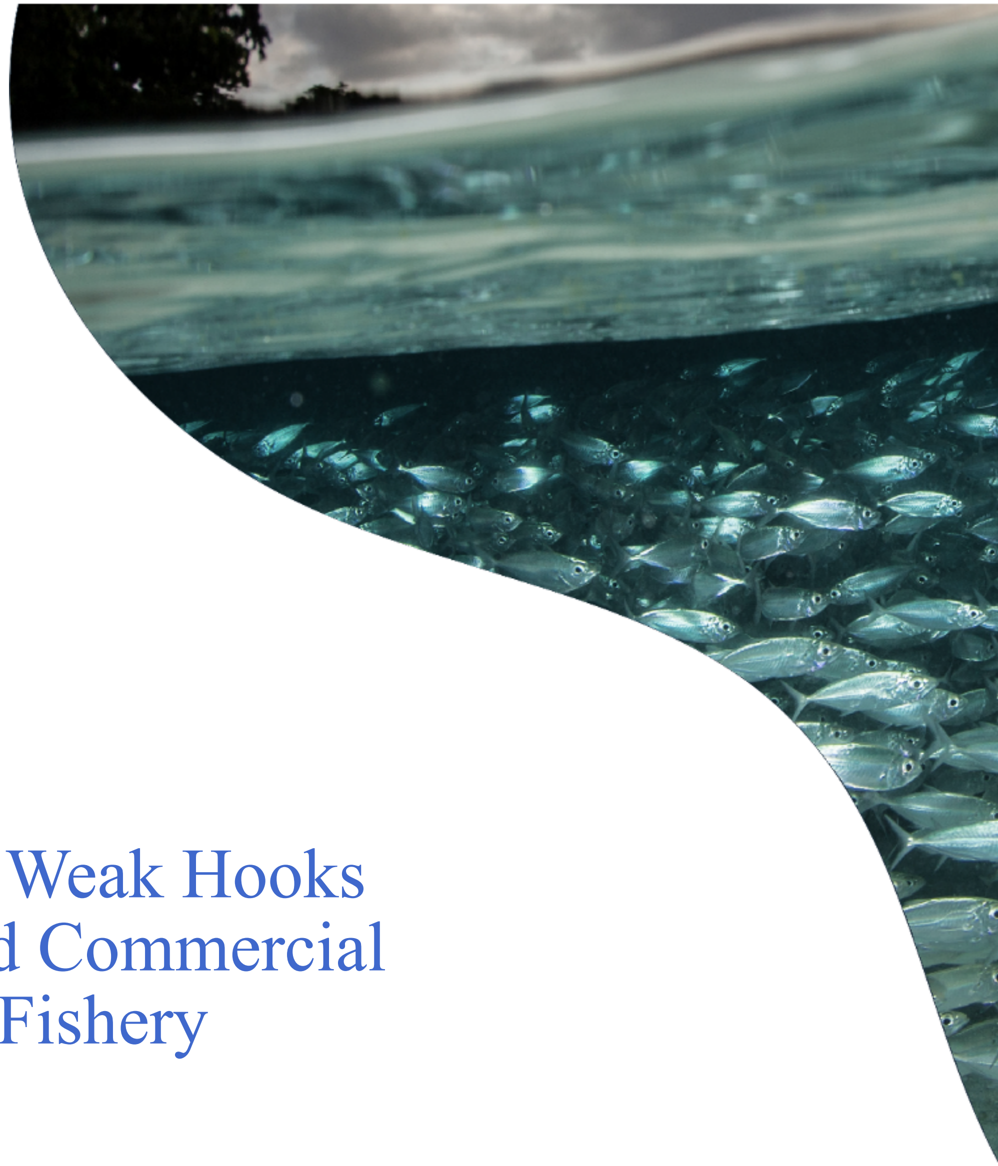
6.A(1)
143rd SSC



FINAL REPORT

Catch Retention of Weak Hooks in the Hawaii-based Commercial Deep-set Longline Fishery

Michael K. Musyl[1*] and Edward Phillips[2]

[1]Pelagic Research Group LLC
P.O. Box 10243, Honolulu, HI 96816 USA

[2]Lynker Technologies
Pacific Region Office
Honolulu, HI 96817 USA

*corresponding author: michael.musyl@gmail.com

TABLE OF CONTENTS

Abstract

To address concerns about serious injury and mortality outcomes of toothed whales (i.e., odontocetes) depredating catch and bait on pelagic longlines and becoming hooked, various mitigation strategies and devices have been developed to reduce the severity of injuries and improve survival outcomes. Potentially, weak circle hooks are a cost-effective solution that are designed to straighten or unbend when cetaceans (or large body mass species) become hooked to free themselves. Essentially the concept is to make the hook the weakest link in the system to take advantage of the tensile force exerted by large species. The concept of weak hooks in fisheries is simple and potentially cost-effective if it can be proven to be operational. It is important to the fishery that the use of the weak hooks must be demonstrated to maintain catch levels of target and marketable species. Weak circle hooks (15/0, 10° offset, 4.2 mm ø) were compared to strong (control) circle hooks (15/0, 10° offset, 4.5 mm ø) in the Hawaii-based commercial deep-set longline fishery targeting bigeye tuna (*Thunnus obesus*) to examine whether bycatch rates, significant injuries and mortality of false killer whales (*Pseudorca crassidens*) could be reduced whilst at the same time maintaining catch rates of target and marketable species. In the present report, we focus on investigating catch rates of target and marketable species on the two hook types. A previous weak hook trial occurred in the Hawaii deep-set longline fishery in 2010 during a seasonal period when larger bigeye tuna are historically absent from the fishery. We corrected this deficiency and replicated the earlier study by specifically sampling larger bigeye tuna from 178 longline sets and could show similar catch rates and body sizes on the two hook types. Though not statistically significant, catch risk of bigeye tuna was higher on weak hooks but mean body length (3.3 cm) and dressed weight (6.8 lb or 3.1 kg) was significantly larger and heavier on strong hooks, respectively. Bigeye caught on strong hooks also fetched a significantly higher mean price per fish at auction ($52.89) but the analysis did not take into account exogenous (i.e., time spent hooked, temperature, dissolved oxygen) and endogenous factors (e.g., stress, parasites, shark damage, bad gaff placement) known to influence flesh quality. Using a meta-analytic approach synthesizing effect sizes on catch rates, body sizes, dressed weights and prices for species at auction, we demonstrated *ex*-vessel revenue was virtually similar on the two hook types.

Keywords: bycatch, cetaceans, common language effect size, depredation, effect size, false killer whale, mean difference, mitigation, ROC, risk ratio

1 Introduction

Globally, odontocetes (i.e., toothed whales) are problematic bycatch in many commercial and artisanal fisheries, including trawl, gillnets, pots/traps and longline gear (Read et al. 2006; Forney et al. 2011; Gilman 2011; Hamer et al 2012; Werner et al. 2015; Hamilton and Baker 2019; Swimmer et al. 2020; Fader et al. 2021a,b). In particular, there is concern about odontocetes depredating the catch and bait on pelagic longline gear and becoming hooked causing serious injuries or even death (Forney et al. 2011; Hamilton and Baker 2019; Fader et al. 2021a,b). As importantly to stakeholders in the fishery, Gilman (2011), Hamer et al. (2012) and Patrick and Benaka (2013) report on the possible loss of revenue of longline fisheries caused by depredation; by missing depredation events, this could jeopardize management of fisheries for target species by under reporting CPUE (Gilman 2011). Consequently, there is biological and economic incentive for stakeholders and management to reduce depredation. At present, there are several possible solutions to reduce odontocete interactions and depredation on longlines (Werner et al. 2006, 2015; Hamer et al. 2012; Hamilton and Baker 2019; Swimmer et al. 2020), but the single most efficacious solution remains in strategies (e.g., "move on")(Hamilton and Baker 2019), although not necessarily the most cost-effective (Forney et al. 2011).

Designing cost-effective bycatch mitigation devices and/or strategies in commercial fisheries is challenging due to competing interests and tradeoffs (e.g., Meyer et al. 2017; Gilman et al. 2019). One promising cost-effective and simple tool is the use of weak circle hooks in longline fisheries to reduce bycatch and mortality outcomes whilst maintaining catch rates for target and retained species (Foster and Bergmann 2010, 2012a,b; Bayse and Kerstetter 2010; Bigelow et al. 2012). Weak circle hooks are designed of a certain tensile strength so that the hook is the weakest part of the gear and large bycaught species of sufficient body mass can free themselves from capture by straightening or unbending the hook. Weak hooks have been successfully applied in the U.S pelagic longline fishery in the Gulf of Mexico (GOM) where mitigation efforts have focused on reducing the catch of spawning size class bluefin tuna (*Thunnus thynnus*). Initial trials reported reductions in bluefin tuna catch rates from 46% to 57% switching to weak hooks (Foster and Bergmann 2010, 2012a,b; Walter 2015, 2017) and the mandatory switch to weak hooks in the fishery occurred in 2011 (Cass-Calay and Walter 2013). In the Eastern Atlantic, Bayse and Kersteter (2010) tested weak hooks in experimental longlines as a possible mitigation strategy to reduce interactions with pilot whales (*Globicephala spp*.) and also to examine if catch rates of yellowfin tuna (*Thunnus albacares*) and other marketable species could be maintained. Similarly, Bigelow et al. (2012) tested weak hooks to examine whether catch rates of bigeye tuna (*Thunnus obsesus*) could be maintained in order for managers to evaluate weak hooks as a possible mitigation strategy for false killer whales (FKW) (*Pseudorca crassidens*) in Hawaii (Table 1).

NOAA Fisheries is responsible for managing species protected under the Marine Mammal Protection Act (MMPA). This includes reducing mortality and serious injury (MSI) to protected species when the harm of the fishery and other actions exceed the capacity of the protected species to survive or recover. As part of this effort, Take Reduction Teams (TRT) are mandated to develop methods to reduce the bycatch, serious injury, and mortality of marine mammals when mortality exceeds Potential Biological Removals (PBR) [16 U.S.C. §1362 (20)]. The False Killer Whale Take Reduction Team was established to reduce mortality and serious injury of false killer whales in the Hawaii longline fishery. The False Killer Whale Take Reduction Plan requires the use of circle hooks with a maximum wire diameter of 4.5 mm, with the goal that when a false killer whale is hooked in the fishery, through appropriate handling, the hooks would straighten and reduce MSI of false killer whales. To date, that goal has not been achieved. This study was conducted to determine whether using a smaller diameter hook with potential to reduce MSI to false killer whales could be used while minimizing impacts to the commercial longline fishery relative to target catch (bigeye tuna) and total catch value. Hook sizes in the study were determined based on the most common hook in use in the fishery (4.5mm), and a comparatively weaker hook (4.2mm) determined through discussion with the Team.

The proposed study will be used to compare previous results of Bigelow et al. (2012) and will specifically test dressed body weight differences between bigeye tuna captured by the two strength hook types. Bigelow et al. (2012) measured body lengths and converted them to weights but did not conduct sampling when larger tuna are historically captured in the fishery. To rectify these deficiencies, it is necessary to sample larger fish to test that actual body weights (from fish at auction) do not significantly deviate between bigeye tuna captured from the 4.2 mm ø (weak) and 4.5 mm ø (control) hook types. We use effect sizes for catch rates, and body sizes in random-effects meta-analysis to synthesize results from Bigelow et al. (2012) in the current study to boost power in the experiment to derive a high level of precision in the point estimates.

2 Materials and Methods

2.1 Experimental Design

The experimental design was provided under the auspices of NOAA Fisheries and required the ability to statistically detect a 10% or smaller reduction in the catch of bigeye tuna and a 5% difference in catch value (determined from body sizes) of target and bycatch species caught on 4.2 mm ø hooks compared to the catch on 4.5 mm ø hooks. The study focused on determining whether bigeye tuna catch rates and value could be maintained by switching to a 4.2 mm ø hook which, in theory, could also mitigate and improve survival outcomes for FKWs (Werner et al. 2015, McLellan et al. 2015, Fader et al. 2021a,b).

Setting the Type I error rate (i.e., rejecting the null hypothesis when it is true) to a conservative value like 10% is often used in pilot studies to detect any evidence that differences exist between populations, groups or treatments (Baverstock and Moritz 1996; Machin et al. 2009; Bigelow et al. 2012; Ryan 2013). Based on Bigelow et al. (2012) that studied 127 longline sets, Pacific Islands Fisheries Science Center (PIFSC) determined that 170 longline sets would be required in the current study to detect a 10% reduction in bigeye tuna catch rates, assuming α=0.10 and β=0.20 (i.e., the same design used in Bigelow et al. 2012 adjusted for changes in CPUE). In Bigelow et al. (2012), 929 bigeye tuna were captured on strong hooks (49.2%) and 948 on weak hooks (50.2%) (11 fish caught on unknown hook type), which translated into a mean CPUE (x 1000 hooks) of 6.1 (±5.02 SD) for strong and 6.2 (±5.39 SD) for weak hooks. From imputed effect sizes in Bigelow et al. (2012), the log odds ratio and log risk ratio on catch rates were both ~0.02 (i.e., odds and risk ratios of 0.980) and the standardized mean difference (*SMD*) (using both Cohen's *d* and Hedges' *g*; see Borenstein et al. 2021 for methods) for bigeye tuna lengths was ~0.03. The small effect sizes indicate the original Bigelow et al. (2012) study was underpowered (Supporting Resource 1; Figs. S1-S3). Testing for a small effect size (0.05) would be cost-prohibitive (Figs. S2-S3), let alone testing at 0.02 or 0.03. Both Bigelow et al. (2012) and Bayse and Kerstetter (2010) reported no significant differences in bigeye tuna catch rates and body sizes using strong and weak hooks in commercial longline fisheries in the Pacific and Atlantic, respectively. This could indicate there were actually no significant differences in the studies or it could mean a failure to detect changes (i.e., Type II errors) using small sample sizes.

Another way to examine the magnitude of effect sizes in Bigelow et al. (2012) was to calculate the common language effect size (*CLES*) for continuous variables (McGraw and Wong 1992). The *CLES* was 0.505 for CPUE, 0.514 for weight, and 0.508 for length. The *CLES* calculates probabilities that a randomly selected score from one of the hook types would be greater than a randomly sampled score from the other hook type (McGraw and Wong 1992). As an example, the probability of a bigeye tuna being larger on a strong hook than a weak hook was 50.8%. Since a certain percentage of false positives (Type I errors) can be expected in null hypothesis significance testing by random sampling (Ellis 2000; Ryan et al. 2013), the estimate of 170 sets in the present study can probably serve as a buffer for unexpected increases or decreases in CPUE. At no cost, power was significantly boosted in the present study (Figs. S1-S3) by incorporating and synthesizing effect sizes and samples from Bigelow et al. (2012) in a random-effects meta-analysis (Musyl et al. 2015; Jackson and Turner 2017; Musyl and Gilman 2019; Borenstein et al. 2021).

2.1.1 Candidate hooks and Strength Tests

Several candidate hooks were examined and tested (Supporting Resource 2; Figs. S4-S6). Destructive failure testing was conducted on 5 samples from each hook type. The round circle hooks chosen for the experiment by NOAA Fisheries was a strong (control) circle hook (15/0, 10° offset/left, ringed, 4.5 mm ø) and a weak circle hook (15/0, 10° offset, ringed, 4.2 mm ø). The control or 4.5 mm ø hook had an average release point (i.e., when it became essentially straightened) at 230 kg (mean = 507.04 lb ± 30.51 (SD)) and the average release point of the 4.2 mm ø hook was 201 kg (443.04 lb ± 35.93). The Δ or the breaking strength difference between 4.5 mm ø and 4.2 mm ø hooks was 29 kg (64 lb)(Table 1). The consistency amongst manufactured hooks used for the study was strong and examined by randomly selecting and measuring the diameter (Figure S12) and gape of (Figure S13).

2.1.2 Field Trials

Four longline vessels (minimum size of 18 m (58 ft.)) were chartered to make 170 experimental deep-set longline gear deployments targeting bigeye tuna with each set deploying >2200 hooks in alternating order (i.e., 4.5 mm ø-4.2 mm ø-4.5 mm ø-4.2 mm ø….and so on for the entire set). All vessels were supplied with the same bait (saury, *Cololabis saira*) and identical gear to make the 2.3 mm ø monofilament gangions and leaders. Bigelow et al. (2012) details the gear used in the longline fishery. Snaps were marked with cable ties by hook type (black for 4.5mm ø & pink for 4.2 mm ø) for crew to assemble into bins in alternating order. Typically each vessel operated 3 baskets of ganglions at a time, as hooks were retrieved from the main line they were distributed according to the needs of each basket in order to maintain the alternating sequence. Both captain and crew confirmed that the alternating pattern was able to be easily and consistently achieved during the study as it was effortlessly adapted by the vessel's crew and confirmed by an observer.

Other considerations required the number of hooks between floats to be equal or greater than 15 (16 U.S.C. 1801, 75 FR 2205 Part 665) and the vessels were required to fish during the "best" catch period defined by NMFS which coincided with experimental fishing (e.g., 24 March 2021 – 31 July 2021)(see also Bigelow et al. 2012). In the experiment, fishermen were allowed to keep and sell their catch and to choose their operational fishing parameters (i.e., setting and hauling times) and fishing areas. All trips accommodated a scientific observer from the Pacific Islands Regional Observer Program who worked with crew to maintain the alternating hook design, recorded catch data, kept straightened hooks, measured fish brought on board and placed metal operculum tags on target and retained catch so these fish could be tracked at auction after the vessel offloaded in port to obtain price and weight information at the United Fishing Agency (UFA), Honolulu, HI.

Captains in the current study indicated to observers what they considered as straightened hooks; Captains defined a 'straightened hook' as having the barb or point of the hook partially or fully opened from the shank (Supporting Resource 2 (Fig.S4) and Supporting Resource 3).Observers were asked to collect all straightened and deformed hooks throughout the study however collection varied amongst vessels and observers. Potential for captains, crew, and observers to have differing interpretations for qualifying straightened or deformed hooks were present and complicated the collection process onboard chartered vessels. Select observers did not collect any hooks due to overbearing workload while at sea. Also, project conflicts complicated diligent hook collection while at sea. On other charters the captain maintained all straightened and deformed hooks throughout the study, in such cases the total number of straightened hooks was obtained from the captain or estimates were given in the event where no hooks were retained.

2.2 Statistical Analysis

To examine for differences in variability of catch rates between species captured on control (4.5 mm ø) and 4.2 mm ø (weak) hooks within and between studies, we calculated Risk Ratios (*RR*) for each species from catch records and the number of hook types deployed and analyzed effect sizes with random-effects models (Borenstein et al. 2021). The *RR* is straightforward to interpret. If the effect size is 1.0, there are no differences between the two groups. Effect sizes >1.0 indicates higher catch risk on 4.5 mm ø hooks whilst <1.0 indicates higher catch risk on 4.2 mm ø hooks. Formally "catch risk" is the risk of being captured on a strong/weak hooks. This calculation is done for each species and the effect size is synthesized by moderator (grouping) variables and weighted by the model. If the *RR* is greater than 1, the catch risk (or catch) is higher on strong hooks. Less than 1, the catch risk is higher on weak hooks. Mean Difference (*MD*) and Standardized Mean Differences (*SMD;* Hedges' *g*) were calculated for comparisons on continuous variables (body length, dressed weight, price per pound and sales price) and analyzed with random-effects models (Borenstein et al. 2021). The *MD* is used when outcomes are measured in the same unit or scale and has greater statistical power than the *SMD*. The *SMD* is used when different scales and units are used to measure outcomes across studies and is more generalizable than the *MD* (Takeshima et al. 2014). To address possible measurement errors in the data, we used the *SMD* as a complimentary analysis to compare results from the *MD*. The *MD* is the raw difference between two means and it is intuitive while the *SMD* is the *MD* divided by the standard deviation from the groups (i.e., either separate or pooled) and is expressed in standard deviation units. For the *MD* and *SMD*, if the effect size is greater than 0.0 (i.e., null), then the comparison favors 4.5 mm ø hooks whilst less than 0.0 favors 4.2 mm ø hooks.

Our hypothesis assumed species to represent random samples (i.e., mixtures of samples, sexes, sizes, hook types) captured under varying environmental and operational conditions (i.e., nuances between fishing practices of crews on different vessels, different fishing locations) over different temporal-spatial scales. In the random-effects model, the underlying (infinite-sample) effect sizes have their own distribution and sampling errors rather than a single value (Sutton et al., 2000; Borenstein et al. 2021). To estimate variability between studies (i.e., tau-squared, τ^2) as T^2, the method of moments (DerSimonian and Laird 1986, Kontopantelis and Reeves 2010) was calculated using Comprehensive Meta Analysis v. 2.2.064 (Borenstein et al. 2021) and the `metafor` package in R (Viechtbauer 2010; Wallace et al. 2012). Cochran's *Q* statistic was used to test for heterogeneity and was also used in mixed-effects analysis of variance (meta-ANOVA) designs to test effect sizes for *RRs* and *MDs* within and across subgroups (e.g., species, ecological subgroups) (Gurevitch and Hedges 1999). The I^2 statistic, derived from *Q*, described the proportion of observed dispersion between studies that was real (Borenstein et al. 2021). Sensitivity analysis was used to look for bias and patterns and the variability in true effect sizes was estimated with 95% prediction intervals (Borenstein et al. 2021). We depicted variability in effect sizes using forest plots where the area of the boxes for each study are proportional to the inverse of the variance, and any side of the box is proportional to the inverse of the standard error (SE). The 95% confidence intervals (horizontal bars) for each study is proportional to the SE and is related to sample size and precision. If the confidence intervals excludes the null (1.0 for *RR*, 0.0 for *MD, SMD*), the *p*-value is less than 0.05 and the study is statistically significant on a *Z* test. In the forest plots, dashed vertical bars represent the summary effect size and the solid vertical line represents the null for both *RR*, *MD* and *SMD*. The diamonds represent the summary effect size, and the width is proportional to the 95% CIs and the *p*-value is from *Q*.

Another diagnostic tool, receiver operator characteristic (ROC) curves and area under the ROC curves (AUC) were calculated to examine signal strength and accuracy of the moderator variables to classify catch by hook types using MedCalc Statistical Software version 17.6 (Ostend, Belgium) and LogXact v.12 (Cytel Inc., Cambridge, MA USA, 2019) (Kleinbaum and Klein, 2010; Hilbe 2016; Umemneku Chikere et al., 2019). We calculated exact 2-tailed Cochran-Mantel-Haenszel (*CMH*) tests to compare mortality outcomes by hook type and species using StatXact v. 12 (Cytel Inc., Cambridge, MA USA, 2019). For consistency across weak hook studies found in the literature, we followed recommendations found in Foster and Bergmann (2010), Kerstetter and Graves (2006) and Bayse and Kerstetter (2010) by only including species with a sample size of at least *n*=10 on one of the two hook types. For consistency, statistical tests were performed at the $p = 0.05$ level of significance to compare results with the species analyzed in Bigelow et al. (2012).

3 Results

3.1 Catch Details

During 24 March 2021 to 31 July 2021, contracted vessels fished an area encompassing 14.27 – 32.10 N, 162.23 – 139.42 W, deploying 251128 4.5 mm ø and 251128 4.2 mm ø circle hooks from 13 trips comprising 178 individual sets (data from an additional 8 sets were also used in the study*). In total, 3536 animals (48%) were caught on control hooks (4.5 mm ø) and 3822 (52%) were caught on 4.2 mm ø (weak) hooks representing 43 species (Table 2). Out of the total catch, lancetfish was caught in the highest numbers on both hook types (40.4% 4.5 mm ø; 41.4% 4.2 mm ø), followed by blue shark (14.7% 4.5 mm ø; 12.6% 4.2 mm ø) and bigeye tuna (12.1% 4.5 mm ø; 11.6% 4.2 mm ø). Three FKW were captured (2 on 4.5 mm ø hooks and 1 on a 4.2 mm ø hook) and one bottlenose dolphin was caught on a 4.2 mm ø hook. In total, 39 4.5 mm ø hooks and 58 4.2 mm ø hooks were considered straightened by vessel captains (Table 2; Figs. S7-S8). In the present study, Captains defined a ‘straightened hook’ as having the barb or point of the hook partially or fully opened from the shank (Supporting Resource 2 (Fig.S4) and Supporting Resource 3).

*Lynker was contracted to collect data from 170 sets aboard chartered longline vessels. Each vessel was contracted to complete approximately 17 sets, however, due to various trip lengths two vessels completed additional sets in response to slow fishing in order to make a profitable fishing trip. Observers were asked to continue to collect catch and hook data with the captain's permission even though they were not contracted to do so. The additional 8 sets were the results of this additional fishing effort and were included in the study with permission of the chartered vessels in hopes to strengthen the statistical power of analysis.

3.2 Analysis of Catch on Strong and Weak Hooks

3.2.1 Catch Rates

The *RR* analysis of catch rates on strong and weak hooks comprised 22 species from the current study (indicated by [5] in figures) and 22 species from Bigelow et al. (2012) (indicated by [3] in figures) (Fig. 1). Within subgroups [5] and [3], no significant differences were detected but studies in subgroup [3] were more heterogeneous (I^2 = 34%) than the current study [5] that was more homogenous (I^2 = 5%). The summary estimate within subgroups [5] and [3] indicated higher catch risk on weak hooks (but not significantly so). Smaller studies are indicated by wide confidence intervals and noticeably smaller effect sizes (i.e., boxes) in the plots indicating less precision and power in the estimate Note especially the much narrower widths of the summary 95% CIs (i.e., diamonds) compared with the individual studies, which indicates more precision in the summary estimates (Fig.1). Overall, there were significant differences between effect sizes in study [3] and the current study [5] (Table 2; Fig.1) ($Q_{(43)} = 60.59$, $p = 0.04$, $T^2 = 0.005$, I^2 = 29%). The summary effect size was 0.959 [95% CI: 0.959 - 1.005] indicating higher catch risk on weak hooks (i.e., <1.0). Both the 95% CI and 95% prediction interval [95% PI 0.820 - 1.110] suggest a propensity for higher catch risk on weak hooks but also that the effect size straddles the null (i.e., vertical line at 1.0). The meta-analysis preserves the information in study [3] where Bigelow et al. (2012) indicated significant pairwise differences for spearfish and yellowfin tuna where the 95% CI s for those species exclude the null.

A comparison of subgroups Target, Retained, Istiophorid billfish, Elasmobranch, and Discard (Fig.2) showed homogeneity in the Target and Elasmobranch subgroups where $I^2 = 0.00$ indicated no dispersion between species ($T^2 = 0.00$). Since by definition the between-studies variance was $T^2 = 0.00$, the effect size can be attributed to random events within studies and that all species share a common effect size (Borenstein et al., 2021). As a subgroup, catch risk was higher on strong hooks for Elasmobranch but all other subgroups showed a higher catch risk on weak hooks. The sensitivity analysis (Fig. 3) indicates effect sizes switching back and forth around the summary effect size and the confidence intervals straddle the null.

Marketable species comprising the Target, Retained and Istiophorid billfish subgroups were analyzed separately for a direct comparison to study [3] (Bigelow et al. 2012)($Q_{(26)} = 36.14$, p = 0.09, $T^2 = 0.006$, I^2 = 28%)(Fig. 4) and the summary effect size was 0.931 [95% CI: 0.874 - 0.992] indicating higher catch risk on weak hooks (i.e., <1.0). Here, the 95% CI excludes the null but the prediction interval [95% PI 0.78 - 1.110] includes it. Most of the species subgroups are homogeneous ($I^2 = 0.00$) but yellowfin tuna, albacore, spearfish, striped marlin, dolphinfish and opah indicate heterogeneity in catch risks (Fig. 4) with spearfish and dolphinfish subgroups indicating significant heterogeneity. For example, study [3] indicates higher catch risk of albacore, spearfish, dolphinfish, and opah on strong hooks compared to higher catch risk for these species on weak hooks in study [5]. The opposite catch risk pattern was observed for yellowfin tuna where higher catch risk was higher on weak hooks in [3] but not in [5]. The wide summary diamonds in Fig. 4, however, indicate less precision and power in most of the comparisons due to small sample sizes.

3.2.2 Analysis of body lengths

The *MD* analysis of body lengths on strong and weak hooks indicated homogeneity (i.e. no significant differences) between groups [5] and [3] with little or no dispersion within subgroups as measured by I^2 (Table 3; Fig.5) ($Q_{(29)} = 27.741$, $p = 0.532$, $T^2 = 0.00$, $I^2 = 0\%$). The summary effect size was 0.281 [95% CI: -0.397 - 0.959] indicating that fish caught on strong hooks were on average 0.281 cm larger than fish caught on weak hooks. The estimates for $T^2 = 0.00$, $I^2 = 0.0$, indicate that all studies share a common effect size (i.e., the 95% CIs and PIs are identical). On average, bigeye tuna in [5] were 3.286 cm significantly larger on strong hooks whereas in study [3] bigeye tuna were larger on weak hooks but the *MD* effect size was smaller (0.70 cm) and non-significant. The analysis using the *SMD* (Fig. 6) indicates the same conclusion but since the scaling provides a less contracted graph, it is often shown in conjunction with *MD* to aid in interpretation but the units are expressed in standard deviations.

3.2.3 Analysis of Auction Weights

Dressed and whole weights for species are provided in Table 3. Overall, the *MD* analysis on the effect size for weight on hook types indicated significant and high dispersion between subgroups (Table 3; Fig. 7) ($Q_{(11)} = 29.881$, $p = 0.002$, $T^2 = 5.807$, $I^2 = 63\%$). The summary effect size was 0.369 [95% CI: -1.597 - 2.335] indicating on average, fish captured by strong hooks were 0.369 lb (0.17 kg) heavier than fish caught on weak hooks. The prediction interval was wide [95% PI -5.45 - 6.18] spanning -5.45 lb (-2.47 kg) on weak hooks to 6.18 lb (2.8 kg) on strong hooks. On average, bigeye tuna captured on strong hooks were significantly heavier (6.834 lb; 3.10 kg) than those caught on weak hooks. Yellowfin tuna were also heavier on strong hooks (3.222 lb; 1.46 kg) but the difference was non-significant (Fig. 7).

Clearly, study [5] swordfish is an outlier (Fig. 7) with fish captured on weak hooks significantly heavier (50.301 lb; 22.82 kg) than those caught on strong hooks. The forest plot using *SMD* showed a similar pattern (Fig. S9). Though the sample size was not large (Table 3), closer inspection indicated that one of the vessels fished north of the island and captured large female swordfish on weak hooks (Fig. 8) presumably during their spawning period (DeMartini et al. 2000; Sculley and Brodziak 2020). The other boats captured larger swordfish on strong hooks but the difference was not significant owing to the dispersion between males and females ($I^2 = 60\%$) and small sample sizes (Fig. 8).

3.2.4 Analysis of Revenue

Table 4 provides summaries of the average sales price per pound and sales per fish for species sold at auction whilst Table 5 provides the gross *ex*-vessel revenue. The *MD* effect size for price per pound on strong and weak hooks indicated homogeneity between groups (Fig. 9; $Q_{(11)} = 11.98$, $p = 0.372$, $T^2 = 0.007$, $I^2 = 8\%$)(*SMD* is given in Fig. S10). The summary effect size was 0.073 [95% CI: -0.098 - 0.245] indicating on average, fish captured by strong hooks generated $0.073 more per pound than fish caught on weak hooks. The prediction interval was [95% PI -0.20 - 0.34]. Bigeye tuna captured on strong hooks produced, on average, $0.075 more pound than bigeye caught on weak hooks. On average, swordfish commanded significantly more per pound on weak hooks ($0.882) but sample sizes were small and variable (Fig. 8; Table 4).

Overall, in the analysis of the *MD* effect size for average sales price per fish (Fig.10; *SMD* given in Fig. S11), there was significant heterogeneity among subgroups ($Q_{(11)} = 24.150$, $p = 0.012$, $T^2 = 211.433$, $I^2 = 54\%$). The summary effect size was 3.665 [95% CI: -9.182 - 16.512] indicating that on average, fish caught on strong hooks sold for $3.665 more per fish than those caught on weak hooks. The prediction interval was [95% PI -31.87 - 16.51]. Bigeye tuna captured on strong hooks averaged significantly more per fish ($52.891) than those caught on weak hooks. Species in the Target subgroup fetched, on average, $12.281 more per fish at auction on weak hooks but the analysis was significantly influenced by the inclusion of swordfish that brought in $408.334 more on weak hooks but the samples were skewed (Fig. 8). Removing [5] swordfish in the analysis dropped I^2 to zero (p=0.669) indicating the inclusion of swordfish contributed significantly to the heterogeneity of the Target subgroup. The Retained ($1.687) subgroup fetched, on average, more on weak hooks but the Istiophorid billfish subgroup brought in $16.513 more on strong hooks but the difference was non-significant and samples were homogeneous (I^2 = 0.0). For all marketable species, *ex*-vessel gross revenue indicates fish captured on strong hooks brought in $204.45 more than fish caught on weak hooks (Table 5).

3.2.5 Common Language Effect Sizes

For bigeye tuna, the *CLES* indicated that body lengths (0.56), dressed weights (0.57), price per pound (0.51) and average sales per fish (0.55) would be higher on strong hooks than weak hooks.

3.2.6 Depredation Rates and Mortality

Incidences of whale depredation were too rare to have a statistical impact (Table 6) and there were no significant differences in at-vessel mortality outcomes for teleosts and elasmobranchs caught on 4.5 mm ø and 4.2 mm ø (weak) hooks (Table 7). No conservation benefit was indicated by the use of 4.2 mm ø hooks for these species.

3.2.7 ROC curves

The various moderator variables in the ROC analysis (from logistic regression models) had low resolving power and were ineffective in classifying catch by hook type (Fig. 11). None of the variables were significant in the logistic regression model and resolving power was essentially "50-50" in terms of correctly assigning a hook type based on these variables.

4 Discussion

The present study found the weak hooks (4.0 mm, 4.2 mm ø) to exhibit similar performance to strong control hooks (15/0, 10° offset, ringed, 4.5 mm ø). The diagnostic tools (i.e., distribution of effect sizes, AUC, sensitivity analysis, confidence intervals, prediction intervals, I^2 and T^2) indicated homogeneity or very close correspondence between species caught on the two hook types. The summary estimate for bigeye tuna reported herein is the best and most precise estimate available to appraise the catch retention rate of 4.2 mm ø hooks in the Hawaii fishery. As many researchers have reported; the ability to replicate findings over temporal and spatial scales is the best and most powerful way to authenticate that study results are real. Single studies are sometimes inconsistent or ambiguous because sample sizes are generally too low which increases the chance of type II errors. By combining information from Bigelow et al. (2012) into the current study in a random-effect meta-analysis, power was significantly increased and the results are generalizable and can be extrapolated. By synthesizing information in a random-effects meta-analysis, random errors were reduced to produce more precise and powerful estimates of the true effect size (Sutton et al. 2000; Welton et al. 2012).

Though larger fish were captured on both hook types in the present study compared to Bigelow et al. (2012) (i.e., strong hooks: mean = 121.95 cm FL (±16.40 SD) *v.* 106.80 (±24.43); weak hooks: 118.66 (±4.29) *v.* 107.5(±24.01)) the differences in size were not statistically significant between the parent studies in the meta-analysis (see Tables 3 and 4 for explanation of percent reduction in price comparing weak and strong hooks as they relate to the TRT set threshold of less than 10% reduction in price or weight). The study achieved the goal of the experiment by analyzing catch rates of larger bigeye tuna that were unavailable to Bigelow et al. (2012) due to seasonal movement patterns. Never-the-less, though not significant, these differences in sizes may have clinical or practical significance. In the overall study, catching smaller bigeye tuna more often on 4.2 mm ø hooks appeared to be ameliorated by catching larger, better quality fish on 4.5 mm ø hooks that fetched more at auction. The percentage of straightened hooks in the current study on 4.5 mm ø (~40% [39]) and 4.2 mm ø hooks (~60% [58]) was not significantly different (2-tailed exact *CMH* test) and suggests the delta breaking strength between hook types (29 kg) probably had little effect on catch rates. In other words, the weak hooks used in Bigelow et al. (2012) and the current study were probably too similar in performance to have a large impact. Though it does not appear wire diameter influenced results between the two Hawaii weak hook studies, Foster and Bergmann (2012b) observed that the shape of the wire stock in the forging process (i.e., laterally compressed) and wire diameter may have promoted release of swordfish from strong hooks. It was noted that strong hooks were more easily torn from the soft flesh located in the jaw than weak hooks.

There are few standard classifications or operational definitions for straightened hooks (i.e., unbent hooks) recognized in the literature. The notable exceptions are Edappazham et al. (2008), Bayse and Kerstetter (2010), Bigelow et al (2012) and McLellan et al. (2015) but they use different schemes and operational definitions. Clearly, a numerical standard or solution must be developed to make meaningful comparisons across studies when comparing straightened hooks and the resistance required to unbend them. Captains in the current study indicated to observers what they considered as straightened hooks (Supporting Resource 3). Bayse and Kerstetter (2010) suggested straightened hooks to be when the barb end of the hook was forced 90° in the "open" position. In this situation, since the barb is not recurved to keep species hooked, the probability of being released would be higher. Bigelow et al. (2012) compared straightened and non-straightened hooks collected by observers and morphometrically compared them. Straightened control (strong) hooks were considered straightened if the hook was deformed (or opened) at ~66% of gape width and for weak hooks this was ~50% of gape width. In the current study, 4.2 mm ø (weak) and 4.5 mm ø hook tandems were tested at incremental force until maximum plastic deformation occurred (Supporting Resources 2). Bigelow et al. (2012) and McLellan et al. (2015) mentioned that hook materials should be considered in defining relative strength which was studied along with tempering in Edappazham et al. (2008). Furthermore, the small number of destructive strength failure tests and samples performed on hooks (Supporting Resource 2) does not approach the testing requirements and precision used in industrial testing procedures and standards to derive tolerances (Meeker and Escobar 1999). As Bigelow et al. (2012) noted, better strength tests conducted at-sea would provide realistic scenarios to test the strength of hooks by retrieving them from gangions under force. Assuming the gangion is not the weakest part of the system, resistance at the hook would not only be determined from the size and weight of the species, but also from factors like surface area and shape of the animal (i.e., compressed, fusiform) and also by exogenous factors like wind, current, wave patterns and angle of attack from the vessel.

There were several potential confounders in the Bigelow et al. (2012) dataset compared to the present study that may have affected selectivity and performance of the hooks (Table 1). The first obvious factors are the delta breaking strength differences between strong and weak hooks (45 kg) and wire diameter of the weak hooks (4.0 mm ø) used in Bigelow et al. (2012) and the weak hooks with a diameter of 4.2 mm ø and a delta breaking difference of 29 kg in the current study (Table 1). Bigelow et al. (2012) also included vessels that fished with a mix of ringed and non-ringed hooks and different leader types (i.e., 2.0 mm ø monofilament, wire). It is not clear if these changes materially affected the catchability and selectivity of the hooks (e.g., Serafy et al. 2012; Ingólfsson et al. 2017; Reinhardt et al. 2017; Gilman et al. 2018, 2020). Piovano and Swimmer (2017) reported catch rates of swordfish were significantly affected by ringed circle hooks. Furthermore, it is not clear if these factors (wire diameter, delta strength, rings, leader material) acted in synergy with minor shape differences of the weak hooks to affect catch rates. Re-analyzing these factors in Bigelow et al. (2012) might help elucidate the possible influence the confounders may have had in catch outcomes on hook types.

Despite using 15/0 to 18/0 strong and weak hooks with vastly different delta breaking strengths (i.e., 29 to 57 kg), different wire diameters (i.e., 3.2 mm to 5.0 mm ø), combinations of offset and non-offset hooks, ringed and non-ringed hooks, and hooks from different manufactures in the comparison tests of hook types on catch rates and body lengths (Table 1); Bayse and Kerstetter (2010), Foster and Bergman (2010) and Bigelow et al. (2012) did not report many significant findings between species comparing weak and strong hooks. Out of a possible 23 pairwise species tests between 15/0 hook types in catch rates, Bigelow et al. (2012) reported two tests as being significantly different (strong hooks caught significantly more spearfish but weak hooks caught significantly more yellowfin). And, out of a possible 15 pairwise tests on body lengths for species between 15/0 hook types, Bigelow et al. (2012) reported no significant differences. Foster and Bergmann (2010) reported that out of a possible 23 pairwise species tests between 16/0 hook types; three tests indicated significantly more fish were caught on strong hooks (lancetfish, wahoo, bluefin tuna) than on weak hooks. From a possible 13 pairwise species tests on 16/0 hook types, Bayse and Kerstetter (2010) documented that significantly more pelagic stingray were captured on strong hooks and in 4 comparisons with 18/0 hook types, significantly more swordfish were caught on strong hooks. Only one significant pairwise species test (out of 7) comparing body lengths and dressed weights in hook types was reported (significantly larger yellowfin were captured on strong 16/0 hooks). The results indicate that hook types in the fisheries probably did not materially affect catch rates or sizes in a large way but it is also likely these studies were underpowered (Musyl et al., in prep). Out of a possible 85 pairwise tests, by chance alone at the $p=0.05$ level, one would expect >4 significant test results and the reported number was 8.

Due to the likely severity of injury between the association of embedded hooks and trailing line left in FKW (and other toothed odontocetes) from longline fishing encounters during depredation events (Gilman 2011; Forney et al. 2011; Hamer et al. 2012; Bradford 2020; Carretta et al. 2021; Fader et al. 2021a,b), current regulatory measures under the Take Reduction Plan (TRP) require use of monofilament nylon leaders and branchline equal to, or greater than, 2.0 mm in diameter to improve health outcomes while removing hooks by unbending them with sufficient force (Bradford 2020; Carretta et al. 2021). For this to work, this assumes the hook is the weakest part of the gear. The combination of small, weak circle hooks is thought to reduce the proportion of interactions that result in serious injury and increase survival outcomes for those cetaceans that are hooked (Forney et al. 2011; McLellan et al. 2015). As pointed out by Baird (2019), however, proper handling techniques by the crew on the longline vessel are needed to remove hooks. Since only 20% of deep-set longline trips are monitored in the Hawaii-based fishery, Baird (2019) contends this is the flaw in the system. Trips that are unmonitored may encourage the crew to cut the line if observers are not present. The use of electronic monitoring systems may help in this regard to encourage and incentivize proper handling techniques.

But there is another potential flaw. To exert sufficient tensile force to unbend the hooks (see Supporting Resource 2; Edappazham et al. 2008; McLellan et al. 2015), the monofilament gangion must not be the weakest link in the system. It has been demonstrated, however, that nylon monofilament yarn in seawater hydrates and can lose up to ~50% of its weight after 30 days and concomitantly be reduced in breaking strength by ~10% after 30 days and up to ~20% after 90 days (Thomas and Hridayanathan 2006; Mondal et al. 2019). Weathering, UV exposure and pollutants (Thomas et al. 2009; Atayeter et al. 2014) can further accelerate the degradation process and weaken the nylon as can the catch of sharks and large fish nicking, abrading and stretching the line (i.e., fishermen call this "smoked" when the line is stretched to a point whereby it becomes opaque). Leaving trailing gear attached to hooks in animals can present challenges. Similar to the scenario mentioned in Musyl et al. (2011) and Musyl and Gilman (2019) to explain variable retention times for pop-up satellite tags attached to a diverse selection of pelagic fishes and sharks, it is possible that fouling organisms could accumulate on the trailing line and exert drag and vibration forces that would be maximized at the hook. Moreover, these forces could prevent or suspend hook wounds from healing, providing a route for infection, inflammation and tissue necrosis. Infusing hooks with antimicrobial agents (Dabrowiak 2009; Shahid et al. 2021; Zhang et al. 2021) and/or making them to corrode faster (Edappazham et al. 2010; McGrath et al. 2011) could reduce infection and necrotizing organisms for those hooks left embedded in tissues. For example, McGrath et al. (2011) documented wire diameter and material as important factors to accelerate hook decay (e.g., nickel-plated carbon-steel hooks decayed faster than stainless hooks). 'Weak split rings' attached to the eyes of strong hooks could also function to eliminate trailing line and lessen tissue damage from hooks being pulled out. McLellan et al. (2015) report on the types of slicing wounds and deep cuts on necropsied cetaceans when hooks are pulled out. Since hooks would be left in place in tissues under this scenario of weak split rings, the system would require some level of antimicrobial protection and (or) hooks degrading over time.

The price analysis did not take into account several influential variables known to affect flesh quality and hence price at auction. Factors such as time spent hooked, species-specific stress responses, parasites, shark damage, temperature, dissolved oxygen, fight time and fishing practices can influence quality and price (e.g., Cramer et al. 1981; Foy et al. 2006; Nóbrega et al. 2014; Khoshnoudi-Nia S, Moosavi-Nasab M 2019; Sogn-Grundvåg et al. 2020; Komolka et al. 2020; FDA 2021) but these factors have no relevance to hook type. Moreover, a poor gaff placement and improper storage and (or) freezer breakdowns can also affect quality and price but these factors have no bearing on hook type. That no significant differences were found in price per pound between animals caught on the two hook types and *ex*-vessel revenue probably reflects a random sample and that the two hook types were similar in performance (see Tables 3 and 4 for explanation of percent reduction in price comparing weak and strong hooks as they relate to the TRT set threshold of less than 10% reduction in price or weight).

Rates of cetacean depredation on longlines have been used to examine movement patterns and possible strategies to mitigate interactions (Forney et al. 2011, Fader et al. 2021a, b). The level of interactions and depredation rates observed in the study were too low to have a statistical impact. Both Bayse and Kersteter (2010) and Bigelow et al (2012) concluded that the level of odontocete interactions were too low in their weak hook studies to judge whether weak hooks reduced interactions. Another plausible way to gauge cetacean interactions on longline gear might be to observe the broken ends of the monofilament gangions to see if the line stretched and ruptured or if it was bitten through. Given the degradation and weakening of monofilament over time, it seems like a logical extension to hypothesize that large animals would occasionally break weathered and degraded monofilament and free themselves.

Lastly, the most cost-effective method to determine the level of cetacean interactions with longlines could be achieved in 2 ways:

1) Shark DNA has been successfully amplified and identified from depredated catch in longline caught fish in Australia (Vardon et al. 2021). Similar studies could be conducted in the Hawaii-based longline fishery to examine species of odontocete depredating the catch and perhaps how many individuals are involved.

2) Using the same techniques, it may be possible to harvest DNA from straightened hooks to see what species are freeing themselves. This is necessary to convince fishermen that large target catch are not escaping from weak hooks. Techniques used to amplify eDNA could be used on the hooks but would require storage at-sea in the appropriate buffer to prevent breakdown and contamination.

Until it can be determined which species are unbending hooks, the efficacy of weak hooks will remain enigmatic.

Acknowledgements
This project was funded by the Protected Resources Division of NOAA Fisheries, Pacific Islands Regional Office through contract no. 135M3319FFT0064P19001 to Lynker Technologies. We thank D. Kramer and K. Brindock from NOAA and J. Vincent, C. Hawkins, and L. Trilles, from Lynker for administration and coordination of the project. We thank F. O'Neil for his efforts in gear procurement and assistance in developing the study. We especially thank fisheries observers M. Clarke, P. Kalahiki, J. Burkhardt, D. Dannis, C. Diaz, and D. Townsend for at-sea data collection, assistance at the auction and data entry. The United Fishing Agency kindly allowed us to track sales data and M. Goto and C. Schrader are thanked for their help and contributions to the project. Lastly, but most importantly, we would like to thank vessel captains and crew; especially T. Ryder, D. Lewis, P. Westbrook, and G. Ching and their crews who made this project possible by fastidiously adhering to the study design by alternating hooks and assisting with data collection.

5 References

Atayeter S, Atar HH, Ören Ö, Meric İ (2014) Determination of mesh breaking strength of polyamide fishing nets under the exposure of different heavy metal concentrations and temperature. Journal of Agricultural Sciences 20:57-62.

Baird RW, Gorgone AM, McSweeney DJ, Webster DL, Salden DR, Deakos MH, Ligon AD, Schorr GS, Barlow J, Mahaffy SD (2008) False killer whales (*Pseudorca crassidens*) around the main Hawaiian Islands: long-term site fidelity, inter-island movements, and association patterns. Marine Mammal Science 24:591-612.

Baird RW, Hanson MB, Schorr GS, Webster DL, McSweeney DL, Gorgone AM, Mahaffy SD, Holzer D, Oleson EM, Andrews RD (2012) Range and primary habitats of Hawaiian insular false killer whales: informing determination of critical habitat. Endangered Species Research 18:47-61.

Baird RW, Oleson EM, Barlow J, Ligon AD, Gorgone AM, Mahaffy SD (2013) Evidence of an island-associated population of false killer whales (*Pseudorca crassidens*) in the Northwestern Hawaiian Islands. Pacific Science 67:513-521.

Baird RW (2019) The perils of relying on handling techniques to reduce bycatch in a partially observed fishery: a potential fatal flaw in the False Killer Whale Take Reduction Plan. PSRG-2019-14. Presented to Pacific Scientific Review Group, March 5-7, 2019, Olympia, WA: Protected species Research Group.

Baverstock PR, Moritz C (1996) Project design. In: Molecular Systematics, 2e (eds: Hillis DM, Moritz C, Mable BK). Sinauer Associates, Sunderland, MA, USA.

Bayse SM, Kerstetter DW (2010) Assessing bycatch reduction potential of variable strength hooks for pilot whales in a western North Atlantic Pelagic Longline Fishery. Journal North Carolina Academy of Science 126, 6–14.

Bigelow KA, Kerstetter DW, Dancho, MG, Marchetti JA (2012) Catch rates with variable strength circle hooks in the Hawaii-based tuna longline fishery. Bulletin of Marine Science 88, 425-447.

Borenstein M, Hedges, LV, Higgins JPT, Rothstein H. (2021) Introduction to meta-analysis, Second Edition. John Wiley & Sons, UK.

Bradford AL, Oleson EM, Baird RW, Boggs CH, Forney KA, Young NC (2015) Revised stock boundaries for false killer whales (*Pseudorca crassidens*) in Hawaiian waters. U.S. Dep. Comm., NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-47, 29p.

Bradford AL (2020) Injury determinations for marine mammals observed interacting with Hawaii and American Samoa longline fisheries during 2018. NOAA Technical Memorandum. NOAA-TM-NMFS-PIFSC-99, 20 p. doi:10.25923/2prh-0z06

Carretta JV, Forney KA, Oleson EM et al. (2021) U.S. Pacific Marine Mammal Stock Assessments: 2020, U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-646.

Cass-Calay SL, Walter JF (2013) Standardized catch rates of large bluefin tuna (*Thunnus thynnus*) from the U.S. pelagic longline fishery in the Gulf of Mexico during 1987-2010. Collect. Vol. Sci. Pap. ICCAT, 69(2): 992-1004 (2013)

Cramer JL, Nakamura RM, Dizon AE (1981) Burnt tuna: conditions leading to rapid deterioration in the quality of raw tuna. Marine fisheries Review 43: 12-16

Dabrowiak JC (2009) Metals in Medicine. John Wiley & Sons, Chichester, UK

Demartini E, Uchiyama JH, Williams HA (2000) Sexual maturity, sex ratio, and size composition of swordfish, *Xiphias gladius*, caught by the Hawaii-based pelagic longline fishery. Fish. Bull. 98:489-506.

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Controlled Clinical Trials, 7, 177–88.

Edappazham G, Thomas SN, Meenakumari B, Ashraf PM (2008) Physical and mechanical properties of fishing hooks. Materials Letters 62: 1543–1546

Edappazham G, Thomas SN, Ashraf PM (2010) Corrosion resistance of fishing hooks with different surface coatings. Fishery Technology 47:121-126.

Ellis PD (2010) The essential guide to effect sizes: statistical power, metaanalysis, and the interpretation of research results. Cambridge University Press, Cambridge, UK.

Fader JE, Elliott BW and Read AJ (2021a) The challenges of managing depredation and bycatch of toothed whales in pelagic longline fisheries: two U.S. case studies. Front. Mar. Sci. 8:618031.doi: 10.3389/fmars.2021.618031

Fader JE , Baird RW, Bradford AL, Dunn DC, Forney KA, Read AJ (2021b) Patterns of depredation in the Hawaiʻi deep-set longline fishery informed by fishery and false killer whale behavior. Ecosphere 12(8):e03682. 10.1002/ecs2.3682

FDA (2021) Fish and Fishery Products Hazards and Controls Guidance Fourth Edition – June 2021. U.S. Department of Health and Human Services Food and Drug Administration, Center for Food Safety and Applied Nutrition. Washington, DC.

Forney KA, Kobayashi DR, Johnston DW, Marchetti JA, Marsik MG (2011) What's the catch? Patterns of cetacean bycatch and depredation in the Hawaii-based pelagic longline fisheries. Marine Ecology 32: 380-391.

Foster F, Bergmann C (2010) 2010 interim report: update on Gulf of Mexico pelagic longline bluefin tuna mitigation research. NOAA Fisheries, Engineering and Harvesting Branch, SEFSC. 11 p.

Foster DG, Bergmann, C (2012a) Bluefin tuna bycatch mitigation research in the Gulf of Mexico pelagic longline yellowfin tuna fishery. Bulletin of Marine Science **88**, 797.

Foster D, Bergmann C (2012b) 2012 Update Gulf of Mexico Weak Hook Research (Preliminary Results). Presentation to the Atlantic HMS Advisory Panel. September 19-21, 2012

Foy RJ, Crapo CA, Kramer DE (2006) Investigating the roles of temperature and exercise in the development of chalkiness in Pacific halibut. International Pacific Halibut Commission, Technical Report No. 50, ISSN: 0579-3920

Gilman EL (2011) Bycatch governance and best practice mitigation technology in global tuna fisheries. Marine Policy, 35, 590-609.

Gilman E, Chaloupka M, Musyl M (2018) Effects of pelagic longline hook size on species- and size-selectivity and survival. Reviews in Fish Biology and Fisheries 28:417-433, doi:10.1007/s11160-017-9509-7

Gilman E, Chaloupka M, Dagorn L, Hall M, Hobday A, Musyl M, Pitcher T, Poisson F, Restrepo V, Suuronen P (2019) Robbing Peter to pay Paul: replacing unintended cross-taxa conflicts with intentional tradeoffs by moving from piecemeal to integrated fisheries bycatch management. Reviews in Fish Biology and Fisheries, doi:10.1007/s11160-019-09547-1

Gilman E, Chaloupka M, Bach P, Fennell H, Hall M, Musyl M, Piovano S, Poisson F, Song L (2020) Effect of pelagic longline bait type on species selectivity: A global meta-analysis of evidence Reviews in Fish Biology and Fisheries *30:535–551*

Gurevitch J, Hedges LV (1999) Statistical issues in ecological meta-analysis. Ecology, 80, 1142–1149.

Hamer DJ, Childerhouse SJ, Gales NJ (2012) Odontocete bycatch and depredation in longline fisheries: A review of available literature and potential solutions. Marine Mammal Science 28: E345-E374.

Hamilton S, Baker GB (2019) Technical mitigation to reduce marine mammal bycatch and entanglement in commercial fishing gear: lessons learnt and future directions. Reviews in Fish Biology & Fisheries https://doi.org/10.1007/s11160-019-09550-6

Hilbe J (2016) Practical guide to logistic regression. CRC Press, New York.

Hughes AR, Inouye BD, Johnson MT, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. Ecol Lett 11(6):609-23

Ingólfsson Ó, Einarsson HA, Løkkeborg S (217) The effect of hook and bait sizes on the size selectivity and capture efficiency in Icelandic longline fisheries. Fisheries research 191: 10-16.

Jackson D, Turner R (2017) Power analysis for random-effects meta-analysis. Research Synthesis Methods, doi:10.1002/jrsm.1240

Kerstetter DW, Graves JE (2006) Survival of white marlin (*Tetrapturus albidus*) released from commercial longline gear in the western North Atlantic. Fisheries Bulletin, 104, 434–444.

Khoshnoudi-Nia S, Moosavi-Nasab M (2019) Prediction of various freshness indicators in fish fillets by one multispectral imaging system. Scientific Reports 9:14704 https://doi.org/10.1038/s41598-019-51264-z

Kleinbaum DG, Klein M (2010) Logistic regression, a self-learning text (3rd ed.). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-1742-3

Komolka K, Bochert R, Franz GP, Kaya Y, Pfuhl R, Grunow B (2020) Determination and comparison of physical meat quality parameters of Percidae and Salmonidae in aquaculture. Foods 2020, 9, 388; doi:10.3390/foods9040388 www.mdpi.com/journal/foods

Kontopantelis E, Reeves D (2010) Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: A simulation study. Statistical Methods in Medical Research, 21, 409–426.

Machin D, Campbell MJ, Tan SB, Tan SH (2009) Sample size tables for clinical studies. 3rd ed. Wiley-Blackwell, Chichester, UK.

Manly BFJ (2005) Multivariate statistical methods: a primer, 3e. Chapman & Hall/CRC, Boca Raton, FL.

Martien KK, Chivers SJ, Baird RW, Archer FI, Gorgone AM, Hancock-Hanser BL, Mattila D, McSweeney DJ, Oleson EM, Palmer C, Pease VL, Robertson KM, Schorr GS, Schultz MB, Webster DL, Taylor BL (2014) Nuclear and mitochondrial patterns of population structure in North Pacific false killer whales (*Pseudorca crassidens*). Journal of Heredity doi: 10.1093/jhered/esu029.

McLellan WA, Arthur LH, Mallette SD, Thornton SW, McAlarney RJ, Read AJ, Pabst DA (2015) Longline hook testing in the mouths of pelagic odontocetes. – ICES Journal of Marine Science, 72: 1706–1713.

McGrath SP, Butcher PA, Broadhurst MK, Cairns SC (2011) Reviewing hook degradation to promote ejection after ingestion by marine fish. Marine and Freshwater Research, doi:10.1071/MF11082

McGraw KO, Wong SP (1992) A common language effect size statistic. Psychological Bulletin, 111, 361–365.

Meeker WQ, Escobar LA (1999). Statistical methods for reliability data. Wiley, New York.

Meyer S, Robertson BC, Chilvers BL, Krkošek M (2017) Marine mammal population decline linked to obscured by-catch. PNAS 114, 44:11781–11786. www.pnas.org/cgi/doi/10.1073/pnas.1703165114

Mondal S, Thomas SN, Manoi KB (2019) Changes in properties of polyamide netting materials exposed to different environments. Journal Fisheries Research 3(2) http://www.alliedacademies.org/journal-fisheries-research/

Musyl M, Domeier M, Nasby-Lucas N, Brill R, McNaughton L, Swimmer J, Liddle J (2011) Performance of pop-up satellite archival tags. Marine Ecology Progress Series, 433, 1–28. https://doi.org/10.3354/meps09202

Musyl MK, Moyes CD, Brill RW, Mourato BL, West A, McNaughton LM, Chiang W-C, Sun C-L (2015) Postrelease mortality in istiophorid billfish. Canadian Journal of Fisheries and Aquatic Sciences **72**, 1–19.

Musyl MK, Gilman EL (2019) Meta-analysis of post-release fishing mortality in apex predatory pelagic sharks and white marlin. Fish and Fisheries, doi:10.1111/faf.12358

Nóbrega CC, Mendes PP, Mendes ES (2014) Factors that determine the quality of bigeye tuna, caught in the western tropical Atlantic Ocean. Arq. Bras. Med. Vet. Zootec. 66: 949-958

Patrick WS, Benaka LR (2013) Estimating the economic impacts of bycatch in U.S. commercial fisheries. Marine Policy 38 (2013) 470–475

Piovano S, Swimmer Y (2017) Effects of a hook ring on catch and bycatch in a Mediterranean swordfish longline fishery: Small addition with potentially large consequences. Aquatic Conservation Marine and Freshwater Ecosystems 27(2):372-380

Read AJ, Brinker P, Northridge S (2006) Bycatch of marine mammals in U.S. and global fisheries. Conservation Biology 20:163-169.

Reinhardt JF, Weaver J, Latham PJ, Dell'Apa A, Serafy JE, Browder JA, Christman M, Foster DG, Blankinship DR (2017) Catch rate and at-vessel mortality of circle hooks versus J-hooks in pelagic longline fisheries: A global meta-analysis. Fish and Fisheries, doi:10.1111/faf.12260

Ryan TP (2013) Sample size determination and power. John Wiley & Sons, Hoboken, New Jersey.

Serafy J E, Orbesen E S, Snodgrass D J, Beerkircher LR, Walter JF (2012) Hooking survival of fishes captured by the United States Atlantic pelagic longline fishery: Impact of the 2004 circle hook rule. Bulletin of Marine Science, 88, 605–621.

Sculley ML, Brodziak J (2020) Quantifying the distribution of swordfish (*Xiphias gladius*) density in the Hawaii-based longline fishery. Version of Record: https://www.sciencedirect.com/science/article/pii/S0165783620301557 Manuscript_18dd8c52eed3422eb61b023be72f89af

Shahid A, Aslam B, Muzammil S, et al. (2021) The prospects of antimicrobial coated medical implants. Journal of Applied Biomaterials & Functional Materials 1-21, sagepub.com/journals-permissions, DOI: 10.1177/22808000211040304, journals.sagepub.com/home/jbf

Sogn-Grundvåg G, Zhang D, Dreyer B (2020) Fishing methods for Atlantic cod and haddock: Quality and price versus costs. Fisheries Research 230 (2020) 10:56-72

Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F (2000) Methods for meta-analysis in medical research. Chichester, UK: John Wiley & Sons.

Swimmer Y, Zollett EA, Gutierrez A (2020) Bycatch mitigation of protected and threatened species in tuna purse seine and longline fisheries. Endangered Species Research 43: 517-542.

Takeshima N, Sozu T, Tajika A, Ogawa Y, Hayasaka Y, Furukawa TA (2014) Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? BMC Medical Research Methodology 2014, 14:30 http://www.biomedcentral.com/1471-2288/14/30

Thomas SN, Hridayanthan C (2006) The effect of natural sunlight on the strength of polyamide 6 multifilament fishing net materials. Fisheries Research 81: 326-330.

Thomas SN, John B, Edappazham G, Kalidas C, Meenakumari B (2009) Standardization of polyamide monofilament yarns for fabrication of gillnets with reference to physical and mechanical properties. Journal Marine Biological Association of India 51: 32-36.

Umemneku Chikere CM, Wilson K, Graziadio S, Vale L, Allen AJ (2019) Diagnostic test evaluation methodology: A systematic review of methods employed to evaluate diagnostic tests in the absence of gold standard – An update. PLoS ONE 14(10): e0223832. https://doi.org/10.1371/journal.pone.0223832

Vardon JL, Williams SM, Bucher DJ, Morgan JAT (2021) Identifying shark species responsible for fisheries depredation off Southeast Queensland, Australia. Molecular Biology Reports https://doi.org/10.1007/s11033-021-06460-4

Viechtbauer W (2010) Conducting meta-analyses in R with the `metafor` package. Journal of Statistical Software 36 (3), http://www.jstatsoft.org/

Wallace BC, Dahabreh IJ, Trikalinos TA, Lau J, Trow P, Schmid CH (2012) Closing the gap between methodologists and end-users: R as a computational back-end. Journal of Statistical Software, 49, 5.

Walter JF (2015) Update of standardized catch rates of large bluefin tuna (*Thunnus thynnus*) from the U.S. pelagic longline fishery in the Gulf of Mexico 1987-2013 with correction for weak hook effects. Collect. Vol. Sci. Pap. ICCAT, 71(3): 1260-1278 (2015)

Walter JF (2017) Update of standardized catch rates of large bluefin tuna (*Thunnus thynnus*) from the U.S. pelagic longline fishery in the Gulf of Mexico 1987-2016. Collect. Vol. Sci. Pap. ICCAT, 73(7): 2515-2527 (2017)

Welton NJ, Sutton AJ, Cooper NJ, Abrams KR, Ades AE (2012) Evidence synthesis for decision making in healthcare. John Wiley & Sons, Chichester, West Sussex, UK.

Werner TB, Kraus S, Read A, Zollett E (2006) Fishing techniques to reduce bycatch of threatened marine mammals. Marine Technology Society Journal 450-68.0:

Werner TB, Northridge S, Press KM, Young N (2015) Mitigating bycatch and depredation of marine mammals in longline fisheries. – ICES Journal of Marine Science, 72: 1576–1586.

Zhang E, Zhao X, Hu J, Wang R, Fu S, Qin G (2021) Antibacterial metals and alloys for potential biomedical implants. Bioactive Materials 6: 2569-2612

Tables

Table 1. Comparison of previous weak studies in the literature.

Source	Circle hook type	Diameter & Fabrication	Pull strength to unbend	Breaking Strength (strong – weak)	Effort (hooks)	No. straight hooks	No. straight x10^3 hooks	Notes
Foster & Bergmann (2010)	16/0 **strong**	4.0 mm ø, Duratin coating, no offset, Mustad model no. 39960D	110–125 kg (243-276lb)	Δ **53 kg (117 lb)**	99,303	63	0.634	Shallow-set longline fishery in the Gulf of Mexico targeting yellowfin tuna (*Thunnus albacares*), 311 sets. Focus of mitigation: spawning size class bluefin tuna (*T. thynnus*)
	16/0 **weak**	3.65 mm ø, Duratin coating, no offset, Mustad model no. 39988D	72 kg (159lb)		99,303	287[A]	2.890	
Bayse & Kerstetter (2010)	16/0 **strong**	3.2 mm x 4.0 mm ø, forged Lindgren-Pitman, oval cross-section	113 kg (249lb)	Δ **45 kg (99 lb)**	7,784	0	0	Shallow-set longline fishery targeting tuna species (yellowfin, bigeye tuna *T. obsesus*) and swordfish (*Xiphias gladius*) in the E. Atlantic (MAB[1]), 21 sets. Potential confounders: testing hooks from different manufacturers (i.e., differences in shape), small sample size. Focus of mitigation: short and longfin pilot whales whales (*Globicephala melas & G. macrorhyncus*)
	16/0 **weak**	3.55 mm ø wire, Mustad model #39960	68 kg (150lb)		7,784	7	0.899	

Table 1 cont.

Source	Circle hook type	Diameter & Fabrication	Pull strength to unbend	Breaking Strength (strong – weak)	Effort (hooks)	No. straight hooks	No. straight x10³ hooks	Notes
Bayse & Kerstetter (2010)	18/0 **strong**	3.6 mm x 5.0 mm ø forged Lindgren-Pitman, oval cross-section	159 kg (350lb)	**Δ 57 kg (125 lb)**	2,327	0	0	Shallow-set longline fishery targeting tuna species (yellowfin, bigeye tuna *T. obsesus*) and swordfish (*Xiphias gladius*) in the E. Atlantic (FEC[1], SAB[1]), 9 sets. Potential confounders: testing hooks from different manufacturers (i.e., differences in shape), small sample size. Focus of mitigation: short and longfin pilot whales whales (*Globicephala melas & G. macrorhyncus*)
	18/0 **weak**	4.95 mm ø wire, Mustad model #39960	102 kg (225lb)		2,327	1[a]	0.439	
Bigelow et al. (2012)	15/0 **strong**	4.5 mm ø, 10° offset, stainless wire	138 kg (304lb)	**Δ 45 kg (99 lb)**	151,369	6	0.040	Hawaii deep-set longline fishery targeting bigeye tuna and yellowfin tuna, 127 sets. Potential confounders: used a combination of ringed and non-ringed hooks, wire and monofilament leaders. Focus of mitigation: false killer whales (*Pseudorca crassidens*)
	15/0 **weak**	4.0 mm ø, 10° offset, stainless wire	93 kg (205lb)		151,369	76[B]	0.462	
Musyl & Phillips (2021)	15/0 **strong**	4.5 mm ø, 10° offset, stainless wire	230 kg (507 lb) *n*=5 tests	**Δ 29 kg (64 lb)**	107345	39	0.363	Hawaii deep-set longline fishery targeting bigeye tuna and yellowfin tuna, 178 sets, 2.3 mm ø monofilament gangions. Focus of mitigation: false killer whales (*Pseudorca crassidens*)
	15/0 **weak**	4.2 mm ø, 10° offset, stainless wire	201 kg 443 (lb) *n*=5 tests		107345D	58[C]	0.540	

[1]Statistical areas in the E. Atlantic: FEC=Florida East Coast, MAB=Mid-Atlantic Bight, SAB=South Atlantic Bight
[A]exact 2-tailed Cochran-Mantel-Haenszel test comparing to a 50:50 proportion was highly significant at $p=2.584 \times 10^{-19}$
[B]exact 2-tailed Cochran-Mantel-Haenszel test comparing to a 50:50 proportion was highly significant at $p=1.004 \times 10^{-9}$
[C]exact 2-tailed Cochran-Mantel-Haenszel test comparing to a 50:50 proportion was not significant at $p=0.196$
[D]number of straightened hooks collected on trips expressed on trips where information was provided.

Table 2. Species captured by hook type, percent change and IUCN status.

Common name	Species	ICUN Global Assessment *	Strong hooks (S)	Weak hooks (W)	Total (N)	% change (S-W)
Albacore	*Thunnus alalunga* (Bonnaterre, 1788)	NT	8	11	19	-0.38
Bigeye thresher	*Alopias superciliosus* (Lowe, 1841)	VU	45	46	91	-0.02
Bigeye tuna	*Thunnus obesus* (Lowe, 1839)	VU	427	444	871	-0.04
Black gemfish	*Nesiarchus nasutus* Johnson,1862	LC		2	2	
Black swallower	*Chiasmodon niger* Johnson, 1864	LC	1		1	
Blue marlin	*Makaira* nigricans Lacépède, 1802	VU	6	11	17	-0.83
Blue shark	*Prionace glauca* (Linnaeus, 1758)	NT	519	482	1001	0.07
Brama pomfret	*Brama japonica* Hilgendorf, 1878		6	5	11	0.17
Cigarfish	*Cubiceps spp.*		1		1	
Common mola	*Mola mola* (Linnaeus,1758)	VU		2	2	
Crocodile shark	*Pseudocarcharias kamoharai* (Matsubara,1936)	NT	4	2	6	0.50
Dagger pomfret	*Taractes rubescens* (Jordan & Evermann, 1887)		26	26	52	0.00
Dolphinfish	*Coryphaena hippurus* Linnaeus, 1758	LC	151	186	337	-0.23
Escolar	*Lepidocybium flavobrunneum* (Smith, 1843)	LC	115	135	250	-0.17
Pacific Fanfish	*Pteraclis aesticola* (Jordan & Snyder, 1901)	LC	1	1	2	0.00
Galapagos shark	*Carcharhinus galapagensis* (Snodgrass & Heller, 1905)	LC	1		1	
Great barracuda	*Sphyraena barracuda* (Edwards, 1771)	LC	1	2	3	-1.00
Hammerjaw	*Omosudis lowii* Günther,1887	LC	1		1	
Lancetfish	*Alepisaurus ferox* Lowe, 1833	LC	1429	1584	3013	-0.11
Longfin escolar	*Scombrolabrax heterolepis* Roule, 1921	LC	11	18	29	-0.64
Louvar	*Luvarus imperialis* Rafinesque, 1810	LC		1	1	
Mobula ray	*Mobula tarapacana* (Philippi, 1892)	VU	1	1	2	0.00
Oceanic whitetip shark	*Carcharhinus longimanus* (Poey, 1861)	CR	2	3	5	-0.50
Oilfish	*Ruvettus pretiosus* Cocco, 1833	LC	6	7	13	-0.17
Opah	*Lampris guttatus* (Brünnich, 1788)	LC	46	47	93	-0.02
Pelagic pomfret	*Brama orcini* Cuvier,1831 & *B. japonica* Hilgendorf,1878			1	1	
Pelagic stingray	*Pteroplatytrygon violacea* (Bonaparte, 1832)	LC	24	35	59	-0.46
Pompano dolphinfish	*Coryphaena equiselis* Linnaeus,1758	LC	3	4	7	-0.33
Roudi escolar	*Promenthichthys prometheus*	LC	1		1	
Rough pomfret	*Taractes asper* (Lowe,1843)	LC	2	1	3	0.5
Shortbill spearfish	*Tetrapturus angustirostris* Tanaka, 1915	DD	33	39	72	-0.18
Shortfin mako shark	*Isurus oxyrinchus* Rafinesque, 1810	EN	16	13	29	0.19
Sickle pomfret	*Taractichthys steindachneri* (Döderlein, 1883)		96	139	235	-0.45

Table 2 cont.

Common name	Species	ICUN Global Assessment *	Strong hooks (S)	Weak hooks (W)	Total (N)	% change (S-W)
Silky Shark	*Carcharhinus falciformis* (Müller & Henle, 1839)	VU	3	1	4	0.67
Skipjack tuna	*Katsuwonus pelamis* (Linnaeus, 1758)	LC	59	65	124	-0.10
Snake mackerel	*Gempylus serpens* Cuvier, 1829	LC	117	114	231	0.03
Striped marlin	*Kajikia audax* (Philippi, 1887)	NT	45	48	93	-0.07
Swordfish	*Xiphias gladius* Linnaeus, 1758	LC	52	52	104	0.00
Tapertail ribbonfish	*Trachipterus fukuzakii* Fitch, 1964	LC	2		2	
Tiger shark	*Galeocerdo cuvier* (Péron & Lesueur, 1822)	NT	1		1	
Unidentified billfish	*Istiophoridae*			3	3	
Unidentified pomfret	*Bramidae* Bonaparte, 1831 ????		1		1	
Unidentified species			1	1	2	
Unidentified tuna	*Thunnini* Starks, 1910		7	4	11	0.43
Velvet dogfish	*Scymnodon squamulosus* (Günther,1887)	DD	10	5	15	0.50
Wahoo	*Acanthocybium solandri* (Cuvier, 1832)	LC	125	148	273	-0.18
Yellowfin tuna	*Thunnus albacares* (Bonnaterre, 1788)	NT	129	125	254	0.03
Black-footed albatros	*Phoebastria nigripes* (Audubon, 1839)	NT		4	4	
Bottlenose dolphin	*Tursiops truncates* Montagu, 1821	LC		1	1	
False killer whale	*Pseudorca crassidens* (Owen, 1846)	NT	2	1	3	0.50
Unidentified shearwater				1	1	
GRAND TOTAL			**3536**	**3822**	**7358**	**-0.08**

DD-Data Deficient, LC- Least Concern, NT- Near Threatened, VU- Vulnerable, EN- Endangered, CR- Critically Endangered

Table 3. Length and weight for species captured in the study and weighed at auction. For istiophorid billfish and swordfish which were measured lower jaw to fork length (LJFK), all other length measurements are fork lengths (FL). The weights for larger fish are the dressed weights provided at the United Fishing Agency, Honolulu, HI.

	Length (cm)							Weight (lbs)						
	Strong			Weak			Dif. between Weak and Strong Hooks	Strong			Weak			Dif. between Weak and Strong Hooks
Species	N	Mean	SD	N	Mean	SD		N	Mean	SD	N	Mean	SD	
albacore	6	94.50	4.93	10	99.00	4.29	4.76%	6	36.33	5.85	10	41.40	4.60	13.96%
bigeye tuna	385	121.95	16.40	397	118.66	16.01	-2.70%	379	80.14	29.84	389	73.30	28.07	-8.54%
blue marlin	6	158.83	19.95	11	166.45	18.27	4.80%	5	83.40	22.71	8	102.50	28.74	22.90%
dolphinfish	124	72.92	15.41	154	71.84	15.23	-1.48%	102	8.35	5.26	117	7.53	4.89	-9.82%
escolar[2]	64	72.65	14.48	84	72.77	17.73	0.17%	33	17.18	9.95	43	21.72	12.64	26.43%
opah[1]	42	106.33	8.18	45	105.56	7.07	-0.72%	38	100.03	24.96	42	104.00	17.97	3.97%
spearfish	28	132.18	7.31	34	130.59	7.50	-1.20%	27	24.93	5.72	31	22.65	4.78	-9.15%
sickle pomfret[2]	87	59.51	10.17	128	60.28	10.49	1.29%	83	16.94	6.24	121	18.93	13.61	11.75%
skipjack tuna[2]	48	70.44	3.74	46	69.72	4.93	-1.02%	26	18.38	12.75	24	17.17	5.04	-6.58%
striped marlin	44	151.50	16.70	44	149.27	17.25	-1.47%	40	65.13	18.14	41	58.39	23.20	-10.35%
swordfish	43	164.21	36.66	42	169.88	38.58	3.45%	34	135.59	93.71	36	185.89	103.43	37.10%
wahoo	111	125.19	12.19	129	125.55	11.93	0.29%	103	25.73	9.62	114	25.82	10.50	0.35%
yellowfin tuna	125	125.66	25.08	114	124.13	20.97	-1.22%	123	81.59	28.08	105	78.37	30.32	-3.95%
Grand Mean							0.4%							5.2%

[1]Whole weights. [2]Species typically sold in lots

Note: The length (cm) and weight (cm) of fish species captured in the study and weighed at auction were compared between strong and weak hooks. Positive values of 'Dif. between Weak and Strong Hooks' indicate a greater mean length or weight caught on weak hooks when compared to strong hooks, while negative values indicate a greater mean length or weight caught on strong hooks when compared to weak hooks. For the targeted species, bigeye tuna, mean length (-2.7%) and weight (-8.54%) of catch was greater on strong hooks when compared to weak hooks, however the percent difference was within the TRT's set threshold for reduction in weight by less than 10%. Collectively examining the size of all species sold at auction, the grand mean difference between strong and weak hooks for length (0.4%) and weight (5.2%) were within TRT's threshold of revenue loss (<10% reduction).

Table 4. Sales information for species sold at the United Fishing Agency (UFA), Honolulu, HI. Data kindly provided by UFA.

	Price per pound (lb.)							Sales Price						
	Strong			Weak			Dif. between Strong and Weak Hooks	Strong			Weak			Dif. between Strong and Weak Hooks
Species	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**		**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	
albacore	6	1.91	0.83	10	2.56	1.28	34.0%	6	71.15	37.79	10	109.56	61.96	54.0%
bigeye tuna	375	6.79	1.98	388	6.72	2.06	-1.0%	375	569.08	307.15	388	516.19	309.60	-9.3%
blue marlin	5	5.42	2.79	8	3.90	1.83	-28.0%	5	445.86	273.30	8	408.91	248.69	-8.3%
dolphinfish	108	4.40	2.70	127	4.24	2.58	-3.6%	101	44.99	41.81	117	40.29	43.88	-10.4%
escolar	24	1.23	1.69	37	0.72	1.19	-41.5%	23	19.92	28.95	35	17.43	38.53	-12.5%
opah	38	4.11	1.43	42	4.22	1.52	2.7%	38	412.53	173.13	42	444.32	177.67	7.7%
spearfish	27	2.96	1.88	31	2.38	1.62	-19.6%	27	71.87	42.00	31	55.54	39.77	-22.7%
sickle pomfret	83	5.87	2.27	121	5.93	2.16	1.0%	83	99.61	51.61	121	114.26	87.81	14.7%
skipjack tuna	35	1.87	1.23	28	1.48	0.52	-20.9%	26	49.09	125.76	23	24.36	14.15	-50.4%
striped marlin	40	3.60	1.10	41	3.44	1.27	-4.4%	40	231.19	85.17	41	213.71	135.65	-7.6%
swordfish	34	4.25	1.82	35	5.13	1.71	20.7%	34	640.50	532.56	35	1048.83	637.92	63.8%
wahoo	102	3.78	2.31	114	3.82	2.26	1.1%	102	99.84	69.71	114	102.89	82.98	3.1%
yellowfin tuna	123	5.95	2.40	105	5.92	2.03	-0.5%	123	515.00	272.60	105	479.89	256.85	-6.8%
Grand Mean							-4.6%							1.18%

Note: The price per pound and sales price of fish species sold at the UFA were compared between strong and weak hooks. Positive values of 'Dif. between Weak and Strong Hooks' values indicates a greater mean price per pound or sales price on weak hooks when compared to strong hooks, while a negative value indicates a greater mean price per pound or sales price on strong hooks when compared to weak hooks. For the targeted species, bigeye tuna, price per pound (-1.0%) and sales price (-9.3%) were on average greater on strong hooks when compared to weak hooks, however the percent difference was within the TRT's set threshold for reduction in value by less than 10%. Collectively examining all species sold at the UFA in Honolulu, the grand mean difference between strong and weak hooks for price per pound (-4.6%) and sales price (1.18%) were within TRT's threshold of revenue loss (<10% reduction).

Table 5. Total *ex*-vessel revenue (gross) for species sold at auction by hook type for the study. Data kindly provided by United Fishing Agency, Honolulu, HI.

Species	Strong Hooks ($)	Weak Hooks ($)
abacore	426.90	1095.60
bigeye tuna	213406.30	200282.50
blue marlin	2229.30	3271.30
blue shark	154.00	
dolphinfish	4543.51	4713.66
escolar	458.20	609.90
oilfish	3.20	
opah	15676.30	18661.30
spearfish	1940.60	1721.70
sickle pomfret	8267.40	13826.00
skipjack tuna	1276.30	560.30
striped marlin	9247.40	8762.10
swordfish	21776.90	36709.10
wahoo	10183.80	11729.50
yellowfin tuna	63345.10	50388.10
Grand Total	**$352,935.20**	**$352,331.10**

Note: The total ex-vessel gross revenue for all species sold at auction that were caught on strong hooks was only $604.15 greater than the total gross revenue for all species sold at auction that were caught on weak hooks.

Table 6. Nominal catch per unit of effort expressed as numbers per 1000 hooks. Depredation damage: DP – species show signs of depredation of unknown origin; MM – marine mammals; SX – shark damage. Descriptions of codes are from the Pacific Islands Regional Observer Program (PRIOP) observer's manual.

	CPUE						Depredation Damage					
							Strong hooks			Weak hooks		
Species	N Strong	Mean Strong	SD	N Weak	Mean Weak	SD	DP	MM	SX	DP	MM	SX
albacore	8	0.029	0.078	11	0.042	0.092		1				1
bigeye thresher	45	0.170	0.344	46	0.162	0.309						
bigeye tuna	427	1.819	0.972	444	1.908	1.103			8	4		5
blue marlin	6	0.022	0.031	11	0.039	0.062						
blue shark	519	2.023	0.618	482	1.918	0.558						1
dagger pomfret	26	0.109	0.098	26	0.133	0.158						
dolphinfish	151	0.550	0.585	186	0.673	0.871	2	1	2	2		
escolar	115	0.469	0.321	135	0.548	0.408	1			4		1
lancetfish	1429	5.804	3.279	1584	6.300	3.683	78		3	96		4
longfin escolar	11	0.052	0.054	18	0.067	0.069				1		
opah	46	0.184	0.264	47	0.196	0.254		3			2	
pelagic stingray	24	0.084	0.143	35	0.119	0.293						
spearfish	33	0.121	0.082	39	0.151	0.149	1	1	1		1	1
shortfin mako	16	0.061	0.052	13	0.048	0.062						
sickle pomfret	96	0.372	0.264	139	0.539	0.348						
skipjack tuna	59	0.227	0.320	65	0.246	0.333	1		1	2		2
snake mackerel	117	0.512	0.350	114	0.460	0.344	11			8		
striped marlin	45	0.163	0.160	48	0.206	0.228			1			1
swordfish	52	0.204	0.142	52	0.192	0.196				1	2	
velvet dogfish	10	0.035	0.059	5	0.016	0.041						
wahoo	125	0.503	0.334	148	0.572	0.429	6	1	3	9	1	3
yellowfin tuna	129	0.487	0.594	125	0.476	0.599			1	1		3

Table 7. Mortality estimates for species caught on strong and weak hooks from available data. The *p*-values are from exact 2-tailed Cochran-Mantel-Haenszel tests. To account for multiple tests of the same hypothesis inflating Type I errors, authors have suggested (e.g., Manly 2005) adjustments to the α level. If 24 independent tests are performed at a starting α of 0.05, then the probability of at least one significant result by chance alone is $1\text{-}0.95^{24} = 0.71$. The Bonferroni corrected α level for this table is 0.0042.

	Strong hooks		Weak hooks		
Species	**Alive**	**Dead**	**Alive**	**Dead**	***p*-value**
albacore	0	8	2	9	0.4854
bigeye thresher	43	2	43	3	0.6655
bigeye tuna	311	114	329	115	0.7587
blue marlin	4	2	5	6	0.6199
blue shark	510	9	481	1	0.0217
brama pomfret	6	0	4	1	0.4545
dagger pomfret	26	0	24	2	0.4902
dolphinfish	53	98	72	114	0.4992
escolar	103	12	117	18	0.5602
lancetfish	560	869	644	940	0.4129
longfin escolar	5	6	9	8	0.7040
oilfish	6	0	7	0	
opah	21	25	22	25	0.9114
pelagic stingray	24	0	35	0	
spearfish	7	26	6	33	0.5538
shortfin mako	14	2	11	2	0.8258
sickle pomfret	93	3	137	2	0.6510
skipjack tuna	1	58	6	59	0.1174
snake mackerel	74	43	87	27	0.0328
striped marlin	17	28	24	24	0.2972
swordfish	33	19	27	25	0.3210
velvet dogfish	9	1	5	0	0.4795
wahoo	17	108	11	137	0.1107
yellowfin tuna	63	66	60	65	0.9008
Grand Total	**2000**	**1499**	**2168**	**1616**	**0.9245**

Figure Legends

Figure 1. Forest plot for the effect size of catch rates by hook type (Tables 2, 6) using the risk ratio (*RR*). Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Effect sizes and 95% CIs are provided for each study. I^2, the amount of variability among studies within species, along with a *p* value testing for heterogeneity (Cochran's Q), are provided. The area of the boxes for each study is proportional to the inverse of the variance, and any side of the box is proportional to the inverse of the standard error. The 95% confidence intervals (horizontal bars) for each study are proportional to the standard error and are related to sample size. The diamonds represent the summary effect size and the width is proportional to the 95% CI. Note especially the much narrower widths of the diamonds indicating more precision in the estimates.

Figure 2. Forest plot for the effect size of catch rates by hook type using the risk ratio (*RR*) on subgroups. Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Descriptions follow Figure 1.

Figure 3. Sensitivity analysis for the effect size of catch rates by hook type using the risk ratio (*RR*). Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Descriptions follow Figure 1. In this analysis, one study is removed and the summary effect is calculated, and then the next study is removed and the summary effect is recalculated and so on. The plot indicated switching back and forth around the summary effect size but no single study had a significant impact on the summary effect because the 95% CIs overlapped

Figure 4. Forest plot for the effect size of catch rates by hook type using the risk ratio (*RR*) grouped by species. Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Descriptions follow Figure 1.

Figure 5. Forest plot for the effect size of body lengths (Table 3) using the mean difference (*MD*). Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Descriptions follow Figure 1.

Figure 6. Forest plot for the effect size of body lengths (Table 3) using the standardized mean difference (*SMD*). Study number is provided (brackets where [3] = Bigelow et al. (2012) and [5] = the current study). Descriptions follow Figure 1.

Figure 7. Forest plot for the effect size of weight at auction (Table 3) using the mean difference (*MD*). Descriptions follow Figure 1.

Figure 8. Forest plot for the effect size of weight at auction (Table 4) using the mean difference (*MD*) for swordfish. Descriptions follow Figure 1.

Figure 9. Forest plot for the effect size of price per pound at auction (Table 5) using the mean difference (*MD*). Descriptions follow Figure 1.

Figure 10. Forest plot for the effect size of average price per fish auction (Table 5) using the mean difference (*MD*). Descriptions follow Figure 1.

Figure 11 Receiver operator characteristic (ROC) curve for the prognostic variables used in the analysis to test whether they were accurate in discriminating catch by hook type. Sensitivity is the true positive rate and 100-specificity is the false positive rate. The further away from the diagonal line represents better discriminatory power of the particular variable in classifying survival outcomes. None of the variables were significant in the logistic regression model and resolving power was essentially "50-50" in terms of correctly assigning a hook type based on these variables. Areas under the curve (AUC) for the moderators were: 0.521 ± 0.0113 (SE) [95% CI 0.499-0.543] for the Group variable (i.e., target, retained, elasmobranch, istiophorid billfish, discards); 0.507 ± 0.0123 [0.485-0.529] for species, 0.514 ± 0.0127 [0.492-0.535] for price per lb., and 0.517 ± 0.0126 [0.496-0.539] for latitude.

Studies	Estimate (95% C.I.)
[3]bigeye tuna 15/0	0.980 (0.895, 1.072)
[3]yellowfin tuna 15/0	0.700 (0.507, 0.966)
[3]albacore 15/0	1.409 (0.816, 2.433)
[3]wahoo 15/0	0.800 (0.487, 1.315)
[3]skipjack tuna 15/0	0.866 (0.684, 1.096)
[3]swordfish 15/0	1.033 (0.626, 1.707)
[3]spearfish 15/0	1.760 (1.077, 2.875)
[3]striped marlin 15/0	1.087 (0.617, 1.915)
[3]sailfish 15/0	0.500 (0.171, 1.463)
[3]blue shark 15/0	1.027 (0.915, 1.152)
[3]pelagic stingray 15/0	1.140 (0.757, 1.716)
[3]bigeye thresher 15/0	1.000 (0.546, 1.831)
[3]shortfin mako 15/0	1.111 (0.452, 2.734)
[3]dolphinfish 15/0	1.114 (0.979, 1.266)
[3]opah 15/0	1.116 (0.835, 1.493)
[3]lancetfish 15/0	1.104 (0.990, 1.231)
[3]snake mackerel 15/0	0.971 (0.841, 1.121)
[3]escolar 15/0	0.813 (0.667, 0.990)
[3]longfin escolar 15/0	0.727 (0.293, 1.808)
[3]sickle pomfret 15/0	0.921 (0.724, 1.171)
[3]dagger pomfret 15/0	0.273 (0.076, 0.978)
[3]great barracuda 15/0	0.923 (0.421, 2.023)
Subgroup [3]Hawaii (I^2=34% , P=0.062)	**0.994 (0.928, 1.063)**
[5]albacore 15/0	0.727 (0.293, 1.808)
[5]bigeye thresher15/0	0.978 (0.649, 1.475)
[5]bigeye tuna 15/0	0.962 (0.842, 1.098)
[5]blue marlin 15/0	0.545 (0.202, 1.475)
[5]blue shark 15/0	1.077 (0.951, 1.219)
[5]dagger pomfret 15/0	1.000 (0.581, 1.722)
[5]dolphinfish 15/0	0.812 (0.655, 1.006)
[5]escolar 15/0	0.852 (0.664, 1.092)
[5]lancetfish 15/0	0.902 (0.840, 0.969)
[5]longfin escolar 15/0	0.611 (0.289, 1.294)
[5]opah 15/0	0.979 (0.652, 1.470)
[5]pelagic stingray 15/0	0.686 (0.408, 1.153)
[5]spearfish 15/0	0.846 (0.532, 1.345)
[5]shortfin mako 15/0	1.231 (0.592, 2.559)
[5]sickle pomfret 15/0	0.691 (0.533, 0.896)
[5]skipjack tuna 15/0	0.908 (0.638, 1.291)
[5]snake mackerel 15/0	1.026 (0.793, 1.328)
[5]striped marlin 15/0	0.937 (0.624, 1.408)
[5]swordfish 15/0	1.000 (0.681, 1.469)
[5]velvet dogfish 15/0	2.000 (0.684, 5.851)
[5]wahoo 15/0	0.845 (0.666, 1.072)
[5]yellowfin tuna 15/0	1.032 (0.807, 1.320)
Subgroup [5]Hawaii (I^2=5% , P=0.392)	**0.926 (0.879, 0.975)**
Overall (I^2=29% , P=0.040)	**0.959 (0.915, 1.005)**

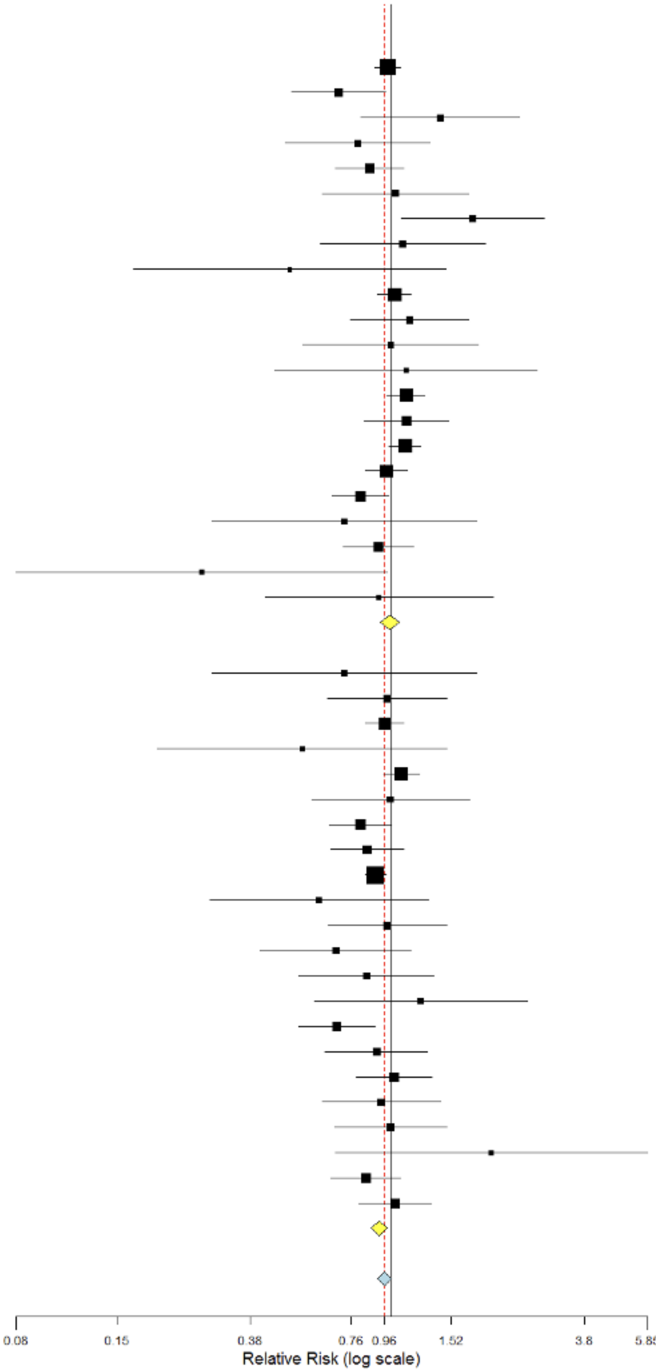


Figure 1

Studies	Estimate (95% C.I.)
[3]bigeye tuna 15/0	0.980 (0.895, 1.072)
[3]yellowfin tuna 15/0	0.700 (0.507, 0.966)
[3]swordfish 15/0	1.033 (0.626, 1.707)
[5]bigeye tuna 15/0	0.962 (0.842, 1.098)
[5]swordfish 15/0	1.000 (0.681, 1.469)
[5]yellowfin tuna 15/0	1.032 (0.807, 1.320)
Subgroup Target (I^2=0% , P=0.504)	**0.966 (0.903, 1.034)**
[3]albacore 15/0	1.409 (0.816, 2.433)
[3]wahoo 15/0	0.800 (0.487, 1.315)
[3]skipjack tuna 15/0	0.866 (0.684, 1.096)
[3]dolphinfish 15/0	1.114 (0.979, 1.266)
[3]opah 15/0	1.116 (0.835, 1.493)
[3]escolar 15/0	0.813 (0.667, 0.990)
[3]sickle pomfret 15/0	0.921 (0.724, 1.171)
[5]albacore 15/0	0.727 (0.293, 1.808)
[5]dolphinfish 15/0	0.812 (0.655, 1.006)
[5]escolar 15/0	0.852 (0.664, 1.092)
[5]opah 15/0	0.979 (0.652, 1.470)
[5]sickle pomfret 15/0	0.691 (0.533, 0.896)
[5]skipjack tuna 15/0	0.908 (0.638, 1.291)
[5]wahoo 15/0	0.845 (0.666, 1.072)
Subgroup Retained (I^2=40% , P=0.059)	**0.902 (0.821, 0.991)**
[3]spearfish 15/0	1.760 (1.077, 2.875)
[3]striped marlin 15/0	1.087 (0.617, 1.915)
[3]sailfish 15/0	0.500 (0.171, 1.463)
[5]blue marlin 15/0	0.545 (0.202, 1.475)
[5]spearfish 15/0	0.846 (0.532, 1.345)
[5]striped marlin 15/0	0.937 (0.624, 1.408)
Subgroup Istiophorid billfish (I^2=43% , P=0.118)	**0.980 (0.714, 1.346)**
[3]blue shark 15/0	1.027 (0.915, 1.152)
[3]pelagic stingray 15/0	1.140 (0.757, 1.716)
[3]bigeye thresher 15/0	1.000 (0.546, 1.831)
[3]shortfin mako 15/0	1.111 (0.452, 2.734)
[5]bigeye thresher15/0	0.978 (0.649, 1.475)
[5]blue shark 15/0	1.077 (0.951, 1.219)
[5]pelagic stingray 15/0	0.686 (0.408, 1.153)
[5]shortfin mako 15/0	1.231 (0.592, 2.559)
[5]velvet dogfish 15/0	2.000 (0.684, 5.851)
Subgroup Elasmobranch (I^2=0% , P=0.784)	**1.045 (0.966, 1.130)**
[3]lancetfish 15/0	1.104 (0.990, 1.231)
[3]snake mackerel 15/0	0.971 (0.841, 1.121)
[3]longfin escolar 15/0	0.727 (0.293, 1.808)
[3]dagger pomfret 15/0	0.273 (0.076, 0.978)
[3]great barracuda 15/0	0.923 (0.421, 2.023)
[5]dagger pomfret 15/0	1.000 (0.581, 1.722)
[5]lancetfish 15/0	0.902 (0.840, 0.969)
[5]longfin escolar 15/0	0.611 (0.289, 1.294)
[5]snake mackerel 15/0	1.026 (0.793, 1.328)
Subgroup Discard (I^2=47% , P=0.058)	**0.967 (0.868, 1.077)**
Overall (I^2=29% , P=0.040)	**0.959 (0.915, 1.005)**

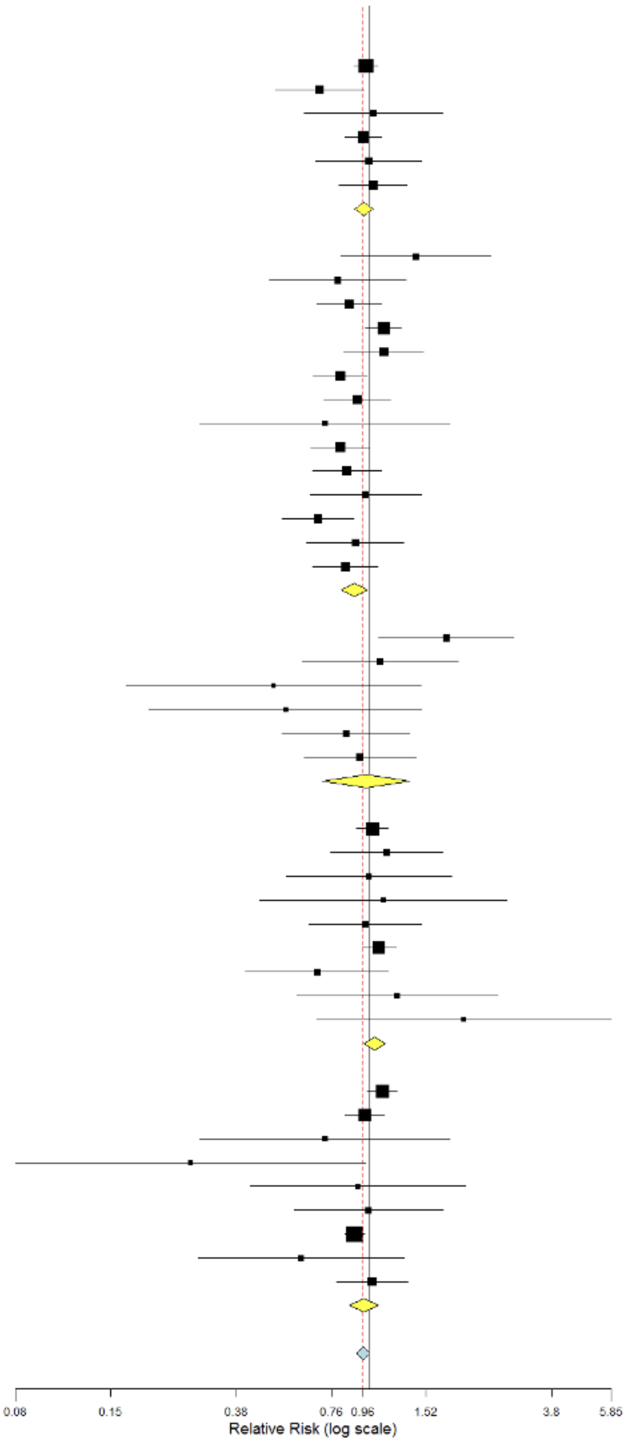


Figure 2

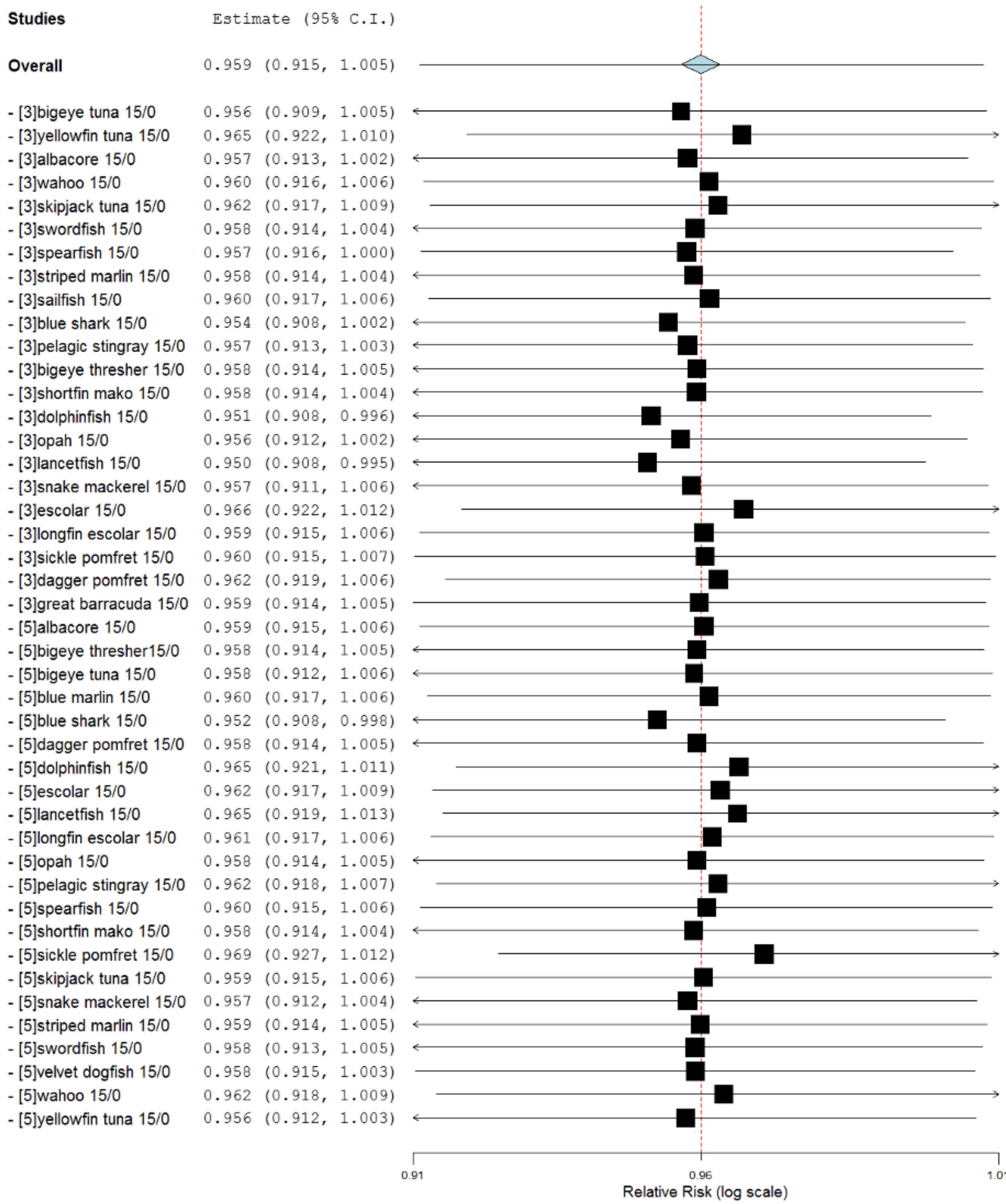

Studies
Estimate (95% C.I.)
Overall 0.959 (0.915, 1.005)
- [3]bigeye tuna 15/0 0.956 (0.909, 1.005)
- [3]yellowfin tuna 15/0 0.965 (0.922, 1.010)
- [3]albacore 15/0 0.957 (0.913, 1.002)
- [3]wahoo 15/0 0.960 (0.916, 1.006)
- [3]skipjack tuna 15/0 0.962 (0.917, 1.009)
- [3]swordfish 15/0 0.958 (0.914, 1.004)
- [3]spearfish 15/0 0.957 (0.916, 1.000)
- [3]striped marlin 15/0 0.958 (0.914, 1.004)
- [3]sailfish 15/0 0.960 (0.917, 1.006)
- [3]blue shark 15/0 0.954 (0.908, 1.002)
- [3]pelagic stingray 15/0 0.957 (0.913, 1.003)
- [3]bigeye thresher 15/0 0.958 (0.914, 1.005)
- [3]shortfin mako 15/0 0.958 (0.914, 1.004)
- [3]dolphinfish 15/0 0.951 (0.908, 0.996)
- [3]opah 15/0 0.956 (0.912, 1.002)
- [3]lancetfish 15/0 0.950 (0.908, 0.995)
- [3]snake mackerel 15/0 0.957 (0.911, 1.006)
- [3]escolar 15/0 0.966 (0.922, 1.012)
- [3]longfin escolar 15/0 0.959 (0.915, 1.006)
- [3]sickle pomfret 15/0 0.960 (0.915, 1.007)
- [3]dagger pomfret 15/0 0.962 (0.919, 1.006)
- [3]great barracuda 15/0 0.959 (0.914, 1.005)
- [5]albacore 15/0 0.959 (0.915, 1.006)
- [5]bigeye thresher15/0 0.958 (0.914, 1.005)
- [5]bigeye tuna 15/0 0.958 (0.912, 1.006)
- [5]blue marlin 15/0 0.960 (0.917, 1.006)
- [5]blue shark 15/0 0.952 (0.908, 0.998)
- [5]dagger pomfret 15/0 0.958 (0.914, 1.005)
- [5]dolphinfish 15/0 0.965 (0.921, 1.011)
- [5]escolar 15/0 0.962 (0.917, 1.009)
- [5]lancetfish 15/0 0.965 (0.919, 1.013)
- [5]longfin escolar 15/0 0.961 (0.917, 1.006)
- [5]opah 15/0 0.958 (0.914, 1.005)
- [5]pelagic stingray 15/0 0.962 (0.918, 1.007)
- [5]spearfish 15/0 0.960 (0.915, 1.006)
- [5]shortfin mako 15/0 0.958 (0.914, 1.004)
- [5]sickle pomfret 15/0 0.969 (0.927, 1.012)
- [5]skipjack tuna 15/0 0.959 (0.915, 1.006)
- [5]snake mackerel 15/0 0.957 (0.912, 1.004)
- [5]striped marlin 15/0 0.959 (0.914, 1.005)
- [5]swordfish 15/0 0.958 (0.913, 1.005)
- [5]velvet dogfish 15/0 0.958 (0.915, 1.003)
- [5]wahoo 15/0 0.962 (0.918, 1.009)
- [5]yellowfin tuna 15/0 0.956 (0.912, 1.003)
0.91
0.96
1.01
Relative Risk (log scale)

Figure 3

Studies	Estimate (95% C.I.)
[3]bigeye tuna 15/0	0.980 (0.895, 1.072)
[5]bigeye tuna 15/0	0.962 (0.842, 1.098)
Subgroup bigeye tuna (I^2=0% , P=0.818)	**0.974 (0.904, 1.050)**
[3]yellowfin tuna 15/0	0.700 (0.507, 0.966)
[5]yellowfin tuna 15/0	1.032 (0.807, 1.320)
Subgroup yellowfin tuna (I^2=72% , P=0.060)	**0.862 (0.590, 1.260)**
[3]albacore 15/0	1.409 (0.816, 2.433)
[5]albacore 15/0	0.727 (0.293, 1.808)
Subgroup albacore (I^2=33% , P=0.222)	**1.124 (0.608, 2.079)**
[3]wahoo 15/0	0.800 (0.487, 1.315)
[5]wahoo 15/0	0.845 (0.666, 1.072)
Subgroup wahoo (I^2=0% , P=0.847)	**0.836 (0.675, 1.036)**
[3]skipjack tuna 15/0	0.866 (0.684, 1.096)
[5]skipjack tuna 15/0	0.908 (0.638, 1.291)
Subgroup skipjack tuna (I^2=0% , P=0.827)	**0.879 (0.722, 1.069)**
[3]swordfish 15/0	1.033 (0.626, 1.707)
[5]swordfish 15/0	1.000 (0.681, 1.469)
Subgroup swordfish (I^2=0% , P=0.919)	**1.012 (0.746, 1.373)**
[3]spearfish 15/0	1.760 (1.077, 2.875)
[5]spearfish 15/0	0.846 (0.532, 1.345)
Subgroup spearfish (I^2=78% , P=0.033)	**1.215 (0.593, 2.490)**
[3]striped marlin 15/0	1.087 (0.617, 1.915)
[5]striped marlin 15/0	0.937 (0.624, 1.408)
Subgroup striped marlin (I^2=0% , P=0.678)	**0.986 (0.709, 1.372)**
[3]blue marlin 15/0	0.889 (0.343, 2.304)
[5]blue marlin 15/0	0.545 (0.202, 1.475)
Subgroup blue marlin (I^2=0% , P=0.487)	**0.704 (0.354, 1.400)**
[3]dolphinfish 15/0	1.114 (0.979, 1.266)
[5]dolphinfish 15/0	0.812 (0.655, 1.006)
Subgroup dolphinfish (I^2=84% , P=0.013)	**0.962 (0.707, 1.311)**
[3]opah 15/0	1.116 (0.835, 1.493)
[5]opah 15/0	0.979 (0.652, 1.470)
Subgroup opah (I^2=0% , P=0.606)	**1.068 (0.843, 1.353)**
[3]escolar 15/0	0.813 (0.667, 0.990)
[5]escolar 15/0	0.852 (0.664, 1.092)
Subgroup escolar (I^2=0% , P=0.772)	**0.828 (0.709, 0.966)**
[3]sickle pomfret 15/0	0.921 (0.724, 1.171)
[5]sickle pomfret 15/0	0.691 (0.533, 0.896)
Subgroup sickle pomfret (I^2=61% , P=0.111)	**0.801 (0.604, 1.062)**
Overall (I^2=28% , P=0.092)	**0.931 (0.874, 0.992)**

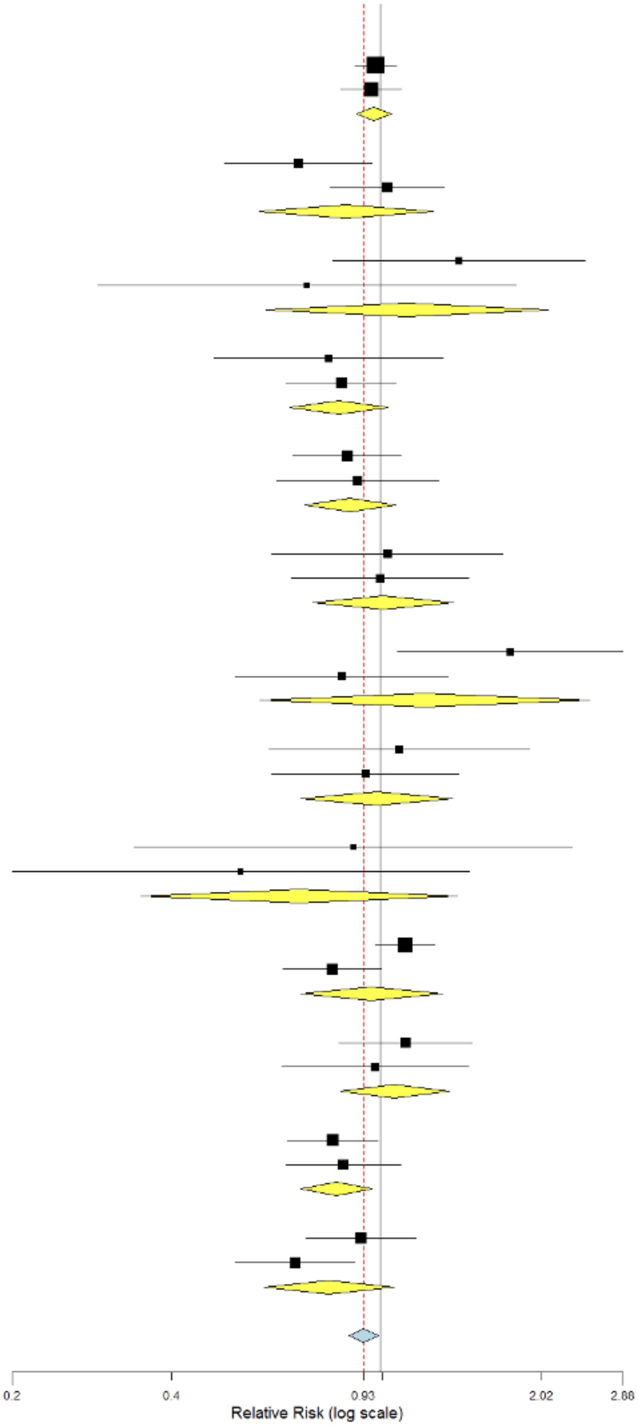


Figure 4

Studies	Estimate (95% C.I.)		
[3]bigeye tuna 15/0	-0.700	(-2.940,	1.540)
[3]yellowfin tuna 15/0	0.000	(-7.339,	7.339)
[3]swordfish 15/0	4.300	(-11.282,	19.882)
[5]bigeye tuna 15/0	3.286	(1.013,	5.558)
[5]swordfish 15/0	-5.672	(-21.680,	10.336)
[5]yellowfin tuna 15/0	1.531	(-4.312,	7.374)
Subgroup Target (I^2=28% , P=0.222)	**1.142**	**(-1.035,**	**3.319)**
[3]albacore 15/0	-2.800	(-6.662,	1.062)
[3]wahoo 15/0	3.100	(-3.195,	9.395)
[3]skipjack tuna 15/0	-0.400	(-2.504,	1.704)
[3]dolphinfish 15/0	1.300	(-1.405,	4.005)
[3]opah 15/0	0.200	(-2.304,	2.704)
[3]escolar 15/0	2.500	(-2.818,	7.818)
[3]sickle pomfret 15/0	0.400	(-4.048,	4.848)
[5]albacore tuna 15/0	-4.500	(-9.258,	0.258)
[5]dolphinfish 15/0	1.075	(-2.550,	4.700)
[5]escolar 15/0	-0.126	(-5.317,	5.066)
[5]opah 15/0	0.778	(-2.446,	4.001)
[5]sickle pomfret 15/0	-0.770	(-3.576,	2.036)
[5]skipjack tuna 15/0	0.720	(-1.054,	2.495)
[5]wahoo 15/0	-0.361	(-3.424,	2.702)
Subgroup Retained (I^2=0% , P=0.711)	**0.109**	**(-0.706,**	**0.924)**
[3]spearfish 15/0	-3.700	(-10.262,	2.862)
[3]striped marlin 15/0	-8.800	(-18.938,	1.338)
[3]blue marlin 15/0	-15.600	(-53.464,	22.264)
[5]blue marlin 15/0	-7.621	(-26.896,	11.654)
[5]spearfish 15/0	1.590	(-2.110,	5.290)
[5]striped marlin 15/0	2.227	(-4.868,	9.323)
Subgroup Istiophorid billfish (I^2=22% , P=0.269)	**-1.080**	**(-4.736,**	**2.576)**
[3]lancetfish 15/0	0.000	(-5.568,	5.568)
[3]snake mackerel 15/0	-2.900	(-8.726,	2.926)
[5]lancetfish 15/0	-0.838	(-7.383,	5.708)
[5]snake mackerel 15/0	7.313	(-3.914,	18.540)
Subgroup Discard (I^2=0% , P=0.466)	**-0.505**	**(-3.784,**	**2.774)**
Overall (I^2=0% , P=0.532)	**0.281**	**(-0.397,**	**0.959)**

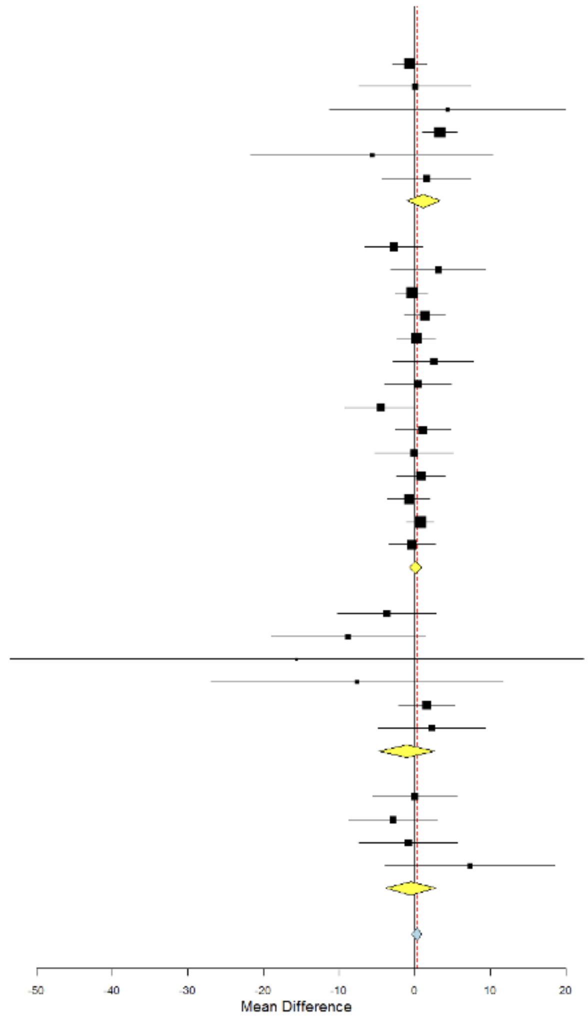


Figure 5

Studies	Estimate (95% C.I.)
[3]bigeye tuna 15/0	-0.029 (-0.121, 0.064)
[3]yellowfin tuna 15/0	0.000 (-0.323, 0.323)
[3]swordfish 15/0	0.143 (-0.387, 0.673)
[5]bigeye tuna 15/0	0.203 (0.062, 0.343)
[5]swordfish 15/0	-0.149 (-0.575, 0.276)
[5]yellowfin tuna 15/0	0.066 (-0.188, 0.320)
Subgroup Target (I^2=40% , P=0.142)	**0.052 (-0.063, 0.167)**
[3]albacore 15/0	-0.393 (-0.945, 0.159)
[3]wahoo 15/0	0.309 (-0.353, 0.970)
[3]skipjack tuna 15/0	-0.046 (-0.291, 0.199)
[3]dolphinfish 15/0	0.100 (-0.109, 0.308)
[3]opah 15/0	0.023 (-0.276, 0.323)
[3]escolar 15/0	0.158 (-0.179, 0.495)
[3]sickle pomfret 15/0	0.033 (-0.339, 0.406)
[5]albacore tuna 15/0	-0.939 (-2.002, 0.124)
[5]dolphinfish 15/0	0.070 (-0.167, 0.307)
[5]escolar 15/0	-0.008 (-0.333, 0.318)
[5]opah 15/0	0.101 (-0.320, 0.522)
[5]sickle pomfret 15/0	-0.074 (-0.346, 0.198)
[5]skipjack tuna 15/0	0.164 (-0.241, 0.569)
[5]wahoo 15/0	-0.030 (-0.284, 0.224)
Subgroup Retained (I^2=0% , P=0.779)	**0.023 (-0.061, 0.107)**
[3]spearfish 15/0	-0.260 (-0.768, 0.248)
[3]striped marlin 15/0	-0.479 (-1.053, 0.095)
[3]blue marlin 15/0	-0.403 (-1.427, 0.622)
[5]blue marlin 15/0	-0.384 (-1.387, 0.619)
[5]spearfish 15/0	0.212 (-0.290, 0.713)
[5]striped marlin 15/0	0.130 (-0.288, 0.548)
Subgroup Istiophorid billfish (I^2=6% , P=0.377)	**-0.094 (-0.337, 0.148)**
[3]lancetfish 15/0	0.000 (-0.206, 0.206)
[3]snake mackerel 15/0	-0.132 (-0.398, 0.134)
[5]lancetfish 15/0	-0.028 (-0.246, 0.190)
[5]snake mackerel 15/0	0.394 (-0.225, 1.012)
Subgroup Discard (I^2=0% , P=0.486)	**-0.023 (-0.151, 0.104)**
Overall (I^2=0% , P=0.600)	**0.018 (-0.030, 0.067)**

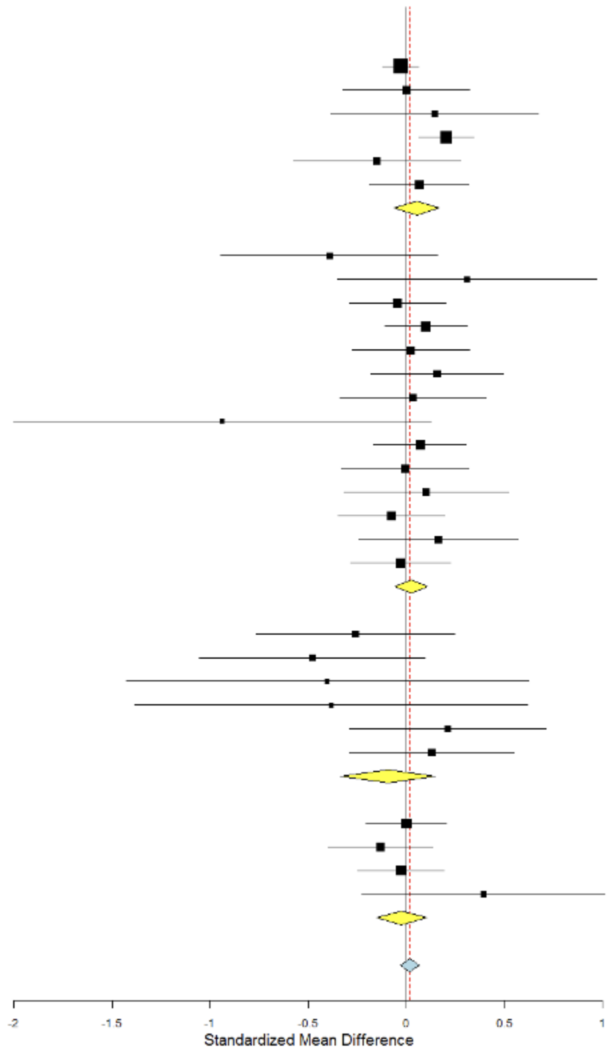


Figure 6

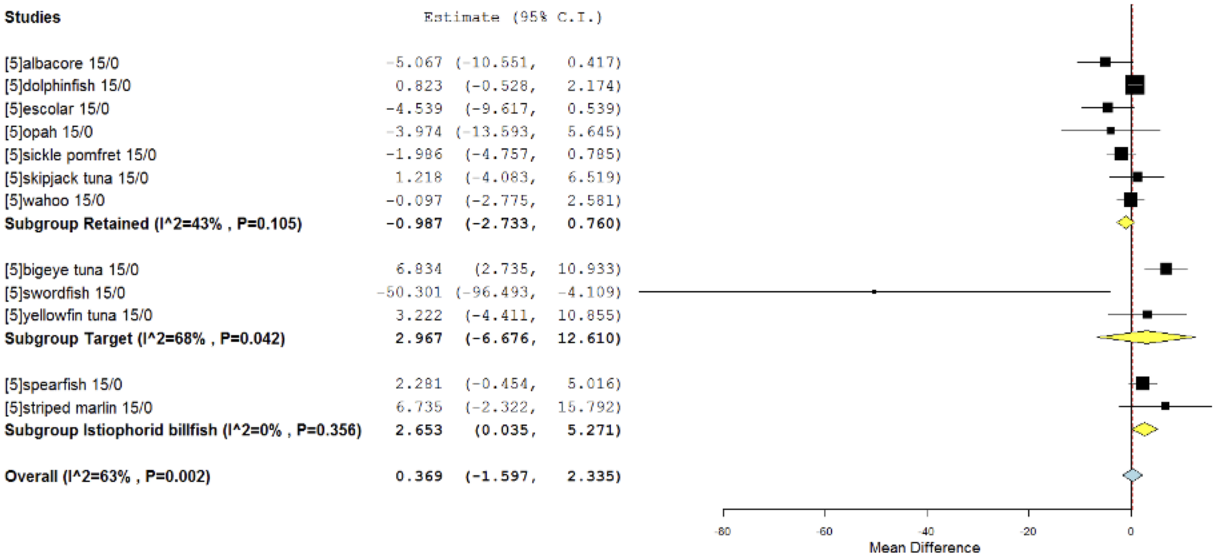


Figure 7

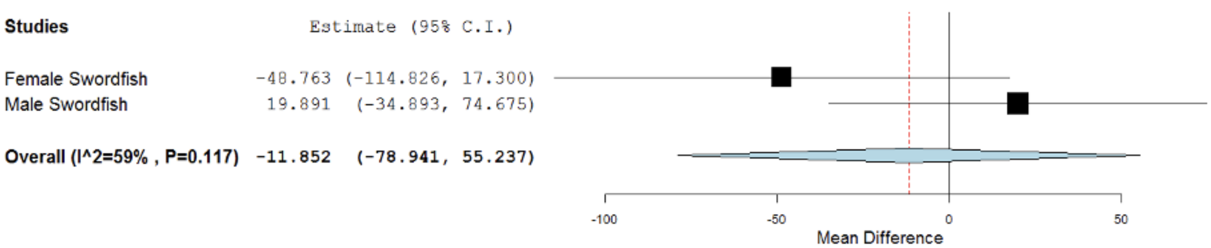


Figure 8

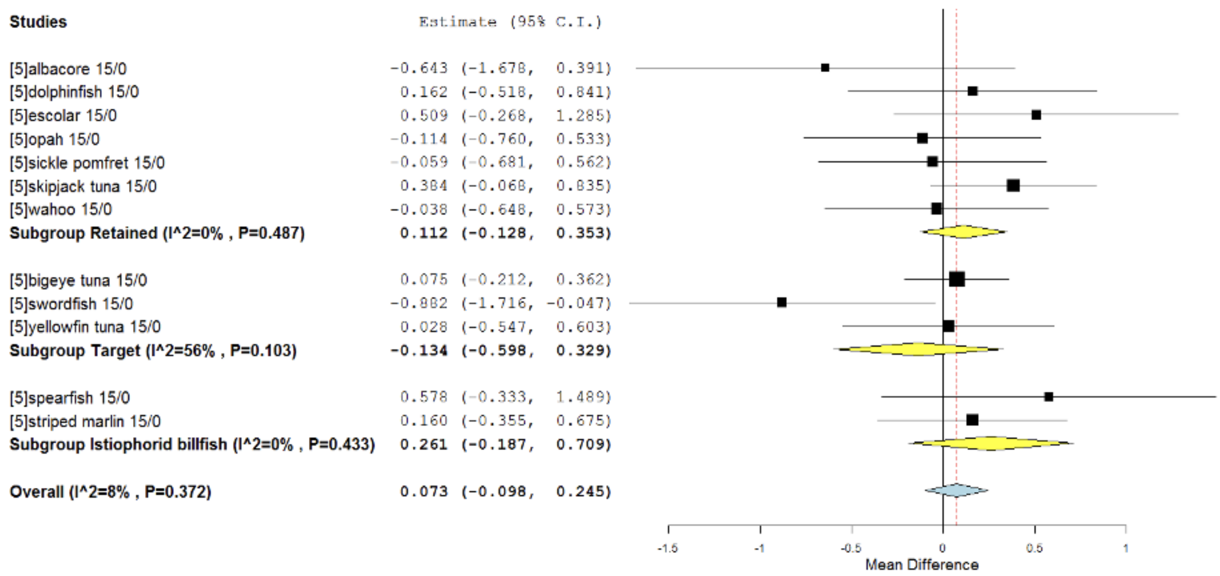


Figure 9

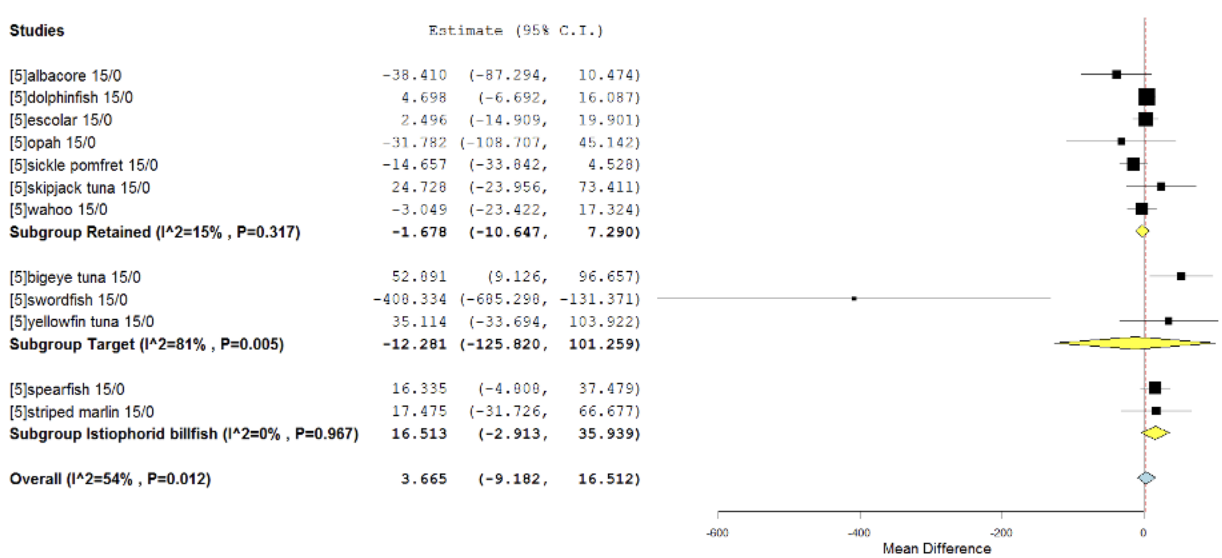


Figure 10

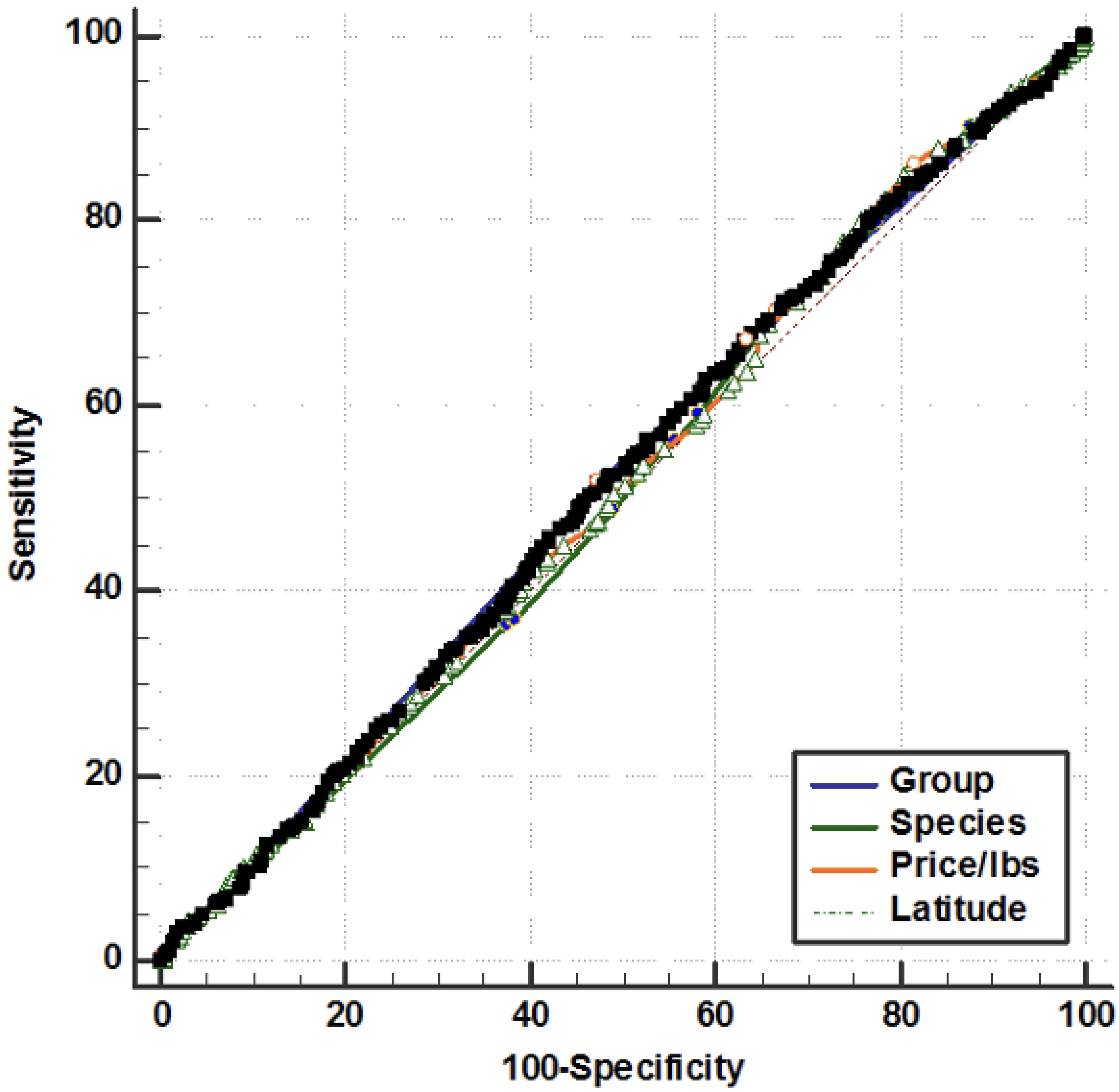

Sensitivity
100
80
60
40
20
0
0
20
40
60
80
100
100-Specificity
Group
Species
Price/lbs
Latitude

Figure 11

APPENDIX A- Supporting Online Materials

Supporting Resource S1: Experimental Design, Power and Cost: Benefit Analysis

The current project was requested to design a weak hook study of appropriate power ($\alpha = 0.10$ and $\beta = 0.20$) with the ability to detect a 0.10 difference (Fig. S1) between capture rates and a 0.05 difference in catch value (determined from body sizes) of target and bycatch species captured on strong (control) and weak hooks in Hawaii-based deep-set commercial longline fisheries. The proposed study deploying 170 sets will be used to compare previous results of Bigelow et al. (2012) that used a similar study design and will specifically test body weight differences between bigeye tuna (*Thunnus obsesus*) captured by the two strength hook types. Bigelow et al. (2012) measured body lengths and converted them to weights but did not conduct sampling when larger tuna are historically captured in the fishery. To rectify these deficiencies, it is necessary to sample larger fish to test that actual body weights (from fish at auction) do not significantly deviate between bigeye tuna captured from the weak and control hook types.

Power

The power of a statistical test is the probability the test will reject the null hypothesis (sometimes called hypothesis of no difference) when the null hypothesis is actually false (probability of not committing a Type II or β error – i.e. accepting a false null hypothesis)(Machin et al. 2009; Ryan 2013). Power is equal to $1 - \beta$ and can be used to calculate the minimum sample sizes needed to detect an effect size (i.e. differences in catch rates and body sizes of target and bycatch species captured on strong (control) and weak hooks) and is predominantly influenced by three parameters: (1) the α level or statistical significance criterion used (in our case, α will be equal to $p = 0.10$ as requested by NOAA Fisheries, (2) effect size, and (3) sample size. Statistical power in studies can be boosted by finding larger effect sizes, increasing sample sizes, relaxing α and also by reducing measurement error (in our case, we assume this will be zero or close to zero). Power can also be increased by using parametric tests instead of non-parametric tests and adopting directional tests (one-tailed) over non-directional tests (two-tailed) but these choices depend on the type of data collected and study design (Ryan 2013). Since a certain percentage of false positives (Type I errors) can be expected in null hypothesis significance testing by random sampling, researchers suggested the ability to replicate findings was the best and most powerful way to authenticate results (e.g., Carver 1978, 1993; Sutton et al. 2000; Ellis 2010; Musyl et al. 2015). Meta-analysis can be used for this purpose (Borenstein et al. 2009).

Meta-analysis

In order to conduct meaningful and realistic power analysis, it is necessary to derive precise estimates of effect sizes which are required as realistic inputs to maximize cost:benefit. In Bigelow et al. (2012), 929 bigeye tuna were captured on control hooks (49.2%) and 948 on weak hooks (50.2%)(11 fish caught on unknown hook type). These catch rates translated into a mean CPUE (x1000 hooks) of 6.1 (±5.02 SD) for strong and 6.2 (±5.39 SD) for weak hooks which indicates effect size was most likely near ~0.03 and that the original study was underpowered (Figs. S1-S3). The sensitivity analysis in Fig. S3 indicates how different effect sizes translate into power and cost:benefit issues. Testing a small effect size (0.05) would be cost-prohibitive (Fig. S3), let alone testing at 0.03. Bigelow et al (2012) could show no significant differences in bigeye tuna catch rates and weights between strong and weak hooks. Bayse and Kerstetter (2010) similarly showed no differences in bigeye tuna catch rates on strong and weak hooks in the Atlantic in a smaller study with 30 longline sets. Collectively, these results indicate the studies were underpowered (i.e., the inability to detect real differences). The conservative estimate of 170 sets can serve as a buffer from unexpected increases or decreases in CPUE. In order to maximise power in the study, however, we strongly advise that doubling the size of bigeye tuna samples (at no cost) could be achieved by including data from Bigelow et al. (2012) in a meta-analysis. Combining Bigelow et al.'s 127 sets into the current study of 170 sets will boost power in the bigeye tuna analysis to nearly 0.90 (Fig. S2). If catch rates are consistent across species and hook types, then the meta-analysis provides a combined estimate that is more precise than any of the individual studies. Estimates of sample sizes and sensitivity analyses were conducted using G*Power 3.1.9.7 (Faul et al. 2009).

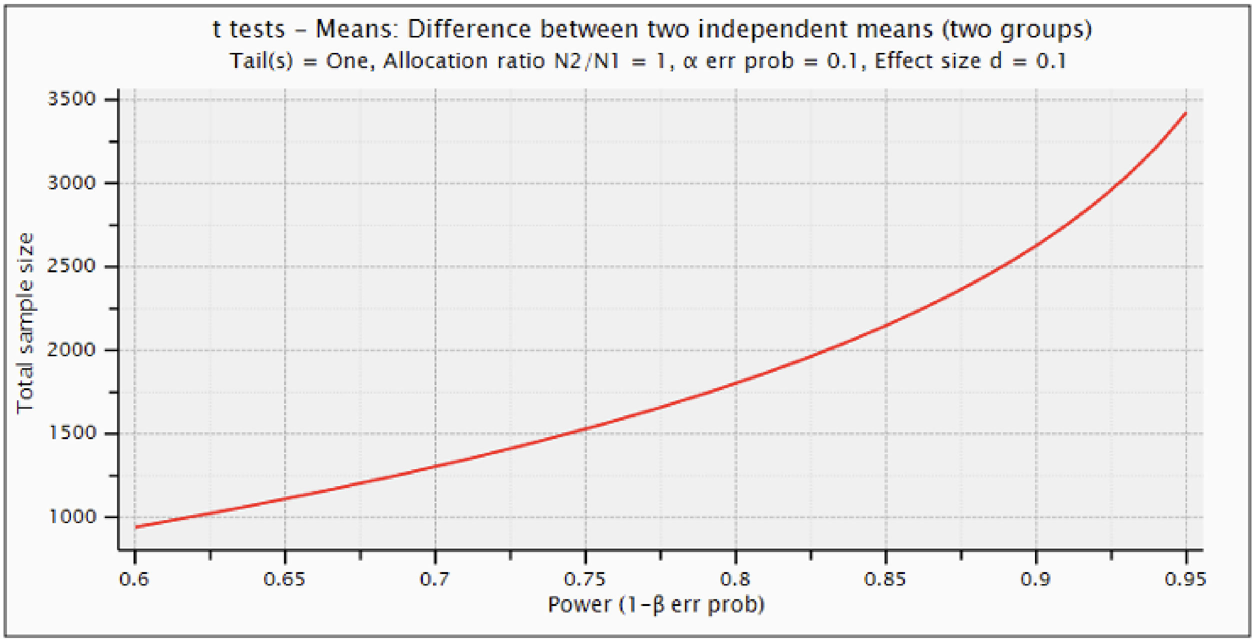


Figure S1. Directional t test with α = 0.10 and β = 0.20 with an **effect size of 0.10.**

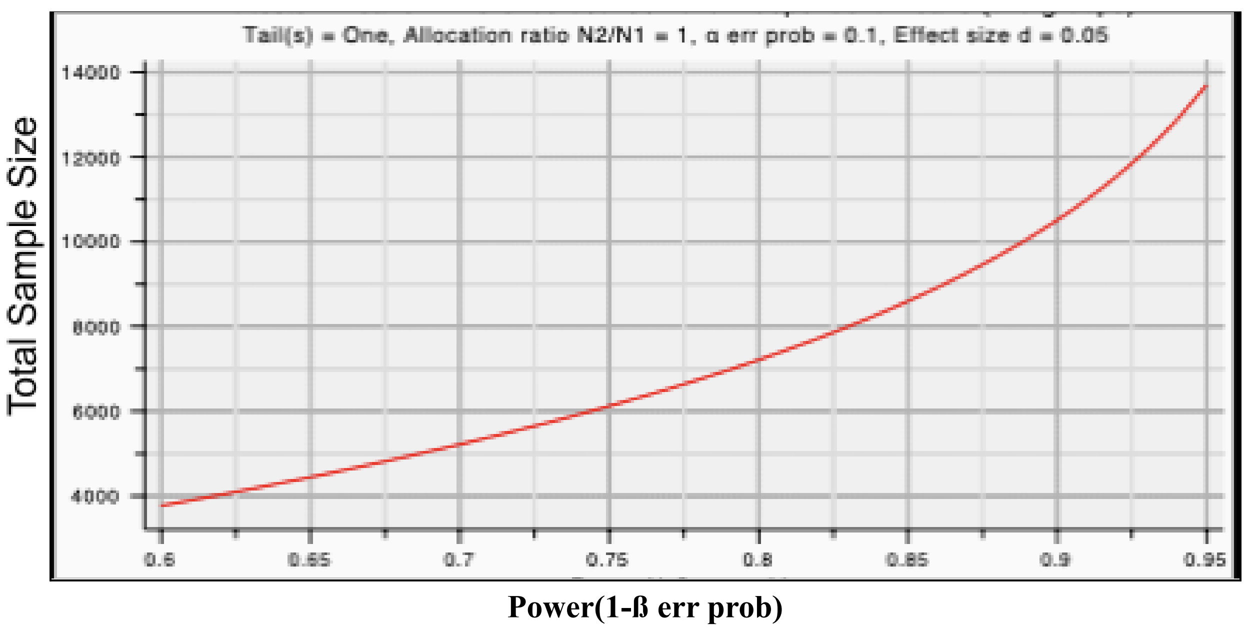
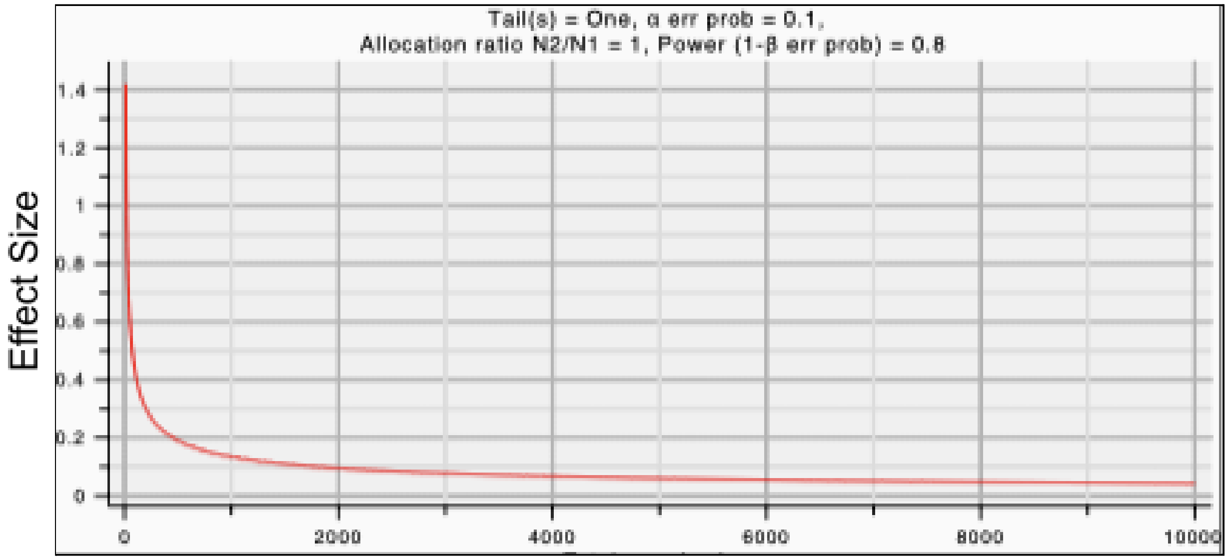


Figure S2. Directional t test with α = 0.10 and β = 0.20 with an **effect size of 0.05**



Total Sample Size

Figure S3. Sensitivity analysis with different **effect sizes** (from Fig. S2).

References

Bayse, S.M., and Kerstetter, D.W. (2010) Assessing bycatch reduction potential of variable strength hooks for pilot whales in a western North Atlantic Pelagic Longline Fishery. Journal North Carolina Academy of Science 126, 6–14.

Bigelow, K.A., Kerstetter, D.W., Dancho, M.G., and Marchetti, J.A. (2012) Catch rates with variable strength circle hooks in the Hawaii-based tuna longline fishery. Bulletin of Marine Science 88, 425-447.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. (2009) Introduction to meta-analysis. John Wiley & Sons, UK.

Carver, R.P. (1978) The case against statistical significance testing. Harvard Educational Review 48, 378-399.

Carver, R.P. (1993) The case against statistical significance testing, revisited. Journal Experimental Education 61, 287-292.

Ellis, P.D. (2010) The essential guide to effect sizes. Cambridge University Press, Cambridge.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41, 1149-1160.

Machin, D., Campbell, M.J., Tan, S.B. and Tan, S.H. (2009) Sample size tables for clinical studies. 3rd ed. Wiley-Blackwell, Chichester, UK.

Musyl, M.K., Moyes, C.D., Brill, R.W., et al. (2015) Postrelease mortality in istiophorid billfish. Canadian Journal of Fisheries and Aquatic Sciences 72, 1–19.

Ryan, T. P. (2013) Sample size determination and power. John Wiley & Sons, Hoboken, New Jersey.

Supporting Resource S2: Tensile strength of hooks and testing protocols *(this section was prepared by personnel from Lynker Technologies and NOAA Fisheries)*

To choose control and experimental hooks for the upcoming field trials; in consultation with personnel from NOAA Fisheries; several candidate hooks were tested for their tensile or straightening strength (below). Gape and "percent open" were measured at 100 lbs, 200 lbs, 300 lbs, 400 lbs, 500 lbs force, and finally at the release point for each hook type where plastic deformation was greatest. The difference between initial "gape" and "percent open" (Fig. S4) at each force level were ultimately used in determining that the 15/0, 4.5 mm ø round circle hook would be used as the strong (control) hook and the 15/0, 4.2 mm ø round circle hook would be used as the weak experimental hook for the study. These hooks exhibited consistent release rates and strength characteristics that were deemed best suited for the study for not being excessively strong or weak (Figs. S5 & S6). Hooks were kindly provided by Neil Kanemoto at POP Fishing and Marine, Honolulu, HI and source material for the hooks was 304 stainless steel hooks manufactured in South Korea by a subsidiary of POP Fishing and Marine.

Candidate hook types tested

14/0 round 4.5 mm diameter offset circle hook, 14/0 flat 4.5 mm diameter offset circle hook,
14/0 round 4.2 mm diameter offset circle hook, 14/0 flat 4.2 mm diameter offset circle hook,
15/0 round 4.5 mm diameter offset circle hook, 15/0 flat 4.5 mm diameter offset circle hook,
15/0 round 4.2 mm diameter offset circle hook, 15/0 flat 4.2 mm diameter offset circle hook,
16/0 round 4.5 mm diameter offset circle hook, 16/0 flat 4.5 mm diameter offset circle hook,
16/0 round 4.2 mm diameter offset circle hook, 16/0 flat 4.2 mm diameter offset circle hook.

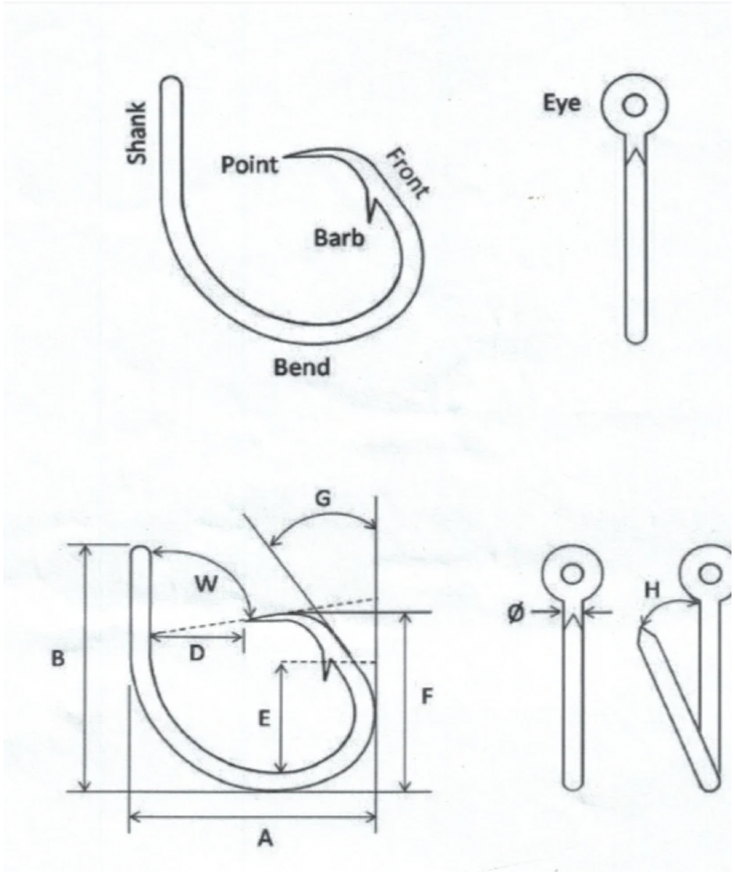


Figure S4. Basic component (upper panel) and measurements (lower) panel where: (A) width, (B) length, (D) gape, (E) throat, (F) front length, (W) point angle, (G) front angle, ø is wire diameter (Lettering conforms to manufacturer conventions).

Protocol for the hook strength pull test

1. In consultation with NOAA staff, various hook parameters (i.e., wire diameter at top of shank, width, length, gape, throat and front length for each hook) were measured (Figure S1).
2. Photographs of each hook prior to testing were used to validate size and type.
3. The hook was inserted into the test machine to perform the test. The test machine was designed by Lingren-Pitman for POP Fishing and Marine using an electric motor to generate force, measured by a Cardinal Model 190 scale.
4. Hook gape was measured approximately every 50-100 lbs of increasing force using a Performance Tool digital caliper (accurate to 1/100th of a millimeter).
5. Photographs of "straightened" hooks were taken and archived.
6. Data were compiled and plotted by pull force in lb. versus percent gape open (Figure S5) and force in lb versus gape measured in millimeters (Figure S6). Percent open calculations
7. Were done to account for size differences in gape widths among the different sized hook types and was calculated as follows:
8. $\Delta G \div G(i) \times 100 = \%\text{ open, where } \Delta G = Gape\ at\ force - G(i)\ or\ Initial\ Gape$
9. Data and photographs were provided to NOAA Fisheries for further analysis.

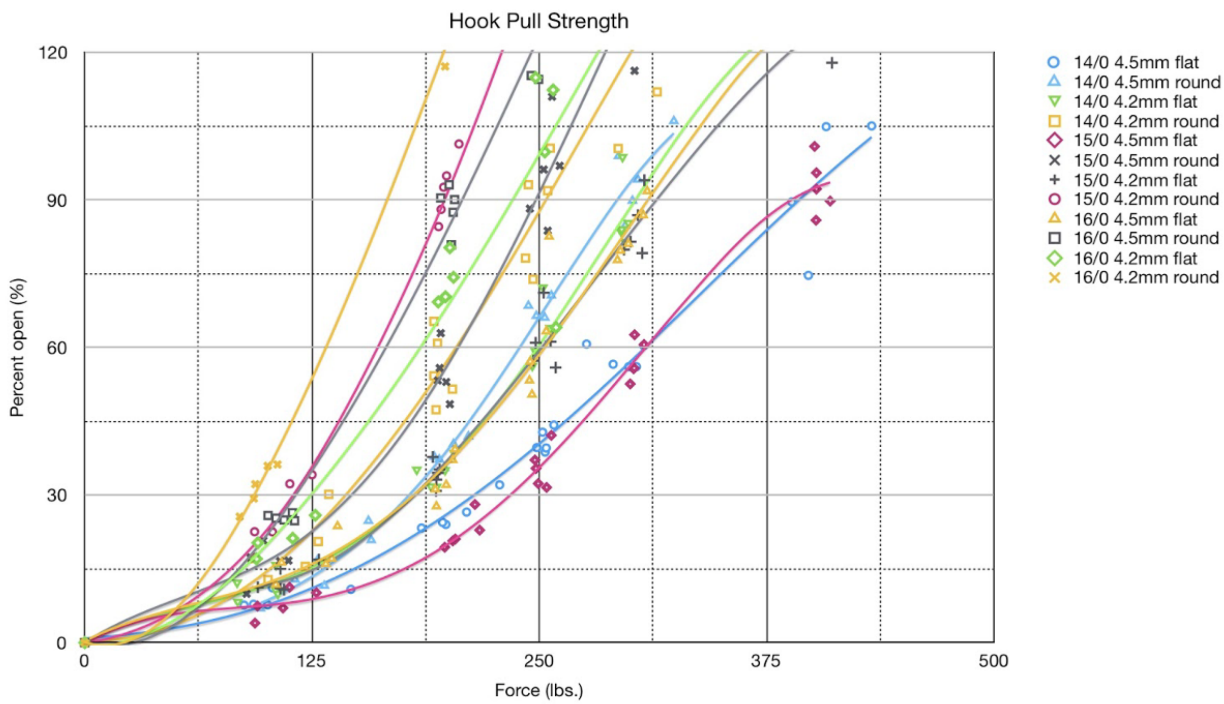


Figure S6. Graph of relative pull or straightening strength tests of the candidate hooks. Data are plotted as force applied to hook in pounds as a function of gape opening measured in millimeters. Best fit lines were determined using a 4th order polynomial.

Brief Results and Discussion of the Strength Test

Though differences in pull strength of hook types and diameters were observed, there was also variability within and between hook types. Preliminary analysis indicated a general trend with 14/0 hooks being the strongest, followed by 15/0 hooks, and the 16/0 hooks being the weakest (Figures S2 & S3). The testing device did not allow for fine-scale force application but several data points were attempted at force stops around 100 lb, 200 lb, 250 lb, 300 lb and every additional 100 lb until the hook came free or broke. Forged or flat shanked hooks also appeared to be stronger than their similarly sized round shank counterparts and wider diameter (4.5 mm) hooks were stronger than 4.2 mm hooks (Figures S2 & S3). At ~90% open gape, there was ~250lb difference between 16/0 4.2 mm round hooks and 14/0 and 15/0 4.5mm flat hooks. The difference between 15/0 flat (4.5 mm) to 15/0 round (4.2 mm) hooks at 90% was ~200 lb. Lastly, in a phone and email survey of the 3 major hook suppliers to the Hawaii-based longline fleet, respondents indicated they primarily provided 15/0 4.5 mm round or flat shanked hooks to the fleet, with a much smaller proportion of 16/0 hooks (Neil Kanemoto, POP Marine; Kim Lu and Dong Dang, fishing longline boat captains & owners, *personal communication*).

Supporting Resource S3: Straightened Hooks

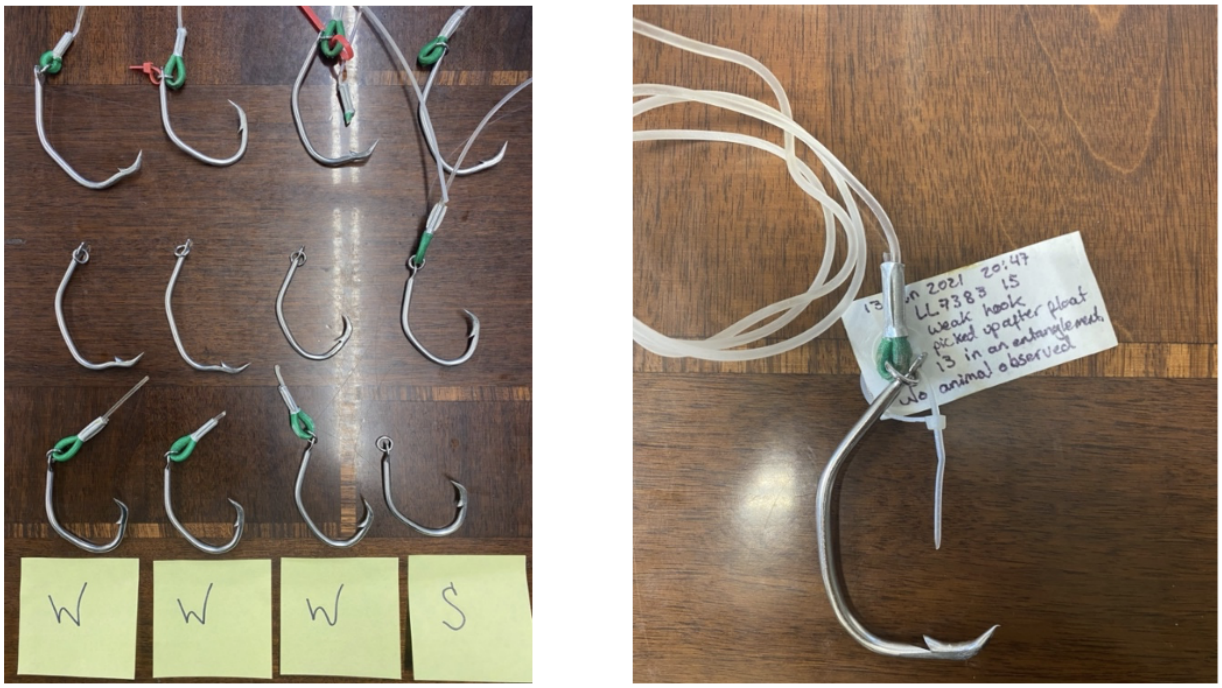


Figure S7. Example of straightened control (strong = S) and experimental (weak = W) circle hooks from the study as identified by vessel captains. A total of 17 bent hooks were collected by observers. Captains of vessels collected 45 and estimated that an average of 12 hooks were bent per trip.

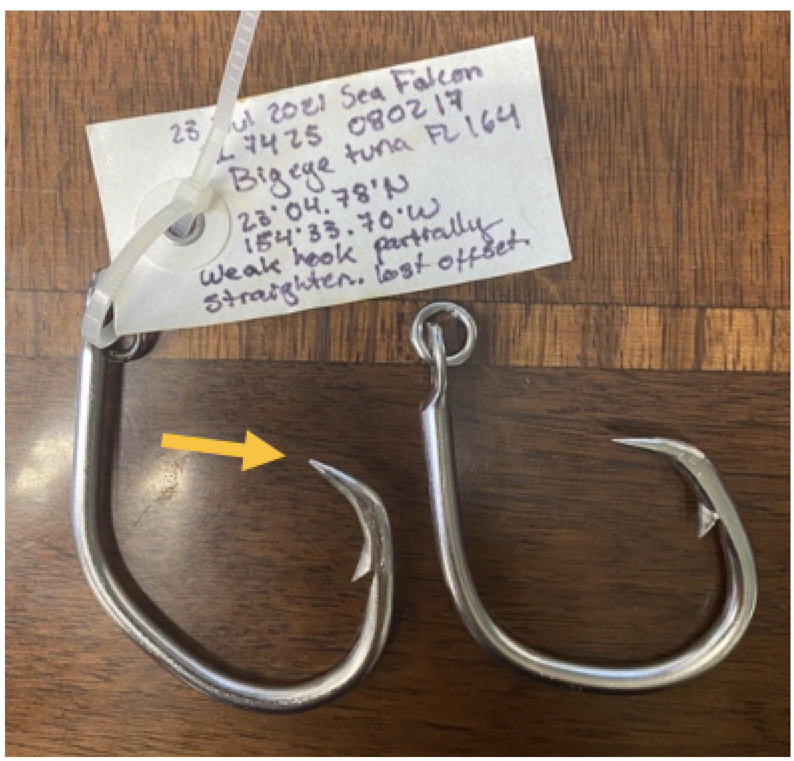


Figure S8. Notice that the barb on the weak hook to the left is partially open (see arrow) compared to the control (strong) hook on the right. Over the duration of the study, only one partially straightened weak hook was recorded with retained catch.

Supporting Resource S4: Complimentary study results

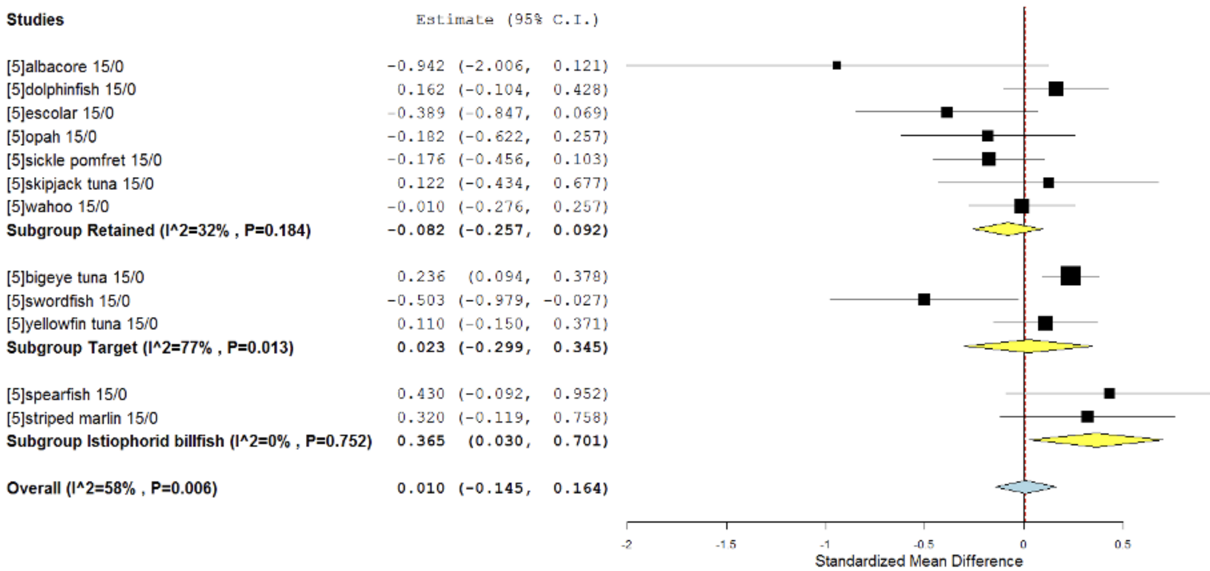


Figure S9. Forest plot for the effect size of weight at auction (see Table 3 in main text) using the standardized mean difference (*SMD*). Effect sizes and 95% CIs are provided for each study. I^2, the amount of variability among studies within species, along with a p value testing for heterogeneity (Cochran's Q), are provided. The area of the boxes for each study is proportional to the inverse of the variance, and any side of the box is proportional to the inverse of the standard error. The 95% confidence intervals (horizontal bars) for each study are proportional to the standard error and are related to sample size. The diamonds represent the summary effect size and the width is proportional to the 95% CI. Note especially the much narrower widths of the diamonds indicating more precision in the estimates.

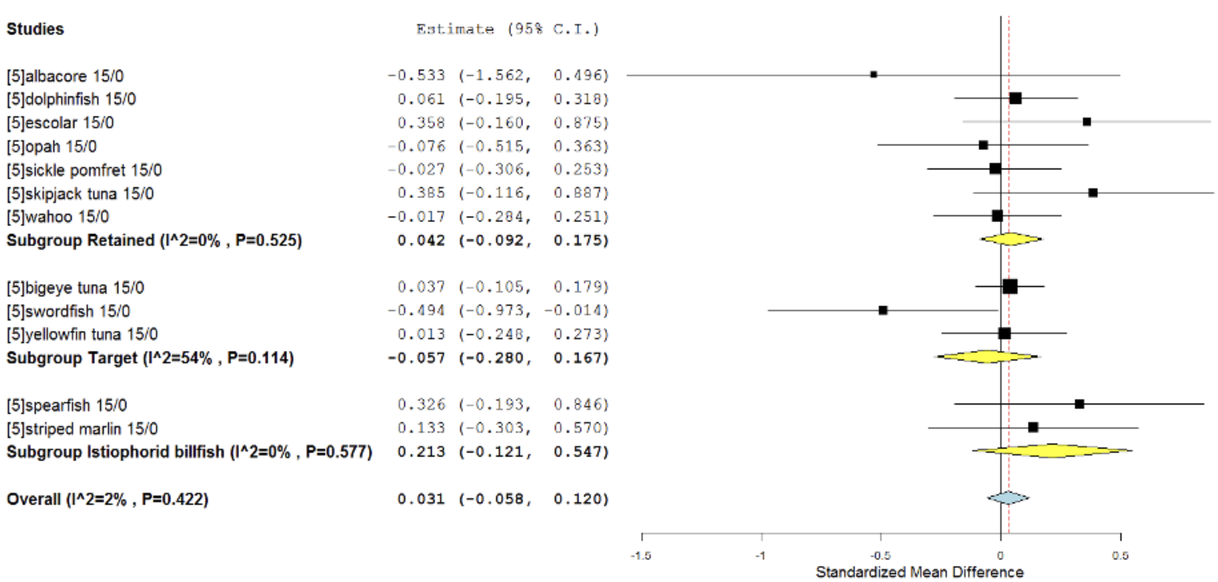


Figure S10. Forest plot for the effect size of price per pound at auction (see Table 4 in main text) using the standardized mean difference (*SMD*). Effect sizes and 95% CIs are provided for each study. I^2, the amount of variability among studies within species, along with a *p* value testing for heterogeneity (Cochran's Q), are provided. The area of the boxes for each study is proportional to the inverse of the variance, and any side of the box is proportional to the inverse of the standard error. The 95% confidence intervals (horizontal bars) for each study are proportional to the standard error and are related to sample size. The diamonds represent the summary effect size and the width is proportional to the 95% CI. Note especially the much narrower widths of the diamonds indicating more precision in the estimates.

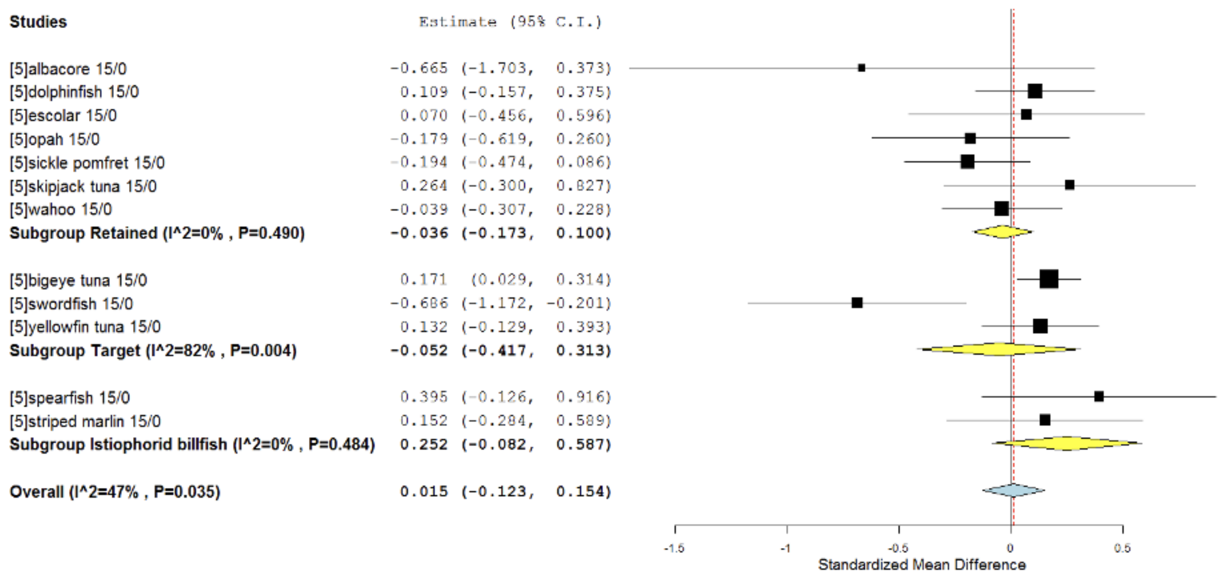


Figure S11. Forest plot for the effect size of average sales per fish at auction (see Tables 4, 5 in main text) using the standardized mean difference (*SMD*). Effect sizes and 95% CIs are provided for each study. I^2, the amount of variability among studies within species, along with a *p* value testing for heterogeneity (Cochran's Q), are provided. The area of the boxes for each study is proportional to the inverse of the variance, and any side of the box is proportional to the inverse of the standard error. The 95% confidence intervals (horizontal bars) for each study are proportional to the standard error and are related to sample size. The diamonds represent the summary effect size and the width is proportional to the 95% CI. Note especially the much narrower widths of the diamonds indicating more precision in the estimates.

Supporting Resource S5: Assuring Consistency of Manufactured Experimental Hooks

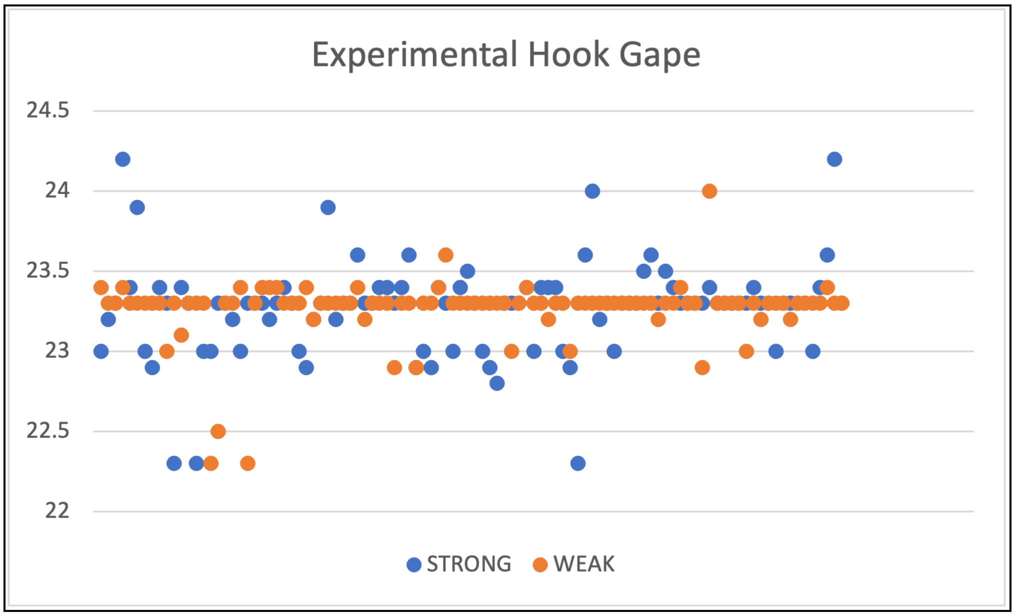


Figure S12. Gape was measured from the point of the hook to the top of the shank on 200 (100 Strong + 100 Weak) hooks to test consistency amongst the manufactured experimental hooks. Both Strong and Weak hooks had an average gape of 23.3mm

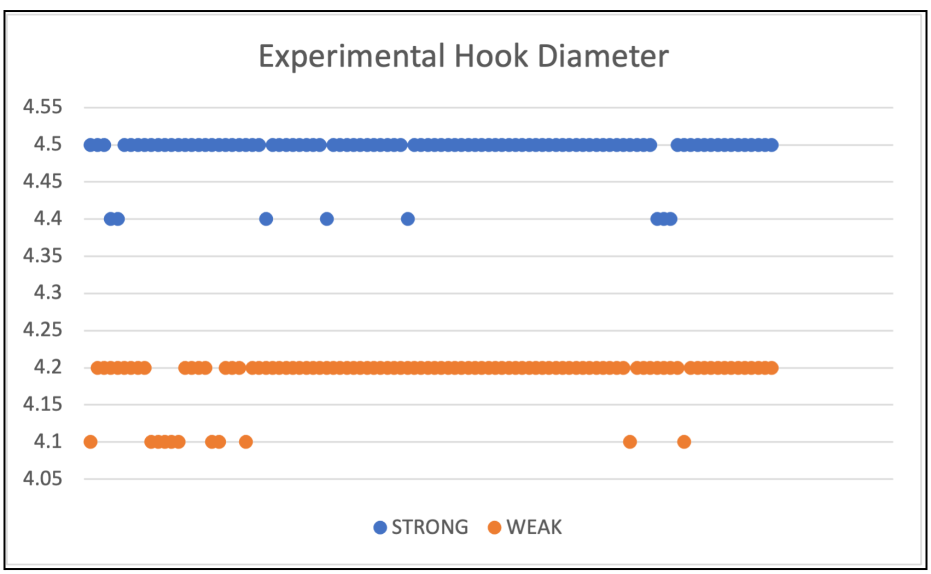


Figure S13. Scatter Plot displaying the consistency of the manufactured hooks. The diameter of 200 (100 Strong + 100 Weak) hooks were measured on the upper portion of the shank. 94% of strong hooks had a diameter of 4.5 mm while weak hooks had a 91% of being 4.3mm.