

Predator control should not be a shot in the dark

Adrian Treves^{1*†}, Miha Krofel^{2†}, and Jeannine McManus^{3†}

Livestock owners traditionally use various non-lethal and lethal methods to protect their domestic animals from wild predators. However, many of these methods are implemented without first considering experimental evidence of their effectiveness in mitigating predation-related threats or avoiding ecological degradation. To inform future policy and research on predators, we systematically evaluated evidence for interventions against carnivore (canid, felid, and ursid) predation on livestock in North American and European farms. We also reviewed a selection of tests from other continents to help assess the global generality of our findings. Twelve published tests – representing five non-lethal methods and 7 lethal methods – met the accepted standard of scientific inference (random assignment or quasi-experimental case-control) without bias in sampling, treatment, measurement, or reporting. Of those twelve, prevention of livestock predation was demonstrated in six tests (four non-lethal and two lethal), whereas counterintuitive increases in predation were shown in two tests (zero non-lethal and two lethal); the remaining four (one non-lethal and three lethal) showed no effect on predation. Only two non-lethal methods (one associated with livestock-guarding dogs and the other with a visual deterrent termed “fladry”) assigned treatments randomly, provided reliable inference, and demonstrated preventive effects. We recommend that policy makers suspend predator control efforts that lack evidence for functional effectiveness and that scientists focus on stringent standards of evidence in tests of predator control.

Front Ecol Environ 2016; 14(7): 380–388, doi:10.1002/fee.1312

Substantial numbers of vertebrate predators have been intentionally killed by government agencies and by private citizens acting legally or illegally (Wirsing and

Ripple 2010; Ripple *et al.* 2014). More recently, however, killing top predators – such as wolves (*Canis lupus*) and leopards (*Panthera pardus*), which occasionally prey on livestock – has prompted concerns associated with ethical issues (Vucetich and Nelson 2014), effectiveness, and ecological impacts. Depletion of apex consumers, which include most large-bodied predators, has led to the degradation of ecosystems and disruption of vital ecological processes worldwide (Estes *et al.* 2011; Ripple *et al.* 2014). As a result, traditional non-lethal methods have been reinstated and new approaches are being developed (Treves *et al.* 2009).

Questions about functional effectiveness center on whether intervening will protect property owners from future losses (“effectiveness” hereafter). The question remains unresolved for many cases but is particularly unclear for killing predators (Mitchell *et al.* 2004; Treves and Naughton-Treves 2005; Woodroffe and Redpath 2015). Although it seems obvious that killing a carnivore about to take a lamb should ensure the latter’s short-term survival, most lethal methods are applied indirectly in wholly different situations. Lethal intervention is usually implemented after carnivores are observed near livestock or days after a predation event has occurred, sometimes far from where the attack occurred (eg Treves *et al.* 2002). Historically, eradication campaigns have been aimed at reducing predation by exterminating species. However, national and global concerns about biodiversity loss have largely discouraged this, when applied to native predators (Treves and

In a nutshell:

- Predator control methods to prevent livestock loss have rarely been subject to rigorous tests using the “gold standard” for scientific inference (random assignment to control and treatment groups with experimental designs that avoid biases in sampling, treatment, measurement, or reporting)
- Across the controlled experiments that we systematically examined, higher standards of evidence were generally applied in tests of non-lethal methods than in tests of lethal methods for predator control
- Non-lethal methods were more effective than lethal methods in preventing carnivore predation on livestock generally; at least two lethal methods (government culling or regulated public hunting) were followed by increases in predation on livestock; zero tests of non-lethal methods had counterproductive effects
- All flawed tests came from North America; 10 of 12 flawed tests were published in three journals, compared to four of 12 tests with strong inference in those same journals
- We recommend suspending lethal predator control methods that do not currently have rigorous evidence for functional effectiveness in preventing livestock loss until gold-standard tests are completed

¹Nelson Institute for Environmental Studies, University of Wisconsin, Madison, WI *(atreves@wisc.edu); ²Biotechnical Faculty, Department of Forestry, University of Ljubljana, Ljubljana, Slovenia; continued on p 388

Karanth 2003; Chapron *et al.* 2014). Furthermore, over time, numerous observers have noted that killing predators could fragment predator social groups or create vacancies in the ecological community, to be filled by more numerous, smaller species of predators that in turn might prey on livestock (Young and Dobyas 1945; Newby and Brown 1958; Haber 1996; Knowlton *et al.* 1999; Prugh *et al.* 2009). Indiscriminate killing was also often ineffective in removing probable culprits (Knowlton *et al.* 1999). Finally, for both lethal and non-lethal interventions, little information was available about the behavioral and population dynamic responses of survivors or any ripple effects, whereby neighboring livestock owners suffer higher costs after predator control was implemented on a nearby property. For example, in response to moderate rates of human-induced mortality, coyotes (*Canis latrans*) frequently showed compensatory reproduction, resulting in higher population growth rates and population densities during subsequent years (Knowlton *et al.* 1999). Controversy and uncertainty about predator control generally persisted for decades in the absence of convincing evidence. Resolving this controversy will help to restore populations of predators and other species in largely undisturbed ecosystems as well as in more developed landscapes with people and domestic animals (Fischer *et al.* 2008).

Prior studies of predator control reviewed evidence for one carnivore species (eg coyotes; Mitchell *et al.* 2004) or a single type of control method (Linnell *et al.* 1997; Mason *et al.* 2001), but general conclusions were elusive because standards of evidence varied or unreliable inferences arose from uncontrolled tests. As the field matured, so did its standards of evidence. Experiments with Australian sheep (*Ovis aries*), for instance, suggested that intense and frequently repeated killing of introduced red foxes (*Vulpes vulpes*) and dingoes (*Canis lupus dingo*) produced only minimal, inconsistent protection for lambs (Greentree *et al.* 2000; Allen and Sparkes 2001). Controlled experiments on three management techniques for European badger (*Meles meles*) – a mustelid – showed that lethal interventions significantly exacerbated disease transmission from badgers to livestock (Vial and Donnelly 2012). Nevertheless, predator control methods have not been subject to comprehensive “clinical trials”, in which interventions that appear effective in “laboratory trials” are tested experimentally on real subjects, to borrow terminology and lessons from the biomedical sciences (Mukherjee 2010). Here we apply uniform criteria and an established standard of evidence to evaluate the effectiveness of various interventions used to prevent predation on livestock by carnivores (ie terrestrial members of Carnivora >5 kg, such as coyotes, wolves, bears, or big cats). We adopted the scientific framework for strong inference first articulated by Platt (1964) to review both the experimental design and the evidence for effectiveness of various, widely used lethal and non-lethal methods.

Strong inference demands the careful avoidance of bias at several stages, primarily through the use of an experimental control with random assignment of treatments, followed by unbiased measurement and reporting subjected to rigorous, anonymous peer review, with disclosure of potential conflicts of interest. For ease of discussion, we refer to random assignment of treatments as the “gold standard” for scientific inference – but we also examine whether study designs included other steps to avoid bias in sampling, measurement, or reporting. We use the scientific terms “bias” and “flawed design” without any suggestion of intentional bias or incompetence. Often well-intentioned and highly competent researchers encounter flaws in research design because of inescapable challenges presented by field conditions. Nevertheless, the gold standard of scientific inference has been embraced by practitioners within the clinical biomedical sciences because of a long history of unreliable inferences from tests that had one or more biases in the sampling of subjects, treatments, measurements, or reporting (Mukherjee 2010). Unlike scholars in the paleosciences (Gould 1980; Biondi 2014) who have made cogent arguments for a lesser standard because studies of the past can never be replicated exactly to the specifications required by scientific experimentation, ecologists have long advocated for controlled experiments in ecological research (Hairston 1989). We therefore hold our subdiscipline to the gold standard. However, the shortage of tests meeting the gold standard (see below) led us to examine an alternative “silver standard” of non-random assignment of treatments, as long as we discerned – from a close reading of the peer-reviewed, published methods – no other biases that might weaken inference. The silver standard included quasi-experimental tests with haphazard assignment of treatments (case-control or Before–After Control–Impact [BACI] designs).

■ Methods

Methods of review

We performed repeated searches of the peer-reviewed literature using Google Scholar, followed by a snowball method using the reference lists of >100 articles identified in the search. We searched with the following keywords: (control, damage, depredation, lethal, non-lethal, removal, or livestock) AND (predat* or carnivor*). For our quantitative summary of results, we included only peer-reviewed, published tests in our native languages (English and Slovenian) that (1) used experimental or quasi-experimental control with a design that allowed strong inference, (2) occurred on working livestock operations with free-ranging, native carnivores of North America or Europe, and (3) verified livestock losses.

Table 1. Tests of interventions to prevent carnivore predation on livestock that met review criteria

		Observed changes (if any) in livestock predation		
		Decrease	No difference	Increase
<i>Lethal methods</i>	Quasi-experimental tests of culling gray wolves (1) and culling, hunting, and poaching Eurasian lynx (2)		Quasi-experimental tests of hunting black bears (3*), hunting and culling brown bears (4), and culling and hunting gray wolves (5)	Quasi-experimental tests of culling coyotes (6) and hunting cougars (7**)
<i>Non-lethal methods</i>	Random assignment test of fladry on gray wolves (8), random assignment test of LGDs on gray wolves and coyotes (9), quasi-experimental tests of LGDs and night enclosures on gray wolves (10), and fladry on gray wolves (11)		Random assignment test of fladry on coyotes (8), quasi-experimental tests of diversionary feeding on brown bears (12)	

Notes: *Some complaints related to livestock predation but some related to property damage. **A quasi-experimental two-county comparison was reported in Peebles *et al.* (2013), based partly on the work of Cooley *et al.* (2009a,b). Sources of evidence are listed by number: 1 = Bradley *et al.* (2015); 2 = Herfindal *et al.* (2005); 3 = Obbard *et al.* (2014) see their Table S1 for use of moving averages; 4 = Sagor *et al.* (1997); 5 = Krofel *et al.* (2011) reanalyzed as after–before measures of livestock losses (WebPanel 1); 6 = Conner *et al.* (1998); 7 = Peebles *et al.* (2013) and Cooley *et al.* (2009a,b) treated as a single test for the two-county comparison, not the state-wide analysis; 8 = Davidson-Nelson and Gehring (2010); 9 = Gehring *et al.* (2010a,b); 10 = Espuno *et al.* (2004); 11 = Musiani *et al.* (2003); 12 = Kavčič *et al.* (2013). LGDs = livestock-guarding dogs. We excluded two studies that used time lags but not BACI designs to infer changes in livestock losses over time (eg Wielgus and Peebles 2014; Fernández-Gil *et al.* 2016).

Regarding criterion (1), we explicitly describe the reasons any test was deemed unreliable based on selection, treatment, measurement, or reporting biases in WebPanel 1. We excluded analyses that were purely correlational, those based only on unverified estimates of livestock loss, and analyses in which $n \leq 4$ subjects (farms or livestock herds) completed the test. Several studies we mention in the footnote to Table 1 came close to the silver standard by calculating time lags in livestock loss following treatments but we omitted them because they failed to estimate change in livestock loss (after–before). We believe several of these might qualify if the data were reanalyzed.

Regarding criterion (2), we defined a working livestock operation as one in which livestock, land, and predators were managed in ways characteristic of a private livestock producer. That criterion excluded tests with captive predators (Jaeger 2004). We did not review qualifying tests from continents other than North America and Europe for two reasons. First, we excluded Australia because dingoes and red foxes are non-native species and their predation on livestock today may have been shaped by domestication and captivity, respectively, as a result of human-associated artificial selection for traits in these canids. Second, by excluding other continents, we avoided biased representation of tests published in languages that we (the authors) could not understand well enough to evaluate the research design. As WebPanel 1 and our descriptions below reveal, careful reading is necessary to understand research design.

Regarding criterion (3), we excluded studies measuring self-reported livestock losses or perceptions of effectiveness from Table 1. Although livestock owners' perceptions of interventions are important for the adoption of effective techniques, the functional effectiveness of candidate solutions should be established first. This exclusion reduced the number of allegedly effective non-lethal methods in particular.

Methods of analysis

Our quantitative summary was limited to counting tests in various categories. We did not attempt to perform a quantitative meta-analysis of effects, because there is no standard for consistent application of treatments and because the variety of methods used even within one category (eg different types of traps, or breeds of livestock-guarding dogs [LGDs]) would introduce uncontrollable variation. Furthermore, tests using the silver standard offer weaker inference than those using the gold standard but to an unknown degree.

We use the terms “culling” to refer to any variety of killing of wild predators by agents of the government and “hunting” to refer to regulated killing by private citizens.

Geographic and taxonomic distribution

The geographic distribution of tests in Europe and North America has been patchy, and the taxonomic distribution has concentrated on canids ($n = 7$), ursids ($n = 3$), and felids ($n = 2$) (Figure 1). The few tests involving wild felids and ursids do not suggest marked differences between taxonomic groups, as detailed below.

Results

Flawed tests

The earliest scientific studies had design flaws and a total of 12 tests (one published as recently as 2008) were excluded despite otherwise meeting our criteria (WebPanel 1); the earliest test with reliable inference was published in 1997 (Sagør *et al.* 1997). Our review of flawed tests revealed two important patterns. First, early investigations with design flaws have been cited

uncritically, even after flaws were identified in peer-reviewed, published articles (eg Mitchell *et al.* 2004). Second, seven tests of lethal methods and four (or five if one counts sterilization) tests of non-lethal methods had design flaws. All flawed tests were conducted in North America.

Lethal methods

Reliable inference was detected in only 7 tests of lethal methods that met the silver standard (those with quasi-experimental designs); tests of lethal methods that might have qualified for the gold standard were flawed (WebPanel 1). Of those 7 tests, only two were shown to reduce livestock losses from predation; in the remaining five, predation on livestock was unaffected in three tests and increased in two tests.

Using a quasi-experimental design to compare Eurasian lynx (*Lynx lynx*) predation on sheep across sites varying in the number of lynx killed over a 6-year period, Herfindal *et al.* (2005) reported that a lethal method (killing by various means) prevented sheep losses but only to a minor degree; prevention differed by site and its duration was short term. The test indicated that for every male and female lynx that were killed, 13 lambs and 2 lambs were saved, respectively. Because the range of each lynx encompassed multiple sheep flocks, the benefits to individual livestock owners averaged <1 lamb saved per lynx killed and were deemed to be “of little practical benefit” (Herfindal *et al.* 2005). Given that individual lynx differed substantially in their tendencies to prey on sheep, benefits were also geographically variable (Herfindal *et al.* 2005).

In three separate investigations of lethal control measures applied to bears, predation on livestock was unaffected or increased. For example, culling Norwegian brown bears (*Ursus arctos*) did not reduce predation on sheep (Sagør *et al.* 1997). Likewise, results from a study of American black bears (*Ursus americanus*) across Ontario, Canada, suggested that neither the number of black bears killed by hunters using various methods, nor bear population size, predicted future bear-related damage; rather, bear food availability was the best predictor (Obbard *et al.* 2014). A similar study in Wisconsin (Treves *et al.* 2010) did not include sufficient numbers of livestock losses among the incidents involving black bears for us to include in Table 1 but the results for agricultural damages of all sorts were similar when the data were reanalyzed as a BACI design.

Most quasi-experimental tests of lethal methods showed no effects or counterproductive effects on livestock loss. Slovenia's nationwide culling of 51 wolves, averaging 4.6 wolves or ~25% of the population annually, was distributed among local management units proportional to the current wolf densities. In an 11-year study in Slovenia, Krofel *et al.* (2011) detected no effect of wolf culling on subsequent livestock losses, even when

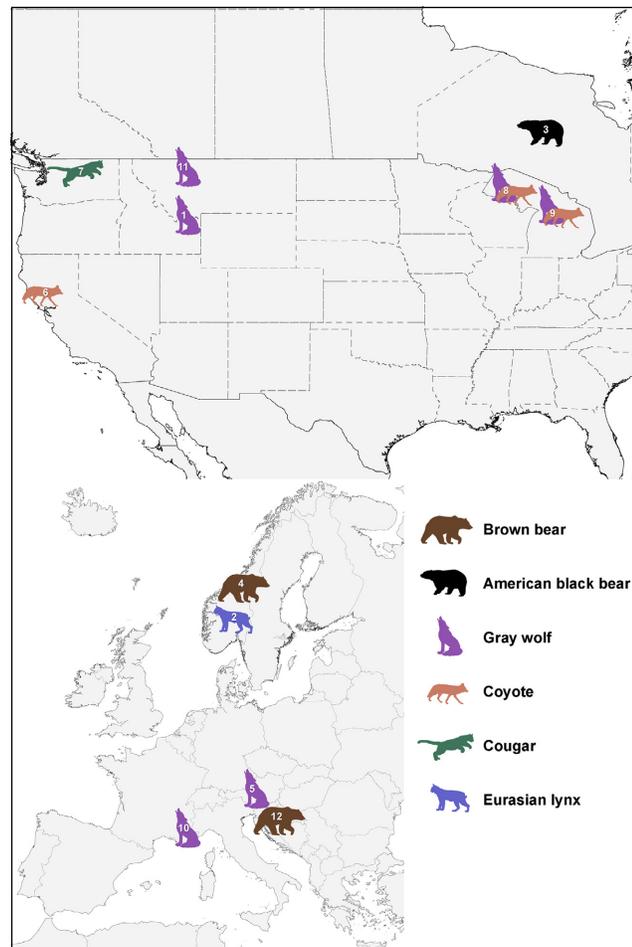


Figure 1. The geographic distribution of tests of lethal and non-lethal methods providing reliable inference about functional effectiveness in preventing carnivore predation on livestock from North American and European livestock farms. Numbers correspond to those in Table 1.

only the years with the most extreme killing rates were compared. The data for this test were reanalyzed in a BACI design to meet the silver standard (WebPanel 1). In 1998, researchers at the University of California's Hopland Research and Extension Center (HREC) investigated preventive effects of coyote killing (Conner *et al.* 1998) conducted using various methods (Figure 2). Conner *et al.* (1998) performed numerous analyses on the same data to test the effect of routine, non-selective coyote killing in preventing sheep predation. We focused only on those analyses that employed BACI designs (comparing lamb losses across consecutive seasons and those with time lags) and these reported counterproductive effects of killing more coyotes (Table 1).

Each of the quasi-experimental tests of lethal methods (Table 1) included unmeasured or uncontrolled variables, which may confound analyses and thereby weaken inference (see below for wolf culling and also WebPanel 1). However, one correlational study we would have excluded (Peebles *et al.* 2013) achieved silver standard when we



Figure 2. M-44 explosive poison (sodium cyanide) delivery device for killing wildlife (Factsheet May 2010, US Department of Agriculture Wildlife Services; <http://1.usa.gov/281v69N>).

considered it in combination with another study to qualify as a BACI design. Cooley *et al.*'s (2009 a,b) study of cougars (*Puma concolor*) strengthened the inference in Peebles *et al.* (2013) when looking only at the two-county comparison in the latter paper. Specifically, Cooley *et al.* (2009 a,b) documented that hunting cougars led to demographic changes in a heavily hunted county and not in another county with much lower rates of cougar hunting (also see White *et al.* 2011). Later, Peebles *et al.* (2013) showed that livestock losses rose annually in correlation with the number of cougars taken by hunters but only in the county that experienced changes in cougar demography. We therefore judged that the two studies together provided the causal mechanism and the BACI design needed to identify it as a silver standard test in Table 1.

Potential confounding variables in two wolf culling studies illustrate how weak inferences from tests using the silver standard impede scientific consensus. Two teams (Wielgus and Peebles 2014; Bradley *et al.* 2015) came to opposite conclusions when analyzing the same data from the northern Rocky Mountain (NRM) wolf population. Although Wielgus and Peebles (2014) found that killing more wolves was followed by more livestock losses during the following year, it did not adequately account for the time series underlying livestock exposure and lethal interventions. We therefore excluded it from Table 1, pending reanalysis. The time series is critical to BACI designs in the silver standard. Namely, as the wolf population increased in size, it also spread geographically, thereby exposing more livestock to wolf predation. Because wolf killing increased over time as recolonizing wolves left strictly protected areas and as policy changes introduced more and more wolf-killing (Bradley and Pletscher 2005; Bradley *et al.* 2008, 2015), one would therefore expect the predictors (wolf killing, livestock exposed, and wolf distribution) to rise over time. This

would create a positive correlation with the observed rise in livestock losses over time. Statistical control for encounters between wolves and livestock would require a measure of geographic spread of wolves, not just wolf and livestock abundance regionally (Wielgus and Peebles 2014). In contrast, Bradley *et al.* (2015) incorporated spatial information in their BACI design but limited their investigation in a critical way: restricting the spatial extent to pack territories.

Bradley *et al.* (2015) reported a reduction in livestock losses subsequent to culling within a wolf pack territory. The reductions were significant after an entire pack was killed, but insignificant when a few wolves were removed; when wolves were neither killed nor removed, no reduction in livestock losses was observed. The analysis was restricted to the affected wolf pack territory, despite the researchers' own work documenting how partial removal of wolves could scatter survivors beyond their original pack range or prompt take-over by a neighboring wolf pack (Bradley 2004; Bradley *et al.* 2008). The analysis should have examined neighboring areas and beyond, including ripple effects, whereby livestock losses recurred up to 16 km from sites of wolf culling (Treves *et al.* 2013). We recommend use of the gold standard for scientific inference to resolve the NRM wolf culling controversy. In sum, we find only weak inference for lethal methods and unconvincing evidence of preventive effects (Table 1).

Non-lethal methods

Non-lethal methods have long been examined but fewer of these studies met our criteria (five tests on six species; Table 1), because the measures of effect often came from livestock owners' perceptions rather than field verification. Of these five tests, four showed preventive effects; one test found preventive effects for wolves but not coyotes and one showed no effect. The latter – a BACI comparison in Slovenia that provided brown bears with livestock carcasses to deter or distract them from attacking sheep – revealed no change in livestock predation regionally (Kavčič *et al.* 2013). A large-scale, long-term study in France evaluated the effectiveness of 0–8 LGDs per pasture, and of mobile electric fences to confine sheep at night, against predation by wolves (Espuno *et al.* 2004). We include their study for the secondary analysis that tested sheep herds and pastures in relation to changes in the number of LGDs over time, not for their primary correlational model, which did not meet our criteria. From that secondary quasi-experimental test, Espuno *et al.* (2004) inferred that a combination of at least five LGDs and night enclosures (but neither in isolation) would prevent virtually all wolf predation on sheep (Figure 3a and b). In addition, two tests of non-lethal methods met the gold standard and showed preventive effects. One conducted on LGDs reported no livestock predation for control or treatment groups but detected an effect

of preventing carnivore incursions into fenced pastures (Gehring *et al.* 2010a, 2010b). We considered prevention of carnivore incursions into livestock pastures to be a relevant measure of effect because incursions are an essential precursor to predation on livestock. Likewise, the technique known as fladry (in which flagging is mounted on fences or ropes as a visual deterrent to predators; Figure 4) also demonstrated preventive effects, in the best random-assignment test that we found (Davidson-Nelson and Gehring 2010). A similar test of fladry used a BACI design (Musiani *et al.* 2003). Fladry was found to be effective against wolves but not coyotes or black bears in the former test and in another random-assignment experiment that we excluded because it did not involve livestock (Shivik *et al.* 2003).

Peer review

Rigorous peer review is a component of the gold standard for scientific inference, but we could not assess the rigor of review in the published tests. Three journals published 10 of the 12 (83%) articles with flawed designs (WebPanel 1), and only four of 12 (33%) tests that were reliable (Table 1). The same society publishes two of the journals, one of which also published a strong critique of several of the flawed tests (Mitchell *et al.* 2004). Yet the three journals continued to publish articles citing the flawed tests as evidence without citing Mitchell *et al.* (2004). Indeed, the latter paper appears to have been cited only once in any of those three journals (<http://bit.ly/28Joqto>, accessed 22 Jan 2016; Web of Science and Science Reports indicated no citations in these journals).

Conclusions

Effectiveness

Tests of effectiveness of interventions to prevent carnivore predation on livestock were consistent across regions. Among 12 North American and European tests that met “gold” or “silver” standards for reliable inference, we found a greater proportion of non-lethal methods were effective in preventing carnivore predation on livestock than lethal methods (80% versus 29%). Quasi-experimental tests of culling and hunting revealed positive, negative, and no effects (Table 1). None of the tests of lethal methods met the gold standard. Indeed, many combined several different methods of killing predators, including unregulated killing that would introduce uncontrolled variables. Culling and hunting appear risky for livestock owners because effects were slight or uncertain and five of seven tests produced no effect or a counterproductive effect (Table 1). This conclusion stands even without the inclusion of four studies that found counterproductive effects of killing wolves, bears, or cougars (Treves *et al.* 2010;



Figure 3. Livestock-guarding dog (a) protecting sheep and (b) charging the approaching photographer.

Peebles *et al.* 2013; Wielgus and Peebles 2014; Fernández-Gil *et al.* 2016). The non-lethal methods that have been tested (LGD, fladry, night enclosures) were not associated with similar negative results.

Two studies – one relying on LGDs (Gehring *et al.* 2010 a,b) and the other on fladry (Davidson-Nelson and Gehring 2010) – provide both strong inference and evidence of effectiveness in preventing predation on livestock. Although fladry may be limited to deterring wolves, LGDs have a long history and detailed technical information on appropriate breeds, husbandry, and deployment.

Our findings for selected sites in North America and Europe are consistent with tests conducted for Asiatic black bears (*Ursus thibetanus*) in Japan (Huygens *et al.* 2004), cougars in Mexico (Zarco-González and Monroy-Vilchis 2014), and canids and felids in South Africa (McManus *et al.* 2015). Using a pseudo-control, case-control design similar to BACI, the latter team found livestock losses and related costs declined for two consecutive years after implementing non-lethal methods (LGDs, alpacas [*Vicugna pacos*], and livestock protective steel collars) as compared with lethal methods (various kill-traps and shooting) in the first year of their study on the same livestock farms. Although the data on livestock losses were self-reported by livestock owners, the research-



Figure 4. An experimental plot containing a road-killed deer carcass surrounded by a treatment of fladry – a flagging method used to deter wolves (Shivik *et al.* 2003).

ers trained the owners in verification techniques and issued field kits to improve verification (McManus *et al.* 2015).

Strength of inference

We found few random-assignment experiments 50 years after its importance for strong inference was first explained (Platt 1964). Of the tests that met our inclusion criteria (Table 1), 83% were quasi-experimental tests using BACI comparisons. We considered that only two tests (17%) both allowed reliable inference and approached the gold standard for experimental design. Other studies were excluded from our quantitative summary because of small sample sizes or unreliable inference (WebPanel 1). The gap between recommended experimental designs (Platt 1964; Hairston 1989) may partly reflect the difficulty of randomizing treatments around working livestock operations. However, the above-mentioned examples of gold standard tests of non-lethal methods emphasize the importance of developing more robust experimental designs for the future.

We recommend an independent scientific panel of experts be convened to conduct a large-scale experiment on predator control, as was done in the UK for badger culling (Vial and Donnelly 2012). Indeed, we suggest that this experiment be subject to an even higher “platinum standard”, which would include “double blinds”, where those measuring effects are unaware of the treatment and where analysts compare results without knowing which data were from treatment or control groups (Mukherjee 2010).

Law, ethics, and ecological side effects

Sound policy should be consistent with law, scientific evidence, and ethical standards of society. The EU

Habitats Directive and various US federal policies and laws (including the Endangered Species Act) require the use of evidence and in some cases specify the best available science (Treves *et al.* 2015). When two or more interventions to control predators are lawful, we recommend that farmers, managers, policy makers, and courts first consider functional effectiveness (will the intervention prevent future threats to human interests?) and the strength of inference for that effect. If two candidate interventions perform equally by those criteria, then we recommend that two additional criteria be considered before implementing predator control: public acceptance (will the intervention be supported by both the complainants and the general public?) and ecological consequences (will the intervention deplete biodiversity or ecosystem services?). We recommend continuing education requirements for wildlife managers to keep up-to-date with the best available science. We also suggest that decision makers should suspend predator control programs that do not meet standards of strong inference about effectiveness, especially if those have legal, ethical, or ecological drawbacks. The burden of proof should rest heaviest on the interventions that have the most serious negative effects on biodiversity, people, and livestock.

Comparisons between non-lethal and lethal methods (such as culling and hunting) reveal how multiple criteria support the use of non-lethal methods. Livestock-guarding dogs and fladry outperformed lethal methods in functional effectiveness and were superior in strength of inference (Table 1). Lethal methods have additional limitations for managing predators and face a legal burden of proof in North America and Europe because of public trust principles or explicit protections (Epstein and Darpö 2013; Treves *et al.* 2015). The Habitats Directive 92/43/EEC, for example, restricts lethal controls to situations with an “absence of a satisfactory alternative” (Article 16, 2). Furthermore, recent court decisions in the US have restricted the use of predator control in several situations (Treves *et al.* 2015; <http://bit.ly/28J2mkq>). Ethical decisions should also consider the values of society at large and the intrinsic worth of all of the individual animals involved (Vucetich and Nelson 2014). For instance, two large-scale studies in the US suggested lower public acceptance of lethal methods than of non-lethal methods and that humanness was important to the public (Reiter *et al.* 1999; Slagle *et al.* in press). Finally, the negative ecological effects of killing carnivores have recently been documented in many regions (Ripple *et al.* 2014; Krofel *et al.* 2015). In many ecosystems, both terrestrial and aquatic, predators appear to play a disproportionate role not only in preventing excessive herbivory, which may result in long-term depletion of vegetation and its associated biodiversity, but also in enhancing species diversity. Regardless of whether predators directly regulate the numerical abundance of their prey or indirectly keep the

survivors fearful, human-induced mortality, translocation, or sterilization methods for predator control may alter predator ecology and ecosystem dynamics with far-reaching effects.

In conclusion, we believe the science of predator control lacks rigor generally – the resulting uncertainty about the functional effectiveness of killing predators should guide evidence-based policy to non-lethal methods until gold standard tests are completed.

■ Acknowledgements

Financial support for research was provided by the Fulbright Program to AT, the Slovenian Research Agency (grant number P4-0059) and the Pahernik Foundation to MK, the Carnegie Corporation of New York to the Global Change and Sustainability Research Institute, University of the Witwatersrand, the ABAX Foundation, and the National Lotteries Distribution Trust Fund (to JM). JV Lopez-Bao and G Chapron provided invaluable feedback.

■ References

- Allen LR and Sparkes EC. 2001. The effect of dingo control on sheep and beef cattle in Queensland. *J Appl Ecol* 38: 76–87.
- Biondi F. 2014. Paleocology grand challenge. *Front Ecol Evol*; doi:10.3389/fevo.2014.00050.
- Bradley EH. 2004. An evaluation of wolf–livestock conflicts and management in the northwestern United States (MS thesis). Missoula, MT: University of Montana.
- Bradley EH and Pletscher DH. 2005. Assessing factors related to wolf depredation of cattle in fenced pastures in Montana and Idaho. *Wildlife Soc B* 33: 1256–65.
- Bradley EH, Pletscher DH, Bangs EE, et al. 2008. Evaluating wolf translocation as a nonlethal method to reduce livestock conflicts in the northwestern United States. *Conserv Biol* 19: 1498–508.
- Bradley EH, Robinson HS, Bangs EE, et al. 2015. Effects of wolf removal on livestock depredation recurrence and wolf recovery in Montana, Idaho, and Wyoming. *J Wildlife Manage* 79: 1337–46.
- Chapron G, Kaczensky P, Linnell JDC, et al. 2014. Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* 346: 1517.
- Conner MM, Jaeger MM, Weller TJ, et al. 1998. Effect of coyote removal on sheep depredation in northern California. *J Wildlife Manage* 62: 690–99.
- Cooley HS, Wielgus RB, Koehler GM, et al. 2009a. Source populations in carnivore management: cougar demography and emigration in a lightly hunted population. *Anim Conserv* 12: 321–28.
- Cooley HS, Wielgus RB, Robinson HS, et al. 2009b. Does hunting regulate cougar populations? A test of the compensatory mortality hypothesis. *Ecology* 90: 2913–21.
- Davidson-Nelson SJ and Gehring TM. 2010. Testing fladry as a nonlethal management tool for wolves and coyotes in Michigan. *Human-Wildlife Interactions* 4: 87–94.
- Epstein Y and Darpö J. 2013. The wild has no words: environmental NGOs empowered to speak for protected species as Swedish courts apply EU and international environmental law. *J European Environ Planning Law* 10: 250–61.
- Espuno N, Lequette B, Pouille ML, et al. 2004. Heterogeneous response to preventive sheep husbandry during wolf recolonization of the French Alps. *Wildlife Soc B* 32: 1195–208.
- Estes JA, Terborgh J, Brashares JS, et al. 2011. Trophic downgrading of planet Earth. *Science* 333: 301–06.
- Fernández-Gil A, Naves J, Ordiz A, et al. 2016. Conflict misleads large carnivore management and conservation: brown bears and wolves in Spain. *PLoS ONE* 11: e0151541.
- Fischer J, Brosi B, Daily GC, et al. 2008. Should agricultural policies encourage land sparing or wildlife-friendly farming? *Front Ecol Environ* 6: 380–85.
- Gehring TM, VerCauteren KC, and Cellar AC. 2010a. Good fences make good neighbors: implementation of electric fencing for establishing effective livestock protection dogs. *Human-Wildlife Interactions* 4: 144–49.
- Gehring TM, VerCauteren KC, Provost ML, et al. 2010b. Utility of livestock-protection dogs for deterring wildlife from cattle farms. *Wildlife Res* 37: 715–21.
- Gould SJ. 1980. The promise of paleobiology as a nomothetic, evolutionary discipline. *Paleobiology* 6: 96–118.
- Greentree C, Saunders G, McLeod L, et al. 2000. Lamb predation and fox control in south-eastern Australia. *J Appl Ecol* 37: 935–43.
- Haber GC. 1996. Biological, conservation, and ethical implications of exploiting and controlling wolves. *Conserv Biol* 10: 1068–81.
- Hairton NG. 1989. Ecological experiments: purpose, design and execution. Cambridge, UK: Cambridge University Press.
- Herfindal I, Linnell JDC, Moa PF, et al. 2005. Does recreational hunting of lynx reduce depredation losses of domestic sheep? *J Wildlife Manage* 69: 1034–42.
- Huygens OC, van Manen FT, Martorello DA, et al. 2004. Relationships between Asiatic black bear kills and depredation costs in Nagano Prefecture, Japan. *Ursus* 15: 197–202.
- Jaeger MM. 2004. Selective targeting of alpha coyotes to stop sheep depredation. *Sheep Goat Res J* 19: 80–84.
- Kavčič I, Adamič M, Kaczensky P, et al. 2013. Supplemental feeding with carrion is not reducing brown bear depredations on sheep in Slovenia. *Ursus* 24: 111–19.
- Knowlton FF, Gese EM, and Jaeger MM. 1999. Coyote depredation control: an interface between biology and management. *J Range Manage* 52: 398–412.
- Krofel M, Cerne R, and Jerina K. 2011. Effectiveness of wolf (*Canis lupus*) culling as a measure to reduce livestock depredations. *Acta Silvae et Ligni* 95: 11–22.
- Krofel M, Treves A, Ripple WJ, et al. 2015. Hunted carnivores at outsized risk. *Science* 350: 518.
- Linnell JDC, Aanes R, Swenson JE, et al. 1997. Translocation of carnivores as a method for managing problem animals: a review. *Biodivers Conserv* 6: 1245–57.
- Mason JR, Shivik JA, and Fall MW. 2001. Chemical repellents and other aversive strategies in predation management. *Endangered Species Update* 18: 175–81.
- McManus J, Dickman A, Gaynor D, et al. 2015. Dead or alive? Comparing costs and benefits of lethal and non-lethal human–wildlife conflict mitigation on livestock farms. *Oryx* 49: 687–95.
- Mitchell BR, Jaeger MM, and Barrett RH. 2004. Coyote depredation management: current methods and research needs. *Wildlife Soc B* 32: 1209–18.
- Mukherjee S. 2010. The emperor of all maladies: a biography of cancer. New York, NY: Scribner.
- Musiani M, Mamo C, Boitani L, et al. 2003. Wolf depredation trends and the use of fladry barriers to protect livestock in western North America. *Conserv Biol* 17: 1538–47.
- Newby F and Brown R. 1958. A new approach to predator management in Montana. *Montana Wildlife* 8: 22–27.
- Obbard ME, Howe EJ, Wall LL, et al. 2014. Relationships among food availability, harvest, and human–bear conflict at landscape scales in Ontario, Canada. *Ursus* 25: 98–110.
- Peebles K, Wielgus RB, Maletzke BT, et al. 2013. Effects of remedial sport hunting on cougar complaints and livestock depredations. *PLoS ONE* 8: e79713.

- Platt JR. 1964. Strong inference. *Science* 146: 347–53.
- Prugh LR, Stoner CJ, Epps CW, *et al.* 2009. The rise of the meso-predator. *BioScience* 59: 779–91.
- Reiter DK, Brunson MW, and Schmidt RH. 1999. Public attitudes toward wildlife damage management and policy. *Wildlife Soc B* 27: 746–58.
- Ripple WJ, Estes JA, Beschta RL, *et al.* 2014. Status and ecological effects of the world's largest carnivores. *Science* 343; doi:10.1126/science.1241484.
- Sagør JT, Swenson JE, and Roskaft E. 1997. Compatibility of brown bear *Ursus arctos* and free-ranging sheep in Norway. *Biol Conserv* 81: 91–95.
- Shivik JA, Treves A, and Callahan M. 2003. Non-lethal techniques: primary and secondary repellents for managing predation. *Conserv Biol* 17: 1531–37.
- Slagle KM, Bruskotter JT, Singh AS, *et al.* Attitudes toward predator control in the United States: 1995 and 2014. *J Mammal*. In press.
- Treves A, Chapron G, López-Bao JV, *et al.* 2015. Predators and the public trust. *Biol Rev*; doi:10.1111/brv.12227.
- Treves A, Jurewicz RL, Naughton-Treves L, *et al.* 2002. Wolf depredation on domestic animals: control and compensation in Wisconsin, 1976–2000. *Wildlife Soc B* 30: 231–41.
- Treves A, Kapp KJ, and Macfarland DM. 2010. American black bear nuisance complaints and hunter take. *Ursus* 21: 30–42.
- Treves A and Karanth KU. 2003. Human–carnivore conflict and perspectives on carnivore management worldwide. *Conserv Biol* 17: 1491–99.
- Treves A and Naughton-Treves L. 2005. Evaluating lethal control in the management of human–wildlife conflict. In: Woodroffe R, Thirgood S, and Rabinowitz A (Eds). *People and wildlife, conflict or coexistence?* Cambridge, UK: Cambridge University Press.
- Treves A, Vucetich JA, Rabenhorst M, *et al.* 2013. An evaluation of localized wolf control efforts to prevent subsequent livestock depredation in Michigan. Natural Resources Report No 2013–14: Little River Band of Ottawa Indians.
- Treves A, Wallace RB, and White S. 2009. Participatory planning of interventions to mitigate human–wildlife conflicts. *Conserv Biol* 23: 1577–87.
- Vial F and Donnelly C. 2012. Localized reactive badger culling increases risk of bovine tuberculosis in nearby cattle herds. *Biol Lett* 8: 50–53.
- Vucetich J and Nelson MP. 2014. Wolf hunting and the ethics of predator control. Political science, comparative politics, political theory. Oxford, UK: Oxford Handbooks Online.
- White KS, Koehler GM, Maletzke BT, *et al.* 2011. Differential prey use by male and female cougars in Washington. *J Wildlife Manage* 75: 1115–20.
- Wielgus RB and Peebles K. 2014. Effects of wolf mortality on live-stock depredations. *PLoS ONE* 9: e113505.
- Wirsing AJ and Ripple WJ. 2010. A comparison of shark and wolf research reveals similar behavioral responses by prey. *Front Ecol Environ* 9: 335–41.
- Woodroffe R and Redpath S. 2015. When the hunter becomes the hunted. *Science* 348: 1312–14.
- Young SP and Dobyas HW. 1945. *Coyote control by means of den hunting*. Washington, DC: US Department of the Interior, Fish and Wildlife Service.
- Zarco-González MM and Monroy-Vilchis O. 2014. Effectiveness of low-cost deterrents in decreasing livestock predation by felids: a case in central Mexico. *Anim Conserv* 17: 371–78.

■ Supporting Information

Additional, web-only material may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/fee.1312/supinfo>

³Research Department, Landmark Foundation, Riversdale, South Africa, and School of Animal, Plants and Environmental Sciences, University of Witwatersrand, Johannesburg, South Africa; † these authors contributed equally to this work



The National Socio-Environmental Synthesis Center

Postdoctoral Fellowship Program

The National Socio-Environmental Synthesis Center (SESYNC), located in Annapolis, Maryland, invites applications from early career scholars (≤ 4 years post PhD) for 2-year postdoctoral fellowships that begin August 2017. Fellows will co-develop research projects with Collaborating Mentors using synthesis methods to address a problem arising from, or associated with, the relationship between humans and nature. Fellows will also participate in the Socio-Environmental Immersion Program, which includes a learning component that will steep participants in theory foundational to understanding socio-environmental systems. For more information, please visit: <http://sesync.us/postdocim>.

In collaboration with the Long-Term Ecological Research (LTER) network, SESYNC also invites applications that propose to use long term data sets, ongoing experiments, or modeling results from LTER sites. For more information about the SESYNC-LTER Postdoctoral Fellowship, please visit: <http://sesync.us/postdoclter>.

Pre-screening applications due October 24, 2016, but are accepted on a rolling basis and should be submitted as soon as possible.

Full applications are due December 5, 2016.

A Treves *et al.* – Supporting information

WebPanel 1. Tests we excluded because the design precluded reliable inference

We use the scientific terms “bias” and “flawed design” without any suggestion of intention or incompetence. Indeed, the flaws we discuss often result from inescapable challenges of running experiments under complex field conditions over many months or years. Several tests were excluded because they were not peer-reviewed, published descriptions of all methods and results. Not all tests conducted at the Hopland Research and Extension Center (HREC) were peer reviewed (HREC 2003), including proceedings of conferences that do not publish the editorial policy on anonymous peer review (eg Proceedings of the Vertebrate Pest Conference; <http://bit.ly/1UycGeA>). Below and in the quotations that follow, we inserted square brackets to identify biases, which we discuss after the quote.

Tests of lethal methods that had flawed designs that precluded strong inference

We found seven tests published since 1978 – each described in the following paragraphs – that evaluated the effectiveness of lethal methods of predator control that fit our criteria but had biases in design that precluded strong inference.

Guthery and Beasom (1978) reported a 17% and 0% decline in predation on goat kids and nanny-goats, respectively, after comparing an untreated pasture to a pasture treated with intense mechanical and explosive trapping (Figure 2), snaring, and shooting. After the test, the authors discovered a decline in native prey species that was twice as large at the untreated pasture as compared with the treated pasture, which unfortunately produced selection bias. Furthermore, it was unclear whether the two pastures received the same level of human stimuli (visits, noise, material left behind, etc), leading to a possible treatment bias.

A test of two lethal methods by Till and Knowlton (1983) came closest to the gold standard in our view. Their experimental design had great potential but because the description of methods and results were flawed, we recommend replicating the test with state-of-the-art reporting. The authors recorded sheep losses 7 days before and 7 days after two treatments and a control. The treatments consisted of one or more technicians back-tracking coyotes to their dens and then either killing only the pups by fumigation (treatment 1) or killing the pups and adults by fumigation and shooting, respectively (treatment 2). In describing this method, Till and Knowlton (1983) cited a manual (Young and Dobyas 1945) that included several alternative treatments. However, Till and Knowlton (1983) did not provide sufficient detail; for instance, back-tracking coyotes is an expert skill but the authors failed to adequately describe who performed the back-tracking, what training they had received, and what actions were taken in various scenarios (eg if technicians lost the coyote trail to the den, if a den was unoccupied, or if adult coyotes did not return to the den). We also noted a discrepancy in their results that confirms that key aspects of the methods were not described: 30 dens should have been destroyed but Till and Knowlton (1983) reported that 40 were destroyed. Till and Knowlton (1983) did not clarify whether dogs were used, as the 1945 methods paper suggested. The study is therefore impossible

to replicate, due in part to these shortcomings, and because the control may have been a pseudo-control (eg no den found), because of ambiguous references to “other control methods” being in place, and because the treatments may have differed from the control in multiple ways. We also detected several types of reporting bias. First, the interval after treatments (7 days) may be too short to represent an unbiased comparison of treatments. The authors’ conclusion – that, during the 7-day period, greater losses of sheep were associated with controls than with treatments – may be correct but was confounded by unclear statistics and by insufficient details, not only regarding the number and reproductive status of adults killed but also on observed pack sizes (hence the numbers of survivors). Finally, although Till and Knowlton (1983) relied on sheep herders to report lamb kills, their paper did not include data on overall losses or losses to other causes. Without data on sheep flock sizes before and after, one cannot determine if disappearances increased or other causes of death confound the results.

O’Gara *et al.* (1983) compared coyote killing on a single property by conducting a before-and-after test, where the researchers initially habituated coyotes for observation and later switched to various lethal methods when funds ran out. A lack of replication and the presence of a pseudo-control (manipulations during the “before” period differed in more than the lethal intervention) preclude reliable inference.

Using a before-and-after test in Alberta, Canada, Bjorge and Gunson (1985) compared 2 years of use of strychnine-laced baits to 4 years pre-poison. Over the course of 2 years, 26 out of 40 wolves were poisoned by researchers and an additional 11 wolves left the study area or died from other causes (in total 93% mortality), resulting in a decline in wolf predation on cattle from 0.7% to 0.3%. The authors warned of non-target mortality (29 non-target animals representing five species also died) and the potential movement of livestock predators when surviving wolves dispersed (Bjorge and Gunson 1985). However, the first two pre-treatment years showed losses that were statistically equivalent to the two treatment years, and the third and fourth pre-treatment years experienced an important change in management, leading to lower cattle density, associated with substantially higher levels of predation in the 2 years before treatment. We are unsure if the control was appropriate or whether it represented a pseudo-control, given that the sample size was one study site over 6 years. Indeed, it is not clear how one should measure the effect of the treatment to avoid pseudo-replication (number of cattle lost, number lost per wolf present, percent of cattle, or years with higher than median losses). Depending on which measure was used, the effect might have been an increase, a decrease, or no change in wolf predation on cattle. Therefore, we find the test inadequate to support reliable inference.

Wagner and Conover (1999) treated several mountain pastures during summer months with mechanical and explosive trapping (Figure 2), snaring, and shooting; subsequently, flocks on those pastures experienced 7.3% verified predation by coyotes. These losses were compared to sheep flocks on another set of pastures treated with those same lethal methods as well as with aerial gunning during winter. The authors claimed a decline to 2.7% losses. The study had five design flaws, some of which were noted by Mitchell *et al.* (2004): (1) control pastures started with 40% higher sheep densities, which has been shown to increase vulnerability to predation by

North American canids (Robel *et al.* 1981; Mech *et al.* 2000; Wydeven *et al.* 2004) and implies a treatment bias; (2) pre-treatment sheep losses were 186% higher in untreated than treated pastures, suggesting selection bias; (3) untreated pastures were subject to twice the lethal effort (excluding aerial-gunning), again suggesting treatment bias; (4) livestock-guarding dogs (LGDs) were apparently matched between treated and untreated pastures but those data were not presented, implying reporting bias; and (5) the authors made an unsupported assumption in their analyses that the ratio of known to unknown losses was constant across treatments and years (measurement bias).

Blejwas *et al.* (2002) tested poison-filled collars on sheep at the HREC. Note the word “control” referred to killing coyotes and other wildlife, not experimental treatments, in the quote that follows:

“Coyote Control. The HREC employed three different control strategies during the course of the study: no control, nonselective control, and selective control. . . . During the no-control periods, animals on the periphery of HREC were still subject to control on adjacent ranches. During nonselective control, the local Wildlife Services specialist attempted to remove as many coyotes as possible from HREC [pseudo-control]. These activities were carried out independently of the ongoing coyote research and without benefit of radiotelemetry locations. During selective control, HREC personnel used [Livestock Protection Collars, LPC] to target depredating coyotes. Once a pattern of coyote predation was established [treatment bias 1], all sheep were removed from the pasture except for a small target flock of 10–30 lambs or yearlings with LPC [treatment bias 2]. Collared lambs were accompanied by uncollared ewes. [An LPC] consists of a pair of toxicant-filled rubber bladders attached to a Velcro collar and placed around the neck of a lamb or small ewe. . . . in some cases, use of the LPC was impractical or unsuccessful, and HREC or Wildlife Services personnel used radiotelemetry to remove these depredating breeders by shooting [treatment bias 3 and reporting bias]” (square brackets added; Blejwas *et al.* 2002).

First, non-selective coyote killing during experiments represents a pseudo-control – allowing only an inference about the addition of LPC to an unmeasured, background level of nonselective coyote killing. The first treatment bias arose from the timing of intervention: “once a pattern of coyote predation was established” (which was undefined). Thus, treated flocks were neither randomly assigned nor selected haphazardly (independent of outcomes), but rather selected based on past vulnerability. In biomedical clinical trials, that step would be analogous to treating patients only when disease symptoms had appeared – and it was not clear how control flocks were managed when a pattern of coyote predation was established. The second treatment bias compounded the latter issue because the vulnerable sheep flock was replaced with a treated one, thereby conflating vulnerability, treatment, and a massive manipulation of the flock. True experimental controls and non-LPC periods should have also had simultaneous flock replacement with lambs wearing dummy collars lacking poison. Finally, the decision to add coyote shooting when LPC was impractical or unsuccessful was the third treatment bias. Because

the latter step was neither quantified nor fully explained, we also find reporting bias.

In a Minnesota study, Harper *et al.* (2008) analyzed the effects on livestock predation in three scenarios: when traps were set and wolves were trapped, when traps were set and no wolves were trapped, and when no traps were set; the authors concluded that the effects of removing wolves by trapping did not differ from trapping without removing wolves. The authors reported exceptions for small effects on sheep farms and when male wolves were removed. However, the test represents a pseudo-control because decisions whether or not to set traps apparently reflected subjective judgments by government trappers, implying possible treatment bias. Also, the authors discarded data points for numerous reasons without citing evidence or by justifying the removal of data post hoc based on results, implying measurement and reporting biases. For example, they excluded farms where trapping was unsuccessful but where dispersing wolves might have been present, which the authors did to “decrease apparent effectiveness of unsuccessful trapping” (Harper *et al.* 2008). Given the Minnesota wolf population size exceeded 1000 individuals, and the very small proportion of marked wolves (www.dnr.state.mn.us/mammals/wolves/mgmt.html), the guesswork required to make such judgments implies possible measurement bias.

We could not draw reliable inference from three or four tests of non-lethal predator control methods (if one counts sterilization as non-lethal).

Tests of sterilization

Bromley and Gese (2009) conducted a well-designed random-assignment experiment to capture what they believed were entire packs of coyotes and surgically sterilize some or conduct sham treatments that were identical except for sterilization. However, we identified a measurement and a reporting bias in this study, which precluded strong inference. First, the position, size, and overlap of territories of the treated, control, and uncaptured packs were potentially important confounding variables. The authors were transparent about the uncaptured coyotes when writing, “In 4 packs, no members were captured or radiocollared, but pack members were observed and the home range boundary was estimated based on the spatial arrangement of adjacent radiocollared packs...many [sheep] kills were located in areas of overlap between territories” (Bromley and Gese 2009). Across both years of the study, the authors reported six sheep kills in core pack areas and 20 on the edge of territories. In 1999 (the year with the best radio-telemetry data), sheep kills were significantly disproportionately on the edge of territories, when accounting for sheep distributions. Therefore, assignment of a sheep kill to a particular coyote pack must have included some uncertainty. Furthermore, that uncertainty was not a random effect because subsequent work showed that the home ranges and core areas of sterilized coyote packs overlapped territories of neighbors significantly more than those of intact coyote packs during the breeding season, when virtually all sheep predation occurred (Seidler and Gese 2012). Thus, assigning sheep kills to a certain pack may have introduced measurement bias to a majority of sheep kills on the edge of territories. Error in classifying even a single sheep kill might alter their results, as evidenced by the slight difference between treatment and control:

“weekly survival rate tended to be higher for sheep in sterile coyote territories (mean = 0.998) than in intact coyote territories (mean = 0.989)” (Bromley and Gese 2009). The authors presented no justification regarding why such a small difference in weekly survival rate was biologically significant, or exceeded the measurement error given uncertainty in assignments described above. Nor did the authors justify why weekly survival was better than other measures. For example, the authors did not emphasize in the abstract or conclusions that they found a counter-productive effect. Namely, they reported that 5 of 9 (56%) sterile packs and 9 of 14 (64%) intact packs were not assigned as having killed sheep. We conclude that strong inference cannot be drawn in either direction, despite the excellent random-assignment of treatment in this study.

Tests of non-lethal methods that had flawed designs that preclude strong inference

We excluded a substantial number of studies of non-lethal methods because they relied on livestock owners to report losses without providing training in verification (Coppinger *et al.* 1988; Meadows and Knowlton 2000). Three additional tests met our criteria except for flaws in research design. Bourne and Dorrance (1982) tested baits laced with an aversive chemical (lithium chloride, LiCl) to deter coyotes and other animals from sheep. This study seemed to present reporting bias: “It seemed improbable that the LiCl baits affected predation in southwestern Alberta because the rate of bait disappearance was so low. Therefore data from the 8 farms in southwestern Alberta were excluded from subsequent analyses of bait disappearance and predation losses” (Bourne and Dorrance 1982). A greater concern was raised by apparent selection bias: “flock size differed markedly between farms treated with placebo and LiCl baits [placebos averaged 123 lambs, LiCl averaged 231 lambs]” (Bourne and Dorrance 1982). Finally, the authors apparently used a pseudo-control that hinders interpretation of the results because lethal controls were implemented throughout the study until depredations stopped, on both treatment and control farms.

Between 1979 and 1992, Linhart *et al.* tested several non-lethal methods. Some of these tests were not peer reviewed and thus did not meet our criteria for inclusion; other tests met our criteria but were flawed. For example, Linhart *et al.* (1984, 1992) tested sound and light devices to prevent coyote predation on sheep. We agree with Mitchell *et al.*'s (2004) reasoning that the BACI design Linhart *et al.* used may have triggered a measurement bias by comparing early losses without treatment to late losses with treatment, within the same year. As time passes, lambs may outgrow the size most coyotes would attack and coyote pups may no longer need the provisioning that seems to prompt alpha breeders to prey on sheep (Knowlton *et al.* 1999). Also Linhart *et al.* (1979) summarized several tests of LGDs on sheep in working farms but relied on various methods that we view as having one or more of the following flaws: pseudo-control, before-and-after comparison with the above-mentioned measurement bias in the timing of comparisons, or small sample size.

Finally, Palmer *et al.* (2010) tested the effects of sheep herders quasi-experimentally. We could not draw strong inference: (a) lethal methods were ongoing in the background against coyotes and cougars; (b) the control (no herder) and treatment (herder or herder and dog)

selected by the owners and treatment flocks were larger than control flocks; (c) bands or flocks of sheep which were the subunits of herds varied in treatment within the same herds, but the analyses were conducted at the level of herds; and (d) although the researchers attended carefully to scavengers (Palmer *et al.* 2010), the quantitative effect of scavengers in relation to different treatments was not adequately described.

Reanalysis

In the main text, we argued that several studies might qualify as “silver” standard tests by our criteria if they re-analyzed data using a BACI design; namely estimating livestock losses minus losses before the treatment. These include those studies listed in a footnote to Table 1. We conducted such a re-analysis of the data presented in Figure 1 in Krofel *et al.* (2011) to illustrate the point. When we recalculated livestock losses each year as a net change in livestock losses over 2 years, we found no effect of wolf culling and hunting, as in Table 1 (Spearman $\rho=0.47$, $p=0.09$. Indeed, there was a trend toward a counterproductive effect that killing more wolves led to more livestock losses the following year).

WebReferences

- Bjorge RR and Gunson JR. 1985. Evaluation of wolf control to reduce cattle predation in Alberta. *J Range Manage* **38**: 483–86.
- Blejwas KM, Sacks BN, Jaeger MM, *et al.* 2002. The effectiveness of selective removal of breeding coyotes in reducing sheep predation. *J Wildl Manage* **66**: 451–62.
- Bourne J and Dorrance MJ. 1982. A field test of lithium chloride aversion to reduce coyote predation on domestic sheep. *J Wildl Manage* **46**: 235–39.
- Bromley C and Gese EM. 2009. Surgical sterilization as a method of reducing coyote predation on domestic sheep *J Wildl Manage* **65**: 510–19.
- Coppinger R, Coppinger L, Langeloh G, *et al.* 1988. A decade of use of livestock guarding dogs. *Proc Vertebrate Pest Conf* **13**: 209–14.
- Guthery FS and Beasom SL. 1978. Effects of predator control on Angora goat survival in South Texas. *J Range Manage* **31**: 168–73.
- Harper EK, Paul WJ, Mech DL, *et al.* 2008. Effectiveness of lethal, directed wolf-depredation control in Minnesota. *J Wildl Manage* **72**: 778–84.
- HREC (Hoplend Research and Extension Center). 2003. Annotated bibliography 1951–2001. Hopland, CA: HREC.
- Knowlton FF, Gese EM, and Jaeger MM. 1999. Coyote depredation control: an interface between biology and management. *J Range Manage* **52**: 398–412.
- Krofel M, Cerne R, and Jerina K. 2011. Effectiveness of wolf (*Canis lupus*) culling as a measure to reduce livestock depredations. *Zbornik Gozdarstva Lesarstva* **95**: 11–22.
- Linhart SB, Dasch GJ, Johnson RR, *et al.* 1992. Electronic frightening devices for reducing coyote depredation on domestic sheep: efficacy under range conditions and operational use. *Proc Vertebrate Pest Conf* **15**: 386–92.

- Linhart SB, Sterner RT, Carrigan TC, *et al.* 1979. Komondor guard dogs reduce sheep losses to coyotes: a preliminary evaluation. *J Range Manage* **32**: 238–41.
- Linhart SB, Sterner RT, Dasch GJ, *et al.* 1984. Efficacy of light and sound stimuli for reducing coyote predation upon pastured sheep. *Protection Ecol* **6**: 75–84.
- Meadows LE and Knowlton FF. 2000. Efficacy of guard llamas to reduce canine predation on domestic sheep. *Wildl Soc Bull* **28**: 614–22.
- Mech LD, Harper EK, Meier TJ, *et al.* 2000. Assessing factors that may predispose Minnesota farms to wolf depredations on cattle. *Wildl Soc B* **28**: 623–29.
- Mitchell BR, Jaeger MM, and Barrett RH. 2004. Coyote depredation management: current methods and research needs. *Wildlife Soc B* **32**: 1209–18.
- O’Gara BW, Brawley KC, Munoz JR, *et al.* 1983. Predation on domestic sheep on a western Montana ranch. *Wildl Soc B* **11**: 253–64.
- Palmer BC, Conover MR, and Frey SN. 2010. Replication of a 1970s study on domestic sheep losses to predators on Utah's summer rangelands. *Rangeland Ecol Manage* **63**: 689–95.
- Robel RJ, Dayton AD, Henderson FR, *et al.* 1981. Relationship between husbandry methods and sheep losses to canine predators. *J Wildl Manage* **45**: 894–911.
- Seidler RG and Gese EM. 2012. Territory fidelity, space use, and survival rates of wild coyotes following surgical sterilization. *J Ethol* **30**: 345–54.
- Till JA and Knowlton FF. 1983. Efficacy of denning in alleviating coyote depredations upon domestic sheep. *J Wildl Manage* **47**: 1018–25.
- Wagner KK and Conover MR. 1999. Effect of preventive coyote hunting on sheep losses to coyote predation. *J Wildl Manage* **63**: 600–12.
- Wydeven AP, Treves A, Brost B, *et al.* 2004. Characteristics of wolf packs in Wisconsin: identification of traits influencing depredation. In: Fascione N, Delach A, and Smith ME (Eds). *People and predators: from conflict to coexistence*. Washington, DC: Island Press.
- Young SP and Dobyas HW. 1945. *Coyote control by means of den hunting*. Washington, DC: US Department of the Interior, US Fish and Wildlife Service.