Endangered Species Recovery Criteria: Reconciling Conflicting Views

STEVEN R. BEISSINGER

Recovery criteria (RC) serve the important purpose of determining when an *endangered* species can be delisted, or removed from protection under the Endangered Species Act (ESA). Although delisting is the ultimate goal for recovering all *threatened* species, it has been a controversial process, because delisted species may lose some protection provided by the ESA, making them susceptible to the same causes of decline that resulted in their initial listing. RC are designated by the US Fish and Wildlife Service (USFWS) in the recovery plan of an endangered species, and the ESA mandates that RC should be based on "objective, measurable criteria." How to designate RC has been a thorny problem since the ESA was enacted in 1972.

In this issue of *BioScience*, we find conflicting views of how RC should be determined offered by Doak and colleagues and by Wolf and colleagues. The authors of both articles champion the application of quantitative RC in place of the heterogeneous approach currently used. But here the similarities end. Doak and colleagues argue that RC should be based on "demographic criteria," emphasizing estimates of the risk of extinction from a population viability analysis (PVA) that projects population size for decades to 100 years (or more) into the future. Under this approach, recovery plans would be required to include RC tied to the probability of meeting specific extinction risk or demographic thresholds from models.

In contrast, Wolf and colleagues espouse a less data-intensive and broader set of methods, based on the ecological principles of representation, resiliency, and redundancy (the "3Rs"). The 3Rs would be evaluated quantitatively or qualitatively using multiple approaches for setting recovery targets, such as the percentage of historic range, population size, the number and spatial distribution of populations, and the risk of extinction from a PVA when adequate data are available.

Both articles provide important insights into the shortcomings of past efforts to delineate RC and discussion of the concepts for delineating RC. But both run head on into the same knotty problems of developing quantitative criteria for RC that relate to population viability and the absence of risk standards.

Different paths, same impasse

For RC, the crucial characteristics related to population viability that can be objectively measured or statistically estimated (i.e., not projected from a simulation model) include population size and trend, the number and distribution of populations throughout the range, the size of the geographic range, and the levels of various threats. In an ideal world, the probability of extinction would also be used as a delisting criterion. Unfortunately, extinction risk is not directly measureable in the field, and it is not an entirely objective quantity, as is discussed below.

Accurate forecasting of a species' risk of extinction has long been a challenge for population and conservation biologists because many factors (foreseen and unforeseen) affect the dynamics of populations, because demography and threats vary stochastically and unpredictably, and because models project population size far into the future (Beissinger and Westphal 1998). As a result, quantitative measures of the risk of extinction are usually produced from simulation (PVA) models that are designed to integrate some portion of these processes. Unfortunately, forecasts of population sizes and extinction risk for 50 or 100 years in the future are often unreliable, unless populations are growing or declining rapidly and environmental stochasticity is small (Ellner and Holmes 2008). Moreover, multiple models applied to the same situation can yield greatly differing results and, indeed, have dueled in the courtroom (Beissinger and Westphal 1998). Few species at the time of listing have adequate data for creating a credible PVA model (see the discussion by Wolf et al.), and better parameter estimates obtained over time can radically change model forecasts of viability (Maehr et al. 2002). Incorporating multiple threats into a PVA is rarely done; for example, few PVAs mechanistically incorporate climate or land-use change, but these processes are potential game changers for all species over the next century.

Model validation is an important step before basing decisions on biological outcomes from a simulation model, but validation is difficult to do for PVA models. The primary prediction from a PVA model—the probability of extinction—is almost impossible to test in the field (Beissinger and Westphal 1998). Comparing the average population trajectory from model outputs with real population trajectories (e.g., Brook et al. 2000) is unlikely to allow the rejection of a model, because the error bars around the predicted population trajectories from stochastic models are often very large (Ellner and Holmes 2008). Moreover, it is the values at one end of the distribution of projected population sizes produced by a PVA (i.e., extinction), which are driven primarily by stochastic processes, that require validation, rather than the average population trajectory, which is driven by mean rates. Secondary predictions from PVA models (e.g., patch occupancy colonization and extinction rates) can be easier to validate, but the time period used to test such predictions is often short (e.g., several years) relative to the time period used for model projections (typically 50–100 years).

Reviews have suggested that PVA models are best used to compare scenarios with each other to inform relative risk, rather than to produce forecasts of the risk of extinction. In this context, models are used to inform decisions in concert with other information, such as expert opinion, rather than becoming the decision-making tool envisioned by Doak and colleagues.

Whether they emerge from a 3Rs, PVA, or expert analysis, the basis for evaluating RC requires the comparison of objective, measurable criteria with standards of acceptable levels of risk. Unlike the risk from exposure to toxins set by the US Environmental Protection Agency, standards for risk of extinction do not exist under US law, and rules have not been promulgated by the USFWS.

Adopting international risk standards

A straightforward solution to the problem of both delisting (recovery) and listing criteria has been around for two decades. The International Union for Conservation of Nature (IUCN) categories of risk were developed with the explicit purpose of producing a relative estimate of the likelihood of extinction for a species (Mace et al. 2008). They are widely acknowledged as useful risk assessment tools for informing conservation when data are limited. These categories are the closest thing in conservation biology to professional risk assessment standards, because they were created and vetted through a rigorous process that included input from a large number of conservation scientists, can be applied consistently, and have stood up well (Mace et al. 2008).

The IUCN categories of risk provide guidelines for the trend, number, size, and distribution of populations in relation to different levels of risk and threat to species. Delisting (i.e., recovery) is readily interpreted as one of those categories; whether the species is subsequently labeled *near threatened* or *vulnerable* may depend on societal preferences. Moreover, the IUCN standards for each category are tied to multiple types of criteria, which allows the USFWS to use a variety of tools to assist in creating RC, including the kinds of analyses envisioned by both Wolf and colleagues and Doak and colleagues. The guidelines are straightforward to apply to species listed under the ESA (Harris et al. 2012).

Setting quantitative criteria for listing and delisting makes the process more objective and less susceptible to political pressures, promotes fairness among species and stakeholders, and increases public trust and political legitimacy (Robbins 2009). Once the criteria are set, evaluating a species' situation using the best available science is a purely scientific problem. The criteria could be treated as guidelines to be followed most of the time, with occasional deviations when a species' biology or situation dictates. Delisting decisions would then be made in combination with a clear understanding of the extent that threats have been ameliorated and laws or processes have been put into effect to maintain protection.

Academia has and will continue to play an important role in assisting conservation, but our capability to be the PVA factory envisioned by Doak and colleagues is limited, and our goals are much broader. If the majority of this effort were performed by professors and graduate students, the analyses would vary greatly in quality, and the risk standards would be inconsistent. The USFWS needs in-house expertise in modeling and quantitative analysis that could be applied across species in a standardized manner to analyze risk and the responses to threats in order to recover *endangered* species. The USFWS would benefit by assembling dedicated regional teams with the appropriate expertise, including that in population biology, genetics, landscape ecology, statistical analysis, and modeling, to perform analyses that could meet a set of agency and professional standards.

References cited

Beissinger SR, Westphal MI. 1998. On the use of demographic models of population viability in endangered species management. Journal of Wildlife Management 62: 821–841.

Brook BW, O'Grady JJ, Chapman AP, Burgman MA, Akcakaya HR, Frankham R. 2000. Predictive accuracy of population viability analysis in conservation biology. Nature 404: 385–387.

Ellner SP, Holmes EE. 2008. Commentary on Holmes et al. (2007): Resolving the debate on when extinction risk is predictable. Ecology Letters 11: E1–E5.

Harris JBC, Reid JL, Scheffers BR, Wanger TC, Sodhi NS, Fordham DA, Brook BW. 2012. Conserving imperiled species: A comparison of the IUCN Red List and U.S. Endangered Species Act. Conservation Letters 5: 64–72.

Mace GM, Collar NJ, Gaston KJ, Hilton-Taylor C, Akcakaya HR, Leader-Williams N, Milner-Gulland EJ, Stuart SN. 2008. Quantification of extinction risk: IUCN's system for classifying threatened species. Conservation Biology 22: 1424–1442.

Maehr DS, Lacy RC, Land ED, Bass OL, Hoctor TS. 2002. Evolution of population viability assessments for the Florida panther: A multiperspective approach. Pages 284–311 in Beissinger SR, McCullough DR, eds. Population Viability Analysis. University of Chicago Press.

Robbins K. 2009. Strength in numbers: Setting quantitative criteria for listing species under the Endangered Species Act. UCLA Journal of Environmental Law and Policy 27: 1–37.

Steven R. Beissinger (beis@berkeley.edu) is a professor of conservation biology in the Department of Environmental Sciences and Museum of Vertebrate Zoology, at the University of California, Berkeley.

doi:10.1093/biosci/biu212