ISAB Response Review of Revised CSS 2018 Annual Report: *Chapter 8. PIT tag and coded-wire tag effects on smolt-to-adult rates for Carson National Fish Hatchery spring Chinook salmon*

(ISAB 2018-4 Update; February 6, 2019)

Background

This addendum is a follow-up review to the ISAB's Review of the Comparative Survival Study (CSS) Draft 2018 Annual Report (ISAB 2018-4; October 18, 2018). In this addendum, the ISAB evaluates the CSS's responses to our comments regarding *Chapter 8. PIT tag and coded-wire tag effects on smolt-to-adult rates for Carson National Fish Hatchery spring Chinook salmon.* The CSS's responses were incorporated in two documents: CSS Responses to comments on the 2018 CSS Annual Report (pages 33-36) and CSS Revised 2018 Report, December 2018 (pages 184-211).

Overall ISAB Response

The revised chapter addresses many of the concerns raised in our review of the preliminary version. The new length data and revised Discussion relative to Knudsen et al. (2009) are vast improvements. Comparing data analyses between studies strengthens the conclusion of no PIT tag effects in the CSS study. Also, the data and discussion of the importance of having an experienced tagging crew is very informative and important information for others.

There are a few minor points that should be considered in any future revision or publication.

The description of how the power analysis was performed was helpful. Both the expected precision (Figures 8.2 and 8.4) and the power (Figure 8.2) assume that no-overdispersion will be present and so may be optimistic. For example, later in the report, the overdispersion factor was found to be around 2, implying that the effective sample size is only about ½ of the actual sample size which will increase standard errors and reduce power. This should be noted in the discussion about power. As well, the model comparison table (Table 8.4) should be adjusted for overdispersion using QAICc. [It is unlikely to materially affect the relative model weights.]

Rather than relying on understanding the coding used to represent the three treatment groups (p. 196), a better method to compare the SARs among the groups would use the *eemeans* package in *R*. This provides estimates of effect sizes on the logit and odds scales without having to know the coding used. It can automatically adjust for overdispersion and can also provide all three pairwise comparisons of the odds ratios (with confidence bounds). [Estimates should match what are presented in the report, but this revised methodology for multiple comparisons is much easier to use.]

Given the three tag groups (CW, PIT, Dual), it is also possible to examine whether the effects of PIT tagging and CW are additive; i.e., is the effect of double handling more/less than the sum of the individual handling/tag effects? [An interaction cannot be fit because there is no group with "no tags".] This is easily done in the current modelling setup.

New Editorial Comments

Figure 8.3. Include the alpha level (0.10) in the legend.

p.192. Last sentence. Reference to Figure 8.2 should be to Figure 8.6.

Table 8.3. It is confusing to have Brood year, N, and CWT code repeated for each brood year for the dual tagged group. Perhaps replace the repeated values by one line in the "middle" of the two recovery rows so that the reader does not inadvertently double count.

ISAB Specific Replies to CSS Responses to ISAB's 2018 Comments on Chapter 8

ISAB 2018: This is a very important study, which appears to be well conceived and executed, but, unfortunately, the reduced returns (compared to projections) reduced the utility of the study.

> **CSS Response:** We disagree that the study has reduced utility. This judgement appears to be based on a misunderstanding of the study design and the return rates that were assumed. The return rates that were realized were well within expectations and we have provided additional details on the return rate assumptions, along with the expected precision and power associated with the tagging levels and study duration.
>
> > **ISAB 2019 Reply:** See the Overall Response discussion above. The revised chapter addresses many of the concerns raised in our review of the preliminary version, but we raise a few minor points that should be considered in future revisions or publications.

ISAB 2018: The original power analysis performed by the authors assumed that SAR values would average 0.37% (0.0037). Instead, for the years they examined, SARs were one-half to one-third of expected (~.0015. or so). This reduces the power to detect effects. While their original power analysis was designed to detect a minimal (relative) difference of 25% in SARs with a power of near 80%, even a 10% effect would be concerning, but an effect of this size will now be very difficult to detect. Indeed, the final estimates of the effect of tag type (Figure 8.9) have very wide confidence bounds on the effect size.

CSS Response: The average SAR of 0.37% represented the total returns to the hatchery and fisheries across brood years 2000-2004, and this value was not used in the power analyses. We used data on returns to the hatchery for the previous 10 brood years to conduct the power analysis. The SARs observed in this study were well within the range of the SARs that were used for the power analysis. We have provided a detailed description of the power analyses in the revised chapter. The power analyses indicate that a 25% difference in SARs would be expected to be detected 94% of the time and a 10% effect is expected to be detected about half of the time. The wide confidence bounds are partially due to the fact that the PIT tag-only fish returned at a higher rate than CW tag-only fish in two years and at a lower rate in two years.

ISAB 2019 Reply: The clarification between the two different endpoints for the SARs (combined returns to the hatchery and captured in the fishery vs. returns to the hatchery only) was helpful. The description of how the power analysis was performed was also helpful. Both the expected precision (Figures 8.2 and 8.4) and the power (Figure 8.2) assume that no-overdispersion will be present and so may be optimistic.

ISAB 2018: The Discussion needs improvement. The Introduction states: "Knudsen et al. (2009) found that smolt-to-adult recruit survivals (SARs) were on average underestimated by 25% for PIT tagged fish due to tag loss and reduced survival of tagged individuals. Further, they found that after correcting for tag loss, mortality caused by tagging was on average 10.3% and as high as 33.3%." It seems that you would want to address the discrepancy in your results as compared to Knudsen et al. (2009). Further, Knudsen et al. (2009) reported differences in growth of PIT- tagged vs non-tagged fish. Does Carson NFH collect lengths and weights at spawning that would allow the CSS to address that question?

CSS Response: The discussion section has been completely re-written in the revised chapter. We have provided an extensive discussion of the differences in results and methods between this study and that of Knudsen et al. (2009). We have also included an analysis of length-at- return for the tag groups in the study, which found no detrimental effect of PIT tags on length compared to CW-tagged fish.

ISAB 2019 Reply: The new length data and revised Discussion relative to Knudsen et al. (2009) are vast improvements. Comparing data analyses between studies strengthens the conclusion of no PIT tag effects. Also, the data and discussion of the importance of having an experienced tagging crew is very informative and important information for others.

ISAB 2018: Rather than simply reporting that no evidence of an effect was found, please *also present point estimates and confidence bounds. For example, Figure 8.9 shows estimates of effect sizes with confidence bounds, but a reader would be hard pressed to interpret the results. All the confidence bounds in Figure 8.9 seem to run from about -1 to 1. Presumably these are on the logit-scale and so represent bounds on the odds-ratio (?) and so the confidence bounds on the effect size are somewhere between exp(-1)=0.36x to exp(1)=2.7x! These are very wide, but many readers may not appreciate just how wide.*

CSS Response: We have provided estimates of the odds ratios and their confidence bounds in the revised chapter.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: The major problem with studies of the kind in this chapter is that very large sample sizes are required and consequently, no one study is likely to give definitive results. It would be an opportune time to do a meta-analysis of the many different studies that have looked at this issue to help guide managers on this important topic.

CSS Response: Unfortunately, there have not been many studies that have looked at this issue on smolt-to-adult return rates. We have provided a review of the only three studies that we could find addressing PIT versus CW tags on SARs (Prentice et al. 1994, Knudsen et al. 2009, Cassinelli et al. 2012).

ISAB 2019 Reply: This was helpful.

ISAB 2018: The chapter fits several Cormack-Jolly-Seber models to selected parts of the data. Is there any advantage to creating a single omnibus model?

CSS Response: We have condensed the number of CJS models down to two. One covers release to adult return to the hatchery and the other covers the time from detection at the Carson NFH adult ladder though spawning. The reason that we have two models is that the data for the second analysis (adult ladder detection through spawning) was aggregated across years due to the smaller sample sizes available.

ISAB 2019 Reply: This response is adequate and well done.

ISAB 2018: The ISAB has the following suggestions and questions to improve the statistics and modeling in the chapter:

p. 176. Figure 8.3 seems to indicate that fish of age 3, 4, and 5 are all pooled to estimate a single "survival" probability. But these three age groups will have quite different survival probabilities from Bonneville (as juveniles) to Bonneville (as *adults). This heterogeneity may affect subsequent estimates of survival. Fortunately, the vast majority of fish are age 4 so any bias is likely to be small. This should be mentioned in the report, and an estimate of potential bias produced.*

CSS Response: Similar to Knudsen et al. (2009), we have calculated SARs based on the total number of age-3, age-4, and age-5 adult returns. Most SARs are calculated this way, although for some analyses SARs are calculated with and without age-3 (jack) returns included. We have stated in the report that SARs include all age-classes. It is correct that almost all of the adults returned at age-4, but we do not agree that summing the age-class returns when calculating overall SARs creates any sort of bias.

ISAB 2019 Reply: This response is adequate.

p. 177. The estimate of tag loss needs a model to justify that these are the estimates of the relevant parameters.

CSS Response: The estimates of tag loss are simply a quotient of the number of fish that were determined to have lost their tag divided by the number of fish effectively sampled for tag loss. This is the model that we used and is identical to Knudsen et al. (2009).

ISAB 2019 Reply: This response is adequate.

p. 178. Was there evidence of over-dispersion? If so, how was this incorporated into the results?

CSS Response: Yes, there was evidence for overdispersion and we used a quasibinomial model to account for the overdispersion and adjust the standard errors.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 179. Doesn't the proportion of release groups in each rearing vessel type add to 1? This would make these three X-variables co-linear with the intercept. How was this handled?

CSS Response: We have removed the comparisons of SARs among release groups, so this issue is no longer relevant.

ISAB 2019 Reply: This response is adequate.

ISAB 2018 Editorial comments

ISAB 2018: p. 171. It states: "… with a power of near 80%." Please give the effect size here rather than making the reader refer to past reports.

CSS Response: We have provided a plot of the power curve for effect sizes ranging from 0% to 30%.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 174, 1st full paragraph. It is not clear when this holding period took place. Was it at the time of smolt release or at the time of tagging?

CSS Response: The short-term holding period was 30-days after tagging and we have provided a better description of this in the revised chapter.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 174. It states: "… these fish were removed from the PTES database ..." It is unclear why these fish had to be removed from the database? Of what is PTES an abbreviation?

CSS Response: PTES was an acronym used by the U.S. Fish and Wildlife Service for this study (the PIT Tag Effects Study). We have removed references to this acronym and clarified that additional data from the study is stored in a database by the U.S. Fish and Wildlife Service at their Columbia River Fish and Wildlife Conservation Office in Vancouver, WA.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 180 (and elsewhere). It states: "… do not indicate significant differences between the two PIT tagged groups." This should be reworded as "indicates that there was no evidence of a difference in survival between the two PIT tagged groups." Similar changes are needed in other conclusions in the chapter.

CSS Response: We have adopted this type of wording throughout the revised chapter.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 184, Table 8.4. How can the number of surplus fish be negative in 2014?

CSS Response: This table represents the difference in timing between the mean date of hatchery entry and the date of surplus, inoculation, and the three spawning periods. The negative number indicates that surplus activity occurred three days before the mean entry date for all fish that returned.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 184, 1st paragraph. How can the total percent returns by age = 104% (7% returned at age-3, 90% returned at age-4, and 7% returned at age-5)? The total number of returning dual tagged fish in Table 8.5 is 77, not 76.

CSS Response: We have removed the percent returns by age and all of the age-at-return data are presented in the table. The number of dual-tagged fish that returned is 77.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 186, Figure 8.9. The X-axis needs units (logits?).

CSS Response: We have removed this figure in the revision and report the odds ratio and its confidence interval in the text.

ISAB 2019 Reply: This response is adequate.

ISAB 2018: p. 187, Figure 8.10. Please place the brood-year of the group of fish next to the corresponding point.

CSS Response: The SARs for each group are provided in a Table 8.3 and provide a way to cross- reference each point. We have also added confidence intervals to the points in the plots.

ISAB 2019 Reply: This response is adequate.