Don’t judge the roar by its echo:

Tests of assumptions, tools and policies for human-carnivore coexistence in North America

By

Naomi X. Louchouarn

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Environment & Resources)

At

THE UNIVERSITY OF WISCONSIN – MADISON

2023

Date of final oral examination: 07/13/2023

The dissertation is approved by the following members of the Final Oral Committee:

Adrian Treves, Professor, Environment & Resources

Zuzana Burivalova, Assistant Professor, Forest and Wildlife Ecology

Holly Gibbs, Professor, Geography

Francisco Santiago-Ávila, Conservation Scientist, Project Coyote, The Rewilding Institute

"Humans have the least experience with how to live and thus the most to learn – we must look to our teachers among the other species for guidance"

- Robin Wall Kimmerer, Braiding Sweetgrass

To my dog, Charlie, you taught me to see the person in non-human animals. You kept me company during long field seasons. You will always be my best dog and my best coauthor.

To my husband, Zach, not sure I would have been able to do any of this without your support, your constant interest, and your willingness to talk through my results ad nauseum.

Table of Contents

Introduction .... *v*
Testing assumptions through best science vii
Objectives x
References x
Chapter 1: Low-stress livestock handling protects cattle in a five-predator habitat *1*
Introduction 2
Materials & Methods 7
Results 15
Discussion 17
Co-benefits 19
Attacks on livestock 20
Wolves 22
Conclusion 24
Tables & Figures 26
References 31
Supplementary Materials for Low-stress livestock handling protects cattle in a five-predator habitat .. 38
R-Code 42
Chapter 2: Evaluating how lethal management affects poaching of Mexican wolves *45*
Introduction 46
Methods 52
Data Collection and Preparation 52
Statistical Methods 56
Diagnostic Step 60
Analyses 60
Results 62
Discussion 64
Implications for endangered species 75
Tables & Figures 79
Data Accessibility 90
Acknowledgments 90
References 91
Chapter 3: Politics and Policy play different roles in predicting Cryptic and reported wolf poaching in Michigan, USA *97*
Introduction 99
Hypotheses and Study Objectives 102
Results 120

Lost-to-follow-up 120
Reported Poaching 122
Nonhuman 123
Discussion 124
Conclusion 131
Tables & Figures 133
References 150
Supplemental Materials 154
Conclusion 170
Randomized-control design 171
Correlation analysis 174
Time-to-event analyses 174
Overall conclusion 176

Introduction

Conservation is a problem-driven field made up of both scientific and practical elements. The field emerged from a need to reverse ecological damage that was occurring at unprecedented rates after the advent of the industrial revolution (Organ *et al.*, 2012b; Soulé, 1985). In democratic jurisdictions across the world, such as the US and Canada, ecological conservation is considered a duty, and natural resources are held in public trust to maintain ecological health in perpetuity (Treves *et al.*, 2017b). However, the need for such actions is driven by human behaviors and natural resource uses, and the why, where, how and when of conservation actions are not immune to history, prejudice and biases. For example, the guiding ethos of North American conservation is the North American Model of Wildlife management, which arose from concerns by wealthy sport hunters and naturalists who worried that game animals and fish would be eradicated by uncontrolled use (Organ *et al.*, 2012a). Therefore, the underlying principle of the approach is sustainable use and hunting culture, not science.

Though the North American model has led to conservation successes for many species, and funding from hunting and fishing licenses has sustained conservation actions for non-hunted species (Organ *et al.*, 2012a; Treves and Martin, 2011), this management approach becomes an issue when the ecological problems in question require conservation of species which compete with humans for natural resources or use of landscapes, such as large carnivorous mammals. Large carnivores, like grey wolves (*Canis lupus)* and grizzly bears (*Ursus arctos*), were not historical game species, and thus these species were not considered worth conserving by the founders of the North American model of wildlife management. Theodore Roosevelt went so far as to call wolves a “waste” (Kellert *et al.*, 1996). Due to this antagonistic history, these species were nearly extinct in Southern Canada and the lower 48 states of the US by the mid-20th century

(Treves *et al.*, 2017b). However, shifts in philosophy and ideas about conservation began in the 1960s, leading to federal protections under the Endangered Species Act (ESA) for many species, including wolves and grizzlies in 1973 (Bruskotter, Schmidt and Teel, 2007). Public attitudes towards wolves and other large carnivores have appeared to follow changes in policy, becoming more positive and protection oriented over time (Bruskotter, Schmidt and Teel, 2007; Kellert *et al.*, 1996; Williams, Ericsson and Heberlein, 2002). Though some interest groups, such as hunters and ranchers, often have less positive attitudes towards wolves in particular (Bruskotter, Schmidt and Teel, 2007; Kellert *et al.*, 1996; Treves and Martin, 2011; Treves, Naughton-Treves and Shelley, 2013; Williams, Ericsson and Heberlein, 2002).

Given the history of large carnivore management before their protection under the ESA, it is unsurprising that hunters appear unwilling to steward large carnivore species in the same way they have for large ungulates, like elk (Treves, 2009; Treves and Martin, 2011). The objective of laws like the ESA is to recover endangered species, allowing the federal government to relinquish management to localized agencies at the state or tribal level. Given the underlying principle of sustainable use guiding North American wildlife management, delisting of these species almost always immediately results in lethal control by state agencies, either through public hunts or government action (Morell, 2008; Treves and Louchouarn, 2022; Vucetich *et al.*, 2017). The assumption being that lethal control is necessary for successful coexistence and conservation, despite a lack of understanding regarding whether lethal control truly improves coexistence with large carnivores, and a history of persecution and extirpation through overhunting.

Testing assumptions through best science

A primary mandate of both the ESA and the North American model of wildlife management is that management and policy be based on ‘the best available science ’(Artelle *et al.*, 2018). This ‘best science ’should test the underlying assumptions of conservation actions, allowing adaptation of these actions to better align with emerging information. However, what makes the ‘best ’science is not defined in the law, and there is no clear guidance on what is required of science-based management (Artelle *et al.*, 2018; Treves, 2022). Artelle et al. (2018) found that more than half of 667 wildlife management systems they examined in North America did not meet the criteria of four hallmarks of science-based management: measurable objectives, evidence, transparency, and independent review. Without these important scientific principles, we cannot be certain that assumptions are being properly tested and conservation actions are being adapted to achieve desired outcomes.

Conservation science, and carnivore coexistence as a subfield, is an emerging field, and has evolved in the past several decades. Many conservation researchers have begun to identify and test the underlying assumptions regarding large carnivore conflict mitigation using higher standards of experimentation (Treves, Krofel and McManus, 2016; Treves *et al.*, 2019; Van Eeden *et al.*, 2018). Two assumptions exist regarding policies which allow lethal control of large carnivores: 1. That lethal control will reduce conflict, and 2. That lethal control will improve tolerance, i.e., coexistence. Therefore, conservation scientists must design experiments to test both assumptions. There is a particular need for high standards of experimentation given that lethal control could undermine recent conservation successes, cost tax-payer money, and is ethically dubious (Wallach *et al.*, 2020).

Regarding conflict mitigation, the objective is to ensure that large carnivores are deterred from attacking or damaging human resources, such as livestock. The highest standard experiments examining the effectiveness of lethal control suggest that it may be counter-productive, ineffective or inconsistent (Van Eeden *et al.*, 2018; Treves, Krofel and McManus, 2016; Treves *et al.*, 2019; Santiago-Avila, Cornman and Treves, 2018). However, no randomized-control trials, likely the highest standard of experimentation available, have been conducted on lethal control of large carnivores (Van Eeden *et al.*, 2018; Treves, Krofel and McManus, 2016; Treves *et al.*, 2019; Santiago-Avila, Cornman and Treves, 2018). Non-lethal methods of predator deterrence do exist, and high standard experiments using randomized control trials of non-lethal methods have found that many such tools are more often effective than lethal methods (Treves, Krofel and McManus, 2016). However, more of these methods exist, such as night penning, light and sound deterrents, husbandry and guardian animal use, and fencing and flagging. Not all methods have been tested with high standards of experimentation, and there is a great deal yet to learn regarding level and time of effectiveness (Van Eeden *et al.*, 2018). Regardless, consensus is emerging that lethal methods may not be the best form of conflict mitigation and may have the opposing effect (Elbroch and Treves, 2023).

The secondary assumption of tolerance for large carnivores is an important consideration because the highest risk of mortality for large carnivores comes from human-causes, usually when these species leave protected areas (Woodroffe and Ginsberg, 1998; Ripple *et al.*, 2014). Mortality from illegal killing of carnivores is of particular concern given the difficulty in estimating this risk, and its poorly understood causal link to human intolerance (Bruskotter and Wilson, 2014; Treves and Bruskotter, 2014). Human intolerance is generally measured through self-reporting and surveying, but whether self-reports of individuals stating an intolerance for carnivores translates to actual increased poaching is unknown (Browne-Nuñez *et al.*, 2015; Bruskotter, Schmidt and Teel, 2007; Ericsson and Heberlein, 2003; Hogberg *et al.*, 2016; Treves, Naughton-Treves and Shelley, 2013; Williams, Ericsson and Heberlein, 2002). The management agencies tend to assume a direct relationship, and suggest that lethal control increases tolerance by showing direct action and reducing conflict (Elbroch and Treves, 2023; Naughton-Treves, Grossberg and Treves, 2003).

However, evidence from studies on wolves in the Midwest and Scandinavia suggest this is not the case. In Wisconsin, longitudinal analyses with repeated surveys of the same respondents before and after wolf hunts found that individuals self-reported as less tolerant of wolves after wolf hunts than they had been before (Treves, Naughton-Treves and Shelley, 2013). Another study examining inclination to poach and support of lethal control by hunters and farmers in Wisconsin found that inclination to poach did not decrease after the implementation of a wolf hunting season (Browne-Nuñez *et al.*, 2015). Support for lethal control also remained high. A study in Finland found that wolf poaching appeared correlated to local intolerance towards wolves, but increasing legal hunting did not necessarily decrease total killing, instead converting what would have been illegal killing into legal killing (Suutarinen and Kojola, 2017). The authors therefore suggest legal hunting is not a solution to improved tolerance.

Studies like Suutarinen & Kojola (2017), and (Chapron and Treves, 2016; Santiago-Ávila, Chappell and Treves, 2020) go beyond examining simple stated tolerance and examined actual population growth rates or individual wolf mortality risk when protection policies change. Santiago-Ávila et al (2020) showed that in a population of wolves in Wisconsin, risk of reported poaching of individual wolves decreased during periods when wolf protections were relaxed, but disappearances of wolves increased. Wolf disappearances can be a result of three primary causes:

migration out of range, collar failure, and cryptic poaching in which evidence is destroyed (Liberg *et al.*, 2012; Treves *et al.*, 2017a). Therefore, changes in wolf disappearance can give us insight into changes in cryptic poaching, even if we cannot observe the exact proportion of disappearances that were cryptically poached.

Objectives

There is a need to use high standards of experimentation to examine the assumptions underlying carnivore management in North America. North American carnivore conservation follows the North American model of wildlife management, and therefore prioritizes hunting over other forms of management. We must therefore clearly understand the risks of this method, given that carnivores have outsized impacts on ecosystem health (Estes *et al.*, 2011; Ripple *et al.*, 2014; Ripple *et al.*, 2022). If lethal control does not reduce conflict, are there non-lethal methods better suited to conflict mitigation? Do reductions in protections of carnivores lead to reductions in poaching, and if so, which protections? Or are other variables more predictive of poaching risk? Poaching is a difficult behavior to understand and mitigate, therefore the more we understand about its predictors the better we can adapt conservation actions meant to mitigate it. In this dissertation I examine these assumptions and the tools and policies used to reduce conflict with and poaching of large carnivores in North America.

References

Artelle, K. A., Reynolds, J. D., Treves, A., Walsh, J. C., Paquet, P. C. and Darimont, C. T. (2018) 'Hallmarks of science missing from North American wildlife management', *Science Advances,* 4(3), pp. eaao0167.

Browne-Nuñez, C., Treves, A., MacFarland, D., Voyles, Z. and Turng, C. (2015) 'Tolerance of wolves in Wisconsin: a mixed-methods examination of policy effects on attitudes and behavioral inclinations', *Biological Conservation,* 189, pp. 59-71.

Bruskotter, J. T., Schmidt, R. H. and Teel, T. L. (2007) 'Are attitudes toward wolves changing? A case study in Utah', *Biological conservation,* 139(1-2), pp. 211-218.

Bruskotter, J. T. and Wilson, R. S. (2014) 'Determining where the wild things will be: using psychological theory to find tolerance for large carnivores', *Conservation Letters,* 7(3), pp. 158-165.

Chapron, G. and Treves, A. (2016) 'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore', *Proceedings of the Royal Society B: Biological Sciences,* 283(1830), pp. 20152939.

Elbroch, L. M. and Treves, A. (2023) 'Perspective: Why might removing carnivores maintain or increase risks for domestic animals?', *Biological Conservation,* 283, pp. 110106.

Ericsson, G. and Heberlein, T. A. (2003) 'Attitudes of hunters, locals, and the general public in Sweden now that the wolves are back', *Biological conservation,* 111(2), pp. 149-159.

Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D. and Jackson, J. B. (2011) 'Trophic downgrading of planet Earth', *science,* 333(6040), pp. 301-306.

Hogberg, J., Treves, A., Shaw, B. and Naughton-Treves, L. (2016) 'Changes in attitudes toward wolves before and after an inaugural public hunting and trapping season: early evidence from Wisconsin's wolf range', *Environmental Conservation,* 43(1), pp. 45-55.

Kellert, S. R., Black, M., Rush, C. R. and Bath, A. J. (1996) 'Human culture and large carnivore conservation in North America', *Conservation Biology,* 10(4), pp. 977-990.

Liberg, O., Chapron, G., Wabakken, P., Pedersen, H. C., Hobbs, N. T. and Sand, H. (2012) 'Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe', *Proceedings of the Royal Society B: Biological Sciences,* 279(1730), pp. 910-915.

Morell, V. 2008. Wolves at the door of a more dangerous world. American Association for the Advancement of Science.

Naughton-Treves, L., Grossberg, R. and Treves, A. (2003) 'Paying for Tolerance: Rural Citizens' Attitudes toward Wolf Depredation and Compensation', *Conservation Biology,* 17(6), pp. 1500-1511.

Organ, J., Geist, V., Mahoney, S., Williams, S., Krausman, P., Batcheller, G., Decker, T., Carmichael, R., Nanjappa, P. and Regan, R. (2012a) 'The North American Model of Wildlife Conservation', *The Wildlife Society Technical Review*, pp. 12-04.

Organ, J. F., Geist, V., Mahoney, S. P., Williams, S., Krausman, P. R., Batcheller, G., Decker, T., Carmichael, R., Nanjappa, P. and Regan, R. (2012b) 'The North American model of wildlife conservation', *The Wildlife Society Technical Review,* 12(04).

Ripple, W. J., Estes, J. A., Beschta, R. L., Wilmers, C. C., Ritchie, E. G., Hebblewhite, M., Berger, J., Elmhagen, B., Letnic, M. and Nelson, M. P. (2014) 'Status and ecological effects of the world's largest carnivores', *Science,* 343(6167).

Ripple, W. J., Wolf, C., Phillips, M. K., Beschta, R. L., Vucetich, J. A., Kauffman, J. B., Law, B. E., Wirsing, A. J., Lambert, J. E. and Leslie, E. (2022) 'Rewilding the American West', *BioScience*.

Santiago-Ávila, F. J., Chappell, R. J. and Treves, A. (2020) 'Liberalizing the killing of endangered wolves was associated with more disappearances of collared individuals in Wisconsin, USA', *Scientific reports,* 10(1), pp. 1-14.

Santiago-Avila, F. J., Cornman, A. M. and Treves, A. (2018) 'Killing wolves to prevent predation on livestock may protect one farm but harm neighbors', *PLoS One,* 13(1), pp. e0189729.

Soulé, M. E. (1985) 'What is conservation biology?', *BioScience,* 35(11), pp. 727-734.

Suutarinen, J. and Kojola, I. (2017) 'Poaching regulates the legally hunted wolf population in Finland', *Biological Conservation,* 215, pp. 11-18.

Treves, A. (2009) 'Hunting for large carnivore conservation', *Journal of Applied Ecology,* 46(6), pp. 1350-1356.

Treves, A. (2022) '"Best available science" and the reproducibility crisis', *Frontiers in Ecology and the Environment,* 20(9), pp. 495-495.

Treves, A., Artelle, K. A., Darimont, C. T. and Parsons, D. R. (2017a) 'Mismeasured mortality: correcting estimates of wolf poaching in the United States', *Journal of Mammalogy,* 98(5), pp. 1256-1264.

Treves, A. and Bruskotter, J. (2014) 'Tolerance for predatory wildlife', *Science,* 344(6183), pp. 476-477.

Treves, A., Chapron, G., López-Bao, J. V., Shoemaker, C., Goeckner, A. R. and Bruskotter, J. T. (2017b) 'Predators and the public trust', *Biological Reviews,* 92(1), pp. 248-270.

Treves, A., Krofel, M. and McManus, J. (2016) 'Predator control should not be a shot in the dark', *Frontiers in Ecology and the Environment,* 14(7), pp. 380-388.

Treves, A., Krofel, M., Ohrens, O. and van Eeden, L. M. (2019) 'Predator control needs a standard of unbiased randomized experiments with cross-over design', *Frontiers in Ecology and Evolution,* 7, pp. 462.

Treves, A. and Louchouarn, N. X. (2022) 'Uncertainty and precaution in hunting wolves twice in a year', *PloS one,* 17(3), pp. e0259604.

Treves, A. and Martin, K. A. (2011) 'Hunters as stewards of wolves in Wisconsin and the Northern Rocky Mountains, USA', *Society & Natural Resources,* 24(9), pp. 984-994.

Treves, A., Naughton-Treves, L. and Shelley, V. (2013) 'Longitudinal analysis of attitudes toward wolves', *Conservation Biology,* 27(2), pp. 315-323.

Van Eeden, L. M., Eklund, A., Miller, J. R., López-Bao, J. V., Chapron, G., Cejtin, M. R., Crowther, M. S., Dickman, C. R., Frank, J. and Krofel, M. (2018) 'Carnivore conservation needs evidence-based livestock protection', *PLoS biology,* 16(9), pp. e2005577.

Vucetich, J. A., Bruskotter, J. T., Nelson, M. P., Peterson, R. O. and Bump, J. K. (2017) 'Evaluating the principles of wildlife conservation: a case study of wolf (Canis lupus) hunting in Michigan, United States', *Journal of Mammalogy,* 98(1), pp. 53-64.

Wallach, A. D., Batavia, C., Bekoff, M., Alexander, S., Baker, L., Ben-Ami, D., Boronyak, L., Cardilin, A. P., Carmel, Y. and Celermajer, D. (2020) 'Recognizing animal personhood in compassionate conservation', *Conservation Biology,* 34(5), pp. 1097-1106.

Williams, C. K., Ericsson, G. and Heberlein, T. A. (2002) 'A quantitative summary of attitudes toward wolves and their reintroduction (1972-2000)', *Wildlife Society Bulletin*, pp. 575-584.

Woodroffe, R. and Ginsberg, J. R. (1998) 'Edge Effects and the Extinction of Populations Inside Protected Areas.', *Science*, 280, pp. 2126-2128. Available at: http://science.sciencemag.org/content/280/5372/2126.abstract.

Chapter 1: Low-stress livestock handling protects cattle in a five-predator habitat

Naomi X Louchouarn[1]*, Adrian Treves[1]

[1]Nelson Institute of Environmental Studies, University of Wisconsin-Madison, Madison, Wisconsin, USA

*Corresponding author: Naomi Louchouarn

Email Address: louchouarn@wisc.edu

Abstract

Given the ecological importance of top predators, societies are turning to non-lethal methods for coexistence. Coexistence is challenging when livestock graze within wild predator habitats. We report a randomized, controlled experiment to evaluate low- stress livestock handling (L-SLH), a form of range riding, to deter grizzly (brown) bears, gray wolves, cougars, black bears, and coyotes in Southwestern Alberta. The treatment condition was supervision by two newly hired and trained range riders and an experienced L-SLH-practicing range rider. This treatment was compared against a baseline pseudo-control condition of the experienced range rider working alone. Cattle experienced zero injuries or deaths in either condition. We infer that inexperienced range riders trained and supervised by an experienced rider did not raise or lower the risk to cattle. Also, predators did not shift to the cattle herds protected by fewer range riders. We found a correlation suggesting grizzly bears avoided herds visited more frequently by range riders practicing L-SLH. More research is required to compare different forms of range riding. However, pending experimental evaluation of other designs, we recommend use of L-SLH. We discuss the cobenefits of this husbandry method.

Keywords: carnivores, human-wildlife coexistence and conflict, randomized, controlled trials, non-lethal methods, predator control

Introduction

Given the important role of top predators in the function and diversity of ecosystems, societies and governments are prioritizing co-existence (Ripple et al., 2014). Human- induced mortality is the dominant cause of mortality for large carnivores across the world. It has resulted in ecosystem degradation (Estes et al., 2011), and extinctions of many populations (Ripple et al., 2014; Estes et al., 2011). Success in preserving carnivore populations depends on converting competition over land and resources from lethal to non-lethal (Ripple et al., 2014; Woodroffe & Ginsberg, 1998; Treves et al., 2017; Van Eeden et al., 2018; Boronyak et al., 2021). Coexistence with bears and wolves in North America is a timely challenge given the protected status of these populations in some jurisdictions and the societal support for livestock grazing on public lands. As these and other carnivores have recolonized their historic ranges they encounter infrequently supervised free-ranging livestock, leading to conflict that can either spark a renewal of eradication campaigns against top predators (e.g., Williams, 2022) or innovative coexistence strategies.

In the United States, agriculture and related industries contribute 1.1 trillion US dollars to the national gross domestic product (GDP), and the highest value livestock sector is cattle production (USDA Economic Research Service, 2021). In Canada, cattle production alone contributes more than 5 billion Canadian dollars to Canada’s GDP and tens of thousands of jobs for the province of Alberta (Lee et al., 2017). Lee et al. (2017) found that more than 60% of Alberta’s beef owners claimed to have lost animals to carnivores. Methods to reduce conflict include education and attractant mitigation programs (e.g., BearSafe program; Alberta

Environment and Parks, 2016) and non-lethal and lethal predator control (e.g., relocations, aversive conditioning and targeted trapping and killing; Treves, Krofel & McManus, 2016; Young, Hammill & Breck, 2019; Khorozyan & Waltert, 2020). However, many of the methods used to deter large carnivores have either never been experimentally tested or have limited supporting data, therefore their users assume effectiveness (Van Eeden et al., 2018).

Livestock owners often perceive that non-lethal methods of predator deterrence are less effective than lethal methods (Scasta, Stam & Windh, 2017). However, recent independent research covering various carnivore species has raised doubts about the effectiveness of killing individual predators. Although there is agreement that predation vanishes when no predators exist, the quantitative relationships between predatory threats to people or property and key environmental variables, such as domestic and wild species ecologies, remain murky. Reviews examining predator removal show that removal efforts are rarely successful, even when efforts are directed at a specific individual predator blamed for property damages (Linnell et al., 1999; Odden et al., 2002; Treves & Naughton-Treves, 2005; Treves, 2009; Treves et al., 2019; Lennox et al., 2018). Multiple studies have found that lethal control of predators often has the counter-productive effect of raising livestock losses or has no effect on losses, e.g., wolves in Michigan, USA (Santiago-Avila, Cornman & Treves, 2018) and in France (Grente et al., 2021), and dingoes in Australia (Wallach et al., 2009). These findings have spurred many independent efforts to find other approaches. In some cases, the partnership between scientists and livestock owners can lead to the transfer and dissemination of scientifically supported innovations (Van Eeden et al., 2018; Ohrens, Bonacic & Treves, 2019; Khorozyan & Waltert, 2021; Khorozyan et al., 2020; Radford et al., 2020).

The effectiveness of many non-lethal methods has not been evaluated using rigorous scientific experiments (Van Eeden et al., 2018). One such method is range riding, i.e., deploying humans using non-lethal methods of predator deterrence and livestock protection. Range riding has two primary elements: the amount of human presence and the behaviors of the humans in the field. Increased human presence among livestock is assumed to deter predators and improve response time if predators are present (Bangs et al., 2006). But how much human presence is effective is still unknown. Furthermore, the most effective behaviors of range riders are not well understood. Indeed range riding is not well-defined and seems to be practiced in myriad forms, each likely to have varying degrees of effectiveness (Parks & Messmer, 2016; Jablonski et al., 2020). The most commonly used forms of range riding are narrowly predator-focused where riders generally hired by government agencies focus on detection and deterrence of predators (Parks & Messmer, 2016; Wilson, Bradley & Neudecker, 2017). Alternatively, riders may focus more on livestock vulnerability and herding practices to foster anti-predator behavior in livestock (Parks & Messmer, 2016; Bruns, Waltert & Khorozyan, 2020). Although the latter involves deterrence of predators, search for predators is lower priority compared to the concentration of effort on livestock behavior, health, and safety. These varied practices are largely driven by anecdotal experience without the benefit of empirical data and are therefore not likely to be equally effective.

We define a particular form of range riding known as low-stress livestock handling (L-SLH) which has been developed among a relatively small group of livestock owners in the North American West (Fig. 1) (Hibbard, 2012). We also examine how increased presence of humans using the method with different levels of experience works to deter predators.

Low-stress livestock handling as predator deterrent – Bud Williams and Temple Grandin first developed L-SLH to reduce stress and improve livestock health. They combined ‘‘pressure and release ’’herding, a form of interacting with livestock that takes advantage of livestock prey responses, to move livestock in a way that both enhances the choices and natural behaviors of the individual animal (Hibbard, 2012; Grandin, 1989). With this technique, handlers move calmly towards livestock, coming into contact with the animal's 'pressure zone', i.e., the distance at which the animal will respond to the handler's presence, and uses this contact to gently move the animal according to the livestock's instinctual responses. As ungulates generally prefer to move as a herd, handlers push individual animals towards the herd, and allow the herd to move together at a calm, yet steady pace. These techniques apparently improve livestock stress, health and yield (Fig. 1) (Hibbard, 2012; Grandin, 1989; Barnes, 2015). This combination of techniques creates a positive association between human actions and herding, which helps make livestock more willing to remain as a herd relative to those who are aggressively handled and therefore associate herding with stress (Hibbard, 2012; Barnes, 2015). Conventional handling generally does not consider the instinctual pressure zones of the animals, and therefore forces animals together in an uncomfortable and rapid way. This can increase stress and produce a tendency towards flightiness (Hibbard, 2012).

Livestock owners who have used the method in regions with high predator numbers report that their livestock behave similarly to wild ungulate herds, which may reduce vulnerability to wild predation (Zaranek, 2016; Mech & Peterson, 2010). Therefore, we hypothesize that L-SLH may deter predator attacks and reduce predation by encouraging natural herding instincts that reduce ungulate vulnerability (Zaranek, 2016). This method may be particularly useful on extensive public lands, where other forms of deterrence may be difficult to implement (Eklund et al., 2017; Stone et al., 2017). Though Barnes (2015) described a quasi-experimental evaluation of L-SLH for livestock herding, it has never, to our knowledge, been experimentally investigated as a form of predator deterrence.

Here we present the first experimentally evaluation of any form of range riding and define a few of the parameters that are important to L-SLH as a non-lethal cattle protection method. We hypothesize that range riders might deter predators from cattle by two primary mechanisms: First, the presence of more humans might deter large carnivores such as grizzly bears (Ursus arctos), black bears (Ursus americanus), wolves (Canis lupus), cougars (Puma concolor) and coyotes (Canis latrans; hereafter LC, for Large Carnivores) from the associated cattle. This predicts that the cattle guarded by a single range rider in our study (pseudo-control) would be more vulnerable than the cattle guarded by several range riders (treatment). The comparison in this study is therefore between the presence of a single rider acting as a baseline (which we define as a pseudo-control), relative to the increased human presence of multiple new riders. Alternatively, a second mechanism might be that herd stress levels predict vulnerability to predation because L-SLH would encourage and reinforce herding behaviors that reduce the risks posed by LCs. This alternative hypothesis predicts that the number of range riders is irrelevant and both pseudo-control and treatment would be effective. Also, the experience level of the range riders might affect the stress and hence vulnerability of the cattle. Accordingly, our experiment was designed to reveal if, counter-intuitively, LCs approached herds with a greater number of less experienced range riders more often, because cattle would be more stressed than cattle exposed to a single experienced range rider. Finally, a frequent unsubstantiated claim about non-lethal methods of predator control is that predators will shift to less-protected neighbors. Our design can detect this effect if herds frequently visited by multiple range riders (treatment) were visited less often by LCs than those supervised by a single range rider (pseudo-control).

Materials & Methods

Study area - We conducted this study on the Spruce Ranching Co-op (hereafter the Co-op), a grazing area used by 38 permitted cattle owners who collectively bring about 2,000 cow- calf pairs and 500 pregnant heifers in June of each year (Fig. 2). The Co-op is located on 22,500 acres (91 km2) of Alberta provincial lease land in the Pekisko Heritage Rangelands area, which is part of the foothills of the southern Canadian Rocky Mountains south of the Banff-Jasper-Yoho National Park complex and north of the Waterton Glacier International Peace Park (Fig. 2). The Co-op overlaps two sources of LCs, and therefore represents a core connectivity area for many species (Proctor et al., 2012). Despite its status as provincial lease land, we did not require permits to access the Co-op as we were not handling animals or collecting samples. Further, we used non-invasive observation methods (details below) and therefore received a study exemption from the Institutional Animal Care and Use Committee at the University of Wisconsin-Madison.

Ranch manager - This study required a pseudo-control, i.e., a baseline condition, which we defined as the presence of the ranch manager. We use a pseudo-control instead of a placebo-control because a placebo-control requires the treatment to be compared against the lack of treatment, which in this case amounts to no human supervision. However, in this study a true control situation cannot be created because livestock owners are unwilling to leave their livestock unattended due to known risk of LC attacks on cattle. At least four individual cattle had been injured or killed by LCs in each of the past three years prior to our study. Provincial statistics on cattle predation in our study region, which is comprised of rugged public lands where human supervision is scarce, reveal that 476 animals were confirmed attacked by carnivores between 2015–2019 (Alberta Environment and Parks, 2020). Also, cattle producers systematically under report suspected predation events in Alberta due to the perceived effort involved in reporting losses (Lee et al., 2017). These data provide some confidence that the absence of human supervision would be riskier for cattle.

The ranch manager represents a pseudo-control because: (1) The ranch manager continued the same practices he has employed for the past 20 years. The ranch manager is highly trained in L-SLH as he has practiced these behaviors on this ranch for two decades. He began learning these methods by attending clinics with Whit Hibbard (see introduction), and now travels throughout the US and Canada discussing L-SLH with interested ranchers through workshops presented by The Working Circle (https://www.workingcircle.org/), a California based non-profit. During our study, he visited every herd (i.e., all eight grazing units) at the same rate he has always done without consideration of the treatment (two range riders) schedule; (2) When the ranch manager required extra help in pseudo-control fields (i.e., where range riders could not go), he hired additional help, as he would normally do. The ranch manager therefore saw each herd on average every 9 days depending on the herd and changing conditions related to weather, his own schedule, and availability of hired helpers. Therefore, we compare the pseudo-control of ranch manager to the treatment of ranch manager with two recently trained range riders. We could only indirectly infer whether range riders are more effective than no protection.

Range riders - We hired two range riders (referred to as Rider A and Rider B). Protocols for this study did not require IRB approval as no data about range riders was collected. The ranch manager trained both range riders. Rider A has run their own ranch in the past and had attended L-SLH clinics with Whit Hibbard. Rider A was also familiar with the landscape but had never spent full seasons working with cattle on the Co-op before a short two-month pilot season in 2019. During the 2019 pilot season, the ranch manager trained Rider A for a week, during which time Rider A shadowed the ranch manager and spent supervised time moving cattle and learned to travel efficiently across the Co-op. Rider A also received a few days of training over the past two decades when the ranch manager hired them. Rider B had worked on the Co-op in the past, with about two decades of experience on their own land and on the Co-op. Therefore, Rider B already had some knowledge of how to travel safely and efficiently across the Co-op. However, Rider B had never been trained formally in L-SLH methods.

During our study, Riders A and B received a week of training from the ranch manager, during which time they shadowed the ranch manager and helped move and manage cattle. This implies some potential ‘‘blurring ’’of treatment and pseudo-control as the idiosyncrasies of the ranch manager’s methods might have been transmitted to the two inexperienced riders. However, the ranch manager visited all herds including the treatment herds so his presence and behavior were a background baseline to the entire experiment and every herd (therefore a pseudo-control) and we were investigating the effect of the supplemental riders on half of the herds. Riders were responsible for the same herds, numbering up to 5 herds at a time. Riders worked together as a pair in areas where bears had recently (within the past 2 weeks) been seen, or that were heavily wooded where bears are difficult to detect. Riders worked alone when they deemed there to be no safety advantage from working in pairs. In other words, in fields with more open space. Riders therefore worked in pairs 38% of the time. All treatment herds received visits from riders working alone and in pairs. Rider A helped train Rider B who was less well versed in L-SLH. Rider pairs or single riders saw each treatment herds every 1–3 days.

Low-stress livestock handling techniques - L-SLH practices included keeping track of cattle behaviors during range riders ’visits and encouraging cattle to bunch together when they appeared to be spread too widely across a pasture. Range riders moved cattle as described on Fig. 1. For example, range riders: opened gates between pastures to allow cattle to move freely at their own pace, herded the cattle slowly to allow all cows to pair up with their calves before moving to new pastures, and used temporary pastures between original and destination pastures for up to a week to keep cattle calm throughout the movement process. Therefore, stressful movements were a minority of study period. Range riders rarely, if ever, rode their horses at more than a walk when near or among cattle. Range riders never used more assertive herding behaviors such as elevated voices or swinging arms. Range riders only used roping when doctoring animals, which was only done within a doctoring pen away from other animals. Range riders moved herds on average 3 times over the 123-day study. The ranch manager was present any time cattle were moved between pastures, but riders only participated in treatment herd moves.

When range riders found dead cattle, they contacted Alberta Fish and Game (AFG) to determine the cause of death. When AFG conservation officers implicated an LC in the death, range riders spent more time with the herd, ensuring that predators did not return. If LCs were spotted, range riders or conservation officers used aversive methods (e.g., warning shots, bear bangers and cracker shells). When possible, the ranch manager moved cattle to a new field further from recent LC presence.

Study design **-** Each herd of cattle, which was independently grazed for the duration of the study period, was one ‘subject ’(Ohrens, Bonacic & Treves, 2019). Each subject herd ranged in size from 150 yearlings to 400 cow-calf pairs. Each treatment condition had an equal number of calves due to the cross-over design (Table 1). However, the average number of calves was somewhat higher in treatment condition during phase 1 (Table 1). In phase 2 that reversed and there were slightly more calves in the pseudo-control condition. By then calves would have gained weight, reducing their vulnerability, so we address this potentially confounding effect below. We conducted the experiment in two phases, each lasting half of the 4-month (July–October 2020) study period, during which we randomly assigned herds to receive ‘treatment ’or ‘pseudo-control’. Phase 1 lasted for half of the grazing season (about 2 months; July–August) and then reversed in phase 2 (September–October) for the crossover. Therefore, treatment herds eventually became pseudo-control and vice versa. We imposed a 7-day washout period between phase 1 and phase 2 to reduce any potential carryover effect. We then compared herds to themselves from phase 1 to phase 2 to evaluate the treatment effect. Our within-subjects design minimized potentially confounding variables between herds and randomization avoided selection bias and potentially confounding order effects due to seasonal changes. We studied eight herds (replicates) and therefore had 16 trials in total, including eight treatments and eight pseudo-controls.

The ranch manager split one herd into two separate herds in phase 1 but combined them in phase 2 because of grass conditions. We treated these herds as one subject in the study: they received the same treatment sequence, and their carnivore presence values are summed during the period in which these herds were grazed separately. The herds were sampled separately when they were separated, therefore summed value of carnivore presence is divided by the summed sampling pressure. Data collection

Trail Camera Data – We used 34 unbaited trail cameras to detect LC presence around herds. We monitored each herd with three cameras. To place cameras, we used ArcGIS 10.7 to create a grid of 40-acre (0.18 km2) cells within each grazed pasture and selected three cells at random. Within each of these three cells, we placed cameras in locations deemed likely to be visited by LCs (e.g., cut lines, cow trails, stream banks, etc.). We moved cameras from pasture to pasture in response to herd movements. Each pasture overlapped an average of 17 grid cells. Therefore, a selection of three grid cells only covered about 17% of each field. However, we augmented this coverage by also conducting predator surveys within pastures; these covered the entire field and included each of the three grid cells with a camera (see section on indirect sign surveys).

We recorded all individual LC visits and numbers of visits by camera-days (i.e., one day per functioning camera per field). We define Individual LC visits as visits to a camera by a single individual of a species. A new visit by an individual of the same species began when the LC was recorded on a camera 1+ hours after the last recorded photo of an LC of that species. Our criterion of one hour between photos served as an index of frequency to estimate LC presence near subject herds. Because we were not concerned with actual abundance but rather presence of LCs around cattle, we did not identify specific individuals within each LC species of interest. Therefore, we did not determine whether individual LC’s were returning. We simply analyzed the presence of LCs.

Indirect sign surveys - The lead author (NXL) was present the majority of the study period (except for the last two weeks of October due to adverse conditions from insurmountable snowfall). NXL completed weekly surveys of LC presence focused on detecting wolf, bear, cougar and coyote activity in pastures with subject herds. These surveys included visits to each of the three trail cameras where NXL conducted systematic 100m transects along the closest animal trail to the camera. Along these transects NXL looked for scat, tracks and other signs. During each visit to a pasture, NXL also examined roads, vehicle trails, creek crossings and barbed wire fence-lines for LC signs. NXL recorded these data as presence or absence of LC by species. We examined tracks and scats whenever possible and identified them using (Halfpenny, 1986). Therefore, to account for uncertainty, we identified wolf and cougar presence solely from track and scat. For bears, we identified generic bear presence from hairs and signs (bear hair, scat, rubbing, and digging signs are distinctive from other LCs) but used tracks to differentiate black from grizzly bears (Halfpenny, 1986). Black bear tracks are smaller, with shorter nails, and the front toes form a clear arc. Grizzly bear tracks have longer nails, and the front toes are straight (Halfpenny, 1986). Therefore, we analyzed three sets of response variables: LC presence data derived from recorded camera visits (see more below), LC presence data from sign survey (Table 1), and the sum of the two datasets.

To reduce possible uncertainty in LC survey data, i.e., from possible misidentification and non-detection of sign, the LC survey data included only NXL's observations and not those of range riders and the ranch manager. This approach also maintained the blinding of range riders to our response variable, carnivore presence. NXL blinded the second author (AT) to treatment condition during analyses. Furthermore, we cross-referenced camera photos with sign data. Not all indirect sign was captured in photos. Other aspects of indirect sign increased our confidence in species identification, such as location or type of sign.

Analysis – We measured the response variables of the number of records of each LC species by a simple sum of the standardized value of the number of visits per day, recorded through photos and sign surveys (binary presence or absence per survey). Data from photos were recorded as number of individual visits per camera trap day, and data from indirect sign surveys were recorded as the presence or absence of predators per sign survey. These two variables were standardized individually and then summed to determine the total presence of each LC species per day (Table 1). Also, we combined all LC presence data into one response variable (pooled LC). We tested normality of the response variables using Shapiro–Wilk tests. We found bear, coyote, and pooled LC presence data to be normally distributed. Therefore, for bears ($W = 0.94$, $p = 0.40$), coyotes ($W = 0.97$, $p = 0.86$), and pooled LC ($W = 0.95$, $p = 0.52$), we used paired test adjusted by Hills and Armitage (Hills & Armitage, 1979) to analyze the effect of the treatment effect and order effects. The numbers of wolf ($W = 0.84$, $p = 0.010$) and cougar ($W = 0.49$, $p = 0.00000019$) records were not normally distributed, therefore for these species we use a non-parametric Wilcoxon sum rank test to evaluate the treatment effect and order effects on wolf and cougar data (Diaz- Uriarte, 2002). We use Hills-Armitage paired t-tests to test the order effect of the phases, i.e., whether having a treatment to pseudo-control (T-PC) sequence results in differing LC presence relative to pseudo-control to treatment (PC-T) sequence (Diaz-Uriarte, 2002). Further, this method allowed us to infer whether phase 1 and 2 were consistently different. For example, if LCs were more active in fall months (September and October of our study), then phase 2 might have had greater LC presence, regardless of a herd's status as a pseudo-control or treatment.

We also used Spearman's rank tests to estimate correlation between the LC presence near herds and the frequency of range rider presence. These correlations do not provide as strong inference as the above tests of treatment effects because the daily schedule of visits by range riders was not under our control. Nevertheless, our two estimates of the change in frequency of range rider visits between phases might reveal if human supervision was associated with changes in LC presence near herds. We defined the dose effect for each herd separately, in two ways: (a) the change in the number of days in which at least one range rider was present in a herd (hereafter range rider days), and (b) the change in the number of range riders summed across each phase in a herd (hereafter dose effect). For both range rider days and dose effect, we calculated the numbers in phase 2 and subtracted the numbers in phase 1 for this within-subject analysis. We correlated both (a) and (b) to the change in presence of bears, wolves, coyotes, and pooled LCs within herds as phase 2 presence − phase 1 presence. To do so, we summed LC presence across days in each phase for each herd separately. We did not analyze cougars separately because of small sample size.

Results

We studied 2,469 adult cattle and 1,928 calves split into eight subject herds on 22,500 acres (91 km2) of public grazing land between July 1 and October 31, 2020 (Table 1). No predation on cattle occurred during the experimental phases in either pseudo-control or treatment herds using a within-subjects test made possible by the crossover design. There was one confirmed livestock attack by a grizzly bear on a calf. This attack occurred during the wash-out period in a herd that was transitioning from pseudo-control to treatment, therefore it is not counted for either condition when we test for treatment effect. This attack was included as one day of grizzly bear presence in the affected herd, for the correlation between the change in grizzly bear presence and the changes in range rider days and dose effects. Eight cows died from ingesting poisonous plants, all during phase 1, and four cows died from other non-predator causes, one occurred during phase 1 and three during phase 2.

LC presence - During our study, we observed every carnivore species within pastures where subject herds were grazed, using cameras or indirect sign surveys for scat, track, hair, etc., regardless of treatment condition or study phase (Table 1; Figs. 3 and 4). We infer every subject herd faced some risk from predators.

Range rider presence - Treated herds experienced on average 2.75 times more human presence than pseudo-control herds (average 15.5 combined visits by range riders per herd vs. average 5.62 visits per herd by the ranch manager alone).

Treatment effect - The treatment conditions (treatment vs pseudo-control) did not predict pooled LC presence near herds in a within-subjects analysis that met the assumptions of normality and equal variance (Fig. 3, t -test $t(3) = -1.53$, $p = 0.89$). Therefore, we find no support for the first hypothesis that the number of range riders had an effect. Likewise, because treatments differed from pseudo-control in the experience of the added range riders, we find no support for an effect of experience or inexperience of range riders. Furthermore, the lack of treatment effect undermines the suggestion that LCs generally were deterred from one type of range riding to the other.

Effect of treatment on separate LC species or genus – Bear ($t(3)=-0.21$, $p=0.85$) and coyote ($t(3) = -0.35$, $p = 0.75$) presence did not differ across treatment condition (Fig. 3). Cougar and wolf presence data required a Wilcoxon two tailed sum rank test ($V = 1$, $p = 0.42$ and $V = 1$, $p = 0.25$ respectively). Therefore, species-specific presence data conform to that for pooled LC showing no treatment effect (Fig. 3).

Period effects might confound comparisons by altering conditions across all replicates in unison. Pooled LC presence did not significantly differ across phases ($t(3) = -0.22$, $p = 0.84$). We observed no difference across phases for wolves ($V = 7$, $p = 0.625$), bears (both species, $t(3) = -0.50$, $p = 0.65$), coyotes ($t(3) = -0.37$, $p = 0.74$) and cougars ($V = 2$, $p = 0.789$).

Carryover effects might arise if the response to the treatment lasted after the herd was no longer being treated. While we cannot rule out the potential for carryover effects because we observed no significant difference between the treatment conditions, we attempted to eliminate any such effect by implementing a 7-day ‘washout ’period.

Correlation of LC presence to range rider days and dose effects - To account for the frequency of range rider presence on change in LC presence we examined the change in ‘dosage ’of range riders within a herd regardless of phase or treatment condition. We include the washout period (Table S2) (e.g., three range riders could have visited a herd in the same day). Using a spearman's rank correlation we found that grizzly bear presence was weakly negatively correlated with days of range rider presence ($p = 0.066$, $rho = -0.67$), but not with dose effect, which accounted for the number of riders present summed across days in each herd ($p = 0.19$, $rho = -0.51$). There was no effect of range rider days or dosage on black bears, wolves or coyotes, nor on pooled LC. Therefore, we infer range riders were effective in protecting cattle, given losses in prior years and other Alberta ranches (see Methods). These data support the hypothesis that L-SLH reduced vulnerability but is not definitive given our pseudo-control and lack of control over the dose of range rider presence in each herd.

Discussion

We report a randomized, controlled experimental trial with crossover design to evaluate range riders practicing low stress livestock handling, L-SLH. Our experiment shows that one can graze thousands of cattle safely on vast public lands hosting grizzly bears and wolves when owners use L-SLH as a method to reduce the vulnerability of domestic animals. We conclude that L-SLH joins a growing number of non-lethal, carnivore control methods proven effective by gold standard experiments.

We find no support for the hypothesis that the risk to cattle from wolves, coyotes, black bears and cougars is affected by the single variables of number of range riders and their experience in practicing L-SLH. However, grizzly bears appeared to avoid herds exposed to regular range rider presence. Further, because L-SLH was being practiced at different dosages (one experienced rider practicing L-SLH vs. three riders practicing L-SLH) on every herd, and no livestock losses occurred during any treatment condition, we can conclude that L-SLH does not increase the risk of attack by LCs on cattle. We cannot rule out that a single experienced range rider practicing L-SLH every 9 days was as effective as one such experienced range rider supplemented by 1-2 inexperienced range riders, all practicing L-SLH and visiting herds every 3 days.

In our study, experienced range riders that had little or no L-SLH experience were trained quickly by an L-SLH experienced range rider. We hypothesize that functionally effective L-SLH can be trained in a short period. L-SLH is a form of livestock handling designed to reduce livestock stress, thereby increasing livestock health and yield. It has been identified by some livestock owners in Western North America as a useful means of retraining herding behavior and reducing livestock vulnerability to carnivore attacks, especially in free-ranging herds (Barnes, 2015; Zaranek, 2016). However, this method is still relatively uncommon, many owners instead preferring handling methods that prioritize speed and efficiency over animal welfare and stress (Hibbard, 2012). Antithetically, when speed is prioritized, animal handling appears more difficult as animals resist herding, leading to increased stress, reduced body condition and yield (Hibbard, 2012; Grandin, 1989). This combination of factors could make free-ranging livestock particularly vulnerable to attack by carnivores (Zaranek, 2016; Mech & Peterson, 2010). Given much livestock grazing worldwide is free-ranging, L-SLH may have wide utility and co-benefits as a mobile deterrence method (Khorozyan et al., 2020; Radford et al., 2020). Range riding comes in many forms, and must be clearly defined and experimentally tested in each new application, if it is to be confirmed as effective (Parks & Messmer, 2016). A lack of consensus on its definition and methods of use reduces its functional and perceived effectiveness, and risks wasting government and community resources (Parks & Messmer, 2016). We propose that further research is needed to compare different forms of animal handling to determine whether

L-SLH reduces vulnerability of livestock when compared to ‘traditional ’handling methods. Furthermore, it should be compared in more species of livestock, though there is some evidence that it may be effective with sheep (Stone et al., 2017), and L-SLH can be used on animals from cattle to chickens (Hibbard, 2012).

Co-benefits

An assumed benefit of range riding is the ability of range riders to quickly observe and manage all sorts of problems affecting the herd. For example, during our study there were eight recorded cattle mortalities from cattle ingesting toxic plants, such as larkspur, water hemlock or saskatoon blooms (Majak, Brooke & Ogilvie, 2008), and four mortalities from other non-predator causes. This is not uncommon in our study area according to the ranch manager, and occurred more regularly earlier in the grazing season, during the first phase of our study, when the majority of poisonous plants are at their most toxic (Majak, Brooke & Ogilvie, 2008). Range riders or the ranch manager identified these areas and were either able to remove dead stock or increase their own presence within these herds.

A primary attractant of grizzly bears in this region of Alberta are dead animals (Morehouse, 2016), so the presence of poison killed cattle might be attracting bears into the study area. This may explain why our results show a non-significant increase in presence of both species of bears during phase 1 of the study (Fig. 3). Attacks by bears on livestock did not increase however, either due to the supplemental feeding provided by the already dead animals, or the increased presence of range riders during these times. We did not observe that these dead animals attracted wolves or other carnivores. Morehouse & Boyce (2017) found that diversion feeding of grizzlies in Alberta resulted in a few dominant males protecting the food source against other individuals. Therefore, depending on the individuals attracted to the dead livestock, bears could have repelled other bears or species from the area. Evidence is mixed regarding whether diversionary feeding of carnivores is effective (Garshelis et al., 2017) at reducing conflict (Morehouse & Boyce, 2017; Steyaert et al., 2014), attracts bears and could increase conflict (Kavčič et al., 2015) or is effective by diverting carnivores to other sources (Stringham & Bryant, 2015). It is difficult to disentangle the effects of human presence, diversion feeding, and competition between bears and other carnivores. We did observe a large negative correlation between the numbers of days with range riders and the presence of grizzly bears, with fewer bears present when range riders increased their presence. However, there was no dose effect of the number of range riders within a herd summed over the study period. In other words, grizzlies might have responded to presence or absence of range riders in a day, but not to the number of individual riders present throughout the study phase. Presumably the direction of causality would not be the reverse, given range riders were employed to deter predators.

Further research is needed to determine the potential for other co-benefits of L-SLH. For example, reduced numbers of livestock losses, either by avoiding poisonous plants or reducing carnivore attacks would increase annual yield for ranchers. However, reduced stress could also significantly increase livestock yields, and therefore improve financial returns at the end of the season (Stringham & Bryant, 2015). Such experiments have never been conducted in conjunction with L-SLH.

Attacks on livestock

The presence of carnivores throughout the study period suggests that risk existed. Furthermore, shortly before the study began, a wolf attack was confirmed within one of the study area's cattle herds. Wolf and grizzly bear attacks on calves were commonly reported throughout the ecosystem in which we worked in the past, with 92 livestock losses reported in the previous 5 years (Alberta Environment and Parks, 2020). While we continued to observe carnivores during our experiment, we received no news of heightened attacks. Carnivores did not on average change their frequency of presence within either treatment or pseudo-control herds (but see discussion of wolves and coyotes below). No carnivores attacked livestock during either treatment or pseudo-control condition (but see below for an attack during the wash-out period). Despite a consistent presence of grizzlies, black bears, wolves, cougars and coyotes throughout our study period, individual predators or their foraging groups did not switch to pseudo-control or treatment herds when confronted with both. Therefore, we find no support for the hypothesis that large carnivores moved from the better-protected herds to less well-protected herds within a large public grazing land allotment. Therefore, our results seem contrary to claims that non-lethal methods simply displace predators to less-protected livestock.

Indeed, such spill-over effects have been reported for lethal control. Several studies found lethal control counter-productive (Van Eeden et al., 2018; Treves et al., 2019). Santiago- Avila, Cornman & Treves (2018) estimated that Michigan farms treated with lethal methods experienced a reduction in the future risk of livestock attacks by approximately 25%, while at the same time, untreated neighboring farms experienced an increase in future risk of approximately 75%. Grente et al. (2021) found a similar failure to attain desired outcomes of lethal control in a majority of nine study areas in France.

During our study's wash-out period, grizzly bear(s) apparently killed a calf in the week preceding when we would switch that herd from pseudo-control to treatment. During the wash-out, all eight herds experienced the lowest frequency of range rider supervision because only the ranch manager visited each herd. Furthermore, the dead calf's herd experienced an average amount of grizzly bear presence in phase 1 (four grizzlies, mean of all herds = 4.75 ±4), it had a lower-than-average number of days with range riders (4 days with range riders, mean of all herds = 10.5 ±6.7) during this phase. This suggests that the lack of human presence provided an opportunity for attack but did not attract grizzlies. While this single attack does not change the statistical significance of the treatment effect, it does suggest that the addition of two newly trained range riders is more effective at reducing risk to cattle than the ranch manager working alone because it increased the number of days riders could be in the herd. Therefore, it is likely that there is a dose effect of L-SLH range riders, but more research is needed to confirm and evaluate how much time should be spent with a herd.

Wolves

Though the effect was not significant, wolves were observed slightly more frequently during phase 2 and near treatment herds during phase 1 of the study (Fig. 3). The increased presence of wolves later in the season may result from wolf breeding behavior. Wolves select breeding sites further from human activity where humans persecute them, as they do on this landscape. This may explain why fewer wolf signs were observed during phase 1, which occurred while pups are young and wolf packs remain close to breeding sites (Mech & Peterson, 2010; Sazatornil et al., 2016). Phase 2 occurred during the autumn (September-October), when pups are older and wolf packs become nomadic, increasing the likelihood that we would observe them in our study area (Wilson, Bradley & Neudecker, 2017; Neufeld, 2006; Packard, 2010). The non-significant increase in the presence of wolves in treatment herds during phase 1 compared to pseudo-control herds may reflect a larger number of calves in the treatment condition during phase 1, or the novelty of the new range riders. First, calves are smaller and weaker than mature cattle, and therefore are more vulnerable to predation. Smaller calves are also less independent and rely on their mothers and the herd for protection (Flörcke & Grandin, 2013). However, if wolves were attracted to the increased number of calves in the treatment condition during phase 1, it is probable that range riders aided in reducing the vulnerability of these herds, leading to no attacks on the calves. We did not observe this same trend in phase 2 when the pseudo-control condition contained more calves. Wolves in Alberta may increase predation on livestock in late summer and fall (during phase 2) (Dorrance, 1982), and appear to preferentially select for cattle less than 9 months old. Therefore, we would expect to observe more wolves in pseudo-control herds in phase 2, where there were more calves. During this time calves are larger in size, but they are also more independent (Reinhardt & Reinhardt, 1981), which may increase their vulnerability in different ways. For example, they wander further from their mothers as they begin to wean off milk and their mothers become less vigilant over them (Flörcke & Grandin, 2013; Reinhardt & Reinhardt, 1981). However, there was no difference in wolf presence between treatment conditions during phase 2, consistent with the novelty of new range riders wearing off.

The second possibility for increased wolf activity in treatment herds in phase 1 is the novelty of new range riders. We presume the intelligence of wolves leads them to investigate novelties such as range riders (Range & Virányi, 2013). Carnivores explore novel situations to gain information about their environments and territories (Much et al., 2018). Repeated exposure to certain circumstances can, however, reduce curiosity and exploratory behavior (Mettke-Hofmann et al., 2006). Many individual wolves avoid new objects and circumstances (Mettke-Hofmann et al., 2006). This avoidance is thought to be a primary driver of the success of fladry, a form of fencing that uses evenly spaced flagging, to reduce wolf encroachment into fladry surrounded areas (Eklund et al., 2017; Davidson-Nelson & Gehring, 2010; Musiani et al., 2003; Musiani & Visalberghi, 2001). However, despite fladry deterrent effects in keeping wolves out, in most studies which have recorded wolf approaches to fladry, more wolf approaches were recorded in proximity to fladry fencing, than to control areas where no fladry was installed, despite wolves rarely if ever crossing the fladry barriers (Davidson-Nelson & Gehring, 2010; Musiani et al., 2003). Therefore, our finding of increased number of wolf visits in treatment herds may have resulted from the novelty of the new range riders. If this conjecture is correct, an increased number of observations of wolves does not imply increased risk to livestock, but instead an opportunity for wolves to learn about their environment. This learning is an important aspect of the deterrence work of range riders (Much et al., 2018). If wolves explore their territory and learn that range riders are a threat or that livestock are not vulnerable, we might expect them to become accustomed to the presence of range riders and focus their energy on hunting wild prey. This may explain why, despite more frequent visits by wolves in phase 2 of our study, there was no increase risk for cattle between the two study phases.

Conclusion

We observe that when properly executed L-SLH protects cattle with fewer riders. Furthermore, it is a method that is quickly taught. Our recommendations are that (1) L-SLH be tested in a randomized controlled experiment against non-L-SLH (i.e., ‘traditional’) livestock handling and/or a true control. This would help to determine whether other forms of human presence deter, attract or have no effect on predator attacks on livestock. True controls might involve a saddled, riderless horse with the cattle herd or visits by humans who do no herding at all. However, we demonstrate in this experiment, how difficult it is to implement a true control, as many livestock owners are unwilling to leave their herds unattended; (2) L-SLH methods should be studied to examine the number of newly trained range riders that are optimal for predator deterrence and cost effectiveness. The results of this study suggest that fewer riders may be just as effective on the predators being deterred, but that some predators, such as grizzly bears, may require a minimum amount of range riders practicing L-SLH. Further, studies comparing numbers of newly trained riders would produce evidence regarding what level of training must be attained for effective predator deterrence, particularly as there are few L-SLH-veteran range riders working today. Therefore, given a lack of further research on range riding, and a lack of consensus on the efficacy of other forms of range riding, L-SLH should be prioritized as the only form of range riding, to our knowledge, to have been experimentally tested.

Methods to reduce risk of attacks on livestock present many benefits, particularly as free-ranging livestock occur throughout the world and present a challenge to co-existence with carnivores. Human-caused mortality, often in response to perceived conflicts, is the primary form of mortality in large carnivore populations, and risks undermining conservation efforts to restore carnivore populations (Ripple et al., 2014). By not harming carnivores through displacement or unbalancing social structures (e.g., wolf packs), L-SLH also presents itself as a non-lethal predator control method which is effective for both target and adjacent properties without harming the carnivores (Haber, 1996).

Tables & Figures

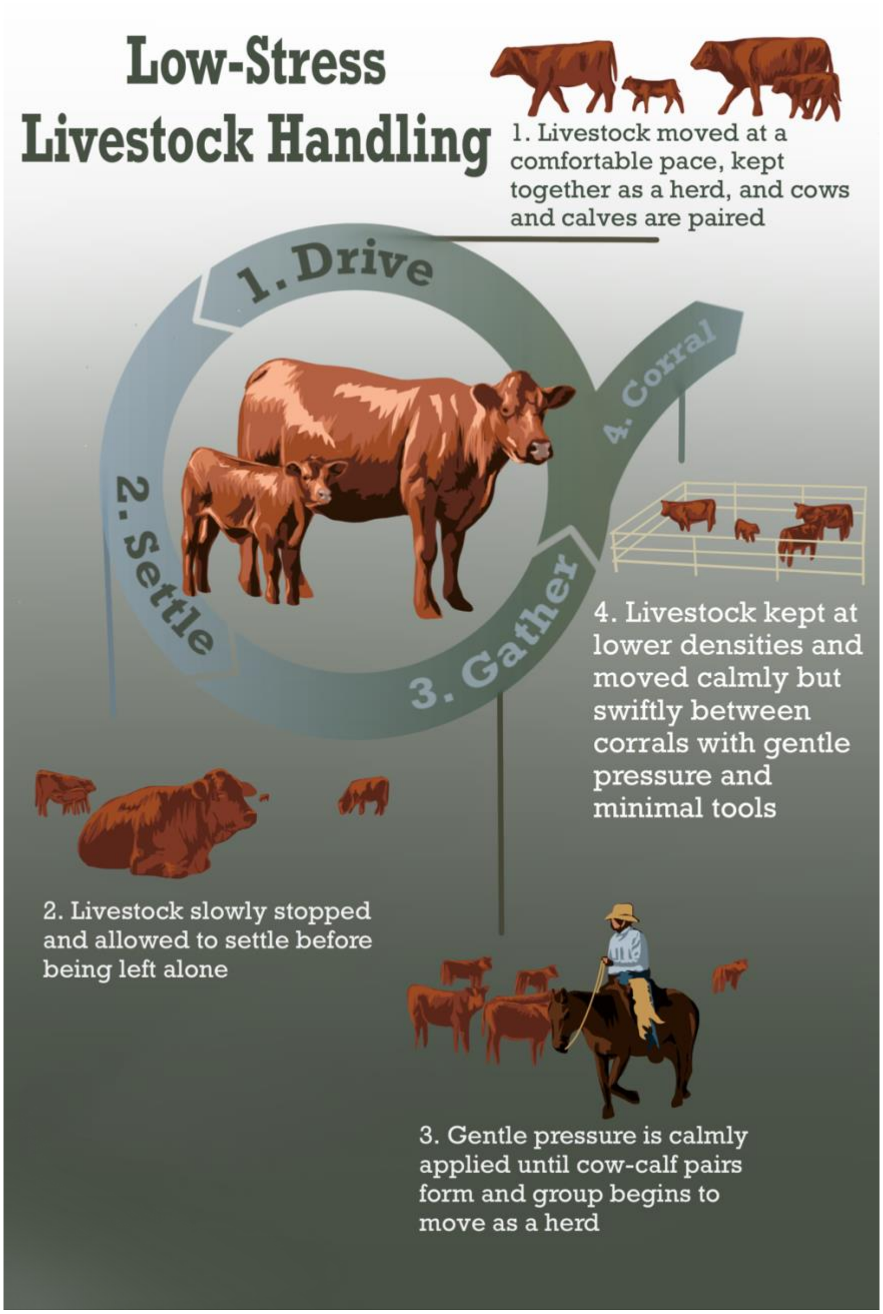


Fig 1. A conceptual diagram of the primary elements of low-stress livestock handling

The major elements of L-SLH are how cattle are moved (Drive), how they are stopped and placed in new pastures (Settle), how they are handled while within pastures or prepared to move out of pastures (Gather), and how cattle are corralled and prepared to me moved to new facilities or sent to feed lots (Corral).

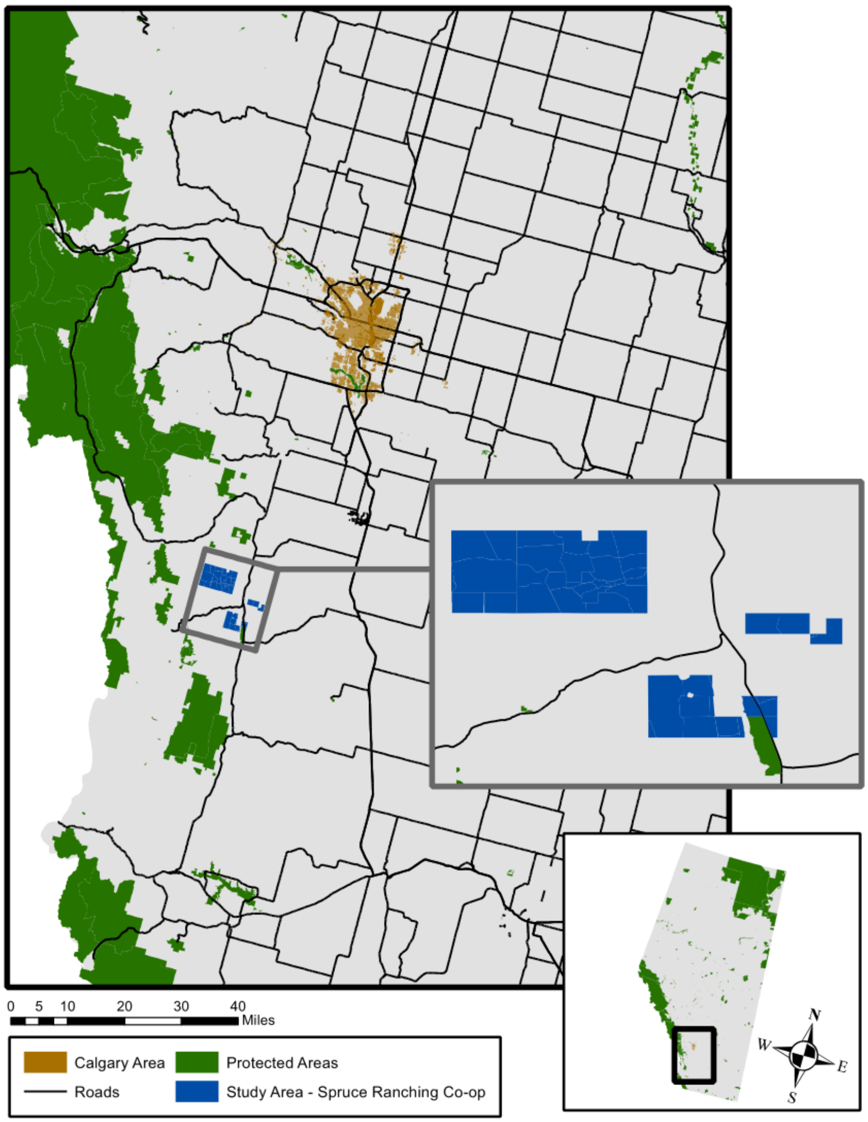
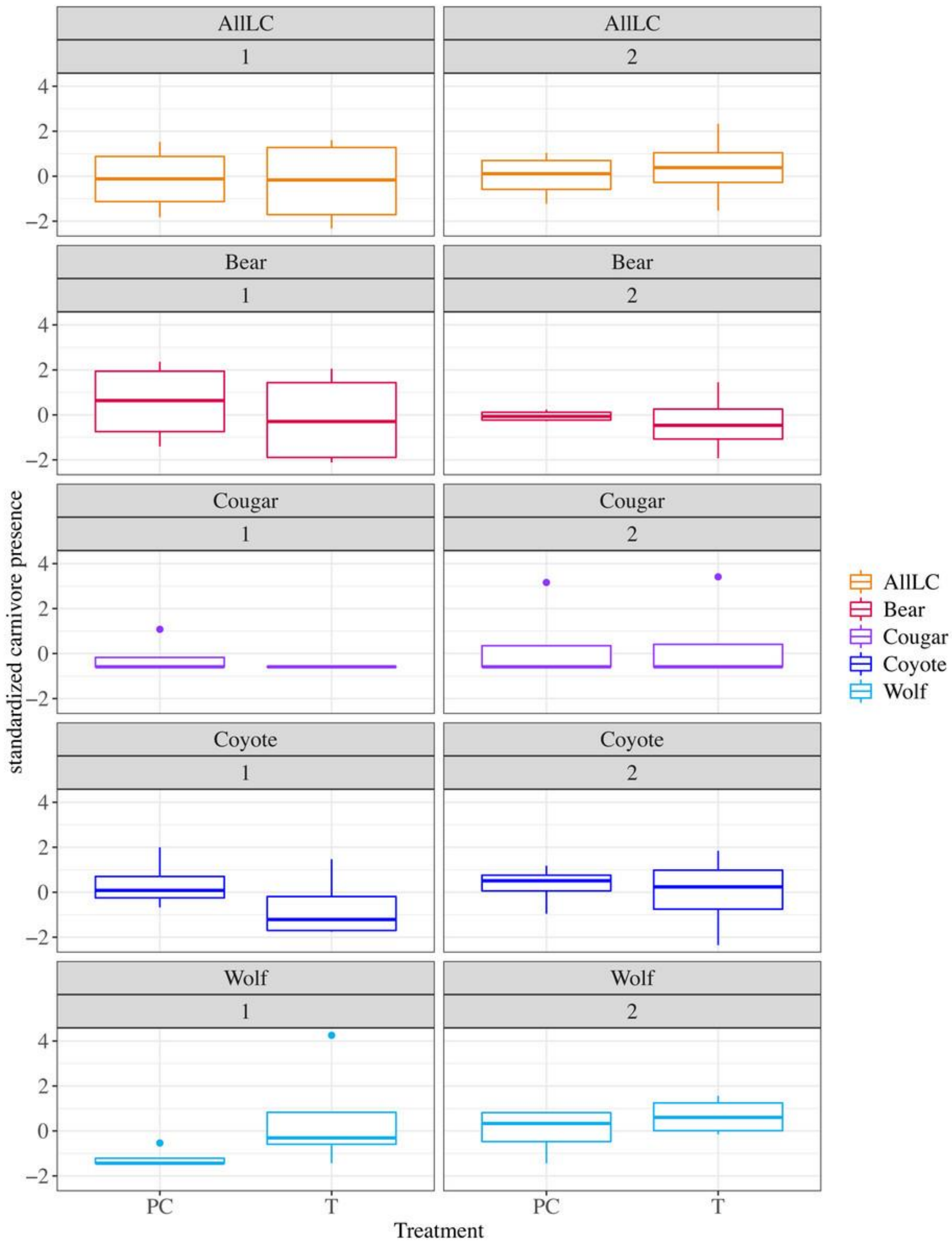


Fig 2. Study area map of the Spruce Ranching Co-op.

Extent map of the Spruce Ranching Co-op, located in the foothills of the Rocky Mountains in Southwestern Alberta, Canada. Protected areas refers to areas designed either by the Canadian federal government or the Alberta provincial government as protected. Base map source: Alberta Environment and Parks, Government of Alberta, 2021



Large carnivore (LC) presence standardized to include trail camera and sign survey data by carnivore species, phase (1 or 2) and treatment condition (PC = pseudo-control, T=treatment). Pooled LC represents the summed presence of all species.

Fig 4. Images of predators captured by trail cameras positioned within pastures with study herds.

(A) Grizzly bear (*Ursus arctos*) (B) Black bear (*Ursus americanus)* (C) Coyote *(Canis latrans*) (D) Wolf (*Canis lupus).* (E) Cougar (*Puma concolor)*

Table 1. Large carnivore visits by species. Number of large carnivore (LC) visits by subject herds, species, and phase, recorded using LC surveys and trail cameras. Number of visits are reported as visits per 10 LC surveys, visits per 100 camera-days. Treatment sequences are pseudo-control- treatment (PC-T) and treatment-pseudo-control (T-PC).

Herd	Treatment Sequence	Number of cattle, number of calves	Wolf Visits		Bear Visits		Cougar Visits		Coyote Visits	
			Phase		*Phase*		*Phase*		*Phase*	
			1	***2***	***1***	***2***	***1***	***2***	***1***	***2***
1	PC-T	400, 400	2.0, 0	5.0, 0	8.0, 5.0	7.5, 1.0	0, 0	2.5, 0	4.0, 0	5.0, 2.1
2	PC-T	355, 355	0, 0	5.0, 0	6.3, 3.9	6.7, 1.9	0, 0	0, 0	2.5, 7.9	3.3, 0
3	PC-T	82, 82	0, 0	2.3, 0	4.4, 2.1	7.1, 0.8	1.1, 0	0, 0	4.4, 0.9	4.3, 2.3
4	PC-T	161, 0*	0, 0	0, 0	2.8, 1.6	8.0, 0	0, 0	0, 0	5.7, 0.6	6.0, 2.4
5	T-PC	250, 250	0, 0	2.9, 0	1.4, 1.3	2.8, 0.63	0, 0	0, 0	1.4, 0	10.0, 0
6	T-PC	480, 480	2.5, 0	3.3, 0	11.3, 2.1	6.7, 4.3	0, 0	0, 2.6	1.3, 0.3	1.7, 3.4
7	T-PC	361, 361	3.8, 0.7	5.7, 0	3.7, 4.0	4.2, 7.5	0, 0	0, 0	2.5, 1.3	4.3, 3.0
8	T-PC	380, 0*	2.5, 0	6.7, 0	2.5, 0	6.7, 0	0, 0	0, 0	2.5, 6.7	0, 0

Notes.

*Herds with no calves were all pregnant yearlings, *i.e.*, heifers.

References

1 Ripple, W. J. et al. Status and ecological effects of the world's largest carnivores. Science 343 (2014).

2 Estes, J. A. et al. Trophic downgrading of planet Earth. science 333, 301-306 (2011).

3 Woodroffe R; Ginsberg, J. Edge Effects and the Extinction of Populations Inside Protected Areas. Science 280, 2126-2128 (1998). <http://science.sciencemag.org/content/280/5372/2126.abstract>.

4 Treves, A. et al. Predators and the public trust. Biological Reviews 92, 248-270 (2017).

5 Van Eeden, L. M. et al. Carnivore conservation needs evidence-based livestock protection. PLoS biology 16, e2005577 (2018).

6 Boronyak, L. et al. Pathways towards coexistence with large carnivores in production systems. Agriculture and Human Values, 1-18 (2021).

7 Williams, P. in The New Yorker (online, 2022).

8 Service, U. E. R. Ag and Food Sectors and the Economu, 2021).

9 Lee, T., Good, K., Jamieson, W., Quinn, M. & Krishnamurthy, A. Cattle and carnivore coexistence in Alberta: the role of compensation programs. Rangelands 39, 10-16 (2017).

10 Parks, A. E. a. Grizzly Bear Response Guide. Report No. 1, 6 (Alberta Environment and Parks, Alberta, Canada, 2016).

11 Treves, A., Krofel, M. & McManus, J. Predator control should not be a shot in the dark. Frontiers in Ecology and the Environment 14, 380-388 (2016).

12 Young, J. K., Hammill, E. & Breck, S. W. Interactions with humans shape coyote responses to hazing. Scientific reports 9, 1-9 (2019).

13 Khorozyan, I. & Waltert, M. Variation and conservation implications of the effectiveness of anti-bear interventions. Scientific reports 10, 1-9 (2020).

14 Scasta, J., Stam, B. & Windh, J. Rancher-reported efficacy of lethal and non-lethal livestock predation mitigation strategies for a suite of carnivores. Scientific reports 7, 1-11 (2017).

15 Linnell, J. D., Odden, J., Smith, M. E., Aanes, R. & Swenson, J. E. Large carnivores that kill livestock: do" problem individuals" really exist? Wildlife Society Bulletin, 698-705 (1999).

16 Odden, J. et al. Lynx depredation on domestic sheep in Norway. The Journal of wildlife management, 98-105 (2002).

17 Treves, A. & Naughton-Treves, L. Evaluating lethal control in the management of human-wildlife conflict. CONSERVATION BIOLOGY SERIES-CAMBRIDGE- 9, 86 (2005).

18 Treves, A. Hunting for large carnivore conservation. Journal of Applied Ecology 46, 1350-1356 (2009).

19 Treves, A., Krofel, M., Ohrens, O. & van Eeden, L. M. Predator control needs a standard of unbiased randomized experiments with cross-over design. Frontiers in Ecology and Evolution 7, 462 (2019).

20 Lennox, R. J., Gallagher, A. J., Ritchie, E. G. & Cooke, S. J. Evaluating the efficacy of predator removal in a conflict-prone world. Biological Conservation 224, 277-289 (2018).

21 Wielgus, R. B. & Peebles, K. A. Effects of wolf mortality on livestock depredations. PloS one 9, e113505 (2014).

22 Santiago-Avila, F. J., Cornman, A. M. & Treves, A. Killing wolves to prevent predation on livestock may protect one farm but harm neighbors. PLoS One 13, e0189729 (2018).

23 Grente, O. et al. Tirs dérogatoires de loups en France: évaluation des effets sur les attaques aux troupeaux. (2022).

24 Wallach, A. D., Ritchie, E. G., Read, J. & O'Neill, A. J. More than mere numbers: the impact of lethal control on the social stability of a top-order predator. PLoS One 4, e6861 (2009).

25 Ohrens, O., Bonacic, C. & Treves, A. Non-lethal defense of livestock against predators: flashing lights deter puma attacks in Chile. Frontiers in Ecology and the Environment 17, 32-38 (2019).

26 Khorozyan, I. & Waltert, M. A global view on evidence-based effectiveness of interventions used to protect livestock from wild cats. Conservation Science and Practice 3, e317 (2021).

27 Khorozyan, I., Ghoddousi, S., Soufi, M., Soofi, M. & Waltert, M. Studded leather collars are very effective in protecting cattle from leopard (Panthera pardus) attacks. Ecological Solutions and Evidence 1, e12013 (2020).

28 Radford, C., McNutt, J. W., Rogers, T., Maslen, B. & Jordan, N. Artificial eyespots on cattle reduce predation by large carnivores. Communications biology 3, 1-8 (2020).

29 Bangs, E. et al. in Proceedings of the Vertebrate Pest Conference.

30 Parks, M. & Messmer, T. Participant perceptions of Range Rider Programs operating to mitigate wolf–livestock conflicts in the western United States. Wildlife Society Bulletin 40, 514-524 (2016).

31 Jablonski, K. E., Merishi, J., Dolrenry, S. & Hazzah, L. Ecological doctors in Maasailand: identifying herding best practices to improve livestock management and reduce carnivore conflict. Frontiers in Sustainable Food Systems 4, 118 (2020).

32 Wilson, S. M., Bradley, E. H. & Neudecker, G. A. Learning to live with wolves: community-based conservation in the Blackfoot Valley of Montana. Human–Wildlife Interactions 11, 4 (2017).

33 Bruns, A., Waltert, M. & Khorozyan, I. The effectiveness of livestock protection measures against wolves (Canis lupus) and implications for their co-existence with humans. Global Ecology and Conservation 21, e00868 (2020).

34 Hibbard, W. Bud Williams 'low stress livestock handling. Stockmanship Journal 1, 6-163 (2012).

35 Grandin, T. Behavioral principles of livestock handling. The Professional Animal Scientist 5, 1-11 (1989).

36 Barnes, M. Livestock management for coexistence with large carnivores, healthy land and productive ranches. Keystone Conservation (2015).

37 Zaranek, H. Stockmanship and Livestock Predation Mitigation. Stockmanship Journal 5, 34-46 (2016).

38 Mech, L. D. & Peterson, R. O. in Wolves 131-160 (University of Chicago Press, 2010).

39 Eklund, A., López-Bao, J. V., Tourani, M., Chapron, G. & Frank, J. Limited evidence on the effectiveness of interventions to reduce livestock predation by large carnivores. Scientific reports 7, 1-9 (2017).

40 Stone, S. A. et al. Adaptive use of nonlethal strategies for minimizing wolf–sheep conflict in Idaho. Journal of Mammalogy 98, 33-44 (2017).

41 Proctor, M. F. et al. Population fragmentation and inter-ecosystem movements of grizzly bears in western Canada and the northern United States: Fragmentation de la Population et Mouvements Inter-Ecosystèmes des Ours Grizzlis dans L'ouest du Canada et le Nord des États-Unis. Wildlife Monographs 180, 1-46 (2012).

42 Parks, A. E. a. Predator Compensation 2015 to 2019: South Saskatchewan Region. (2020).

43 Halfpenny, J. C. A field guide to mammal tracking in North America. (Big Earth Publishing, 1986).

44. Hills, M. & Armitage, P. The two-period cross-over clinical trial. British journal of clinical pharmacology 8, 7 (1979).

45 Diaz-Uriarte, R. Incorrect analysis of crossover trials in animal behaviour research. Animal behaviour (2002).

46 Majak, W., Brooke, B. M. & Ogilvie, R. T. Stock-poisoning plants of western Canada. Canadian Department of Agriculture, Ottawa (2008).

47 Morehouse, A. T. Grizzly bear population ecology and large carnivore conflicts in southwestern Alberta. (2016).

48 Morehouse, A. T. & Boyce, M. S. Troublemaking carnivores: conflicts with humans in a diverse assemblage of large carnivores. Ecology and Society 22 (2017).

49 Garshelis, D. L., Baruch-Mordo, S., Bryant, A., Gunther, K. A. & Jerina, K. Is diversionary feeding an effective tool for reducing human–bear conflicts? Case studies from North America and Europe. Ursus 28, 31-55 (2017).

50 Steyaert, S. M. et al. Behavioral correlates of supplementary feeding of wildlife: can general conclusions be drawn? Basic and Applied Ecology 15, 669-676 (2014).

51 Kavčič, I. et al. Fast food bears: brown bear diet in a human-dominated landscape with intensive supplemental feeding. Wildlife Biology 21, 1-8 (2015).

52 Stringham, S. F. & Bryant, A. Distance-dependent effectiveness of diversionary bear bait sites. Human–Wildlife Interactions 9, 12 (2015).

53 Smith, B. J. Moving Em': A Guide to Low Stress Animal Handling. 1 edn, (Graziers Hui, 1998).

54 Sazatornil, V. et al. The role of human-related risk in breeding site selection by wolves. Biological Conservation 201, 103-110 (2016).

55 Neufeld, L. M. Spatial dynamics of wolves and woodland caribou in an industrial forest landscape in west-central Alberta. (University of Alberta Edmonton, Alberta, Canada, 2006).

56 Packard, J. M. in Wolves 35-65 (University of Chicago Press, 2010).

57 Flörcke, C. & Grandin, T. Loss of anti-predator behaviors in cattle and the increased predation losses by wolves in the Northern Rocky Mountains. Open Journal of Animal Sciences 3, 248 (2013).

58 Dorrance, M. J. Predation losses of cattle in Alberta. Rangeland Ecology & Management/Journal of Range Management Archives 35, 690-692 (1982).

59 Reinhardt, V. & Reinhardt, A. Natural sucking performance and age of weaning in zebu cattle (Bos indicus). The Journal of Agricultural Science 96, 309-312 (1981).

60 Range, F. & Virányi, Z. Social learning from humans or conspecifics: differences and similarities between wolves and dogs. Frontiers in Psychology 4, 868 (2013).

61 Much, R. M., Breck, S. W., Lance, N. J. & Callahan, P. An ounce of prevention: Quantifying the effects of non-lethal tools on wolf behavior. Applied Animal Behaviour Science 203, 73-80 (2018).

62 Mettke-Hofmann, C., Rowe, K., Hayden, T. & Canoine, V. Effects of experience and object complexity on exploration in garden warblers (Sylvia borin). Journal of Zoology 268, 405-413 (2006).

63 Moretti, L., Hentrup, M., Kotrschal, K. & Range, F. The influence of relationships on neophobia and exploration in wolves and dogs. Animal Behaviour 107, 159-173 (2015).

64 Davidson-Nelson, S. J. & Gehring, T. M. Testing fladry as a nonlethal management tool for wolves and coyotes in Michigan. Human-Wildlife Interactions 4, 87-94 (2010).

65 Musiani, M. et al. Wolf depredation trends and the use of fladry barriers to protect livestock in western North America. Conservation Biology 17, 1538-1547 (2003).

66 Musiani, M. & Visalberghi, E. Effectiveness of fladry on wolves in captivity. Wildlife Society Bulletin, 91-98 (2001).

67 Haber, G. C. Biological, conservation, and ethical implications of exploiting and controlling wolves. Conservation Biology 10, 1068-1081 (1996).

Acknowledgements - We would like to thank Joe Engelhart and the members of the Spruce Ranching Co-op for being our research partners and allowing us to design this experiment with their livestock. We would like to thank our range riders for their hard work and dedication to this project.

Supplementary Materials for Low-stress livestock handling protects cattle in a five-predator habitat

Naomi X Louchouarn, Adrian Treves

Correspondence to: louchouarn@wisc.edu

Table S1. Expanded data table for use with R code listed below. Approaches are defined as visits per camera trap day. Indirect sign are calculated as visits per large carnivore survey. Seq refers to the treatment sequence of either pseud-control-treatment (PC-T) or treatment-pseudo-control (T-PC).

Herd	Seq	Treatment. Control	Phase	Wolf		Bear		Cougar		Coyote	
				Approaches	Indirect sign	Approaches	Indirect sign	Approaches	Indirec t sign	Approaches	Indirect sign
1	PC-T	PC	1	0.0	0.20	0.022	0.80	0.0	0.0	0.0	0.40
1	PC-T	T	2	0.0	0.50	0.0	0.75	0.0	0.25	0.021	0.50
2	PC-T	PC	1	0.0	0.0	0.033	0.63	0.0	0.0	0.079	0.25
2	PC-T	T	2	0.0	0.50	0.0099	0.67	0.0	0.0	0.0	0.33
3	PC-T	PC	1	0.0	0.0	0.0	0.44	0.0	0.11	0.0091	0.44
3	PC-T	T	2	0.0	0.29	0.0076	0.71	0.0	0.0	0.023	0.43
4	PC-T	PC	1	0.0	0.0	0.016	0.29	0.0	0.0	0.0055	0.57
4	PC-T	T	2	0.0	0.0	0.0	0.80	0.0	0.0	0.024	0.60
5	T-PC	T	1	0.0	0.0	0.0066	0.14	0.0	0.0	0.0	0.14

5	T-PC	PC	2	0.0	0.29	0.0063	0.29	0.0	0.0	0.0	1.0
6	T-PC	T	1	0.0	0.25	0.014	1.125	0.0	0.0	0.0035	0.13
6	T-PC	PC	2	0.0	0.33	0.043	0.67	0.026	0.0	0.034	0.17
7	T-PC	T	1	0.0067	0.38	0.040	0.38	0.0	0.0	0.013	0.25
7	T-PC	PC	2	0.0	0.57	0.023	0.43	0.0	0.0	0.030	0.43
8	T-PC	T	1	0.0	0.25	0.0	0.25	0.0	0.0	0.067	0.25
8	T-PC	PC	2	0.0	0.67	0.0	0.67	0.0	0.0	0.0	0.0

Table S2. Table of range rider days (number of days with at least one range rider present) and range rider dose (summed number of range rider visits by phase. LC presence columns are based on number of days with at least one grizzly bear, black bear, wolf and coyote presence by phase is also included.

	Range Rider Days		Dose Effect		Grizzly Bear		Black Bear		Wolf		Coyote	
Herd	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
1	4	22	4	33	4	0	3	1	1	2	2	3
3	7	20	15	32	6	2	3	1	0	4	13	2
3	4	12	4	15	0	1	5	1	0	2	5	5
4	6	13	6	21	1	0	0	0	0	0	5	7
5	10	5	12	5	1	2	3	2	0	2	2	6
6	16	10	19	12	4	6	3	1	1	2	2	2
7	16	12	24	11	5	3	4	1	6	5	8	8
9	14	6	18	4	0	0	1	0	1	2	5	0

R-Code

1. Standardizing the direct and indirect predator signs

subtract from the mean of all observations and divide by the SD of all observations. Do this for both response variables; approaches (trail camera data) and indirect sign (large carnivore surveys)

```
Data$Approaches<- Data$BA+Data$WA+Data$CA+Data$BBA
approaches_mean <- mean(Data$Approaches)
approaches_sd <- sd(Data$Approaches)
Data$Standard_Approaches <- (Data$Approaches-approaches_mean)/approaches_sd
Data$Indirect <- Data$BI+Data$WI+Data$CI
indirect_mean<- mean(Data$Indirect)
indirect_sd<-sd(Data$Indirect)
Data$Standard_Indirect<- (Data$Indirect-indirect_mean)/indirect_sd
hist(Data$Standard_Indirect)

Data$pooled <- Data$Standard_Approaches+Data$Standard_Indirect
```

2. Subset the data by treatment sequence for analysis

```
PCT<- subset(Data, Data$seq == "PCT")

TPC<- subset(Data, Data$seq == "TPC")
```

3. Examine each predator species for individual response to treatment

Pooled LC

```
Shapiro.test(Data$Pooled.Standardized)

#treatment effect

t.test(PCT$Pooled.treatmenteffect, TPC$Pooled.treatmenteffect, paired = T)

#period effect

t.test(PCT$Pooled.periodeffect, TPC$Pooled.periodeffect, paired = T)
```

Bear

```
Shapiro.test(Data$Bear.Standardized)

#treatment effect

t.test(PCT$Bear.treatmenteffect, TPC$Bear.treatmenteffect, paired = T)

#period effect

t.test(PCT$Bear.periodeffect, TPC$Bear.periodeffect, paired = T)
```

Wolf

```
shapiro.test(Data$Wolf.Standardized)
wilcox.test(PCT$Wolf.treatmenteffect, TPC$Wolf.treatmenteffect, paired = T)
wilcox.test(PCT$Wolf.periodeffect, TPC$Wolf.periodeffect, paired = T)
```

Cougar

```
shapiro.test(Data$Cougar.Standardized)
wilcox.test(PCT$Cougar.treatmenteffect, TPC$Cougar.treatmenteffect, paired = T)
wilcox.test(PCT$Cougar.periodeffect, TPC$Cougar.periodeffect, paired = T)
```

Coyote

```
Shapiro.test(Data$Coyote.Standardized)
#treatment effect
t.test(PCT$Coyote.treatmenteffect, TPC$Coyote.treatmenteffect, paired = T)
#period effect
t.test(PCT$Coyote.periodeffect, TPC$Coyote.periodeffect, paired = T)
```

4. Spearman’s Rank Correlations

```
#range rider data
RRDays <- Data$RRdays.Phase2-Data$RRdays.Phase1
RRDose <- Data$RRdose.Phase2-Data$RRdose.Phase1
#Grizzly Bear
GBChange <- Data$GB.Phase2-Data$GB.Phase1
cor.test(RRDays, GBChange, method="spearman")
cor.test(RRDose, GBChange, method="spearman")
#Black Bear
BBChange <- Data$BB.Phase2-Data$BB.Phase1
cor.test(RRDays, BBChange, method="spearman")
```

```
cor.test(RRDose, BBChange, method="spearman")
#Wolf
WChange <- Data$W.Phase2-Data$W.Phase1
cor.test(RRDays, WChange, method="spearman")
cor.test(RRDose, WChange, method="spearman")
#Coyote
CoyChange <- Data$Coy.Phase2-Data$Coy.Phase1
cor.test(RRDays, CoyChange, method="spearman")
cor.test(RRDose, CoyChange, method="spearman")
```

Chapter 2: Evaluating how lethal management affects poaching of Mexican wolves

Naomi Louchouarn[1]*, Francisco J. Santiago-Ávila[1]*, David R. Parsons[2], Adrian Treves[1]

[1]Nelson Institute for Environmental Studies, University of Wisconsin-Madison

[2] Project Coyote Science Advisory Board

*Equal first co-authors

Abstract

Despite illegal killing (poaching) being the major cause of death among large carnivores globally, little is known about the effect of implementing lethal management policies on poaching. Two opposing hypotheses have been proposed in the literature: implementing lethal management may decrease poaching incidence (‘killing for tolerance’) or increase it (‘facilitated illegal killing’). Here, we report a test of the two opposed hypotheses that poaching (reported and unreported) of Mexican grey wolves (*Canis lupus baileyi*) in Arizona and New Mexico, USA, responded to changes in policy that reduced protections to allow more wolf-killing. We employ advanced biostatistical survival and competing-risk methods to data on individual resightings, mortality and disappearances of collared Mexican wolves, supplemented with Bayes Factors to assess strength of evidence. We find strong evidence that Mexican wolves were 121% more likely to disappear during periods of reduced protections than during periods of stricter protections, despite no change in legal removals by the agency. We also find inconclusive evidence for any decreases in reported poaching. Therefore, we find strong support for the ‘facilitated illegal killing ’hypothesis and none for the ‘killing for tolerance ’hypothesis. We provide recommendations for improving the effectiveness of US policy on environmental crimes, endangered species, and protections for wild animals. Our results have implications beyond the

USA or wolves because the results suggest transformations of decades-old management interventions against human-caused mortality among wild animals subject to high rates of poaching.

Keywords: Conservation, endangered species, poaching, policy signal, survival analysis, large carnivore, Mexican wolf, Canis lupus baileyi

Introduction

Human-caused mortality is the major cause of death among large, terrestrial, mammalian carnivores worldwide [1], including the USA [2-5]. Anthropogenic mortality has precipitated the decline and extirpation of carnivore populations worldwide both indirectly and directly through the often coinciding threats of habitat loss and degradation, prey depletion, and killing [6]. Indeed, reported and unreported poaching is the major form of human-caused mortality for large carnivore populations in several regions [7,8], including five U.S. wolf populations [4]. Such mortality raises individual and societal concerns because poaching is an environmental crime, harms individual animals, and undermines restoration and conservation efforts.

Identifying and estimating poaching is hindered by concealment of evidence. Estimating, concealed, illicit killing rates has recently been transformed by two analyses that used data that had previously been ignored. Liberg et al. [7] estimated the hazard rate of cryptic (i.e., unreported or concealed) poaching by considering slow-downs in population growth and accounting for the disappearances of marked grey wolves in Scandinavia. Similarly, Treves et al. [5] re-calculated the risk of poaching relative to other causes of death by considering missing, marked animals, which had previously been excluded from analyses under an erroneous assumption that marked animals that disappeared would have died of similar causes as those marked animals found dead. Therefore, investigators can now better estimate heretofore under-appreciated variables that are essential to understanding population dynamics and individual animal life histories. However, the latter study admittedly did not directly estimate poaching, instead using estimates from other populations (Scandinavia and Wisconsin) as multipliers to indirectly quantify cryptic poaching, and did not measure policy effects on poaching or consider time to exposure of wolves to policies. Its objective was strictly to estimate the risk of poaching in a population regardless of policy period. Therefore, here we propose an important advance to estimate the relationships between policy interventions and fates of marked carnivores, while controlling for spatiotemporal covariates. We test opposed hypotheses from the literature explained next.

The scientific literature has recently addressed the question of if and how policies may influence the hazard and incidence of poaching. The usual assumption (despite lack of empirical evidence) is that some predator-killing (e.g.: government permits for killing or public hunting seasons) might increase tolerance for a species (and thus reduce poaching); an argument first articulated in federal court in 2005 (this should be citation 51 to HSUS 2006) and summarized more generally in [9]. We call this first hypothesis ‘killing for tolerance’, which predicts legal killing will reduce poaching through the following mechanism: legalizing or liberalizing killing of controversial species will lead would-be perpetrators to desist from poaching because of increased tolerance for the species or approval for protectionist policies. Early tests of this notion of ‘killing for tolerance ’include [10-16]. Olson et al. [10] examined correlations between documented (i.e., reported) poaching of Wisconsin’s wolves and management policies between 2003-2011. They suggested that the incidence of known poaching events was inversely related to the proportion of each year with state management associated with liberalized killing periods, and hypothesized that frustration with protections for wolves led to increased poaching. Studying the same population albeit with more sophisticated modeling of demographic processes, Stenglein et al. [17] estimated an additional mortality of 4% was necessary to explain the observed slow-down in the population's annual growth rate within that same time period. These early tests of the killing for tolerance hypothesis attribute rates of poaching and population dynamic changes to illegal actions motivated by inconsistent management and protections for controversial wolves. By contrast, Chapron & Treves [14] reported serial slow-downs of wolf population growth during 6 non-consecutive policy periods in Wisconsin and Michigan from 1995-2012, which seemed attributable to unreported wolf-killing. They proposed an explanation we refer to as 'facilitated illegal killing'. Three social scientific studies published between 2013 and 2015 [11-13] examined attitudes towards wolves in Wisconsin and found that tolerance decreased as wolf killing was progressively liberalized or intention to poach wolves increased as wolf-killing was progressively liberalized from 2003-2013. Considering such evidence, the alternative hypothesis of 'facilitated illegal killing 'suggests that liberalized killing might decrease the value of wolves to would-be perpetrators of poaching, or decreasing the risk of being caught [14]. A 2019 re-analysis using the methods proposed below found liberalized killing policy periods in Wisconsin, USA (1979-2012), were associated with increases in hazard and incidence of wolf disappearances that outweighed by five-fold any decreases in reported poaching, undercutting the 'killing for tolerance 'hypothesis [18]. Despite the lack of a clear causal connection between attitudes and poaching, the study describe here tries to establish a closer mechanistic link between policies and poaching behavior. In sum, two published hypotheses make opposed predictions about the rates of poaching in relation to policies for liberalizing legal killing of controversial species.

Other research linking wolf mortality to population growth rates in a hunted Finnish population found increases in population size were positively associated with increases in poaching [8]. Using generalized linear models focused on predictors of poaching, the same team later found the number of legally hunted wolves both across the country and at the local scale was associated with a decrease in the probability of poaching, while increases in the number of wolves that could be legally killed (the 'bag limit') were associated with increases in the probability of poaching [19]. Additionally, the authors suggest that declines in poaching following higher levels of legal hunting might be an artifact of a decrease in the individual wolves exposed to poaching [19, p. 7]. They concluded that "tolerance for carnivores cannot be promoted by legal hunting alone, so more comprehensive conservation efforts are needed" [8,19].

The most recent publication on this topic for grey wolves in Scandinavia (20), suggested that more territorial breeding individuals removed legally, fewer such animals disappeared (presumably poached), but their analysis has been questioned on the grounds of inappropriate statistical analyses and incomplete treatment of the apparent rise in disappearances during years with legal wolf-killing [21].

However, there remain unresolved concerns about omitted methods and the statistical approaches and assumptions made in all three studies from the U.S. upper Midwest and Nordic countries [18,22]. None of the studies [8,14,15,17] explicitly modelled survival in relation to the amount of time wolves were exposed to liberalized killing policies that changed 12 times between 1995–2012 in the US and several times in the Nordic countries [14,23]. Here we build on these analyses by including the amount of time that radio-collared wolves were exposed to liberalized killing policies to re-estimate risk of poaching.

Indeed, a simple reduction in poaching may not be equivalent to ‘tolerance’; that is, greater acceptance of wolves on the landscape. The ‘killing for tolerance ’hypothesis suggests a cognitive mechanism that implies something broader than a reduction in poaching, which is only one anthropogenic endpoint affected by tolerance. However, poaching is not the only human behavior affected by tolerance causing wolf mortality. Human tolerance would arguably affect other anthropogenic endpoints such as legal killings or (in this particular population) final removals (e.g., through increased legal killings by citizens or requests for action to government agents leading to increased mortality, similar to that found in Scandinavia [19]). Thus, any comprehensive exploration of ‘tolerance ’affecting wolf mortality should examine the interactions between the different anthropogenic endpoints and their resulting incidences.

Opposing views of the relationship between legal killing and poaching of wolves can be tested if we analyze individual wolf survival in relation to the timing and duration of their exposure to periods with different policies for legal killing. Here we will test the specific hypothesis that rates of hazard and incidence of mortality or disappearance of wild Mexican grey wolves (Canis lupus baileyi) in Arizona and New Mexico, USA, changed after policies altered the legality of killing or harassment of Mexican wolves by the public and government agencies. Table 4 provides predictions for the two response variables of hazard ratios and competing risk subhazard ratios. There has been no research on the opposing hypotheses of “killing for tolerance” or “facilitated illegal killing”, within this population. Such a test is particularly important as Mexican wolves are a highly endangered subspecies of grey wolves. Previous estimates [4] indicate poaching rates in this population have been high and under-estimated by traditional methods.

The subspecies was functionally extinct in the USA by the 1970s due to extermination efforts by state, federal, tribal and private actors [24]. A captive breeding program began in 1977, and the US government began releasing Mexican wolves to the Blue Range Wolf Recovery Area in southern Arizona and New Mexico in 1998. From 1998 to 2016, all Mexican wolves released to the recovery area were fitted with a radio collar and were closely monitored (n=279 radio-collared). Here we examine the survival and disappearances of adult marked Mexican wolves before, during, and after two policy periods, one in 2005-2009 and another starting in 2015, that liberalized killing or removal of wolves by government or private actors.

We examine data on radio collared animals using competing risk analyses that allow the modelling of hazards and incidences of multiple fates (i.e., various causes of death or lost to monitoring; endpoints, hereafter) while controlling for multiple covariates. We model exposure time to policy changes. This analysis allows us to make inferences beyond the cursory examination of compensatory mortality causes due to changes in policy, to examine how policies might impact an individual wolf's probability of succumbing to a cause of death. Furthermore, using a competing risks analysis allows us to include disappearance as an endpoint which is crucial given that prior work shows that censoring LTF led to systematic under-estimates of poaching in other grey wolf populations [4,18]. The results of this analysis can provide recommendations for improving the effectiveness of US policy on environmental crimes, endangered species, and protections for wild animals. Therefore, our analyses have implications beyond the USA or wolves because the methods promise to transform scientific understanding of processes and patterns in human-caused mortality among wild animals subject to high rates of unregulated killing.

Methods

Data Collection and Preparation

We analyzed data acquired from the Department of Interior U.S. Fish & Wildlife service (USFWS) Mexican Wolf Recovery Program (MWRP) and their Office of Law Enforcement (OLE) in separate but overlapping datasets on (marked (hereafter collared), monitored Mexican wolves in the wild. The MWRP survival data include the monitoring history for all collared and monitored adult Mexican wolves in the wild since the beginning of the recovery program, 29 March1998- 31 December 2016; n=279 (monitored wolf pups were excluded from this dataset).

Because of the small wild population size and the captive breeding program, the majority of wolves in the Mexican wolf recovery area were collared or marked for monitoring. Only wild-born individuals that eluded capture remained unmonitored. Therefore, our analysis has the benefit of reducing (but not completely eliminating) a common bias in monitored animal studies when the marked subsamples are non-random, unrepresentative samples of the wild population, and may not have the same mortality risk as unmarked individuals [25-29].

The MWRP survival data contain the following individual covariates we used in our statistical analysis: monitoring start date, sex, and endpoint (i.e., end of monitoring time by: cause of death, lost to follow-up LTF, or removal by agency action). The endpoint ‘removal ’by agency action typically involved USFWS live-capture of a wolf from the wild followed by either placement in captivity or killing. The endpoint of LTF occurred when the telemetry equipment affixed to a wolf in the wild stopped functioning and the collar was never recovered. This could happen from mechanical/battery failure or destruction by external causes such as humans. Some wolves had multiple collars during their monitoring history as a result of recapture and recollaring. The vast majority of monitored time intervals (87.6%) were obtained using VHF collars, while the remaining 13% of monitored intervals were obtained with GPS collars. In our data, the average amount of time to LTF for wolves wearing VHF collars was 621 days, with a range of 7 to 3,079 days. The average battery life of a VHF radio collar is about 1,095 days, but wolves were often recollared. Only one individual disappeared while wearing a GPS collar, and this individual went LTF after 169 days. For recovered collars, cause of death was estimated by USFWS using standard methods following necropsy and radiography.

For each of the 279 wolves in the MWRP survival data (1998-2016), we estimated the time between collaring (monitoring start date) and endpoint in days (t), but we did so differently for surviving, dead, and LTF endpoints. We censored surviving wolves at the end of our study period. For LTF endpoints, we used the date of last telemetry contact. Inclusion of LTF as a separate, explicitly modelled endpoint was crucial for our inferences because of prior work showing that censoring LTF led to systematic under-estimates of the risk and hazard of poaching in other grey wolves [4,18]. Some wolves might have lived on for a time after their telemetry contact was lost, so LTF represents a systematic under-estimate of survival and hence of our parameter, t. We address the consequences and magnitudes of that bias in the Results. For our 'mortality 'and 'agency removal 'endpoints, we estimated t for wolves monitored by telemetry until death and the date of final removal to captivity by agency action, respectively.

Mortality endpoints obtained from the MWRP survival data were classified only as 'human 'or 'nonhuman 'in our first analysis step. The human-caused endpoint was identified in the MWRP data as mortality with 'likely and known human causes', without a more specific cause of death (e.g., vehicle collision, poaching). In the second analysis step, we turned to the OLE data, which categorized human-caused mortality by the following causes of death: vehicle collisions, trap, gunshot, blunt force trauma, 'unknown 'or 'other'. Using this data, we classified human-caused deaths as either poached (trap, gunshot, blunt force trauma) or non-criminal (vehicle collisions, 'unknown', 'other 'with no evidence of human intent). We used all human-caused deaths recorded up to and including 31 December 2016.

We focused our analysis on a time-varying covariate for policy-period (the policy intervention or IV in Table 4). Policy-period was binary for period of liberalized wolf-killing ('1') or stricter protections ('0') following exact policy change dates. Our policy covariate changed from 0 to 1 on October 10, 2005 when Standard Operating Procedure 13.0 (SOP 13) 'Control of Mexican wolves', was implemented by the Mexican Wolf Blue Range Reintroduction Project Adaptive Management Oversight Committee (AMOC) and changed back to 0 on December 2, 2009 (Table 1). SOP 13 liberalized wolf killing ('1') by establishing a "three-strikes" policy requiring the permanent removal of wolves implicated in three instances of predation on domestic ungulates during a one-year period. During SOP13, removals of wolves more than doubled relative to the previous seven years [30]. The policy was challenged in court and terminated by a federal judge on 2 December, 2009. However, a subsequent change in policy would again liberalize the killing of wolves. Thus, our policy covariate changed from 0 to 1 again from 16 January 2015 to 31 December 2016, which is the last date of our Freedom of Information Act data request (and the end of our study period). On 16 January 2015, the USFWS implemented a modification to the 1998 Endangered Species Act (ESA) 10j rule that expands the area where Mexican wolves can be released, allows permitted private entities to kill wolves on non-federal land if wolves are deemed to be a danger to domestic animals, and allows killing by government agents on private and state lands if wolves cause unacceptable predation on big game animals.

We also modeled season as a time-dependent covariate using an October-March (winter) and April-September (summer) split, because elsewhere season is known to mediate mortality in wolves [8,18,31,32]. For example, preliminary analysis of a population of Wisconsin wolves revealed winter periods were associated with increases in the hazard and incidence of various endpoints (LTF, poaching, nonhuman death) and at different rates [18]. To model both time-dependent covariates (policy and season), we created splits in each collared wolf's monitoring history. We refer to these splits as 'spells 'given they refer to briefer periods within an individual's monitoring time (see Table 1). In selecting the covariates of interest, we are following best practices of having at least 10 endpoint events per covariate [33-35]. We have therefore excluded from our models any covariates unless they are essential to control.

Nuisance variables are unlikely to confound our analyses as we discuss next. A hypothetical nuisance variable would have to not only correspond to the various changes in policy (2-3) but also be widespread across both NM and AZ across multiple jurisdictions (tribal, state, federal, county lands), and affect multiple independent adult wolves in packs occupying virtually exclusive home ranges. That leaves a climatic event or other widespread biotic event such as a disease with more than one change (to correspond with the policy changes of interest). We have searched both USFWS program documents and the scientific literature and have found no evidence of changes in environmental events or onsets of disease. Moreover, the covariates that may impact our hypotheses would need to affect the poaching, LTF and legal killing risks. Instead, environmental changes that may covary with the policy may in fact show changes to the 'nonhuman 'endpoint hazard and incidence, while perhaps affecting the changes in incidence of our anthropogenic endpoints (through endpoint interactions, see below) but not their hazards.

Statistical Methods

We employed endpoint-specific hazard and subhazard models in a competing risk framework, which are extensions of survival (or 'time-to-event') analyses, and a special case of multi-state models [36]. Survival analyses estimate the probability of observing a time interval from the start of monitoring (in our case, release to the wild with a functioning transmitter) to an endpoint, T, greater than some stated value t, S(t) =P(T>t) within a specified analysis time (our study period above). These techniques allow for calculating the (endpoint-specific) hazard function, hk (t), or the instantaneous rate of occurrence of a particular endpoint k conditional on not experiencing any endpoint until that time [37-40]. We used the hazard function to estimate the relative hazard of a collared wolf reaching an endpoint such as LTF given its survival to a particular date. We used semi-parametric Cox proportional hazard models to estimate covariate hazard ratios (HR) to model how endpoint-specific hk(t) changes as a function of survival time and model covariates. The estimation of covariate effects on the endpoint-specific hazard is modeled as $h_k(t) = h_{0k}(t)e^{(\beta_1 x_1 + \cdots + \beta_j x_j)}$, where $h_{0k}(t)$ is an unestimated baseline hazard function (i.e., semi-parametric) and β_j represent estimates of hazard ratios (HRs) for each covariate x_j (HR<1 represents a reduction in hazard and HR>1 an increase in hazard).

However, hazard rates do not consider competing risks. Competing risk analyses go beyond standard survival analyses by considering multiple endpoints simultaneously (e.g.: multiple causes of death, agency removal, or LTF). These models are useful for estimating the relative incidence of a particular endpoint, while accounting for other competing endpoints (e.g., the occurrence of human-caused death in the presence of a risk of nonhuman-caused death and LTF). In a competing risk framework, individuals can potentially experience the event of interest (i.e., end of monitoring time) by multiple, mutually exclusive endpoints, although only one is observed. Because the event of interest can only occur due to one endpoint, we refer to the endpoints as 'competing 'to bring about that event, and to the respective probabilities over time of that occurring as 'competing risks'.

Competing risk techniques estimate the cumulative incidence (CIF) curve for each endpoint, defined by the failure probability Prob(T<t, D=k); that is, the cumulative probability of endpoint k having occurred first (element D is an index variable that specifies which endpoint occurred) at time T, which specifies when the event happened within the study period interval defined over time t in the presence of other competing endpoints (i.e., subjects experiencing other endpoints are still considered at risk as individual wolves entered and left the risk set throughout the study period) [36,40,41].

Within the competing risks framework, Fine-Gray (FG) subhazard models estimate differences in CIFs for a given endpoint conditional on covariates [41,42]. FG models use regression techniques similar to the Cox model, except parameter interpretation changes as follows: estimates are interpreted as subhazard ratios (SHRs) or relative incidence (rather than HRs) in the presence of other endpoints (i.e., for each covariate x_j: SHR<1 represents a reduction in incidence and SHR>1 an increase in incidence). Although both hazard and competing risk models are informative, the competing risk models consider more information and provide greater predictive power [40,41,43].

Hence, while endpoint-specific Cox models and their HRs allowed us to test the hypothesis that liberalized wolf-killing affected the rate of occurrence (i.e., hazard or risk) of any endpoint relative to policy periods, the FG models and their SHRs allowed us to test if and how much liberalized killing affected the probability and incidence of endpoints, in addition to the potential simultaneous effects of other covariates. CIFs allowed us to visualize those effects on incidence while considering the prevalence of each endpoint in the population. Therefore, we used both hazard and incidence to infer the changes due to policy period and test the opposed hypotheses (Table 4).

Our Cox proportional hazards and Fine-Gray subhazard models comply with the appropriate number of events per variable recommended in the scientific literature to ensure accurate estimation of regression coefficients and their associated quantities for the endpoints of interest (poached, agency removals, LTF) (see Tables 2 and 3 below) [33,34,44].

Following recommendations for rigorous competing risk analysis [40,41,43,45], we reported results on all endpoint-specific hazards and CIFs to elucidate how hazards and incidences of multiple endpoints interact. For example, analysis of Wisconsin wolves suggested the increases in both the hazard and incidence of LTF during liberalized killing periods offset and potentially overcompensated for the smaller decreases in hazard and incidence of reported wolf-poaching estimated during those same periods [18].

Finally, we used Bayes 'Factor (BF) [46] to assess the strength of evidence for each of our alternative hypotheses and the null hypothesis for each poaching endpoint. We used the free BF online calculator found at http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/Bayes.htm, which assumes parameter estimates are normally distributed with known variance. The parameters used for each endpoint will be its point estimate and SE derived from the final Cox and FG models for testing HRs and SHRs, respectively. To assess the robustness of our conclusion to our prediction of the population distribution given our hypotheses and because prior theoretical support for any particular BF specification is scant, we assumed three different likelihood functions for the hypotheses 'predicted effect as recommended by [46]: (1) a half-normal function using the legal removal endpoint’s point estimates to model the expected standard deviation as SD=point estimate (2) a uniform function using the legal removal endpoint for Mexican wolves as the upper bound and 0 as the lower bound, and (3) a half-normal using endpoint-specific estimates from [18] to calculate the SD in the same manner as (1) [46, 47]. In doing so, we follow [46]’s recommendations to use likely values while keeping our predictions blind to the data as required by a registered report. In the absence of a base of prior literature to guide us, our use of these default variances maintains the required ‘blind ’to our data with a reasonable estimate of variability in the distribution around the point estimate. We use the legal removal endpoint estimates (rather than other imperfectly reported endpoints) in two specifications of our hypotheses because we know there is an effect (i.e., more wolves are killed legally during legalized killing periods). We report BFs for all HRs and SHRs of interest (i.e., reported poached and LTF). BFs strength of evidence for each hypothesis (or null) was interpreted as follows: $1/3<BF<3$, would be inconclusive evidence; $BF>3$, would be substantial evidence for the alternative hypothesis; $BF<1/3$, would be substantial evidence for the null hypothesis (i.e., no effect) [46, 47]. Given our three BF specifications for each endpoint parameter, we would conclude in favor of a particular hypothesis if most BF specifications (2 out of 3 BFs) support that conclusion. Accordingly, we might generate contradictory evidence (strong support for each hypothesis and the null) or inconclusive evidence (failure of any hypothesis to survive a majority of the BF specifications). For purposes of comparison, and to provide estimates of policy effects on ‘total potential ’(cryptic + reported) poached, we also aggregate and run all analysis on the new endpoint LTF+POA (including the procedure for pre-specified BFs as above).

Diagnostic Step

Because we used information from two data sources (MWRP and OLE datasets), we analyzed the data in two steps to provide more nuance about the effect of the policy intervention on mortality and disappearance. Both steps employed all survival and competing risk analyses previously described. We analyzed four endpoints: 'human, 'nonhuman', 'LTF', and 'agency removal '(Table 2). The drawback from this endpoint breakdown is the inability to conclude anything directly regarding any policy effects on subsets of anthropogenic mortality (e.g., poaching or non-criminal human-caused deaths).

Analyses

We added OLE data on human-caused endpoints and further specified which were poaching and which were deemed non-criminal (Table 3). However, this comes at the expense of lower number of observations in each human-caused endpoint and reduced statistical power (fewer events per variable, see previous section and Tables 2 and 3). Thus, results from the diagnostic step above will prove more statistically robust, but the exploration of the various anthropogenic endpoints is imperative, given evidence of different policy effects on each [18].

By evaluating the effect(s) of both liberalized killing periods (SOP 13 and revised 10(j) Rule), we will strengthen the inference about the policy intervention with a better case-control design (reverse-treatment or before-during-after-during). Therefore, both policy periods will be sampled twice.

Two divisions of the USFWS did the preliminary quality check on data. First the Mexican Wolf Recovery team collected mortality and disappearance data with a simple endpoint classification as 'died of nonhuman cause, died of human cause, legal removal by agency, disappeared (LTF). ’Next, for the subset of human-caused deaths above, the independent USFWS OLE classified deaths by cause (vehicle, poaching, accidental) and occasionally reclassified a human-caused death as nonhuman after detailed investigations, some or all of which included necropsy, radiography, or field investigation. Within our team, the two co-lead authors performed interobserver reliability tests by independently taking the data provided by the USFWS and summarizing it (Tables 2 and 3), then comparing and resolving discrepancies (arbitrated by the senior author in case of disagreements).

In Phase 2, the two co-lead authors separately prepared the data for analysis by aligning individual wolf survival histories with covariates and the intervention (policy period). The senior author Dr. Treves is Dr. Santiago-Ávila’s post-doc supervisor and Ms. Louchouarn’s PhD advisor. He and David R. Parsons served as a skeptical trouble-shooter for analyses, interpretation, and writing – blind to results until the Phase 2 analysis was considered complete by the two co-lead authors. This team organization and separation of powers was intended to reduce bias and improve the strength of inference.

Results

The two policy periods we examined which liberalized killing of Mexican wolves resulted in various changes in the hazard (risk) and incidence of endpoints of collared wolves relative to the two periods of stricter protection. We report changes in hazards from Cox models, subhazards from Fine-Grey (FG) competing risk models, and cumulative incidence functions (CIFs). We present point estimates and compatibility intervals in accordance with recent recommendations on the communication of results by moving beyond discussions of traditional 'statistical significance '[18,48]. We assessed the strength of evidence for our two sets of null and alternative hypotheses ('killing for tolerance 'or 'facilitated illegal killing') using Bayes Factors [46].

Models were run first on only the MWRP data, which had fewer mortality categories because the human caused mortality endpoint combined the criminal and non-criminal endpoints from the OLE data (Table 2). The only difference in model results during this portion of our analysis was represented in the human-caused endpoint and is reported in supplemental materials. Covariates of winter and sex did not significantly affect the results of any models, and therefore the most parsimonious model included the policy intervention without either covariate. The proportional hazard assumption of Cox models was met for all endpoints (see Suppl. Fig. S7-S8). For information about each model and their parameters, see supplemental materials (Suppl. Table S1). The best models revealed the following changes for collared Mexican wolves:

*Lost to follow-up (LTF)***.** Periods of liberalized killing were associated with a 121% increase in the hazard for the endpoint of LTF, relative to periods of stricter protection, compatible with a positive range (does not include zero) of +36% to +260% (HR = 2.21, $p < 0.001$; Table 5, Fig. 1-B). The proportion of collared wolves (CIF) with the endpoint of LTF increased by 128% (SHR

= 2.28, compatible interval= 38–276%, $p<0.001$) during periods of liberalized killing relative to periods of stricter protection (Table 6, Fig 1-A).

Reported poached. Periods of liberalized killing were associated with a 22% decrease in hazard for the endpoint of reported poached, relative to periods of stricter protection, compatible with a range that overlaps zero of -56% to +39% ($p = 0.407$; Table 5, Fig. 1-D). The proportion of collared wolves (CIF) with the endpoint of reported poached decreased by 31% (SHR=0.69, compatible interval = -62% to +25%, $p=0.226$) during periods of liberalized killing relative to periods of stricter protection (Table 6, Fig. 1-A).

Agency Removal. Periods of liberalized killing were associated with a 5% increase in hazard for the endpoint of agency removal, relative to periods of stricter protection (HR = 1.05, compatible interval includes zero, -41% to +88%, $p=0.863$, Table 5, Fig 1-C). The proportion of collared wolves (CIF) with endpoint of agency removal decreased by 4% (SHR = 0.959, compatible interval includes zero, -48% to +75%, $p=0.892$) during periods of liberalized killing relative to periods of stricter protection (Table 6, Fig 1-A).

Non-criminal. Periods of liberalized killing were associated with a 42% increase in hazard for the endpoint of non-criminal causes, relative to periods of stricter protection (HR=1.42, compatible interval includes zero, -25% to +171%, $p= 0.278$, Table 5). The proportion of collared wolves (CIF) with the endpoint of non-criminal increased by 27% (SHR=1.27, compatible interval includes zero, -33.7% to +146%, $p=0.463$) during periods of liberalized killing relative to periods of stricter protection (Table 6).

Natural. Periods of liberalized killing were associated with a 28% increase in hazard for the endpoint of natural causes relative to periods of stricter wolf protections (HR=1.28, compatible interval includes zero, -47% to +208%, p= 0.582, Table 5). The proportion of collared wolves (CIF) with the endpoint of natural increased by 32%, compatible interval= -44% to +211% (SHR=1.32, p=0.526; Table 6).

*Total potential poached***.** Periods of liberalized killing were associated with a 42% increase in hazard for the endpoint of total potential poached (aggregated endpoint of LTF and reported poached), relative to periods of stricter protections (HR=1.42, compatible interval = -0.7% to +103%, p= 0.05). The proportion of collared wolves (CIF) with endpoint of total potential poached increased by 38% (SHR=1.38, compatible interval = -5% to +101%, p= 0.095).

Bayes factor. We calculated Bayes Factor (BF) in three ways to model the expected predicted and maximum effect (Suppl. Table S2). At least two out of three BFs were 1<BF<3 for all endpoints (LTF, poached, total potential poached), suggesting our data is inconclusive as to our opposed hypotheses following our interpretation criteria (Table 4). Indeed, all three BF specifications proved inconclusive for the reported poached and total potential poached endpoints. However, our presumption that agency removal increased during periods of liberalized killing was not supported (HR = 1.05 and SHR=0.96), therefore specifications 1 and 2 seem meaningless (Table 7). The third scenario, based on endpoint-specific estimates of predicted effect as recommended by Dienes [46,47] remained meaningful (Table 7). See supplemental materials for information about the inputs used to calculate BFs (Suppl. Table S2).

Discussion

Here, we report a replication of the findings of Santiago-Ávila et al. [19] that grey wolves that disappeared from monitoring did so at higher rates during periods of reduced protections (i.e.

liberalized killing) than during periods of stricter protections under the US Endangered Species Act (ESA). We find stronger evidence for this pattern among collared Mexican wolves (C. l. baileyi) than was found among collared grey wolves in Wisconsin, USA [19]. Because the disappearance of collared wolves that are being monitored by VHF or GPS is most often caused by illegal activities [5,7,49], the present study further undermines the common assumption that animals lost to monitoring suffer from the same hazards and endpoints as those animals that are perfectly monitored [4].

In the paragraphs below, we first justify the assertion that the observed pattern in disappearances results from increases in cryptic poaching. Second, we conclude that our results support the 'facilitated illegal killing 'hypothesis and do not support the 'killing for tolerance ' hypothesis. Third, we discuss strength of inference in this subfield of wildlife science including protections against bias used in this study for the first time in this subfield to strengthen inference. Fourth, we propose that legal killing, non-lethal removal from the wild, and facilitation of cryptic poaching are all impediments to endangered wolf recovery under the ESA. Finally, we discuss the general lessons we draw from this study for the US federal agency implementing the ESA (USFWS) for wolves generally, for other countries and for anti-poaching research and intervention.

The relative stability in hazard and incidence of all known fates (not including the subset of radio- collared Mexican wolves that were designated LTF, lost to follow-up) between policy periods would suggest that if LTF wolves were in fact lost due to the same endpoints as monitored individuals with known fates, then we should observe relative stability in hazard and incidence of LTF between policy periods. This is not the case in either this study or [19]. Past work provides numerous independent lines of evidence that the majority of LTF could not be emigrants nor transmitter failures. First and most importantly, in both this study and that of Santiago-Ávila et al. [19], hazard and incidence rates of LTF changed in correlation with policies on legal killing, which could not plausibly have caused transmitter failures; see also [49] on different rates of LTF between hunting and non-hunting seasons. Also, battery life might be confounding LTF that occurred long after collaring. Contrary to this expectation, LTF was much shorter (average 788 days) than the average length of time to the natural mortality endpoint (1175 days). If battery life were the confounding factor, we would expect the average time to LTF to be more similar to the average time to the natural mortality endpoint. Second, if LTF were largely made up of emigrants from the Mexican wolf recovery area, some of these individuals would probably have been found dead in surrounding areas by citizens with nothing to hide who presumably would have reported their observations to authorities. The USFWS databases we used contained no such cases. Therefore, LTF wolves were most likely killed and the evidence of the illegal action was concealed, e.g. by destruction of transmitters. Such cryptic poaching was first estimated by Liberg et al. [7], and subsequently explored in Treves et al. [4,5] and Santiago-Ávila et al. [19], and could have been exacerbated in the Mexican wolf recovery area by the policy of sharing radiofrequencies of collared wolves with members of the public [50]. We conclude that our results on the disappearances of collared Mexican wolves reinforce those first reported by Treves et al. [5] which demonstrated the bias introduced by excluding disappearances of marked wolves in mortality analyses for four endangered wolf populations.

We tested opposed hypotheses about the effect of legalizing killing or removal of individuals of an endangered species on the survival of collared individuals remaining in the wild. The USFWS, responsible for implementing the ESA for terrestrial species, has been a particular supporter of the ‘killing for tolerance ’hypothesis. It has repeatedly invoked this hypothesis in its endangered grey wolf management under the assumption that government-permitted killing of grey wolves would mitigate or prevent illegal killing and raise public tolerance for wolves, so ongoing recoveries would not be stopped or slowed by illicit resistance [9,51]. A federal court rejected that argument as a speculative approach to abridging the ESA prohibitions that are explicitly aimed at preventing killing [9] but in 2017, another court seemed to defer to the agency on this point when it wrote '… it is clear that in drafting the present Section 10(j) rule, the take provisions are critical to conciliating those opposed to the reintroduction effort,…' [52, p. 43]. Despite this deference, the court remanded the rule to the USFWS to repair its deficiencies. We recommend the USFWS abandon the expectation and repudiate the oft-repeated and unsupported notion that liberalizing killing would reconcile opponents of wolf protection. The results of this study support the alternative and mutually exclusive hypothesis that liberalized killing policies facilitate illegal killing and join a growing body of evidence that suggests liberalized killing policies lower tolerance for wolves and slow wolf population growth substantially more than expected from legal killing rates.

This study, therefore, adds to the literature regarding policy effects on anthropogenic causes of wolf mortality. Human dimensions research using focus groups and mail-back surveys measuring attitudes of Wisconsin residents in and out of wolf range found that respondents ' tolerance for wolves decreased and reported respondents 'willingness to poach wolves increased after wolf-killing was liberalized seven times between 2003 and 2013 [12,14]. Moreover, calls for more killing of wolves followed relaxing ESA protections [13]. Further, a study of population dynamics showed with 92% certainty that there had been an unexplained decrease in the growth of wolf populations in both Wisconsin and Michigan after wolf-killing was liberalized, independent of the number of wolves killed legally [15,16]. The latter authors could only explain the repeated, parallel slow-downs by the existence of undetected mortality [15,16], a result which withstood a series of attempts at rebuttal that did not include new data [17] and in one case muddied the waters with errors and omissions [23]. The conclusion from Chapron & Treves [15,16] that the length of the policy period was predictive of the population growth slow-down independent of the reported number of wolves killed is consistent with our finding that exposing Mexican wolves to liberalized killing was associated with higher hazard and incidence of LTF, not predicted by the hazard or incidence of wolves legally killed. Therefore, three independent lines of evidence point in the same direction and opposite to the USFWS hypothesis about liberalizing killing.

Moreover, here we provide more direct, stronger inference than ever before, against the government's 'killing for tolerance 'hypothesis in a new population of wolves. Given that periods of reduced protections allowed for greater flexibility in legal lethal actions towards Mexican wolves, we expected to observe a higher increase in the hazard of agency removals during periods of reduced protections (HR = 1.05). Therefore, we reject the possible explanation that increased hazard or incidence of agency removal (SHR = 0.95) during liberalized killing periods is somehow leading to an increase in disappearances (such as emigration or super-additive mortality). Rather, it was the policy period announcement or its duration per se that had the effect of increasing collared wolf disappearances. Chapron & Treves [15] proposed that reducing protections for wolves sends a policy signal lowering the value of wolves to the public including would-be poachers or reducing the likelihood of enforcement against poaching. The hypothesized 'policy signal 'seems to convey that either the lives of individual wolves are perceived as less valuable, the benefit of wolves has declined, or prosecution of poachers will relax. We find little evidence to support the latter, because the LTF endpoint represents destruction of evidence of poaching. Instead, would-be poachers appeared to have opted to act cryptically or increase their concealment of evidence during periods of reduced protection, an inference that is supported by a recent news report [53]. The inference that would-be poachers became more concerned with law enforcement while increasing their poaching rates is consistent with the current study in Mexican wolf range and that of Santiago-Ávila et al. [19] for Wisconsin wolves.

Until sophisticated, replicable studies of confirmed poachers and their attitudes are conducted, we cannot know if would-be poachers responded to policy signals by repeating past poaching behaviour with the addition of more concealment of evidence, or if new actors began poaching with concealment. We predict a mix of both patterns, but a preponderance of the poaching during periods of liberalized killing was by individuals who now chose to conceal evidence. That pattern would be supported if the USFWS began to give out radiofrequencies of collared wolves or otherwise changed agency conduct in the field in such a way as to expose collared wolves to higher risk.

In past studies, including those from the US Midwest, as well as one performed in Scandinavia [8], periods of reduced wolf protection were associated with significant increases in hunting or government lethal removal of wolves. Rates of wolf disappearances or poached wolves also increased, but not as drastically as we observe here. These less drastic changes for other endpoints may be a result of 'cleaning up the numbers '[8,19]; more wolves reach the endpoint of legal killing before they can succumb to some other endpoint, such as reported or cryptic poaching, thereby muddying our understanding of the effect of legal killing policies on other fates of collared wolves. Our study is not confounded by any effect of 'cleaning up the numbers 'because Mexican wolves were not subject to higher hazard of legal removal and the incidence of wolves lost by agency removal decreased, yet radio-collared wolves disappeared at higher rates. Rather, we detected significantly more hazardous conditions for the critically endangered Mexican wolf when the USFWS reduced ESA protections independent of agency removal.

Regarding the unexpected finding that the rate of agency removal changed little as policy periods changed from stricter protection to reduced protection, we re-examined our starting assumption. We based our starting assumption of a higher increase in agency removals during two periods of reduced protection (SOP 13 and revised 10(j) rule) on two pieces of likely misleading information. First, by examining the raw numbers of agency removals in tables 3 and 4, it appears that the prevalence of wolves being removed during periods of liberalized killing is about 150% of that being removed during periods of stricter protections. However, hazard and incidence, as we calculate here with survival analysis methods, are based off the sum of all the days each wolf was exposed to each endpoint (i.e. their aggregate time-at-risk). Therefore, time-at-risk is much greater than the number of days over particular policy periods (as reported in tables 3 and 4). We further based our expectations of the change in agency removals over policy periods on a source which claimed that agency removals more than doubled during the SOP 13 policy relative to prior years (see Methods) [31]. However, Fitzgerald [31] lacks any accompanying data.

An often-overlooked aspect of wolf mortality reporting is how the agency classifies cause of death when human-caused. Treves et al. [4] mentioned consolidating causes of death such as non-permitted trapping, shooting, poisoning into one category of poaching, especially when the agency might otherwise misidentify the primary cause of death. We observed a pattern in the data for Mexican wolves that was not detected in Santiago-Ávila et al. [19]. In particular, the incidence of the non- criminal endpoint increased by 27% during periods of reduced wolf protections (table 6). Our results (HR = 1.42 and SHR = 1.27) for the non-criminal endpoint suggest the possibility that USFWS staff classified a greater number of anthropogenic causes of death as non-criminal during periods of reduced protections. This may be a logical result of liberalizing killing, as less killing is legally classified as 'poaching'. Indeed, some poaching could merely have been reclassified as non-criminal by USFWS staff using subjective definitions and thereby confirming the (erroneous) perception that poaching had diminished because of 'increased tolerance'. However, the non-criminal endpoint includes vehicle collisions, and other human-caused mortalities that were classified as non-criminal after an investigation by the USFWS Office of Law Enforcement (OLE). Therefore, it is impossible to know the real cause of the increase in hazard and incidence of the non-criminal endpoint during periods of reduced protections without knowing more about the OLE investigations. We received two datasets for Mexican wolves from the USFWS. One set pooled all anthropogenic mortality in one category, which is clearly less useful for analyses such as ours that can tease apart the effect of policy interventions on specific endpoints. Hence, the second dataset from OLE which assigned criminal and non-criminal causes of death and also distinguished further subcategories was much more useful to us. We surmise it would have been more useful to managers and the public also. Therefore, we recommend the USFWS share data on mortality of endangered species that are disaggregated into no fewer than four categories (legal, illegal, vehicle collision and natural), report disappearances (LTF) systematically in the same tables along with start and end dates for time on the air. By the same token, too many categories of poaching as a cause of death can obscure the priority of illegal killing.

Some readers might wonder if frustration among would-be poachers rose in the Mexican wolf range because agency removals did not change between periods, despite the policy signal that protections were loosened. However, we argue that if this were a valid explanation, we would expect frustration with the lack of change in the rate of lethal actions to be prevalent during both policy periods, and we would anticipate the rate of disappearance of wolves to be comparable during the two periods, which we do not see. We would, therefore, expect a different pattern than in the US Midwest where the agency did use liberalized killing periods to lethally remove wolves at higher and increasing rates [19]. Proponents of the frustration hypothesis claimed Wisconsin's would-be poachers were frustrated when protections were tightened [11]. No plausible cognitive mechanism that would differ between would-be poachers of Wisconsin and those of New Mexico/Arizona has been presented. The frustration hypothesis requires that two different cognitive behavioural mechanisms exist in the two populations, which does not seem parsimonious and is not consistent with the attitudinal data from Wisconsin (see above). This Mexican wolf study cannot support the USFWS idea that without legal recourse, actors would take matters into their own hands, because would-be poachers observing agency removals would be expected to be as frustrated in and out of the policy periods examined here. Because this study and Santiago-Ávila et al. [19] used time-series analysis for before-and-after comparison of interventions (BACI) without randomization to control or treatment, it provided stronger inference than prior work that relied on correlation and single point estimates [11,15,16], by accepted standards from other fields [54,55]. The standards have been explained at length in [56–58] in relation to their application to evaluation of methods to prevent carnivore attacks on livestock. The current study also integrated several novel protections against bias that further strengthen the inference. An additional reduction in confounding variables in this study is that just over half of the wild Mexican wolves were marked and monitored, compared to the average of 13% marked in the Wisconsin population. That increases the strength of generalizations about Mexican wolves as a whole and increases our confidence in parameter estimates for known and unknown fates in the present study. Therefore, only evidence using experimental controls could achieve stronger inference than does the current study.

Furthermore, this study directly attempted to reduce bias in the following ways. First, we used official data as classified by the management agencies in charge (both USFWS MWRP and its OLE), not by us. Therefore, if any bias exists in the classification scheme, it reflects classification decisions by the agencies and probably random error given the agencies were apparently unaware of the hypotheses. Second, by publishing methods before completing an analysis, we ensured that the analyses could not be amended to support one or another hypothesis. Further, peer-review of the methods helped our team develop stronger analysis methods that would allow us to better interpret the strength of the evidence. Fourth, we used BF rather than arbitrary traditional significance thresholds to assess the relative strength of evidence for our hypotheses. Our team also developed internal safeguards by having separate members of the team independently interpret the results without taking part in analysis. The ESA requires the use of best available scientific and commercial data, hence policymakers wishing to implement the ESA as intended by Congress can take comfort that the science has advanced to the highest level, rather than continuing to debate with imperfect evidence as the scientific community did from 2013 to 2019 (reviewed in the section on hypothesis tests)

We conducted a BF analysis to quantitatively assess the strength of evidence for our two competing hypotheses and the null hypothesis. Calculating BFs for each endpoint allows us to go beyond significant or non-significant results to examine whether non-statistically significant results truly represented evidence against either hypothesis [47]. To determine whether our results represent evidence for or against either opposing hypothesis or the null hypothesis, Dienes [47] recommends using prior published research to determine what our theory predicts. The theories we test here have not been widely tested; therefore, our best source for endpoint-specific parameter estimates came from Santiago-Ávila et al. [19].

Those prior results from an unrelated dataset provided a comparison for the Mexican wolf results in a registered report but prior to analysing the data. We calculated the BFs of our results using three specifications (defined in Methods) and, following our stated interpretation criteria for BFs detailed in the Methods and table 1, none of our endpoints of interest (LTF, poached, total poached) provided conclusive evidence for either of our alternative hypotheses (electronic supplementary material, table S2). However, we have greater confidence in the results calculated using the third specification; the half- normal distribution calculated using prior data; Santiago-Ávila et al.'s [19] HR and SHR values for the LTF and 'reported poached ' endpoint. Our rationale for our confidence in the third specification is as follows: (i) the effects observed in both studies are endpoint-specific; therefore, the estimates used ([19]'s HR and SHR) are a result of similar mechanisms, rather than using estimates from different endpoints that probably result from different mechanisms, as in our other two specifications, both based on using the Mexican wolf agency removal endpoint for comparison; (ii) when submitting our methods as a registered report, we had to make a 'blind '(prior to analyses) assumption regarding the change in agency removal endpoint for Mexican wolves (i.e. that agency removals would increase with liberalized protections), which proved counterintuitively unchanged between policy periods, thereby eliminating its potential predictive power; and (iii) the agency removal and reported poached HRs are opposite in direction and the same occurs with the agency removal and LTF SHRs, so those estimates do not provide plausible parameters for the reported poached and LTF endpoints.

The only BF that was conclusive was the support for the 'facilitated illegal killing ' hypothesis shown by the increase in disappearances of collared Mexican wolves (LTF, table 7). By contrast, the corresponding BF for the total potential poached endpoint fell below the criterion level, because the aggregated LTF and reported poached endpoints ran in opposite directions, but the increase in the proportion of wolves with fate LTF was nearly three times greater than the decrease in the proportion of wolves reported poached. None of the BF analyses supported the killing for tolerance hypothesis (table 7). We conclude the USFWS claim in federal court that lessening ESA protections with the 10(j) rule would reconcile opponents of reintroduction and in turn be harmless for the Mexican wolves in the wild [59] now seems untenable.

Implications for endangered species

Policy interventions should be effective, i.e. achieve their goals, without serious, unwanted side effects. This study finds that for Mexican wolves, there were serious side effects of the liberalized killing policies. The increase in disappearances of Mexican wolves we detected was substantial during those periods of reduced protections, despite a lack of change in the rate of government removal of wolves. Unplanned, unregulated disappearances are wasteful: a waste of taxpayer money spent on telemetry and relocation to the wild; a loss of individual animals that are unique and irretrievable by known technology; a waste of private resources used in their captive breeding; and undermines the role of the federal government as trustee of US wildlife since 1842 [3]. The effect we found also demonstrates widespread unlawful disregard for the most popular environmental law ever passed in the USA [60].

Further, the policy of liberalizing killing cannot be justified by the vague and indirect claim that it speeds population growth at the expense of individual survival because USFWS data show that the Mexican wolf population declined from 55 to 42 during a 6-year period of liberalized killing from 2004 to 2009 [50]. Similarly, after the implementation of the 10(j) revised rule in 2015, the Mexican wolf population in the wild declined 12%, partially rebounded the next year, and did not change by 2017 when the court order remanded the revised 10(j) rule to the USFWS. Thereafter, growth continued at the prior rate averaging 22% per year (fig. 5 in [50]). The latter finding replicates that of Chapron & Treves [15–17,23] for Wisconsin's and Michigan's grey wolves. Currently, the Mexican wolf population numbers 163 [61]. In view of these results, we hypothesize that population growth will slow, halt and maybe even reverse, given currently authorized liberalized killing, with the effect on growth mediated by the magnitude of the policy signal on disappearances; that is, on cryptic poaching rather than agency removals. However, protections for Mexican grey wolves could be strengthened in a revised 10(j) rule being considered by USFWS.

In this context, the balance tilts towards the Mexican wolves by law (quoting the court in Center for Biological Diversity v. Jewell (2018) 'Harm to endangered or threatened species is considered irreparable harm, and the balance of hardships will generally tip in favor of the species. See Marbled Murrelet v. Babbitt, 83 F.3d 1068, 1073 (9th Cir. 1996) ("Congress has determined that under the ESA the balance of hardships always tips sharply in favour of endangered or threatened species."); Amoco Prod. Co. v. Vill. of Gambell, AK, 480 U.S. 531, 545 (1987) ("Environmental injury, by its nature, can seldom be adequately remedied by money damages and is often permanent or at least of long duration, i.e. irreparable. If such injury is sufficiently likely, therefore, the balance of harms will usually favor the issuance of an injunction to protect the environment.") 'p. 23, Docket CV-15-00019-TUC-JGZ, U.S. District Court Arizona, 2018).

The issue goes beyond Mexican wolves. As recently as 14 December 2020, the USFWS continued to espouse the unsupported view that reducing or removing ESA protections will help individual wolves to survive and help wolf populations to recover [51]. In the latter letter to the State of California Fish & Game Commission, the USFWS cited outdated studies that have been superseded since 2016, with frequent communications [62–64].

There are lessons in the current work that have implications beyond the USA and beyond wolves. We suggest other national policies for killing large predators (or other non-humans) to raise tolerance or lower poaching should be scrutinized for strength of inference and the quality of evidence (e.g. [21,22]). The notion that protection for large carnivores generates poaching as a form of rural resistance has merit, but the suggestion that relaxing protections is the solution is no longer credible (contra Kalternborn & Brainerd [65]). Similarly, the long-held notion that without compensation for losses, affected people will kill wildlife illegally, needs re-examination in the light of our current results. The alternative is that compensation might encourage hostage-taking, i.e. an escalation by affected parties to threatening endangered species if they are not better compensated for property damages. Our findings join an active debate about leniency versus enforcement as more functionally effective conservation interventions.

In cases of wildlife trade, such as ivory or rhino horn, arguments for leniency are focused around creating a legal trade to inhibit the lucrative illegal trade [66,67]. However, sources of wildlife crime stemming from conflict, such as in the case of grey wolves, may not be effectively understood, nor managed in the same ways. Leniency has been tried for grey wolves, and evidence suggests leniency fails to achieve greater tolerance and reduced wildlife crime. On the contrary, leniency is associated with increased poaching of wolves in the USA. We encourage the scientific evaluation of all candidate interventions as experiments, preferably with suitable comparisons or even experimental controls, with safeguards against bias.

We suspect wolves are not exceptional among large carnivores regarding the effect of relaxing or fortifying legal protections, because the same justifications for liberalizing killing of brown bears and lions have been used whenever prohibitions on hunting or other lethal management are proposed [10,68–70]. We urge that similar studies be completed to examine whether there is in fact a difference in how liberalized killing policies affect other large carnivores. Further, we hypothesize that it is the attitudes and values of the human actors that are the unifying variable, not the nature of the environmental component or species of wildlife. Our interpretation of these findings, given their consistency with past studies, is that when policies are implemented which reduce the value of non- human beings, such as policies which enable their killing for the sole benefit of human actors, there will be increased harm to those beings and damage to the environment, including crimes.

Tables & Figures

Table 1. Example of monitoring history of a hypothetical wolf ID, broken up into spells for the integration of time-dependent covariates. We use “analysis time” for the time intervals and order of spells, as covariates change (either policy or season). The endpoint categorical variable is only reflected for the last spell, which corresponds to when monitoring ended (at t=250 in this hypothetical case).

wolf ID	analysis time when spell begins	analysis time when spell ends	policy	season	endpoint
MX1209	0	57	1	1	
MX1209	57	140	1	0	
MX1209	140	350	0	0	2

Table 2. Number of endpoints (unique wolf IDs) during periods of liberalized killing or periods of stricter protections for step 1 (diagnostic step). Wolves that survived to the end of the study period (n= 52) are omitted here and censored in analyses. The study period spanned March 29, 1998 to December 31, 2016 inclusive.

Endpoint	Stricter protection, policy period = 0 (t = 4,621days)	Liberalized killing policy period = 1 (t = 2,230 days)
Agency Removal	28	20
Non-human cause	11	11

Human cause	55	35
LTF	27	40

Table 3. Number of endpoints (unique wolf IDs) during periods of liberalized killing or periods of stricter protections for step 2 using OLE data from investigations of suspicious deaths. Wolves that survived to the end of the study period (n= 52) are omitted here and censored in analyses. The study period spanned March 29, 1998 to December 31, 2016 inclusive.

Endpoint	Stricter protection, policy period = 0 (t = 4,621days)	Liberalized killing policy period = 1 (t = 2,230 days)
Agency Removal	28	20
Poaching	35	17
Natural Death	11	11
Non-criminal	20	18
LTF	27	40

Table 4. Relationship between our hypotheses, proposed analyses and interpretation of outcomes (including contingent interpretation and synthesis of model results). HRpoa refers to the hazard ratio of the poaching endpoint, while HRltf refers to the hazard ratio of the LTF endpoint.

Question	Hypotheses	Sampling plan (e.g. power analysis)	Analysis Plan	Interpretation given different outcomes (note see footnote and main text for Bayes factor specifications †)

Do hazard rates or cumulative incidence of death by poaching or disappearance (DV) of wild, collared adult Mexican grey wolves change after policies change (IV) from strict protection to liberalized killing and back again.	'Killing for tolerance' predicts the hazard and incidence decline for the endpoint 'poached' (poa) or the endpoint LTF when the IV of policy period liberalizes wolf-killing.	All collared wild Mexican gray wolves from MWRP and OLE 1998-2016 (n=279) A diagnostic test is run on the samples with four endpoints (human, nonhuman, removal, LTF) before proceeding to the analysis plan (See Diagnostic Step). See Tables 2 and 3 for endpoint-specific sample sizes split by the IV of policy period.	For MWRP and OLE datasets: Endpoint-specific Cox multiple regression models (for each endpoint) on the IV of policy period and other covariates. Competing risk Fine & Gray multiple regression models (for each endpoint) on the IV of policy period and other covariates. CIFs allow for analysis of population effects (incidence) while considering the prevalence of each endpoint in the population.	HRpoa and HRltf are <1 OR (HRpoa has to be <1 and greater in magnitude than any increase in HRltf OR HRltf has to be <1 and greater in magnitude than any increase in HRpoa) AND endpoint-specific CIFs estimate which endpoint has a greater effect on the population (from Fine-Gray models of competing risks) The criterion for determining if TOTAL potential poached' probability f declined is a decline in the combined incidence of LTF and POA.

	'Facilitated illegal killing' predicts the hazard and incidence increase for the endpoint 'poached' (poa) or the endpoint LTF when the IV of policy period liberalizes wolf-killing.	All collared wild Mexican gray wolves from MWRP and OLE 1998-2016 (n=279) A diagnostic test is run on the samples with four endpoints (human, nonhuman, removal, LTF) first before proceeding to the analysis plan (See Diagnostic Step). See Tables 2 and 3 for endpoint-specific sample sizes split by the IV of policy period.	For MWRP and OLE datasets: Endpoint-specific Cox multiple regression models (for each endpoint) on the IV of policy period and other covariates. Competing risk Fine & Gray multiple regression models (for each endpoint) on the IV of policy period and other covariates. CIFs allow for analysis of population effects (incidence) while considering the prevalence of each endpoint in the population.	HRpoa and HRltf are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) AND endpoint-specific CIFs estimate which endpoint has a greater effect on the population (from Fine-Gray models of competing risks). The criterion for determining if 'TOTAL potential poached' probability for wolves declined is a decline in the combined incidence of LTF and POA.

†Following reviewer recommendations, we will use Bayes 'Factor (BF) using three specifications. BF estimates the strength for our alternative and null hypotheses for particular endpoints, and allows us to assess insensitivity of the data to resolve differences between hypotheses. For purposes of comparison, and to provide estimates of policy effects on 'total potential '(cryptic + reported) poaching,

we proceed to aggregate poaching endpoints and run all analysis on the new endpoint LTF+POA (including BFs) (see Statistical Methods section above).

Table 5. Cox model of cause-specific hazards for each endpoint for 279 collared Mexican wolves. HR<1 represents a reduction in hazard during periods of liberalized killing (lib_kill =1) and HR > 1 denotes an increase in hazard, HR < 1 a decrease, and HR=1 no change in hazard. Only the most parsimonious model is presented (Suppl. Mat. For all models), In all cases the proportional hazard assumption of the Cox models was met. Comp. Int. = compatible interval around point estimates.

	Endpoint									
	Lost to follow-up (LTF)		Agency Removal		Reported Poached		Non-Criminal		Natural	
Variable	HR	Comp. Int.	HR	Comp. Int.	HR	Comp. Int.	HR	Comp. Int.	HR	Comp. Int.
Liberalized Killing Periods (lib_kill)	2.21*	1.36-3.60	1.05	0.59-1.88	0.78	0.44-1.39	1.42	0.75-2.71	1.28	0.53-3.08

* $p < 0.001$, all other results had $p > 0.05$

Table 6. Fine Grey competing risk models of cause-specific subhazard (SHR) for each endpoint for 279 collared Mexican wolves. SHR < 1 represents a reduction in the incidence of the endpoint during periods of liberalized killing (lib_kill = 1) and SHR > 1 an increase in incidence. SHR = 1 would represent no change in relative incidence. Only the most parsimonious model is presented (see Suppl. Mat. Comp. Int. = compatible intervals around point estimates.

Variable	Endpoint									
	Lost to follow-up (LTF)		Agency Removal		Reported Poached		Non-Criminal		Natural	
	SHR	Comp. Int.	SHR	Comp. Int.	SHR	Comp. Int.	SHR	Comp. Int.	SHR	Comp. Int.
Liberalized Killing Periods (lib_kill)	2.28*	1.38-3.76	0.96	0.53-1.75	0.69	0.38-1.25	1.27	0.66-2.46	1.32	0.56-3.11

* p-value < 0.001

Table 7. Bayes Factor (BF) calculations for reported poached. LTF and aggregated 'total potential poached '(LTF+POA) endpoints for collared Mexican wolves using three specifications; (1) a half-normal distributions using the Mexican wolf agency removal endpoint point estimate of HR and SHR; (2) a uniform function using the agency removal endpoint for Mexican wolves as the upper bound and 0 as the lower bound, and (3) a half normal distribution and the analogous estimates of HR and SHR from Santiago-Avila et al. [18]; see Suppl. Mat. For all parameters). BFs strength of evidence for each hypothesis (or null) was interpreted as follows: 1/3<BF<3 (ref), would be inconclusive evidence; BF>3 would represent substantial evidence for the alternative hypothesis; BF<1/3, would represent substantial evidence for the null hypothesis of no association.

	Endpoint					
	LTF		POA		LTF+POA	
BF Specifications	HR	SHR	HR	SHR	HR	SHR
(1) half-normal w/MX-agency removal	1.8	0.69	0.89	1.14	1.15	0.76
(2) uniform w/upbound-MX agency removal	1.41	1.30	0.93	0.92	1.09	1.19
(3) half-normal w/WI POA	8.08	8.08	1.25	1.63	1.35	0.47

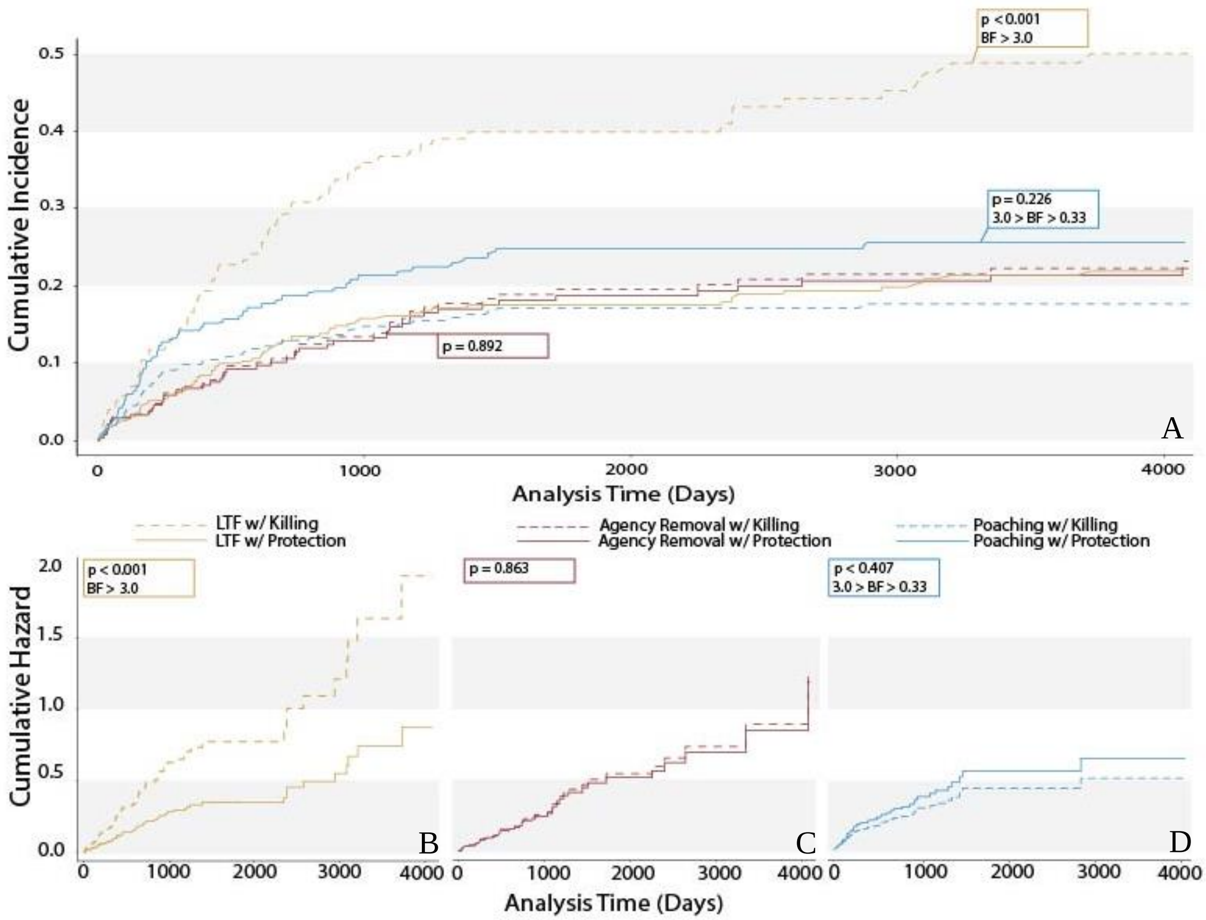


Figure 1. Cumulative incidence functions (CIF, panel A) derived from Fine-Grey subhazard models and hazard functions derived from univariate Cox models (panels B-D) for 279 collared Mexican wolves during periods with reduced protections for wolves (liberalized killing periods = 1, dashed lines) and periods with stricter protections (solid lines) for three independent endpoints (LTF, n=67; reported poached, n=52; and agency removal, n= 48). See Supplemental Materials for the cause-specific hazard functions (Suppl. Fig S1-S2) and CIFs (Suppl. Fig. S3-S4) for the natural and non-criminal endpoints. Bayes Factors (BF) support a hypothesis over the null when >3, support the null over a hypothesis when <0.33 and represent inconclusive evidence for either hypothesis with 0.33<BF<3 [46]. Panel A. CIF curves show the proportion of collared wolves

disappearing (LTF, yellow line) was significantly greater than other endpoints, during periods of liberalized killing (SHR = 2.28, Compatible interval = +0.38 to +2.76, p <0.001). Panels B-D. Lines show cumulative hazard over analysis time (days of monitoring each wolf). B. LTF HR = 2.21 (compatible interval = +0.36 to +2.60). C. Agency removal HR = 1.05 (Compatible interval = -0.41 to +0.88). D. Reported poached HR= 0.78 (Compatible interval = -0.56 to +0.39).

Data Accessibility

The pre-registered Stage 1 report can be found on the Open Science Framework at the following link: https://osf.io/f2kmb/?view_only=28de338783724f4e96c4bede49d795d1

The raw datasets sources from the U.S. Fish and Wildlife Service and the Office of Law Enforcement have been submitted to Dryad and can be found at the following link: https://datadryad.org/stash/share/6MNn9bAx93rtciLPNSFZN0DG0cGgoI7Z-OSZ83DN2LQ

We have also included the prepared data in our Dryad submission which is ready to be run through the STATA code included starting on p. 11 of Supplemental Materials.

Acknowledgments

We thank the US Fish & Wildlife Service Mexican Wolf Recovery Program and Office of Law Enforcement staff, especially Maggie Dwire and John Oakleaf, for data collection, provision and assistance in data interpretation. We thank Judy Calman for assistance in obtaining agency data. We thank the UCLA Law School Animal Law & Policy Grants Program and Therese Foundation for funding This article does not necessarily reflect the views of the institutions or agencies involved.

References

1. Woodroffe R, Ginsberg JR. Edge Effects and the Extinction of Populations Inside Protected Areas. Science [Internet]. 1998 Jun [cited 2020 Jan 8]; 280:2126-2128. Available from http://science.sciencemag.org/content/280/5372/2126.abstract
2. Wydeven A, Mladenoff DJ, Sickley TA, Kohn B, Thiel P, Hansen JL. Road density as a factor in habitat selection by wolves and other carnivores in the Great Lakes region. Endangered Species Update [Internet]. 2001 [cited 2020 Jan 8]; 18(4):110–114. Available from https://pdfs.semanticscholar.org/ffc8/2dbae01f91009758343c0cdd6c6edeb9d573.pdf
3. Treves A, Chapron G, López-Bao J V, Shoemaker C, Goeckner AR, Bruskotter JT. Predators and the public trust. Biol. Rev. [Internet]. 2017 Nov [cited 2020 Jan 8]; 92(1): 248–270. Available from https://doi.org/10.1111/brv.12227
4. Treves A, Artelle KA, Darimont CT, Parsons DR. Mismeasured mortality: correcting estimates of wolf poaching in the United States. J Mammal [Internet]. 2017 Oct [cited 2020 Jan 8]; 98(5): 1256-1264. Available from https://doi.org/10.1093/jmammal/gyx052
5. Treves A, Langenberg JA, López-Bao J V., Rabenhorst MF. Grey wolf mortality patterns in Wisconsin from 1979 to 2012. J Mammal [Internet]. 2017 Feb [cited 202 Jan 8]; 98(1):17–32. Available from 573http://dx.doi.org/10.1093/jmammal/gyw145
6. Ripple WJ, Estes JA, Beschta RL, Wilmers CC, Ritchie EG, Hebblewhite M, Berger J, Elmhagen B, Letnic M, Nelson MP et al. Status and ecological effects of the world's largest carnivores. Science [Internet]. 2014 Jan [Cited 2020 May 19]; 343 (6167); 1241484. Available from https://science.sciencemag.org/content/343/6167/1241484.full
7. Liberg O, Chapron G, Wabakken P, Pedersen HC, Hobbs NT, Sand H. Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. P Roy Soc B-Biol Sci [Internet]. 2012 Mar [cited 2020 Jan 8]; 279:910–915. Available from http://www.ncbi.nlm.nih.gov/pubmed/21849323
8. Suutarinen J, Kojola I. Poaching regulates the legally hunted wolf population in Finland. Biol Conserv [Internet]. 2017 [cited 2020 Jan 8]; 215:11–18. Available from http://www.sciencedirect.com/science/article/pii/S0006320717302148
9. Treves A. Hunting for large carnivore conservation. J Appl Ecol [Internet]. 2009 Nov [cited 2020 Jan 8]; 46(6):1350–5641356. Available from https://doi.org/10.1111/j.1365-2664.2009.01729.x
10. Olson ER, Stenglein JL, Shelley V, Rissman AR, Browne-Nuñez C, Voyles Z, et al. Pendulum swings in wolf management led to conflict, illegal kills, and a legislated wolf hunt. Conserv Lett [Internet]. 2015 Sept [cited 2020 Jan 8]; 8(5): 351-360. Available from https://doi.org/10.1111/conl.12141
11. Treves A, Naughton-Treves L, Shelley V. Longitudinal analysis of attitudes toward wolves. Conserv Biol [Internet]. 2013 Apr [cited 2020 Jan 8]; 27(2):315–323. Available from http://www.ncbi.nlm.nih.gov/pubmed/23293913

12. Browne-Nuñez C, Treves A, MacFarland D, Voyles Z, Turng C. Tolerance of wolves in Wisconsin: A mixed-methods examination of policy effects on attitudes and behavioral inclinations. Biol Conserv [Internet]. 2015 Sept [cited 2020 Jan 8]; 189: 59–71 Available from https://doi.org/10.1016/j.biocon.2014.12.016

13. Hogberg J, Treves A, Shaw B, Naughton-Treves L. Changes in attitudes toward wolves before and after an inaugural public hunting and trapping season: early evidence from Wisconsin's wolf range. Environ Conserv [Internet]. 2015 May [cited 2020 Jan 8]; 43(1): 45–55. Available from https://doi.org/10.1017/S037689291500017X

14. Chapron G, Treves A. Blood does not buy goodwill: allowing culling increases poaching of a large carnivore. P Roy Soc B-Biol Sci [Internet]. 2016 May [cited 2020 Jan 8]; 283(1830): 20152939. Available from http://rspb.royalsocietypublishing.org/content/royprsb/283/1830/20152939.full.pdf

15. Chapron G, Treves A. Correction to 'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore.' P Roy Soc B-Biol Sci [Internet]. 2016 [cited 2020 Jan 8]; 283(1845): 20162577. Available from http://rspb.royalsocietypublishing.org/lookup/doi/10.1098/rspb.2016.2577

16. Chapron G, Treves A. Reply to comments by Olson et al. 2017 and Stien 2017. P Roy Soc B-Biol Sci [Internet]. 2017 [cited 2020 Jan 8]; 284: 20171743. Available from http://rspb.royalsocietypublishing.org/content/royprsb/284/1867/20171743.full.pdf

17. Stenglein JL, Zhu J, Clayton MK, Van Deelen TR. Are the numbers adding up? Exploiting discrepancies among complementary population models. Ecol Evol. 2015 Jan [cited 2020 Jan 8]; 5(2): 368–553376. Available from http://dx.doi.org/10.1002/ece3.1365

18. Santiago-Ávila FJ. An interdisciplinary evaluation of large carnivore management: the grey wolf in the Western Grey Lakes [PhD dissertation]. University of Wisconsin-Madison; 2019.

19. Suutarinen J, Kojola I. One way or another: predictors of wolf poaching in a legally harvested wolf population. Anim Conserv [Internet]. 2018 [cited 2020 Jan 8]; 21(5):1–9. Available from https://doi.org/10.1111/acv.12409

20. Liberg O, Suutarinen J, Akesson M, Andren H, Wabakken P, Wikenros C, Sand H. Poaching-related disappearance rate of wolves in Sweden was positively related to population size and negatively to legal culling. Biol Conserv [Internet]. 2020 Mar [Cited 2020 May 19]; 243: 108456. Available from https://www.sciencedirect.com/science/article/abs/pii/S0006320719311498

21. Treves A, Louchouarn N, Santiago-Avila S. Response to Liberg et al. 2020: Modelling concerns confound evaluations of legal wolf-killing. Biol Conserv. In Review.

22. Chapron G, Treves A. Reply to comment by Pepin et al. 2017. P Roy Soc B-Biol Sci [Internet]. 2017 Mar [cited 2020 Jan 8]; 284(1851): 20162571. Available from http://rspb.royalsocietypublishing.org/content/royprsb/284/1851/20162571.full.pdf.

23. Refsnider RL. The role of the Endangered Species Act in Midwest wolf recovery. In: Wydeven AP, Van Deelan TR, Heske E, editors. Recovery of grey wolves in the Great Lakes Region of the United States. New York: Springer; c2009. p. 311–329

24. US Fish and Wildlife Service [Internet]. Mexican Wolf Recovery Efforts; 2019 Apr 8 [cited 2020 Jan 8]. Available from https://www.fws.gov/southwest/es/mexicanwolf/Recovery.html

25. Wydeven AP, Treves A, Brost B, Wiedenhoeft JE. Characteristics of wolf packs in Wisconsin: identification of traits influencing depredation. In: People and Predators: From Conflict to Coexistence. Washington, DC: Island Press; c2004. p. 28–50

26. Mladenoff DJ, Clayton MK, Pratt SD, Sickley TA, Wydeven AP. Change in occupied wolf habitat in the northern Great Lakes region. In: Wydeven AP, Van Deelan TR, Heske E, editors. Recovery of grey wolves in the Great Lakes Region of the United States. New York: Springer; c2009. p. 119–138

27. Ruid DB, Paul WJ, Roell BJ, Wydeven AP, Willging RC, Jurewicz RL, Lonsway DH. Wolf–human conflicts and management in Minnesota, Wisconsin, and Michigan. In: Wydeven AP, Van Deelan TR, Heske E, editors. Recovery of grey wolves in the Great Lakes Region of the United States. New York: Springer; c2009. p. 279–295

28. Thiel RP, Hall W, Heilhecker E, Wydeven AP. An Isolated Wolf Population in Central Wisconsin. In: Wydeven AP, Van Deelan TR, Heske E, editors. Recovery of grey wolves in the Great Lakes Region of the United States. New York: Springer; c2009. p. 107–117

29. Wydeven AP, Wiedenhoeft JE, Schultz RN, Thiel RP, Jurewicz RL, Kohn BE, Van Deelen TR. History, population growth, and management of wolves in Wisconsin. In: Wydeven AP, Van Deelan TR, Heske E, editors. Recovery of grey wolves in the Great Lakes Region of the United States. New York: Springer; c2009. p 87–105

30. Fitzgerald EA. The Lobo Limps on from Limbo: A History, Summary, and Outlook for Mexican Wolf Recovery in the American Southwest. Colo. Nat. Resources, Energy & Envtl. L. Rev [Internet]. 2018 [cited 2020 Jan 8]; 29(3): 223-284. Available from https://www.colorado.edu/law/sites/default/files/attached-files/fitzgerald_online_copy.pdf

31. Schmidt JH, Johnson DS, Lindberg MS, Adams LG. Estimating demographic parameters using a combination of known-fate and open N-mixture models. Ecology [Internet]. 2015 Oct [cited 2020 Jan 8]; 56: 2583–2589. Available from: https://esajournals.onlinelibrary.wiley.com/doi/10.1890/15-0385.1

32. Stenglein JL, Wydeven AP, Van Deelen TR. Compensatory mortality in a recovering top 549carnivore: wolves in Wisconsin, USA (1979–2013). Oecologia [Internet]. 2018 [cited 2020 Jan 8]; 187: 99–111. Available from https://doi.org/10.1007/s00442-018-4132-4

33. Concato J, Peduzzi P, Holford TR, Feinstein AR. 1995. Importance of Events per Independent Variable in Proportional Hazards Analysis I. Background, Goals and General Strategy. J Clinical Epidemiol [Internet]. 1995 [cited 2020 Feb 25]; 48 (12): 1495–1501. Available from https://www.ncbi.nlm.nih.gov/pubmed/8543963

34. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of Events per Independent Variable in Proportional Hazards Regression Analysis II. Accuracy and Precision of Regression Estimates. J Clinical Epidemiol [Internet]. 1995 [cited 2020 Feb 25]; 48 (12): 1503–1510. Available from https://www.ncbi.nlm.nih.gov/pubmed/8543964

35. Vittinghoff E, Mcculloch CE. Original Contribution Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. Am J Epidemiol [Internet]. 2007 [cited 2020 May 18]; 165 (6): 710–18. https://doi.org/10.1093/aje/kwk052

36. Putter H, Fiocco M, Geskus RB. 2007. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007 Oct [cited 2020 Jan 8]; 26(11): 2389–2430.

37. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd edn. Hoboken, NJ: Wiley, 2002c.

38. Heisey DM, Patterson BR. A Review of Methods to Estimate Cause-Specific Mortality in Presence of Competing Risks. J Wildlife Manage [Internet]. 2006 [cited 2020 Jan 8]; 70:1544–1555. Available from http://www.jstor.org/stable/4128086

39. Hosmer Jr DW, Lemeshow S, May S. Applied survival analysis: Regression modelling of time to event data. 2nd edn. Hoboken, New Jersey, USA: Wiley-Interscience, c2008.

40. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation [Internet]. 2016 [cited 2020 Jan 8]; 133: 601–609. Available from http://circ.ahajournals.org/content/circulationaha/133/6/601.full.pdf

41. Dignam JJ, Kocherginsky MN. Choice and Interpretation of Statistical Tests Used When Competing Risks Are Present. J Clin Oncol [Internet]. 2008 [cited 2020 Jan 8]; 26:4027–4034. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2654314/

42. Fine JP, Gray RJ. A proportional hazards model for the sub-distribution of a competing risk. J American statistical Association [Internet]. 1999 [cited 2020 Jan 8]; 94:496–509. Available from https://amstat.tandfonline.com/doi/abs/10.1080/01621459.1999.10474144#.XhZmVfyIb_U

43. Dignam JJ, Zhang Q, Kocherginsky MNM. The Use and Interpretation of Competing Risks Regression Models. Clin Cancer Res [Internet]. 2012 [cited 2020 Jan 8]; 18:2301–2308. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3328633/

44. Austin PC, Allignol A, Fine JP. The Number of Primary Events per Variable Affects Estimation of the Subdistribution Hazard Competing Risks Model. J Clin Epidemiol [Internet]. 2017 [cited 2020 Feb 25]; 83: 75–84. Available from https://doi.org/10.1016/j.jclinepi.2016.11.017

45. Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. J Clin Epidemiol [Internet]. 2013 [cited 2020 Jan 8]; 66:648–653. Available from http://dx.doi.org/10.1016/j.jclinepi.2012.09.017

46. Dienes Z. Using bayes to get the most out of non-significant results. Front Psychol [Internet]. 2014 [cited 2020 May 18]; 5: 781. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4114196/

47. Dienes Z. How do I know what my theory predicts? Advances in Methods and Practices in Psychological Science [Internet]. 2019 [cited 2020 Jun 25]; 2(4): 364-377. Available from: https://doi.org/10.1177/2515245919876960. \

48. Amrhein V, Greenland S, McShane B. Retire statistical significance. Nature [Internet]. 2019 [cited 2020 Dec 18]; 567:305–307

49. Agan S. The human dimensions and spatial ecology of poaching and implications for red wolf survival [Dissertation]. New Hampshire: Antioch New England Graduate School; 2020. 181 p.

50. U.S. Fish and Wildlife Service. Mexican wolf recovery: progress report #21. Albuquerque; U.S. Fish and Wildlife Service Southwest Region; 2018. 71 p. Progress Report No.: 21.

51. Humane Society of the U.S. et al. v Dirk Kempthorne, Secretary of the Interior, et al. (2006). 481 F. Supp. 2d 53 (DDC 2006)

52. Anderson G. Mexican wolf killings expose a dark underbelly of Western culture. Wildlife News [Internet]. 2020 Dec 21 [cited 2020 Dec 23]. Available from: http://www.thewildlifenews.com/2020/12/21/mexican-wolf-killings-expose-a-dark-underbelly-of-western-culture/

53. Platt JR. Strong inference. Science [Internet]. 1964 [cited 2020 Dec 18]; 146(3642): 347-353. Available from: https://science.sciencemag.org/content/146/3642/347

54. Ioannidis JPA. Why most published research findings are false. PLoS Med [Internet]. 2005 [cited 2020 Dec 18]; 2(8): e124. Available from: https://doi.org/10.1371/journal.pmed.0020124

55. Treves A, Krofel M, McManus J. Predator control should not be a shot in the dark. Front Ecol Environ [Internet]. 2016 [cited 202 Dec 18]; 14(7): 1-9. Available from: https://doi.org/10.002/fee.1312

56. Treves A, Krofel M, Ohrens O, van Eeden LM. Predator control needs a standard of unbiased randomized experiments with cross-over design. Front Ecol Environ [Internet]. 2019 [cited 2020 Dec 18]; 7(462). Available from: https://doi.org/10.3389/fevo.2019.00462

57. van Eeden LM, Eklund A, Miller JRB, Lopez-Bao JV, Chapron G, Cejtin MR, et al. Carnivore conservation needs ecidence-based livestock protection. PLOS Biol [Internet]. 2018 [cited 2020 Dec 18]; 16(9): e2005577. Available from: https://doi.org/10.1371/journal.pbio.2005577

58. Bruskotter JT, Vucetich JA, Slagle KM, Berardo R, Singh AS, Wilson RS. Support for the U.S. Endangered Species Act over time and space: Controversial species do not weaken public support for protective legislation. Conserv Letters [Internet]. 2018 [cited 2020 Dec 18]; e12595: 1-7. Available from: https://doi.org/10.1111/conl.12595

59. U.S. Fish and Wildlife Service. Mexican wolf population rises to at least 163 animals [Press Release]. Phoenix: U.S. Fish and Wildlife Service; 2020 Mar 18 [cited 2020 Dec 18]. https://www.fws.gov/news/ShowNews.cfm?ref=mexican-wolf-population-rises-to-at-least-163-animals-&_ID=36531

60. Frazer G. Letter to California Fish and Game Commission in response to FWS/AES/DCC/BDFS/073807. 2020 Dec 14. Washington, D.C.: U.S. Fish and Wildlife Service. 4 pp.

61. Atkins North America, Inc. Summary report of independent peer review for the U.S. Fish and Wildlife Service gray wolf delisting review. Tampa: Atkins North America, Inc; 2019. 245 p.

62. Treves, A. (2020). Memo: RIN:1018-BD60 proposed rule to remove federal protections for gray wolves nationwide. Office of Management and Budget Office of Information and regulatory Affairs, 29 sep 2020 (telecommuting: list if attendees: Maricela Constantino - DOI Teleconference; Sean Gallagher - DOI Teleconference; Bivan Patnaik - DOI Teleconference; Kristen Floom - DOI Teleconference; Austin Mudd - OIRA Teleconference; Maureen Trnka - OIRA Teleconference; Julie Hewitt - OIRA Teleconference; Matthew Oreska - OMB Teleconference; Ellen VanGelder - DOI Teleconference; Dr. Adrian Treves - University of Wisconsin - Madiso

63. Kalternborn BP, Brainerd SM. Can poaching inadvertently contribute to increased public acceptance of wolves in Scandinavia? Eur J Wildlife Res [Internet]. 2016 [cited 202 Dec 18]; 62: 179-188. Available from: https://link.springer.com/article/10.1007/s10344-016-0991-3

64. Haas TC, Ferreira SM. Combating rhino horn trafficking: the need to disrupt criminal networks. PLOS One [Internet]. 2016 [cited 2020 Dec 18]; 11(11): e0167040. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5117767/

65. Biggs D, Cook C, Redford K, Holden MH. How to overcome fierce debates about banning all trade in ivory. The Conversation [Internet]. 2018 [cited 2020 Dec 18]. Available from: https://theconversation.com/how-to-overcome-fierce-debates-about-banning-all-trade-in-ivory-95318

66. Epstein Y. 2017. Killing wolves to save them? Legal responses to ‘tolerance hunting’ in the European Union and United States. RECIEL [Internet]. 2017 [cited 2020 Dec 18]; 26(1): 19-29. Available from: http://onlinelibrary.wiley.com/wol1/doi/10.1111/reel.12188/abstract

67. Mincher BJ. Harvest as a component of Greater Yellowstone Ecosystem grizzly bear management. Wildlife Soc B [Internet]. 2002 [cited 2020 Dec 18]; 30(4): 1287-1292. Available from: https://www.jstor.org/stable/3784302

68. Loveridge AJ, Searle AW, Morindagomo F, Macdonald DW. The impact of sport-hunting on the population dynamics of an African lion population in a protected area. Biol Conserv [Internet]. 2007 [cited 2020 Dec 18]; 134(4): 548-558. Available from: https://doi.org/10.1016/j.biocon.2006.09.010

69. Ctr. for Biological Diversity v. Jewell, No. CV-15-00019-TUC-JGZ (l) (D. Ariz. Mar. 30, 2018).

Chapter 3: Politics and Policy play different roles in predicting Cryptic and reported wolf poaching in Michigan, USA

Naomi X Louchouarn[1*], Francisco Santiago-Ávila[2], Adrian Treves[1]

[1]University of Wisconsin – Madison

[2]Project Coyote and the Rewilding Institute

*Corresponding Author: louchouarn@wisc.edu

Abstract: Poaching, both reported and unreported, i.e., cryptic, has been identified as the primary cause of mortality for multiple wolf populations globally. Estimation of poaching causes and impact is difficult because poachers hide or destroy evidence. Survival analyses that have examined cause-specific mortalities and disappearances in multiple wolf populations have found that wolf disappearances and reported poaching do not correlate to the same variables such as policy changes. Here we use proportional hazard cox models to examine how hazard of reported poaching, nonhuman (i.e., natural mortalities) mortalities and disappearances of monitored wolves in Michigan (n = 501) from 1992-2022 correlated to changes in wolf management policies, partisan political changes, biological and seasonal variables. We examine how federal and state executive administrations may predict poaching hazards. In addition, we test 5 hypotheses from the literature regarding causes of poaching: 1. Liberalizing wolf killing can improve tolerance for wolves among would-be poachers and therefore reduce poaching (*killing for tolerance*), 2. Liberalizing wolf killing affects would-be poacher behavior by increasing poaching (*facilitated poaching)* 3. inconsistent wolf management produces frustration which

leads to increased poaching (*frustration in management*). 4. poaching increases after the first wolf delisting, regardless of protection status thereafter (*tipping point*) and 5. Wolf poaching increases during and following regulated public hunting periods without without subsequent return to pre-harvest levels (*ratcheting hazard)*. We find that wolf disappearances increase during democratic presidential administrations, but that certain administrations are associated with greater increases in hazard than others. We find that hazard of wolf disappearances in Michigan significantly increases after wolf protections are relaxed, but we cannot discriminate between the 5 hypotheses. We also find that reported poaching and disappearances of wolves were not correlated with the same variables. We find partial evidence for correlation between all of our policy variables and wolf disappearances. However, reported poaching appeared to correlate to opportunity and detectability of wolves during annual hunting and snow seasons, while disappearances correlated to the broad-scale federal political variables and policy. These results suggest that cryptic poaching and reported poaching are different behaviors driven by different motivators. Investigators should not omit policy and partisan political variables in analyses of wolves' survival or mortality patterns, and possibly other controversial wildlife,

Introduction

Humans and large wild carnivores compete for space and resources, often resulting in unsustainable human-induced mortality and near extinction of carnivore populations or entire species. Globally governments and organizations have made efforts to conserve and rehabilitate these carnivore populations (Chapron et al., 2014, Ripple et al., 2016). In the United States, legal protections conferred on large carnivores such as grey wolves (*Canis lupus)* and grizzly bears (*Ursus arctos*) with the Endangered Species Act (ESA) have resulted in these species returning to parts of their historic ranges. However, humans remain the leading cause of mortality for all large carnivores globally (Estes et al., 2011, Ripple et al., 2014, Woodroffe and Ginsberg, 1998). Poaching, both reported and unreported, i.e., cryptic, has been identified as the primary cause of mortality for multiple wolf populations globally and estimation of its causes and impact is difficult because poachers hide or destroy evidence (Liberg et al., 2012, Suutarinen and Kojola, 2017, Treves et al., 2017). Ensuring continued conservation successes therefore requires understanding and limiting poaching. In the U.S., the assumption by federal and state management agencies is that greater flexibility in managing conflict, usually by relaxing federal protections, will improve tolerance toward wolves among would-be poachers, reducing poaching and improving conservation outcomes (Bruskotter *et al.*, 2014; Morell, 2008; *Humane Society of U.S. v. Kempthorne*, 2006).

Actions management agencies have taken to improve tolerance for wolves and hopefully reduce poaching include reimbursement for lost domesticated animals, the use of lethal control and public hunts (Naughton-Treves et al., 2003, Treves and Naughton-Treves, 2005, Treves and Bruskotter, 2014). Because poaching is so difficult to measure, evaluations of interventions to reduce poaching require data on reported poaching and wolf disappearances. A persistent problem for agencies monitoring marked wolves is the disappearance of many individuals marked with telemetry collars (Chapron and Treves, 2016a, Treves et al., 2017, Stenglein et al., 2018, Santiago-Ávila et al., 2020, Agan et al., 2021, Louchouarn et al., 2021, Santiago-Ávila and Treves, 2022). These disappearances generally approach or exceed 50% of marked animals. Three causes of disappearance of marked animals exist: migration out of telemetry range, collar failure and cryptic poaching.

Recent survival analyses conducted separately on multiple populations of grey wolves (Santiago-Ávila et al., 2020, Santiago-Ávila and Treves, 2022, Louchouarn et al., 2021) and one population of red wolves (*Canis rufus*) (Santiago-Ávila *et al.*, 2022) showed that wolf disappearances increased during policy periods in which protections for wolves were reduced. In the case of Mexican grey wolves (*Canis lupus baileyi*), the hazard of wolves disappearing increased by 121% during two periods of reduced federal protections (Louchouarn et al., 2021). Increases in hazard and incidence of disappearance were also observed during four periods of reduced protections for coyotes and red wolves in North Carolina (Santiago-Ávila *et al.*, 2022), and six periods of reduced protections for grey wolves in Wisconsin (Santiago-Ávila, Chappell and Treves, 2020). If battery failure or migration were major causes of collared wolves' disappearances, we wouldn't expect to observe changes in disappearances aligning closely with repeated policy changes. Further, the intensively monitored small populations of Mexican gray and red wolves have virtually no records of long-distance migration (Louchouarn *et al.,* 2021; Santiago-Ávila *et al.*, 2022). In Wisconsin, while disappearances did appear to increase in correlation with policy periods, seasonal changes and non-wolf hunting seasons for big game such as deer or black bears was more strongly correlated, leading to documented increases in both reported poaching and wolf disappearances in multiple studies (Stenglein et al., 2018, Santiago-Ávila et al., 2020, Santiago-Ávila and Treves, 2022). The neighboring jurisdictions of Michigan and Wisconsin have been cooperating in wolf monitoring so human-caused deaths of collared wolves that came to either agency's attention would not go unreported, therefore migration between states was unlikely to lead to disappearances. Wolf disappearances also tended to occur earlier in the life of batteries than other endpoints such as collisions or natural causes of mortality, suggesting that battery failure was not a frequent cause of these endpoints (Santiago-Ávila and Treves, 2022). While wolf disappearances are made up of a combination of cryptic poaching, migration and collar failure, the accumulated evidence suggests that cryptic poaching makes up most wolf disappearances. Therefore, disappearances should be analyzed to examine when and possibly why cryptic poaching changes over time.

The studies summarized above document how different policy periods, big mammal hunting seasons or the durations of changes in management may correlate with cause-specific mortality and population dynamics of grey wolves. They also exemplify the importance of including disappearances into survival analyses. Such studies provide us with important insights as to the possible predictors of poaching risk and hazard among wolves. However, these variables may not be the only predictors of poaching patterns and rates. Support for gray wolf conservation and protective management also appear linked to political ideology in the US (van Eeden et al., 2021, Williams, 2022, Ditmer et al., 2022, Carlson et al., 2020, Hamilton et al., 2020).

In the US, democratic voting was strongly positively correlated with support for wolf restoration (Ditmer et al., 2022). However, another study found that jurisdictions with generally strong support for conservation also had stronger anti-conservation factions (Manfredo et al., 2017)..This phenomenon has also been studied in Scandinavia and Germany with wolves and in

England with badgers (Enticott, 2011, Von Essen et al., 2015, Clemm von Hohenberg and Hager, 2022). In Germany, wolf predation on livestock was linked with marked increases in far-right voting, and in verbal attacks by populist politicians framing wolves as a threat to rural economies (Clemm von Hohenberg and Hager, 2022). Positions towards controversial species such as wolves are linked to identity and cultural association, more so than direct experience with conflict (Bruskotter and Wilson, 2014, Dickman, 2010, Treves and Bruskotter, 2014, Clemm von Hohenberg and Hager, 2022). We therefore hypothesize that wolf poaching may occur as a form of political expression and rebellion (Muth and Bowe Jr, 1998, Enticott, 2011, Serenari and Peterson, 2016, Von Essen et al., 2015). However, to our knowledge, no survival analyses have included political variables in their cause-specific mortality hazard models.

Hypotheses and Study Objectives

Here we use survival analyses to explore changes in cause-specific mortality hazard for monitored grey wolves in Michigan from 1992-2022. We test hypotheses about factors such as policy changes, annual exposure to risk, and politics might affect collared wolf survival, and whether different mortality hazards were correlated to the same or different factors. First, we hypothesize that federal and state level political variables will correlate with poaching risk for wolves in Michigan, given the social science findings summarized above. We do not, however, predict a direction of this effect, given this variable has never been included in survival analyses of wolves. We therefore test these hypotheses against the null hypothesis that no relationship exists between hazard of cause-specific endpoints for collared wolves and changes in partisan politics such as federal presidential administrations.

Second, we replicate Santiago-Ávila and Treves (2022)'s methods to examine how annual changes in big-game hunting and snow cover may relate to wolf cause-specific mortality.

Given that Michigan and Wisconsin are similar and share a wolf population, we predict the same outcome as reported by Santiago-Ávila and Treves (2022). Namely, that snow cover periods will experience increased hazard of poaching endpoints (disappearance and reported poaching). We further predict that periods with the combination of snow cover and hunting will demonstrate the greatest increase in hazard of poaching endpoints.

Third, we provide several hypotheses regarding policy correlates of wolf survival. Federal classification of grey wolves in Michigan and its neighboring jurisdictions of the Upper Midwest USA changed from ESA listed to delisted 15 times from 1992 to 2022. Some periods within that time were characterized by a change in status every several months. The literature contains five hypotheses to explain the changes in wolf survival and poaching.

The first hypothesis, *killing for tolerance,* derives from a phenomenon first articulated during Humane Society vs. Kempthorne (2006), in which the federal government suggested, with no evidence, that allowing some legal killing of wolves would help to deter illegal killing by would-be poachers and improve tolerance (Treves, 2009). This hypothesis was tested by Chapron & Treves (Chapron and Treves, 2016a, Chapron and Treves, 2016b) who found serial slow-downs of 4 and 5% in the Wisconsin and Michigan wolf population growths respectively. These slow-downs correlated with periods of reduced wolf protections independent of the number of wolves killed legally. They inferred that cryptic poaching accounted for the slow-downs. That was disputed without presenting new data (Olson et al., 2017, Pepin et al., 2017, Stien, 2017). Their objections were rebutted with data leaving the original findings untouched (Chapron and Treves, 2017a,(Chapron and Treves, 2017). Santiago-Ávila et al., (2020) confirmed Chapron and Treves (2016a)'s inference that cryptic poaching rose during periods of liberalized wolf-killing but used survival analysis and Fine-Gray competing risk models to evaluate marked wolf cause-specific survival, including disappearances. These authors suggested that reported poaching and cryptic poaching correlated differently to liberalized wolf killing policies and also suggested the hypothesis of *facilitated poaching.* Louchouarn & Santiago-Ávila et al. (2021) also found support for the *facilitated poaching* hypothesis in their study on Mexican grey wolves summarized above. Here we attempt to replicate the above findings and re-evaluate the two hypotheses by analyzing federal and state-level changes in wolf protections.

Olson et al. (2015) suggest another alternative hypothesis, that periods of inconsistent management have a greater effect of increasing poaching hazard than simply periods of reduced wolf protection. These authors suggest that frustration born of inconsistent management of wolves in Wisconsin lead to unintended backlash and increases in illegal killing (Olson et al., 2015). We call this hypothesis the *frustration in management* hypothesis. However, Olson et al., (2015)'s analysis omitted disappearances of marked wolves and did not account for the timing of individual wolf mortality in relation to liberalized wolf-killing periods and incidents. Stenglein et al., (2015) conducted an integrated population model on Wisconsin wolves and found a 4% decrease in wolf survival that they suggest is best explained by increased illegal killing. They observed this trend between 2003-2011, a period of frequent changes in wolf management. While Stenglein et al., (2018) conducted a survival analysis of this same Wisconsin wolf population which did include disappearances as a cause-specific endpoint, and showed a marginal increase in disappearances after 2003, their analysis did not explicitly explore the connection between policies and wolf disappearances or the difference between hazard of reported poaching and hazard of disappearance. Therefore, here we test the suggestion that inconsistent management correlates to cryptic and reported poaching in Michigan.

Finally, two recent survival analyses from Minnesota found that hazard of poaching increased after a regulated public hunting and trapping season on wolves was legalized, and continued to increase after the wolf hunt ended (Barber-Meyer et al., 2021). Oliynyk et al. (2023) supported these findings in their study, using linear regression, but see Chakrabarti et al., (2022) for a smaller sample of mortality they interpret as unchanged by Minnesota's wolf-hunt. However, none of these authors include disappearances of wolves in their analyses. From the combination of these Minnesota analyses we propose a fourth and fifth hypothesis. The fourth, *ratcheting hazard*, comes directly from these analyses which suggest that creating public hunting and trapping seasons leads to a ratcheting increase in risk of poaching, during and after the actual hunting seasons (Oliynyk et al. 2023). This hypothesis specifically relates to public hunting and not to exact changes in wolf listing status under the ESA. This hypothesis suggests that protection policies do not all have the same effect, and willingness to kill wolves may increase over time. This is supported by longitudinal research in Wisconsin which found that hunters were more willing to kill wolves illegally the longer wolves had been present on the landscape (Treves et al., 2013) and escalated demands for additional methods to kill wolves once legalized (Browne-Nuñez et al., 2015).

The fifth hypothesis, *Tipping point*, is related to the ratcheting hazard hypothesis, though it is not directly explored in any previous analyses. This hypothesis suggests that the first delisting event, regardless of following protection statuses, lead to increased risk of poaching thereafter.

Here we examine the five policy related hypotheses defined above on the population of monitored wolves in Michigan by analyzing survival of individual wolves in relation to changes in policies that delisted wolves and liberalized killing of wolves. Table 1 provides predictions of how hazard of poaching endpoints (reported poaching and disappearance) must respond to policy changes to support each of the five policy hypotheses. This study would be the first to test all five of these hypotheses within the same wolf population, and the first to examine these hypotheses on Michigan's population of wolves.

Materials & Methods

Data collection and preparation

We analyzed data on radio-collared grey wolves in Michigan collected by the Michigan Department of Natural Resources (MIDNR). These data consisted of all collared wolves monitored through VHF or GPS telemetry from 14 July 1992- 12 August 2022 (n=501). Three percent of the sample (n=14) were alive at the end of the study period on 12 August 2022, and were therefore right censored in the analysis. Following Santiago-Ávila and Treves (2022), we assigned the following mutually exclusive endpoints to the remainder of the sample (97%, n=487), as estimated by MIDNR biologists (Division, 2022): collisions (vehicle strikes, n = 43, 9%), lethal control (by agency personnel, n= 16, 3%), HHS (control action by agency personnel for health and human safety concerns, or other reasons, n= 14, 3%), nonhuman (intraspecies strife, illness or some other natural cause of death, n=66, 13%), reported poached (as determined by MIDNR investigation, n = 105, 21%), lost-to-follow-up (LTF, wolves that disappeared from monitoring, n=206, 41%), and unknown (the cause of death could not be confirmed, n=37, 7%). We then built Cox proportional hazards models using a set of overlapping predictors defined below (Table 2) (Kalbfleisch and Prentice, 2011), to examine the cause-specific hazard of wolves reaching 3 endpoints of interest: LTF, reported poached and nonhuman. Using these endpoints we explore whether there is evidence to support any of the hypotheses described in the introduction, and how policies, and other overlapping predictor variables correlate with hazard of poaching (both reported and unreported through disappearances) and nonhuman mortality. We were unable to examine the legal killing endpoints (HHS and lethal control) due to the small sample sizes for these endpoints which limited our ability to use cox models.

Hypothesis Covariates

We begin by listing our primary covariates of interest to our hypotheses described in the introduction, then define additional covariates that may act as nuisance factors, for which we should control.

Policy- We examined how wolf management changes may have correlated to wolf mortality hazard rates from various causes. In this analysis we examine both federal and state protections. Federally, Grey wolves in Michigan had full protection under the ESA from the start of wolf monitoring in the state until March 31, 2003, summarized in (Chapron and Treves, 2016a). From April 2003 to 2022, wolves in Michigan have experienced more than a dozen periods of non-overlapping changes in federal policies which alternated between full protections and loosened protections which liberalized rules on wolf killing (Table 2, Table 3). These periods federally down-listed wolves from ‘endangered’ to ‘threatened’, delisted wolves from the endangered species list altogether, allowing more killing, or issued permits to allow managers and private landowners more freedom to kill wolves.

During this same period the state of Michigan also down-listed wolves from state endangered to state threatened on June 17, 2002. While wolves were classified as state threatened, lethal control was only legal when wolves were also federally down-listed. Wolves in Michigan were then delisted from the state endangered species list on April 27, 2009, at which point they were classified as a protected animal, which essentially meant that hunting and culling could occur if regulated and established through a legislative process. In October and November of 2008 two state laws, Public Aacts (PA) 318 and 290, were passed which allowed hunters and farmers to kill wolves when they deemed their hunting dogs or livestock were at risk (MIDNR, 2022). However, these laws were only applicable when wolves were federally delisted. Wolves remained classified as a state protected animal until December 12, 2012, when Governor Snyder signed PA 520, which classified wolves as a game species for the first time, beginning the process of legalizing a public wolf hunt. While there were challenges to PA 520, and a following law PA 521 (see *political* section for details), wolves remained classified as game species, a classification which was finally maintained through PA 281 (Vucetich et al., 2017). PA 281 was deemed unconstitutional and overturned on November 23, 2016, but was swiftly replaced with PA382 on December 21, 2016, again classifying wolves as a game species (MIDNR, 2022). When wolves were federally delisted, the state designation of wolves as a game species allows for regulated public hunting. However, only one public hunt, in 2013, has occurred in Michigan thus far. To better understand how cause-specific mortality may change with changing management, we used five schemes for examining policy changes, which we use as covariates in our models (Tables 2-4, Figure 1).

Policy schemes 1 (*federal_list)* and 2 (*state_list)* follow (Chapron and Treves, 2016a), which proposed the novel hypothesis that governments send a signal to (would-be) poachers when they change policy to liberalize wolf killing and examine the '*killing for tolerance'* hypothesis (Figure 1A, 1B). These authors did not distinguish state or federal policies that liberalized wolf-killing nor did they distinguish the actors doing that killing. Their analysis at a population-scale was later corroborated and refined by Louchouarn and Santiago-Ávila et al., (2021) and Santiago-Ávila et al., (2020) used survival analysis methods and found evidence that federal policies liberalizing killing of wolves correlates with a higher risk of poaching-related endpoints. They concluded that killing does not improve tolerance, because it did not reduce hazard or incidence of poaching endpoints (i.e., reported poaching and disappearance). Therefore, here we test the hypothesis of *'killing for tolerance'* using the same covariate of federal liberalized killing vs full protection periods (*federal_list* '1' = wolf ESA delisted or downlisted periods, '0' = wolves listed as endangered under the ESA) based on the legal start and end dates of federal policy changes (Table 3, Figure 1A).

Policy scheme 2 does the same, but examines state listing, with different categories based on state classification of wolves, and whether wolf killing was legal, given federal listing status (Table 4, Figure 1B).

Our other three policy variables treat changes in policy as a correlate to risk for wolf cause-specific mortality. To test how inconsistent management relates to cause-specific mortality of wolves in Michigan, our third policy variable (*inconsistent)* comprised 3 categories (Table 3, Figure 1C). Inconsistent = 0 combines all the periods during which wolves were federally listed as endangered for two or more years. Inconsistent = 1 combines all the times during which wolves were consistently federally de- or downlisted for two or more years. Inconsistent = 2 spans a period of highly inconsistent management from 2003 to 2009, during which time a single policy status never lasted more than one year. During this time, wolf protection oscillated between ESA listing and down or delisting 10 times. Using this policy scheme, we test the '*frustration in management*' hypothesis which assumes that hazard of poaching endpoints (LTF and reported poaching) both increase during inconsistent management periods (category 2), relative to consistent management periods (categories 0 and 1).

Policy scheme four (*relist*) assumes that the period from the start of the study to the first wolf down-listing in April 2003 (*relist = 0)* represents a different level of risk than the following periods (Figure 1D). Thereafter, relist = 1 is the combination of all the listed periods after April 2003, and relist = 2 is all the delisted or downlisted periods after April 2003 (Table 3). This scheme tests the hypothesis of '*Tipping point*', which assumes that hazard of endpoints increases after a downlisting, and never returns to its original level after relisting occurs.

Finally, our fifth policy scheme examines how a regulated public wolf hunting period related to cause-specific mortality of Michigan's monitored wolves. This scheme depends on state policies, however, wolves could not be hunted when they were listed as federally endangered. This scheme consists of five categories, designed to replicate those used in Minnesota analyses while also examining how listed and delisted periods correlate to hazard (Figure 1E; (Barber-Meyer et al., 2021, Chakrabarti et al., 2022, Oliynyk, 2023). Hunt = 0 combines all listed pre-hunting periods from 1992 at the start of the study period to December 11, 2012, before wolves were classified as a game species in Michigan (Table 3). Hunt = 1 combines all the delisted pre-hunting periods during this same time frame. Hunt = 2 represents the period of time from December 12, 2012 to December 19, 2014 when wolves were classified as a game species. We chose to use December 12, 2012 as the start of this period as this is the date on which Governor Snyder signed PA 520, designating wolves as a game species for the first time and beginning the process of legalizing wolf hunting *(Vucetich et al., 2017).* Hunt =3 combines all the listed periods after the public wolf hunting period. Hunt = 4 combines all the delisted periods during this same time interval. With policy scheme 5, we test the '*ratcheting hazard*' hypothesis, which posits that hazard of poaching endpoints increases during and following wolf hunting, relative to pre-hunting periods. Results that would support each policy hypothesis are listed in Table 1.

Political – political alignment and ideology may be correlated to support of conservation (Manfredo et al., 2017) and willingness to poach certain animals (Von Essen et al., 2015, Muth and Bowe Jr, 1998). To test wether poaching rates correlate with political partisanship, we explored four political partisan covariates. This assumes that the election of one or the other party's top executive sends an influential policy signal to would-be poachers who are partisans. Because the wolf has negative connotations for partisan conservatives and positive connotations for partisan liberals, we predict a change in poaching endpoints when one party is in power over another, though we do not predict a direction of this effect. To designate periods for survival analyses, we used election dates in November. Thus, we assumed that election dates are more of a public signal than are inauguration dates. We also include a binary party variable where Republicans = 0 and Democrats = 1 (Table 2).

Because politics are also local, we also included Michigan gubernatorial administrations and party as covariates (Table 2). The governors of Michigan did play significant roles in wolf management, particularly when wolves were de- or downlisted from the ESA. For example, Governor Snyder's administration, from 2010 to 2018, signed or pushed through multiple laws which affected wolf management. In December of 2012, Governor Snyder signed PA 520, a public act that classified grey wolves as a game species, a requirement needed to allow a public hunt. However, after the signing of PA 520, a petition from concerned Michiganders garnered enough signatures to challenge PA 520 through a referendum on the 2014 general election ballot. If the referendum succeeded, the public would have effectively repealed PA 520. However, in the interim between submission of petition signatures and the general election, PA 521 was enacted, which shifted the authority of 'game' species designation away from the state legislature and gave this authority to the Natural Resource Commission, a panel of 7 governor-appointed members (Vucetich et al., 2017). A second petition then garnered enough signatures to challenge PA 521 through another referendum on the general election ballot. However, again before the general election took place, a third petition was created, this time for a voter-proposed law that would again give the authority to designate game species to the natural resources commission, undermining any referendum against PA 521. This law went further by providing an appropriation to the MIDNR for Asian carp management. Any Michigan state law including a budgetary appropriation is exempt from ballot referendums (MI Const. Article II § 9). This voter-proposed law was then passed by the state legislature as PA 281 in August of 2014, and could therefore not be challenged through ballot referendum. During the November 2014 general election, both PA 520 and PA 521 were struck down by the Michigan voters (Vucetich et al., 2017).

Risk Seasons - Santiago-Ávila & Treves (2022) found that hazard of reported poaching and disappearances (LTF) changed in parallel with intra-annual changes in snow cover and hunting seasons for other large mammals in Wisconsin. Hazard of these endpoints increased during snow cover, deer or black bear hunting and hounding periods. We replicate their methods by creating 4 categorical risk seasons, according to Michigan's big game hunting, hounding and snow cover periods (Table 2). Training big game hounds is legal from July 8 to April 15 of the following year, except for a short period at the beginning of bear hunting season from Sept 2-11.

Training and hunting with big game hounds has been known to lead to conflicts with wolves, particularly in the summer months, when wolf pups are young and remain near the den site, and wolves are more territorial (Bump et al., 2013, Treves and Menefee, 2022, Wydeven et al., 2004). These conflicts could lead to reported or cryptic poaching (Treves et al., 2017). Big game hunting of black bears or white-tailed deer could also increase risk for wolves, as there are more humans with firearms present to encounter wolves. During a longitudinal study on attitudes towards wolves in Wisconsin, deer hunters reported an increased inclination to poach wolves (Treves et al., 2013, Treves et al., 2017). Michigan and Wisconsin share a similar hunting culture; therefore, we examine a similar risk season in Michigan. Firearms season in the UP of Michigan for bears or deer span Sept 7-Nov 15 and Dec 12-Jan 1 of each fall and winter throughout our study period.

Finally, snow cover may also affect wolf survival. Certainly, snow cover is correlated with winter weather, which is the period during which natural mortality is highest for grey wolves in the midwest (Stenglein et al. 2018). However, snow cover also increases detectability of wolves, which may make it easier for would-be poachers. Santiago-Ávila & Treves (2022) and Stenglein et al. (2018) both found an increase in reported poaching and disappearances when snow was present, therefore we include snow cover in our risk seasons. We defined snow 'on' and 'off' periods based on first and last day each season with >1 inch of snow reported by four weather stations in the UP: Herman, Marquette, Ironwood and Sault Ste. Marie. Using AWSSI data and the average first and last date of 1+ inch of snow cover for the UP we calculated the snow 'on' and 'off' period began Nov 9 and ended April 15 each year (Boustead, 2023).

Using the snow cover and intra-year wildlife management dates, we created 4 risk seasons: *risk_season* = 0 from April 16-July 7, and Sept 2-6 represents the period with no hounding/no hunting but no snow. *Risk_season*=1, July 8 – Sept 1 and Sept 7-Nov 8 represents the period with hounding or hunting but no snow. This variable encompasses sub-periods in which only hound training is allowed, or only big game hunting is allowed and when both are allowed. We have combined these under the assumption that hound training for big game is likely more common before the big game hunting seasons, and that 'only big game hunting' or 'only hounding' are equally risky. *Risk_season* = 2, Nov 9 – Nov 14, Dec 1 – Dec 11 and Jan 2 – April 15, represent periods with hounding/snow. There is no time in which there is snow on the ground, but hound training is not permitted, therefore, the closest we can get to examining only the effect of snow is to differentiate this period of hounding and snow cover from *risk_season* = 3, which represents the periods with hounding or hunting and snow. By comparing these four risk seasons, we can estimate if hazard of endpoints increases in correlation with these changing activities and weather conditions.

Additional Covariates

To ensure that we accounted for potential nuisance factors, we also tested the following biological and monitoring related variables. We examined three biological variables: proportion of wolf population being monitored, years since wolf recolonization and winter severity (Table 2). We collected wolf population size data from the MIDNR reports on overwinter, minimum wolf population, which was estimated annually until 2014 (except for 2012), and every other year thereafter.

We accounted for the fact that the whole wolf population could not have been monitored throughout the study period by dividing the number of wolves being monitored (i.e., being followed via telemetry) each year by the wolf population estimate for that year. The MIDNR did not follow a consistent number of wolves each year. The annual distribution of these data are normally distributed. At the beginning of the study period only one wolf was being monitored, representing 5% of the estimated wolf population at that time. In 2002, 71 wolves were being monitored representing 26% of the wolf population. However, this number dropped back to 7% in 2021, the last year for which we have an entire years-worth of telemetry data. We divided this variable into four categories to facilitate comparison across the study period (Table 2).

We collected winter severity data from the accumulated winter season severity index (AWSSI) published by the Midwest regional climate center (Boustead, 2023) for four weather stations in the UP: Herman, Marquette, Ironwood and Sault Ste. Marie. We averaged these values to estimate a winter severity for each year of the study in all of MI UP wolf range. We binned winter severity values by using the AWSSI's winter severity categories of mild to extreme (Table 2).

We test the effect of years since wolf recolonization to examine whether year contributes to model fit as a proxy for some other variable we have not otherwise measured. This variable is also relevant to an alternative hypothesis that would posit that wolf poaching increases with time exposed to wolves.

We also examined wolf collar type as a binary inter-year variable: VHF radio-telemetry (collar = 0), or a GPS (collar = 1, Table 2). While GPS collars have greater range, allowing agency staff to monitor wolf movements across larger areas, the battery requirements are greater, which may affect collar failure rates (Hebblewhite and Haydon, 2010, Tomkiewicz et al., 2010, Williams et al., 2020). Therefore, we examine the effect of collar type on endpoints, particularly LTF. While some LTF endpoints are likely to be explained by collar failures of both types of collars, including this variable should reveal if collar type significantly predicts occurrence of this endpoint.

Statistical analyses

We conducted a survival analysis using endpoint-specific Cox proportional hazard (PH) models. Survival analyses estimate the probability of observing a time interval from when a wolf is first captured and fitted with a radio collar to an endpoint, T, greater than some stated value t, $S(t) = P(T > t)$ within a specified analysis time (our variable periods). These methods allow for calculating the (endpoint-specific) hazard function, $h_k(t)$, or the instantaneous rate of occurrence of a particular endpoint *k conditional on not experiencing any endpoint until that time* (Hosmer et al., 1999, Kalbfleisch and Prentice, 2011). We used the semi-parametric Cox PH models to estimate the relative hazard of a collared wolf reaching an endpoint, given its survival to a particular date and a set of covariates. Using these models, we estimate covariate hazard ratios (HR) to model how endpoint-specific $h_k(t)$ changes as a function of survival time and model covariates. We modeled the covariate effects on the endpoint-specific hazard as $hk(t)=h0k(t)\mathrm{e}(\beta 1x1+\ldots+\beta jxj)$, where $h0k(t)$ is a not estimated baseline hazard function (i.e. semi-parametric) and βj represent the estimates of HRs for each covariate xj (HR < 1 represents a reduction in hazard and HR > 1 an increase in hazard, relative to that covariate's baseline condition, the 0 categories in Table 2).

Diagnostic Step

For each endpoint we began by analyzing each variable with univariate analyses; log-rank tests for categorical covariate and univariate cox models for continuous variables. In this way, we could identify variables likely to improve the fit of the models for that endpoint. We considered variables for inclusion in the final model if the univariate test produced a p-value at or below 0.25, following (UCLA statistical consulting Group). Of the variables which passed the univariate analyses, we tested each pair of covariates for collinearity using variance inflation factors (VIF). We then built univariate Cox models for each endpoint using the covariates relevant to our hypotheses. If any of the additional covariates were considered to be relevant to model fit and were not colinear, we controlled for this variable by including it in multivariate models or stratifying by that variable. Cox models require at least 10 events per covariate used in a multivariate model (Bradburn et al., 2003). Therefore, the maximum complexity of multivariate models built for each endpoint was dependent on the sample size of wolf mortality or disappearance events for that endpoint.

Model selection

After building each model, we tested the proportionality assumption of our Cox models by examining Schoenfeld residuals, and if necessary, we use time-varying coefficients (tvc) to control for non-proportionality. If a model meets the assumption of proportionality, Schoenfeld residuals will appear random against time. Tvcs represent covariate effects which interact with a function of monitoring time, in this case ln(t), to account for such non-proportionality.

We also used Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC) and log-likelihood as model statistics to compare and select the best models. We examine differences in these statistics for all models. Models are considered to have higher empirical support if the AIC and BIC values are lower, and a model has very little empirical support if it has an AIC value more than 10 and/or a BIC value more than 5 points higher than the model with the lowest AIC and/or BIC (Guthery, 2003, Raftery, 1995). If all values were equal, we chose the most parsimonious model. We used Cox-Snell residual plots to visually evaluate goodness-of-fit for the best models. A model with a high goodness-of-fit will create a

Cox-Snell residual plot in which the Nelson-Aelen cumulative function closely aligns with the line of Cox-Snell residuals.

Results

Here we report the results of cox model analyses for 3 endpoints of wolf mortality in Michigan from 1992 to 2022. Collinearity of variables for each endpoint are reported in supplemental Tables SM2, SM7 and SM11.

Lost-to-follow-up

Univariate analyses revealed that policy scheme 1, gubernatorial party and the proportion of the wolf population being monitored were unlikely to predict the hazard of the LTF endpoint (Table SM1).

The risk season variable was examined in a univariate model because risk season is an intra-year variable and therefore does not interact with our inter-year variables. Risk season was associated with changes in LTF hazard. While the *'hunting or hounding-no snow'* and the *'hunting or hounding-snow'* risk seasons did not show substantial change in LTF hazard, hazard increases by 60% with a positive compatible interval of +2% to +151% during the *'hounding-snow'* risk season (HR = 1.60, p = 0.04, Table SM3, Figure 2A).

GPS collars appeared to non-proportionally increase hazard of LTF by 46% relative to VHF collars (HR = 0.20, tvc = 1.46, p =0.026, Table SM3). Therefore, we stratified by collar type in all our remaining cox models that include the hypothesis variables. According to model statistics, the best model to predict LTF hazard was the stratified cox model built with the presidential administration variable (Table 5, Figure 3A). Relative to President Clinton's administration, hazard of LTF increased by 105% during Obama's administration (HR = 2.05, p = 0.023, CI95% 1.11 to 3.80). President Trump's administration was associated with a non-proportional 256% increase in hazard of LTF endpoints (HR = 0.00058, tvc = 3.56, $p < 0.001$). Therefore, hazard of LTF increased over time during Trump's administration. There was no substantial change in LTF hazard during Bush or Biden's administrations, though the study period only included a short period of Biden's administration, therefore the sample size during this period is very small. A post-hoc pairwise analysis found that there was no substantial difference in hazard between the Obama and Biden administrations, nor during the Biden and Bush administrations. This model was associated with the lowest AIC and BIC values (Table SM5), therefore it is used as the comparison model for all other models. Difference in AIC and BIC are reported for all models in Table SM5. We find that hazard of monitored wolves going LTF increased 55% during democratic presidents, relative to republican presidents with a positive compatible range of +16% to +107% (HR = 1.55, $p = 0.003$, Table SM4).

The Gubernatorial administration covariate also appeared to correlate to LTF, with substantial increases in hazard in every subsequent gubernatorial administration (Table SM4). However, this model had a difference in AIC > 10 and a difference in BIC > 5, suggesting this variable had a lack of empirical evidence and was not a good predictor of LTF hazard.

The second-best stratified cox model was built with policy scheme 2, which tested the '*killing for tolerance*' hypothesis at the state level and had a difference in AIC of 5, but a difference of BIC of 18 relative to the best model reported above (Figure 4A; Table SM5). This model showed no substantial change in LTF hazard during the '*state threatened – killing*' and '*state protected species – killing*' periods. However, LTF hazard appears to increase by 106% during the '*State threatened – no killing*' period (HR = 2.06, $p < 0.05$, CI 95% 1.18 to 3.63), 205% during the '*state protected – no killing*' period (HR = 3.05, $p<0.005$, CI 95% 1.69 to 5.49), and 191% during the '*state game – killing*' period (HR = 2.91, $p < 0.001$, CI 95% 1.52 – 5.55).

Hazard of LTF was associated with a non-proportional increase of 136% during the '*State game – no killing*' period.

All remaining stratified models built with policy schemes 3-5 had differences in AIC > 10 and BIC >5, suggesting these variables were less explanatory of changes in LTF hazard (Table SM5). A model built with policy scheme 3, which tested the '*frustration in management*' hypothesis, and stratified by collar type shows a substantial 578% increase in hazard of LTF during inconsistent management periods (HR = 6.78, $p < 0.005$, CI 95% 1.79 to 25.64), but no change in LTF hazard during consistently delisted periods relative to consistently listed period (Figure 4B, Table SM4). Policy scheme 4, which tested the '*tipping point*' hypothesis, was associated with a 103% increase in LTF hazard during listed periods after 2003 (HR = 1.03, p <0.005, CI 95% 1.24 to 3.33) and a 104% increase in LTF hazard during delisted periods (HR = 1.04, $p < 0.005$, CI95% 1.26 to 3.32; Figure 4C, Table SM4). Policy scheme 5, which tested the '*ratcheting hazard*' hypothesis was associated with a nonproportional 127% increase in hazard of LTF during '*listed post-harvest*' periods (HR = 0.0059, tvc = 2.27, $p < 0.001$), but showed no substantial change in hazard during the '*delisted pre-harvest*', '*harvest*' or '*delisted post-harvest*' periods (Figure 4D, Table SM4).

Reported Poaching

The univariate diagnostic step revealed that only policy scheme 1, policy scheme 4, presidential administration, risk season, proportion of the population monitored and time since wolf return were likely to predict hazard of reported poaching (Table SM6).

Reported poached had little correlation with any covariates explored. All the models had AIC values within 10 points of one another, though BIC values had larger differences, likely because some variables had more complexity than others and the BIC statistic gives more weight to parsimony. Reported poached was associated with a 94% increase in hazard during the '*hunting or hounding – snow season*' (HR = 1.94, p = 0.06, CI 95% 0.95 to 3.94), but no change in hazard during the '*hunting or hounding – snow*' and '*hounding – snow*' periods (Table 5, Figure 2B).

Policy Schemes 1 and 4 showed no substantial changes in reported poaching hazard between different policy periods (Table SM 8). A cox model built with presidential administration shows a 133% increase in reported poaching hazard during the Bush administration (HR = 2.33, p = 0.11, CI 95% 0.82 to 6.61), a 166% increase during the Obama administration (HR = 2.66, p = 0.073, CI 95% 0.91 to 7.75) and a 223% increase during the Trump administration (HR = 3.23, p = 0.037, CI 95% 1.08 to 9.70), but no change in hazard during the Biden administration (Figure 3B).

Nonhuman

The diagnostic univariate tests determined that only three covariates were likely to be predictive of nonhuman mortality hazard: winter severity, risk season and proportion of the population monitored. None of these variables were colinear. Only the risk season variable produced significant Cox PH model results for the nonhuman mortality endpoint. Hazard of the nonhuman endpoint increases substantially in every risk season relative to the baseline season of '*no hunting or hounding-no snow*'. The greatest increases occur during the snow seasons (Table 5, Figure 2C). Hazard of nonhuman mortality increases 628% during the '*hounding-snow*' season (p < 0.001 with CI 95% of 2.22 to 23.85) and 799% during the '*hunting or hounding-snow*' season (p < 0.001 with CI 95% 2.68 to 30.08). A post-hoc pairwise analysis of this model finds that the '*hounding-snow*' and the '*hunting or hounding – snow*' periods are not substantially different from one another (p = 0.55), but both are associated with greater hazard of the nonhuman endpoint than the '*hunting or hounding-no snow'* period ($p = 0.049$ and $p = 0.027$, respectively, Table 5).

Discussion

We tested various hypotheses regarding how policies, seasonal variation and political partisanship correlated to cause-specific mortality of grey wolves. We tested these hypotheses by conducting proportional hazard cox regression analyses on a sample of 501 monitored grey wolves in Michigan from 1992 to 2022. We found that our three endpoints of interest (disappearance, reported poached and nonhuman mortality) were predicted by different variables, and that most of our hypothesized predictors were only weakly predictive of cause-specific hazard. However, we begin by highlighting that we observed a consistent non-proportionality in the disappearance (i.e., LTF) endpoint data during some combination of years from 2014 to 2020. Any variable, either political, policy-related, or otherwise, which included a subcategory which spanned these years was non-proportional. 2014-2020 was a period when wolves were consistently listed as endangered federally, but we could not otherwise identify why such non-proportionality in the LTF endpoint would be occurring. We therefore report our findings with the caveat that this non-proportionality suggests an increase in hazard of LTF over time from 2014 to 2020.

Federal partisan political variables were most predictive of changes in hazard of wolf disappearance, particularly when controlling for collar type. We found that hazard of wolf disappearances increased by nearly 50% during the administration of Democratic presidents, but that the highest increase in hazard of wolf disappearance occurred during a Republican's, namely President Trump's administration (Table 5, Figure 2). Hazard of reported poaching also appeared to increase by the greatest margin during President Trumps administration (Figure 2). This correlation could be related to the president, or some other factor we did not measure. However, there is evidence from social science literature that suggests political affiliations correlate with attitudes towards wolves (van Eeden et al., 2021, Carlson et al., 2020, Clemm von Hohenberg and Hager, 2022, Ditmer et al., 2022, Hamilton et al., 2020, Manfredo et al., 2017) and other wildlife (Manfredo et al., 2017, Dickman, 2010, Enticott, 2011). Therefore, these results suggest support for the hypothesis that willingness to poach may be related to political partisans' attitudes (Von Essen et al., 2015, Dickman, 2010, Enticott, 2011, Muth and Bowe Jr, 1998). As politics have become progressively more polarized and linked to identity, there is some evidence that poaching has become an act of political resistance. In Scandinavia, disgruntlement among hunters who felt they had been excluded from conservation discourse lead to a resistance movement of illegal hunting (Von Essen et al., 2015). In England, farmers admitted to killing badgers out of frustration towards urban culture and resistance to conservation initiatives (Enticott, 2011). In Idaho, evidence has emerged that wolf hunting is a direct response to perceived overreach by liberal interests (Williams, 2022), while in the southeast US hunting countercultures have been linked to poaching (Serenari and Peterson, 2016). Manfredo et al., (2017) found that pro-hunting individuals in states with majority pro-conservation value orientations had less trust in natural resource management agencies than similarly minded individuals in states with less pro-conservation values. These authors inferred that this clash in values could lead to greater social conflict and cultural backlash.

U.S. Democrats are more likely to show positive attitudes towards wolves than Republicans (van Eeden et al., 2021, Ditmer et al., 2022). The perception that Democratic presidents or governors will restrict hunting or otherwise favor wolves could motivate would-be poachers, though this would not explain the large and non-proportional increase in hazard of disappearance we observe during President Trump's administration. We recommend that further research quantitatively explore the connection between federal partisan politics and wolf disappearances and reported poaching.

The relationship between LTF hazard and political variables cannot be explained by wolf migration, nor by battery failure. First, we controlled for collar type because GPS collars did correlate with higher hazard of disappearance. Therefore, any reported correlation between disappearances and political or policy variables are independent of collar type and hence some malfunctions. Migration of wolves out of telemetry range is also unlikely to explain the increase in LTF hazard because the variable of time since wolf recolonization of Michigan did not increase the hazard of disappearances of wolves (Table SM3). Wolves have become more established over time, leading to a greater density across Michigan wolf range. As wolf density reaches its maximum, more wolves may migrate out of the study region to seek open territory (Stenglein et al., 2018, Barber-Meyer et al., 2021). During our study period, the wolf population plateaued around 2010-2011, and subsequently remained relatively constant at around 650 wolves (MIDNR, 2022). If migration was the primary cause of wolf disappearance, we would expect a number of disappearances to plateau when the wolf population size did. Furthermore, we would expect to see an increase in disappearances between October and November, the '*hunting or hounding-no snow*' risk season, (Santiago-Ávila and Treves, 2022) when migration is expected to rise. Instead, we observe no change in LTF hazard during the *'hunting or hounding-no snow* 'and '*hunting or hounding-snow*' periods, and a 60% increase in LTF during the '*hounding-snow*' risk season (Figure 3A, Table SM3). In addition, Michigan and Wisconsin share a wolf population, therefore migration out of range is likely to result in Michigan wolves migrating into Wisconsin more so than the Canadian province of Ontario which is separated from the UP by a water barrier. However, Michigan and Wisconsin share information about collared individuals, therefore individuals migrating into Wisconsin are likely to be found and reported by Wisconsin DNR as occurred 26 times between 1979-2012(Treves *et al.*, 2017c). Nor is it likely that the observed increase in LTF hazard during the '*hounding – snow'* risk season is a result of collar battery failure from cold temperatures given we see no change during the other snow season (Table SM3). Once again, these results replicate those found by Santiago-Ávila and Treves, (2022), who observed a similar increase in collared wolf LTF during Wisconsin's no hunting/snow season. These authors suggest this rise in LTF hazard is a result of increased cryptic poaching as there are fewer people on the landscape to report wolf carcasses. We do not dispute that migration and collar failure make up some percentage of LTF endpoints, but it is unlikely that they approach a majority even when combined, given the changes we observe in LTF hazard in response to political variables, of which wolves have no knowledge.

We found only partial evidence to support the '*killing for tolerance*' hypothesis (policy schemes 1 and 2), and only weak evidence to support the '*frustration in management*' (policy scheme 3), '*tipping point*' (policy scheme 4) and '*ratcheting hazard*' (policy scheme 5) hypotheses. The policy schemes we designed to allow us to examine these hypotheses provided weaker models of changes in the hazard of disappearance (Tables SM5 and SM8) relative to the political partisanship variables.

Though policy scheme 1, which divided the study period into binary periods of reduced or full protections, did not provide any significant results to support the '*killing for tolerance*' hypothesis, policy scheme 2, which examined these policy changes at the state level, found some mixed evidence for this hypothesis. In particular, the periods of time when wolves were down-listed as state threatened or categorized as a state protected species (i.e., delisted) but killing was not legal had substantially higher hazard of wolf disappearances. However, hazard of disappearances was greatest when wolves were listed as game species in Michigan, regardless of whether killing was legal. Therefore, there is mixed support for the alternative *'facilitated poaching'* hypothesis also.

Though the strength of the models was weaker, we did observe some support for the *'frustration in management'*, *'tipping point'* and *'ratcheting hazard'* hypotheses. Olson et al. (2015) suggested that reported poaching of wolves may increase during consistent wolf protection and inconsistent management periods in Wisconsin due to frustration with management inconsistency. However, these authors did not examine wolf disappearances. We designed policy scenario 3 to allow us to compare hazard of poaching endpoints (disappearance and reported poaching) during consistent and inconsistent management periods. We observed a substantial increase in hazard of disappearances during inconsistent management periods, but no difference in hazard during both consistent management periods (fully protected or delisted) which suggests some support for the *'frustration in management'* hypothesis, though we observe no similar correlation with the reported poaching endpoint (Figure 4C). However, though it appears that inconsistency in management correlates with increases in the hazard of wolf disappearance, this inconsistency period also came directly following the first delisting, after which we see there is a general increase in hazard of LTF. Therefore, we cannot determine whether the increase comes from inconsistency in management or a first delisting event.

Multiple groups out of Minnesota (Barber-Meyer, Wheeldon and Mech, 2021; Oliynyk, 2023) have observed increases in wolf mortality rates during and after wolf hunting. We explored this connection but went a step further by also examining whether a first delisting was correlated with increased hazard of poaching endpoints in Michigan. We find some support for the *'tipping point'* hypothesis, as we observed an equal increase in hazard of disappearances after the first delisting in 2003, regardless of the status of wolf protections thereafter. This variable did not produce significant hazard ratios for the reported poached endpoint (Figure 4D). Policy scheme 5 was designed to directly examine the pattern observed by (Barber-Meyer, Wheeldon and Mech, 2021; Oliynyk, 2023) by analyzing how hazard of cause-specific mortality or disappearance changed before, during and after wolf hunting. A wolf hunt occurred in Michigan in 2013 but had been in planning process from 2012. This hunt was never repeated as wolves were federally relisted on the endangered species list in 2014. We find some weak support for the *'ratcheting hazard'* hypothesis as hazard of disappearance appeared to increase non-proportionally during the period when wolves were listed immediately post-hunting, but showed no change during hunting, or when wolves were delisted after hunting (Figure 4E).

If disappearances mostly comprise cryptic poaching, then we observe a general increase in cryptic poaching after relaxing wolf protections. We infer partial support for the *facilitated poaching* hypothesis, because cryptic poaching rose after delisting and increased legal control events, even if wolves were again protected afterwards.

Overall, our results do not clarify an exact policy mechanism for the observed changes in reported and cryptic poaching. Therefore, we cannot identify a single unifying theory to explain the connection between policy and poaching, beyond observing that there appears to be a connection between increased hazard of disappearances and policies which reduce protections or liberalize wolf killing. Further, the inference that none of the policy variables were predictive of reported poaching suggests that cryptic and reported poaching do not act in the same way and therefore cannot be predicted by the same variables.

Indeed, reported poaching was not clearly correlated to any variables, though it did appear to have a connection to intra-annual risk season. We observe a 94% increase in hazard for reported poaching during the big game hunting season in which snow is present (Table 5, Figure 2B). However, we observed no similar increase of disappearances during this period (Table SM3, Figure 2A). Hazard of disappearances appears to increase during the '*hounding – snow*' period when big game hunting is prohibited. A decoupling of reported poaching and cryptic poaching is supported by and duplicates a prior finding for Wisconsin (Santiago-Ávila and Treves 2022). During the '*hunting or hounding- snow*' season, increased human activity and increased ability of humans to track wolves seems to increase hazard of poaching, but may also increase the detectability of poachers or wolf carcasses (Stenglein et al., 2018, Santiago-Ávila and Treves, 2022).

Cryptic poaching which includes concealment of evidence is the more clearly criminal behavior because reported poaching may occur when a perpetrator turns themselves in for accidentally shooting a collared wolf (Treves et al., 2017). Alternatively, some perpetrators leave a wolf carcass without tampering with telemetry under the misimpression it was a coyote which can be legally shot under many circumstances in Michigan. Either way, cryptic poaching involves theft or tampering with publicly owned equipment. Therefore, it is reasonable to assume that cryptic and reported poaching are different behaviors motivated differently.

For comparison we also examine changes in hazard of nonhuman mortality and found this hazard more than quadrupled in all risk seasons relative to baseline, with the highest hazards occurring during snow seasons. This is likely a result of natural annual mortality trends, in which wolf mortality increases during winter (Stenglein et al., 2018). Other wolf management and political variables did not have a significant impact on hazard of nonhuman mortality. The nonhuman endpoint was not predicted by any political or policy variables, which is to be expected given non-human mortality should not respond to non-biotic variables.

Conclusion

The best models for hazard of each of the examined endpoints (deaths or disappearances of collared wolves) are related to different variables. This approach has allowed us to examine how different mortality causes likely result from combinations of different variables, and to highlight why each mortality endpoint should be examined individually and considered as such in policy decisions (Santiago-Ávila and Treves, 2022). Past survival analyses on wolves have analyzed cause-specific mortality hazard using some of the same variables to produce cumulative incidence functions (Santiago-Ávila et al., 2020, Santiago-Ávila and Treves, 2022, Agan et al., 2021, Louchouarn et al., 2021). However, there is an equal value to examining each endpoint's hazards individually to determine if different variables explain changes in hazard.

US presidential administrations, but not Michigan's gubernatorial administrations, substantially correlated to the hazard of radio-collared wolf disappearance. Further, we report Democratic presidential administrations have higher hazard of poaching, but that particular presidential administrations correlated to the greatest increases in hazard. Despite some mixed evidence for each of our policy-related hypotheses, the general pattern appears to suggest increases in hazard of wolf disappearances, but not reported poaching, when protections for wolves were relaxed.

For Michigan wolves, reported poaching and disappearances of wolves were not correlated with the same variables. Reported poaching appears to relate to opportunity and detectability of wolves, while disappearances were related to broad-scale political variables. We conclude firmly that cryptic poaching and reported poaching may respond differently, and we suggest that political partisanship be examined as a possible cause of cryptic poaching.

Tables & Figures

Table 1. Relationship between our hypotheses, analyses and interpretation of outcomes (including contingent interpretation and synthesis of model results). HRpoa refers to the hazard ratio of the reported poaching endpoint, while HRltf refers to the hazard ratio of the LTF endpoint.

Question	Hypotheses	Sampling plan	Interpretation given different outcomes
Do hazard rates of death by poaching or disappearance of wild, collared adult grey wolves in Michigan change after policies change from strict protection to liberalized killing and back again.	'Killing for tolerance' predicts the hazard declines for the endpoint 'reported poached' (poa) or the endpoint LTF when the IV of federal and/or state policy period liberalizes wolf-killing.	All collared gray wolves in Michigan with the endpoints of interest collected from MIDNR (n=497) from 1992 - 2022 A diagnostic test is run on the samples with three endpoints (reported poached, LTF and nonhuman) before proceeding to the analysis plan (See Diagnostic Step).	For federal_list variable: HRpoa and HRltf are <1 for OR (HRpoa has to be <1 and greater in magnitude than any increase in HRltf OR HRltf has to be <1 and greater in magnitude than any increase in HRpoa) AND/OR for state_list variable: HRpoa and HRltf are <1 for subcategories with killing OR (HRpoa has to be <1 and greater in magnitude than any increase in HRltf OR HRltf has to be <1 and greater in magnitude than any increase in HRpoa)

	‘Facilitated poaching’ predicts the hazard increases for the endpoint ‘reported poached’ (poa) or the endpoint LTF when the IV of federal and/or state policy period liberalizes wolf-killing.		For the federal list variable: HRpoa and HRltf are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) AND/OR for state_list variable: HRpoa and HRltf are >1 for subcategories with killing OR (HRpoa has to be >1 and greater in magnitude than any decrease in HRltf OR HRltf has to be >1 and greater in magnitude than any decrease in HRpoa)
Do hazard rates of death by poaching or disappearance (DV) of wild, collared adult grey wolves in Michigan change during policy periods (IV) of consistent protection relative to periods when wolves are consistently delisted, and periods when wolf management is inconsistent	‘Frustration in management’ predicts the hazard increases for the endpoint ‘reported poached’ (poa) or the endpoint LTF when the IV of policy period is inconsistent (changes more frequently than every two years), while periods with consistent management (wolves either listed or delisted) have decreased hazard for the endpoints of reported poaching and/or LTF		HRpoa and HRltf for inconsistent management period are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) AND HRltf and/or HRpoa are = 1 for periods when wolves are consistently (2+ years in a row) delisted This would mean that periods when wolves are consistently delisted have equal hazard for poaching endpoints as periods when wolves are consistently listed

	‘Facilitate poaching’ predicts in this case the hazard increases for ‘reported poached’ (poa) or LTF when wolves are consistently delisted		HRpoa and HRltf for consistently delisted period are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) This would mean that periods when wolves are consistently delisted have greater hazard for poaching endpoints than periods when wolves are consistently listed. Inconsistent protections may have increased or decreased hazard for these endpoints, and therefore support for this hypothesis is not exclusive of support for the “frustration in management hypothesis”.
Do hazard rates of death by poaching or disappearance (DV) of wild, collared adult grey wolves in Michigan change during policy periods (IV) after Michigan’s wolves were first delisted in 2003.	‘Tipping point’ predicts the hazard increases for ‘reported poached’ (poa) or LTF during periods when wolves are delisted after 2003, and during periods when wolves are listed after 2003.		HRpoa and HRltf for listed AND delisted periods after 2003 are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) This would mean that poaching hazard increases after a first delisting, regardless of policy changes thereafter *Facilitated poaching* is partially supported in this case because a first delisting leads to increased hazard of poaching endpoints

	‘facilitated poaching’ in this case predicts hazard increases more for ‘reported poached’ or LTF during periods when wolves are delisted after 2003		HRpoa and HRltf for delisted periods after 2003 are >1 AND HRpoa and HRltf for delisted periods after 2003 are > HRpoa and HRltf for listed periods after 2003
Do hazard rates of death by poaching or disappearance (DV) of wild, collared adult grey wolves in Michigan change during periods (IV) with legal wolf harvest (2013-2014), and post-harvest (2014-2022)	‘Ratcheting hazard’ predicts the hazard increases for ‘reported poached’ (poa) or LTF during periods when public wolf harvest is legalized, and periods after harvest.		HRpoa and HRltf harvest and both post-harvest periods are >1 OR (HRpoa has to be >1 and greater than any decrease in HRltf OR HRltf has to be >1 and greater than any decrease in HRpoa) Partial support for this hypothesis if HRpoa and/or HRltf > 1 for all harvest and post-harvest periods but post-harvest period HR is not greater than harvest HR. Full support for this hypothesis if HRpoa and/or HRltf during post-harvest periods > HRpoa and/or HRltf during harvest periods > 1 *Facilitated poaching* is partially supported in this case because a public harvest leads to increased hazard of poaching endpoints Greater support for *facilitated poaching* hypothesis if hazard is greater during delisted periods relative to listed periods throughout.

Table 2. Inter- and Intra-year categorical and continuous variables tested using univariate analyses for use in cox proportional hazard models.

Category	**Citation/rationale**
Political Variables	
Presidential Administrations	
Clinton administration (Start of study to Nov 6, 2000)	that political ideology is correlated to support for wildlife conservation(Clemm von Hohenberg and Hager, 2022; Ditmer *et al.*, 2022; Hamilton *et al.*, 2020; van Eeden *et al.*, 2021), and in some cases stated willingness to poach (Enticott, 2011; Von Essen *et al.*, 2015)Williams, 2022 #152]) We assumed that election dates are a stronger signal than inauguration dates.
Bush Administration (Nov 7, 2000 to Nov 3, 2008)	
Obama Administration (Nov 4, 2008 to Nov 7, 2016)	
Trump Administration (Nov 8, 2016 to Nov 2, 2020)	
Biden Administration (Nov 3, 2020 to end of study)	
Presidential Party	
Republican (party = 0)	
Democrat (party = 1)	
Michigan Gubernatorial Administration	
Engler (start of study to Nov 4, 2002)	
Granholm (Nov 5, 2002 to Nov 1, 2010)	
Snyder (Nov 2, 2010 to Nov 5, 2018)	
Whitmer (Nov 6, 2018 to end of study)	
Governor's Party	
Republican (party = 0)	
Democrat (party = 1)	
Policy Variables	
Policy Scheme 1 (federal_list)	
Periods of reduced protection (federal_list = 0)	

Periods of full protection (federal_list = 1)	Following (Chapron and Treves, 2016a; Santiago-Ávila, Chappell and Treves, 2020), period duration based on policy start and end dates. See table 2 for exact dates)
Policy scheme 2 (state_list)	
State endangered (state_list = 0)	Policy durations based on dates public acts were signed by governors. No killing periods are based on when wolves were classified as federally endangered which made the state laws liberalizing killing inactive.
State threatened – no killing (state_list = 1)	
State threatened – killing (state_list = 2)	
State protected species – no killing (state_list = 3)	
State protected species – killing (state_list = 4)	
State game – no killing (state_list = 5)	
State game – killing (state_list = 6)	
Policy Scheme 3 (relist)	
Period wolves listed before 2003 (relist = 0)	Separating periods in which wolves were ESA listed into pre- and post-2003 to examine whether hazard changes based on past protection status
Periods wolf listed after 2003 (relist = 1)	
Periods wolves de- or downlisted (relist = 2)	
Policy Scheme 4 (inconsistent)	
Periods of consistent wolf protection (inconsistent = 0)	Periods in which a single federal wolf protection status were maintained for more than two years were considered 'consistent' management. 2022 onward is considered consistent because the study period ends but the protection status for wolves has not changed.
Periods of inconsistent management (inconsistent = 1)	
Periods of consistent reduced protection (inconsistent = 2)	
Policy Scheme 5 (harvest)	
Listed periods before any public harvest legalized (harvest = 0)	The policy signal for the beginning of the wolf harvest period is when wolf harvesting is legalized by designating Michigan's wolves as a game species.
Delisted periods before any public harvest legalized (harvest = 1)	

Legal public harvest period (harvest = 2)	
Listed periods after public harvest no longer legal (harvest = 3)	
Delisted periods after public harvest no longer legal (harvest = 4)	
Intra-year Variable – Risk Season	
No hounding, no big game hunting, no snow (risk_season = 0)	Hound training or hunting with hounds is legal in Michigan from July 8 to April 15 of the following year, with the exception of Sept 2-11. Big game hunting with firearms is legal for bears and deer from Sept 7-Nov 15 and Dec 12-Jan 1. The average first date of Snow in the UP is Nov 9, and the average last date of snow is April 15 (AWSSI).
Hounding or big game hunting, no snow (risk_season = 1)	
Hounding and snow (risk_season = 2)	
Hounding, big game hunting and snow (risk_season = 3)	
Additional variables	
Winter Severity	
Mild (0 – 1859)	Collected from the AWSSI categories named for quantiles of all past data for each of the four weather stations in Michigan's upper peninsula. (AWSSI)
Moderate (1860 – 2103)	
Average (2104 – 2103)	
Severe (2104-2505)	
Extreme (2506-3206)	
Proportion of wolf population monitored	Number of wolves monitored per year divided by minimum annual wolf population count
<5%	
5-10%	
10-15%	
>15%	
Time since wolf return to michigan	

Collar Type	
VHS collar (collar = 0)	While GPS collars have greater range, allowing agency staff to monitor wolf movements across larger areas, they have larger battery requirements which may lead to greater failure rates (Hebblewhite and Haydon, 2010; Tomkiewicz *et al.*, 2010; Williams *et al.*, 2020). There are more VHS collars at the beginning of our study, and more GPS collars toward the end of our study period.
GPS collar (collar = 1)	

Table 3. Start and end dates for each period under each of four policy schemes used as covariates in cox PH models. Policy scheme 1 follows (Santiago-Ávila et al., 2020)**,** and is based on the start and end dates of federal policies which listed or delisted wolves from the endangered species list. Policy Scheme 3 we combine all periods of wolf delisting longer than two years into a single category. We do the same with all listing periods longer than two years. And lastly, we combine all periods of listing or delisting shorter than 1 year into a single category. Policy scheme 4 separates wolf listing periods from before and after 2003 into separate categories, while maintaining delisting periods as its own category. Finally, in policy scheme 5 we separate periods when wolves were listed and delisted before a wolf harvest occurred, and after a wolf harvest.

Dates	**Policy Scheme 1**	**Policy Scheme 3**	**Policy Scheme 4**	**Policy Scheme 5**
7/14/1992 to 3/31/2003	Listed (lib_kill = 0)	Consistently Listed (inconsistent = 0)	Listed (relist = 0)	Pre-harvest listed (harvest = 0)
4/1/2003 to 1/30/2005	Delisted (lib_kill = 1)	inconsistent (Inconsistent = 1)	Delisted (relist = 2)	Pre-harvest delisted (harvest = 1)
1/31/2005 to 3/31/2005	Listed (lib_kill = 0)		Relisted (relist = 1)	Pre-harvest listed (harvest = 0)
4/1/2005 to 9/13/2005	Delisted (lib_kill = 1)		Delisted (relist = 2)	Pre-harvest delisted (harvest = 1)
9/14/2005 to 5/5/2006	Listed (lib_kill = 0)		Relisted (relist = 1)	Pre-harvest listed (harvest = 0)
5/6/2006 to 7/31/2006	Delisted (lib_kill = 1)		Delisted (relist = 2)	Pre-harvest delisted (harvest = 1)
8/1/2006 to 3/11/2007	Listed (lib_kill = 0)		Relisted (relist = 1)	Pre-harvest listed (harvest = 0)
3/12/2007 to 9/28/2008	Delisted (lib_kill = 1)		Delisted (relist = 2)	Pre-harvest delisted (harvest = 1)
9/29/2008 to 5/3/2009	Listed (lib_kill = 0)		Relisted (relist = 1)	Pre-harvest listed (harvest = 0)
5/4/2009 to 6/30/2009	Delisted (lib_kill = 1)		Delisted (relist = 2)	Pre-harvest delisted (harvest = 1)
7/1/2009 to 1/26/2012	Listed (lib_kill = 0)	Consistently Listed	Relisted (relist = 1)	Pre-harvest listed (harvest = 0)

		(inconsistent = 0)		
1/27/2012 to 12/18/2014	Delisted (lib_kill = 1)	Consistently Delisted (inconsistent = 2)	Delisted (relist = 2)	Harvest* (harvest = 2)
12/19/2014 to 11/2/2020	Listed (lib_kill = 0)	Consistently Listed (inconsistent = 0)	Relisted (relist = 1)	Post-harvest listed (harvest = 3)
11/3/2020 to 2/9/2022	Delisted (lib_kill = 1)	Consistently Delisted (inconsistent = 2)	Delisted (relist = 2)	Post-harvest delisted (harvest = 4)
2/10/2022 to 9/14/2022	Listed (lib_kill = 0)	Consistently Listed (inconsistent = 0)	Relisted (relist = 1)	Post-harvest listed (harvest = 3)

*Harvest period begins when wolves are classified as a game species and plans for harvest are legalized on 12/12/2012

Table 4. Start and end dates for each period under policy scheme 2 used as covariates in cox PH models. Policy scheme 2 is based on the start and end dates of state policies which classified wolves as endangered, threatened, protected animal (ie delisted) and game species. This policy scheme also specifies times in which wolves could or could not be killed because of state laws or federal listing as endangered.

Dates	**Policy Scheme 2**
7/14/1992 to 6/16/2002	State Endangered (state_list = 0)
6/17/2002 to 4/30/2003	State threatened – no killing (state_list = 1)
5/1/2003 to 1/31/2005	State threatened – killing (state_list = 2)
2/1/2005 to 5/31/2005	State threatened – no killing (state_list = 1)
6/1/2005 to 9/13/2005	State threatened – killing (state_list = 2)
9/14/2005 to 5/5/2006	State threatened – no killing (state_list = 1)
5/6/2006 to 7/31/2006	State threatened – killing (state_list = 2)
8/1/2006 to 4/26/2009	State threatened – no killing (state_list = 1)
4/27/2009 to 5/3/2009	State protected species – no killing (state_list = 3)
5/4/2009 to 6/30/2009	State protected species – killing (state_list = 4)
7/1/2009 to 1/26/2012	State protected species – no killing (state_list = 3)
1/27/2012 to 12/11/2012	State protected species – killing (state_list = 4)
12/12/2012 to 12/18/2014	State game – killing (state_list = 6)
12/19/2014 to 11/22/2016	State game – no killing (state_list = 5)
11/23/2016 to 12/20/2016	State protected species – no killing (state_list = 3)
12/21/2016 to 11/2/2020	State game – no killing (state_list = 5)
11/3/2020 to 2/9/2022	State game – killing (state_list = 6)
2/10/2022 to 9/14/2022	State game – no killing (state_list = 5)

Table 5. The best cox PH models, chosen based on adherence to proportionality assumptions, AIC, BIC, log-likelihood and parsimony, for each of the 3 cause-specific mortality endpoints. Hazard ratios, standard error (in parentheses) and 95% confidence intervals (in brackets) are reported by category for each covariate and time-varying covariate. Letters are applied as superscript to represent significant difference between categories according to post-hoc pairwise comparison.

Lost-to-follow-up (LTF) Endpoint			
Best Model: presidential administration + tvc(Trump) (stratified by collar type)			
AIC = 2128, BIC = 2164, LL= -1059			
Presidential administration (baseline: Clinton)			
Bush	Obama	Trump	Biden
1.17[a] (0.37) [0.63-2.17]	2.05[b]*(0.65) [1.11-3.80]	0.00058** (0.0011) [0.00025-0.14]	2.10[a,b] (0.52) [0.55-2.83]
Tvc (Trump)	2.27 ***(0.93) [1.42-5.94]		
Reported Poached Endpoint			
Best Model: risk season			
AIC = 1098, BIC = 1120, LL = -546			
Risk Season			
Hunting or hounding-no snow	Hounding-snow	Hunting or hounding-snow	
0.65 (0.23) [0.32-1.30]	1.21[a] (0.41) [0.62-2.35]	1.94[a] (0.41) [0.95-3.94]	
Nonhuman Endpoint			

Best Model: risk season		
AIC = 659, BIC = 680, LL = -326		
Risk Season		
Hunting or hounding-no snow	Hounding-snow	Hunting or hounding-snow
3.66[a]* (2.34) [1.04-12.86]	7.28[b]***(4.41) [2.22-23.85]	8.99[b]*** (5.54) [2.68-30.08]

*$p \leq 0.05$ ** $p \leq 0.005$ *** $p \leq 0.001$

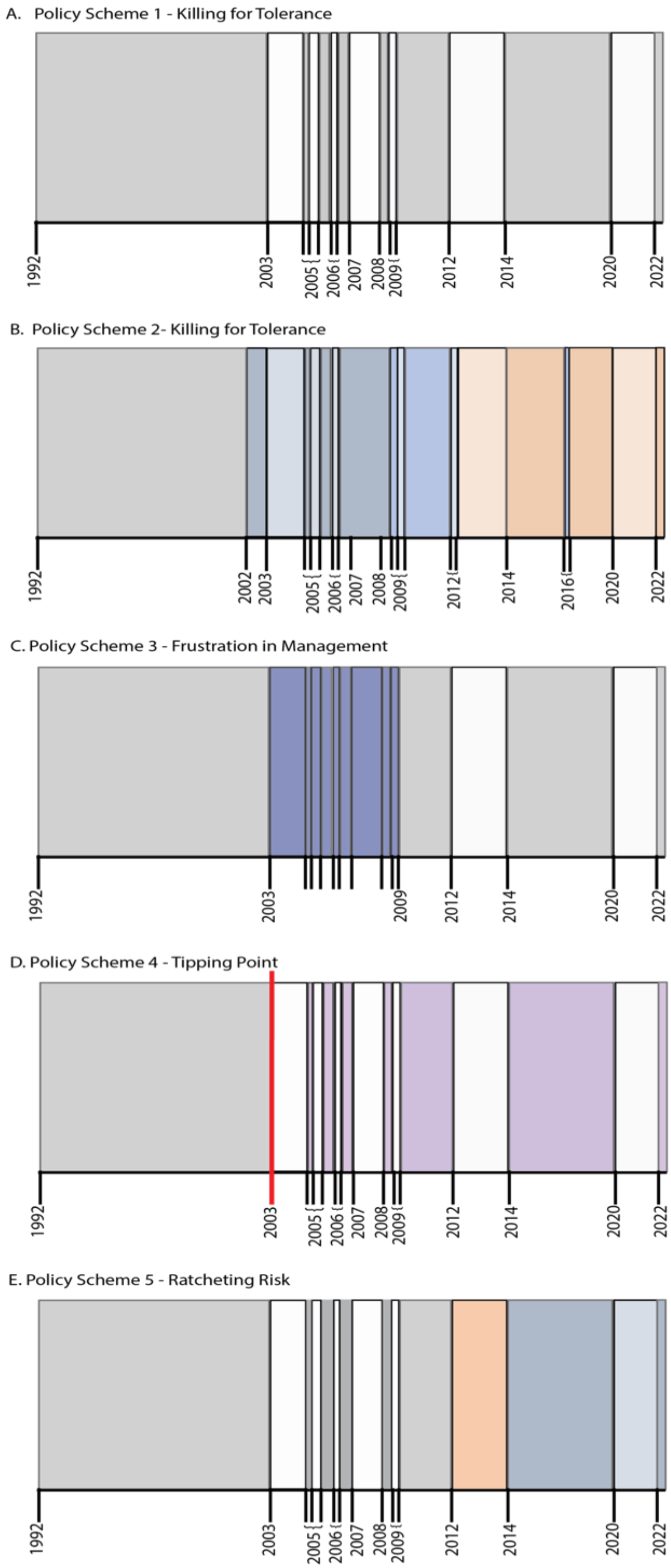


Figure 1. A visual representation of the study timeline and each policy scheme based on date of policy change. Ticks along the timeline represent policy changes, but only years are shown. For exact dates see Tables 3 and 4. A. Policy scheme 1 compares hazard of cause-specific endpoints in federally listed (grey) vs delisted (white) periods. B. Policy scheme 2 compares hazard of cause-specific endpoints in state listed (grey), state threatened (grey-blues), state protected species (light blues), state game (oranges). Darker shades represent periods with no killing, and lighter shades are periods with killing. C. compares hazard of cause-specific endpoints in periods with consistent protections for wolves (grey), periods with reduced protections (white) and periods when policies were inconsistent between 2003 and 2009 (dark blue). D. Policy scheme 4 compares hazard of cause-specific endpoints during the period of full wolf protection (grey) from 1992 to 2003 (red line), to periods of reduced protection (white) and periods of full protection after 2003 (purple). E. Policy Scheme 5 examines hazard of cause-specific endpoints during listed and delisted periods before a wolf harvest (in orange), and after a wolf harvest.

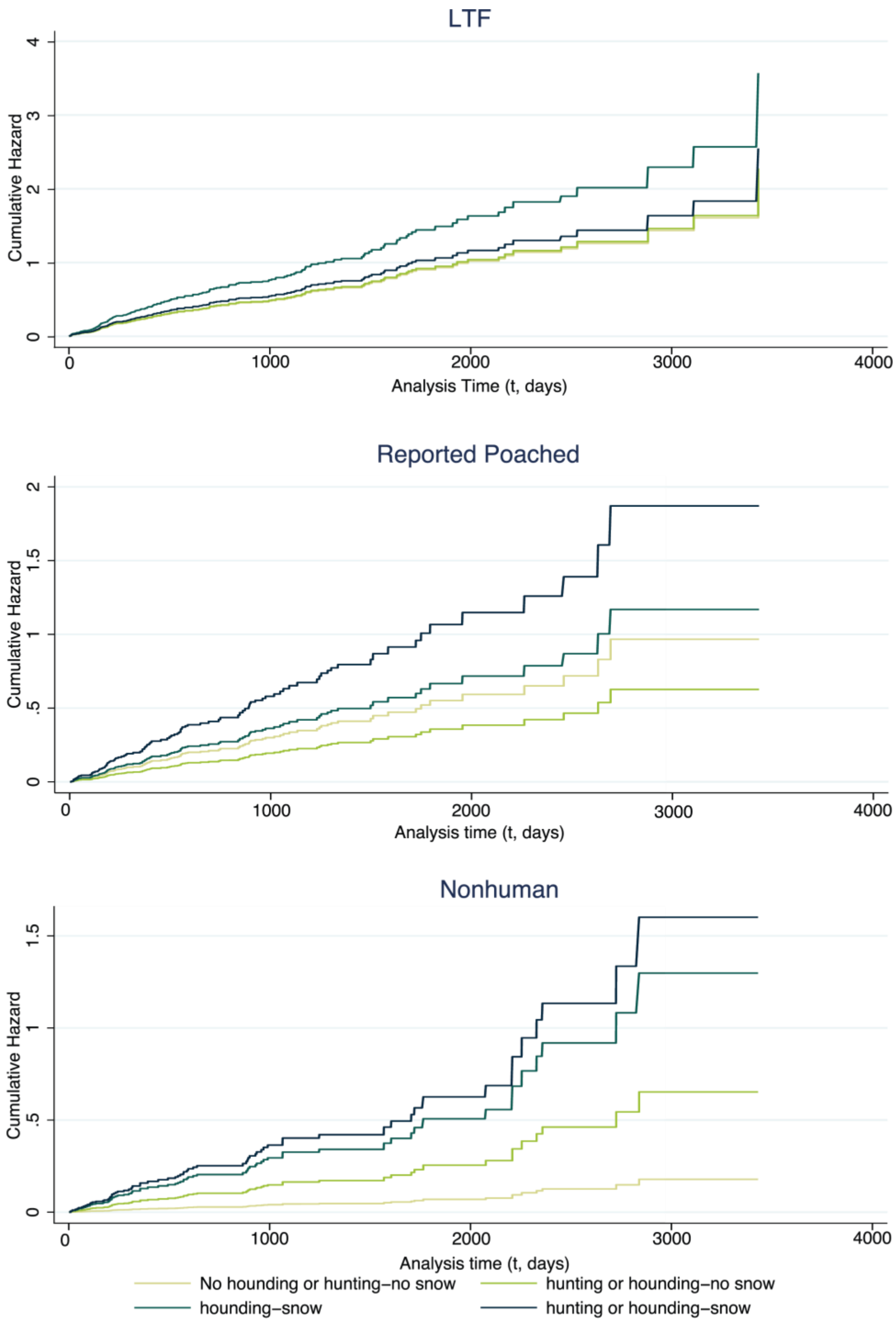


Figure 2. Endpoint-specific cumulative hazard for risk season derived from a univariate Cox model for the LTF (n=206), reported poached (n=105) and nonhuman (n=66) endpoints.

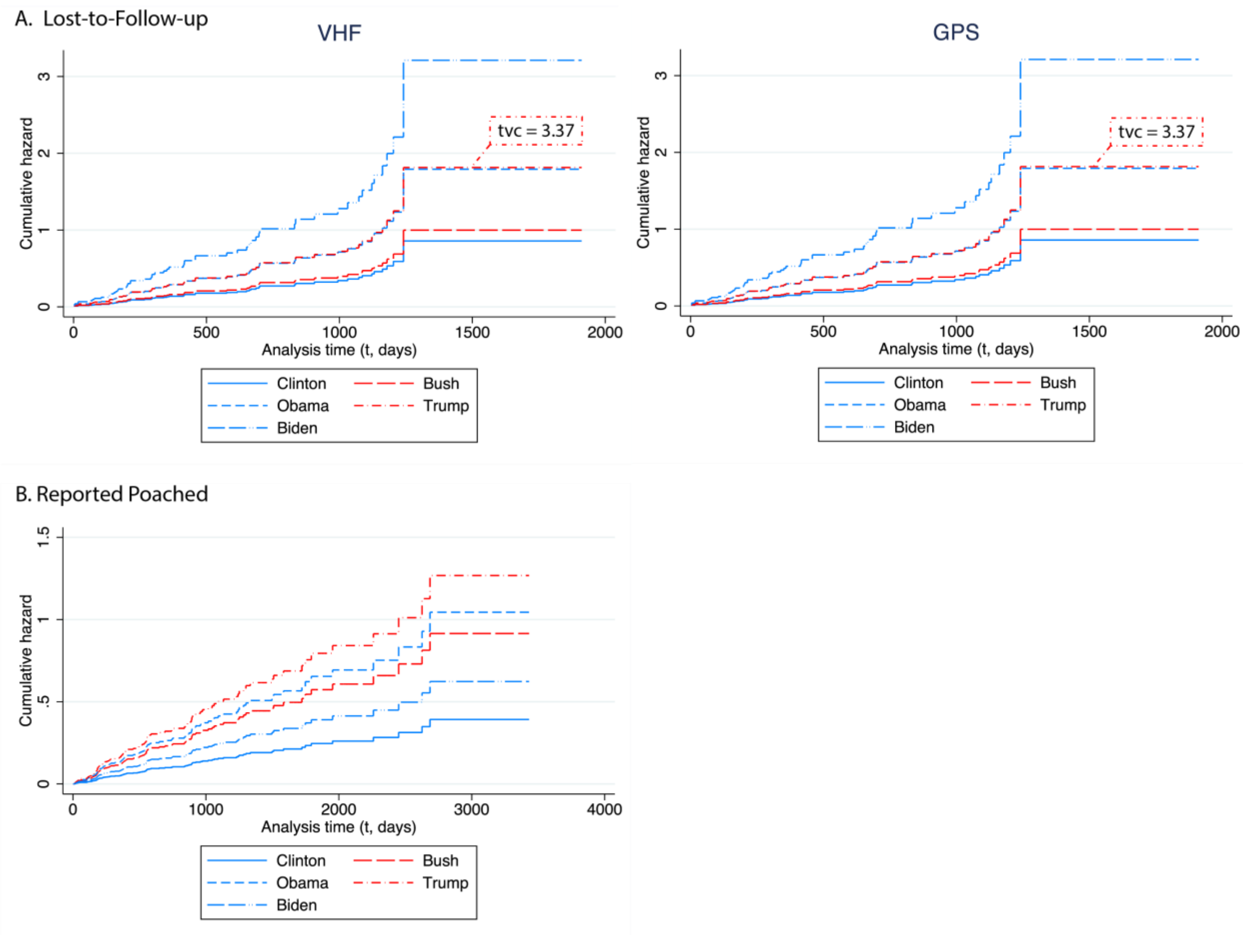


Figure 3. Endpoint-specific cumulative hazard for presidential administration derived from univariate Cox models for the LTF (n=206) and reported poached (n=105) endpoints. A. The LTF model is stratified by collar type. Cox model graphs only represent a single point in time and therefore cannot visually represent tvc. The tvc value is added for the Trump administration. B. Collar type did not impact the hazard of the reported poached endpoint, therefore the reported poached model is not stratified.

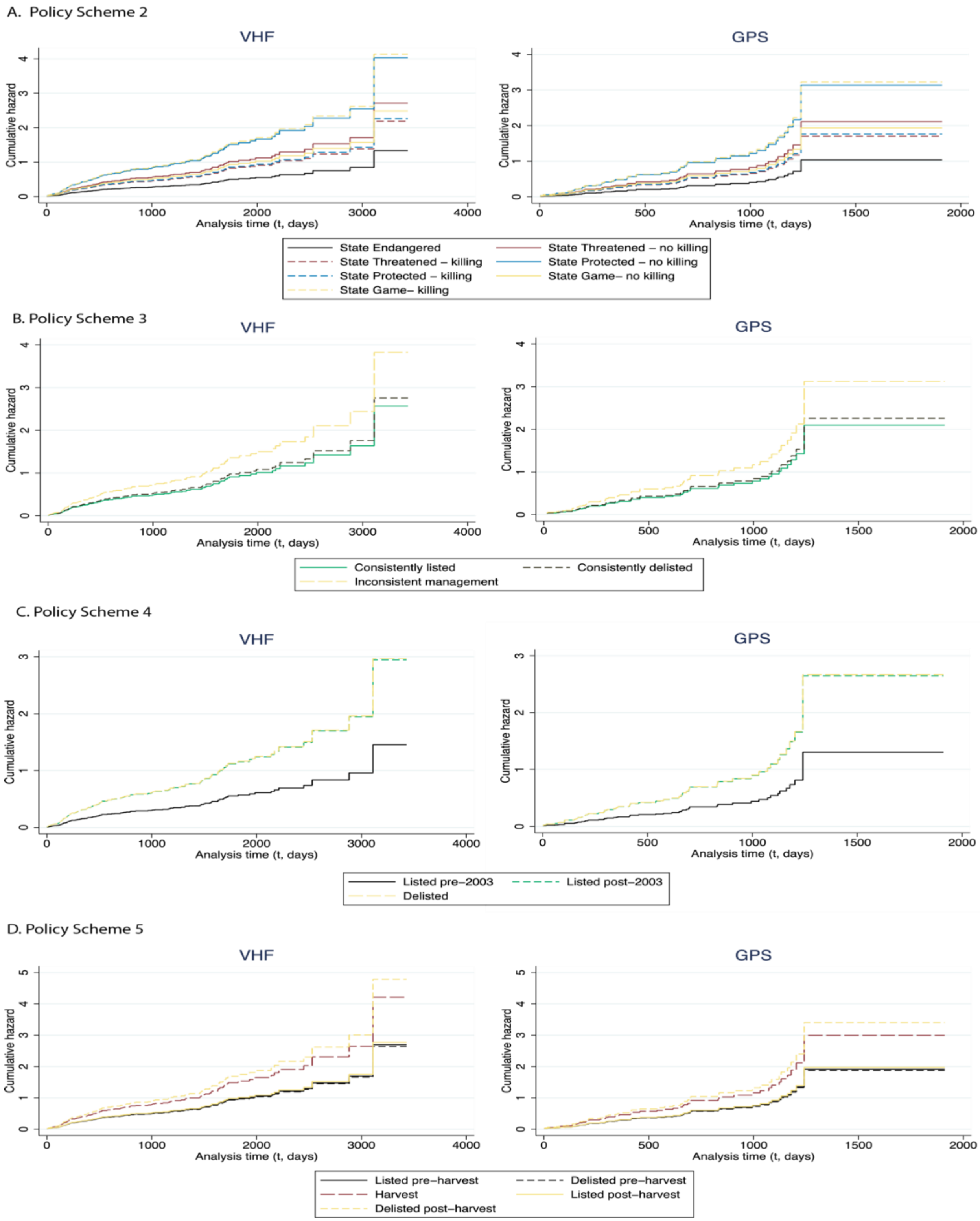


Figure 4. Endpoint-specific cumulative hazard for policy schemes 2-5 derived from univariate Cox models for the LTF (n=206) endpoint. Models are stratified by radio-collar type to control for the 40% increase in hazard of LTF with GPS collars.

References

2006. Humane Society of U.S. v. Kempthorne. United States District Court, D. Columbia.

AGAN, S. W., TREVES, A. & WILLEY, L. L. 2021. Estimating poaching risk for the critically endangered wild red wolf (Canis rufus). *Plos one,* 16**,** e0244261.

BARBER-MEYER, S., WHEELDON, T. & MECH, L. D. 2021. The importance of wilderness to wolf (Canis lupus) survival and cause-specific mortality over 50 years. *Biological Conservation,* 258**,** 109145.

BOUSTEAD, B., HILBERG, S 2023. Accumulated Winter Season Severity Index. *In:* PURDUE UNIVERSITY, N. (ed.). Midwest Regional Climate Center: Midwest Regional Climate Center.

BRADBURN, M. J., CLARK, T. G., LOVE, S. B. & ALTMAN, D. G. 2003. Survival analysis Part III: multivariate data analysis–choosing a model and assessing its adequacy and fit. *British journal of cancer,* 89**,** 605-611.

BROWNE-NUÑEZ, C., TREVES, A., MACFARLAND, D., VOYLES, Z. & TURNG, C. 2015. Tolerance of wolves in Wisconsin: a mixed-methods examination of policy effects on attitudes and behavioral inclinations. *Biological Conservation,* 189**,** 59-71.

BRUSKOTTER, J. T. & WILSON, R. S. 2014. Determining where the wild things will be: using psychological theory to find tolerance for large carnivores. *Conservation Letters,* 7**,** 158-165.

BUMP, J. K., MURAWSKI, C. M., KARTANO, L. M., BEYER JR, D. E. & ROELL, B. J. 2013. Bear-baiting may exacerbate wolf-hunting dog conflict. *PLoS One,* 8**,** e61708.

CAPPELLETTI, J. A. W., COREY. 2022. Co-leader of Witmer kidnapping plot gets 16 years in prison. *Associated Press,* December 27, 2022.

CARLSON, S. C., DIETSCH, A. M., SLAGLE, K. M. & BRUSKOTTER, J. T. 2020. The VIPs of wolf conservation: How values, identity, and place shape attitudes toward wolves in the United States. *Frontiers in Ecology and Evolution,* 8**,** 6.

CHAKRABARTI, S., O'NEIL, S. T., ERB, J., HUMPAL, C. & BUMP, J. K. 2022. Recent Trends in Survival and Mortality of Wolves in Minnesota, United States. *Frontiers in Ecology and Evolution,* 10**,** 542.

CHAPRON, G., KACZENSKY, P., LINNELL, J. D., VON ARX, M., HUBER, D., ANDRÉN, H., LÓPEZ-BAO, J. V., ADAMEC, M., ÁLVARES, F. & ANDERS, O. 2014. Recovery of large carnivores in Europe's modern human-dominated landscapes. *science,* 346**,** 1517-1519.

CHAPRON, G. & TREVES, A. 2016a. Blood does not buy goodwill: allowing culling increases poaching of a large carnivore. *Proceedings of the Royal Society B: Biological Sciences,* 283**,** 20152939.

CHAPRON, G. & TREVES, A. 2016b. Correction to 'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore'. *Proceedings of the Royal Society B: Biological Sciences,* 283**,** 20162577.

CHAPRON, G. & TREVES, A. 2017. Reply to comments by Olson et al. 2017 and Stien 2017. *Proceedings of the Royal Society B: Biological Sciences,* 284**,** 20171743.

CLEMM VON HOHENBERG, B. & HAGER, A. 2022. Wolf attacks predict far-right voting. *Proceedings of the National Academy of Sciences,* 119**,** e2202224119.

DEPETRIS-CHAUVIN, E. 2015. Fear of Obama: An empirical study of the demand for guns and the US 2008 presidential election. *Journal of Public Economics,* 130**,** 66-79.

DICKMAN, A. J. 2010. Complexities of conflict: the importance of considering social factors for effectively resolving human–wildlife conflict. *Animal conservation,* 13**,** 458-466.

DITMER, M. A., NIEMIEC, R. M., WITTEMYER, G. & CROOKS, K. R. 2022. Socio-ecological drivers of public conservation voting: Restoring gray wolves to C olorado, USA. *Ecological Applications,* 32**,** e2532.

DIVISION, W. 2022. Michigan Wolf Mortality Database. *In:* RESOURCES, M. D. O. N. (ed.).

ENTICOTT, G. 2011. Techniques of neutralising wildlife crime in rural England and Wales. *Journal of Rural Studies,* 27**,** 200-208.

ERICSSON, G. & HEBERLEIN, T. A. 2003. Attitudes of hunters, locals, and the general public in Sweden now that the wolves are back. *Biological conservation,* 111**,** 149-159.

ESTES, J. A., TERBORGH, J., BRASHARES, J. S., POWER, M. E., BERGER, J., BOND, W. J., CARPENTER, S. R., ESSINGTON, T. E., HOLT, R. D. & JACKSON, J. B. 2011. Trophic downgrading of planet Earth. *science,* 333**,** 301-306.

GROUP, U. S. C. *Survival Analysis with Stata* [Online]. UCLA: Statistical Consulting Group. [Accessed June 6 2023].

Guthery, F. S. 2003. Model selection and multimodel inference: a practical information-theoretic approach. JSTOR.

HAMILTON, L. C., LAMBERT, J. E., LAWHON, L. A., SALERNO, J. & HARTTER, J. 2020. Wolves are back: Sociopolitical identity and opinions on management of Canis lupus. *Conservation Science and Practice,* 2**,** e213.

HEBBLEWHITE, M. & HAYDON, D. T. 2010. Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365**,** 2303-2312.

HOGBERG, J., TREVES, A., SHAW, B. & NAUGHTON-TREVES, L. 2016. Changes in attitudes toward wolves before and after an inaugural public hunting and trapping season: early evidence from Wisconsin's wolf range. *Environmental Conservation,* 43**,** 45-55.

HOSMER, D. W., LEMESHOW, S. & MAY, S. 1999. Regression modeling of time to event data. *New York*.

HOUSTON, M. J., BRUSKOTTER, J. T. & FAN, D. 2010. Attitudes toward wolves in the United States and Canada: a content analysis of the print news media, 1999–2008. *Human Dimensions of Wildlife,* 15**,** 389-403.

KALBFLEISCH, J. D. & PRENTICE, R. L. 2011. *The statistical analysis of failure time data*, John Wiley & Sons.

LAKES, G., TIMES, G. & OUTDOORS, G. 2006. REVIEW OF SOCIAL AND BIOLOGICAL SCIENCE RELEVANT TO WOLF MANAGEMENT IN MICHIGAN.

LIBERG, O., CHAPRON, G., WABAKKEN, P., PEDERSEN, H. C., HOBBS, N. T. & SAND, H. 2012. Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. *Proceedings of the Royal Society B: Biological Sciences,* 279**,** 910-915.

LOUCHOUARN, N. X., SANTIAGO-ÁVILA, F. J., PARSONS, D. R. & TREVES, A. 2021. Evaluating how lethal management affects poaching of Mexican wolves. *Royal Society open science,* 8**,** 200330.

MANFREDO, M. J., TEEL, T. L., SULLIVAN, L. & DIETSCH, A. M. 2017. Values, trust, and cultural backlash in conservation governance: The case of wildlife management in the United States. *Biological Conservation,* 214**,** 303-311.

MIDNR (2022) *Michigan Wolf Management Plan,* Lansing, Michigan: Michigan Department of Natural ResourcesWildlife Division Report No. 3703).Morell, V. 2008. Wolves at the door of a more dangerous world. American Association for the Advancement of Science.

Morell, V. 2008. Wolves at the door of a more dangerous world. American Association for the Advancement of Science.

MUTH, R. M. & BOWE JR, J. F. 1998. Illegal harvest of renewable natural resources in North America: Toward a typology of the motivations for poaching. *Society & Natural Resources,* 11**,** 9-24.

NAUGHTON-TREVES, L., GROSSBERG, R. & TREVES, A. 2003. Paying for Tolerance: Rural Citizens 'Attitudes toward Wolf Depredation and Compensation. *Conservation Biology,* 17**,** 1500-1511.

NIEMIEC, R., BERL, R. E., GONZALEZ, M., TEEL, T., CAMARA, C., COLLINS, M., SALERNO, J., CROOKS, K., SCHULTZ, C. & BRECK, S. 2020. Public perspectives and media reporting of wolf reintroduction in Colorado. *PeerJ,* 8**,** e9074.

OLSON, E. R., CRIMMINS, S. M., BEYER JR, D. E., MACNULTY, D. R., PATTERSON, B. R., RUDOLPH, B. A., WYDEVEN, A. P. & VAN DEELEN, T. R. 2017. Flawed analysis and unconvincing interpretation: a comment on Chapron and Treves 2016. *Proceedings of the Royal Society B: Biological Sciences,* 284**,** 20170273.

OLSON, E. R., STENGLEIN, J. L., SHELLEY, V., RISSMAN, A. R., BROWNE-NUÑEZ, C., VOYLES, Z., WYDEVEN, A. P. & VAN DEELEN, T. 2015. Pendulum swings in wolf management led to conflict, illegal kills, and a legislated wolf hunt. *Conservation Letters,* 8**,** 351-360.

PEPIN, K. M., KAY, S. L. & DAVIS, A. J. 2017. Comment on:'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore'. *Proceedings of the Royal Society B: Biological Sciences,* 284**,** 20161459.

PIERRE, J. M. 2019. The psychology of guns: risk, fear, and motivated reasoning. *Palgrave Communications,* 5.

RIPPLE, W. J., CHAPRON, G., LÓPEZ-BAO, J. V., DURANT, S. M., MACDONALD, D. W., LINDSEY, P. A., BENNETT, E. L., BESCHTA, R. L., BRUSKOTTER, J. T. & CAMPOS-ARCEIZ, A. 2016. Saving the world's terrestrial megafauna. *Bioscience,* 66**,** 807-812.

RIPPLE, W. J., ESTES, J. A., BESCHTA, R. L., WILMERS, C. C., RITCHIE, E. G., HEBBLEWHITE, M., BERGER, J., ELMHAGEN, B., LETNIC, M. & NELSON, M. P. 2014. Status and ecological effects of the world's largest carnivores. *Science,* 343.

ROELL, B. J., BEYER JR, D. E., HOGREFE, T. C., LONSWAY, D. H., SITAR, K. L. & LEDERLE, P. E. 2009. Michigan wolf management, 2008 report. *Michigan Department of Natural Resources Wildlife Division Report,* 3498.

Raftery, A. E. (1995) 'Bayesian model selection in social research', *Sociological methodology*, pp. 111-163.

SANTIAGO-ÁVILA, F. J., CHAPPELL, R. J. & TREVES, A. 2020. Liberalizing the killing of endangered wolves was associated with more disappearances of collared individuals in Wisconsin, USA. *Scientific reports,* 10**,** 1-14.

SANTIAGO-ÁVILA, F. J. & TREVES, A. 2022. Poaching of protected wolves fluctuated seasonally and with non-wolf hunting. *Scientific reports,* 12**,** 1-10.

SERENARI, C. & PETERSON, M. N. 2016. A sociopolitical perspective on the illegal take of wildlife in the southeastern, USA.

STENGLEIN, J. L., WYDEVEN, A. P. & VAN DEELEN, T. R. 2018. Compensatory mortality in a recovering top carnivore: wolves in Wisconsin, USA (1979–2013). *Oecologia,* 187**,** 99-111.

STENGLEIN, J. L., ZHU, J., CLAYTON, M. K. & VAN DEELEN, T. R. 2015. Are the numbers adding up? Exploiting discrepancies among complementary population models. *Ecology and evolution,* 5**,** 368-376.

STIEN, A. 2017. Blood may buy goodwill: no evidence for a positive relationship between legal culling and poaching in Wisconsin. *Proceedings of the Royal Society B: Biological Sciences,* 284**,** 20170267.

SUUTARINEN, J. & KOJOLA, I. 2017. Poaching regulates the legally hunted wolf population in Finland. *Biological Conservation,* 215**,** 11-18.

TOMKIEWICZ, S. M., FULLER, M. R., KIE, J. G. & BATES, K. K. 2010. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365**,** 2163-2176.

TREVES, A. 2009. Hunting for large carnivore conservation. *Journal of Applied Ecology,* 46**,** 1350-1356.

TREVES, A., ARTELLE, K. A., DARIMONT, C. T. & PARSONS, D. R. 2017. Mismeasured mortality: correcting estimates of wolf poaching in the United States. *Journal of Mammalogy,* 98**,** 1256-1264.

TREVES, A. & BRUSKOTTER, J. 2014. Tolerance for predatory wildlife. *Science,* 344**,** 476-477.

TREVES, A. & MENEFEE, L. 2022. Adverse effects of hunting with hounds on participants and bystanders. *bioRxiv***,** 2022.08. 16.504031.

TREVES, A. & NAUGHTON-TREVES, L. 2005. Evaluating lethal control in the management of human-wildlife conflict. *CONSERVATION BIOLOGY SERIES-CAMBRIDGE-,* 9**,** 86.

TREVES, A., NAUGHTON-TREVES, L. & SHELLEY, V. 2013. Longitudinal analysis of attitudes toward wolves. *Conservation Biology,* 27**,** 315-323.

VAN EEDEN, L. M., S RABOTYAGOV, S., KATHER, M., BOGEZI, C., WIRSING, A. J. & MARZLUFF, J. 2021. Political affiliation predicts public attitudes toward gray wolf (Canis lupus) conservation and management. *Conservation Science and Practice,* 3**,** e387.

VON ESSEN, E., HANSEN, H. P., KÄLLSTRÖM, H. N., PETERSON, M. N. & PETERSON, T. R. 2015. The radicalisation of rural resistance: How hunting counterpublics in the Nordic countries contribute to illegal hunting. *Journal of rural studies,* 39**,** 199-209.

VUCETICH, J. A., BRUSKOTTER, J. T., NELSON, M. P., PETERSON, R. O. & BUMP, J. K. 2017. Evaluating the principles of wildlife conservation: a case study of wolf (Canis lupus) hunting in Michigan, United States. *Journal of Mammalogy,* 98**,** 53-64.

WILLIAMS, H. J., TAYLOR, L. A., BENHAMOU, S., BIJLEVELD, A. I., CLAY, T. A., DE GRISSAC, S., DEMŠAR, U., ENGLISH, H. M., FRANCONI, N. & GÓMEZ-LAICH, A. 2020. Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology,* 89**,** 186-206.

WILLIAMS, P. 2022. Killing wolves to own the libs? *The New Yorker.* online.

WOODROFFE, R. & GINSBERG, J. R. 1998. Edge Effects and the Extinction of Populations Inside Protected Areas. *Science* [Online], 280. Available: http://science.sciencemag.org/content/280/5372/2126.abstract.

WYDEVEN, A. P., TREVES, A., BROST, B. & WIEDENHOEFT, J. E. 2004. Characteristics of wolf packs in Wisconsin: identification of traits influencing depredation. *People and predators: from conflict to coexistence***,** 28-50.

Supplemental Materials

Table SM1. Univariate analysis results for the LTF endpoint (n=206). Categorical variables are tested using logrank tests, while continuous variables are tested using univariate cox models. P values >0.25 signify a variable that is unlikely to contribute substantially the fit and prediction capabilities of a multivariate model.

Variable	P value
Categorical (logrank tests)	
Winter Severity	**0.11**
Collar Type	**0.0033**
Policy Scheme 1	0.26
Policy Scheme 2	**<0.00000001**
Policy Scheme 3	**0.0099**
Policy Scheme 4	**0.0019**
Policy Scheme 5	**0.0006**
Presidential Administration	**<0.00000001**
Presidential Party	**0.0006**
Gubernatorial Administration	**<0.00000001**
Governor’s Party	0.39
Risk Season	**0.06**
Proportion of Population Monitored	0.33
Continuous	
Time since wolf return	**<0.00000001**

Table SM2. Examination of collinearity of each pair of variables which passes the univariate analysis step in Table SM1. Variance inflation factor (VIF) is reported for each pair of variables. VIF values greater than 5 represent higher rate of collinearity.

	Collar Type									
Winter Severity	1.38	Winter								
Policy Scheme 2	1.71	3.65	Policy Scheme 2							
Policy Scheme 3	1.09	1.83	-	Policy Scheme 3						
Policy Scheme 4	1.18	3.16	-	-	Policy Scheme 4					
Policy Scheme 5	2.32	2.30	-	-	-	Policy Scheme 5				
President	1.98	4.48	**9.54**	1.65	2.74	3.86	President			
Presidential Party	1.08	1.77	1.91	1.35	1.44	1.27	-	Presidential Party		
Governor	1.90	4.34	**13.14**	1.78	3.26	4.38	**14.98**	1.69	Governor	
Risk Season	1.15	2.35	1.75	1.41	1.81	1.41	2.07	1.43	1.98	Risk Season
Time since return	1.64	**7.27**	**8.70**	1.83	3.82	3.37	**21.01**	1.87	**14.97**	2.46

Table SM3. Univariate cox model results for the LTF endpoint for potential nuisance variables and the risk season variable

Winter severity		
0.69 (0.17) [0.42-1.13]		
Tvc(winter) 1.07 (0.046) [0.99-1.17]		
Collar type		
0.20 (0.19) [0.028-1.37]		
Tvc (collar) 1.46* (0.25) [1.05-2.03]		
Risk season (baseline: no hunt/no hound/no snow)		
Hunting or hounding/no snow	hounding/snow	Hunting or hounding/snow
1.02 (0.22) [0.65-1.59]	1.60* (0.37) [1.02-2.51]	1.14 (0.34) [0.63-2.05]
Proportion of population monitored (baseline <5%)		
5-10%	10-15%	>15%
0.77 (0.17) [0.50-1.20]	0.87 (0.21) [0.54-1.40]	0.67 (0.16) [0.42-1.07]
Time since return		
0.93 (0.044) [0.84-1.01]		
Tvc(time since return) 1.02* (0.084) [1.00-1.04]		

*P $\leq$ 0.05 ** p $\leq$ 0.005 *** p $\leq$ 0.001

Table SM4. Cox model results for the LTF endpoint for policy and politics related hypotheses, stratified by collar type.

Policy Scheme 2 (baseline: state endangered)					
State threatened – no killing	State threatened - killing	State protected species – no killing	State protected species – killing	State Game – no killing	State Game – killing
2.06* (0.59) [1.18-3.63]	1.67 (0.56) [0.86-3.23]	3.05*** (0.91) [1.69 – 5.49]	1.75 (0.85) [0.68 – 4.54]	0.0095*** (0.012) [0.00080 – 0.11]	2.91*** (0.96) [1.52 – 5.55]
Tvc(State game – no killing)				2.36*** (0.45) [1.62-3.43]	
Policy Scheme 3 (baseline: consistently listed)					
Consistently delisted		Inconsistent management			
1.07 (0.18) [0.76-1.50]		6.78** (4.60) [1.79-25.64]			
Policy Scheme 4 (baseline: listing pre-2003)					
Listing post-2003			Delisting		
2.03** (0.51) [0.24-3.33]			2.04**(0.51) [1.26-3.32]		
Policy Scheme 5 (baseline: listed pre-harvest listed)					
Delisted pre-harvest	Harvest	Listed post-harvest		Delisted post-harvest	
0.99 (0.19) [0.68-1.44]	1.61 (0.42) [0.97-2.68]	0.0059** (0.0096) [0.00025-0.14]		1.25 (0.52) [0.55-2.83]	
Tvc(listed post-harvest)		2.27*** (0.54) [1.42-3.62]			
Presidential administration (baseline: Clinton)					
Bush	Obama		Trump	Biden	
1.17 (0.37) [0.63-2.16]	2.05* (0.65) [1.11-3.80]		0.00058*** (0.0011) [0.00016-0.021]	2.10 (1.09) [0.76-5.80]	
Tvc (Trump)			3.56 ***(0.93) [2.13-5.94]		

Presidential Party (baseline: republican)		
1.55** (0.23) [1.16-2.07]		
Gubernatorial administration (baseline: Englert)		
Granholm	Snyder	Whitmer
1.68* (0.43) [1.02-2.77]	2.53***(0.69) [1.48-4.31]	3.55** (1.52) [1.53-8.23]

*P ≤ 0.05 ** p ≤ 0.005 *** p ≤ 0.001

Table SM5. Model selection process for the LTF endpoint. Only models with significant results (p<0.05) are shown. Number of variables, AIC, BIC and log-likelihood values are shared for every model built. Change in AIC and BIC values from the first and most complex model are also shared. The best model (bolded) is chosen based on AIC and BIC values, as well as parsimony. Proportionality of each model was checked, and the relevant tvc were added when nonproportionality of the model without tvc was found. All models are stratified by collar type.

Model Number	Model Structure	# variables	AIC	DAIC	BIC	DBIC	Log-likelihood
1	***Presidential administration + tvc (Trump administration)***	**2**	**1896**	**0**	**1932**	**0**	**-943**
2	*Presidential party*	1	1920	24	1928	-4	-959
3	*gubernatorial administration*	1	1918	22	1939	7	-956
4	*Policy Scheme 2 + tvc (state game – no killing)*	2	1901	3	1952	20	-943
5	*Policy Scheme 3*	1	1925	29	1947	15	-959
6	*Policy Scheme 4*	1	1921	25	1935	3	-958
7	*Policy Scheme 5 + tvc(listed post-harvest)*	2	1913	17	1950	18	-951
8	*Policy Scheme 5 + tvc(listed post-harvest) + gubernatorial administration*	3	1903	7	1961	29	-943

9	*Policy Scheme 4 + Presidential Party*	2	1914	18	1935	3	-954
10	*Policy Scheme 5 + tvc(listed post-harvest) + Presidential party*	3	1911	15	1955	23	-949
11	*Presidential party + Gubernatorial Administration*	2	1915	19	1944	12	-953
12	*Policy Scheme 3 + tvc (Delisted) + gubernatorial administration*	3	1919	23	1963	31	-953

Table SM6. Univariate analysis results for the reported poached endpoint (n=105). Categorical variables are tested using logrank tests, while continuous variables are tested using univariate cox models. P values >0.25 signify a variable that is unlikely to contribute substantially the fit and prediction capabilities of a multivariate model.

Variable	P value
Categorical (logrank tests)	
Winter Severity	0.53
Collar Type	0.79
Policy Scheme 1	**0.18**
Policy Scheme 2	0.44
Policy Scheme 3	0.60
Policy Scheme 4	**0.10**
Policy Scheme 5	0.61
Presidential Administration	**0.20**
Presidential Party	0.45
Gubernatorial Administration	0.56
Governor’s Party	0.94
Risk Season	**0.00076**
Proportion of Population Monitored	**0.089**
Continuous	
Time since wolf return	**0.18**

Table SM7. Examination of collinearity of each pair of variables which passes the univariate analysis step in Table SM6. Variance inflation factor (VIF) is reported for each pair of variables. VIF values greater than 5 represent higher rate of collinearity.

	Policy Scheme 4	Policy scheme 1	President	Risk Season	Proportion of population Monitored
Policy Scheme 1	-				
President	2.74	1.32			
Risk season	1.81	1.23	2.07		
Proportion of population Monitored	1.76	1.35	1.42	2.07	
Time since return	3.82	1.46	**21.01**	2.46	1.76

Table SM8. Univariate cox model results for the reported poached endpoint for each variable that passed the diagnostic step

Policy Scheme 1 (baseline: full protections)			
Reduced Protections			
0.75 (0.16) [0.49 – 1.15]			
AIC = 1104 BIC = 1111 LL = -551			
Policy Scheme 4 (baseline: listing pre-2003)			
Listing post-2003		Delisting	
1.53 (0.42) [0.91-2.61]		1.01 (0.30) [0.57-2.84]	
AIC = 1103 BIC = 1118 LL = -549			
Presidential administration (baseline: Clinton)			
Bush	Obama	Trump	Biden
2.33 (1.24) [0.82-6.61]	2.66 (1.45) [0.91-7.75]	3.23* (1.81) [1.08-9.70]	1.59 (1.14) [0.39-6.51]
AIC = 1105 BIC = 1135 LL = -548			
Risk season (baseline: no hunt/no hound/no snow)			
Hunting or hounding/no snow	hounding/snow	Hunting or hounding/snow	
0.65 (0.23) [0.32-1.30]	1.21 (0.41) [0.62-2.35]	1.94 (0.70) [0.95-3.94]	
AIC = 1098 BIC = 1120 LL = -546			
Proportion of population monitored (baseline <5%)			
5-10%	10-15%	>15%	
3.04 (1.84) [0.93-9.98]	2.74 (1.71) [0.81-9.28]	3.83 * (2.30) [1.18-12.44]	
AIC = 1102 BIC = 1124 LL = -548			
Time since return			
1.02 (0.014) [0.99-1.05]			
AIC = 1104 BIC = 1111 LL = -551			

*$P \leq 0.05$ ** $p \leq 0.005$ *** $p \leq 0.001$

Table SM9. Model selection process for the reported poached endpoint. Number of variables, AIC, BIC and log-likelihood values are shared for every model built. Change in AIC and BIC values from the first and most complex model are also shared. The best model (bolded) is chosen based on AIC and BIC values, as well as parsimony. Proportionality of each model was checked, and the relevant tvc were added when nonproportionality of the model without tvc was found. All models are stratified by collar type.

Model Number	Model Structure	# variables	AIC	DAIC	BIC	DBIC	Log-likelihood
1	*Risk Season*	1	1098	0	1120	0	-546
2	*Presidential administration*	1	1105	7	1135	15	-548

Table SM10. Univariate analysis results for the nonhuman endpoint (n=66). Categorical variables are tested using logrank tests, while continuous variables are tested using univariate cox models. P values >0.25 signify a variable that is unlikely to contribute substantially the fit and prediction capabilities of a multivariate model.

Variable	P value
Categorical (logrank tests)	
Winter Severity	**0.0019**
Collar Type	0.34
Policy Scheme 1	0.73
Policy Scheme 2	0.91
Policy Scheme 3	0.66
Policy Scheme 4	0.78
Policy Scheme 5	0.85
Presidential Administration	0.75
Presidential Party	0.34
Gubernatorial Administration	0.50
Governor’s Party	0.99
Risk Season	**0.0011**
Proportion of Population Monitored	**0.014**
Continuous	
Time since wolf return	0.89

Table SM11. Examination of collinearity of each pair of variables which passes the univariate analysis step in Table SM10. Variance inflation factor (VIF) is reported for each pair of variables. VIF values greater than 5 represent higher rate of collinearity.

	Winter	
Risk Season	2.35	Risk Season
Proportion population monitored	1.82	2.07

Winter severity (baseline:average)			
Mild	Moderate	Severe	Extreme
0.72 (0.27) [0.34-1.50]	1.55 (0.53) [0.79 – 3.04]	0.45 (0.20) [0.18 – 1.10]	0.41 (0.20) [0.15 – 1.09]
AIC = 664 BIC = 693 LL = -328			
Risk season (baseline: no hunt/no hound/no snow)			
Hunting or hounding/no snow	hounding/snow	Hunting or hounding/snow	
3.66* (2.34) [1.04-12.86]	7.28*** (4.41) [2.22-23.85]	8.99*** (5.53) [2.68-30.08]	
AIC = 659 BIC = 680 LL = -326			
Proportion of population monitored (baseline <5%)			
5-10%	10-15%	>15%	
0.88 (0.45) [0.33-2.40]	0.54 (0.32) [0.17-1.73]	1.68 (0.80) [0.66-4.28]	
AIC = 668 BIC = 689 LL = -331			

Table SM12. Univariate cox model results for the nonhuman mortality endpoint for each variable that passed the diagnostic step

Table SM13. Model selection process for the nonhuman endpoint. Number of variables, AIC, BIC and log-likelihood values are shared for every model built. Change in AIC and BIC values from the first and most complex model are also shared. The best model (bolded) is chosen based on AIC and BIC values, as well as parsimony. Proportionality of each model was checked, and the relevant tvc were added when nonproportionality of the model without tvc was found. All models are stratified by collar type.

Model Number	Model Structure	# variables	AIC	DAIC	BIC	DBIC	Log-likelihood
1	***Risk Season***	**1**	**659**	**0**	**680**	**0**	**-326**

2	*Winter Severity*	1	664	5	693	13	-328
3	*Proportion of Population Monitored*	1	668	9	689	9	-331

Conclusion

Conservation science is a problem-oriented field with an element of practical application. The objective of the ‘science ’portion of conservation is to test the implemented solutions, understand causes of problems and mechanisms leading to desirable or undesirable effects and suggest better solutions (Williams, Balmford and Wilcove, 2020). The field is highly interdisciplinary and can suffer from challenges common to many scientific disciplines. Sound scientific analysis in any field requires the following elements: reproducibility, transparency and independent review (Munafò *et al.*, 2017; Powers and Hampton, 2019; Treves, 2022). Lack of standardization and randomization in data collection can severely bias study designs and make results non-reproducible. This is true of any field, especially those plagued with small sample sizes and low power experiments (Ioannidis, 2005). Most environmental science fields suffer from these challenges, as many questions must be tested at large spatial and temporal scales, in systems that are not easy to manipulate (Sagarin and Pauchard, 2010). Even at smaller scales, funding limitations and the relative infancy of the sub-field may limit sample sizes, statistical power or overall ability to draw conclusions from results (Christie *et al.*, 2020). Replication of results in various conditions and through differing study designs improves confidence in the findings, and reduces the risk of false positives (Christie *et al.*, 2020).

Human-predator coexistence as a sub-field is still relatively young, as many large carnivores have only begun to return to historic ranges in the past few decades. Therefore, many hypotheses in the subfield of human-carnivore coexistence have yet to be tested or have little evidence to support or refute predictions. Given this, some biases will exist until the body of literature grows, and more study designs can replicate or refute findings across a range of conditions (Christie *et al.*, 2020). It is therefore imperative that the science produced is held to a high standard, to minimize as many biases as possible. Doing this may require an assessment of study designs and examination of false positive rates (FPR) for those studies that exist.

A better understanding of false positive rates is emerging, and suggests that the conventional definition of statistical significance has led to a lack of reproducibility and an over-estimation of effect sizes across scientific disciplines (Benjamin *et al.*, 2018; Colquhoun, 2014). In other words, many studies claim to observe an effect, when in truth none exists. This is likely a result of a misunderstanding of true FPR, and low power experiments (Benjamin *et al.*, 2018; Colquhoun, 2014; Khorozyan, 2021). Colquhoun (2014) suggests that the use of significance values of $p = 0.05$, which is the standard significance value across scientific disciplines, will result in false positives 30% of the time. However, Khorozyan (2021) estimated that FPR of studies on electric fencing effectiveness to deter large carnivores could be as high as 30 to 50%.

In this conclusion to my dissertation, I explore how the different study designs I use in my chapters minimize biases, and how they may have suffered from additional biases. I will then make recommendations about how future analyses in the subfield could improve upon those I describe here.

Randomized-control design

The first chapter of this dissertation describes a randomized-control trial with cross-over design. This study design was intended to maximize the quality of the collected data while minimizing possible biases. This was achieved in several ways. First, all cow herds present on our study ranch were included in the study, therefore we limited the risk of self-selection bias by ranchers (Treves *et al.*, 2019). Second, randomization is an important tool for reducing selection bias, i.e., which study subject receives a treatment and when (Treves *et al.*, 2019). Without randomization, we risk increasing the probability that observed effects are occurring by random chance instead of the treatment in question (Colquhoun, 2014). In this study, we randomly allocated treatment or control condition to each of our 8 sample herds by using a random number generator. Each herd was allocated a number from 1-8, the first four randomly selected herds started the study in control conditions, and the second four randomly selected herds started the study in treatment conditions. We further limited the selection bias, and the possible biasing effect of time, by using a cross-over design, which meant all subjects switched from their first tt condition to the other condition midway through the experiment (Treves *et al.*, 2019). We ensured that each herd was exposed to each treatment condition for an equal amount of time. We further allocated a one week (7 day) ‘wash-out ’period between treatments, during which time no treatment was applied to the herds, ideally allowing any treatment effect to dissipate (Treves *et al.*, 2019; Diaz-Uriarte, 2002). We then compared subject herds to themselves in each condition, reducing confounding differences between herds that may not have been controlled for otherwise (Treves *et al.*, 2019).

We further reduced treatment bias, i.e., the standardization, or lack there-of, of treatment application, by using the same range riders, trained together by our co-op manager, and ensuring that they were consistent about their visits to treatment herds (Treves *et al.*, 2019). Range riders recorded their visits to herds in logs and with audio recorders, so that I, as the field researcher, could ensure that visits to herds were consistent. Measurement bias, i.e., methods for recording response variables are inconsistent, was reduced by using multiple forms of response variable data collection. First, an equal number of trail cameras were used in each herd, and these were placed using a systematic random sampling scheme within each pasture. Additional data were collected using predator sign surveys which were conducted along main trails in each pasture and around each camera. Bias was minimized in these surveys by ensuring only one person conducted surveys, and the same methods for surveying were used in every herd, regardless of treatment condition. The best way to reduce measurement bias in this case would have been to ensure the field researcher conducting surveys did not know whether the herd was a treatment or control herd (Treves *et al.*, 2019). One way to do this would be to have one researcher manage the subject treatments and sampling schedule, and another field researcher to do the tasks of field sampling while being ‘blinded ’to the treatment condition. In this experiment this was impossible given this approach requires at least two field researchers, which we did not have the funds to support.

While the methods described above eliminated some biases, after publication of these results we came across literature by Colquhoun (2014) and Benjamin et al. (2018) which demonstrated this study did not fully eliminate all likely statistical biases. Benjamin et al. (2018) state that it is imperative that we as scientist go beyond a reliance on p-values and improve study designs to reduce biases. We have overall done so in this case, but we fall short in two important ways. First, the power of our experiment was low because our sample size was only 8 herds. Although each herd did receive each treatment condition, to achieve a statistical power of 80% (i.e., to ensure that if an effect exists it will be observed 80% of the time) a minimum sample size of 16 is required (Colquhoun, 2014). Mitigating those pitfalls, we note that we had very few ‘statistically significant ’results, therefore we generally did not report a discovery which may have been a false positive (but see below for correlation analysis). Second, we examined the responses of multiple species to range riding. When studies examine multiple outcomes in response to a single treatment, it can lead to increased rates of false positives. Therefore, we would have been better served examining the response from only one species of large carnivore instead of five.

Correlation analysis

Correlation analyses are known to introduce additional biases and weaken our ability to make inferences about results given that these analyses do not clarify true causality of observed effects (Treves *et al.*, 2019). In our first chapter, our only reported significant results were a dose effect of number of range riders and grizzly bear presence around herds. This effect was reported and deemed to be significant with a p-value < 0.1. In the chapter we note that this inference was weakest, given the added confounding effect of time on the analysis, though the step that herds had been randomly assigned to treatment conditions, and treatments were applied at different times did improve the level of inference for this analysis (Treves *et al.*, 2019). However, after further reading, we find that our significance level of $p < 0.1$ was far too high, and would have introduced an unacceptably high risk of false positive (Benjamin *et al.*, 2018; Colquhoun, 2014). According to Benjamin et al. (2018), the best way to reduce this risk, especially when reporting a new finding, as we have done here, is to limit the significance level to $p \leq 0.005$. By this standard, therefore, we cannot reliably conclude that grizzly bears responded to number of days with range rider visits. However, we are replicating a claimed result (Stone *et al.*, 2017), so our results seem suggestive of a treatment effect.

Time-to-event analyses

In the second and third chapters of this dissertation I use time-to-event analysis methods to examine cause-specific mortality of radio-collared wolves in two populations. Any analysis which includes cross-sectional measurements over time and is not able to control for the effect of time, must consider the potential confounding effect of temporal autocorrelation (Underwood, 1992). Time-to-event analyses, by contrast, use time as a response variable because they examine the amount of time elapsed until an event of interest occurs (Kalbfleisch and Prentice, 2002).

Furthermore, these models are autoregressive, and therefore account for the effect of time on the outcome variable, which a typical linear regression would not do (Guo, 2009; Underwood, 1992).

Although many biomedical studies use time-to-event analyses to compare across groups and therefore can examine differences between treatment and control groups, our analyses could not do this. First, the entire wolf population in a given jurisdiction would have been exposed to the same intervention, which in this case was changes in policy or partisan political variables. There would have been no way to separate a portion of each wolf population and keep them from being exposed to these interventions. This analysis therefore cannot rise to the highest level of inference, which would require a control group. Once more similar studies have emerged across jurisdictions, we will be able to judge replication and at least compare groups, though we have not yet determined an approach that would produce a true control.

Furthermore, any study of radio-collared animals will suffer from some unknown level of selection bias. Marked animals are a non-random subsample of a population, determined by the individual animals willingness to enter a trap, and the trap location and sampling effort. The study we conducted on Mexican wolves likely goes as far as possible towards reducing this bias because the population was so small, and nearly 75% of the population was monitored by collaring.

We did however reduce other potential biases in our time-to-event analyses. First we reduced reporting bias by submitting the Mexican wolf analysis as a registered report. Registered reports allow researchers to receive peer-review on introduction and methods before conducting data analyses. If accepted, the journal agrees to publish any results, regardless of statistical significance, thereby reducing the bias toward positive results in the literature (Treves *et al.*, 2019; Chambers and Tzavella, 2022). Second, we reduced the risk of false positives by using bayes factors (BF) to interpret our Mexican wolf results in relation to hypothesized outcomes. This represents an effort at replication. BFs allow researchers to more directly explore the strength of evidence to support their hypothesis by comparing their results to some informative prior based on validated previous results (Benjamin *et al.*, 2018).

Overall conclusion

As in any scientific field, biases, funding and data collection challenges can present barriers to the correct interpretation of scientific results. Because conservation issues are time sensitive, it behooves us as researchers to rise to these challenges and mitigate their impact by being cognizant of possible biases and reducing them to the best of our ability. This will involve improving study designs and being transparent about possible unavoidable biases. Here I have explored both how we attempted to reduce biases, and how certain biases may have entered these analyses to illustrate how high standard experiments can be developed, and when these analyses have fallen short. My hope is that the methodology and this conclusion is as useful to the subfield of human-carnivore coexistence science as are the findings of my experiments.

References

Artelle, K. A., Reynolds, J. D., Treves, A., Walsh, J. C., Paquet, P. C. and Darimont, C. T. (2018) 'Hallmarks of science missing from North American wildlife management', *Science Advances*, 4(3), pp. eaao0167.

Barber-Meyer, S., Wheeldon, T. and Mech, L. D. (2021) 'The importance of wilderness to wolf (Canis lupus) survival and cause-specific mortality over 50 years', *Biological Conservation*, 258, pp. 109145.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L. and Camerer, C. (2018) 'Redefine statistical significance', *Nature human behaviour*, 2(1), pp. 6-10.

Browne-Nuñez, C., Treves, A., MacFarland, D., Voyles, Z. and Turng, C. (2015) 'Tolerance of wolves in Wisconsin: a mixed-methods examination of policy effects on attitudes and behavioral inclinations', *Biological Conservation,* 189, pp. 59-71.

Bruskotter, J. T., Schmidt, R. H. and Teel, T. L. (2007) 'Are attitudes toward wolves changing? A case study in Utah', *Biological conservation,* 139(1-2), pp. 211-218.

Bruskotter, J. T., Vucetich, J. A., Enzler, S., Treves, A. and Nelson, M. P. (2014) 'Removing protections for wolves and the future of the US Endangered Species Act (1973)', *Conservation Letters,* 7(4), pp. 401-407.

Bruskotter, J. T. and Wilson, R. S. (2014) 'Determining where the wild things will be: using psychological theory to find tolerance for large carnivores', *Conservation Letters,* 7(3), pp. 158-165.

Chambers, C. D. and Tzavella, L. (2022) 'The past, present and future of registered reports', *Nature human behaviour,* 6(1), pp. 29-42.

Chapron, G. and Treves, A. (2016) 'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore', *Proceedings of the Royal Society B: Biological Sciences,* 283(1830), pp. 20152939.

Chapron, G. and Treves, A. (2017) 'Reply to comment by Pepin et al. 2017', *Proceedings of the Royal Society B: Biological Sciences,* 284(1851), pp. 20162571.

Christie, A. P., Abecasis, D., Adjeroud, M., Alonso, J. C., Amano, T., Anton, A., Baldigo, B. P., Barrientos, R., Bicknell, J. E. and Buhl, D. A. (2020) 'Quantifying and addressing the prevalence and bias of study designs in the environmental and social sciences', *Nature communications,* 11(1), pp. 6377.

Colquhoun, D. (2014) 'An investigation of the false discovery rate and the misinterpretation of p-values', *Royal Society open science,* 1(3), pp. 140216.

Diaz-Uriarte, R. (2002) 'Incorrect analysis of crossover trials in animal behaviour research', *Animal behaviour*.

Elbroch, L. M. and Treves, A. (2023) 'Perspective: Why might removing carnivores maintain or increase risks for domestic animals?', *Biological Conservation,* 283, pp. 110106.

Ericsson, G. and Heberlein, T. A. (2003) 'Attitudes of hunters, locals, and the general public in Sweden now that the wolves are back', *Biological conservation,* 111(2), pp. 149-159.

Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D. and Jackson, J. B. (2011) 'Trophic downgrading of planet Earth', *science,* 333(6040), pp. 301-306.

Guo, S. (2009) 'Survival Analysis'.

Hogberg, J., Treves, A., Shaw, B. and Naughton-Treves, L. (2016) 'Changes in attitudes toward wolves before and after an inaugural public hunting and trapping season: early evidence from Wisconsin's wolf range', *Environmental Conservation,* 43(1), pp. 45-55.

Humane Society of U.S. v. Kempthorne [2006].

Ioannidis, J. P. (2005) 'Why most published research findings are false', *PLoS medicine,* 2(8), pp. e124.

Kalbfleisch, J. D. and Prentice, R. L. (2002) *The Statistical Analysis of Failure Time Data.* John Wiley & Sons.

Kellert, S. R., Black, M., Rush, C. R. and Bath, A. J. (1996) 'Human culture and large carnivore conservation in North America', *Conservation Biology,* 10(4), pp. 977-990.

Khorozyan, I. (2021) 'Dealing with false positive risk as an indicator of misperceived effectiveness of conservation interventions', *Plos one,* 16(8), pp. e0255784.

Liberg, O., Chapron, G., Wabakken, P., Pedersen, H. C., Hobbs, N. T. and Sand, H. (2012) 'Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe', *Proceedings of the Royal Society B: Biological Sciences,* 279(1730), pp. 910-915.

Louchouarn, N. X., Santiago-Ávila, F. J., Parsons, D. R. and Treves, A. (2021) 'Evaluating how lethal management affects poaching of Mexican wolves', *Royal Society open science,* 8(3), pp. 200330.

MIDNR (2022) *Michigan Wolf Management Plan*, Lansing, Michigan: Michigan Department of Natural ResourcesWildlife Division Report No. 3703).

Morell, V. 2008. Wolves at the door of a more dangerous world. American Association for the Advancement of Science.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J. and Ioannidis, J. (2017) 'A manifesto for reproducible science', *Nature human behaviour,* 1(1), pp. 1-9.

Naughton-Treves, L., Grossberg, R. and Treves, A. (2003) 'Paying for Tolerance: Rural Citizens' Attitudes toward Wolf Depredation and Compensation', *Conservation Biology,* 17(6), pp. 1500-1511.

Oliynyk, R. T. (2023) 'Human-caused wolf mortality persists for years after discontinuation of hunting', *Scientific Reports,* 13(1), pp. 11084.

Organ, J., Geist, V., Mahoney, S., Williams, S., Krausman, P., Batcheller, G., Decker, T., Carmichael, R., Nanjappa, P. and Regan, R. (2012a) 'The North American Model of Wildlife Conservation', *The Wildlife Society Technical Review,* pp. 12-04.

Organ, J. F., Geist, V., Mahoney, S. P., Williams, S., Krausman, P. R., Batcheller, G., Decker, T., Carmichael, R., Nanjappa, P. and Regan, R. (2012b) 'The North American model of wildlife conservation', *The Wildlife Society Technical Review,* 12(04).

Powers, S. M. and Hampton, S. E. (2019) 'Open science, reproducibility, and transparency in ecology', *Ecological applications,* 29(1), pp. e01822.

Ripple, W. J., Estes, J. A., Beschta, R. L., Wilmers, C. C., Ritchie, E. G., Hebblewhite, M., Berger, J., Elmhagen, B., Letnic, M. and Nelson, M. P. (2014) 'Status and ecological effects of the world's largest carnivores', *Science,* 343(6167).

Ripple, W. J., Wolf, C., Phillips, M. K., Beschta, R. L., Vucetich, J. A., Kauffman, J. B., Law, B. E., Wirsing, A. J., Lambert, J. E. and Leslie, E. (2022) 'Rewilding the American West', *BioScience*.

Sagarin, R. and Pauchard, A. (2010) 'Observational approaches in ecology open new ground in a changing world', *Frontiers in Ecology and the Environment,* 8(7), pp. 379-386.

Santiago-Ávila, F. J., Agan, S., Hinton, J. W. and Treves, A. (2022) 'Evaluating how management policies affect red wolf mortality and disappearance', *Royal Society Open Science,* 9(5), pp. 210400.

Santiago-Ávila, F. J., Chappell, R. J. and Treves, A. (2020) 'Liberalizing the killing of endangered wolves was associated with more disappearances of collared individuals in Wisconsin, USA', *Scientific reports,* 10(1), pp. 1-14.

Santiago-Avila, F. J., Cornman, A. M. and Treves, A. (2018) 'Killing wolves to prevent predation on livestock may protect one farm but harm neighbors', *PLoS One,* 13(1), pp. e0189729.

Soulé, M. E. (1985) 'What is conservation biology?', *BioScience,* 35(11), pp. 727-734.

Stone, S. A., Breck, S. W., Timberlake, J., Haswell, P. M., Najera, F., Bean, B. S. and Thornhill, D. J. (2017) 'Adaptive use of nonlethal strategies for minimizing wolf–sheep conflict in Idaho', *Journal of Mammalogy,* 98(1), pp. 33-44.

Suutarinen, J. and Kojola, I. (2017) 'Poaching regulates the legally hunted wolf population in Finland', *Biological Conservation,* 215, pp. 11-18.

Treves, A. (2009) 'Hunting for large carnivore conservation', *Journal of Applied Ecology,* 46(6), pp. 1350-1356.

Treves, A. (2022) '"Best available science" and the reproducibility crisis', *Frontiers in Ecology and the Environment,* 20(9), pp. 495-495.

Treves, A., Artelle, K. A., Darimont, C. T. and Parsons, D. R. (2017a) 'Mismeasured mortality: correcting estimates of wolf poaching in the United States', *Journal of Mammalogy,* 98(5), pp. 1256-1264.

Treves, A. and Bruskotter, J. (2014) 'Tolerance for predatory wildlife', *Science,* 344(6183), pp. 476-477.

Treves, A., Chapron, G., López-Bao, J. V., Shoemaker, C., Goeckner, A. R. and Bruskotter, J. T. (2017b) 'Predators and the public trust', *Biological Reviews,* 92(1), pp. 248-270.

Treves, A., Krofel, M. and McManus, J. (2016) 'Predator control should not be a shot in the dark', *Frontiers in Ecology and the Environment,* 14(7), pp. 380-388.

Treves, A., Krofel, M., Ohrens, O. and van Eeden, L. M. (2019) 'Predator control needs a standard of unbiased randomized experiments with cross-over design', *Frontiers in Ecology and Evolution,* 7, pp. 462.

Treves, A., Langenberg, J. A., López-Bao, J. V. and Rabenhorst, M. F. (2017c) 'Gray wolf mortality patterns in Wisconsin from 1979 to 2012', *Journal of mammalogy,* 98(1), pp. 17-32.

Treves, A. and Louchouarn, N. X. (2022) 'Uncertainty and precaution in hunting wolves twice in a year', *PloS one,* 17(3), pp. e0259604.

Treves, A. and Martin, K. A. (2011) 'Hunters as stewards of wolves in Wisconsin and the Northern Rocky Mountains, USA', *Society & Natural Resources,* 24(9), pp. 984-994.

Treves, A., Naughton-Treves, L. and Shelley, V. (2013) 'Longitudinal analysis of attitudes toward wolves', *Conservation Biology,* 27(2), pp. 315-323.

Underwood, A. (1992) 'Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world', *Journal of experimental marine biology and ecology,* 161(2), pp. 145-178.

Van Eeden, L. M., Eklund, A., Miller, J. R., López-Bao, J. V., Chapron, G., Cejtin, M. R., Crowther, M. S., Dickman, C. R., Frank, J. and Krofel, M. (2018) 'Carnivore conservation needs evidence-based livestock protection', *PLoS biology,* 16(9), pp. e2005577.

Vucetich, J. A., Bruskotter, J. T., Nelson, M. P., Peterson, R. O. and Bump, J. K. (2017) 'Evaluating the principles of wildlife conservation: a case study of wolf (Canis lupus) hunting in Michigan, United States', *Journal of Mammalogy,* 98(1), pp. 53-64.

Wallach, A. D., Batavia, C., Bekoff, M., Alexander, S., Baker, L., Ben-Ami, D., Boronyak, L., Cardilin, A. P., Carmel, Y. and Celermajer, D. (2020) 'Recognizing animal personhood in compassionate conservation', *Conservation Biology,* 34(5), pp. 1097-1106.

Williams, C. K., Ericsson, G. and Heberlein, T. A. (2002) 'A quantitative summary of attitudes toward wolves and their reintroduction (1972-2000)', *Wildlife Society Bulletin*, pp. 575-584.

Williams, D. R., Balmford, A. and Wilcove, D. S. (2020) 'The past and future role of conservation science in saving biodiversity', *Conservation Letters,* 13(4), pp. e12720.

Woodroffe, R. and Ginsberg, J. R. (1998) 'Edge Effects and the Extinction of Populations Inside Protected Areas.', *Science*, 280, pp. 2126-2128. Available at: http://science.sciencemag.org/content/280/5372/2126.abstract.