



*Independent Scientific Advisory Board
for the Northwest Power and Conservation Council,
Columbia River Basin Indian Tribes,
and National Marine Fisheries Service
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204
ISAB@nwcouncil.org*

ISAB Review of the 2005 Comparative Survival Studies' Annual Report and Applicability of Comparative Survival Studies' Analysis Results

Robert Bilby
Susan Hanna
Nancy Huntly, Chair
Stuart Hurlbert
Roland Lamberson
Colin Levings
William Percy
Thomas Poe
Peter Smouse

Charles Coutant, Ad Hoc Member
Richard Alldredge, Ad Hoc Member, ISRP

ISAB 2006-3
March 15, 2006

ISAB Review: The 2005 CSS Annual Report and Applicability of CSS Analysis Results

Executive Summary.....	1
I. Introduction and Background.....	5
Review Assignment	5
Previous Reviews.....	6
II. Review of 2005 CSS Annual Report.....	7
Methods (Chapter 2).....	7
Results (Chapter 3)	9
Adult Drop-out Rates (Chapter 4).....	11
Hatchery-to-Hatchery SARs for Various Hatcheries (Chapter 5)	11
Upriver-Downriver Comparisons (Chapter 6).....	12
Simulated PIT-tag data to test CJS survival estimates (Chapter 7)	12
ISAB Evaluation of Comments by BPA and NOAA Fisheries.....	13
<i>BPA Comments</i>	13
<i>NOAA Fisheries Comments</i>	14
III. ISAB Answers to Council’s Questions	15
IV. ISAB Conclusions and Recommendations.....	17
References.....	19
Appendix A: Previous Review Comments by ISAB and ISRP	20

ISAB Review: The 2005 CSS Annual Report and Applicability of CSS Analysis Results

Executive Summary

On December 20, 2005, the Council requested that the Independent Scientific Advisory Board (ISAB) review the 2005 Annual Report for the Comparative Survival Study (CSS) prepared by the Fish Passage Center (FPC) and the Comparative Survival Study Oversight Committee, as well as critical comments on the draft of that report by the Bonneville Power Administration (BPA) and NOAA Fisheries. The CSS is a field study, begun in 1996, that addresses important and technically complex issues regarding the survival of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead through the Columbia River hydrosystem from juveniles through returning adults. The study focuses on relative survival of fish that traveled downstream as juveniles by alternative routes (e.g., in river, transported, different routes of dam passage, and different numbers of dams passed). The results can have important implications for operation of the hydrosystem to ensure protection and propagation of anadromous salmonids. The Council expressed a desire to aid resolution of disputes over the study by obtaining the ISAB review.

The Council asked that the ISAB assess the overall integrity and scientific soundness of the CSS report and address the following specific questions:

- 1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?*
- 2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?*

The ISAB accepted the assignment on January 12, 2006 and received a briefing on the CSS Annual Report from the study's Principal Investigators on January 27th. The ISAB considers that there are two parts to this review: (1) review of the 2005 CSS Annual Report and (2) a determination of the utility of the CSS comparative survival estimates for various management and hydrosystem operational decisions.

The ISAB finds that the CSS is an ambitious, long-term study that is being criticized because its objectives are not yet fully met, despite prodigious efforts in both the field and in complex data analyses. The CSS has used the PIT-tag technology to mark and track individual salmon and steelhead through their smolt-to-adult life stages. Expectations of this mark-recapture technology exceed the results that are practically attainable, and its use is still evolving. The CSS study participants have been major players in this evolution. We find the present annual report to be a further incremental step in the direction of documenting different survival rates of different stocks under different migration conditions. That the present report is not a perfect reconstruction of

differential survival histories is largely a result of the current analytical capabilities and available sample sizes. The deficiencies seem to be highlighted in some aspects because of experimental design and analytical approaches taken by the authors. The ISRP comment from their 2002 review still applies that “the formulas [used to compute relative survival rates] are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view.”

Specific Responses to the Council’s Questions

1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?

All in all, the design, implementation, and interpretation of the *statistical analyses* underpinning the report are very good. Nonetheless, there are broader concerns over the design of the study such as sample size, sampling sites, time periods for analyses, and other features. Improvements can be made, and our recommendations follow.

Since the region is unwilling to conduct the manipulative experiments in the hydrosystem that the ISAB and ISRP have recommended for many years, the CSS is doing the next best thing. That is, the study is following as many fish through their life cycle as possible, calculating the survival, and comparing outcomes.

2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

The ISAB believes the Council should view the CSS as a good, long-term monitoring program, the results of which should be viewed with increasing confidence as years pass. Under scrutiny from periodic peer reviews and agency comments, the methods should improve and the results become ever more valuable. The project is definitely worthy of Council support.

The Council’s question is difficult to answer with the present annual progress report. The project needs a synthesis report that clearly describes the analytical methods and summarizes the project results in a holistic way for its decade of effort.

The ISAB recognizes a disconnect between the present status of results and much of the decision-making that takes place regarding hydrosystem operations and fish protection. Although the project is making good progress at addressing such issues as the value of transportation and the relative survival from different passage routes, many relationships between survival and specific operational alternatives or environmental features during migration cannot be resolved when data are aggregated simply by year of migration. For this information to be most useful for making management decisions, aggregations of

data within years and across years for different operational options and environmental constraints should be pursued. We encourage the project to move in that direction.

The results of the CSS appear to indicate that PIT-tagged fish do not have the same survival rate as untagged fish. This conclusion is not emphasized by the current progress report, but it has major implications for many uses of the PIT-tag technology. Comparisons among PIT-tagged groups of fish are probably appropriate, but extrapolations of the results from PIT-tagged fish to untagged populations should be made with caution.

Recommendations

- It has been ten years since the CSS was initiated. The report the ISAB reviewed was the latest in a series of annual progress reports, and thus lacking a holistic perspective. The ISAB recommends that the CSS produce a ten-year summary report providing an in-depth description of methods and detailed analyses and interpretation of the data in a retrospective style.
- The CSS needs to more effectively present the methodologies used in their analyses so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated in the ten-year summary recommended above.
- The ISAB agrees with critics who express concern that two downriver sites (Carson Hatchery and John Day River) are probably insufficient to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.
- Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.
- Aggregation of data solely by juvenile migration year should be supplemented with analyses that group data on environmental and operational factors that may be amenable to control.
- Assumptions inherent in the analyses should be specifically tested, with continued vigilance toward avoiding bias.
- Pre-assigning the intended routes of passage at the time of release into inriver and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the

study design is scheduled for implementation in 2007, but should begin in 2006, if feasible.

- Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (*D*). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.
- The CSS should be supplemented by funded research into analytical methods that can improve, and hopefully simplify, the mathematical and statistical approaches currently in use. It is not clear from available information whether the problem is that the formulas are unnecessarily complicated, inappropriately specified, or just not well explained (see bullet #2 above).
- More attention should be given by the CSS and the region as a whole to the apparent documentation that PIT-tagged fish do not survive as well as untagged fish. This point has major implications for all uses of PIT-tagged fish as surrogates for untagged fish.

I. Introduction and Background

Review Assignment

On December 20, 2005, the Council requested that the Independent Scientific Advisory Board (ISAB) review the 2005 Annual Report for the Comparative Survival Study (CSS) prepared by the Fish Passage Center (FPC) and the Comparative Survival Study Oversight Committee. The CSS is a field study of the survival of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead through the hydrosystem from juveniles through returning adults, with a focus on relative survival of fish that traveled as juveniles by alternative routes (e.g., in river, transported, different routes of dam passage, and different numbers of dams passed). The annual report reviews recent mark/recapture activities and bootstrap analysis for generating confidence intervals.

The CSS is important, as it is one of the few organized attempts to systematically release PIT-tagged, hatchery-reared fish, and wild smolts into the Columbia River for the purpose of monitoring and evaluation. Most aspects of the study, from its design and methods to the analytical results, have been strongly debated in the Region because the relative survival rates of salmonids under different hydrosystem operations and environmental constraints is at the heart of water and fish management policies.

In response to the release of the draft version of this annual progress report, both the Bonneville Power Administration and NOAA Fisheries provided the FPC with letters setting forth both broad concerns and detailed criticisms of the findings and results reported in the draft report. Before finalizing the report, the FPC provided detailed responses to both Bonneville and NOAA Fisheries addressing their concerns. The Council expressed its wish to contribute to the resolution of these important and technically complex issues by having the ISAB conduct its own review of the final progress report and the attendant letters. In conducting the review, the Council asked that the ISAB assess the overall integrity and scientific soundness of the CSS report and address the following specific questions.

- 1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?*
- 2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?*

The ISAB accepted this important assignment on January 12, 2006 and received a briefing on the CSS Annual Report from the study's Principal Investigators on January 27th. The ISAB considers that there are two parts to this review: (1) review of the 2005 CSS Annual Report and (2) a determination of the utility of the CSS comparative survival estimates for various management and hydrosystem operational decisions.

The CSS was initiated in 1996 by the Northwest fishery agencies and tribes as a long-term study to estimate survival rates over different life stages of spring and summer Chinook salmon produced in hatcheries in the Snake River basin and selected lower hatcheries in the lower Columbia River. The study has expanded somewhat to encompass wild Chinook salmon and steelhead, and the mix of hatcheries has changed with experience. The premise of the research was that, through use of PIT tags implanted in juveniles at the point of release from hatcheries or rearing facilities, the survival of unique groups of fish could be determined as they passed through PIT-tag detectors in juvenile bypasses at dams or in adult fish ladders on their return. From these survival rates it was hypothesized that one could quantify differential survival according to passage route. Of particular interest were differences in survival related to distance from the ocean, between transported and in-river fish and the delayed effects of hydrosystem passage (by juveniles) on adult returns.

Previous Reviews

Both the ISAB and the ISRP previously reviewed the CSS study proposals in 1998 (ISAB 1998) and 2002 (ISRP 2002) and the recommendations from those reviews were generally as follows (recommendations are provided in full in Appendix A):

In 1998, the ISAB supported funding of the study. They recommended including naturally reproducing populations as well as hatchery fish and suggested that other life-history types of Chinook salmon and steelhead be included. They recommended quantifying survival from tributary hatcheries to Lower Granite Dam and McNary Dam, and through the entire hydrosystem when sufficient detectors were functional. They encouraged attempts to compare survival of PIT-tagged fish to untagged fish or fish tagged by other methods. The ISAB also saw this as a way to coordinate the PIT-tagging efforts of many agencies and to provide an opportunity for periodic workshops to review results.

The ISRP reviewed the continuation proposal in 2002 and also recommended funding. The “best” formulas for calculating smolt-to-adult survival rates from then-available data were judged “complicated, convoluted, and in general, very unsatisfactory from a statistical point of view.” It was noted that arguments over these methods would likely continue and spawn even more detailed arguments and counter-arguments. Much of the difficulty lies in small sample sizes due to both numbers of fish tagged and the number of detections. Improved detection at Bonneville Dam was recommended. The ISRP recommended more research on mathematical and statistical methods both within this project and outside it for estimating life-cycle survival.

II. Review of 2005 CSS Annual Report

Methods (Chapter 2)

There are three principal issues over the study's methods. One concerns the selection of hatcheries (or other release sites), especially for comparisons between smolts with long passage routes through the hydrosystem and those migrating from lower in the basin with few dams to pass. Another relates to the mathematical and statistical methods employed in the analyses, including potential biases and the types of aggregation of data for summaries. A major point raised by NOAA Fisheries is the unreliability of the PIT-tag method to represent the survival of untagged fish (the CSS data indicate that PIT-tagged fish do not survive as well as untagged fish, and therefore are not adequate surrogates for untagged fish in the population).

Some study methods are not fully described in this annual progress report. We did not seek out previous annual progress reports to fill in the information gaps. This difficulty begs for a summary report that can provide a more complete description of methods.

It would be useful to have the SARs analyzed as a function of size at release. This could be tested for rather than just presenting size data. Also, data on size of all PIT-tagged fish from hatcheries and other release sites should be included in much greater detail than median lengths at tagging reported in Table 2 (e.g., include mean lengths, weights, and ranges). Sizes at release may be a significant factor in differential SARs from various sources. Fish size is generally not accorded much significance in the CSS studies despite a well-known survival advantage for larger fish. As raised in comments by NMFS, these size effects need to be given more consideration in further analyses. The ISAB recommends including a specific analyses focusing on the effects of size at release on SAR values of all PIT-tagged fish.

The numbers of fish available for tagging is a major constraint. As tables 2-5 demonstrate, the number of tagged fish vary considerably by location and year. The study participants have had to be opportunistic despite an intended experimental design. To their credit, they appear to have been quite successful in obtaining numerous stocks and years to compare.

Holdovers (fish not migrating fully through the hydrosystem in the year of initial outmigration; Connor et al. 2002) cause methodological problems. The authors have tried to account for these fish in different ways in this and the previous annual report. They believe the present method has less bias for estimating survival. This needs to be evaluated in later years.

We admire the study participants for attempting to segregate fish among their several migration-route histories. Although the term "destined" seems too strongly pre-ordained for the current methods of release and tracking, fish do have the three options listed: in-river by non-bypass routes, in-river through dam bypasses, or routed to transportation at the collector dams. They have these options at most dams (not all dams have facilities to

collect fish for transportation), thus expanding the number of possible migration histories. Equipment failures, changes in protocols at a particular dam from year to year, and other irregularities complicate matters even more. This is a real “haystack” of PIT-tag data from which to extract the key “needles” in the form of meaningful comparisons of survival among both source groups and passage histories.

As in the comments by BPA and NMFS, we are critical of the authors’ choice to summarize SAR results only on an annual basis. The determinants of SAR likely vary as much with the environment within a migration year as between years, and these could be tested. The environmental status and hydrosystem-operating mode at the specific time a fish migrates through the system represents the features that are most relevant to survival and are specific targets for modification, rather than average conditions over a migration year. It has been an ongoing criticism of the FPC that they do not further refine their data analyses to within-year conditions (e.g., the ISAB’s comments on the FPC flow augmentation analyses reported in ISAB flow augmentation reviews (ISAB 2004-2)).

We recognize the problems presented by segregating migration histories within years. For example, fish from a release batch disperse in the river and do not all pass a dam at the same time, and therefore individuals experience different environmental and operational histories. However, further breakdown by operational modes or environmental features (such as temperature ranges) could greatly enhance the value of further analyses of the CSS data. The annual summaries can be considered as broad “first cuts” that may be modified by these additional analyses.

The evolving nature of these analyses is reflected in Table 8, which shows older and more recent estimates of the comparison of the differential delayed mortality between transported and in-river fish (D). Despite the number of significant figures reported, the overall number can change, as the influences on it are better understood and included in calculations. Although labeled as a “correction” based on comments on the draft report we see the change as progressive improvement (they may change again).

The study has necessarily aggregated batches of tagged fish, as described at the bottom of page 12. The authors seem to have accounted for this in a reasonable way.

As an overall perspective, there is no way of avoiding the realization that there are a lot of assumptions inherent in the study, from tagging through analyses and presentation of data. Further research should test these assumptions, or tag a sufficient number of appropriate fish so that empirical data can replace assumptions.

Much of the continuing controversy is related to the mathematical and statistical methods employed. We agree with the earlier ISAB comment that the “formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view.” That said, we think the FPC response to the issues raised by NMFS and BPA is quite good. Where questions of bias in estimators are raised, the primary issue appears to be estimating SAR starting from the population at Lower Granite Dam rather than from

other projects. However, the ISAB found the explanation by the CSS scientists as to why the estimate was made in this manner to be reasonable.

There are assumptions made no matter which method is proposed for estimation. For example, the CSS makes the assumption that the transportation proportion for the unmarked population of each hatchery group and the aggregate wild group is approximately the same. Also, it is assumed that the PIT tagged and untagged smolts have the same probability of surviving to and being collected at the dams in the hydro system. These assumptions should be tested.

With respect to the assertion that the PIT tagging reduces survival (see NOAA Fisheries' comments below), we are concerned about the basic premise of the CSS, namely that PIT-tagged fish can serve as surrogates for the unmarked population. If this assertion stands up to further scrutiny, then use of PIT tags should be restricted to comparisons among PIT-tagged groups, and not with unmarked fish.

The use of the bootstrap method to estimate confidence intervals is appropriate. The methodology is now widely used in many statistical applications.

The ISAB hopes the sponsors will more effectively present the methodologies used in the next (2006) Annual Report or in the 10-year summary report we recommend so the criticism of complicated and convoluted formulas can be avoided.

Results (Chapter 3)

The level of scientific satisfaction with the results varies among the species and stocks analyzed. In some cases the results as presented are fairly robust; in other cases where data are scant, trends may be visible but lack statistical significance. The authors present what they have.

Wild Chinook

The problem of small sample sizes for wild Chinook is clearly illustrated by Table 9, which presents the age composition of their PIT-tagged returns. Although a few years had three-digit numbers per age category (1999, 2000, 2002), other years had single- or double-digit numbers. Expansions, while logical, still do not avoid the problem of having few adult returns. Regrettably, it is the wild Chinook that suffer most severely from this concern.

The low return rates of tagged wild Chinook cause the SAR estimates to be very uncertain. The 90% confidence limits of the transport SAR calculations (Table 11) show very wide ranges. What reasonable conclusions can one make when the 90% confidence ranges from zero to over 3? The results do more to demonstrate the *lack* of ability to determine the true SAR than anything. The authors recognize this difficulty in the text on

page 15, and we can take their analyses as a straightforward presentation of the SAR values they calculated using limited data.

The authors were criticized for comparing their calculated SAR values (inexact as they probably are) to the 2% for stable stocks and 4% for recovery recommended by Marmorek et al. (1998). We find no fault with their flagging their calculated values near 1% as a likely problem. We agree with critics of the study that there are better estimates now of stock-specific returns needed for stable populations and recovery, and better calculations of SAR values would be an improvement. But the general trend is unsettling and the CSS results should be taken in their intended context.

The consistent trend in the comparison of SAR values for smolts collected at a collector dam (C_1) and those not detected (C_0) (page 16) also is troubling, despite understood problems with the data. A difference of 25% might just be real. (The table referred to should be Table 12, not Table 10).

In our view, the scant data provide essentially no meaningful information on the relative survival of transported smolts and in-river migrants (T/C ratio) for wild Chinook salmon in all years except 2001 (Figure 4). That year most smolts were transported because of extremely low river flows and high temperatures for in-river migrants, and the transport SAR was high. The values of the differential delayed mortality between transported and in-river migrants (D) have a similar limitation, as the authors note.

We are inclined to view the further analysis of wild Chinook data on pages 19-24 as not warranted based on the scant amount of data available. Perhaps we do not follow the intent of the authors in this section. Further combining of SARs, T/Cs, and D s to come up with sample sizes suitable for statistical analysis seems to us to be inappropriate. The more fruitful direction for the longer term would seem to be to tag more fish in order to match these values with specific operational and environmental regimes that could (at least for operations) be modified to obtain better survival.

Hatchery Chinook

The foundation of data for hatchery Chinook salmon is much better than for wild Chinook (Table 17). However, when taken to the level of specific source hatchery (Table 19), in many cases the data look nearly as sparse as for wild Chinook.

We did not specifically critique the authors' results or discussion of each specific hatchery. The variation among hatcheries is rather expected, based on different rearing conditions, fish size at release, distance from the ocean, etc. The authors seem to have made logical attempts to explain differences in SAR performances. It is interesting that the Rapid River Hatchery seems to be the closest surrogate for wild Chinook. Size effects noted earlier probably deserve more attention.

The T/C ratios among hatcheries are nearly all above 1, indicating superior survival of the transported fish. The ratios are not far above 1, however, and only the estimated error bounds get above 2 (the expected T/C in the absence of *D*).

Wild Steelhead

The numbers of returning adult steelhead are even fewer than for wild Chinook, and thus the results are even less reliable. We view these results as merely presentation of what is available, rather than providing a strong case for any conclusion. Within the limitations of the data, some of the same trends appear as for Chinook, such as higher SAR values for fish not detected as smolts, somewhat higher SARs for transported fish (for steelhead this was above 2 three of 5 years, excluding 2001), and widely varying *D* values. The issue of residualism is important for steelhead, as the authors point out.

Hatchery Steelhead

Low numbers of fish make this analysis problematic. Small sample sizes yield no statistically significant results. However, the authors carry through with the same analyses as for the other groups. The most interesting suggestion is that a possible relationship between fish detected at collector dams and those undetected through the hydrosystem appears to have disappeared in 2000 and 2002.

Adult Drop-out Rates (Chapter 4)

The potential for loss of adults migrating upstream being influenced by the outmigration experiences of the fish as smolts has been raised in the region. We were pleased to see the adult PIT-tag detection data used to track adult upstream movements and losses. The data seem to support conclusions that dropout is higher where there is a fishery (not unexpected), hatchery fish dropped out somewhat more than wild (not stressed by the authors), and that transported fish had a somewhat higher dropout rate than in-river fish. The comparisons in this report just scratch the surface of what can be learned from these data. More important than the Transport/In-river comparisons are potential insights into migration rates at different flows and other environmental differences. Perhaps the emphasis on “survival” in the CSS led to the more narrow focus.

Hatchery-to-Hatchery SARs for Various Hatcheries (Chapter 5)

A basic premise of the CSS was that different survival rates could be calculated for each hatchery from which smolts were released. After many adjustments for terminal fisheries and other factors, this chapter seems to be a straightforward presentation of the SAR values from hatchery back to hatchery for five hatcheries. The problem of small sample sizes is evident. In order to have enough fish for hatchery comparisons, the authors did not do a transported vs. in-river comparison.

Upriver-Downriver Comparisons (Chapter 6)

A prime motivation for the CSS was the hypothesis that the SARs for salmonids that must pass downstream through the hydropower system as juveniles would be lower than those for fish passing no or few dams. To test this hypothesis, there must be adequate representation from both upriver and downriver fish sources.

We concur with critics who express concern that the two downriver sites (Carson Hatchery and John Day River) are probably too few to give accurate upriver-downriver comparisons. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be reliable, it seems prudent to add more downriver sites in the future.

Partition of results into common-year effects and differential mortality as carried out by Deriso et al. (2001) and this study appears reasonable and justified, despite criticisms from Williams et al. (2005). As an editorial note, “fig.y” and later “fig yy” need their numbers.

Estimates of differential upriver-downriver mortality based on spawner-recruit and PIT-tag SAR values provide useful confirmation during the one year of overlap (2000). It would be useful to continue these parallel analyses. We do not understand, however, how averaging 1.48, 0.78, and 1.18 supports the conclusion that upriver stocks survive “about 1/3 as well as John Day populations for these years.”

We were puzzled that the conclusions listed for this chapter did not mention the upriver-downriver comparison for which the chapter was titled. Instead, the conclusions relate to common survival patterns estimated by the two techniques, comparison of wild and hatchery fish, and high correlations among populations. It would have been informative and appropriate to include the comparative survival information (upriver populations survived about 1/3 as well) in the conclusions.

Simulated PIT-tag data to test CJS survival estimates (Chapter 7)

In principle, one can test the reliability of analytical methods by developing simulated data sets and conducting analyses on them. We generally concur that testing the analytical approach with simulated data should provide a useful evaluation of the approach. The present section provides insufficient information, however, to understand what is being done. The abbreviation CJS needs to be defined.

ISAB Evaluation of Comments by BPA and NOAA Fisheries

BPA Comments

BPA was critical of the observational nature of the CSS, the use of a “heuristic analytical approach” devoid of a statistical model, bias in the estimates that lead to incorrect conclusions, misguided emphasis on *D*, a misguided upriver-downriver comparison, and generally flawed and skewed interpretations that minimize the benefits of transportation and the return rates of salmonids. It provided its own mathematical derivation of transported SAR as an appendix.

BPA’s initial criticism that the CSS cannot make direct causal inferences about any particular natural or anthropogenic factor is technically correct, as is the need for manipulative and replicated experiments in order to do so. However, the ISAB and its precursor advisory bodies have requested such manipulative and replicated experiments in the FCRPS for more than a decade, and the requests have been refused by BPA and other action agencies as impractical. BPA is criticizing the CSS for deficiencies in their study when these deficiencies have been caused largely by BPA policy decisions. What the CSS is doing is consistent with its initial study proposal, continuing objectives, and periodic technical reviews.

We do not fault the CSS for its empirical approach. First, the CSS authors do not merely compare hatchery-to-hatchery SAR values, but try several measures of survival along the migration corridor. Survival to Lower Granite Dam is used as a more reliable measure than returns to the hatchery of origin, for example. The CSS has standardized much of its data to the LGR site. We do not see that the approaches used in the CSS analysis are appropriately characterized as biased. As the BPA commenter notes, the issue is somewhat moot because the CSS results do show advantages for transportation in some years, especially in the drought year of 2001.

We do not see that the CSS has focused on *D* as a primary gauge of the effectiveness of transportation. It seems to be presented as one measure along with others. We believe that use of multiple metrics benefits the comparisons. In addition, delayed mortality is real. Therefore, why shouldn’t one calculate the difference in this delayed mortality between transported and in-river fish? We note that the CSS has updated its estimates of *D* based on comments, which we take as a sign of continual improvement.

Some inconsistency between earlier progress reports and this one are to be expected. That’s why they are “progress reports.” This criticism is one reason why the ISAB sees the need for a ten-year summary report as well as the incremental annual reports.

We concur that the upriver-downriver comparison has problems. The BPA commenter correctly criticizes the CSS for relying on just one downstream hatchery when the upstream hatcheries showed such wide variation in results. But the BPA comment does not acknowledge that the CSS also used the John Day River stock for the downriver set. The Hilborn et al. (1993) paper cited by BPA (without reference) does not eliminate the

possibility that information other than that used by Hilborn et al. could show differences between upriver and downriver performance. We would encourage the CSS participants to build on this critique and bolster the downriver samples.

NOAA Fisheries Comments

The NOAA Fisheries comments reflected their belief that the analyses in the progress report are incomplete, do not fully support the findings in the executive summary and chapters, and lack a holistic approach to analyzing all available data. They argue for more in-depth analyses and broader discussion of all relevant data on the effects of the hydropower system on salmonid stocks. They opine that PIT-tagged fish do not represent the untagged populations, that the CSS made selective use of data, that statistical significance is used inconsistently, and that there are biases in the comparisons between treatments and controls. A major point is that the PIT-tagged fish really do not provide a true representation of the untagged population, based on the CSS data. In addition to these general topics, they provided detailed comments by section.

The ISAB suggests that the NOAA Fisheries' expectation that the present annual progress report be a holistic evaluation of all data is unrealistic. That criticism would be more appropriate for a final or periodic summary report. An annual progress report is, by design, of more limited scope. We do agree, however, that a holistic summary is sorely needed after 10 years of work and incremental progress reports.

The NOAA commenter states that the PIT-tagged fish do not represent the survival of the untagged population, while the CSS premise is that they would and the report implies that they do. This is an important difference. In the NOAA Fisheries' comments (and in the technical memo they cite), they note that the PIT-tagged fish returned at about ½ the rate of untagged fish. The data to make these comparisons is in the CSS report, but the CSS authors do not make the comparisons. We agree with NOAA Fisheries that this difference is not trivial and that the CSS must discuss it as well as simply present results. In our view, however, the CSS quite fairly presents the PIT-tag data as its best estimate, although admittedly imperfect. The difficulty comes from comparing the results to the published 2% value for sustainability of a population (tagged and untagged).

We concur that there is some vagueness in statements about statistical significance. On some points, the CSS report simply relies on overlap of the 90% confidence limits. In other places it is not so clear. The CSS could improve this aspect of its reporting. Statistical significance should be tested for and the nature and level of significance of the tests reported.

We concur that size of fish matters and that more attention should be placed on fish sizes in subsequent CSS analyses.

We agree that the Executive Summary could better reflect the results of Chapter 3 in regard to the degree to which hatchery fish can be used as surrogates for wild fish. Nonetheless, the statement that the CSS continues to evaluate this seems appropriate.

As NOAA Fisheries comments, the bullets for Chapter 3 could better represent the text. But these bullets need to be understood as brief summaries of what the text reports.

As we noted before, we concur that use of only one hatchery for the downriver comparison is not good practice, considering the variation seen in results for upstream hatcheries.

The detailed comments are valuable for the CSS to consider as it moves along with the work.

III. ISAB Answers to Council's Questions

1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?

All in all, the design, implementation, and interpretation of the *statistical analyses* underpinning the report are very good. Nonetheless, there are broader concerns over the design of the study such as sample size, sampling sites, time periods for analyses, and other features. Improvements can be made, and our recommendations follow.

Since the region is unwilling to conduct the manipulative experiments in the hydrosystem that the ISAB and ISRP have recommended for many years, the CSS is doing the next best thing. That is, the study is following as many fish through their life cycle as possible, calculating the survival, and comparing outcomes.

The study design could be improved in several ways. Adding more downriver hatcheries to make more valid upstream/downstream survival comparisons. Much more attention should be given to the size of tagged fish at various release locations, because survival is known to be affected strongly by fish size. The data could be aggregated to more closely meet the needs of hydrosystem managers. Whether by design or implementation, the aggregation of data simply by year of outmigration is insufficient to resolve many of the important issues related to environmental influences and hydrosystem operations. The numbers of fish tagged may never be sufficient for resolving in-season patterns of survival. However, as data are accumulated over more years, it may be feasible to partition analyses into environmental or operational categories across years to obtain more functional correlations. Having a controlled and manipulated experimental design would be preferable (as BPA asserts), but the chance of this happening is slim. Repeated entreaties by the ISAB, its predecessor advisory bodies and the ISRP have all been met with objections to the effect that such a system wide experiment is not possible to manage (although we note that the region managed to implement high spill in 2005 on court order, although no planned experiments were conducted). The opportunistic approach of documenting survival under whatever conditions are dealt seems to be the only alternative.

Implementation would be improved by tagging more fish (particularly wild), but there is likely a limit to the amount that can be accomplished due to manpower limitations. The study managers have been quite opportunistic in arranging tagging and in coordinating tagging efforts among many different entities. Pre-assignment of fish to either inriver or transport passage routes at the time of release would greatly improve study design and make the analyses and results more transparent. Assignment of passage route at release is planned for implementation in 2007 (i.e., a given tag number would really be “destined” to be shunted to a particular route, if possible). This modification should be implemented in 2006, if possible.

The data analyses require extensive statistical manipulations to extract useful information from the mass of PIT-tag detections. We can only agree with the earlier ISRP comment that the "formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view." Pre-assignment of fish to inriver and transport groups at time of release should help. The study participants have gone to great lengths to seek ways to analyze the data appropriately. Bootstrapping confidence limits is a major improvement. We do not find any particular bias in the analyses or interpretations. Likewise, we see no inherent problem with the assumptions, and some assumptions will always have to be made. These assumptions should be tested as the project progresses.

Taken alone, the current progress report does not adequately present the analytical methods and some data presentations are difficult to follow (e.g., labeling axes as log survival instead of actual survival). The ISAB encourages the sponsors to more effectively present the methodologies in a summary report (perhaps as part of the 2006 Annual Report) so the methods of analysis can be better understood.

2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

The Council’s question is difficult to answer with just the present annual progress report. The value of this project for informing management decisions on the hydropower system would be greatly enhanced if a synthesis report were produced that clearly describes the analytical methods and summarizes the project results in a holistic way for its decade of effort. We recognize that this is what NOAA Fisheries hoped to see.

The CSS is providing long-term monitoring of lifetime survival of salmon and steelhead stocks using a technology that the region has spent a great deal of money developing and implementing. As an ongoing effort, subject to periodic review and comment, it is providing an evolving picture. It would be wrong to believe that the results as of today are the end-all for making decisions about the operation of the hydrosystem. The CSS is learning as it goes, which is to be expected. More years and more analyses of specific questions are needed.

Because the CSS is focusing on annual data, the relationships to specific operational and environmental factors within years are not addressed. As commenters have pointed out, these more specific correlations would be more useful for guiding operational decisions. The ISAB recognizes a disconnect between the present status of results and much of the decision-making that takes place regarding hydrosystem operations and fish protection. Although the project is making good progress at addressing such issues as the value of transportation and the relative survival from different passage routes, many relationships between survival and operational or environmental features during migration cannot be resolved when data are aggregated simply by year of migration. For this information to be most useful for making decisions, aggregations of data within years or across years for different operational options and environmental conditions need to be pursued. Even after aggregating the available, relevant data across several years, there may not be a sufficient number of tag detections to make such correlations for all important combinations of operational status and environmental conditions. Either more fish need to be tagged or correlations made after more years of data for which operational and environmental modes can be grouped. The former would be the more expeditious approach.

IV. ISAB Conclusions and Recommendations

The CSS is an ambitious, long-term study that is being criticized because its objectives are not yet fully met, despite prodigious efforts in both the field and in complex data analyses. It has used the PIT-tag technology to mark and track individual salmon and steelhead through their smolt-to-adult life stages. Expectations of this mark-recapture technology exceed the results that are practically attainable, and its use is still evolving. The CSS study participants have been major players in this evolution. We find the present annual report to be a further incremental step in the direction of documenting different survival rates of different stocks under different migration conditions. That the present report is not a perfect reconstruction of differential survival histories is largely a result of the current analytical capabilities and available sample sizes. The deficiencies seem to be highlighted in some aspects because of experimental design and analytical approaches taken by the authors. The ISRP comment from their 2002 review still applies that “the formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view.”

The Council should view the CSS as a good, long-term monitoring program the results of which will become increasingly valuable to managers as years pass. Scrutiny from periodic peer reviews and agency comments will help ensure that the methods and analytical approaches improve. The project is definitely worthy of Council support.

Recommendations

- It has been ten years since the CSS was initiated. The report the ISAB reviewed was the latest in a series of annual progress reports, and thus lacking a holistic perspective. The ISAB recommends that the CSS produce a ten-year summary report providing an

in-depth description of methods and detailed analyses and interpretation of the data in a retrospective style.

- The CSS needs to more effectively present the methodologies used in their analyses so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated in the ten-year summary recommended above.
- The ISAB agrees with critics who express concern that two downriver sites (Carson Hatchery and John Day River) are probably insufficient to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.
- Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.
- Aggregation of data solely by juvenile migration year should be supplemented with analyses that group data on environmental and operational factors that may be amenable to control.
- Assumptions inherent in the analyses should be specifically tested, with continued vigilance toward avoiding bias.
- Pre-assigning the intended routes of passage at the time of release into in-river and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the study design is scheduled for implementation in 2007, but should begin in 2006, if feasible.
- Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (*D*). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.
- The CSS should be supplemented by funded research into analytical methods that can improve, and hopefully simplify, the mathematical and statistical approaches currently in use. It is not clear from available information whether the problem is that the formulas are unnecessarily complicated, inappropriately specified, or just not well explained (see bullet #2 above).

- More attention should be given by the CSS and the Region as a whole to the apparent documentation that PIT-tagged fish do not survive as well as untagged fish. This point has major implications for all uses of PIT-tagged fish as surrogates for untagged fish.

References

ISAB 1998. Review of Comparative Survival Rate Study of Hatchery PIT Tagged Chinook. Report ISAB 98-1, Northwest Power Planning Council and National Marine Fisheries Service, Portland, Oregon.

ISRP 2002. Final Review of Fiscal Year 2003 Mainstem and Systemwide Proposals. Report ISRP 2002-14. Northwest Power and Conservation Council, Portland, Oregon.

Appendix A: Previous Review Comments by ISAB and ISRP

ISAB Comments (ISAB 1998)

- Fund the proposed study.
- So long as the present configuration and operation of the federal hydroelectric system exists, extend (or continue) PIT tagging to include naturally reproducing populations of spring chinook whenever population sizes may permit. Continue PIT tagging other chinook life history types, and extend PIT tagging to other life history types of other species of salmon, including steelhead, whenever possible.
- Apply enough PIT tags to spring chinook production from Kooskia, Pahsimeroi, McCall, Sawtooth, and Clearwater (Powell, Crooked River and Red River Ponds) hatcheries to estimate survival to Lower Granite Dam. Whenever possible apply enough PIT tags to spring chinook at these hatcheries to estimate survivals to McNary Dam.
- Compare rates of return to each hatchery of PIT tagged and untagged adults to establish degree of comparability of survivals of PIT tagged juvenile salmon to survivals of juveniles not PIT tagged. To investigate rate of shedding of PIT tags through the adult stage, and where straying of adults from another hatchery is possible, investigate thermal mass marking of all hatchery production. Where smolt to adult survival of PIT tagged fish is compared to that of coded wire tagged (CWT) fish, develop a procedure to study tag loss and to compare rate of return of PIT to CWT within the hatchery release.
- Make estimates of survival applicable to the entire Snake-Columbia River federal hydroelectric system as soon as possible.
- Promote coordination and cooperation among agencies applying PIT tags and other marks by including a list of other agencies marking salmon and steelhead of the same origin in the proposal, along with comments from those other agencies. Sponsor an interagency workshop on the use of tagging data at five-year intervals. The workshop would produce consensus recommendations and procedures for coordinating tagging activities.

ISRP Comments (ISRP 2002)

Various scientists in the region, in particular scientists from the Comparative Survival Study project and NMFS, have considered the problems in estimating the LGD to LGD smolt-to-adult survival rates (SARs) from currently available data and have apparently arrived at what they consider to be the “best” formulas. Unfortunately, the formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view. Accordingly, there is high probability that these methods will continue to spawn arguments and counter-arguments over trivial issues that will occupy the resources of the

region, because the stakes are high (e.g., high costs of spill, high costs of transportation, unknown long term effects of the non-normative transportation, high costs of flow augmentation, etc).

The long-term solutions to the mathematical and statistical problems in estimation of smolt-to-adult return rates (Bonneville to Bonneville and Bonneville to Low Granite SARs) appear to be: 1) detection of sufficient numbers of PIT tagged juveniles passing Bonneville Dam Powerhouse II at the planned corner collector; 2) estimates of mortality of fish passing via that route; 3) and/or sufficiently large sample sizes of PIT tagged fish downstream of Bonneville. The ISRP recommends that these sampling efforts for PIT tagged juveniles be given high priority by the Council and the Corps of Engineers. In particular, Task 2 of NMFS proposal #198331900 for development of PIT tag detection in the corner collector at Bonneville Dam Powerhouse II should be given high priority.

We do not provide unqualified endorsement of the particular estimation formulas that are proposed, and we recommend that continuing statistical methods research be directed at investigating the performance of various proposed estimators and possible alternatives, including but not limited to the proposed methods and planned bootstrapping. Such research on mathematical and statistical methods could be pursued by the sponsors of this project, and by others. As an aid to clarity in comparison among possible alternative analyses, we recommend that the FPC make available a single reference data set which includes all the necessary interpretation of route of passage of PIT tagged fish and culls any suspect or ambiguous data that might be subject to further interpretation. The budget for the recommended mathematical and statistical analyses is relatively minor compared to the total cost of the project so investigation of our unresolved questions about statistical methods should not require substantial reallocation of the budget in this project to ensure compatibility of objectives, common methods and protocols. This coordination could be accomplished under the favorably reviewed CBFWA proposal #35033.