



Trade-offs in experimental designs for estimating post-release mortality in containment studies



Mark W. Rogers^{a,*}, Andrew B. Barbour^b, Kyle L. Wilson^c

^a U.S. Geological Survey, Great Lakes Science Center, Lake Erie Biological Station, 6100 Columbus Avenue, Sandusky, OH 44870, United States

^b School of Forest Resources and Conservation, University of Florida, 7922 NW 71st Street, Gainesville, FL 32653, United States

^c Department of Biological Sciences, University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N 1N4, Canada

ARTICLE INFO

Article history:

Received 13 May 2013

Received in revised form 23 October 2013

Accepted 25 October 2013

Keywords:

Tagging

Containment

Post-release mortality

Experimental design

ABSTRACT

Estimates of post-release mortality (PRM) facilitate accounting for unintended deaths from fishery activities and contribute to development of fishery regulations and harvest quotas. The most popular method for estimating PRM employs containers for comparing control and treatment fish, yet guidance for experimental design of PRM studies with containers is lacking. We used simulations to evaluate trade-offs in the number of containers (replicates) employed versus the number of fish-per container when estimating tagging mortality. We also investigated effects of control fish survival and how among container variation in survival affects the ability to detect additive mortality. Simulations revealed that high experimental effort was required when: (1) additive treatment mortality was small, (2) control fish mortality was non-negligible, and (3) among container variability in control fish mortality exceeded 10% of the mean. We provided programming code to allow investigators to compare alternative designs for their individual scenarios and expose trade-offs among experimental design options. Results from our simulations and simulation code will help investigators develop efficient PRM experimental designs for precise mortality assessment.

Published by Elsevier B.V.

1. Introduction

Numerous studies have sought to estimate post-release effects of fishery or research activities on fish that are otherwise presumed to survive. Multiple examples of post-release mortality (PRM) estimation come from “catch-and-release science” (Cooke and Schramm, 2007), accounting for tagging mortality in mark-and-recapture studies (e.g., Brenden et al., 2010), and incorporating parameters for “cryptic mortality” into population models (e.g., Coggins et al., 2007). Freshwater catch-and-release (CR) mortality studies have been a large contributor to understanding handling effects. For example, Hühn and Arlinghaus (2011) identified 107 studies with 252 individual CR mortality estimates in a meta-analysis limited to genera popular in European fisheries. Efforts aimed at decreasing unintended mortality (e.g., circle hook requirements) and accounting for unintended mortality (e.g., tagging studies, fishery bycatch) necessitate robust PRM experimental designs.

Pollock and Pine (2007) provided guidance for estimating mortality and uncertainty in CR studies and discussed trade-offs among approaches (e.g., containment versus telemetry). For the most com-

mon method, containment studies (Cooke and Schramm, 2007; Raby et al., in press), Pollock and Pine (2007) recommended that: (1) control fish always be used, (2) individual fish not be treated as replicates, and (3) precision be considered before experimentation to determine if estimates would be informative. In studies with randomized designs, container-to-container effects can bias estimates of mortality for tagged compared with control containers. To account for the potential for container-specific effects, Pollock and Pine (2007) encouraged the use of stocking equal numbers of control and treatment fish in each container (hereafter called ‘paired-design’); thus any container-to-container effects will be equally distributed between control and tagged fish. Pollock and Pine (2007) focused on CR studies, yet their recommendations extend to other PRM studies with objectives of quantifying and comparing treatment effects.

Cowx et al. (2010) stated that a “plethora” of studies have aimed to measure lethal and sublethal handling effects. However, we are not aware of any in-depth exploration of Pollock and Pine’s (2007) warning for careful planning and study design. Cooke et al. (2013) provided perspectives on measuring physiological consequences of CR fishing and common limitations, yet gaps still remain for quantifying treatment effects. For fish tagging studies, in particular, guidance exists for tag selection and ensuring that the marked population is representative of the unmarked population (e.g., Pine et al., 2003) and mark-recapture literature explicitly details the

* Corresponding author. Tel.: +1 419 625 1976; fax: +1 419 625 7164.

E-mail address: mwrogers@usgs.gov (M.W. Rogers).

need and methods for accurate estimates of capture probability. Yet, methods for quantifying mortality due to handling and tagging have received less attention and could result in bias in fish tagging studies if the number-at-large is not adequately assessed. We were challenged with designing an experiment to measure PIT-tagging mortality for a yellow perch movement study at Lake Erie, and wanted to extend our result to create a broadly applicable simulation tool to inform future PRM study designs. Our objective was to use simulations to evaluate trade-offs in the ability to detect mortality in paired-design containment experiments across a range of design scenarios in which sample size (number of fish per container) and replication (number of containers) were independent variables. We also evaluated the effect of control-fish survival and among-container survival variability on the precision of mortality estimates. Lastly, we provided Program R code (R Development Core Team, 2012) that can be modified to address other scenarios.

2. Materials and methods

Our simulations followed a recommended experimental design with control fish (handled) and treatment fish (handled and tagged) that were then held together within containers (Pollock and Pine, 2007). We simulated survival for individuals from each treatment group as a binomial process. Each individual had a probability of surviving determined by treatment-dependent handling and tagging mortality inputs. The number of tagged and untagged fish in each enclosure were equal, thus the number of total fish per container had to be an even number. We chose to use a paired-design to ensure container-specific effects would not bias estimates of mortality between control and tagged fish. The number of containers was specified and individual containers were treated as replicates. Tag survival was estimated for each group as the difference between handling survival and tagging mortality. We calculated standard errors of the mean tagging survival for each container to determine 95% confidence intervals. To determine the average expectation for each scenario, we repeated simulations 1000 times.

Our simulation inputs were founded on a PIT-tagging mortality study for juvenile bluegill and yellow perch (Kaemingk et al., 2011). Our baseline handling survival was 95% for our simulations. We simulated a range of additive tagging mortality from 5% to 20% in increments of 5%. The number of fish in each simulated container ranged from 10 to 100 by 2, equally divided between the two treatments (i.e., 10 fish per container equated to 5 tagged and 5 control fish). We simulated a range of containers from 2 to 25. Our evaluation metrics included (1) relative bias in tagging mortality estimates relative to known input value, (2) average standard error in tagging mortality estimate, (3) percent coverage of tagging mortality estimates, and (4) the mean p -value from simulation replications for an analysis of variance (ANOVA) that tested for differences in survival rates between control and tagged fish with containers as replicates for each iteration. We calculated relative bias using the formula:

$$(\hat{M}_t - M_t) \times M_t^{-1} \quad (1)$$

where $\hat{M}_t$ is estimated tagging mortality for the i th iteration and M_t is the true mortality level. We then derived the average of the relative departures for all iterations for a given parameter scenario and reported that value as bias. We determined percent coverage as the number of iterations during which the 95% confidence interval of estimated tagging mortality included the true value for a given parameter set. We applied these outputs to select the design for our yellow perch scenario and simulated data under that design to determine if the data conformed to ANOVA assumptions. We assessed homoscedasticity assumptions with Levene's test and

normality assumptions with a Shapiro–Wilk test from 10,000 iterations (sensu Cooke et al., 2003).

We evaluated the sensitivity of simulation metrics to changes in control fish survival and among container variation in control fish survival. For sensitivity of designs to control fish survival, we compared a baseline scenario where handling survival was 95% to a scenario where handling survival was 80% with additive tagging mortality ranging from 5% to 20% in increments of 5% for both. To evaluate the effect of among container variation, we used a scenario with 10 replicates (i.e., containers), ranging from 12 to 200 fish-per-container in increments of 4, and a mean handling survival of 95%. We applied random deviates to the mean handling survival for each container. Random deviates were drawn from a normal distribution with standard deviations ranging from 0% to 25% of mean handling survival in increments of 2%. Simulation output metrics included mean relative bias, mean percent coverage, and mean ANOVA p -values as described above.

3. Results

Our simulations allowed for direct comparison of trade-offs in number of containers (i.e., replicates) versus total number of fish per container for informing PRM experimental designs. As expected, experiments with few individual fish and little replication were most biased. In some cases with few fish, the binomial process resulted in a higher number of fish dying from handling than tagging despite our known inputs with increased mortality from tagging. In experiments with low numbers of replicates, increasing the number of fish per container was marginally effective for improving bias (Fig. 1). As tagging mortality increased, bias and standard error estimates improved at lower numbers of replicates and fish per container relative to low tagging mortality (Figs. 1 and 2). For all tagging mortality values we evaluated, acceptable coverage levels (i.e., $\geq 90\%$) were only obtained with >6 containers (Fig. 3). The number of fish per container had little effect on coverage (Fig. 3) thereby requiring the use of >6 replicates when estimating tagging mortality under this survival scenario. Experiments with low tagging mortality and few replicates had a low ability to detect tagging effects with ANOVA. For our yellow perch scenario, we concluded that 8 replicates with 100 fish per container would adequately allow us to detect additive tagging mortality ≥ 0.05 (Fig. 4). This choice was informed by simulation outputs as well as considerations for the number of containers available ($N=4$) and the expected tagging period (one week). Shapiro–Wilk and Levene's tests to assess violations of normality and homoscedastic assumptions for the ANOVA determined that normality and homoscedastic violations occurred in only 13% and 6%, respectively, of 10,000 simulations in the yellow perch scenario.

Proper experimental design was dependent upon mean control fish survival and variability of mean control fish survival across containers. Lower handling survival of control fish resulted in a need for more experimental effort to maintain precision. For example, the ability to detect a significant difference between handling and tagging mortality when alpha was equal to 0.10 could be achieved when handling survival was 95% and tagging mortality was 5% by using 8 containers and 100 fish per container, but required 7 additional containers when handling survival decreased to 80% with 5% tagging mortality. When the standard deviation around the mean survival for control fish was higher than 10%, average ANOVA p -values exceeded 0.20 in simulations with 100 fish per container and 10 replicates (Fig. 5). Therefore, experimental design requires an expectation of what baseline handling mortality may be and careful consideration to minimize among container variation.

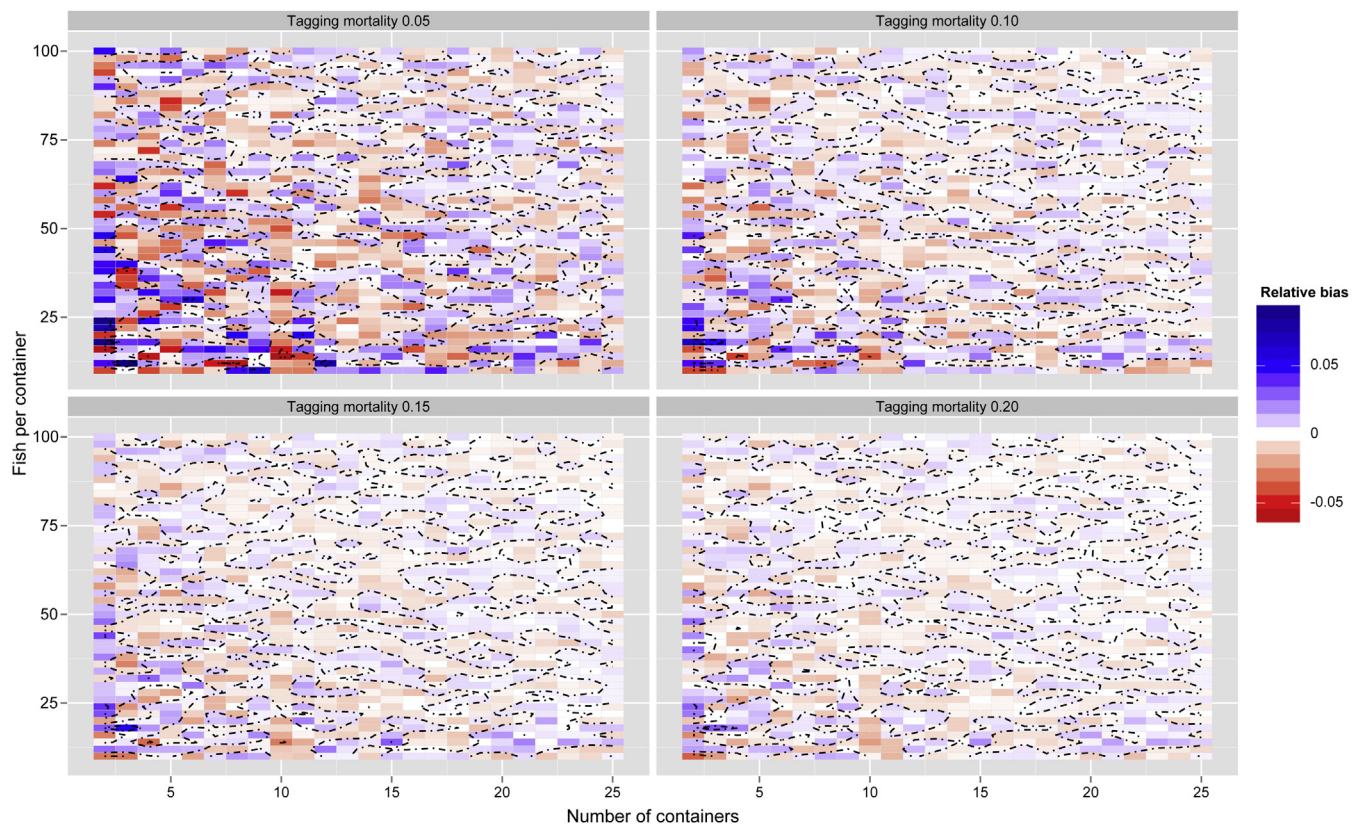


Fig. 1. Simulated relative bias in tagging mortality estimates relative to input value (panels) across a range of containers and fish per container when handling survival = 0.95.

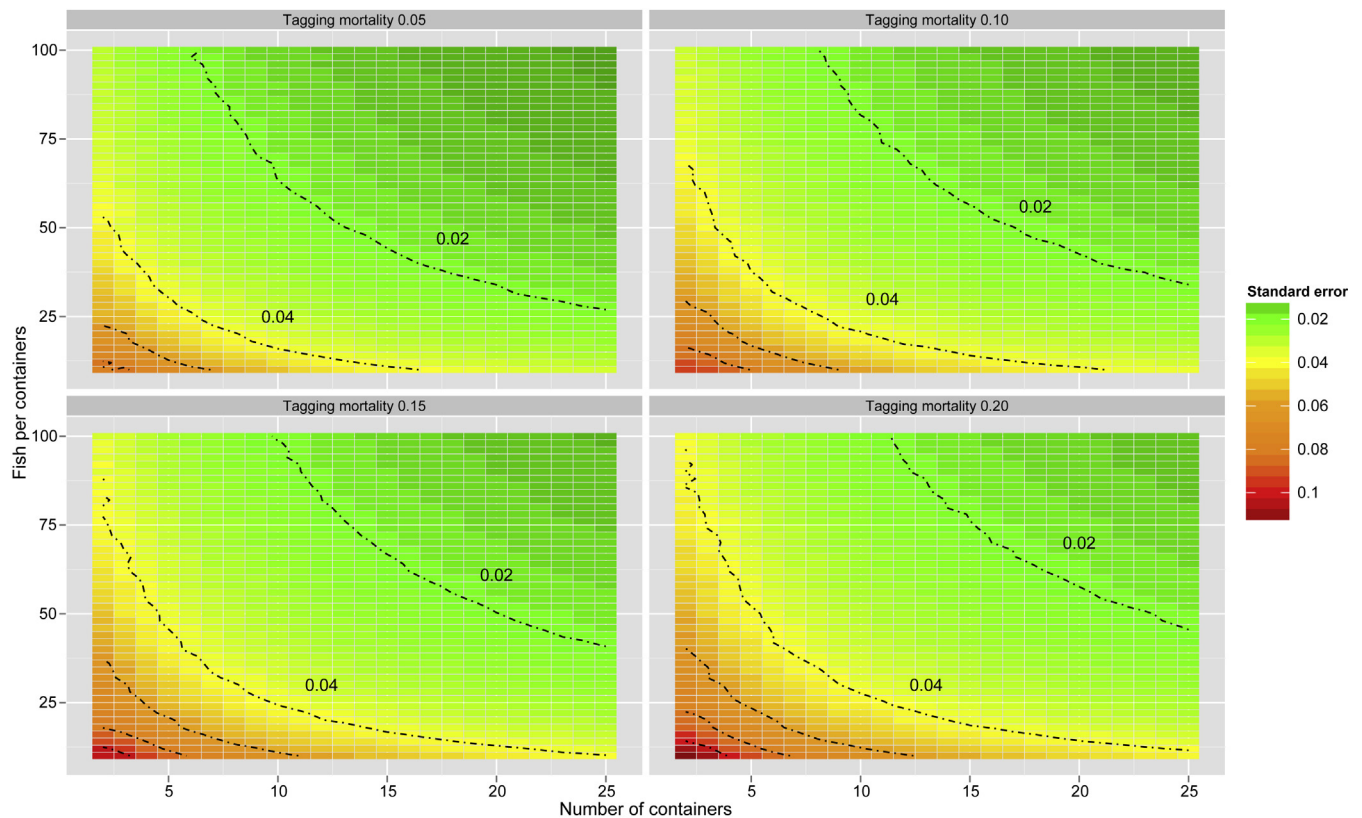


Fig. 2. Standard error estimates across a range of containers and fish per container when handling survival = 0.95.

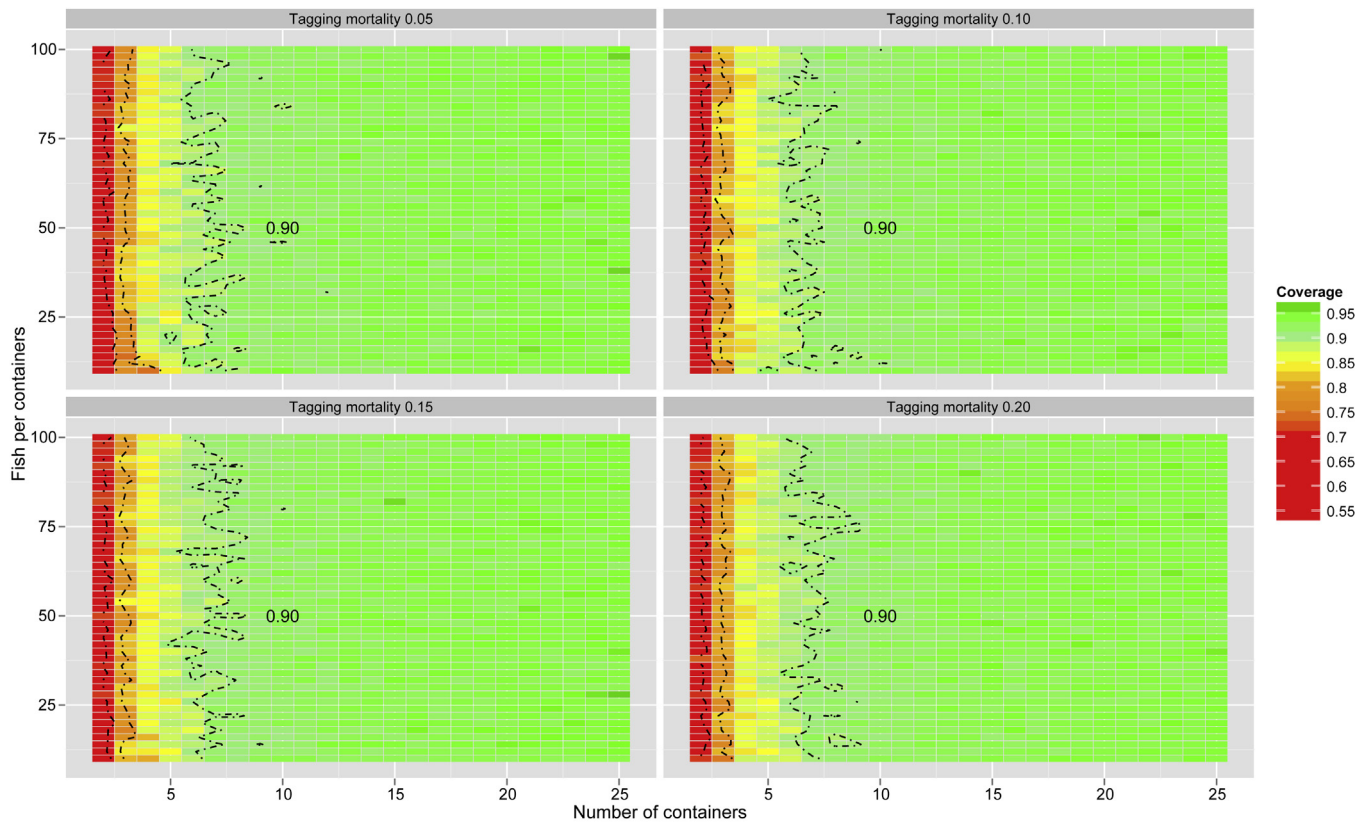


Fig. 3. Percent coverage of estimated 95% confidence intervals for tagging mortality across a range of containers and fish per container when handling survival = 95%.

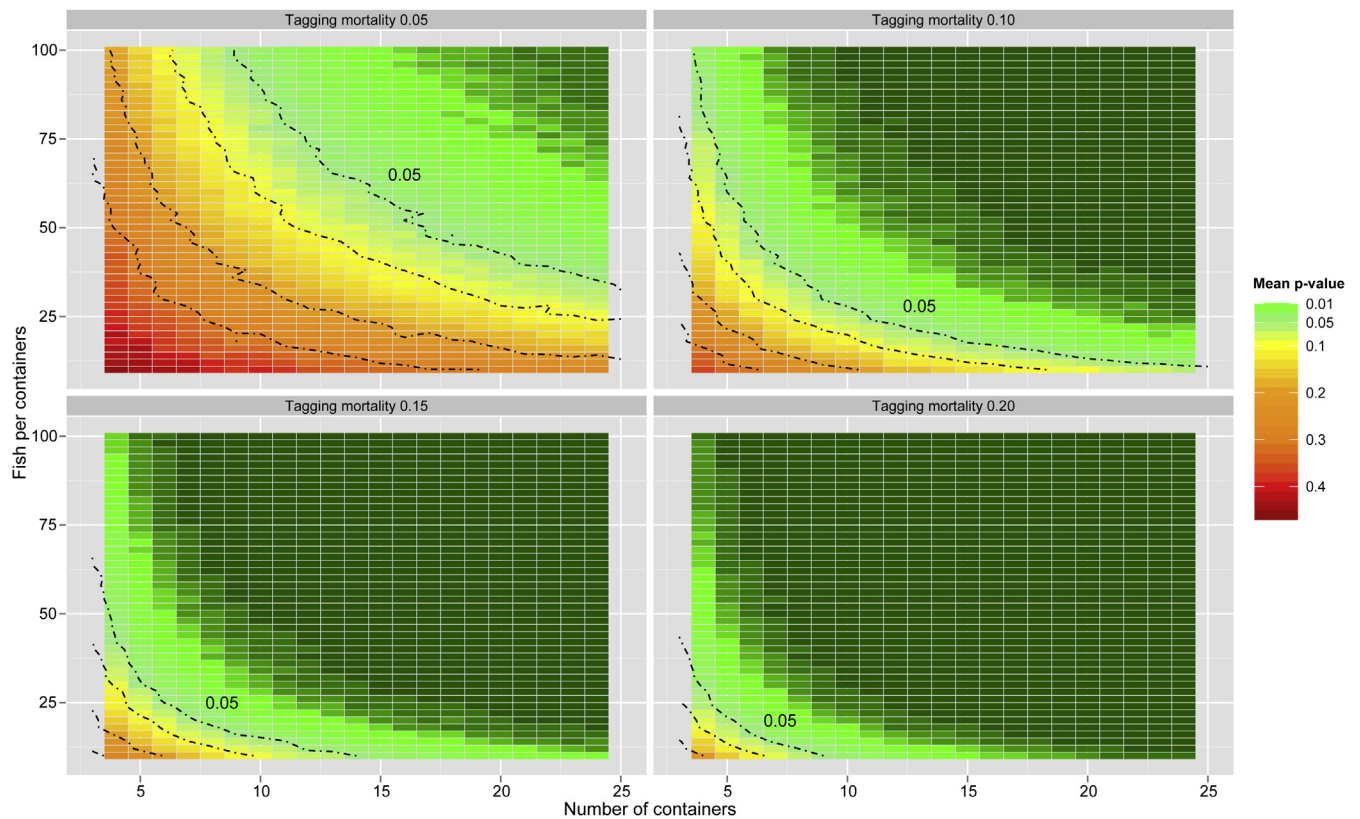


Fig. 4. Average p -value resulting from analysis of variance tests for differences in mortality between control and treatment fish across a range of containers and fish per container when handling survival = 0.95.

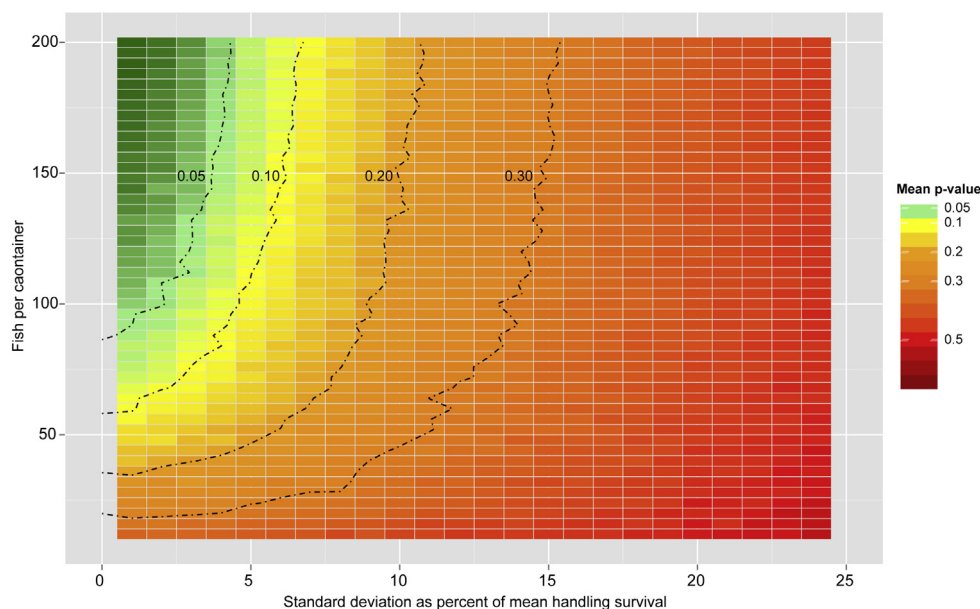


Fig. 5. Effect of among-container variability on the average p -value resulting from analysis of variance tests for difference in mortality between control and treatment fish across a range of fish per container using ten container replicates and handling survival = 95%.

4. Discussion

Developing a tool to assess trade-offs in the number of containers employed and the number of fish per container revealed the ability to maximize design efficiency relative to desired precision. Our simulations confirmed our expectation that increased numbers of replicates (i.e., containers) and fish per container would provide less biased and more precise estimates of tagging mortality. Much of these results are statistically intuitive in nature, but prior PRM investigations had not formally quantified the trade-offs in containment study designs that could be considered along with logistical limitations (e.g., numbers of containers available, number of fish that can be collected or held) in PRM studies.

Trade-offs in study design were largely driven by the interaction of control and handling survival rates, where some values of these vital rates affected statistical trends or assumptions. For an expected handling survival, a more intensive experimental effort was required as additive tagging mortality decreased. For fisheries management evaluations, the ability to detect strong post-release effects at lower experimental effort may be advantageous because the most deleterious treatments are the most easily detected. In contrast, research studies that strive to impose minimal post-release effects would require more intensive efforts to quantify treatment effects. The ANOVA assumptions of normality and homogenous variance were infrequently violated during our yellow perch scenario evaluation (5–13% of iterations). Violations of normality mainly occurred when control fish survival was relatively high (~95%), thereby resulting in a skewed distribution in survival due to limitations caused by the upper bound (i.e., 100%) of survival. Our usage of the mean p -value from the ANOVA illustrated on average the level of significance between differences in mortality between tagged and control fish and this metric reflects statistical power among experimental designs (see [Hoenig and Heisey, 2001](#)). Investigators also could consider outputting the maximum P value (i.e., rather than the mean) in the supplemental code to evaluate very conservative estimates of detectable differences.

Sensitivity of control-fish survival to estimating treatment mortality has not been evaluated, but our simulations showed it can highly influence estimation. Multiple authors have pointed out that true controls are nearly impossible to obtain for PRM studies

because there is always some unnatural handling, transport, or containment effects in experiments. Collection of control fish usually employs a gear that is believed to be innocuous, but some species may require collection with a riskier gear (e.g., short-term gill net sets) or experiments may need to take place at times when environmental conditions (e.g., water temperature) may not be optimal for survival of control fish. Strong influences of control fish survival on tagging-mortality estimation suggested a need for an a priori expectation of handling survival. Evaluations of among-container variability in handling survival highlighted the importance of controlling for container-specific influences. We only evaluated effects of random variation among containers, but suggest that investigators consider the potential for systematic variation (e.g., effects across a nearshore to offshore gradient) that could occur in some experiment settings.

The design of PRM studies includes many more logistical considerations than those included in our simulations. For example, the duration of PRM experiments can include immediate, short-term, and long-term periods ([Pollock and Pine, 2007](#)). For our yellow perch experiment at Lake Erie, potential for high wave events will likely limit us to short-term (i.e., ≤ 72 h) periods. One limitation to short-term experiments on PRM mortality is that results are often treated as finite rates by assuming that all tagging mortality was accounted for within the evaluation experimental period, whereas other methods (e.g., telemetry tags) may provide estimates that are more representative of total mortality from handling. Species-specific characteristics may also limit options for experimental design. For example, [Schmitt and Shoup \(2013\)](#) evaluated hooking mortality of blue catfish *Ictalurus furcatus* and used one treatment and one control fish, similar in size, per container due to their potential aggressiveness. Container and species-specific characteristics may also limit the number of fish per container before crowding effects become a concern. However our simulation program can incorporate these considerations by allowing adjustment of input ranges of fish per container and number of containers to facilitate design decisions.

[Pollock and Pine \(2007\)](#) concluded that CR studies are easy to design. We argue that they are easy to conceive, but tools for explicit evaluation of alternative designs were lacking. Although our simulations focused on the most popular design for PRM studies

(i.e., container experiments), we recognize that estimates resulting from these experiments may be an underestimate owing to the unnatural treatment of experimental fish in natural settings. Barbour et al. (2012) documented a significant effect of pulsed gastric lavage on apparent survival in a natural setting despite no evidence of an effect in earlier caging trials, but could not definitively differentiate lethal (predation) from sublethal (emigration) effects. Raby et al. (in press) highlighted that post-release predation, which is not measured in containment experiments, has largely been ignored. Despite this deficit, we expect that PRM studies employing containers will continue to be important for directing fisheries practices that consider handling and post-release effects (e.g., development of a global code of practice, see Arlinghaus et al., 2010). We encourage the continued exploration of alternative study designs to maximize experimental efficiency for producing precise PRM estimates.

Acknowledgements

We thank Joe Hightower of the U.S. Geological Survey and two anonymous reviewers for comments that helped improve this article. ABB was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-E802270. This article is Contribution 1805 of the United States Geological Survey Great Lakes Science Center. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fishres.2013.10.020>.

References

- Arlinghaus, R., Cooke, S.J., Cowx, I.G., 2010. Providing context to the global code of practice for recreational fisheries. *Fish. Manage. Ecol.* 17, 146–156.
- Barbour, A.B., Boucek, R.E., Adams, A.J., 2012. Effect of pulsed gastric lavage on apparent survival of a juvenile fish in a natural system. *J. Exp. Mar. Biol. Ecol.* 422/423, 107–113.
- Brenden, T.O., Jones, M.L., Ebener, M.P., 2010. Sensitivity of tag-recovery mortality estimates to inaccuracies in tag shedding, handling mortality, and tag reporting. *J. Great Lakes Res.* 36 (2010), 100–109.
- Coggins Jr., L.C., Catalano, M.J., Allen, M.S., Pine III, W.E., Walters, C.J., 2007. Effects of cryptic mortality and the hidden costs of length limits in fishery management. *Fish. Fish.* 8, 196–210.
- Cooke, S.J., Schramm, H.L., 2007. Catch-and-release science and its application to conservation and management of recreational fisheries. *Fish. Manage. Ecol.* 14, 73–79.
- Cooke, S.J., Suski, C.D., Siepker, M.J., Ostrand, K.G., 2003. Injury rates, hooking efficiency and mortality potential of largemouth bass (*Micropterus salmoides*) captured on circle hooks and octopus hooks. *Fish. Res.* 61, 135–144.
- Cooke, S.J., Raby, G.D., Donaldson, M.R., Hinch, S.G., O'Connor, C.M., Arlinghaus, R., Danylchuk, A.J., Hanson, K.C., Clark, T.D., Patterson, D.A., 2013. The physiological consequences of catch-and-release angling: perspectives on experimental design, interpretation, extrapolation, and relevance to stakeholders. *Fish. Manage. Ecol.* 20, 268–287.
- Cowx, I.G., Arlinghaus, R., Cooke, S.J., 2010. Harmonizing recreational fisheries and conservation objectives for aquatic biodiversity in inland waters. *J. Fish Biol.* 76, 2194–2215.
- Hoening, J.M., Heisey, D.M., 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. *Am. Statistician* 55, 1–6.
- Hühn, D., Arlinghaus, R., 2011. Determinants of hooking mortality freshwater recreational fisheries: a quantitative meta-analysis. In: Beard Jr., T.D., Arlinghaus, R., Sutton, S.G. (Eds.), *The Angler in the Environment: Social, Economic, Biological, and Ethical Dimensions*. Proceedings of the Fifth World Recreational Fishing Conference, Symposium 75. American Fisheries Society, Bethesda, Maryland, pp. 141–170.
- Kaemingk, M.A., Weber, M.J., McKenna, P.R., Brown, M.L., 2011. Effect of passive integrated transponder tag implantation site on tag retention, growth, and survival of two sizes of juvenile bluegills and yellow perch. *N. Am. J. Fish. Manage.* 31, 726–732.
- Pine, W.E., Pollock, K.H., Hightower, J.E., Kwak, T.J., Rice, J.A., 2003. A review of tagging methods for estimating fish population size and components of mortality. *Fisheries* 28, 10–23.
- Pollock, K.H., Pine III, W.E., 2007. The design and analysis of field studies to estimate catch-and-release mortality. *Fish. Manage. Ecol.* 14, 1–8.
- R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Raby G.D., Packer J.R., Danylchuk A.J., Cooke S.J. The understudied and underappreciated role of predation in the mortality of fish released from fishing gears. *Fish. Fish.*, <http://dx.doi.org/10.1111/faf.12033> (in press).
- Schmitt, J.D., Shoup, D.E., 2013. Delayed hooking mortality of blue catfish caught on juglines. *N. Am. J. Fish. Manage.* 33, 245–252.