Please draw on the citations and logic below, which represent my interpretation of evidence relating to various issues in the wolf science-policy interface. This commentary addresses common policies and claims made by wildlife agencies about their jurisdiction’s wolf populations. I have attempted two steps: (1) disentangling value judgments from scientific claims, and (2) my perception of the consensus in the scientific community if any. Paraphrasing Oreskes N. 2019: No single study should be considered reliable, even replicated studies await consensus in the scientific community, and only qualified experts with a track record in the particular field can weigh in on that evidence. Sometimes it is too early to claim consensus and sometimes consensus in the scientific community should be reached but the intrusion of claims by those who have financial or non-financial competing interests will distort or delay scientific consensus and its application to policy. Therefore, I offer below my own view of consensus within the scientific community about three specific bodies of evidence in wolf science. Where I cannot find consensus I try to illuminate the standards of evidence one should employ to weigh contrasting research and select the best available science.

I want to echo a comment by colleagues with which I agree and which requires the strongest emphasis: "In general, the systematic omission of scientific literature on various topics, the arbitrary prioritization of policy objectives, and the resulting biased perspectives on wolves, interventions and impacts result in a prejudiced and unscientific Plan." (Santiago-Ávila & Lute 2023 provided to CPW during this public comment period). I would add that selective citation biased towards one worldview is a breach of scientific integrity that violates federal regulations and policy on agency science as I explain further in Section 5 below. This shortcoming of the wolf management plan also makes the science not the best available. Given federal involvement in wolf reintroduction to Colorado, federal standards for best available science apply.

Also, in case the reader is tempted to dismiss my work as less objective than someone else’s because I work in the science-policy interface (or other less diplomatic formulations of this canard) point them here: Treves A, Paquet PC, Artelle KA, Cornman AM, et al. 2021. Transparency about values and assertions of fact in natural resource management. Frontiers in Conservation Science: Human-Wildlife Dynamics 2: e631998. 10.3389/fcosc.2021.631998.

1. Population viability assessments (PVAs), minimum viable populations, and strict numerical delisting targets.

First and most importantly, value judgments arising from personal, organizational, or governmental commitments, investments, and preferences for certain outcomes are the first steps in PVAs because the following inputs or decisions are value-based choices not scientific decisions: (A) how far ahead in time should one forecast (e.g., perpetuity or a few years, see Frankham et al. below for preserving evolutionary potential and consider whether the jurisdiction is fulfilling the trustee duty to preserve the asset for future generations); (B) how much risk of extinction is the public willing to tolerate and how much is the agency allowing? (e.g., near zero? Or a predetermined level of risk?). (C) Did the modelers exclude any threats (e.g., super-additive mortality from human-caused killing, illegal killing, catastrophic disease)? (D) Most PVAs risk being misapplied to jurisdictions rather than to actual biological populations. (E) The decision to focus on census population size (all individuals) or effective population size (Ne, shorthand definition is all breeding individuals). Each address different aspects of viability. The former addresses demographic viability without guarantees that the surviving animals will be healthy and reproducing effectively, whereas the latter is more precautionary and addresses evolutionary potential. If evolutionary potential is preserved, one can generally assume demographic survival has also been preserved. In wolves, where virtually all packs contain only two breeders, effective population size can be estimated by the number of packs in one year

Most of the cited articles in this letter are available for free download at the following institutional server without pass wall or paywall: http://faculty.nelson.wisc.edu/treves/archive_BAS/Pubs_for_public_comments.zip. Failure to include these documents in the administrative public record is unlawful.
(although this too may be an over-estimate because not all pairs breed each year) but the rest of the population provides a source of new breeders to replace those that die.

References and notes for section 1 only

Consider the distinction between science-informed decisions and making decisions based solely on science from the quote from Vucetich JA, Nelson MP and Phillips MK. 2006. The normative dimension and legal meaning of endangered and recovery in the U.S. Endangered Species Act. Conserv Biol 20: 1383-1390. “The ESA's requirement that endangerment be determined ‘solely on the basis of the best scientific and commercial data available’ does not mean scientists have exclusive right to determine the normative dimensions of specifying the conditions of extinction. This mandate merely provides science the exclusive right to determine whether specified conditions for endangerment are met by particular species.”

Carroll C, Lacy RC, Fredrickson RJ, Rohlf DJ, et al. 2019. Biological and sociopolitical sources of uncertainty in population viability analysis for endangered species recovery planning. Scientific Reports 9: e10130. https://doi.org/10.1038/s41598-019-45032-2. Abstract: Although population viability analysis (PVA) can be an important tool for strengthening endangered species recovery efforts, the extent to which such analyses remain embedded in the social process of recovery planning is often unrecognized. We analyzed two recovery plans for the Mexican wolf that were developed using similar data and methods but arrived at contrasting conclusions as to appropriate recovery goals or criteria. We found that approximately half of the contrast arose from uncertainty regarding biological data, with the remainder divided between policy-related decisions and mixed biological-policy factors. Contrasts arose from both differences in input parameter values and how parameter uncertainty informed the level of precaution embodied in resulting criteria. Policy-related uncertainty originated from contrasts in thresholds for acceptable risk and disagreement as to how to define endangered species recovery. Rather than turning to PVA to produce politically acceptable definitions of recovery that appear science-based, agencies should clarify the nexus between science and policy elements in their decision processes. The limitations we identify in endangered-species policy and how PVAs are conducted as part of recovery planning must be addressed if PVAs are to fulfill their potential to increase the odds of successful conservation outcomes.

Frankham R, Bradshaw CJA and Brook BW. 2014. Genetics in conservation management: Revised recommendations for the 50/500 rules, red list criteria and population viability analyses. Biol Conserv 170: 56-63. https://doi.org/10.1016/j.biocon.2013.12.036 Abstract: Conservation managers typically need to make prompt decisions based on limited information and resources. Consequently, generalisations have essential roles in guiding interventions. Here, we (i) critique information on some widely accepted generalisations and variables affecting them, (ii) assess how adequately genetic factors are currently incorporated into population viability analysis (PVA) models used to estimate minimum viable population sizes, and (iii) relate the above to population size thresholds of the IUCN Red List criteria for threatened species that were derived from genetic considerations. Evidence accumulated since 1980 shows that genetically effective population size ($Ne$) = 50 is inadequate for preventing inbreeding depression over five generations in the wild, with $Ne > 100$ being required to limit loss in total fitness to <10%. Further, even $Ne = 500$ is too low for retaining evolutionary potential for fitness in perpetuity; a better approximation is $Ne > 1000$. Extrapolation from census population size ($N$) to $Ne$ depends on knowing the ratio of $Ne/N$, yet this information is unavailable for most wild populations. Ratio averages (~0.1–0.2) from meta-analyses are sufficient, provided adjustments are made for dissimilar life histories. Most PVA-based risk assessments ignore or inadequately model genetic factors. PVA should routinely include realistic inbreeding depression, and genetic impacts on evolutionary potential should be incorporated where appropriate. Genetic generalisations used in conservation, the treatment of
Public comment by Adrian Treves, PhD

genetics in PVAs, and sections of the IUCN Red List criteria derived from genetic considerations, all require revision to be more effective conservation tools.

Finally, the selection of a specific numerical target or even a range of targets presumes a value judgment at some point in the process. For an explanation of how this plays out in sustainable harvest models applied to wolves, please see Treves A, Paquet PC, Artelle KA, Cornman AM, et al. 2021. Transparency about values and assertions of fact in natural resource management. Frontiers in Conservation Science: Human-Wildlife Dynamics 2: e631998. 10.3389/fcosc.2021.631998.

2. Lethal methods are *valuable* tools in the toolkit to prevent adverse predation by wolves. (* you can replace valuable with other value judgments such as best, essential, necessary*)

None of the terms surrounded by * are scientific in this context, they are all value judgments. A scientist would write lethal wolf control is a tool in the toolkit. The adjective is a flag that someone is attempting to impose their values on their readers. A scientist would add the adjectival phrase "effective to achieve x" or "ineffective to achieve x" or related measurable criteria that do not presuppose a normative judgment about whether one should use lethal control (as do the * adjectives). Note the word control can be misleading if not defined as it may convey to some audiences that prevention has been achieved even when it has not.

In most cases in wolf policy, effectiveness of lethal or non-lethal methods is oriented to reducing future losses of domestic animals or wild animals. Starting with domestic animals, beware of certain value judgments inherent to this question: (A) do the domestic animals or wild ungulates take priority over wolves, especially when lethal control is envisioned? That would also be a value judgment. (B) Also beware of the value judgment that all lethal methods can be meaningfully grouped into a single category. Scientifically, neither lethal nor non-lethal interventions can be grouped for a claim of effectiveness because each variety and each implementation may differ from the next in effectiveness based on design, location, conditions, effort, etc. Only when followed by the details of studies of effectiveness can one group two or more methods together to make claims (i.e., the advocate for either needs to anchor their conclusion about a method in a study specific to that method). (C) Some methods for protecting domestic and wild animals from wolves are in a gray area between lethal and non-lethal. This and other false dualisms are common in this field. (D) Effectiveness as I have defined it above has no spatial or temporal specification but should. No method is always effective, therefore the claim of effectiveness should be applied to a particular time, place, and design of the intervention. When scientists communicate clearly, we cite a study when making such a claim. All implementations are experiments, no single study will predict with perfect certainty whether a new application is effective. (E) Another subtler value judgment is whether the effectiveness of lethal methods and effectiveness of non-lethal methods are commensurable, because the effects of one cannot be reversed while the other may be.

Given the caveats above, single studies are no basis for confidence until replicated. Therefore the methods that have been replicated by two or more studies show they are effective should give the most confidence. Therefore, systematic reviews or meta-analyses are our best basis for predicting whether a given intervention is effective. Even systematic reviews and meta-analyses are prone to error if they have a small sample size to draw from, authors have competing interests, or the analysis did not consider strength of inference.

The strongest inference is drawn from randomized, controlled trials (RCTs) of methods to prevent predation on domestic or wild ungulates. Preferably those RCTs include crossover designs, blinding, and are subjected to authentic independent review and efforts at replication. The next and much lower strength of inference is the silver standard of before-and-after
comparison without randomization. That is the highest standard reached for lethal control of wolves and was only reached in 3 studies to my knowledge.


In the below graphic, I summarize the three studies' findings on the effects of lethal methods on wolves but again note that we still have no RCT on killing wolves to protect domestic animals and note the methods for killing wolves in the three studies differed somewhat.

My co-authors and I do not consider the analysis by Bradley et al. (2015) in the Journal of Wildlife Management to be reliable because (a) the authors could not explain several steps in the methods to us, (b) they did not share the data for us so we could use our own methods, (c) their methods biased the results toward favoring lethal control by extending the time horizon for livestock losses beyond the point where wolves held a territory. Thereby, they counted vacant territories as if territories can kill livestock, rather than packs. Their approach seems analogous to a study of a hospital treatment that measured filled and vacant hospital beds rather than the survival or death of patients) - see detailed explanation and discussion in Santiago-Ávila et al. 2018. Furthermore, the journal in which Bradley et al. 2015 was published, only subscribed to the Committee on Publication Ethics (COPE, https://publicationethics.org/) in 2022: Krausman PR. 2022. Improving the journal of wildlife management: A response to the perspectives of Johnson et al. (2021) and Gould et al. (2021). The Journal of Wildlife Management 86: e22167. https://doi.org/10.1002/jwmg.22167. Therefore, articles in that journal, the Wildlife Society Bulletin and Wildlife Monographs did not have the guardrails on publication ethics recommended by the Committee on Publication Ethics, which include scrutiny of potentially competing interests among authors, peer reviewers, editors, and the publisher, the Wildlife Society. Nor did the journals have guardrails for correction and retraction of misleading or fabricated science.

I asserted above that meta-analyses and systematic reviews are relatively more valuable than single studies for drawing generalizable inferences about the effectiveness of any method for preventing wolf predation on wild or domestic animals. Although the 11 studies below differ in standards of evidence, sample sizes, and inclusion criteria, one consistent conclusion emerges: without RCTs or other robust designs that control confounding variables, it is impossible to claim with confidence the effectiveness of any method. That does not mean we have zero information or knowledge, but rather that your caution should be heightened when we lack RCTs for this field of science.

Moreover, van Eeden et al. 2018 in PLoS Biology showed how differently four studies regarding lethal methods published between 2016-2018 drew from the literature despite similar search criteria. "Our four reviews [6–9] jointly screened >27,000 candidate studies. The four sets of inclusion criteria differed in geographic coverage, carnivore species, and standards of evidence and research design (see S1 Table), which limited overlap in the studies that passed screening (only 19% of studies were included in two or more of the four reviews; no study was included in all four, S1 Fig). The differing inclusion criteria also meant that it was not possible to conduct a
Lethal management of wild wolves to protect domestic animals has only been evaluated with before-and-after comparisons without randomization (silver standard). The field should aim for the higher strength of inference provided by gold-standard randomized, controlled trials.

<table>
<thead>
<tr>
<th>Effect</th>
<th>France</th>
<th>Slovenia</th>
<th>Michigan U.P.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired reduction in livestock deaths</td>
<td>33%</td>
<td>28%</td>
<td>25%</td>
</tr>
<tr>
<td>Undesirable increase in livestock deaths</td>
<td>11%</td>
<td>65%</td>
<td>75%</td>
</tr>
<tr>
<td>No effect</td>
<td>55%</td>
<td>7%</td>
<td>-</td>
</tr>
</tbody>
</table>


Note: Bradley et al. (2015) does not meet the criteria for reliability until they remedy irreproducible methods, a bias towards lethal control, and share their data for replication (Bradley et al. 2015. Effects of wolf removal on livestock depredation recurrence and wolf recovery in Montana, Idaho, and Wyoming. J Wildl Manage 79: 1337–1346.)

quantitative comparison (meta-analysis) combining the data from our four reviews, but we suggest that such an analysis should be conducted in the future as evidence increases. Nonetheless, our reviews came to remarkably similar conclusions, irrespective of methods, suggesting that our conclusions are robust."p.3 van Eeden et al. 2018 PLoS Biology. Because we brought together almost two dozen authors from 11 countries for van Eeden et al. 2018 in PLoS Biology, it is the leading review that drew the clearest consensus. That consensus included (a) the field needs stronger inference and (b) that lethal methods have not been studied with as high standards of inference as non-lethal methods.

Moreover, combining van Eeden et al. 2018 with more recent work indicates that several non-lethal methods are more effective in protecting domestic animals than lethal methods appear to be. Note that even though these meta-analyses do not concern only wolves, one can learn from studies of the effectiveness of interventions against other predators to draw inference about how these would work against wolves.

Among those non-lethal methods shown to be more effective and studied multiple times are fladry and livestock-guarding dogs when deployed and maintained correctly as explained in the source articles reviewed. The studies also endorse effective fencing albeit fewer RCTs have evaluated it.

References and notes for section 2 only (in addition to those cited above)


Most of the cited articles in this letter are available for free download at the following institutional server without pass wall or paywall: http://faculty.nelson.wisc.edu/treves/archive_BAS/Pubs_for_public_comments.zip. Failure to include these documents in the administrative public record is unlawful.


Other promising methods tested by RCT only once against wolves includes range riding using low-stress livestock handling. Likewise, I recommend consideration of methods that proved effective from RCT studies with other predators, despite never having been tested on wolves, such as painted eyespots on livestock and humans assisted by deterrent noise-makers and dogs.


Recent meta-analyses of lethal methods against predators of wild ungulates suggests unpredictable outcomes. Although it did not focus on wolves, many of the included studies were of killing wolves to protect wild ungulates. The authors below review the many decades of research on this question and note the shortage of RCTs or the poor quality of controlled studies used to address the question. Therefore, I see no scientific consensus on the effectiveness of killing wolves to protect wild ungulates. I do see consensus on a value-based issue relating to evidence; namely that the design of predator-killing programs should be treated as experiments and monitored scientifically by independent uninterested parties. A recent study in Alberta, Canada also shows that "increasing large-predator populations do not necessarily reduce
hunter harvest of elk” and that sustainable hunting of elk has continued, and populations have increased with increasing large predator populations (Trump et al. 2022). Another recent study analyzing 4 decades of efforts in Alaska, US to reduce abundance of large predators, including gray wolves, brown and black bears, found: (1) no positive correlations between killing of bears and subsequent moose hunting, (2) moose hunting was negatively correlated with the prior year’s wolf killing (weak relationship) and (3) no differences in mean moose hunting during periods of recent liberalized killing of predators relative to prior periods (Miller et al. 2022).


Even when interventions to protect other animals from wolves are effective, there is little consensus on how long effects persist. As the authors below have shown, virtually all studies are one grazing season or briefer. Few interventions have been studied long-term. One has to look at single studies to understand the likely short-term and long-term effects of an intervention, but then I caution the results apply only to that design and experimental set up.


3. Estimating wolf abundance with methods other than validated mark-recapture methods
Recent work by Creel (2021) and Treves & Santiago-Ávila (2023) points out the many scientific shortcomings in the recent approaches taken by the states of Idaho, Montana, and Wisconsin to estimate their statewide wolf abundances. Before shortcuts can be safely taken to estimate wolf abundance accurately, precisely, reproducibly, and with sensitivity to changing conditions, the new methods should be validated by third-party, independent scientists comparing new methods to old methods. Thus far, no shortcut to mark-recapture methods has proven reliable.


4. Killing for tolerance
Finally, the claim that killing wolves improves public tolerance has failed multiple tests by multiple lead authors using different datasets and entirely different approaches to the question (social scientific, population ecology, wolf survival). Unfortunately, I am a co-author on almost every study, so opponents may cite non-independence of the studies. That assumes I can
somehow persuade or compel other scientists to do my bidding. That is ridiculous on its face but is probably best disputed by pointing to the diversity and independence of these authors compared to the homogeneity of the opposing side.

Moreover, disputing the independence of the many studies refuting the idea that killing improves tolerance also ignores the Nordic studies that do not involve me (see below) and ignores the weakness of evidence that killing improves tolerance. The contrasting views are either not peer-reviewed, have not been replicated, have shortcomings that have been exposed in subsequent peer-reviewed work, or did not address the question.

Furthermore, some advocates for killing for tolerance have pivoted to claiming killing wolves opens a space for dialogue with a few powerful, narrow interests. Setting aside the ethics of killing wolves for that purpose, the evidence from Hogberg et al. 2015 does not support the claim because the prime target demographic group in Wisconsin (men residing in wolf range who have familiarity with hunting) had the sharpest decline in tolerance for wolves after wolf-killing was liberalized to include public hunting and trapping.

In summary

- Attitudes to wolves became more negative or did not improve when protections for wolves were reduced [1-5].
- Poaching was higher when wolf protections were reduced, measured by individual wolf survival rates [6-9].
- Poaching was higher when wolf protections were reduced, measured by wolf population dynamics [10-13]. Attempts were made to challenge the latter results, all of which failed because they lacked data to support their claims [14, 15] or had shortcomings that made them irreproducible [16-23]. Our calls for corrections of the latter studies have yielded one correction thus far [24-26]. We await further corrections and retractions.
- Poaching is the major cause of wolf mortality and it is mismeasured or under-reported by agencies because of cryptic poaching [27-29].
- Miscellaneous work on poaching and the effects of lethal management, to guide more or better enforcement and also more effective management policies: [24, 28-38].
- Note a debate in the Nordic countries remains unresolved pending sharing of all data and transparent debate about statistical methods. The Scandinavian analysis that claimed that liberalizing wolf-killing will reduce wolf-poaching has been challenged [39, 40]. Similarly, two papers by the same pair of authors studying the Finnish wolf population concluded, “We conclude that tolerance for carnivores cannot be promoted by legal hunting alone…” [41] but in 2018, they concluded, “Our results provided evidence that poaching is a matter of local intolerance toward wolves and that the problem is mainly related to wolf hunting.” [42] but suggests legal killing pre-empted illegal killing by removing wolves that would have been poached [41, 42]. Clearly, further investigation with accepted methods is warranted.

References for section 4 only

7. Santiago-Ávila, F.J., R.J. Chappell, and A. Treves, Liberalizing the killing of endangered wolves was associated with more disappearances of collared individuals in Wisconsin, USA. Scientific Reports, 2020. 10:pp13881. /10.1038. | https://doi.org/10.1038/s41598-020-70837-x.


Most of the cited articles in this letter are available for free download at the following institutional server without pass wall or paywall: http://faculty.nelson.wisc.edu/treves/archive_BAS/Pubs_for_public_comments.zip. Failure to include these documents in the administrative public record is unlawful.
5. How research integrity influences the quality of science

Certain principles of research integrity affect the quality of science so strongly that failure to adhere to minimum standards of scientific integrity result in unreliable, irreproducible, or fabricated scientific findings. The US federal government has repeatedly tightened regulations and policies to avoid breaches of research integrity, so I will restrict myself here to point out the most common and problematic research misconduct that would render policy based on such science unsound and vulnerable to legal jeopardy.

- **Falsification, fabrication, or plagiarism is illegal for any recipient of federal monies**, which could place state and tribal wildlife agencies in legal jeopardy if they produce such work, because of federal support for fish and wildlife agencies. Financial penalties may be due upon conviction. A portion off such penalties can be won by whistle-blowers who report research misconduct (see [https://www.whitehouse.gov/ostp/](https://www.whitehouse.gov/ostp/) for federal regulations and policies including relevant Congressional Acts).

- **Non-disclosures of potentially competing interests, whether financial or non-financial.** Authors of scientific work whose institutional affiliations are state or tribal wildlife agencies should beware of this breach because it can lead to correction or retraction of scientific articles, following the Committee on Publication Ethics guidelines on transparent disclosures [https://publicationethics.org/](https://publicationethics.org/). Similar recommendations apply to US government agencies NAS National Academies of Sciences EM. 2017. Fostering integrity in research. Washington, DC: The National Academies Press. & Biden JR. Order PE (Ed). 2021. Memorandum on restoring trust in government through scientific integrity and evidence-


- A general lack of transparency, failure to share data, failure to disclose assumptions, methods, value judgments, and failure to embrace open, independent review can all affect the reliability of science. Indeed, articles published by the Wildlife Society before 2022 were published without the ethical guardrails of the Committee on Publication Ethics, COPE (Krausman PR. 2022. Improving the journal of wildlife management: A response to the perspectives of Johnson et al. (2021) and Gould et al. (2021). The Journal of Wildlife Management 86: e22167. https://doi.org/10.1002/jwmg.22167), signifying that such articles should be considered less reliable until proven each article followed COPE best practices on its own. This also means the Journal of Wildlife Management is considered less reliable than journals or publishers that have been signatories to COPE for longer.

- Finally, government agencies in the USA are trustees of nature including wildlife. They are trustees for current and future generations, so their duty is to preserve first and regulate current uses second, to prevent substantial impairment of the trust assets. Catering to a subset of the public or interest groups is unlawful. Failure to regulate illegal uses or over-use is likewise unlawful.

References for section 5 only


Open Science Collaboration. 2015. Reproducibility project: Psychology. OSF: 10.17605/OSF.IO/EZCUJ.


I have submitted an archive with the following scientific publications also stored permanently at the URL in the footer to this document


- Treves A, MichaelppNelson PD, Ruth H. Spaniol Chair of Renewable Resources and Professor, Oregon State University, Jonathan Way PD, Eastern Coyote Research and Clark University (Worcester, MA), Guillaume Chapron PD, Associate Professor, Swedish University of Agricultural Sciences, et al. 2014. Memo in support of minnesota bill sf2256 and specifically pertaining to “modifying wolf management provisions”.
- Treves A and scholars asa. 2016. Petition to: Secretary sally jewell and secretary penny pritzker: Implement the endangered species act using the best available science. Union of Concerned Scientists.

Most of the cited articles in this letter are available for free download at the following institutional server without pass wall or paywall: http://faculty.nelson.wisc.edu/treves/archive_BAS/Pubs_for_public_comments.zip. Failure to include these documents in the administrative public record is unlawful.
• Treves A and Menefee L. 2022. Adverse effects of hunting with hounds on participants and bystanders. Biorxiv, https://www.biorxiv.org/content/10.1101/2022.08.16.504031v2


• . 2013. Evaluating the scientific soundness of plans for harvesting wolves to manage depredations in michigan. Little River Band of Ottawa Indians Natural Resources Report.

Most of the cited articles in this letter are available for free download at the following institutional server without pass wall or paywall: http://faculty.nelson.wisc.edu/treves/archive_BAS/Pubs_for_public_comments.zip. Failure to include these documents in the administrative public record is unlawful.