Received: 2 November 2022 | Revised: 9 March 2023 | Accepted: 18 March 2023

DOI: 10.1002/jwmg.22412

COMMENTARY

THE JOURNAL OF WILDLIFE MANAGEMENT

THE WILDLIFE SOCIETY

# Causality and wildlife management

**Jim Hone**[1] | **Charles J. Krebs**[1,2]

[1]Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601, Australia

[2]Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

**Correspondence**
Jim Hone, Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601 Australia.
Email: jim.hone@canberra.edu.au

## Abstract

Establishing cause and effect (i.e., causality) is a fundamental aim in science and important for wildlife management, as we need to know the cause of an event if we seek to reproduce, enhance, or diminish it. We review 13 alternative approaches for applying 12 criteria to establish causality. Strength of causal inference is greater when more causal criteria are applied so we propose a scaffolding set of criteria to clearly establish causality. We recommend validating predicted outcomes of wildlife management efforts when predictions are based on a unique mechanism and temporality, especially when manipulative experimental studies are not feasible. We use 3 case studies, namely of lamb predation by feral pigs (*Sus scrofa*), causes of trends in northern spotted owls (*Strix occidentalis caurina*), and causes of trends in mallards (*Anas platyrhynchos*), to illustrate these topics, which are of wide relevance in wildlife management. We recommend greater use of causality and relative strength of causal inference to improve adaptive wildlife management.

**KEYWORDS**

applied ecology, bias, causal criteria, experiments, hypotheses, precision, predictions, strength of causal inference, wildlife management

A fundamental aim of science is to establish that occurrence of an event is caused by another event, and is known variously as cause and effect or causality (Hill 1965, Harre 1972, Ford 2000). How evidence of causality is perceived and used, was recognized decades ago by Aldo Leopold when writing about wildlife management: "We are… trying to … qualify our minds to comprehend the meaning of evidence" (Leopold 1933:231, Anderson 2008:2).

Causal inference uses logic and evidence to determine whether a proposed cause has the observed effect, and statistical inference uses parameter estimation and inferences from a random sample to a population (Williams 1997, Cox 2006). Statistical inference, in regression, focuses on bias and precision of the outcome (Y) and causal inference on the change in outcome (ΔY) given a change in management (ΔX). For example, statistical inference estimates the number of livestock killed (Y) by a specified abundance of predators (X), whereas causal inference estimates the change in livestock kills (ΔY) for a given change in abundance of predators (ΔX). Also, evidence for causality consists of showing that a change in management (ΔX), such as trapping of predators, leads to a change in outcomes (ΔY), such as a change in the number of livestock killed. Prediction estimates what outcome (Y) is predicted from management (X). Causal predictions combine these to ask what the outcome (Y) and the change in outcome (ΔY) are given the change in management (ΔX). These are similar, but not identical, to the expression of Arif and MacNeil (2022) who described causal inference as asking, what is the effect of X on Y?

Wildlife management resources (e.g., money, time, personnel) are limited, so there is a strong incentive to predict outcomes accurately, and to avoid taking actions that have negative effects. The effort-outcome principle in wildlife management was stated as a cause and effect relationship between management efforts (inputs or cause) and outcomes (effects; Hone et al. 2017, 2023). Conservation scientists use the declining population paradigm (Caughley 1994) to identify the cause(s) of population decline and reduce them to reverse the decline. Wildlife scientists and managers would benefit from more reliable knowledge to aid these efforts (Romesburg 1981, Sells et al. 2018). Causal evidence or inference was not included in indicator criteria for evidence in a review of some wildlife management programs in North America (Artelle et al. 2018). We recommend that management programs include cause and effect relationships as an additional indicator criterion.

Medicine, epidemiology (Cox and Donnelly 2011), and wildlife management (Caughley and Sinclair 1994) are analogous. For example, dose-response (medicine) or effort-outcome (wildlife management) relationships, the design and interpretation of studies of many people (common diseases) or many wildlife (common and overabundant species), and the study of few people (rare diseases) or few wildlife (threatened and endangered species) are similar. Randomized control trials to adaptive experimental designs (Adaptive Platform Trials Coalition 2019) to infer effects of interventions (medical treatments) are similar to wildlife management designs to infer effects of efforts. Disease and wildlife-threatening processes may have a single cause or multiple simultaneous causes. Medicine describes a prognosis for future disease and wildlife management makes predictions, projections, and forecasts about the future. We use the terms predictions, projections, forecasts, and prognoses as synonyms. Both disciplines can require ethics committee approval, which may limit some experimental designs (Ware 1989). Considering the similarities, we expect medical causality can be usefully applied in wildlife management.

Manipulative experiments with replicated, randomized treatments and controls are often considered the gold standard for demonstrating causality (Larsen et al. 2019). Experiments require assumptions that should be verified to increase the chance that observed results are caused by treatments rather than uncontrolled covariates (Kimmel et al. 2021). Causality has developed over the past 150 years, and recent ecological examples include identifying causal influences on species traits (Cronin and Schoolmaster 2018), biodiversity, and ecosystem functioning (Schoolmaster et al. 2020). Researchers have examined how much can be causally inferred from observational studies (Rubin 2005; Pearl 2009*a*, *b*; Imbens 2020). We do not comprehensively review the assumptions, strengths, and limitations of these alternative approaches, though Imbens (2020) has a useful recent review. Rather, we highlight key features relevant to wildlife management. In particular, using causal inference, we estimate effects of causes in experiments in contrast to inferring the causes of effects in observations (Holland 1986).

Our objective is to collate published causal criteria used to infer causality in a range of scientific disciplines. We collate and discuss qualitative rankings of levels of causal inference, and propose that validating predictions represents a high level of causal inference when used with a known unique mechanism and temporality, that is, the cause occurs before the effect. We then propose a future study that aims to estimate the relative sensitivity of various analyses for demonstrating causality, in the apparent absence of such a published study. We apply causal criteria and levels of causal inference to 3 case studies of wildlife management involving multiple field studies: lamb predation by feral pigs (*Sus scrofa*) in southern Australia, causes of population trends in northern spotted owls (*Strix occidentalis caurina*) in western North America, and effects of harvesting on abundance of mallards (*Anas platyrhynchos*) in parts of North America.

## CAUSALITY

In the late 1800s, a study established whether a microbe caused a disease, as described in Koch's postulates (Byrd and Segre 2016). These postulates involve isolating and identifying the microbe in a diseased host, infecting a different healthy host that later develops the same disease and from which the same microbe can be isolated, and observing transfer to additional hosts thereby serially demonstrating disease causality (Byrd and Segre 2016). The declining population paradigm in conservation biology (Caughley 1994) uses analogous steps of identifying a likely cause(s) of decline, reducing the likely cause(s), verifying if the decline reverses, and often introducing the population to an area free of the likely cause, as demonstrated in wildlife examples (Leopold 1933, Caughley and Gunn 1996).

We identified causal criteria described in the scientific literature. We conducted online searches in Google using the terms causal criteria, causal guidelines, and causality. We also searched Google Scholar for cited publications, and read 15 statistics, wildlife, and ecology textbooks. We found causal criteria in 13 publications (Table 1). The criteria are summarized as:

Plausible mechanism. The proposed mechanism is biologically realistic.
Consistency. The effect of the proposed cause is repeatable and hence predictable.
Experiment. A comparison of effects with and without a treatment (a proposed cause).
Temporality. The cause occurs before the effect.
Dose-response relationship. A range of levels of the cause has a clear relationship to the levels of effects.
Specificity. The effect is particular to the cause.
Coherence. Multiple lines of evidence give the same conclusion of causality.
Strength. The relationship between cause and effect is statistically strong such as a high correlation and a precisely estimated slope of a relationship.
Evidence of response. An effect is observed not assumed.
Analogy. A cause and effect relationship is like a similar ecological process.
Evidence of stressor. A cause is observed not assumed.
Convergence of predictions. With an increasing sample size there is an increase in the correlation between observed and predicted values of effects.

The 4 criteria described most frequently were plausibility of the mechanism (10 studies), consistency (9 studies), experiments demonstrating change (8 studies), and temporality (8 studies). Some criteria appear redundant (e.g., evidence of stressor and evidence of response are subsumed into temporality). Temporality and mechanism can be linked by checking that the time scale between cause and effect is consistent with the mechanism (Howick et al. 2009). No criteria were nominated by all studies and no study nominated all criteria (Table 1).

A distinction is made often between causality and correlation. The latter describes the extent and direction (positive or negative) of association without implying causality. A common criticism of observations, in contrast with randomized manipulative experiments, is that observed effects could be caused by some unmeasured, confounding factor. Using Cornfield's inequality can assess the relative contribution of a possible confounding factor to an observed effect (Cornfield et al. 1954, Greenhouse 2009). Researchers used this to conclude the contribution of genes to development of lung cancer in smokers was minor (Cornfield et al. 1954). We can also use propensity scoring to adjust for potential confounders, as in studies of the effects of invasive mammalian herbivores (Ramsey

**TABLE 1** Examples of published criteria for establishing causality. Association is not listed here as its interpretation is ambiguous. Criteria are listed in rows by frequency of occurrence with most frequent at the top, and sources are listed in columns chronologically with oldest on the left. Criteria are shown as reported (+) or not reported (−). Sources are 1 = Hill (1965), 2 = Granger (1969), 3 = Harre (1972), 4 = Susser (1991), 5 = Ford (2000), 6 = Adams (2005), 7 = Cox (2007), 8 = Pearl (2009*a*, *b*), 9 = Howick et al. (2009), 10 = Cox and Donnelly (2011), 11 = Norris et al. (2012), 12 = Sugihara et al. (2012), and 13 = Nichols et al. (2017).

| | Sources | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Criteria** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** |
| Plausible mechanism | + | − | + | − | + | + | + | + | + | + | + | − | + |
| Consistency | + | + | + | − | + | + | − | − | + | + | + | − | + |
| Experiment | + | + | + | + | + | + | − | − | − | + | − | − | + |
| Temporality | + | − | + | + | − | + | + | + | + | + | − | − | − |
| Dose-response relationship | + | − | − | − | − | + | − | − | + | + | + | − | + |
| Specificity | + | − | − | − | − | + | + | − | − | + | − | − | − |
| Coherence | + | − | − | − | − | − | − | − | + | − | + | − | + |
| Strength | + | − | − | − | − | + | − | − | + | + | − | − | − |
| Evidence of response | − | − | − | − | − | − | − | − | − | − | + | − | + |
| Analogy | + | − | − | − | − | − | − | − | + | − | − | − | − |
| Evidence of stressor | − | − | − | − | − | − | − | − | − | − | + | − | + |
| Convergence of predictions | − | − | − | − | − | − | − | − | − | − | − | + | − |

et al. 2019), but some authors have discouraged the analysis (King and Nielsen 2019). An assumption of causal sufficiency, namely that all common causes are included (Runge et al. 2019) is required in such analyses, as strong inferences are limited by unobserved variables acting as confounders.

Greater use in wildlife management of explicit causal criteria to complement adaptive management has been proposed to increase the strength of causal inference (Hone et al. 2023). So how do studies rank the strength of causal inference? We now briefly answer this question using published rankings from various scientific disciplines.

## STRENGTH OF CAUSAL INFERENCE

Strength of inference is also known as levels of evidence (McArdle 1996) or strength of evidence (Anderson 2008) and describes our confidence in having demonstrated a relationship. We present 4 previously published systems of ranking the levels, or strength, of causal inference proposed in science and statistics (Table 2). The ranking schemes use a scaffolding approach so any level assumes evidence of each of the lower levels. One system features association, temporality, and direction (consequential change; Susser 1991). A second system proposes 3 levels: zero level is based on evidence of association, first level is based on experiment, and the second level uses evidence of a mechanism (Cox and Wermuth 2004:287). A third system uses a hierarchy of direct, then mechanistic and parallel evidence (Howick et al. 2009). A fourth system proposes a ladder of causality with 3 rungs: association (seeing), interventions (doing), and counterfactuals (imagining; Pearl and Mackenzie 2018). We propose related ranking systems (Table 2, numbers 5 and 6) building on the previously proposed causal criteria (Tables 1 and 2) and the observation that validated predictions, based on a unique mechanism and temporality, that are consistent (repeated), unbiased, and precise provide strong evidence of causality (Table 2).

**TABLE 2** Examples of comparative elements of systems for ranking strength or levels of causal inference. Sources are 1 = Susser (1991), 2 = Cox and Wermuth (2004), 3 = Howick et al. (2009), 4 = Pearl and Mackenzie (2018), and 5 (manipulative experiments) and 6 (observations) are proposals from this commentary.

| | Sources | | | | | |
|---|---|---|---|---|---|---|
| **Strength** | **1** | **2** | **3** | **4** | **5** | **6** |
| Zero | Association | Association | | Association | Association | Association |
| First | Temporality | Experiment | Direct evidence | Intervention | Experiment with change and temporality | Unique mechanism and temporality |
| Second | Direction (change) | Mechanism | Mechanistic evidence | Counter-factual | Unique mechanism | Validated predictions |
| Third | | | Parallel evidence | | Validated predictions | |

**TABLE 3** Examples of validating predictions involving some active or passive wildlife management. The parameters estimated, the analysis used to assess prediction bias, and sources are shown. Predictions were point, not interval, estimates. Studies that included statistical analysis reported no significant bias.

| Parameters | Analysis | Sources |
|---|---|---|
| Waterfowl mortality in response to hunting | Many analyses | Nichols (1991:509–517) |
| Prevalence of bovine tuberculosis in possums | 95% CI coverage | Caley et al. (1999:figure 5) |
| Abundance of harvested kangaroos | 95% CI coverage | Morgan and Pegler (2010:figure 30.5) |
| Abundance of harvested mallards | 95% CI coverage | Nichols et al. (2015:figure 1a), Nichols et al. (2019:figure 1) |
| Densities of black and white rhinoceros | 95% CI coverage | Ferreria et al. (2015:text) |
| Trends in elephants | Resilience framework | Guldemond et al. (2022:figure 4) |

Published proposals discussed in the previous paragraph do not include validated predictions, in contrast to historical examples in the broader fields of science. Historical scientific examples include the correct prediction by Halley of the year of the next return sighting of what is now called Halley's comet (Lipton 2005); the insertion of gaps in Mendeleev's periodic table of the chemical elements followed later by independent discovery of gallium and germanium, the properties of which matched those expected of elements in the table's gaps (Ball 2019); and 7 predictions of Einstein's general theory of relativity being validated, including the bending of light (Anon 2019). In epidemiology, the trend in cases of foot and mouth disease in livestock in the United Kingdom was predicted by a mechanistic model of host-disease status and dynamics including livestock culling (Government Office for Science 2018). In wildlife management, predictions of abundance of mallards one year ahead were often unbiased, and were supported by increasing evidence of unique mechanisms, namely additive mortality and weak density-dependent reproduction in the presence of hunting mortality (Nichols et al. 2015, 2019). Those 5 predictions used unique mechanisms to generate the predictions, namely Newtonian physics, chemical properties, bending of space by mass, changes in host-disease status, and hunting mortality, respectively. Our proposed ranking scheme for strength of causal inference without experiments but with a unique mechanism, temporality, and validated predictions (Table 2, source 6) occurred in the 5 examples of predictions. Hence, these were all causal, not statistical, predictions that were validated. Other wildlife examples of validated predictions are present in the literature (Table 3).

Experimental comparisons with and without the hypothesized cause allow estimation of a change in effects (ΔY) and such a change is to be greater than the influence of plausible confounders. Evidence of change is more important than for the study to be experimental (Howick et al. 2009). The criterion of experiment with change, also known as difference-making (Russo and Williamson 2007), combines the criteria (Table 1) of experiment, strength, dose-response, and temporality. Evidence that a possible cause occurred after an event leads to refutation or rejection of a hypothesis of cause and effect (Popper 1965, Weed 1997). The ranking systems of strength for causal evidence complement but differ from the useful levels of evidence proposed for environmental management decisions based on accumulated studies and their review in Dicks et al. (2014) because that review does not include explicit consideration of causality.

Wildlife and environmental management efforts are implemented to produce particular outcomes, such as increasing the abundance of an endangered species by predator control or providing nest boxes, or they aim to allow harvesting wildlife sustainably (Caughley 1994). Researchers and managers would like to be confident that the management action produced the observed outcome. Many wildlife management studies have assessed strength of causal inference, such as using differing efforts to increase bird abundance (Newton 1998), conserve species (Rehme et al. 2011), and remove non-native fish in aquatic systems (Rytwinski et al. 2019). Validating predictions was not used to assess causal evidence in any of the above studies. In contrast, studies in disease ecology and control (Plowright et al. 2008), environmental impact assessments (Suter et al. 2010, Norris et al. 2012, Nichols et al. 2017), and community ecology (Barton et al. 2015, Schoolmaster et al. 2020), explicitly emphasized the causal criteria approach of Hill (1965) or Pearl (2009*a*). Alternative models, as hypotheses, can be evaluated using information-theoretic analyses (Anderson 2008), noting that those analyses evaluate relative association and not necessarily causality. We now examine one topic in the ranking systems in more detail, namely validating predictions, as this criterion has been prominent in recent research. Of course, unbiased predictions do not by themselves imply causality; however, a lack of predictive accuracy demonstrates possible confounding or additional causal factors are obscuring the causal effect of interest.

## VALIDATING PREDICTIONS

The Green List (Akçakaya et al. 2018), now called Green Status of species (Grace et al. 2021, International Union for Conservation of Nature 2022), incorporates expected trends under various management scenarios and hence includes predictions of wildlife trends and management effects. The criterion of predictability combines the criteria of consistency, coherence, and convergence and can be based upon analogy (Tables 1 and 2). Obtaining similar predictions from multiple sources of evidence demonstrates coherence. Unbiased prediction implies the ability to generate reliable knowledge (Romesburg 1981, Mouquet et al. 2015, Sells et al. 2018). Ideally, prediction uses data independent of the formulation of the theory or hypothesis (Lloyd 2010), and hence is out-of-sample, and is called confirmation (Hempel 1966:37), verification (Garton et al. 2005), or validation (Mouquet et al. 2015). The utility of predictions is especially important in wildlife management, which can involve considerable costs. Managers need to make decisions and predictions need to be useful, unbiased, and precise. Sometimes the best predictions may not be the most useful (Boettiger 2022).

Explanation of past events (i.e., accommodation) provides weaker inference than prediction of independent future events (Lipton 2005). But such accommodation can be useful as shown by attempts to reconstruct past global and continental-scale temperatures with (less biased) and without (more biased) effects of human-released greenhouse gases (Intergovernmental Panel on Climate Change [IPCC] 2007:figure SPM4, 2021:figure SPM1b). Those IPCC studies used observational data, and a simple experimental comparison of 2 sets of models, namely with and without effects of greenhouse gases (Lloyd 2010). Classical experiments manipulating greenhouse gases are clearly not possible on a continental or global scale. Prediction is a key component of Granger causality (Granger 1969, Sugihara et al. 2012), which proposes that removal of a putative cause results in poorer predictions than in the presence of the cause. The analysis inferred a causal relationship between greenhouse gas levels, especially carbon dioxide in the atmosphere and temperatures (Kodra et al. 2011). More recently, causal discovery has been used to infer causality in weather conditions (Runge et al. 2019). Convergent cross mapping proposes that predictions improve with larger datasets as demonstrated by an increasing correlation between observed and predicted parameter values (Sugihara et al. 2012, Chang et al. 2017), though without the experimental comparison. Convergent cross mapping and dynamic modeling have been used in fisheries science and management and demonstrated predictability (Sugihara et al. 2012, Ushio et al. 2018); however, dynamics of mammal, bird, and bony fish populations were often unpredictable beyond 2 years ahead (Clark and Luis 2020). Clark and Luis (2020) suggest that predictions should be short-term and updated regularly as in predictions of mallard abundance (Nichols et al. 2015, 2019) and in weather forecasting.

The techniques used for validating predictions in medicine need further research (Lin et al. 2021) and guidelines for such validation have been proposed (Riley et al. 2016). The metric used to evaluate predictions in wildlife management can be examined more closely. Use of the correlation coefficient (Sugihara et al. 2012) assesses the closeness of the observed and predicted values. But correlation assesses joint distributions (Pearl 2009*a*) not change. Correlation has no units but estimates of change ($\Delta Y$) have units. The correlation does not assess bias; a high correlation can occur with quite biased predictions. Bias is assessed by estimating if the slope of a linear regression of observed and predicted values equals 1.0, and the intercept on the *y*-axis is the origin (0, 0). An alternative metric that combines bias and a measure of variation around the prediction is mean squared error (MSE), which is defined as bias$^2$ + variance (Cochran 1963), and has been used to evaluate forecasts (predictions) of mallard dynamics (Zhao et al. 2016). The best model has the lowest MSE. Predictions that use regression assume the *x* variable(s) are measured with no error. Presence of such error lowers the estimated slope (Snedecor and Cochran 1967) and increases the width of prediction intervals (Behney 2020). Model selection methods such as Akaike's Information Criterion assess relative goodness of fit of models and have a strong theoretical basis (Anderson 2008). Such analyses and causal inference, however, are not the same though can be considered complementary (Stewart et al. 2023).

# PROPOSAL FOR STUDIES OF THE COMPARATIVE SENSITIVITY OF ANALYSES TO DETERMINE CAUSALITY

Several experimental and observational analyses now exist to evaluate causality as outlined above. Has the relative sensitivity, or comparative ability, of these approaches to infer causality been evaluated? Sensitivity is defined here as the ability to detect causality when it occurs, and as the change in outcomes ($\Delta Y$) for a unit change in management efforts ($\Delta X$). Both components of the definition are relevant to wildlife scientists and managers. Ecological interactions in small mammals have been investigated using classical experiments and path analysis (Wright 1921) of observations and experimental data (Smith et al. 1997). The path analysis yielded mixed results, identifying some causal interactions and not others, whereas analysis of experimental data identified more causal relationships (Smith et al. 1997). These interpretations were criticized (Grace and Pugesk 1998) and the criticisms were recognized and refuted (Smith et al. 1998). In human medicine, the results of randomized controlled trials and observational studies on the same topics have reported similar results (Benson and Hartz 2000, Concato et al. 2000), though their results and interpretations have also been disputed (Kunz et al. 2000, Pocock and Elbourne 2000).

We propose similar comparative studies using newer alternative analyses to estimate the effects of causes and infer causes of effects (Holland 1986) and the relative sensitivity of different analyses to infer causality. Observational studies that provide data on wildlife management efforts and outcomes could infer causes of effects using various analyses, including structural equation modeling (Smith et al. 1997, Grace et al. 2015), path analysis (Shipley 2009), the methods of Pearl (2009*a*, *b*; Arif and MacNeil 2022), convergent cross mapping (Sugihara et al. 2012), empirical dynamic modeling (Ushio et al. 2018), propensity scores (Ramsey et al. 2019), potential outcomes analysis (Rubin 2005, Imbens 2020), and causal discovery (Runge et al. 2019). The proposed studies should evaluate the assumptions of each alternative analysis to help identify reasons why some analyses generate stronger to weaker causal inferences. A simulation study would be very useful, as logistics may limit a large-scale field study, unless a space-for-time substitution can be used for small datasets (Clark et al. 2015). The proposed sensitivity study would go some way to addressing the concern of Shadish et al. (2002) that observational studies may lead to unrealistic expectations or inferences about the strength of causal inference.

The topics of causality and strength of causal inference can be applied to specific wildlife management topics. We illustrate the application with 3 case studies, 2 of which demonstrate strength of causal inference based on manipulative experiments and 1 based on observations and validated predictions.

## CASE STUDY 1: LAMB PREDATION BY FERAL PIGS

Lamb production can be adversely influenced by predation in parts of southern Australia (Rowley 1970, Pavlov et al. 1981) and managers need to know how much lamb production could be changed by reducing such predation. The features of 4 studies are evaluated here using causal criteria (Table 1) and the ranking of strength of causal inference (Table 2). Association and temporality were evident; studies reported that feral pigs were present before and during lambing (Table 4). Comparative manipulative experiments had features of controls (simultaneous not sequential with and without feral pigs), randomization, replication, analysis, and temporality (Table 4). Randomization consisted of random allocation of pregnant sheep to treatment sites (Plant et al. 1978, Pavlov et al. 1981, Choquenot et al. 1997), rather than random allocation of the treatment, feral pigs, to sites. A positive dose-response relationship (Table 4) with diminishing returns was demonstrated between the number of female sheep that had lambed and lost their lamb (inferred predation) and feral pig abundance (Choquenot et al. 1997). As predicted by predator-prey theory (Krebs 2009), there was a positive relationship between lamb kills and lamb abundance (Pavlov et al. 1981, Hone 1994). Reports of lambs killed by feral pigs (Rowley 1970) were strengthened by repeated observations of lamb kills by feral pigs (Pavlov and Hone 1982). The strength of evidence was the difference (37 lambs produced per 100

**TABLE 4** Strength or level of causal inference reported (+) or not (−) in studies of lamb predation by feral pigs in southern Australia. Examples of studies are 1 = Rowley (1970); 2 = Plant et al. (1978); 3 = Pavlov et al. (1981), Pavlov and Hone (1982), and Hone (1994); and 4 = Choquenot et al. (1997).

| Strength | Features of studies | Sources 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Zero (association) | Feral pigs with live lambs and then lambs killed by feral pigs | + | + | + | + |
| First (manipulative experiment with change and temporality) | Experimental control | − | + | + | + |
| | Randomization | − | + | + | + |
| | Replication | − | + | + | + |
| | Statistical analysis | − | − | + | + |
| | Dose-response relationship shows change | − | − | + | + |
| Second (unique mechanism) | Predation of lambs by feral pigs observed with unique feeding behavior | + | − | + | − |
| Third (validated predictions) | Validation with independent data<br>Granger causality<br>Convergence with larger dataset | − | − | − | − |
| Levels demonstrated | | 2 | 2 | 3 | 2 |

pregnant sheep) between flocks with (80) and without (117) feral pigs (Plant et al. 1978) and averaged 42% and 61% across 4 lambing periods (Pavlov et al. 1981). The average number of lambs at the end of lambing season was 111 with feral pigs and 138 without feral pigs across 3 densities of feral pigs, with each flock having approximately 150 pregnant sheep at the start of the lambing season (Choquenot et al. 1997, their experiment 2). The behavior of feral pigs killing and feeding on lambs showed a unique mechanism (Table 4) that allowed discrimination between predation by feral pigs and a possible alternative predator, namely the red fox (*Vulpes vulpes*; Pavlov and Hone 1982). There was no validation of predictions of lamb predation by feral pigs and no evidence of convergence of predictions with a larger dataset (Table 4). Predation by feral pigs is analogous to some other mammalian predators, namely coursers, who chase and run down their prey. Evidence of the stressor (feral pigs) and the stress (lamb deaths, or female sheep that had lambed and lost their lambs) were reported. The results of the above studies and their causal features were combined to estimate the benefits and costs of feral pig control (Choquenot and Hone 2002). The 4 studies of lamb production demonstrated 2 to 3 levels of causality (Table 4).

## CASE STUDY 2: MANAGEMENT OF NORTHERN SPOTTED OWLS

The northern spotted owl occurs in forests of western North America. Detailed demographic studies have reported no trend (Lande 1988) and later a slow decline in abundance across their range (Anthony et al 2006, Forsman et al. 2011, Dugger et al. 2016, Franklin et al. 2021). Observational studies have hypothesized (Kelly et al. 2003) or identified effects of many environmental factors on owl demography, including some negative effects of associated barred owl (*Strix varia*) on spotted owl survival and colonization rates (Gutiérrez et al. 2007, Dugger et al. 2016), and on survival and recruitment (Franklin et al. 2021). These latter studies are evidence of association (Table 5). More recently, researchers using experimental removals of barred owls from parts of the spotted owl range have reported an increase in annual finite population growth rate of spotted owls (Diller et al. 2016, Dugger et al. 2016, Wiens

**TABLE 5** Strength or level of causal inference reported (+) or not (−) in studies of causes of population trends in northern spotted owls in western North America. Examples of studies are 1 = Lande (1988); 2 = Kelly et al. (2003), Gutiérrez et al. (2007), Anthony et al. (2006), Forsman et al. (2011), and Franklin et al. (2021); 3 = Diller et al. (2016), Dugger et al. (2016), and Wiens et al. (2021).

| | | Sources | | |
|---|---|---|---|---|
| **Strength** | **Features of studies** | **1** | **2** | **3** |
| Zero (association) | Barred owl present with northern spotted owl | − | + | + |
| First (manipulative experiment with change and temporality) | Experimental control | − | − | + |
| | Randomization | − | − | − |
| | Replication | − | − | + |
| | Statistical analysis | + | + | + |
| | Dose-response relationship shows change | − | − | − |
| Second (unique mechanism | Negative effects of barred owls on spotted owl apparent survival and recruitment rates | − | + | + |
| Third (validated predictions) | Validation with independent data | − | − | − |
| | Granger causality | − | − | − |
| | Convergence with larger dataset | − | − | − |
| Levels demonstrated | | 1 | 3 | 3 |

et al. 2021). The comparative manipulative experiments had features of controls, replication (3 to 5 areas with barred owls removed and 3 to 5 areas as controls with no barred owls removed), analysis, and temporality but not randomization (Table 5). A dose-response relationship between levels of removals and change in growth rate was not reported. A mechanism of negative effects on spotted owl survival, colonization, and recruitment rates were reported (Table 5). There were no validated predictions, though a prediction occurred of continued spotted owl decline if barred owls were not removed (Franklin et al. 2021). The studies demonstrated 1 to 3 levels of causality (Table 5).

## CASE STUDY 3: MANAGEMENT OF MALLARD HARVEST

Mallards are harvested in North America as part of recreational hunting. Studies facilitate answering the question, does mallard harvest cause variation in mallard abundance? Researchers using observations have reported a negative relationship between annual survival rates and harvest rates of mallard (Reynolds and Sauer 1991), hence demonstrating association (Table 6). There are no published manipulative, randomized, replicated experiments on the effects of harvest (Table 6), as noted previously (Johnson et al. 2002, Conn and Kendall 2004). A mechanism (Table 6) for an effect of harvest on abundance is a theoretical and observed relationship between survival rates and harvest mortality rate (Nichols 1991; Johnson et al. 2002; Nichols et al. 2015, 2019). Validated predictions (Table 6) of mallard abundance have been reported (Cooch et al. 2014; Johnson et al. 2015; Nichols et al. 2015, 2019; Zhao et al. 2016). Early studies reported predictions for each of 4 models of dynamics, and later studies reported those and model-averaged predictions. Adaptive management (Williams et al. 1996) is used to learn the effects of changes in harvest levels and regulations. The studies demonstrated 2 to 3 levels of causality (Table 6).

**TABLE 6** Strength or level of causal inference reported (+) or not (−) in studies of causes of population trends in mallard in parts of North America. Examples of studies are 1 = Reynolds and Sauer (1991); 2 = Nichols (1991); and 3 = Johnson et al. (2002), Cooch et al. (2014), Johnson et al. (2015), Nichols et al. (2015, 2019), and Zhao et al. (2016).

| | | Sources | | |
|---|---|---|---|---|
| Strength | Features of studies | 1 | 2 | 3 |
| Zero (association) | Negative relationship between survival rates and hunting rates in observations | + | + | + |
| First (manipulative experiment with change and temporality) | Experimental control | − | − | − |
| | Randomization | − | − | − |
| | Replication | − | − | − |
| | Statistical analysis | − | − | − |
| | Dose-response relationship shows change | − | − | − |
| Second (unique mechanism) | Negative effects of hunting mortality on mallard survival rates, with temporality shown by time difference between harvest and survival rate estimation | + | + | + |
| Third (validated predictions) | Validation with independent data | − | − | + |
| | Granger causality | − | − | − |
| | Convergence with larger dataset | − | − | − |
| Levels demonstrated | | 2 | 2 | 3 |

## SUMMARY

Causality is fundamental in science (Hill 1965, Harre 1972), including wildlife management (Williams 1997, Ford 2000), and is central to our ability to infer that management efforts cause observed outcomes. We show that strength of causal inference and causal criteria are central to helping us, in Leopold's (1933:231) words, "to comprehend the meaning of evidence." Many previous studies have stimulated much useful discussion about strength of inference (Platt 1964) and scientific rigor (Romesburg 1981, Sells et al. 2018), and their relationship to strength of causal inference. Analogous issues in medicine and epidemiology suggest wildlife management could benefit from using similar approaches when making causal inferences; each discipline tries to make causal inferences about how to manage common to rare occurrences (diseases or wildlife). We recommend future studies evaluate the relative sensitivity of experiments and more recent analyses of observations, using multiple levels of management efforts or actions to estimate effort-outcome relationships, recognizing few such studies have occurred. Simulation studies could be very useful approaches.

## MANAGEMENT IMPLICATIONS

The management implications of weaker versus stronger causal inference are several. First, research or management studies demonstrating weaker causal inference (zero level; Table 2) should be interpreted as hypothesis generating for further study and should not be used in management policy and practice as evidence of causality. Studies leading to weaker inference should be described appropriately. Those of intermediate strength (first level; Table 2) should be interpreted cautiously and used in adaptive management and research to increase knowledge of causality. Studies with stronger causal inference (second and third levels; Table 2) could be used in management policy and practice, with on-going assessment in an adaptive management framework. Second, many studies in wildlife management cannot do manipulative experiments because of logistical, financial, or ethical constraints but can use natural experiments created by events such as highway construction or bushfires, or make observations, and then generate and validate causal predictions. We recommend managers and scientists validate predictions of the outcomes of wildlife management efforts, noting that some studies have done so already (Table 3). Third, the topics listed here (Tables 1 to 6) should enable scientists and managers to quickly assess the desirable features of studies or management programs, identify their relative strengths and weaknesses, and develop improved management. Such assessment would also include costs, which we do not evaluate here. We recommend greater use of causality to improve adaptive wildlife management. We also recommend a close examination of the extent to which causality is taught in wildlife courses at colleges and universities to future wildlife scientists and managers.

### ACKNOWLEDGMENTS

We acknowledge the innovative thinking of our late colleague Graeme Caughley. We thank colleagues B. Allen, S. Box, P. Bridgewater, P. Caley, C. A. Donnelly, V. A. Drake, J. Li, S. Nichols, and A. Upitis for useful comments or discussions. We thank anonymous reviewers and an editor for useful comments on an earlier manuscript. We thank the Institute for Applied Ecology and University of Canberra for support. Open access publishing facilitated by University of Canberra, as part of the Wiley - University of Canberra agreement via the Council of Australian University Librarians.

### CONFLICTS OF INTEREST STATEMENT

The authors declare no conflicts of interest.

### ETHICS STATEMENT

Animal Ethics Committee approval was not required for this research.

## DATA AVAILABILITY STATEMENT

Data sharing not applicable - no new data generated.

## ORCID

*Jim Hone* http://orcid.org/0000-0002-8104-8852

## REFERENCES


Adams, S. M. 2005. Assessing cause and effect of multiple stressors on marine systems. Marine Pollution Bulletin 51: 649–657.

Adaptive Platform Trials Coalition. 2019. Adaptive platform trials: definition, design, conduct and reporting considerations. Nature Reviews 18:797–807.

Akçakaya, H. R., E. L. Bennett, T. M. Brooks, M. K. Grace, A. Heath, S. Hedges, C. Hilton-Taylor, M. Hoffmann, D. K. Keith, B. Long, et al. 2018. Quantifying species recovery and conservation success to develop an IUCN green list of species. Conservation Biology 32:1128–1138.

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer Verlag, Berlin, Germany.

Anon. 2019. A century of correct predictions. Nature Physics 15:415.

Anthony, R. G., E. D. Forsman, A. B. Franklin, D. R. Anderson, K. P. Burnham, G. C. White, C. J. Schwarz, J. D. Nichols, J. E. Hines, G. S. Olson, et al. 2006. Status and trends in demography of northern spotted owls, 1985–2003. Wildlife Monographs 163:1–48.

Arif, S., and M. A. MacNeil. 2022. Predictive models aren't for causal inference. Ecology Letters 25:1741–1745.

Artelle, K. A., J. D. Reynolds, A. Treves, J. C. Walsh, P. C. Paquet, and C. T. Darimont. 2018. Hallmarks of science missing from North American wildlife management. Scientific Advances 4:eaao0167.

Ball, P. 2019. Bringing order to chaos. New Scientist 421:34–36.

Barton, P. S., J. C. Pierson, M. J. Westgate, P. W. Lane, and D. B. Lindenmayer. 2015. Learning from clinical medicine to improve the use of surrogates in ecology. Oikos 124:391–398.

Behney, A. C. 2020. Ignoring uncertainty in predictor variables leads to false confidence in results: a case study of duck habitat use. Ecosphere 11:e03273.

Benson, K., and A. J. Hartz. 2000. A comparison of observational studies and randomized, controlled trials. New England Journal of Medicine 342:1878–1886.

Boettiger, C. 2022. The forecast trap. Ecology Letters 25:1655–1664.

Byrd, A. L., and J. A. Segre. 2016. Adapting Koch's postulates. Science 351:224–226.

Caley, P., G. J. Hickling, P. E. Cowan, and D. U. Pfeiffer. 1999. Effects of sustained control of brushtail possums on levels of *Mycobacterium bovis* infection in cattle and brushtail possum populations from Hohotaka, New Zealand. New Zealand Veterinary Journal 47:133–142.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology 63:215–244.

Caughley, G., and A. Gunn. 1996. Conservation biology in theory and practice. Blackwell, Oxford, United Kingdom.

Caughley, G., and A. R. E. Sinclair. 1994. Wildlife ecology and management. Blackwell, Oxford, United Kingdom.

Chang, C.-W., M. Ushio, and C.-H. Hseih. 2017. Empirical dynamic modeling for beginners. Ecological Research 32:785–796.

Choquenot, D., and J. Hone. 2002. Using bioeconomic models to maximise benefits from vertebrate pest control: lamb predation by feral pigs. Pages 65–79 *in* L. Clark, editor. Human conflicts with wildlife: economic considerations. USDA, Fort Collins, Colorado, USA.

Choquenot, D., B. Lukins, and G. Curran. 1997. Assessing lamb predation by feral pigs in Australia's semi-arid rangelands. Journal of Applied Ecology 34:1445–1454.

Clark, A. T., H. Ye, F. Isbell, E. R. Deyle, J. Cowles, D. Tilman, and G. Sugihara. 2015. Spatial convergent cross mapping to detect causal relationships from short time series. Ecology 96:1174–1181.

Clark, T. J., and A. D. Luis. 2020. Nonlinear population dynamics are ubiquitous in animals. Nature Ecology and Evolution 4:75–81.

Cochran, W. G. 1963. Sampling techniques. Second edition. Wiley, New York, New York, USA.

Concato, J., N. Shah, and R. I. Horwitz. 2000. Randomized, controlled trials and observational studies, and the hierarchy of research designs. New England Journal of Medicine 342:1887–1892.

Conn, P. B., and W. L. Kendall. 2004. Evaluating adaptive mallard management models with time series. Journal of Wildlife Management 68:1065–1081.

Cooch, E. G., M. Guillemain, G. S. Boomer, J.-D. Lebreton, and J. D. Nichols. 2014. The effects of harvest on waterfowl populations. Wildfowl Special Issue 4:220–276.

Cornfield, J., W. Haenszel, E. C. Hammond, A. M. Lilienfeld, M. B. Shimkin, and E. L. Wynder. 1954. Smoking and lung cancer: recent evidence and a discussion of some questions. Journal of the National Cancer Institute 22: 173–203.

Cox, D. R. 2006. Principles of statistical inference. Cambridge University Press, Cambridge, United Kingdom.

Cox, D. R. 2007. Applied statistics: a review. Annals of Applied Statistics 1:1–16.

Cox, D. R., and C. A. Donnelly. 2011. Principles of applied statistics. Cambridge University Press, Cambridge, United Kingdom.

Cox, D. R., and N. Wermuth. 2004. Causality: a statistical view. International Statistical Review 72:285–305.

Cronin, J. P., and D. R. Schoolmaster. 2018. A causal partition of trait correlations: using graphical models to derive statistical models from theoretical language. Ecosphere 9:e02422.

Dicks, L. V., J. C. Walsh, and W. J. Sutherland. 2014. Organising evidence for environmental management decisions: a '4S' hierarchy. Trends in Ecology and Evolution 29:607–613.

Diller, L. V., K. A. Hamm, D. A. Early, D. W. Lamphear, K. M. Dugger, C. B. Yackulic, C. J. Schwarz, P. C. Carlson, and T. L. McDonald. 2016. Demographic response of northern spotted owls to barred owl removal. Journal of Wildlife Management 80:691–707.

Dugger, K. M., E. D. Forsman, A. B. Franklin, R. J. Davis, G. C. White, C. J. Schwarz, K. P. Burnham, J. D. Nichols, J. E. Hines, C. B. Yackulic, et al. 2016. The effects of habitat, climate, and barred owls on long-term demography of northern spotted owls. Condor 118:57–116.

Ferreria, S. M., C. Greaver, G. A. Knight, M. H. Knight, I. P. J. Smit, and D. Pienaar. 2015. Disruption of rhino demography by poachers may lead to population declines in Kruger National Park, South Africa. PLoS One 10:e0127783.

Ford, E. D. 2000. Scientific method for ecological research. Cambridge University Press, Cambridge, United Kingdom.

Forsman, E. D., R. G. Anthony, K. M. Dugger, E. M. Glenn, A. B. Franklin, G. C. White, C. J. Schwarz, K. P. Burnham, D. R. Anderson, J. D. Nichols, et al. 2011. Population demography of northern spotted owls. Studies in Avian Biology Number 40. University of California Press, Berkeley, USA.

Franklin, A. B., K. M. Dugger, D. B. Lesmeister, R. J. Davis, J. D. Wiens, G. C. White, J. D. Nichols, J. E., Hines, C. B. Yackulic, C. J. Schwarz, et al. 2021. Range-wide declines of northern spotted owl populations in the Pacific Northwest: a meta-analysis. Biological Conservation 259:1091168.

Garton, E. O., J. T. Ratti, and J. H. Giudice. 2005. Research and experimental design. Pages 43–71 *in* C. Braun, editor. Techniques for wildlife investigations and management. Sixth edition. The Wildlife Society, Washington, D.C., USA.

Government Office for Science. 2018. Computational modelling: technological futures. Government Office for Science, London, United Kingdom.

Grace, J. B., and B. H. Pugesk. 1998. On the use of path analysis and related procedures for the investigation of ecological problems. American Naturalist 152:151–159.

Grace, J. B., S. M. Scheiner, and D. R. Schoolmaster. 2015. Structural equation modeling: building and evaluating causal models. Pages 168–199 *in* G. A. Clark, S. Negrete-Yankelevich, and V. J. Sosa, editors. Ecological statistics: contemporary theory and applications. Oxford University Press, Oxford, United Kingdom.

Grace, M. K., E. L. Bennett, H. R. Akçakaya, C. Hilton-Taylor, M. Hoffmann, R. Jenkins, E. J. Milner-Gulland, A. Nieto, R. P. Young, and B. Long. 2021. IUCN launches Green Status of species: a new standard for species recovery. Oryx 55: 651–652.

Granger, C. W. J. 1969. Investigating causal relations by using econometric models and cross-spectral methods. Econometrica 37:424–438.

Greenhouse, J. B. 2009. Commentary: Cornfield, epidemiology, and causality. International Journal of Epidemiology 38: 1199–1201.

Guldemond, R. A. A., C. J. Louw, C. Mare, C. Norgaard, and R. J. van Aarde. 2022. Demographic responses of an insular elephant population to removal as a management intervention. Conservation Science and Practice 4:e12741.

Gutiérrez, R. J., M. Cody, S. Courtney, and A. B. Franklin. 2007. The invasion of barred owls and its potential effect on the spotted owl: a conservation conundrum. Biological Invasions 9:181–196.

Harre, R. 1972. The philosophies of science: an introductory survey. Oxford University Press, Oxford, United Kingdom.

Hempel, C. G. 1966. Philosophy of natural science. Prentice-Hall Inc., Englewood Cliffs, New Jersey, USA.

Hill, A. B. 1965. The environment and disease: association or causation? Proceedings of the Royal Society of Medicine 58: 295–300.

Holland, P. W. 1986. Statistics and causal inference. Journal of the American Statistical Association 81:945–960.

Hone, J. 1994. Analysis of vertebrate pest control. Cambridge University Press, Cambridge, United Kingdom.

Hone, J., V. A. Drake, and C. J. Krebs. 2017. The effort-outcomes relationship in applied ecology: evaluation and implications. BioScience 67:845–852.

Hone, J., V. A. Drake, and C. J. Krebs. 2023. Evaluation options for wildlife management and strengthening of causal inference. BioScience 73:48–58.

Howick, J., P. Glasziou, and J. K. Aronsen. 2009. The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? Journal of the Royal Society of Medicine 102:186–194.

Imbens, G. W. 2020. Potential outcome and directed acyclic graph approaches to causality: relevance for empirical practice in economics. Journal of Economic Literature 58:1129–1179.

Intergovernmental Panel on Climate Change [IPCC]. 2007. Intergovernmental Panel on Climate Change. Fourth Assessment Report. Summary for policy makers. IPCC, Geneva, Switzerland.

Intergovernmental Panel on Climate Change [IPCC]. 2021. Climate Change 2021. The physical science basis. Sixth Assessment Report. Summary for policy makers. IPCC, Geneva, Switzerland.

International Union for Conservation of Nature. 2022. The Green Status of Species. IUCNredlist.org/about/green-status-species. Accessed 2 Nov 2022.

Johnson, F. A., W. L. Kendall, and J. A. Dubovsky. 2002. Conditions and limitations on learning in the adaptive management of mallard harvests. Wildlife Society Bulletin 30:176–185.

Johnson, F. A., G. S. Boomer, B. K. Williams, J. D. Nichols, and D. J. Case. 2015. Multilevel learning in the adaptive management of waterfowl harvests: 20 years and counting. Wildlife Society Bulletin 39:9–19.

Kelly, E. G., E. D. Forsman, and R. G. Anthony. 2003. Are barred owls displacing spotted owls? Condor 105:45–53.

Kimmel, K., L. E. Dee, M. L. Avolio, and P. J. Ferraro. 2021. Causal assumptions and causal inference in ecological experiments. Trends in Ecology and Evolution 36:1141–1152.

King, G., and R. Nielsen. 2019. Why propensity scores should not be used for matching. Political Analysis 27:435–454.

Kodra, E., S. Chatterjee, and A. R. Ganguly. 2011. Exploring Granger causality between global average time series of carbon dioxide and temperature. Theoretical and Applied Climatology 104:325–335.

Krebs, C. J. 2009. Ecology. The experimental analysis of distribution and abundance. Sixth edition. Benjamin Cummings, New York, New York, USA.

Kunz, R., K. S. Khan, and H.-H. Neumayer. 2000. Observational studies and randomized trials. New England Journal of Medicine 343:1194–1195.

Lande, R. 1988. Demographic models of the northern spotted owl (*Strix occidentalis caurina*). Oecologia 75:601–607.

Larsen, A. E., K. Meng, and B. E. Kendall. 2019. Causal analysis in control-impact ecological studies with observational data. Methods in Ecology and Evolution 10:924–934.

Leopold, A. 1933. Game management. Charles Scribner's Sons, New York, New York, USA.

Lin, L., M. Sperrin, D. A. Jenkins, G. P. Martin, and N. Peek. 2021. A scoping review of causal methods enabling predictions under hypothetical interventions. Diagnostic and Prognostic Research 5:1–16.

Lipton, P. 2005. Testing hypotheses: predictions and prejudice. Science 307:219–221.

Lloyd, E. A. 2010. Confirmation and robustness of climate models. Philosophy of Science 77:971–984.

McArdle, B. H. 1996. Levels of evidence in studies of competition, predation, and disease. New Zealand Journal of Ecology 20:7–15.

Morgan, D. G., and P. Pegler. 2010. Managing a kangaroo population by culling to simulate predation: the Wyperfeld trial. Pages 349–359 *in* G. Coulson and M. Eldridge, editors. Macropods. The biology of kangaroos, wallabies and rat-kangaroos. CSIRO Publishing, Melbourne, Australia.

Mouquet, N., Y. Lagadeuc, V. Devictor, L. Doyen, A. Duputie, D. Eveillard, D. Faure, E. Garnier, O. Gimenez, P. Huneman, et al. 2015. Predictive ecology in a changing world. Journal of Applied Ecology 52:1293–1310.

Newton, I. 1998. Population limitation in birds. Academic Press, London, United Kingdom.

Nichols, J. D. 1991. Responses of North American duck populations to exploitation. Pages 498–525 *in* C. M. Perrins, J.–D. Lebreton, and G. J. M. Hirons, editors. Bird population studies. Relevance to conservation and management. Oxford University Press, Oxford, United Kingdom.

Nichols, J. D., F. A. Johnson, B. K. Williams, and G. S. Boomer. 2015. On formally integrating science and policy: walking the walk. Journal of Applied Ecology 52:539–543.

Nichols, J. D., W. L. Kendall, and G. S. Boomer. 2019. Accumulating evidence in ecology: once is not enough. Ecology and Evolution 9:13991–14004.

Nichols, S. J., M. Peat, and J. A. Webb. 2017. Challenges for evidence-based environmental management: what is sufficient and acceptable evidence for causation? Freshwater Science 36:240–249.

Norris, R. H., J. A. Webb, S. J. Nichols, M. J. Stewardson, and E. T. Harrison. 2012. Analyzing cause and effect in environmental assessments: using weighted evidence from the literature. Freshwater Science 31:5–21.

Pavlov, P. M., and J. Hone. 1982. The behaviour of feral pigs, *Sus scrofa*, in flocks of lambing ewes. Australian Wildlife Research 9:101–109.

Pavlov, P. M., J. Hone, R. J. Kilgour, and H. Pedersen. 1981. Predation by feral pigs on Merino lambs at Nyngan, New South Wales. Australian Journal of Experimental Agriculture and Animal Husbandry 21:570–574.

Pearl, J. 2009*a*. Causal inference in statistics: an overview. Statistics Surveys 3:96–146.

Pearl, J. 2009*b*. Causality. Models, reasoning and inference. Second edition. Cambridge University Press, Cambridge, United Kingdom.

Pearl, J., and D. Mackenzie. 2018. The book of why. The new science of cause and effect. Penguin, London, United Kingdom.

Plant, J. W., R. Marchant, T. D. Mitchell, and J. R. Giles. 1978. Neonatal lamb losses due to feral pig predation. Australian Veterinary Journal 54:426–429.

Platt, J. R. 1964. Strong inference. Science 146:347–353.

Plowright, R. K., S. H. Sokolow, M. E. Gorman, P. Daszak, and J. E. Foley. 2008. Causal inference in disease ecology: investigating ecological drivers of disease emergence. Frontiers in Ecology and Environment 6:420–429.

Pocock, S. J., and D. R. Elbourne. 2000. Randomized trials or observational tribulations? New England Journal of Medicine 342:1907–1909.

Popper, K. R. 1965. Conjectures and refutations: the growth of scientific knowledge. Harper & Row, New York, New York, USA.

Ramsey, D. S. L., D. M. Forsyth, E. Wright, M. McKay, and I. Westbrooke. 2019. Using propensity scores for causal inference in ecology: options, considerations, and a case study. Methods in Ecology and Evolution 10:320–331.

Rehme, S. E., L. A. Powell, and C. R. Allen. 2011. Multimodel inference and adaptive management. Journal of Environmental Management 92:1360–1364.

Reynolds, R. E., and J. R. Sauer. 1991. Changes in mallard breeding populations in relation to production and harvest rates. Journal of Wildlife Management 55:483–487.

Riley, R. D., J. Ensor, K. I. E. Snell, T. P. A. Debray, D. G. Altman, K. G. M. Moons, and G. S. Collins. 2016. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. British Medical Journal 353:i3140.

Romesburg, H. C. 1981. Wildlife science: gaining reliable knowledge. Journal of Wildlife Management 45:293–313.

Rowley, I. 1970. Lamb predation in Australia: incidence, predisposing conditions, and the identification of wounds. CSIRO Wildlife Research 15:79–123.

Rubin, D. B. 2005. Causal inference using potential outcomes: design, modeling, decisions. Journal of the American Statistical Association 100:322–331.

Runge, J., P. Nowack, M. Kretschmer, and D. Sejdinovic. 2019. Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances 5:eaau4996.

Russo, F., and J. Williamson. 2007. Interpreting causality in the health sciences. International Studies in the Philosophy of Science 21:157–170.

Rytwinski, T., J. J. Taylor, L. A. Donaldson, J. R. Britton, D. R. Browne, R. E. Gresswell, M. Lintermans, K. A. Prior, M. G. Pellatt, C. Vis, and S. J. Cooke. 2019. The effectiveness of non-native fish removal techniques in freshwater ecosystems: a systematic review. Environmental Reviews 27:91–94.

Schoolmaster, D. R., C. R. Zirbel, and J. P. Cronin. 2020. A graphical model for resolving species identity effects and biodiversity-ecosystem function correlations. Ecology 101:e03070.

Sells, S. N., S. B. Bassing, K. J. Barker, S. C. Forshee, A. C. Keever, J. W. Goerz, and M. S. Mitchell. 2018. Increased scientific rigor will improve reliability of research and effectiveness of management. Journal of Wildlife Management 82:485–494.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin, Boston, Massachusetts, USA.

Shipley, B. 2009. Confirmatory path analysis in a generalized multilevel context. Ecology 90:363–368.

Smith, F. A., J. H. Brown, and T. J. Valone. 1997. Path analysis: a critical evaluation using long-term experimental data. American Naturalist 149:29–42.

Smith, F. A., J. H. Brown, and T. J. Valone. 1998. Path modeling methods and ecological interactions: a response to Grace and Pugesk. American Naturalist 152:160–161.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods. Sixth edition. Iowa State University Press, Ames, USA.

Stewart, P. S., P. A. Stephens, R. A. Hill, M. J. Whittingham, and W. Dawson. 2023. Model selection in occupancy models: inference versus causation. Ecology 104:e3942.

Sugihara, G., R. May, H. Ye, C.-H. Hsieh, E. Deyle, M. Fogarty, and S. Munch. 2012. Detecting causality in complex ecosystems. Science 338:496–500.

Susser, M. 1991. What is a cause and how do we know one? A grammar for pragmatic epidemiology. American Journal of Epidemiology 133:635–648.

Suter, G. W., S. B. Norton, and S. M. Cormier. 2010. The science and philosophy of a method for assessing environmental causes. Human and Ecological Risk Assessment 16:19–34.

Ushio, M., C.-H. Hsieh, R. Masuda, E. R. Deyle, H. Ye, C.-W. Chang, G. Sugihara, and M. Kondoh. 2018. Fluctuating interaction network and time-varying stability of a natural fish community. Nature 554:360–363.

Ware, J. H. 1989. Investigating therapies of potentially great benefit: ECMO. Statistical Science 4:298–306.

Weed, D. L. 1997. On the use of causal criteria. International Journal of Epidemiology 26:1137–1141.

Wiens, J. D., K. M. Dugger, J. M. Higley, D. B. Lesmeister, A. B. Franklin, K. A. Hamm, G. C. White, K. E. Dillone, D. C. Simon, R. B. Bow, et al. 2021. Invader removal triggers competitive release in a threatened avian predator. Proceedings of the National Academy of Sciences USA 118:e2102859118.

Williams, B. K. 1997. Logic and science in wildlife biology. Journal of Wildlife Management 61:1007–1015.

Williams, B. K., F. A. Johnson, and K. Wilkins. 1996. Uncertainty and the adaptive management of waterfowl harvests. Journal of Wildlife Management 60:223–232.

Wright, S. 1921. Correlation and causation. Journal of Agricultural Research 20:557–585.

Zhao, Q., E. Silverman, K. Fleming, and G. S. Boomer. 2016. Forecasting waterfowl population dynamics under climate change–does the spatial variation of density-dependence and environmental effects matter? Biological Conservation 194:80–88.

*Associate Editor: Quresh Latif*.

**How to cite this article:** Hone, J., and C. J. Krebs. 2023. Causality and wildlife management. Journal of Wildlife Management e22412. https://doi.org/10.1002/jwmg.22412