0.5850	0.5551	292 (2256)
<i>Perameles eremiana</i> (1943)	0.5793	0.3494	146 (2089)
<i>Pteropus brunneus</i> (1874)	0.5442	0.1974	206 (2080)
<i>Nyctophilus howensis</i> (1915)	0.5399	0.4619	291 (2206)
<i>Zaglossus attenboroughi</i> (1961)	0.5267	0.2925	103 (2064)
<i>Uromys porculus</i> (1888)	0.4729	0.4498	184 (2072)
<i>Melanomys zunigae</i> (1949)	0.4461	0.3692	182 (2131)
<i>Notomys longicaudatus</i> (1901)	0.4206	0.1974	206 (2107)

<i>Leimacomys buettneri</i> (1890)	0.3239	0.2302	292 (2182)
<i>Hybomys basillii</i> (1965)	0.3171	0.2703	206 (2171)
<i>Dendromus vernayi</i> (1937)	0.3108	0.2798	292 (2229)
<i>Leporillus apicalis</i> (1933)	0.3068	0.1974	206 (2139)
<i>Myosorex eisentrauti</i> (1968)	0.2930	0.1340	292 (2260)
<i>Myosorex rumpii</i> (1967)	0.2930	0.2798	292 (2259)
<i>Crateromys paulus</i> (1953)	0.2900	0.1443	114 (2067)
<i>Uromys emmae</i> (1946)	0.2743	0.1886	155 (2101)
<i>Tylomys tumbalensis</i> (1968)	0.2725	0.2389	206 (2174)
<i>Octodon pacificus</i> (1959)	0.2603	0.1974	206 (2165)
<i>Pipistrellus sturdeeii</i> (1891)	0.2169	0.1698	292 (2183)
<i>Dipodomys gravipes</i> (1986)	0.2162	0.1974	206 (2192)
<i>Nesoryzomys darwini</i> (1930)	0.1450	0.1012	292 (2222)
<i>Mysateles meridionalis</i> (1978)	0.1353	0.0898	114 (2092)
<i>Pseudomys glaucus</i> (1956)	0.1257	0.1012	292 (2248)
<i>Mystacina robusta</i> (1967)	0.1162	0.1012	292 (2259)
<i>Conilurus albipes</i> (1875)	0.1043	0.0233	206 (2081)
<i>Oryzomys antillarum</i> (1877)	0.1016	0.0233	206 (2083)
<i>Phalanger matanim</i> (1997)	0.0974	0.0898	146 (2143)
<i>Peromyscus mekisturus</i> (1948)	0.0822	0.0557	206 (2154)
<i>Oryzomys nelsoni</i> (1897)	0.0660	0.0233	206 (2103)
<i>Nyctimene santacruensis</i> (1907)	0.0650	0.0413	292 (2199)
<i>Chaeropus ecaudatus</i> (1907)	0.0624	0.0233	206 (2113)
<i>Notomys amplius</i> (1896)	0.0537	0.3743	206 (2102)
<i>Thallomys shortridgei</i> (1922)	0.0507	0.0233	206 (2128)
<i>Nilopegamys plumbeus</i> (1928)	0.0447	0.0233	206 (2134)
<i>Paracoelops megalotis</i> (1947)	0.0378	0.0284	292 (2239)
<i>Juscelinomys candango</i> (1960)	0.0350	0.0233	182 (2142)
<i>Tylomys bullaris</i> (1968)	0.0311	0.0233	206 (2174)
<i>Abditomys latidens</i> (1972)	0.0298	0.0233	206 (2178)

<i>Pteropus tokudae</i> (1974)	0.0290	0.0233	206 (2180)
<i>Murina tenebrosa</i> (1962)	0.0286	0.0240	292 (2254)
<i>Pteralopex pulchra</i> (1990)	0.0253	0.0233	206 (2196)
<i>Pharotis imogene</i> (1890)	0.0121	0.0058	292 (2182)
<i>Harpiola grisea</i> (1904)	0.0101	0.0058	292 (2196)
<i>Cryptochloris wintoni</i> (1907)	0.0100	0.0058	292 (2199)
<i>Chrysochloris visagiei</i> (1950)	0.0077	0.0058	292 (2242)
<i>Cryptonanus ignitus</i> (1962)	0.0072	0.0058	292 (2254)
<i>Dendromus kahuziensis</i> (1967)	0.0068	0.0058	292 (2259)
<i>Crocidura wimmeri</i> (1976)	0.0065	0.0058	292 (2268)
<i>Myotis hajastanicus</i> (1989)	0.0015	0.0015	413 (2402)

Table 3: The 99 missing mammals arranged in descending order of the probability that the species is extinct in 2014 (E. 2014), calculated from the model. For each species we provide the extinction rate (E. rate) and maximum time to rediscovery (M. time) from our model. The year in which the maximum time to rediscovery occurs is provided in brackets. Note that the cure rate model does not refer specifically to the individual species on that line, but to a the set of species with the same combination of traits as the species in question (see Methods and Appendix B for further information). When the time missing is greater than the maximum time to rediscovery, the model predicts the species to be truly extinct - regardless of the extinction rate (maximum time to rediscovery in bold). Species with a very low chance of being extinct despite ≥ 50 years since last reporting (with a $< 10\%$ probability of being extinct under our model although they have been missing for > 50 years - since 1966 or before) are indicated by a maximum time to rediscovery in italics.