Collaborative Systemwide Monitoring and Evaluation Project (CSMEP)

Collaborative Systemwide Monitoring and Evaluation Project (CSMEP)

Project No. 2003-036-00

Snake River Basin Pilot Report

Volume 1

COLUMBIA BASIN FISH AND WILDLIFE AUTHORITY

Collaborative Systemwide Monitoring and Evaluation Project (CSMEP)

Project No. 2003-036-00

Snake River Basin Pilot Study

Volume 1

Submitted to

Tracy Yerxa
KEWB-4
Bonneville Power Administration
P. O. Box 3621
Portland, OR 97208-3621

Compiled by

David R. Marmorek, Marc Porter, Darcy Pickard and Katherine Wieckowski
ESSA Technologies Ltd.
1765 West 8th Avenue, Suite 300
Vancouver, BC V6J 5C6

On behalf of the

Columbia Basin Fish and Wildlife Authority
851 SW Sixth Avenue, Suite 260
Portland, OR 97204

Nov. 15, 2007

Contributors:

Robert Al-Chokhachy (Eco Logical)
John Arterburn (CTCR)
Chris Beasley (QC-NP)
Tom Berggren (FPC)
Bill Bosch (YN)
Nick Bouwes (Eco Logical)
Alan Byrne (IDFG)
Rich Carmichael (ODFW)
Tim Copeland (IDFG)
Tim Dalton (ODFW)
Dave Fast (YN)
Jeff Fryer (CRITFC)
Peter Galbreath (CRITFC)
Peter Hahn (WDFW)
Jay Hesse (NPT)
Annette Hoffman (WDFW)
Saang-Yoon Hyun (CRITFC)
Chris Jordan (NOAA)
Steve Katz (NOAA)
Lyman MacDonald (WEST)
Ken MacDonald (CBFWA)
Dave Marmorek (ESSA)
Claire McGrath (USFS)
Rick Orme (NPT)
Charlie Paulsen (PES-BPA)
Charlie Petrosky (IDFG)
Darcy Pickard (ESSA)
Marc Porter (ESSA)
Dan Rawding (WDFW)
Michael Rayton (CTCR)
Kris Ryding (WDFW)
Tom Rien (ODFW)
Phil Roger (CRITFC)
Kris Ryding (WDFW)
Sam Sharr (IDFG)
Eric Tinus (ODFW)
Earl Weber (CRITFC)
Paul Wilson (USFWS)
Keith Wolf (CTCR)
Frank Young (CBFWA)

Citation: **Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) - Marmorek, D.R., M. Porter, D. Pickard and K. Wieckowski (eds.).** 2007. Snake River Basin Pilot Study: Volume 1. Prepared by ESSA Technologies Ltd., Vancouver, B.C. on behalf of the Columbia Basin Fish and Wildlife Authority, Portland, OR. 2007. Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) Snake Basin Pilot Report. Prepared by ESSA Technologies Ltd., Vancouver, B.C. on behalf of the Columbia Basin Fish and Wildlife Authority, Portland, OR. 47 pp.

Executive Summary

The Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) was created for the shared, multi-agency development of a regional monitoring and evaluation (M&E) program for fish populations. It is a bottom-up effort to build consensus to ensure technically and consistently sound programmatic decisions on M&E. Specific goals for CSMEP are to: 1) document, integrate, and make available existing monitoring data on listed salmon, steelhead and other fish species of concern, 2) critically assess strengths and weaknesses of these data for answering high priority monitoring questions, and 3) collaboratively design and help agencies implement improved monitoring and evaluation methods related to key decisions in the Columbia Basin.

CSMEP adopted the Environment Protection Agency's (EPA) Data Quality Objectives process (DQO) to guide development and evaluation of alternative designs within the five M&E domains (Status & Trends, Harvest, Hydrosystem, Habitat and Hatcheries). The DQO process helped CSMEP to clarify program objectives, define the appropriate types of data to collect/analyze and specify tolerable limits on potential decision errors. This provided a basis for establishing the quality and quantity of data needed to support management decisions. For habitat action effectiveness M&E, CSMEP additionally developed a 'Question Clarification' process that provided some greater flexibility in identifying information needs. In conjunction with the DQO, CSMEP has been using a structured decision analysis approach to help evaluate trade-offs across the M&E design alternatives. CSMEP's evolving quantitative tools and analyses allow assessment of a variety of M&E design alternatives, in terms of both qualitative and quantitative evaluative criteria.

Systematically developing and evaluating alternative M&E designs is complex. CSMEP, therefore, initially focused on spring/summer Chinook in the Snake River Basin ESU, as a test case to refine design methods and analytical tools. The Snake River Basin was considered large enough to present many of the M&E challenges typical of the entire Columbia River Basin, including consideration of tradeoffs among monitoring objectives, and forced CSMEP scientists to use relevant data from other regions, particularly for hydro, hatchery and harvest questions that are Columbia River Basin-scale in nature. CSMEP's design evaluations within the Snake River Basin pilot study are described for each of the five M&E domains.

Status and Trends

Status Quo monitoring for Snake Basin Spring Summer Chinook contains weaknesses for assessing viability at the population level as per IC-TRT viability criteria. The current monitoring does not assess spatial structure information in many populations and lacks abundance estimates in non-index areas for populations without weirs or spatially representative redd counts. CSMEP's recommended 'Medium' design would cost considerably less than the Status Quo monitoring, yet should perform better in answering the question: is the ESU viable? It must be recognized that Status Quo monitoring has not been developed to address only this single viability question, but is rather a consolidation of weirs, redd counts, and other monitoring that is being done to address a variety of questions. However, it appears that a simple reallocation of resources to Status Quo monitoring in the Snake River Basin could address current weaknesses and improve viability assessments. This would require; (1) changing the redd survey program to CSMEP's 'Medium' design where all populations have multiple redd counts and spatial structure assessed, and (2) installing a weir in the Middle Fork Salmon River MPG.

The IC-TRT rule set is conservative, so high uncertainty generally results in underestimating viability. The most likely error from CSMEP simulation models was in depicting a population as 'Not Viable', when the population is in fact 'Viable'. This common result must be considered when evaluating the tradeoffs among designs. While simpler designs for monitoring viability may be less costly in the short term, inferior data resulting from such designs may incur higher costs over the long term due to the inability to make a correct assessment of the ESU.

Harvest

Status Quo harvest monitoring generally does not provide precision estimates around harvest impacts. Such estimates, however, would improve the ability of managers to quantify risks of harvest management decisions. Uncertainty around harvest impact estimates can result in overharvest of listed stocks or conversely in lost harvest opportunities. It can also contribute to uncertainty around evaluation of status, trends and viability. New analytical techniques are required for preseason and in-season abundance forecasts, although improvements to run size estimates and inseason forecasts may be possible at modest cost with available data and methods. There is a need to evaluate new technologies/techniques for improved stock identification and composition estimates (e.g., PIT tags, GSI). These techniques may be suitable to improve stock identification resolution. Ultimately, there is a considerable need to further improve coordination between entities collecting fisheries harvest monitoring and evaluation information.

Hydro

Status Quo monitoring has allowed a good estimate of annual compliance with the SAR target for wild spring-summer Chinook, but this is partly because SARs have historically been so far below the target. If SARs get closer to the 2-6% target range, higher precision estimates may be required to definitively assess compliance. CSMEP's 'High' design improves the precision of estimates of SARs and in-river survival for wild spring-summer Chinook, allowing more definitive evaluations of annual compliance with targets than is possible with Status Quo monitoring. CSMEP's 'Medium' design enables more *representative* estimates of hatchery survival than is possible with Status Quo monitoring, but has little effect on statistical reliability. CSMEP's 'Low' design, which drops CSS tagging of hatchery fish, would substantially reduce the current ability of managers to assess annual compliance of in-river survival targets (wild plus hatchery fish), and the ability to assess transportation effectiveness for hatchery fish.

Multiple-year estimates should be used for assessing compliance, in addition to annual estimates. Multiple-year estimates can provide insights on compliance with only a relatively small number of PIT-tags (e.g., 5,000 tags), which permits analyses on smaller spatial scales (e.g., MPGs, some large populations) and smaller temporal scales (in-season patterns). Increasing the number of tags per year will improve the precision of annual and seasonal estimates, but for transportation evaluations a very large increase in tags would be required to make substantive improvements over the Status Quo, and is likely not cost-effective. For multiple-year estimates, statistical precision increases with increasing tag numbers up to 5,000 tags, but beyond this level little further benefit is seen. Adding more years to those averages can significantly improve statistical precision. But there is a tradeoff however, in that longer durations of monitoring (e.g., beyond 5-10 years) may be beyond the time scales of interest for some decisions.

Habitat

Various issues must be resolved in creating designs for habitat action effectiveness monitoring. Practical action effectiveness monitoring designs must first incorporate sufficient analytical flexibility to compensate for less than complete control over action implementation. Also it is likely that long term Status Quo designs (generally intended for status and trends monitoring), cannot provide adequate information at the temporal and spatial scales required for efficient implementation of action effectiveness evaluations. Thus, implementation of action effectiveness evaluations will necessitate both new sampling effort and the modification of existing sampling efforts. Further targeted research on the mechanistic linkages between habitat restoration actions and fish population responses is also still needed.

Any of CSMEP's designs for monitoring the effectiveness of habitat actions in the Lemhi River watershed (their pilot area for developing designs) would provide better information than the ongoing Status Quo monitoring in the watershed. Although each CSMEP design alternative would allow quantitative evaluations of the effects of reconnection projects on fish populations to varying degrees of accuracy and precision, CSMEP's more intensive and costly 'Medium' or 'High' designs would likely be required for discerning the mechanistic connections between restorative actions and fish response (i.e., why actions worked or did not). While simpler designs for monitoring effectiveness may appear less expensive in the short term, they are likely to be ultimately more costly as monitoring will need to be continued longer to detect effects. Simpler designs will also lack the added benefit of providing transferable mechanistic information on the benefits of specific projects or project types that can inform cost savings in other watersheds.

As one moves to other subbasins where habitat management issues are diverse, there are likely to be potentially large differences in design elements; in particular, where and when to deploy monitoring resources. It will be impossible to predict this ahead of consideration of the mature scientific questions specific to those locations. Consideration of those questions will in turn require a unique rather than template process that is informed by the management history and management plans in those new locations.

Hatcheries

Columbia River Basin status quo hatchery RME is primarily focused at the scale of individual projects. At that scale, the existing RME is likely to provide adequate information to evaluate hatchery mitigation goals and to address the impacts of hatchery supplementation on abundance and productivity of targeted populations. Alternatively, little existing research is focused on the aggregate impact of hatcheries at larger spatial scales (drainage or basin level), particularly in regard to the impact of hatchery straying and relative reproductive success (RRS) in non-target populations. The current non-random distribution of straying and RRS monitoring precludes statistically valid inference from sampled to un-sampled populations. As a result, under the Status Quo, monitoring effort must be deployed wherever we want an answer. Methods for collecting, analyzing, and reporting data also vary significantly among agencies. Thus, even if effort were representatively distributed, it is unclear whether the resulting information could currently be aggregated and analyzed to enable statistically valid inference to un-sampled populations.

CSMEP's recommended 'Medium' stray ratio design provides stray ratio estimates at the population scale and enables estimates of precision and bias in carcass recovery methods, while the recommended 'Medium' RRS design ensures that RRS can be calculated over the entire life-cycle, although it will not give comparable productivity estimates in un-supplemented populations. Implementation of any of CSMEP's designs for stray ratio and relative reproductive success (RRS) offers substantial improvement over the Status Quo. While RME costs would increase over the short-term, in the longer-term the inferential ability afforded by even the low designs will significantly reduce RME expenditures within the Columbia River Basin. Under the Status Quo, RME is required for every program/population for which information is desired. While the CSMEP designs do not supplant the need for all program specific RME, they do significantly reduce the breadth of RME that would otherwise be required to accompany all programs. In addition, the CSMEP designs enable an evaluation of the aggregate impacts of hatcheries, which cannot be achieved given existing RME. Perhaps most importantly, the CSMEP designs enable informed decisions with regard to the use of hatcheries, and achieve this goal by building on existing RME effort, thus affording substantial cost-efficiency.

Integration

Monitoring and evaluation involves systematic long-term data collection and analysis to measure the state of the resource, detect changes over time and test action effectiveness. Currently, fish populations in the Columbia River Basin are monitored by a number of separate programs established by different agencies. Most of the fish monitoring programs were designed to answer specific management questions at small spatial and temporal scales, and utilize different measurement protocols and sampling designs. This has resulted in an inability to efficiently integrate monitoring at larger spatial scales required for ESU or regional fish population assessment. There is a need for consistent, long-term integrated monitoring of Columbia River Basin fish populations.

Developing a workable plan for efficiently integrating Columbia Basin-wide M&E (spatially, temporally, ecologically and programmatically) will likely involve multiple, simultaneous strategies, which CSMEP has been pursuing in their Snake River Basin pilot. These strategies include:

1. *Building on a Status & Trends foundation.* Layering of action effectiveness M&E alternatives on a consistent foundation of spatially representative Status and Trends monitoring
2. *Integration within domains*. Evaluating how alternative designs could best address multiple questions within a particular M&E domain (i.e., Hydrosystem, Hatchery, Harvest, Habitat, or Status & Trends specific)
3. *Integration across domains*. Evaluating how alternative designs could best address multiple questions across M&E domains (e.g., what elements of each subgroup's designs can serve multiple functions)
4. *Maximizing benefits of monitoring techniques*. Evaluating how any particular monitoring technique can help address multiple questions across M&E domains (e.g., PIT tagging to address a suite of questions)
5. *Maximizing sampling efficiencies and minimizing redundancies in designs*. Evaluating shared costs and data gathering opportunities across overlapping designs.

General CSMEP recommendations

Regional M&E for fish populations should be developed through a long-term, systematic process that involves dialogue with Columbia River Basin fish managers and decision makers to identify the key management decisions, spatial and temporal scales of decisions, information needs, time frame for actions, and the level of acceptable risks when making the decisions. It should be recognized that monitoring and evaluation are absolutely critical to the region's adaptive management cycle.

Decisions on regional M&E designs need to be based on a quantitative evaluation of the costs and benefits of the Status Quo and alternative designs to answer management questions. It will likely be much more cost-effective to build on the strengths of the region's existing monitoring infrastructure, rather than applying a uniform "cookie-cutter" approach throughout the Columbia River Basin. Each region in the Columbia River Basin has invested considerable resources to develop a monitoring infrastructure that is primarily adapted to address local needs. Improved designs that can overcome weakness in the existing M&E programs should allow assessments at larger spatial and longer temporal scales.

The development and implementation of sound M&E designs must be accompanied by strong data management systems which facilitate the sharing, analysis and synthesis of data across agencies, spatial and temporal scales, and disciplines. Without a strong investment in data management, even the best monitoring designs will falter.

Status and trends monitoring of fish populations must satisfy the needs of population and ESU level assessments (for both listed and unlisted species) of viability, as well as assessments of overall trends in population abundance and productivity at larger spatial and longer temporal scales. It must also meet the needs of multiple agencies with different objectives, questions, and scales of interest.

Status and trends monitoring can provide the foundation of a regional M&E program but it must be integrated with action effectiveness monitoring. An integrated M&E program provides economy of scale, prevents duplicative efforts, and is cost effective. Action effectiveness monitoring is more focused on specific questions that influence fish populations hence, it is typically of fixed duration and usually provides more precision. It can respond to adaptive management needs by focusing its efforts to address the mechanistic causes of uncertainty in the relationship between management actions and fish population responses. Action effectiveness monitoring designs must respond to highly varied M&E needs. M&E designs under development must also be integrated across species.

Agencies should evaluate hybrid sampling designs to improve fish population monitoring that is based on fixed index sites. A hybrid sampling design would supplement the existing non-random, index monitoring sites with spatially representative sites. This approach would allow agencies to assess the bias in index sites, get reliable estimates of population abundance for viability assessments, permit aggregation to a variety of larger spatial scales (e.g., MPG, sub-basin), support the sharing of data collected by different agencies with different interests, and facilitate data analyses.

Table of Contents

List of Tables vii

List of Figures vii

1. Overview of the CSMEP Snake Basin Pilot 1

1.1 Introduction 1
1.2 Process of developing and evaluating alternative M&E designs 1
1.3 CSMEP's Strategic Approach 4
1.3.1 CSMEP's Snake River Basin Pilot 5

2. Specific Results 7

2.1 Status and Trends 7
2.1.1 Priority question 7
2.1.2 What are the consequences of making the wrong decision? 7
2.1.3 Monitoring design alternatives and trade-off analyses 7
2.1.4 Design alternatives 8
2.1.5 Tradeoff analyses 8
2.1.6 Conclusions and recommendations 11
2.2 Harvest 12
2.2.1 Priority questions 12
2.2.2 Related decisions 12
2.2.3 What are the consequences of making a wrong decision? 12
2.2.4 Monitoring design alternatives and trade-off analyses 12
2.2.5 Conclusions/recommendations 13
2.3 Hydro 18
2.3.1 Priority question 18
2.3.2 Related Decision 18
2.3.3 Consequences of wrong decisions 19
2.3.4 Monitoring design alternatives and trade-off analyses 19
2.3.5 Conclusions and recommendations 24
2.4 Habitat 26
2.4.1 Priority questions and the question clarification process 26
2.4.2 Related decisions 27
2.4.3 Consequences of wrong decisions 27
2.4.4 Monitoring design alternatives and trade-off analyses 28
2.4.5 Conclusions & recommendations 29
2.5 Hatcheries 32
2.5.1 Related decisions 32
2.5.2 Monitoring design alternatives and trade-off analyses 33
2.5.3 Conclusions 36
2.5.4 Design recommendations 37
2.6 Integrated Monitoring 38
2.6.1 Integration Strategies 39
2.7 Summary of general recommendations 44

References 47

List of Tables

Table 1.1. Examples of M&E design objectives and evaluative criteria. 4
Table 2.1.1. Description of four monitoring design alternatives and how they differ for each performance measure. 9
Table 2.1.2. Trade-off analyses of each design for assessing the viability of the Snake River spring/summer Chinook salmon ESU. 10
Table 2.2.1. Description and evaluation of harvest monitoring design alternatives. 15
Table 2.3.1. Description and evaluation of hydro monitoring design alternatives (2 page table). 21
Table 2.4.1. Example of a key general question about habitat effectiveness and the nested "question clarification" process used to precisely determine the specific information needs required to address this question sufficiently for management purposes. 27
Table 2.4.2. Alternative sampling and response designs for evaluating Lemhi River subbasin habitat actions (**what**, **how, where** data are collected). 30
Table 2.4.3. Overall effectiveness monitoring designs for evaluating effectiveness of Lemhi River watershed habitat restoration actions, and qualitative assessment of design alternatives. 31
Table 2.4.4. Costs of alternative CSMEP habitat action effectiveness monitoring designs for the Lemhi River subbasin. 31
Table 2.5.1. Objectives by alternatives matrix for hatchery stray ratio designs. 35
Table 2.5.2. Objectives by alternatives matrix for the relative reproductive success designs. 36
Table 2.6.1. Abbreviated list of questions answerable in whole or in part with PIT-tagged fish. 41

List of Figures

Figure 1.1. The EPA's Data Quality Objectives process (DQO) 2
Figure 1.2. Flow of the PrOACT decision process recommended by CSMEP to narrow the range of acceptable M&E designs. 3
Figure 1.3. Insights gained from the CSMEP Snake River Basin Pilot study (blue shaded area) will have applications to other areas of the Columbia River Basin (CRB) and will similarly benefit from analyses being undertaken elsewhere in the CRB. 5
Figure 1.4. Anadromous and resident fish lifecycles and associated M&E domains. 6
Figure 2.4.1. Map of the Lemhi River watershed denoting Sections A (migration corridor), B (action area), and C (potential reference area). 28
Figure. 2.6.1. Conceptual illustration of identification of opportunities and subsequent development of integrated monitoring designs across CSMEP subgroups. 40
Figure 2.6.2. Monitoring techniques and potential linkages across status & trends and action effectiveness monitoring. 42
Figure 2.6.3 Front-end user interface for CSMEP's Cost Integration Database Tool....43
Figure 2.7.1. The adaptive management cycle, with example Columbia Basin entities included. 44

Glossary

BLM	Bureau of Land Management
BONN	Bonneville Dam
BoR	Bureau of Reclamation
CBFWA	Columbia Basin Fish and Wildlife Authority
CRB	Columbia River Basin
CSMEP	Collaborative Systemwide Monitoring and Evaluation Project
CSS	Comparative Survival Study
CTUIR	Confederated Tribes of the Umatilla Indian Reservation
CWT	coded wire tags
BiOp	Biological opinion for operation of the Federal Columbia River Power System.
DIT	double index tagging
EPA	US Environmental Protection Agency
ESA	Endangered Species Act
ESU	Evolutionary Significant Unit
FCRPS	Federal Columbia River Power System
GSI	genetic stock identification
HCP	Habitat Conservation Plan
IC-TRT	Interior Columbia Technical Recovery Team
ISEMP	Integrated Status and Effectiveness Monitoring Program
IDFG	Idaho Department of Fish and Game
ISS	Idaho Supplementation Studies
LCR	Lower Columbia River
LGR	Lower Granite Dam
LOM	looking outward matrix
M&E	monitoring and evaluation
MaSA	major spawning area
MiSA	minor spawning area
MR	mark-recapture
MPG	major population group
NOAA	National Oceanic and Atmospheric Administration
NPCC	Northwest Power and Conservation Council
NPT	Nez Perce Tribe
ODFW	Oregon Department of Fish and Wildlife
PIT tags	Passive Integrated Transponder tags
PNI	proportionate natural influence
PTAGIS	PIT Tag Information System
RME	research, monitoring, and evaluation
RRS	relative reproductive success

SAR	smolt-to-adult return rate
SBT	Shoshone Bannock Tribe
SR	Snake River
TIR	transport to in-river ratio
USACE	US Army Corps of Engineers
USFS	US Forest Service
WDFW	Washington Department of Fish and Wildlife
VSI	visual stock identification
VSP	viable salmonid population

1. Overview of the CSMEP Snake Basin Pilot

1.1 Introduction

The Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) was created to involve federal, state and tribal scientists and managers in the collaborative, multi-agency development of a regional monitoring and evaluation (M&E) program for fish populations. It is a bottom-up effort to build consensus across multiple agencies to ensure technically and consistently sound programmatic decisions on M&E. Specific goals for CSMEP are to: 1) document, integrate, and make easily available existing monitoring data on listed salmon, steelhead and other fish species of concern, 2) critically assess strengths and weaknesses of these data for answering high priority monitoring questions, and 3) collaboratively improve design of M&E related to key decisions in the Columbia Basin.

1.2 Process of developing and evaluating alternative M&E designs

An M&E design is the description of the combination of logical, statistical, logistical, and cost components associated with a particular approach to answering management questions. General design strategies have been prepared for other programs in the Columbia River basin. For example, Hillman (2004) describes an overall monitoring and evaluation strategy for the Upper Columbia Basin using four components: 1) a “statistical” design, which provides the logical structure and identifies the minimum requirements for status/trend and effectiveness monitoring; 2) a “sampling” design which describes the process for selecting sampling sites; 3) a “measurement” design outlining the specific performance measures and how to monitor them; and 4) a “results” design that explains how the monitoring data will be analyzed to make inferences. Consistent with this approach CSMEP has adopted the US Environmental Protection Agency’s DQO (EPA 2000) process to guide the development and evaluation of alternative M&E designs (Figure 1.1).

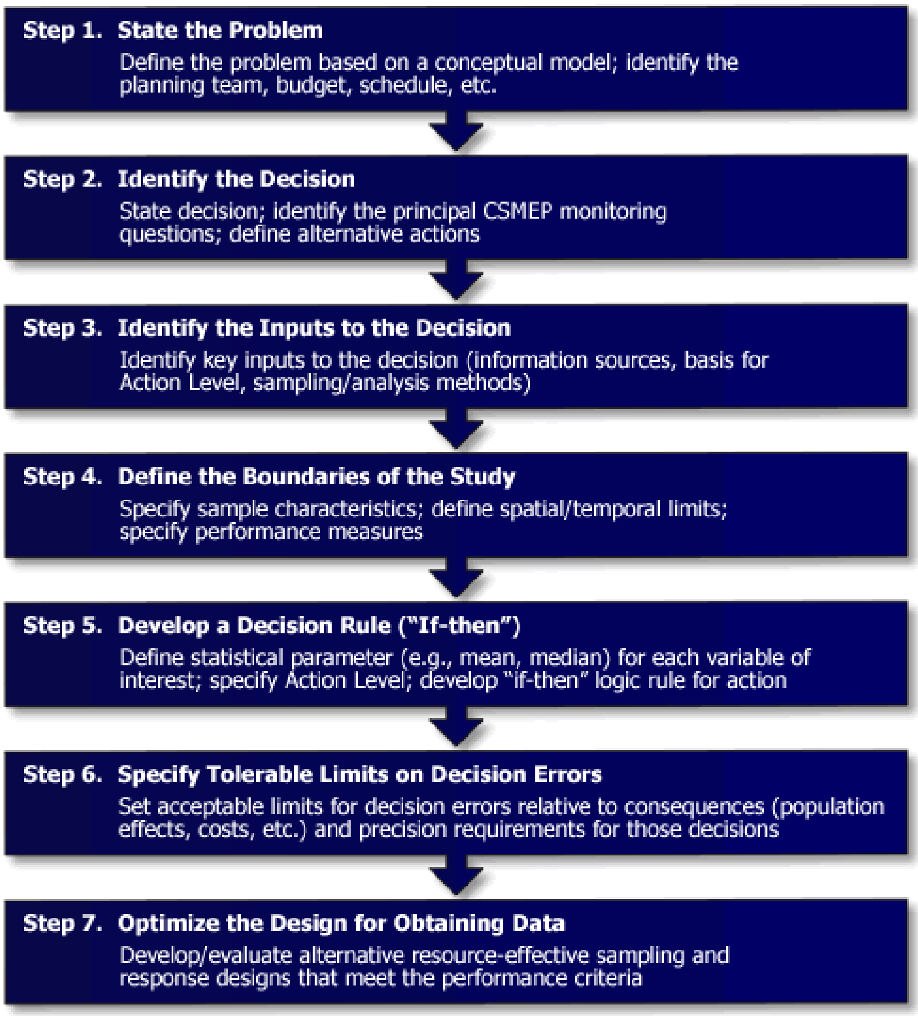

Figure 1.1. The EPA's Data Quality Objectives process (DQO) (source: EPA 2000). The DQO process is a collection of qualitative and quantitative statements that help to clarify program objectives, define the appropriate types of data to collect/analyze and specify tolerable limits on potential decision errors. This provides a basis for establishing the quality and quantity of data needed to support decisions. The DQO approach has forced CSMEP scientists to consult with program managers on the management decisions to be made, explore alternative analytical/evaluation approaches to those decisions, define the performance measures required to feed those analytical approaches, and design the sampling required to generate the data for the key performance measures. For habitat action effectiveness M&E, we used a 'Question Clarification' process that provided greater flexibility in identifying information needs.

Although development of effective designs within M&E domains is critical it does not of itself provide Columbia River Basin agencies with the information to converge on an 'optimal' M&E program. Ultimately, this involves analyzing the benefits and costs of different designs across multiple client agencies, objectives and M&E domains. It is not an easy problem. CSMEP has been applying the PrOACT approach (Hammond et al 1999) for evaluating cost-effective M&E design alternatives within the five M&E domains, and recommends applying this across domains. ProACT (Figure 1.2) is a simplified approach to multi-objective decision analysis. The acronym stands for ***P***roblem definition, determination of ***O***bjectives, development of ***A***lternatives (M&E designs), calculation or assessment of the ***C***onsequences associated with each alternative across the set of objectives, and evaluation of

Tradeoffs between alternatives for particular objectives, or between objectives within a particular alternative.

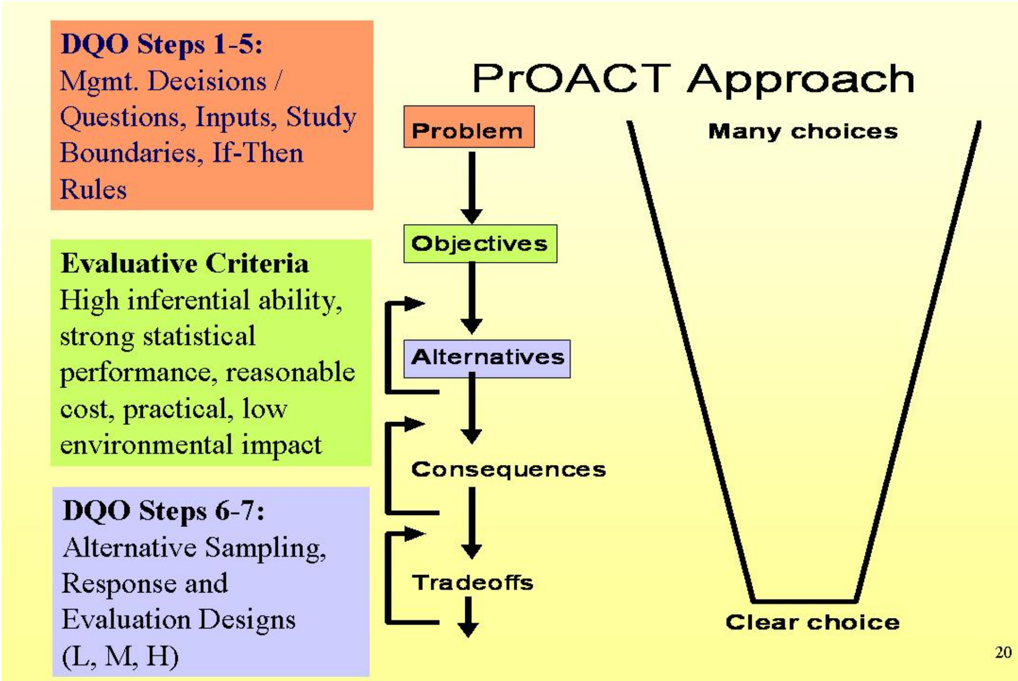

Figure 1.2. Flow of the PrOACT decision process recommended by CSMEP to narrow the range of acceptable M&E designs.

PrOACT is an iterative process that involves cycling over the development of alternatives, evaluating them, assessing tradeoffs, then starting again with better alternatives. One begins with a broad set of alternatives that gradually narrows to an acceptable choice or set of choices. Consultation with programmatic levels is critical throughout this process, so that the appropriate objectives and alternatives are considered (Table 1.1). CSMEP has begun to apply this approach as it moves to integrate designs from each domain into a holistic Columbia River basinwide M&E program that addresses multiple management questions.

Table 1.1. Examples of M&E design objectives and evaluative criteria.

CSMEP design objective	Potential evaluative criteria for design objective
High inferential ability	- Ability to answer questions at appropriate scale. - Ability to supply adequate information for clients' decisions. - Spatially representative of larger unit of interest. Ability to legitimately aggregate data required for decisions.
Strong Statistical Performance	- Precision (relative to required precision for management decisions). - Statistical power to detect various effect sizes of management importance over relevant time periods. - Coverage i.e., how often does the true value fall within the 95% confidence interval of the estimate. This depends on both bias and precision of the method used. - Bias (estimated by comparisons to very best measurement possible, close to census).
Reasonable Cost	- Cost/year at scale of interest. Cost for duration of M&E program. - Hybrids: Precision / cost, coverage/cost, accuracy/cost. - Ability to leverage other funding sources. Use overlapping domains of interest from different agencies.

1.3 CSMEP's Strategic Approach

Decisions on regional M&E designs need to be based on a quantitative evaluation of the costs and benefits of alternative designs, including Status Quo approaches. Alternative designs should build on the strengths of each subbasin's existing monitoring infrastructure and data, remedy some of the major weaknesses, and adapt to regional variations that affect monitoring protocols. Selected designs should improve the reliability of management decisions related to the status and trends of fish populations and should also improve evaluations of the effectiveness of habitat, harvest, hatchery and hydrosystem recovery actions within the Columbia River Basin.

CSMEP assembled detailed inventories[1] of fish population data for thirteen subbasins in Washington, Oregon and Idaho, and completed rigorous assessments of the strengths and weaknesses of these data for addressing high priority questions about salmon populations. These inventories were not intended to document all M&E actions everywhere – rather they were intended to evaluate the quality of information available by subsampling among the various subbasins. We have been exploring how best to integrate the most robust features of these existing monitoring programs with new approaches, and implementing the structured processes described in Section 1.2 to evaluate the costs, benefits and tradeoffs of different M&E designs.

Systematically developing and evaluating alternative M&E designs is complex. CSMEP, therefore, initially focused on spring/summer Chinook salmon in the Snake River Basin ESU, as a test case to refine design methods and analytical tools that will ultimately benefit the entire Columbia River Basin and Pacific Northwest (see Figure 1.3).

[1] CSMEP's metadata inventories are available at http://csmep.streamnet.org/ (CSMEP/CSMEP)

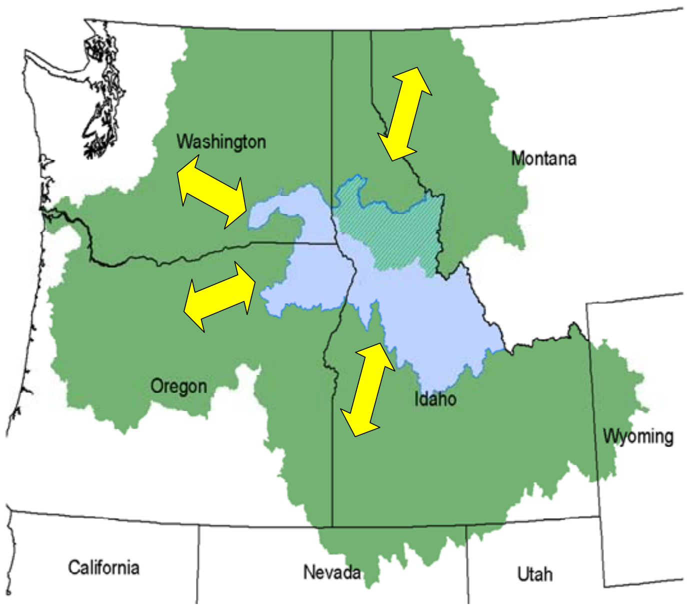

Figure 1.3. Insights gained from the CSMEP Snake River Basin Pilot study (blue shaded area) will have applications to other areas of the Columbia River Basin (CRB) and will similarly benefit from analyses being undertaken elsewhere in the CRB.

1.3.1 CSMEP's Snake River Basin Pilot

Salmon and steelhead occupying the Snake River Basin have declined precipitously to abundances warranting protection under the Endangered Species Act (ESA). The causes most commonly cited for these declines are grouped into four domains:

- *Habitat:* historical spawning areas have been isolated and degraded by human activities.
- *Hydropower:* the construction and operation of mainstem and tributary hydropower structures has altered population connectivity, altered life-history timing and increased mortality.
- *Harvest:* fisheries have exerted mortality on targeted and non-targeted stocks of anadromous, adfluvial, and resident species.
- *Hatcheries:* although intended to provide mitigation and/or conserve salmonid resources, hatcheries pose a multitude of potential risks to extant salmon and steelhead populations as well as other taxa of concern.

CSMEP chose the Snake River Basin as pilot study to develop M&E designs for the following reasons:

- In addition to salmon, there are ESA listed steelhead and bull trout populations, so it presents the challenge of integrating designs across multiple species.
- It has a broad diversity of current monitoring activities and has undergone a thorough CSMEP inventory of existing data, as well as detailed strengths and weaknesses assessments of these data for answering key questions.
- It provides an opportunity to explore an approach with Basin-wide applicability: 'hybrid' sampling designs that build on the existing strengths of monitoring data (e.g., long time series of index counts), but supplement current efforts with more representative sampling.

- It lies within the states of Idaho, Oregon and Washington and is an area of great interest to various client groups (e.g., NOAA, USFWS, NPT, CTUIR, SBT, IDFG, ODFW, WDFW, USFS, BLM, BoR, USACE)
- It is large enough to present many of the M&E challenges typical of the entire Columbia River Basin, including consideration of tradeoffs among monitoring objectives.
- There are hydro, hatchery, habitat and harvest actions requiring evaluation.
- It is one of the three pilot study areas (together with the John Day and Wenatchee subbasins) to be addressed by NOAA as part of their Integrated Status and Effectiveness Monitoring Program (ISEMP).
- The Snake River Basin forces CSMEP scientists to use relevant data from other regions, particularly for hydro, hatchery and harvest questions that are Columbia River Basin-scale in nature. For these domains CSMEP designs must, by necessity, extend beyond the bounds of the Snake River Basin.

For each of the five M&E domains illustrated in Figure 1.4 CSMEP biologists have developed quantitative tools and analyses to project the consequences and tradeoffs of alternative M&E designs in their Snake River Basin pilot, in terms of both the qualitative and quantitative evaluative criteria outlined in Table 1.1. For each domain an 'Objectives by Alternatives' matrix has been developed that provides managers a useful way to organize and assess the performance of each alternative design (i.e., Status Quo, 'Low', 'Medium', 'High') across a suite of critical objectives, and to identify trade-offs for making decisions on monitoring designs. These evaluations are described in Chapter 2.

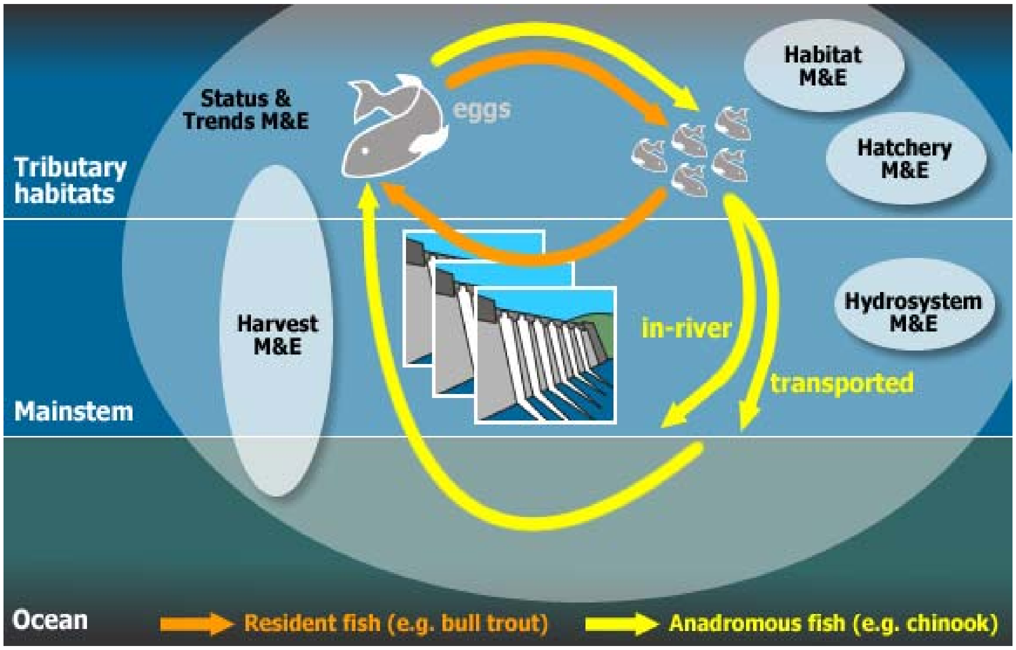

Figure 1.4. Anadromous and resident fish lifecycles and associated M&E domains. Status & Trends M&E (larger darker colored ellipse) encompasses the full range of habitats utilized within fish lifecycles and can be informed by the monitoring being undertaken within the other four M&E domains.

2. Specific Results

2.1 Status and Trends

The Interior Columbia–Technical Recovery Team (IC-TRT) has developed viability criteria for Interior Columbia Basin Salmonid ESUs. The viability assessment combines the four Viable Salmonid Population (VSP) performance measures that describe abundance, productivity, spatial structure and diversity at the population level to evaluate the status of the ESU (IC-TRT 2005). We use the IC-TRT viability criteria as a framework for assessing alternative monitoring strategies in the Snake River spring/summer Chinook salmon ESU.

2.1.1 Priority question

Using the IC-TRT viability criteria, are Snake River spring/summer Chinook salmon viable?

Related Decision: Has there been sufficient improvement in the population status of Snake River spring/summer Chinook salmon to meet the biological de-listing criteria (abundance, productivity spatial structure and diversity)? The biological de-listing criteria combined with the administrative de-listing criteria are conditions that must be met to allow removal of ESA restrictions (NMFS 2000).

2.1.2 What are the consequences of making the wrong decision?

Incorrectly concluding that the delisting criteria have been achieved:

- decisions to relax ESA restrictions increase the risk of extinction; and
- socio-economic consequences of stock collapse.

Incorrectly concluding that the delisting criteria have not been achieved:

- minimal biological impact given that decisions do not relax ESA restrictions;
- unnecessary listing and restrictive measures; and
- loss of harvest opportunity.

2.1.3 Monitoring design alternatives and trade-off analyses

We used a model to test the ability of monitoring programs to correctly assess spring/summer Chinook salmon population viability in the Snake River ESU using a simulated spawner abundance dataset. We assessed the monitoring currently being done in the basin (Status Quo), a ‘Low” design that relies on M&E methods that are less precise than used in the Status Quo design, a ‘Medium’ design that strengthens some of shortcomings of the Status Quo design, and a ‘High’ design that incorporates more precise M&E methods in all populations. The model inputs were based on the best available information on the precision and bias of monitoring methods used in the designs. The simulation results are summarized as the probability of making a correct viability assessment. Table 2.1.1 summarizes the monitoring designs and Table 2.1.2 summarizes the trade-off analyses of each design for assessing the viability of the Snake River spring/summer Chinook salmon ESU.

2.1.4 Design alternatives

The ability to correctly evaluate viability using the IC-TRT criteria depends on the accuracy and precision of the data needed to assess the VSP parameters. Our ‘Low’, ‘Medium’, and ‘High’ designs were constructed to evaluate the viability of the Snake River ESU. They were not constructed to answer any other management decision. The Status Quo design was an assemblage of all monitoring being done annually in the Snake River ESU, for any reason, that could be used in a viability assessment.

The Status Quo monitoring design has good quality information in some populations for some of the VSP criteria and very poor quality information in others. Populations with effective weirs have good abundance and diversity data, but may not assess spatial structure. Many populations use index counts to estimate abundance hence, there is no estimate of bias or precision. Index redd counts in populations with more than one Major Spawning Area (MaSA) or Minor Spawning Area (MiSA) usually do not assess spatial structure.

The ‘High’ design collects abundance and life-history diversity data (age structure, length, sex ratio, proportion natural origin) for all 32 populations using weirs. In five populations where weirs would likely capture < 40% of the spawners (due to location or size of the river), multi-pass index redd counts supplement the abundance and diversity estimates. The spatial structure of each population was obtained from a single census redd survey through out the entire spawning area. This design collects the most precise and accurate data from all populations.

The ‘Medium’ design uses only five weirs, but ensures that each MPG had a weir. The reduction in weirs increases the uncertainty of the age-structure, proportion natural origin, and other life-history diversity statistics at the population level since life-history data collected at each weir will be assumed to represent all of the populations within the MPG. Abundance in the remaining populations is estimated using multi-pass redd counts in index areas plus a one-time census redd count. The single pass spatial census redd count provides a ratio of redds within and outside of the index sites, improving the estimate of abundance as well as providing spatial structure information for each population.

The ‘Low’ design has no weirs and abundance estimates are based on a single redd count in index areas expanded to the entire population using IC-TRT assumptions. The population abundance estimates have the highest uncertainty in this design. The limited field sampling provides no estimates of spatial structure in populations with more than one MaSA or MiSA, and the number of carcasses recovered may not be representative of the population life-history diversity parameters.

2.1.5 Tradeoff analyses

A correct viability assessment was made 60% of the time with the Status Quo M&E. There was an improvement in the percent of correct decisions from the Status Quo using the ‘Medium’ (73% correct) and high (84% correct) designs. The ‘Low’ design correctly assessed the viability 41% of the time. A larger proportion of correct viability assessments of the ESU were made using the ‘Medium’ design than the Status Quo and at a lower cost. The ‘High’ design correctly assessed the viability 84% of the time but it was nearly 3 times the cost of the ‘Medium’ design and 1.7 times the cost of the Status Quo design.

Table 2.1.1. Description of four monitoring design alternatives and how they differ for each performance measure.

Performance Measures Required	Description of Monitoring Design Alternatives			
	Status Quo	Low	Medium	High
Abundance of Fish	Weir with Mark-Recapture (MR) in 13 populations, weir count only in one population.	No weirs (however there are hatchery weirs in 12 populations that will be operating).	Weir with MR in one population for each of 5 MPGs. (an additional 8 populations have a hatchery weir that will be operating)	Weir with MR in all 32 populations.
Abundance / Spatial Distribution of Redds	Single pass aerial index redd counts in 15 populations. Single pass ground index redd counts in 5 populations. Multi pass ground census redd counts in 8 populations. Single pass census redd count in 2 populations. No redd counts in 2 populations.	Fixed single redd counts for all 32 populations, using index sites. 26 aerial & 6 ground (2 wilderness, 4 road access)	Multi-pass (3x) index redd sites in all populations. Includes 18 aerial and 14 ground counts with a one-time census of the entire spawning area of the population to address spatial structure (6 ground and 27 aerial census surveys). The one time pass provides a ratio of redds within and outside of the index sites, improving the estimate of abundance as well.	Multi-pass redd counts in 5 populations where the weir captures < 40% of spawners in the population (two raft surveys and 3 ground surveys). A one time census survey of the entire spawning area of each population will be done to assess spatial structure (6 ground and 26 aerial census surveys).
Age Structure of Spawners (for the initial run, we are using a fixed age-structure for the simulated data)	Scale analyses in 13 populations with a weir and 10 populations having multi-pass redd counts (9 populations done by the ISS study that are not considered Status Quo redd counts for abundance estimates).	Representative samples taken at Lower Granite Dam provide a single estimate for age structure for all populations in the ESU.	Age structure estimated in 5 populations (one population in each MPG) from adults sampled at the weir. In addition, age structure estimated in 14 other populations surveyed with ground redd counts. Age-structure data collected at each weir will be assumed to represent all of the populations within the MPG.	Age structure estimated in all 32 populations from adults sampled at weirs and during ground redd counts where this occurs. Each population will have a unique age-structure estimate.
Origin of Spawners (for the initial simulation we are assuming we know the origin of spawners)	Examine hatchery marks on carcasses or at weirs in 21 populations (plus an additional 5 populations surveyed by ISS); detect pit-tags at each weir	Examine hatchery marks on carcasses in 6 populations.	Examine fish for hatchery marks at weir for 5 populations; examine carcasses during all ground redd counts (14 populations).	Examine fish for hatchery marks at weirs and during ground and raft redd counts where they occur.
Sex Ratio of Spawners (We are not considering this parameter explicitly-next round)	Carcass survey or handle at weir in 21 populations (5 additional populations are surveyed by ISS).	Samples taken at Lower Granite Dam for entire ESU. Single estimate for sex ratio for all populations in ESU.	Examine fish at weir in 5 populations; examine carcasses in the 14 populations surveyed with ground redd counts.	Examine fish at weirs and during ground and raft redd counts where they occur

ISS = Idaho Supplementation Study. This is a BPA funded Chinook supplementation research project being done in Idaho. It began in 1992 and is funded at least until December 31, 2009 (funded for the BPA FY07-09 proposal period).

Table 2.1.2. Trade-off analyses of each design for assessing the viability of the Snake River spring/summer Chinook salmon ESU.

Performance Measures	Evaluation of Monitoring Design Alternatives Qualitative evaluations (Q): 5 = excellent; 4 = very good; 3= good; 2= fair; 1=poor; ?= Unknown; n.a. not applicable.			
	Status Quo	Low	Med	High
Ability to make viability assessments for each population in an ESU (Q)	(2) Spatial structure insufficient in 10 populations, incomplete in 1 population, complete in 21 populations. Good estimates of abundance and age structure in 10 populations. Potentially biased abundance and age estimates in the rest.	(1) Spatial structure insufficient in 17 populations, complete in 15 populations. Potentially biased abundance estimates in all populations with no estimate of precision. Age structure estimated for the entire ESU.	(4) Complete spatial structure and unbiased estimates of abundance in all populations. Each MPG has its own age structure determined from the adults sampled at the five weirs.	(5) As close to census counts as possible for abundance (weirs with multiple redd counts in populations where the weir captures < 40% of the spawners). One-time census redd count also provides abundance and complete spatial structure information for each population. Population specific estimates for: age structure, origin of spawners, and sex ratio.
Ability to estimate long term trends, continue time series (Q)	(3) Can continue current time series; however don't have estimates of bias & precision in most cases, so minimal ability to correctly detect trends among populations.	(1) Would only be able to continue time series of redd counts. Don't have estimates of bias and precision so poor ability to detect trends among populations.	(3) Can be done, but will require extra expense in the short term. Need to implement both the current Status Quo monitoring and the new proposed design for a few years in order to determine how well the Status Quo abundance estimates correlate with those derived from this design.	(3) Can be done, but will require extra expense in the short term. Need to implement both the current Status Quo monitoring and the new proposed design for a few years in order to determine how well the Status Quo abundance estimates correlate with those derived from this design.
Ability to aggregate status and trend data to multiple scales for regional scale, high level assessments (Q)	(2) Possible (e.g., Status of Resource Report) but no precision estimates for MPG's in Idaho. Estimates for the Lower Snake and Grande Ronde/Imnaha MPGs can be made with precision estimates.	(3) This design provides poor viability assessments at the population scale. For any scale greater than the population scale, it would be easy to aggregate the results; however the results would not be expected to be as accurate as in the M & H designs).	(3) This design provides fair viability assessments at the population scale. At the MPG level it should be quite good as there are some MPG specific measurements. Should be able to aggregate the results to MPG or ESU scales.	(4) This design provides good viability assessments at the population scale. For any scale greater than the population scale, it would be easy to aggregate the results.
Costs for Status and Trends	$1,282,497	$175,197	$709,900	$2,124,715
Correct viability assessment	59.5%	40.9%	72.9%	84.1%
Underestimate viability	32.7%	54.5%	17.5%	10.1%
Overestimate viability	7.8%	4.6%	9.6%	5.8%

The intent of the IC-TRT was to ensure a precautionary approach when making viability assessments. Our results confirm that incorrect decisions tend to be conservative for all four designs and if the data is poor the tendency to underestimate viability increases. When an incorrect viability assessment of the ESU was made, the error was usually caused by underestimating the viability. For example, in the 'Low' design the viability decisions[2] were 41% correct, 55% underestimated, and 5% overestimated. In the 'High' design where more precise methods were used to collect better quality data the viability decisions were 84% correct, 10% underestimated, and 6% overestimated. The same trend in the percent of correct, underestimated, and overestimated viability assessments was observed in the Status Quo and 'Medium' design results (Table 2.1.2).

2.1.6 Conclusions and recommendations

The Status Quo Snake Basin Spring Summer Chinook monitoring design contains weaknesses for assessing viability at the population level as per the IC-TRT viability criteria. The current monitoring does not assess spatial structure information (not all MaSA and MiSA are surveyed) in 11 populations. It lacks an abundance estimate in the non-index areas for populations without weirs or spatially representative redd counts (22 populations) preventing the assessment of bias inherent in index counts. The Middle Fork Salmon MPG lacks a weir, but all other MPGs have at least one weir providing life-history data (also referred to as diversity) such as sex ratio, percent female, percent natural origin, length, age, tissue samples for genetics in addition to abundance information.

The cost of the 'Medium' design is significantly less than the Status Quo, yet performs better to answer the question: is the ESU viable? Although the 'Medium' design cost less than Status Quo, the Status Quo design is a consortium of weirs, redd counts, and other monitoring that is being done for many different purposes. The major difference in cost between the Status Quo and the 'Medium' design is the number of weirs (14 vs. 5). Although, it may not be necessary to have 14 weirs to answer this one question, these weirs can be used to answer other management questions. Most of the weirs in the Status Quo design are associated with hatchery programs and will operate yearly. If the hatchery weirs were included in the 'Medium' design we would expect to see a higher percentage of correct viability assessments (somewhere between the 'Medium' and the 'High' design). A reallocation of resources in the Status Quo design could address its weaknesses and improve the viability assessments. This would require: (1) changing the redd survey program to the 'Medium' design where all populations have multiple redd counts and spatial structure assessed and (2) installing a weir in the Middle Fork Salmon River MPG. Index surveys by nature are not representative samples and so estimates based on these surveys are likely to be biased and unable to provide spatial structure information. However, these weaknesses can be addressed by supplementing the index surveys with some form of spatially representative sampling.

The IC-TRT rule set is conservative, so high uncertainty generally results in underestimating viability. Our results confirm that the most likely error was finding a population not viable when the population was in fact viable. This result, in addition to the cost and the consequences of incorrect viability assessments, should be considered when evaluating the tradeoffs among designs. While a lower cost design may save money in the short term, if the resulting data is of lower quality then there is the possibility of incurring higher costs over the long term due to the inability to make a correct assessment of the ESU. The simulation model we developed is an example of a tool managers can use to assess current monitoring programs and evaluate alternative monitoring strategies.

[2] Percentages may not add to 100% due to rounding error.

2.2 Harvest

2.2.1 Priority questions

1. What are the in-season estimates of run size and escapement for each stock management group (target and non-target) and how do they compare to preseason estimates?
2. What is the target and non-target harvest and when is it projected to meet allowable levels?
    - Species and Stock Groups of interest: wild and hatchery Upper Columbia spring Chinook and Snake River spring/summer Chinook.
    - Spatial Scales of interest: fisheries in the mainstem and major tributaries.
    - Time Scales of interest: Seasonal (January – June 15), Annual, Multi-Year.

Snake River spring/summer Chinook are harvested in the lower mainstem Columbia River (along with other listed Chinook from the upper Columbia and mid-Columbia) so it is important to have an accurate pre-season estimate (to structure the fisheries), and reliable determinations of the number of Chinook harvested and the stock composition so as to stay within ESA guidelines and fisheries quotas, and ensure that upriver hatcheries meet their mitigation goals. These questions focus attention on the key issues of identifying the number of fish that are impacted by fisheries while working toward recovery of the stocks, and improving how managers project when that number is achieved.

2.2.2 Related decisions

An important consideration in managing fisheries is the timing of harvest of stocks of concern. Fisheries are not only managed for total catch, but for duration of season, which directly controls total catch. Managers must therefore project what a fishery will catch over some time period. Stock composition of catch is a second critical component in projecting the impacts of a fishery. Most often, managers utilize the most recent available stock composition information to project expected composition of upcoming fisheries. This may be replaced by pre-season composition estimates if projections for early-season fisheries are needed and tag recovery information is not yet available. Managers may also adjust expected composition based on historic information regarding the run timing of key stocks.

2.2.3 What are the consequences of making a wrong decision?

If harvest is overestimated, fisheries may be constrained in order to remain below the impact guideline, creating lost opportunities. If harvest is underestimated, fisheries may be allowed to exceed the impact guideline, delaying recovery of the stocks.

2.2.4 Monitoring design alternatives and trade-off analyses

Abundance of these stocks are calculated by adding Bonneville Dam fishway counts to losses from lower river fisheries. If either the fishway counts or estimates of losses from lower river fisheries are inaccurate, estimates of run size will be incorrect. Estimates for number of fish released by anglers in the lower river are derived from creel interviews. Because released fish cannot be examined by surveyors, stock identification of released fish is not possible by direct examination, and estimates of the number of fish released are dependent upon the angler’s memory/honesty. Estimated harvest of spring Chinook in lower river commercial fisheries is derived from landing tickets submitted to ODFW and WDFW. Fish are sampled by agency staffs at buying stations to collect biological and mark sample data. Average weights per species are applied to the total reported pounds landed from all landing tickets to estimate the total number of fish landed.

Once fish pass Bonneville Dam, they are known to be upriver spring Chinook. However, fish encountered in the lower river may be from upriver or lower river populations. Stock identification (upper or lower river) of kept fish is derived using Visual Stock Identification (VSI). CWTs are applied to many lower river Chinook stocks and a few upriver stocks, but are not applied to upriver fish at rates high enough to provide a robust stock identification by CWTs alone. Released fish are not examined by surveyors, and because of differences in the percentage of marked fish between upriver and lower river stocks, stock identification of released fish is assumed to not be equal to the proportions of upriver versus lower river fish in the kept catch. Managers must use preseason expectations of abundance of upper and lower river spring Chinook, combined with the expected marking rates for each group, to estimate the composition of released fish. Fisheries above Bonneville Dam encounter only upriver spring Chinook, and tribal commercial fisheries retain all fish caught, meaning that stock identification of released fish is not necessary for either of these fisheries. Commercial fisheries below Bonneville Dam are examined during fishing seasons by onboard observers to estimate the number of upriver spring Chinook released from these fisheries. The ratio of unmarked to marked Chinook for all observations from a fishery is multiplied by the total number of marked fish landed to estimate the total number of Chinook released.

Although PIT tags are widely used in the Columbia Basin, their use in monitoring fisheries is limited, primarily due to the large number of fish that would need to be tagged for sufficient recoveries in fisheries. Genetic stock identification (GSI) could offer an alternative.

Individual fisheries may be managed on harvest rates as low as 0.01%, so small changes in estimates of run size can have a large effect. Higher-than-expected catches in lower river fisheries also can force closures of upriver fisheries in order to maintain total impacts below ESA limits. Estimates of precision are not provided with these projections. Adding estimates of precision, and if possible, indications of directional biases, would aid managers in determining how much weight to put on individual estimates, and in weighing the likelihood of over- or under-estimating run size. Additionally, new methods for projecting run size, such as relationships to environmental variables, may be available to help improve forecasting accuracy.

In Status Quo monitoring, in-season post-release mortality rates are not monitored. Instead standard rates from previous studies are applied. Conducting long-term, fishery-specific mortality studies is inherently difficult and expensive. Double Index Tagging (DIT) is a method that has been proposed for use in assessing mortality of fish stocks. DIT release groups are most useful if they are representative of unmarked wild fish that co-occur within fisheries.

2.2.5 Conclusions/recommendations

Status Quo harvest monitoring generally does not provide precision estimates; however, such estimates would be useful for managers in allowing them to quantify the risk of available harvest management decision options.

- Uncertainty or errors in harvest impact estimates can also limit evaluation of status, trends, and viability.
- Uncertainty or errors in harvest impact estimates can result in lost harvest opportunities or over harvest of listed stocks.

- Include estimates of precision in vital estimates.
- Develop new analytical techniques for preseason and in-season abundance forecasts.
- Continue to evaluate new technologies/techniques for stock identification and composition estimates (PIT tags, GSI).
- Evaluate and refine methods for estimating number of fish released from selective fisheries.
- Evaluate the potential development of an indicator stock to represent Snake River spring/summer Chinook in in-river fisheries.
- Improve coordination between entities collecting fisheries monitoring and evaluation information.

Table 2.2.1. Description and evaluation of harvest monitoring design alternatives.

Description of Design Alternatives

Questions: 1) What are the in-season estimates of run size and escapement for each stock management group (target and non-target) and how do they compare to preseason estimates?
2) What is the target and non-target harvest and when is it projected to meet allowable levels?

Performance Measures	Status Quo	Low	Medium	High
Pre-season forecast of adult abundance at the Columbia River mouth				
	Cohort regression and expert opinion	Cohort regression and expert opinion	Cohort regression and expert opinion	Cohort regression and expert opinion
		Incorporate precision estimates in addition to point estimates.	Estimate precision	Estimate precision
				Apply new methods, and/or incorporate additional data (e.g., autocorrelation and/or incorporating environmental data).
Stock Composition of the Catch				
<u>Lower Columbia River Commercial Fishery</u>				
	Visual Stock Identification (VSI)	VSI	Estimate precision	Estimate precision
		Estimate precision	PIT-tag sampling of kept and released catch from a PIT-tagged wild fish population large enough ensure adequate recovery information (10 recoveries/tag group/year) ESU-level resolution. [186K juveniles/year= hydro 'High' design]	GSI sampling of released catch sufficient to describe MPG-level stock composition.
		PIT-tag sampling of kept and released catch under current tagging programs. (86K juveniles/year at LGR = hydro 'Medium' design)	Development of CWT-indicator stock(s) to represent wild Sp/Su Snake River Chinook ESU(s).	
			Genetic Stock Identification – sampling of released catch to describe ESU-level stock composition.	

Performance Measures	Status Quo	Low	Medium	High
Mainstem Sport Fishery				
	Inseason stock comp from preseason estimates and mark rates; verified post season.	Inseason stock comp from preseason estimates and mark rates; verified post season.	Incorporate precision estimates in addition to point estimates.	Incorporate precision estimates in addition to point estimates.
		Incorporate precision estimates in addition to point estimates.	GSI ESU-level stock composition from commercial fishery	GSI MPG-level stock composition from commercial fishery.
			Development of CWT indicator stock to represent wild Sp/Su Snake River Chinook.	
Zone 6 Tribal Fishery				
	Dam counts inform stock composition and Unmarked:Marked fish ratio? Post season run reconstruction eventually addresses this through CWT-tag recoveries.	Dam counts inform stock composition and Unmarked:Marked fish ratio? Post season run reconstruction eventually addresses this through CWT-tag recoveries.	Dam counts inform stock composition and Unmarked:Marked fish ratio? Post season run reconstruction eventually addresses this through CWT-tag recoveries.	Dam counts inform stock composition and Unmarked:Marked fish ratio? Post season run reconstruction eventually addresses this through CWT-tag recoveries.
		Incorporate precision estimates in addition to point estimates.	Incorporate precision estimates in addition to point estimates.	Incorporate precision estimates in addition to point estimates.
		PIT-tag sampling of harvested catch under current tagging programs. (86K juveniles/year at LGR = hydro 'Medium' design)	PIT-tag sampling of harvested catch from a PIT-tagged wild fish population large enough ensure adequate recovery information (10 recoveries/tag group/year) ESU-level resolution. [186K juveniles/year= hydro 'High' design]	GSI sampling of released catch sufficient to describe MPG-level stock composition.
			Development of CWT-indicator stock(s) to represent wild Sp/Su Snake River Chinook ESU(s).	
			Genetic Stock Identification – sampling of released catch to describe ESU-level stock composition.	

<table>
<tr><th>Evaluation Criteria</th><th colspan="4">Qualitative Evaluations: 5=excellent, 4=very good, 3=good, 2=fair, 1=poor</th></tr>
<tr><th></th><th>Status Quo</th><th>Low</th><th>Medium</th><th>High</th></tr>
<tr><td colspan="5">Ability to estimate stock specific run size pre-season and in-season (pre-season / in-season)</td></tr>
<tr><td>LCR Commercial Fishery</td><td colspan="3" rowspan="2">(2-3)</td><td rowspan="3">(3-4)</td></tr>
<tr><td>Mainstem Sport Fishery</td></tr>
<tr><td>Zone 6 Tribal Fishery</td><td colspan="3">(3-4)</td></tr>
<tr><td colspan="5">Ability to estimate stock specific escapement</td></tr>
<tr><td>LCR Commercial Fishery</td><td rowspan="3">(2)
Dam counts and hatchery returns inform escapement. Stock ID is up- or downstream not ESU</td><td colspan="2">(3)
ESU ID</td><td>(4)
MPG ID</td></tr>
<tr><td>Mainstem Sport Fishery</td><td colspan="3">(3)
Interview bias?</td></tr>
<tr><td>Zone 6 Tribal Fishery</td><td colspan="3">(3)
Between dam conversion losses uncertain.</td></tr>
<tr><td colspan="5">Ability to estimate target harvest</td></tr>
<tr><td>LCR Commercial Fishery</td><td rowspan="3">(3)</td><td rowspan="3">(3)
ESU ID</td><td colspan="2" rowspan="3">(4)
MPG ID</td></tr>
<tr><td>Mainstem Sport Fishery</td></tr>
<tr><td>Zone 6 Tribal Fishery</td></tr>
<tr><td colspan="5">Ability to estimate non-target harvest</td></tr>
<tr><td>LCR Commercial Fishery</td><td>(2)
ID up- or downstream not ESU</td><td>(3)
ESU ID</td><td colspan="2">(4)
MPG ID</td></tr>
<tr><td>Mainstem Sport Fishery</td><td>(2)
Based on commercial fishery and angler creel.</td><td colspan="3">(3)
Interview bias?</td></tr>
<tr><td>Zone 6 Tribal Fishery</td><td>(2)
ID up- or downstream not ESU</td><td colspan="3">(3)
Bio-sampling rate of direct sales uncertain.</td></tr>
<tr><td colspan="5">Ability to project when harvest impact will meet allowable levels</td></tr>
<tr><td>LCR Commercial Fishery</td><td rowspan="3">(2)</td><td colspan="3">(4)</td></tr>
<tr><td>Mainstem Sport Fishery</td><td colspan="3">(3)</td></tr>
<tr><td>Zone 6 Tribal Fishery</td><td colspan="3">(3-4)</td></tr>
</table>

2.3 Hydro

2.3.1 Priority question

Are mainstem survival rates (Lower Granite to Bonneville, LGR to BONN), Smolt-to-Adult Return Rates (SARs; LGR to LGR), and important SAR comparisons relating to the effectiveness of transportation and overall hydrosystem operations, meeting targets set by the Northwest Power and Conservation Council (NPCC)[3], and the Biological Opinion on operation of the Federal Columbia River Power System (FCRPS BiOp)[4]?

- Species and Stock Groups of interest: wild and hatchery Snake River spring/summer Chinook; wild and hatchery steelhead.
- Spatial Scales of interest: Snake River aggregate; Major Population Groups (MPGs) in Snake Basin, downstream stocks for contrast.
- Time Scales of interest: Seasonal, Annual, Multi-Year.

This question is of interest for three reasons: 1) evaluation of the effectiveness of hydrosystem operations in meeting survival goals; 2) understanding the extent to which mainstem, estuary and ocean life history stages are limiting the recovery of different MPGs (status and trend question); and 3) for understanding the effectiveness of hatchery operations (hatchery action effectiveness).[5] The existence and operation of the Federal Columbia River Power System (FCRPS) is one of the more important factors influencing mainstem survival of three ESUs of concern to this Snake River [SR] pilot study: SR spring/summer Chinook, SR fall Chinook, and SR steelhead. ESA-listed bull trout are also effected. This pilot study focuses on spring/summer Chinook, with some steelhead results. There is a need to assess what quality of data are required to: 1) reliably detect the effects of FCRPS actions on fish survival rates; and 2) reliably compare survival rates to pre-defined goals.

2.3.2 Related Decision

Decisions on FCRPS actions directly or indirectly affecting survival of these stocks are conducted to meet the requirements of the ESA to minimize take and contribute towards recovery of listed fish. These actions include juvenile collection, bypass, and transportation; water management; and offsite mitigation. Information on the expected and actual effectiveness of these actions is essential for reliable decisions on how to manage the hydrosystem seasonally (e.g., should spring-summer Chinook be transported earlier in the season?), annually (e.g., how should water management and transportation strategies change in wet vs. average vs. dry years?), and over multiple years (e.g., is the system configuration and operation providing sufficient survival to support stock recovery?).

---

3 Pg. 13 of NPCC mainstem amendments of 2003-2004. www.nwcouncil.org/library/2003/2003-11.pdf has an interim objective of 2-6% SAR for listed Snake River and Upper Columbia chinook and steelhead (minimum 2%, average 4%).

4 For example, the 2000 FCRPS BiOp set a standard of 49.6% survival for Snake River spring/summer Chinook from LGR to BONN (Table 9.2-3, pg. 9-14, NMFS 2000).

5 For example, the McCall hatchery has SARs that are about three times those of the Dworshak hatchery, suggesting differences in the health, size, timing or other attributes of released fish (Fig 3.7, CSS Draft 10-year report, 2007).

2.3.3 Consequences of wrong decisions

Making hydrosystem decisions that harm fish could significantly reduce the chances that ESA listed wild stocks will persist and recover, and could significantly harm other non-listed stocks. Making the wrong decisions on limiting factors (due to inadequate M&E) could lead to cost-ineffective mitigation strategies. Wrong decisions that reduced power generation but had little or no survival benefit would have unnecessary economic impacts, and potentially other environmental impacts, depending on how the lost power generation was replaced.

2.3.4 Monitoring design alternatives and trade-off analyses

The main information used as input to these decisions are analyses of PIT-tag[6] recoveries, largely through the Comparative Survival Study (CSS 2007). The Status Quo monitoring is intended to acquire estimates of mainstem in-river survival, SARs, TIRs (Transport to Inriver Ratios[7]) for the Snake River aggregate (all spring-summer Chinook that spawn and rear above LGR) as a whole (generally using wild plus hatchery fish). We examined existing data to assess the ability of managers to make reliable decisions on whether or not different groups of fish were meeting established survival targets, in different years and for different stock groups. We considered weaknesses in the Status Quo M&E identified by agency fish managers and scientists (e.g., much less certainty for wild spring-summer Chinook than for hatchery Chinook evaluations, samples that are not representative of the run at large, much less certainty for MPGs than for Snake Basin aggregate). We then developed M&E alternatives to overcome these weaknesses (e.g., increase the number of PIT-tagged wild fish, distribute tags in proportion to hatchery releases, apply multi-year average SARs and SAR ratios for MPG-level evaluations). While there are many strengths to Status Quo monitoring, it has some weaknesses that could lead to erroneous decisions on hydrosystem action effectiveness, particularly for wild spring-summer Chinook and MPG scales, where current sample sizes are too low to ensure high precision, reliable inferences. (The same is true for steelhead, but this report is focused on spring-summer Chinook.) In addition to investigating improvements to the Status Quo, we also explored the implications of reducing M&E investments (i.e., a 'Low' Alternative without the CSS hatchery fish). Table 2.3.1 describes and evaluates these alternatives.

The ability to reliably assess compliance with a survival rate target depends partly on how close survival rates are to that target, as well as on the M&E methods. For example, Status Quo monitoring shows that SARs of wild spring/summer Chinook during 1994-2004 have generally been well below the 2-6% target recommended by the Northwest Power and Conservation Council (NPCC). While this target is primarily for listed (i.e., wild) populations, we can also examine the performance of hatcheries against this same SAR goal. There are four hatcheries with 8 years of SAR / TIR data, and one hatchery with 4 years of SAR data, for a total of 36 years of data. SARs of hatchery[8] Chinook were definitively below the 2% level in 28 of these 36 data-years, and definitively above 2% in only 6 of these 36 data-years (i.e., definitive evaluations were possible in 34 of 36 data-years). By contrast, definitive evaluations[9] with the SAR target (i.e., compliance or not) can be determined from Status Quo monitoring in only 6 of 10 years for wild spring/summer Chinook (Table 2.3.1). TIRs have often been close to 1 for wild spring/summer

6 Passive Induced Transponder (PIT) tags uniquely identify individual fish. They're inserted into salmon parr or smolts, which are subsequently detected at weirs, dam bypasses, or harvest. As discussed in section 2.6, PIT-tagged fish can be used for multiple purposes (status and trend, hydro, hatchery, habitat and harvest M&E).

7 TIRs are the ratio of the SAR of transported fish to the SAR of in-river fish. A TIR greater than 1 means that transported fish have survived better than in-river fish; a TIR less than 1 means the opposite.

8 The five hatcheries are Dworshak, Rapid River, Catherine Creek, McCall, and Imnaha River.

9 "Definitive evaluations" are considered to occur when the 90% confidence interval for the estimate does not overlap the target. This is used as an example decision criterion. See bottom of Table 2.3.1 for quantitative metrics.

Chinook and for three of the five hatcheries. Because TIR estimates are less precise than SARs (i.e., wider confidence intervals), definitive evaluations of whether transported fish survived better than in-river fish can be made in only 3 of 10 years for wild spring/summer Chinook, and in only 20 of 36 years of TIR data for the five hatcheries.

Our cost estimates for the Status Quo, 'Medium' and 'High' M&E alternatives include the costs of long-running foundational projects (i.e., Smolt Monitoring Program – NPCC project # 198712700, PTAGIS - #199008000, UW Statistical Support - # 198910700, Passage Survival Estimates - #199302900, CSS - #199602000). The variable costs of PIT-tagging across alternatives are estimated assuming $2.10 for the tag itself, $1.16 labor cost per tagged hatchery fish, and $12.36 labor cost per tagged wild fish.

Table 2.3.1. Description and evaluation of hydro monitoring design alternatives (2 page table).

Description of Design Alternatives

Question: Are mainstem survival rates, SARs and important SAR ratios relating to the effectiveness of transportation and overall hydrosystem operations, meeting NPCC and BiOp targets?

Performance Measures	Status Quo	Low	Medium	High
SARs, TIRs, mainstem survival	*SR Hatchery Chinook:* • # tags= 255,000 *\SR Wild Chinook:* • # tags=66,000 (29 stream RSTs) *Lower and Mid-Col R Hatchery Chinook:* • # tags=70,000 *Lower and Mid-Col R Wild Chinook:* • 6,000 PIT-tags @ John Day River	Background level of PIT-tagging. *SR Hatchery Chinook:* • # tags=40,000 *SR Wild Chinook:* • Same as Status Quo • # tags=66,000 (29 stream RSTs) *Lower and Mid-Col R Hatchery Chinook:* • Same as Status Quo but drop Carson • # tags=55,000	*SR Hatchery Chinook:* • Distribute tags in proportion to hatchery releases across all SR hatcheries; distribute fish (i.e.,% transported) according to run at large • # tags=275,000 *SR Wild Chinook:* • # tags=86,000 (40 stream RSTs) *Lower and Mid-Col R Hatchery Chinook:* • Same as Status Quo • # tags=70,000 *Lower and Mid-Col R Wild Chinook:* • # tags=6,000	*SR Hatchery Chinook:* • Distribute tags proportionately as for 'Medium'; increase # • # tags=375,000 *SR Wild Chinook:* • # tags=186,000 (29 stream RSTs + 8 large traps to cover 6 MPG strata, incl. Clearwater; <u>not</u> by population) *Lower and Mid-Col R Hatchery Chinook:* • # tags=100,000 *Lower and Mid-Col R Wild Chinook:* • # tags=6,000
Abundance	*Snake Basin:* as described for Status Quo alternative under Status and Trend (section 2.1)	*Snake Basin:* as described for Low alternative under Status and Trend (section 2.1); SARs estimated from run reconstructions. *Downstream stocks:* John Day redd counts to provide contrast.	*Snake Basin:* as described for 'Medium' alternative under Status and Trend (section 2.1). *Downstream stocks:* one population / regional stock group in Lower & Mid Columbia (John Day, Deschutes/Warm Springs, Yakima, Wind, Klickitat).	*Snake Basin:* as described for 'High' alternative under Status and Trend (section 2.1). *Downstream stocks:* Possibly weirs John Day, Wind and Klickitat (not essential if 'High' level PIT-tagging is implemented, which is more precise).

Evaluation of Monitoring Design Alternatives for Snake River spring-summer Chinook

Qualitative evaluations (**Q**): **5** = excellent; **4** = very good; **3**= good; **2**= fair; **1**=poor; **?=** Unknown; **n.a.** not applicable.

Evaluation Criteria	Status Quo	Low	Med	High
Ability to assess compliance with SAR targets (e.g., NPCC target of 2-6%)	(3)-(4) Very good (4) for hatchery fish (can assess compliance in nearly all years); good (3) for wild fish alone (can currently assess compliance in most years).[10]	(1)-(3) Poor (1) for hatchery fish because no longer representative of run at large. Good (3) for wild fish.	(3)-(4) Similar to Status Quo, but more representative of run at large	(4) Very good (4) for both hatchery and wild fish. Able to assess SARs at MPG level.
Ability to assess compliance with BiOp in-river survival targets	(1)-(3) Good (3) for hatchery + wild fish combined; poor (1) for wild fish alone. But since in-river survival rates for wild and hatchery fish are similar this is OK.	(1)-(2) Fair (2) for hatchery + wild fish combined; poor (1) for wild fish alone.	(2)-(4) Very good (4) for hatchery + wild fish combined; fair (2) for wild fish alone.	(4)-(5) Excellent (5) for hatchery + wild fish combined; very good (4) for wild fish alone.
Ability to assess transportation effectiveness	(1)-(4) Poor (1) for wild fish on an annual basis (e.g., can only tell if TIR significantly different from 1 in 3/10 years); very good (4) for Rapid River and McCall hatcheries; fair for Imnaha (2); poor for Dworshak and Catherine hatcheries (1).	(1) Same as SQ for wild fish. Unable to do TIRs for hatchery fish.	(1)-(4) Similar to Status Quo, but more representative of run at large.	(3-4) Some improvement in precision of wild and hatchery TIRs (3). Able to assess TIRs at MPG level.
Ability to analyze upstream-downstream contrasts in survival	(2 -3) Current contrasts use ~40 yrs of Spawner-Recruit data (R/S) from index sites, with weaknesses described in Status and Trend section. SAR estimates for Snake Aggregate and John Day provide stronger inferences, but no other wild downstream stocks have SAR estimates.	(2-3) Differential mortality estimates updated from Spawner-Recruit data only (not SAR measurements) for index stocks. Much weaker inferences about differential mortality.	(3) Improved SAR precision on wild fish provide a better check on SR derived differential mortality estimates. SR data improved, but weaknesses of historical SR data remain. Uncertain how many wild fish can be tagged.	(3-4) More representative R/S monitoring (abundance and productivity) and tagging for SAR estimates; keep index sites to maintain time series (historical weaknesses remain). Improved statistical inferences for Snake and downriver aggregates.
Costs[11] (% change from Status Quo)	$8,548,334 (0%)	$7,848,426 (-8.2%)	$8,989,417 (+5.2%)	$11,050,222 (+29%)
Fraction of years for which definitive[12] evaluations can be made of compliance with SAR target[13]	Snake Basin level: **wild spring/summer:** 6/10 years **hatchery stocks:** Rapid River 6/8; McCall 8/8; Imnaha 8/8; Dworshak 8/8; Catherine 4/4 MPG level: *in prep*	Snake Basin level: **wild spring/summer:** 6/10 yrs **hatchery stocks:** Rapid River 6/8; McCall 7/8; Imnaha 7/8; Dworshak 8/8; Catherine 4/4 MPG level: *in prep*	Snake Basin level: **wild spring/summer:** 7/10 yrs **hatchery stocks:** Rapid River 6/8; McCall 8/8; Imnaha 8/8; Dworshak 8/8; Catherine 4/4 MPG level: *in prep*	Snake Basin level: **wild spring/summer:** 8/10 yrs **hatchery stocks:** Rapid River 6/8; McCall 8/8; Imnaha 8/8; Dworshak 8/8; Catherine 4/4 MPG level: *in prep*

[10] These evaluations are based on 1994-2004 data, with SARs generally far below the 2% minimum goal. As SARs approach 2%, it will be more difficult to assess compliance.

[11] See text preceding Table 2.3.1 for an explanation of how costs were derived. Various long term foundational tagging programs are assumed to continue.

[12] "Definitive evaluations" are considered to occur when the 90% confidence interval for the estimate does not overlap the target. This is used as an example decision criterion.

Evaluation Criteria	Status Quo	Low	Med	High
… of in-river survival target	**wild spring/summer**: 2/9 years **wild + hatchery stocks**: 7/9	**wild spring/summer**: 2/9 years **wild + hatchery stocks**: 3/9	**wild spring/summer**: 2/9 years **wild + hatchery stocks**: 7/9	**wild spring/summer**: 7/9 years **wild + hatchery stocks**: 7/9
… of transportation effectiveness	**wild spring/summer**: 3/10 years **hatchery stocks**: Rapid River 6/8; McCall 7/8; Imnaha 4/8; Dworshak 2/8; Catherine 1/4 **MPG level**: *in prep*	**wild spring/summer**: 3/10 yrs **hatchery stocks**: Rapid River 2/8; McCall 4/8; Imnaha 2/8; Dworshak 1/8; Catherine 0/4 **MPG level**: *in prep*	**wild spring/summer**: 3/10 yrs **hatchery stocks**: Rapid River 6/8; McCall 7/8; Imnaha 4/8; Dworshak 2/8; Catherine 1/4 **MPG level**: *in prep*	**wild spring/summer**: 3/10 yrs **hatchery stocks**: Rapid River 6/8; McCall 7/8; Imnaha 5/8; Dworshak 3/8; Catherine 1/4 **MPG level**: *in prep*

[13] Source for Status Quo: Figures 4 and 6-10, Table F-1 in CSS (2006); Tables D-13 to D-18 in CSS (2007)

2.3.5 Conclusions and recommendations

The optimal design and confidence level in answer to hydrosystem action effectiveness questions depends on four factors: 1) how the question is asked; 2) the decision criteria used, 3) the spatial and temporal scale of interest; and 4) the true value of the parameter being estimated relative to the target. Factors 2 and 3 are particularly important in determining how precisely one can estimate a performance measure, and factors 1, 2 and 4 determine how precise you need to be. This summary is meant to provide a *demonstration* of a systematic process for converging to a reliable M&E program for hydrosystem questions, based on our understanding of agency mandates and performance standards, questions of interest, and an assumed set of decision rules. Table 2.3.1 uses the decision rule that "definitive evaluations" occur when the 90% confidence interval for an estimated SAR or TIR is entirely above or entirely below a target, and does not overlap it. We have explored other decision rules (e.g., chances that the 5-year average for a performance measure exceeds a standard), and summarize our results in Volume 2 of this report. We hope that this demonstration will catalyze further dialogue with decision makers and program managers on their requirements and possible decision rules.

The benefits and costs of the different alternatives are outlined in Table 2.3.1. In general, the main benefit of the 'High' alternative is that it improves the precision of estimates of SARs and in-river survival for wild spring-summer Chinook at multiple spatial scales (i.e., ESU and MPG), allowing more definitive evaluations of annual compliance with targets. The ability to estimate annual compliance with the SAR target is already good under the Status Quo M&E, but this is partly because SARs have been so far below the target. If SARs approach the lower limit of the target range (i.e., 2%) higher precision estimates may be required to definitively assess compliance. If SARs were to rise significantly above the 2% level in future, then less precise estimates might be sufficient to assess compliance. The 'High' alternative does not provide substantial improvements in evaluating transportation effectiveness on an annual basis, for three reasons. First, TIRs are a ratio of SARs, and in some years TIRs appear to be much more variable than SARs alone (see confidence intervals on the graphs in Volume 2). Second, the 'High' alternative increases the number of transported fish much more than the number of in-river fish, which constrains how much improvement in precision occurs. Third, estimated TIRs are frequently close to 1, the assumed "threshold", consequently it is not possible to determine whether in-river or transportation was better regardless of CI width. The 'Medium' alternative makes hatchery TIR estimates more representative of the total population (compared to the Status Quo alternative), but makes little difference to our quantitative metrics of statistical reliability. The 'Low' alternative, which drops CSS tagging of hatchery fish, would substantially reduce the ability of managers to definitively assess annual compliance with in-river survival targets (wild plus hatchery fish), and the ability to assess transportation effectiveness for hatchery fish.

We strongly recommend using multiple-year estimates for assessing compliance, in addition to annual estimates. Multiple-year averages can provide insights on compliance with only a relatively small number of PIT-tags (e.g., 1,000 to 5,000 tags), which permits analyses on smaller spatial scales (e.g., MPGs, some large populations) and smaller temporal scales (in-season patterns). Increasing the number of tags/year can help to improve the precision of annual and seasonal estimates, but for transportation evaluations a very large increase in tags would be required to make substantive improvements over the Status Quo. For multiple-year averages, statistical precision improves up to the level of 5,000 PIT-tags, beyond this level there isn't much benefit. However, adding more years to those averages can significantly improve statistical precision. There is a tradeoff however, in that longer durations of monitoring (e.g., beyond 5–10 years) might be beyond the time scales of interest for some decisions.

Ultimately, the most cost-effective approach is to integrate tags from multiple sources for multiple management questions, which we discuss in Section 2.6. The level of integration possible is highly dependent on the questions, and how they are framed.

2.4 Habitat

The goal of CSMEP's Habitat Subgroup was originally to develop a generic template that could be modified and applied to different design situations within the Columbia River Basin. However the group identified several challenges to this:

1. Habitat conditions vary greatly across subbasins in terms of their natural biogeoclimatic regimes, the status of their fish populations, the degree of human impact and management, and the number and nature of restoration actions that have been implemented, or are being considered for implementation within them.
2. Habitat effectiveness questions encompass different scales of inquiry, which imply different scales of monitoring.
3. There has been, to date, a lack of specific policy input/guidance on habitat effectiveness questions.[14] Given the range of habitat conditions and various scales of interest, this input is crucial for narrowing the range of possible habitat action effectiveness designs.

CSMEP's Habitat Subgroup endeavored to work beyond the original plan of developing a generic template design, and instead tried to provide decision-makers with practical examples of why particular types of information are so important for quantitative design. This compromise provides a way of moving beyond a general discussion of design considerations and avoids developing a generic design that provides a precise answer to the wrong question.

As a pilot evaluation of this approach the Habitat Subgroup designed several alternative plans for monitoring the effectiveness of restoration actions prescribed in the Lemhi Habitat Conservation Plan (HCP). The planned duration of the Lemhi HCP is 35 years, during which time a number of water conservation projects will be implemented. Although a number of restoration activities are planned as part of the Lemhi HCP, the most significant projects will consist of actions designed to reconnect isolated tributaries to the main stem Lemhi River and reestablish historic temporal hydrographic patterns. This series of approximately 10 to 16 restoration actions are expected to improve access to historical habitat for Chinook salmon, steelhead, and bull trout.

2.4.1 Priority questions and the question clarification process

The priority questions identified within the Lemhi Habitat Conservation Plan were:

1. Have reconnection projects increased the distribution and density of Chinook juveniles?
2. Have reconnection projects increased number and size of juvenile Chinook outmigrants?
3. Have reconnection projects changed timing of Chinook outmigration?
4. Have reconnection projects increased Chinook parr-smolt survival?
5. Have reconnection projects increased Chinook adult returns?
6. Have reconnection projects increased distribution and abundance of bull trout?
7. Have reconnection projects improved bull trout survival?

As these initial questions were considered far too generic to adequately address the specific responses to tributary reconnections, the Habitat Subgroup created a series of nested subquestions that could further

[14] For example many Habitat Conservation Plans (HCPs) lack specific biological criteria for success.

clarify the information needs. Although intended for policy makers, the Habitat Subgroup applied this "Question Clarification" process to their interpretation of the intent of the Lemhi HCP. This process produced a suite of clarified questions for the HCP around which the Habitat Subgroup could develop their designs. Table 2.4.1 provides an example of how the Question Clarification process incrementally refines the information requirements for fully addressing a habitat effectiveness question for management purposes.

Table 2.4.1. Example of a key general question about habitat effectiveness and the nested "question clarification" process used to precisely determine the specific information needs required to address this question sufficiently for management purposes.

Key general habitat action effectiveness question (example)
1. Have specific habitat projects affected Chinook population abundance or condition in the Lemhi River subbasin?
Question clarification process:
• What are all the species, including life-history type and gender, of interest?
• What is the spatial boundary of the population for which inferences will be made?
• What is the population response variable you want to evaluate to determine whether a change has occurred?
• Define change in the population response variable (i.e., what is the reference and final condition)?
• What is the size of change in population response you want to be able to detect?
• Over what time period(s) do you want to describe this population response?
• Are there surrogate measures that you can use to answer this question?
• To what factors do you want to be able to attribute the observed population response?
• What tradeoffs between uncertainty, errors, and costs are you willing to accept?
Etc.

2.4.2 Related decisions

Determining whether goals are met prior to the full implementation of the Lemhi HCP (35 year time frame) will require frequent review of the information collected by the effectiveness monitoring program. In the event of underperformance of current HCP prescriptions and schedules, a related decision is whether interim goals should be established under an adaptive management framework that will prescribe more aggressive actions, or alternatively, continuation of scheduled activities or even scaling back if objectives are being achieved as planned.

2.4.3 Consequences of wrong decisions

If a conclusion is reached that the Lemhi HCP has resulted in a benefit to the target populations when, in fact, the actions have had no beneficial effect (termed a Type I error), recommendations for these types of restoration efforts to be undertaken elsewhere could be erroneous. Given limited resources for restoration projects, funds used for non-beneficial actions are wasteful and may exclude implementation of other truly useful strategies. Alternatively, concluding that no benefit of the HCP exists, when a benefit has actually occurred (termed a Type II error), may result in the termination of actions that actually work well. This type of error could be potentially harmful to local fish populations in the Lemhi River watershed, and a consequent failure to apply these types of restoration actions to similar habitat problems elsewhere would be lost opportunities for other high risk populations.

2.4.4 Monitoring design alternatives and trade-off analyses

Low', 'Medium', and 'High' intensity effectiveness monitoring design alternatives are presented in Table 2.4.2, for addressing the Lemhi HCP habitat restoration effectiveness questions (Table 2.4.3). "Intensity" refers to the relative density and distribution of sampling within areas A, B and C of the Lemhi watershed (see Figure 2.4.1). The "Status Quo" is an alternative that represents current monitoring in the Lemhi Basin and was not designed to detect fish responses to habitat restoration projects implemented in the Lemhi watershed. Instead, it provides some basic information for evaluating the status and trends of Chinook.

Building on existing monitoring programs and data, the 'Low' alternative makes relatively minor adjustments in the current monitoring regime in order to provide a basic design that would detect the effects of the Lemhi HCP on steelhead and Chinook. It is not intended to provide information about the cause-effect relationships that drive observed changes and thus will provide no objective basis by which managers can improve existing actions, or those implemented in the future. The 'High' design alternative is an "ideal" design that should be capable of providing precise answers as well as feedback to managers to improve both how actions are implemented and monitored. The 'Medium' alternative falls in between the 'High' and 'Low' alternatives with respect to criteria such as precision, cost, and the ability to provide adaptive feedback.

Ideally, explicit statistical models of the 'High', 'Medium' and 'Low' designs should be developed in conjunction with test monitoring data. This would allow evaluation of precision and bias in the performance measures captured in Table 2.4.2, while also determining minimum sample sizes necessary to achieve a given statistical power to detect effects of importance. Test data was not yet available within the Lemhi River watershed to allow the Habitat Subgroup to make such evaluations and, to date, CSMEP has not completed a formal analysis for estimating trade-offs between precision and sample size. Current trade-off comparisons between the alternative designs for the Lemhi (see Table 2.4.3) are therefore primarily qualitative and based on the practical experience of CSMEP analysts

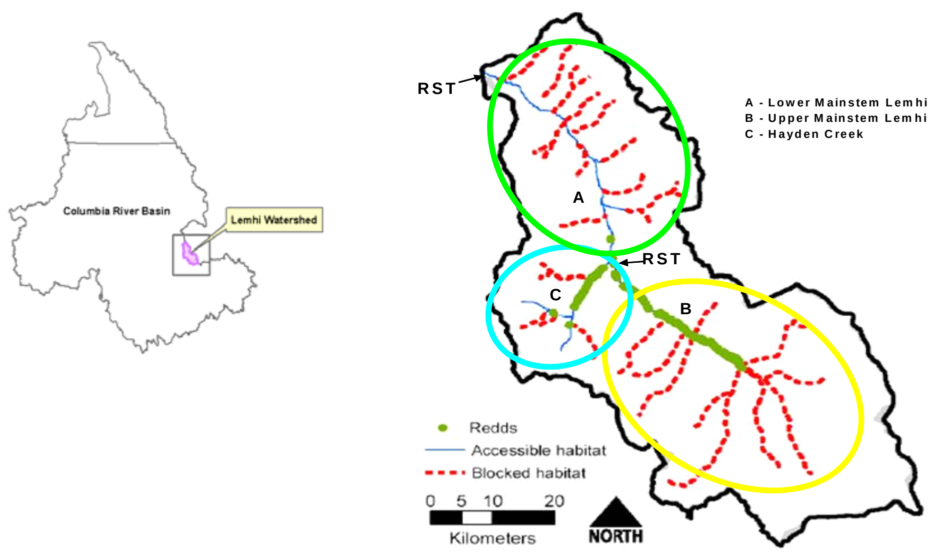

Figure 2.4.1. Map of the Lemhi River watershed denoting Sections A (migration corridor), B (action area), and C (potential reference area). RST - location of existing rotary screw traps.

Cost models were estimated for each of the ‘Low’, ‘Medium’ and ‘High’ designs in the Lemhi using both a “Top-down” and a “Bottom-up” approach (see Table 2.4.4). The “Top-down” approach was based on per project costs and contracting history for previous projects. The “Bottom-up” approach is based on unit costs (e.g., costs per sample) times the number of units (e.g., number of samples) and is thus explicitly linked to the differences in sample size and monitoring protocol. Using the two approaches simultaneously provides a means of “bounding” the annual costs of each alternative while creating a useful cross-check between practical experience and design-driven costs.

2.4.5 Conclusions & recommendations

CSMEP’s Habitat Subgroup, through their Question Clarification process, developed the ‘Low’, ‘Medium’, and ‘High’ intensity design alternatives for monitoring the effectiveness of habitat restoration actions in the Lemhi River watershed. Although each alternative would allow for quantitatively evaluating the effects of HCP reconnection projects on fish populations to varying degrees of bias and precision, the more involved and costly ‘Medium’ and ‘High’intensity designs would likely be required for discerning the mechanistic connections between restorative actions and fish response (i.e., why actions worked or did not). Implementing any one of the design alternatives would provide better information than the current and ongoing Status Quo alternative in the Lemhi River watershed (which currently monitors only the status and trends of Chinook) while simultaneously monitoring the effectiveness of habitat restoration actions for the duration of the effectiveness monitoring program.

The Habitat Subgroup has also identified a number of pragmatic issues regarding the Lemhi HCP that must be resolved in any technical “template” for habitat action effectiveness monitoring. Practical action effectiveness monitoring designs must first incorporate sufficient analytical flexibility to compensate for less than complete control over action implementation. Second, it is likely that existing, but disparate, sampling efforts cannot provide adequate information at the temporal and spatial scales required for efficient implementation of action effectiveness evaluations. Thus, it is likely that the efficient implementation of action effectiveness evaluations will necessitate both a new sampling effort and the modification of existing sampling efforts. Third, it is clear that targeted research for illuminating the mechanistic linkages between habitat restoration actions and fish population responses is still needed. Resource managers must have the tools necessary for making the correct tactical monitoring decisions and properly prescribing habitat restoration actions. As one moves to other subbasins where habitat management issues are diverse, there are likely to be potentially large differences in design elements; in particular, where and when to deploy monitoring resources. It will be impossible to predict this ahead of consideration of the mature scientific questions specific to those locations. Consideration of those questions will in turn require a unique rather than template process that is informed by the management history and management plans in those new locations.

Table 2.4.2. Alternative sampling and response designs for evaluating Lemhi River subbasin habitat actions (**what**, **how, where** data are collected).

Performance Measures	Status Quo (SQ)	Low	Medium	High
1. Spatial distribution (Chinook par, steelhead parr/smolts, all bull trout)	Snorkel counts conducted in (A) and (C)	SQ + Hayden Creek	'Low' + snorkel counts in all tribs with higher intensity.	'Medium' + and in mainstem below all trib junctions for abundance estimates.
2. Parr density (Chinook)	Snorkel counts conducted in (A) and (C)	SQ + Hayden Creek	'Low' + snorkel counts in all tribs with higher intensity.	'Medium' + fixed sites within tribs, and in mainstem below all trib junctions for abundance estimates
3. Smolts per redd (Chinook)	One screw trap located in (A).	Screw traps in (A), (B) and (C).	Same as 'Low'	'Medium' + PIT tag detectors at the mouths of (B) (A) and (C).
4. Migratory timing & size (Chinook)	One screw trap located in (A).	Screw traps in (A), (B) and (C).	Same as 'Low'.	'Medium' + PIT tag detectors at the mouths of (B) (A) and (C)..
5. Parr-to-smolt survival (Chinook)	Survival from trap in Lower Lemhi to LGR.	Some tagging from fish captured through seining throughout drainage. Screw trap at mouth of (B) (A) and (C).	'Low' + More extensive tagging from fish captured through seining throughout drainage.	'Medium' + PIT tag detector in all reconnected tribs and in mainstem below all tribs.
6. Redd counts (Chinook)	Redd counts conducted in upper Lemhi.	Full (A+B+C) redd surveys.	Same as 'Low'	Same as 'Low'.
7. Spawning adults (Chinook)	Inferred from redd counts	Full mainstem (A+B+C) carcass surveys.	Same as 'Low'	'Low' + weirs at (B) and just below confluence of (A)and (B). PIT tag adults and recapture with carcass surveys and PIT tag antenna.
8. Population abundance (bull trout)	Redd counts conducted in some tribs in (C) and (A).	Redd counts in paired tribs containing bull trout in the lower (B) and upper (A) Lemhi, and control tribs in Hayden Creek (C).		Extensive mark-recapture data collected in paired tribs throughout the Lemhi Basin and control tribs in Hayden (C) to estimate abundance and bias in redd counts. Use of PIT-tag detectors at key migration points
9. Survival of juvenile and adult migratory bull trout	N/A	N/A	N/A	Extensive mark-recapture data collected in paired tribs throughout the Lemhi Basin and control tribs in (C) to estimate survival across life stages. Use of PIT-tag detectors, weirs, and screw traps at key migration points to provide additional recapture events.

Table 2.4.3. Overall effectiveness monitoring designs for evaluating effectiveness of Lemhi River watershed habitat restoration actions, and qualitative assessment of design alternatives. Quality of information: 5 = excellent; 4 = very good; 3= good; 2= fair; 1=poor; N/A = not applicable.

Questions evaluated	Status Quo (SQ)	Low	Medium	High
1. Have projects increased the distribution and density of Chinook juveniles?	(1) Presence/absence only, area limited	(3) Qualitative differences in density, limited habitat information	(4) Detect effects, Improved spatial resolution vs. 'Low'	(5) Most powerful design. Mark-recapture estimates of density. Should demonstrate project effects.
2. Have projects increased number and size of juvenile Chinook outmigrants?	(1) Area limited, cannot detect effects	(3) Improved design, but still limited ability to detect effects	(3) Detect effects, habitat surveys increase likelihood of identifying cause/effect relationship	(4) Detect effects, screw trap and PIT tag antennas will increase accuracy & precision of population estimates.
3. Have projects changed timing of Chinook outmigration?	(2) Same as Question 2	(2) Same as Question 2	(2) Same as Question 2	(2) Same as Question 2
4. Have the projects increased Chinook parr-smolt survival?	(1) Before/after possible, unlikely to detect effects	(2) Same as Question 2	(2) Same as Question 2	(2) Same as Question 2
5. Have the projects increased Chinook adult returns?	(1) Area limited, cannot detect effect	(3) Better design, but still unlikely to detect effect.	(4) Detect effects, habitat surveys increase likelihood of identifying cause/effect relationship	(5) Weirs, carcass surveys and PIT tag antennas increases precision & accuracy.
6. Have projects increased distribution and abundance of bull trout?	(1) Area limited, no pre-project data exists for treatment tribs	(3) Improved design, some pre-treatment data, migratory bull trout only	N/A	(5) Abundance for resident & migratory bull trout, evaluation of redd count bias
7. Have the projects improved bull trout survival?	NA	NA	NA	(5) Good design, estimates of density. Should demonstrate project effects.

Table 2.4.4. Costs of alternative CSMEP habitat action effectiveness monitoring designs for the Lemhi River subbasin.

Cost estimate method	Status Quo (SQ)	Low	Med	High
Top-Down = based on per project costs and contracting history	125,000/yr	$323,000/yr	377,000/yr	$580,000/yr
Bottom-up = based on cost per unit time per person multiplied by the sample sizes identified in the plans.	125,000/yr	$354,000/yr	$493,400/yr	$643,600/yr

2.5 Hatcheries

Questions around the effectiveness of hatcheries are Columbia River Basin-scale in nature. CSMEP hatchery designs consequently needed to extend beyond the boundaries of the Snake River Basin. Among the various questions and uncertainties (CSMEP 2006) which surround the use of hatcheries in the Columbia River basin, CSMEP's Hatchery Subgroup identified the following as the highest priority question:

What is the distribution and relative reproductive success of hatchery origin adults in target and non-target Columbia River Basin populations?

- Species and populations of interest: interior Columbia River Basin stream-type Chinook salmon populations (see CSMEP 2007 for a table of populations).
- Spatial Scales of interest: Designs target the interior Columbia River Basin, but results are applicable at scales as small as individual populations.
- Time Scales of interest: Annual or by generation (approximately six years).

Target populations are defined as those that are deliberately supplemented by hatchery production, and non-target populations as those that are not deliberately supplemented but may receive *de facto* supplementation in the form of stray hatchery origin adults. Strays are defined as any hatchery origin adult from a supplementation program that returns to a population other than its target. Conversely, any adult from a harvest augmentation hatchery is considered a stray if it is not harvested or collected for broodstock but instead attempts to spawn in any stream (supplemented or otherwise).

The distribution and relative reproductive success of hatchery origin adults is of key importance when evaluating the net benefits of hatcheries either individually or cumulatively. In general terms, the effectiveness of hatcheries rests on their ability to either increase harvest and/or to increase the abundance of adults in target populations without decreasing productivity. For both types of programs, the potential for negative impacts can be assessed at a coarse scale by evaluating stray ratios, defined as the relative abundance of stray hatchery origin adults, and by understanding the reproductive success of those strays.

2.5.1 Related decisions

The ability to monitor and estimate stray ratios and the relative reproductive success of hatchery origin adults in target and non-target populations informs numerous management questions, including but not limited to:

1. Is supplementation effective at increasing adult abundance without impacting natural productivity in targeted populations?
2. Do hatcheries, either individually or cumulatively, reduce productivity of non-target populations?
3. How should production within a mixed (hatchery and natural) population, major population group (MPG), or evolutionarily significant unit be apportioned between hatchery and natural origin adults? In short, how do hatchery fish "count" in delisting decisions?

Can we separate the confounding effects of stray hatchery origin adults in hatchery and habitat effectiveness evaluations?

What are the consequences of making the wrong decision?

With regard to the primary question and most of the related decisions, poor information would lead to either: 1) continued or expanded use of hatcheries despite substantial deleterious impacts or 2) decreased use of hatcheries despite their ability to increase harvest and/or decrease extinction risk without substantial impacts to non-target populations. Current knowledge is insufficient to guide decisions regarding the appropriate role of hatcheries in harvest augmentation or recovery, leading to potential paralysis in management decisions and/or management based on best professional judgment.

2.5.2 Monitoring design alternatives and trade-off analyses

Evaluations require two types of information:

1. estimates of the relative abundance of strays in a "representative" group of Columbia River Basin populations and
2. estimates of the reproductive success of hatchery origin adults relative to natural origin adults in target and non-target populations.

Although the two types of information are most informative when utilized simultaneously, sampling challenges preclude the formulation of a single design to generate representative estimates for both. The next two sections therefore develop proposed 'Low', 'Medium', and 'High' level designs separately for each type of information.

Stray ratio design

The relative abundance of strays, hereafter "stray ratio" is calculated as the number of stray hatchery origin adults within a population divided by total adult abundance in that population. These numbers can be obtained either by direct total counts, or as estimated total counts. Secondarily, information on the origin of strays is useful in identifying the spatial extent of straying and the types of hatcheries and/or individual facilities that contribute to observed stray ratios to the greatest degree. The primary source of information used to calculate stray ratios is returns of coded wire tags (CWTs) and external marks such as fin clips that are applied at hatchery facilities. We have identified four primary weaknesses with existing mark recovery data (see also PSC 2005). First, recovery effort is not randomly distributed, with greater effort occurring in supplemented populations. Second, existing reporting mechanisms (i.e., the Regional Mark Information System) often lack the necessary metadata to calculate stray ratios. For example, records may indicate the number of tags recovered from a location but may not include information on the number of carcasses surveyed for tags. Third, recovered tags must be "expanded" based on survey effort (e.g., percentage of handled carcasses that were scanned for a CWT), tagging effort (fraction of fish tagged in the release group), and the probability of detecting a tag if one is actually present, which differs depending on the interrogation technique employed. These "expansions" add substantial variance to estimates of stray ratios. Finally, there is no existing mechanism to report missing data, and thus no means to determine the quality of existing data.

Following interim guidance from NOAA Fisheries, our designs target the ability to detect a stray ratio as small as 5% (Grant 1997) with a coefficient of variation equal to or less than 20% in all populations. If we assume that all hatchery origin adults are 100% externally marked with an adipose fin clip and that 50% of hatchery origin adults are marked with a CWT and that recovery data are perfect (e.g., CWT detection is 100%), simulations suggest that existing (*Status Quo*) recovery efforts will return stray ratio estimates with a coefficient of variation between 13% and 81%, depending on survey effort, when the true stray ratio is 5% (CSMEP 2007). If the total number of carcasses can be estimated (e.g., via sight/re-sight methods) the CV improves slightly, potentially yielding CVs in the range of 10% to 79%. Nonetheless, once all sources of error are accounted for, the precision accompanying stray ratio estimates based on *Status Quo* sampling is unlikely to be sufficient to make sound management decisions.

CSMEP design alternatives (Table 2.5.1) to estimate stray proportions at the population and basin scale will utilize a rotating panel design that will distribute effort in a systematic-random fashion both spatially and temporally in all major population groups. All designs estimate stray ratios for all populations in the interior Columbia Basin, but differ with regard to the frequency of sampling. The ‘Low’ design estimates stray-ratios in one population within each MPG annually using carcass surveys, with the remaining populations sampled approximately every third year using a rotating panel design. The ‘Medium’ design maintains annual sampling in one population and increases the frequency of sampling to approximately every two years in the remaining populations. Additionally, from among the populations sampled annually, bi-directional weirs will be operated on three of them in order to estimate precision and bias in carcass survey techniques. The ‘High’ design builds on the ‘Medium’ design by employing one bi-directional weir in each of the eight interior Columbia River Basin MPGs.

Relative reproductive success design

The greatest uncertainty accompanying the operation of hatcheries regards the impacts of hatchery origin adults on productivity in target and non-target populations. Numerous existing and proposed hatchery research, monitoring, and evaluation (RME) projects have been designed to assess long-term changes in productivity. However, these efforts typically focus only on the target population(s), and thus provide little information to evaluate potential impacts on non-target populations. Likewise, an observed change in productivity when assessed using common performance metrics such as juveniles per adult or adult per adult ratios is only sufficient to indicate that a change occurred, but not why the change occurred. For example, if a decrease in *per capita* productivity were observed, it might be difficult or impossible to determine whether that result was a function of some deleterious impact accompanying supplementation, or any number of other alternatives such as a reduction in habitat quality or density dependence. Molecular genetic techniques can be employed to directly estimate the amount of production that can be attributed to individual naturally spawning hatchery origin adults relative to natural origin adults (relative reproductive success; RRS), thus enabling a direct evaluation of the impacts of hatchery origin adults on *per capita* productivity.

The CSMEP designs (Table 2.5.2) seek to evaluate RRS in target and non-target populations selected to represent the range of hatchery management paradigms in the interior Columbia River Basin. A few RRS studies are underway or proposed, however they do not represent the range of hatchery management paradigms, and they typically focus only on heavily supplemented populations. Given the diversity of broodstock management and escapement protocols utilized by supplementation programs, we have ranked populations based on their average “proportionate natural influence” (PNI) scores for target populations and by stray ratio for non-target populations (CSMEP 2007). PNI is calculated as (HSRG 2004):

PNI = (proportion of naturally produced fish in the broodstock (pNOB))/
(pNOB + proportion of hatchery fish on the spawning grounds (pHOS))

We propose to distribute RRS efforts across the range of population average PNI values using a systematic random approach, thus enabling the results of the studies to be applied to the collection of supplemented Columbia River Basin population whether or not all are included in the study. Inferences to individual supplemented populations, that are not included in the study, can be made by use of models developed from observed data. The proposed ‘Low’ design utilizes RRS in six supplemented populations and apportions juvenile production to naturally spawning hatchery and natural origin adults, thus estimating juveniles per adult separately for naturally spawning hatchery and natural origin adults. The proposed project will generate estimates over three successive brood years, approximately ten years.

Passive integrated transponder (PIT) tags will be implanted in all sampled juveniles to monitor the subsequent survival of juveniles based on the origin of their parents. The shortcoming of this approach is that juvenile tagging effort may be insufficient to estimate survival to adult return. The ‘Medium’ design builds on the ‘Low’ design by directly estimating RRS of the progeny through adult return. The ‘High’ design is identical to the ‘Medium’ design, but includes a sample of six un-supplemented populations selected using a systematic random sampling approach across the range of stray ratios observed in Columbia River Basin populations. While the ‘Low’ and ‘Medium’ designs provide estimates of the RRS of strays only in supplemented populations, the ‘High’ design also provides direct estimates of the RRS of stray hatchery origin adults in un-supplemented streams.

Table 2.5.1. Objectives by alternatives matrix for hatchery stray ratio designs. For the purposes of cost estimation, the study is assumed to have a ten year duration. Cost estimates include total annual cost, percentage of total annual cost covered by existing programs (e.g., weirs currently operated under other projects), and total annual cost adjusted for existing effort (i.e., net “new” expenditures). Qualitative evaluations (Q): 5 = excellent; 4 = very good; 3 = good; 2 = fair; 1 = poor; ? = Unknown; n.a. not applicable.

		Design Alternatives			
Design objectives	Performance measures	Status Quo	Low	Med	High
Inferential ability (Qualitative)	Ability to representatively estimate stray ratios and origin of strays	(1)	(3) provides only ratios	(4)	(4)
	Frequency of sampling	Varies	(3)	(4)	(4)
Cost (x $1,000)	Average total annual cost	n.a.	$357,000	$551,000	$873,000
	(% of cost covered by existing operations)		(85%)	(60%)	(50%)
	Adjusted total annual cost		$54,000	$220,000	$437,000
Statistical Reliability (N)	Bias estimation	(1)	(3)	(4)	(5)
	Maintain coefficient of variation < 0.2	(1)	(3)	(4)	(4)

Table 2.5.2. Objectives by alternatives matrix for the relative reproductive success designs. The ten year duration of the designs is sufficient to return RRS estimates for three brood years of stream-type Chinook salmon. The 'Low' design is based on parent to progeny ratios, and thus has a five year sampling duration as opposed to a ten year sampling duration for the 'Medium' and 'High' designs, which require parent to progeny and recruit per spawner ratios. Per site sampling costs for the 'Low', 'Medium', and 'High' designs are identical for the first three years, in subsequent years the 'Low' design costs decrease because only juveniles are sampled and the operation of weirs can be discontinued (for the purposes of this study). Cost estimates include total annual cost, percentage of total annual cost covered by existing programs (e.g., weirs currently operated under other projects), and total annual cost adjusted for existing effort (i.e., net "new" expenditures). Qualitative evaluations (Q): 5 = excellent; 4 = very good; 3 = good; 2 = fair; 1 = poor; ? = Unknown; n.a. not applicable.

Design objectives	Performance measures	Design Alternatives			
		Status Quo	Low	Med	High
Inferential ability (Qualitative)	Ability to representatively estimate relative reproductive success across PNI	n.a. or ?	(3) Adult to juvenile only	(4)	(4)
	Ability to estimate RRS of strays in non-target populations	1	(3) Hatchery influenced only	(3) Hatchery influenced only	(5) Supplemented and un-supplemented
	Life stage specific impact assessment	Varies	(3) Juvenile/Adult	(5) Juvenile/Adult and Adult/Adult	(5) Juvenile/Adult and Adult/Adult
Cost	Average total annual cost	N/A	$241,000	$469,000	$938,000
	(% of cost covered by existing operations)		(85%)	(85%)	(42%)
	Adjusted total annual cost		$36,000	$70,000	$544,000
Statistical Reliability (N)	Robust to changes in overall productivity	N/A	(3)	(3)	(5)

2.5.3 Conclusions

As described in previous CSMEP hatchery subgroup documents (CSMEP 2006), current (*Status Quo*) Columbia River Basin hatchery RME is primarily focused at the scale of individual projects. At that scale, existing RME is likely to provide adequate information to address the impacts of hatcheries on abundance and productivity of those specific targeted populations. Alternatively, little existing research is focused on the aggregate impact of hatcheries, particularly with regard to non-target populations. After extensively reviewing existing hatchery RME, we have found that the most intensive RME projects (e.g., those employing RRS) generally tend to accompany the most innovative supplementation projects. Likewise much less intensive RME, with regard to genetically-based RRS or simple mark recovery effort, accompanies non-target populations. This non-random distribution of effort precludes statistically valid inference from sampled to un-sampled populations. As a result, under the *Status Quo*, monitoring effort must be deployed wherever we want an answer. Additionally, we have determined that methods for collecting, analyzing, and reporting data vary significantly among agencies. Thus, even if effort were representatively distributed, it is unclear whether the resulting information could be aggregated and analyzed to enable statistically valid inference to un-sampled populations.

CSMEP hatchery subgroup efforts have thus focused on the development of systematic sampling designs that representatively sample populations and enable strong statistical inference for un-sampled populations. Likewise, we have identified the need for standardized sampling, analysis, and reporting methods.

For both the stray ratio and RRS design alternatives the differences between the ‘Low’, ‘Medium’, and ‘High’ designs developed by the hatchery subgroup are best illustrated by considering the secondary management questions that could be informed by the designs. For example, while it is true that selecting the ‘Medium’ or ‘High’ level straying design offers improved precision relative to the ‘Low’ design, the ‘Medium’ and ‘High’ level designs have a secondary benefit in that they provide additional information – namely, an improved ability to identify where strays originate, as opposed to simply their number. The ‘High’ design alternative provides information at the MPG scale, and thus may be more useful for de-listing decisions based on IC-TRT criteria. Similarly, the ‘High’ level RRS design alternative yields direct estimates of the RRS of stray hatchery origin fish in un-supplemented populations, whereas that information must be inferred for either the ‘Medium’ or ‘Low’ design alternatives. Although not directly required *per se* to address the primary management question, that information is likely to be useful in de-listing evaluations and as a means to control for the effect of strays for habitat or hatchery action effectiveness evaluations that rely on treatment versus reference comparisons.

Lastly, the implementation of even the ‘Low’ stray ratio and RRS hatchery designs offers substantial improvement over the *Status Quo*. While RME costs would increase over the short-term, in the long-term the inferential ability afforded by even the ‘Low’ designs will significantly reduce RME expenditures within the Columbia River Basin. This statement follows from the simple fact that under the *Status Quo,* RME is required for every program/population for which information is desired. Thus any new propagation program would have to be accompanied by substantial RME. While the CSMEP designs do not supplant the need for all program specific RME, they do significantly reduce the breadth of RME that would otherwise be required to accompany all programs. In addition, the CSMEP designs enable an evaluation of the aggregate impacts of hatcheries, which cannot be achieved given existing RME. Perhaps most importantly, the CSMEP designs enable informed decisions with regard to the use of hatcheries, and achieve this goal by building on existing RME effort, thus affording substantial cost-efficiency.

2.5.4 Design recommendations

Stray Ratio Design
The consensus opinion of the hatchery subgroup is to recommend implementation of the medium-level stray ratio design alternative. The medium-level design alternative provides stray ratio estimates at the population scale and enables estimates of precision and bias in carcass recovery methods for a single population within each of three MPGs. However, if there is reason to believe that the precision and/or bias of carcass recovery efforts would vary among MPGs, it may be prudent to implement the high design and/or to move the three experimental bi-direction weirs periodically to evaluate bias and precision within each MPG.

Relative Reproductive Success Design
The consensus opinion of the hatchery subgroup is to recommend implementation of the medium-level RRS design alternative. The medium level design ensures that RRS can be calculated over the entire life-cycle, although it will not give comparable productivity estimates in un-supplemented populations. If there are reasons to suspect that the reproductive success of naturally spawning hatchery origin fish might change in the presence of greater numbers of hatchery origin adults, it would be prudent to implement the high level design.

2.6 Integrated Monitoring

Monitoring and evaluation involves systematic long-term data collection and analysis to measure the status of the resource, detect changes over time and test action effectiveness. These efforts can be used to evaluate the success of management strategies, potentially revise these strategies, or to focus research on determining the reason for observed changes. Currently, fish populations in the Columbia River Basin are monitored by a number of separate programs established by different agencies. Most of the fish monitoring programs were designed to answer specific management questions at small spatial and temporal scales (e.g., targeting a particular stream or a particular component of the life cycle) and utilize different measurement protocols and sampling designs. This has resulted in an inability to efficiently integrate monitoring at larger spatial scales required for ESU or regional fish population assessment. There is a need for consistent, long-term integrated monitoring of Columbia River Basin fish populations. However, integrated monitoring cannot be carried out by one organization or agency alone. The design and implementation of integrated monitoring at the Columbia Basin scale is problematic, not least because of the constraints imposed by the need to make maximum use of existing monitoring sites and networks. Major program design issues with truly integrated monitoring include the need to address multiple objectives across agencies, the role of existing monitoring sites and operational aspects of integrating program infrastructures.

One of the most difficult aspects of designing a comprehensive monitoring program is integration of many different monitoring projects so that the interpretation of the whole monitoring program yields information more useful than that of individual parts (NPS 2006). Full integration requires consideration of five dimensions, including space, time, life history stages, multiple species, and multiple programs:

- *Spatial integration* involves establishing linkages of measurements made at different spatial scales within a monitoring network, or between individual programs and broader regional programs. It requires understanding of ecological processes, spatially representative monitoring sites, and the design of statistical sampling frameworks that permit the extrapolation and interpolation of data.
- *Temporal integration* involves linking measurements made at various frequencies (e.g., daily flow and temperature measurements, annual redd counts, channel and vegetation assessments every few years). Temporal integration requires nesting the more frequent (and often more intensive sampling) within the context of less frequent sampling.
- *Life history integration* involves assessing survival and habitat requirements throughout the entire life cycle of the fish.
- *Species integration* involves efficiently collecting information for multiple species present in the system
- *Programmatic Integration* involves the coordination and communication of monitoring activities within and among federal, state and tribal agencies, to promote broad collaborative participation in monitoring designs, consistent monitoring protocols wherever feasible, and multiple uses of the resulting data.

CSMEP has begun to explore alternative approaches for integrating designs across M&E domains within its Snake River Basin Pilot Study. These efforts are intended to identify strategies and develop analytical tools to assist integration efforts. Improved monitoring efficiencies through integrated designs across multiple questions and scales, is a common challenge and goal in all basins; hence the results from CSMEP's pilot work will benefit the entire Columbia River Basin.

2.6.1 Integration Strategies

CSMEP subgroups have each developed M&E designs to address specific questions relevant to decision makers in their particular domain. These designs have (to date) been developed separately from the designs of the other domains, with only limited effort to integrate them. Now that subgroup-specific designs have been formulated for identified priority questions, CSMEP can assess where elements of these designs may converge (spatially, temporally, ecologically and programmatically). Identification of the common elements within the designs will provide the 'building blocks' to develop a Columbia River Basin-wide integrated M&E program to address a suite of management questions. This will be an iterative learning process, through which CSMEP will identify workable strategies for simultaneously addressing multiple questions across domains.

Strategies for integration that CSMEP is pursuing include:

1. *Building on a Status & Trends foundation.* Layering of action effectiveness M&E alternatives on a consistent foundation of spatially representative Status and Trends monitoring
2. *Integration within domains*. Evaluating how alternative designs could best address multiple questions within a particular M&E domain (i.e., Hydrosystem, Hatchery, Harvest, Habitat, or Status & Trends specific)
3. *Integration across domains*. Evaluating how alternative designs could best address multiple questions across M&E domains (e.g., what elements of each subgroup's designs can serve multiple functions)
4. *Maximizing benefits of monitoring techniques*. Evaluating how any particular monitoring technique can help address multiple questions across M&E domains (e.g., PIT tagging to address a suite of questions)
5. *Maximizing sampling efficiencies and minimizing redundancies in designs*. Evaluating shared costs and data gathering opportunities across overlapping designs.

CSMEP is consolidating an initial set of base designs for the five M&E domains and beginning to identify opportunities to address specific questions in multiple domains simultaneously (Figure 2.6.1). For example, CSMEP's hydrosystem and hatchery stray monitoring strategies are building on the preliminary designs developed by the Status and Trend group. Ultimately, it is CSMEP's intent to develop examples of integrated sets of 'Low', 'Medium', 'High' designs across all five M&E domains to illustrate various dimensions of M&E tradeoffs (i.e., cost, precision, monitoring objectives).

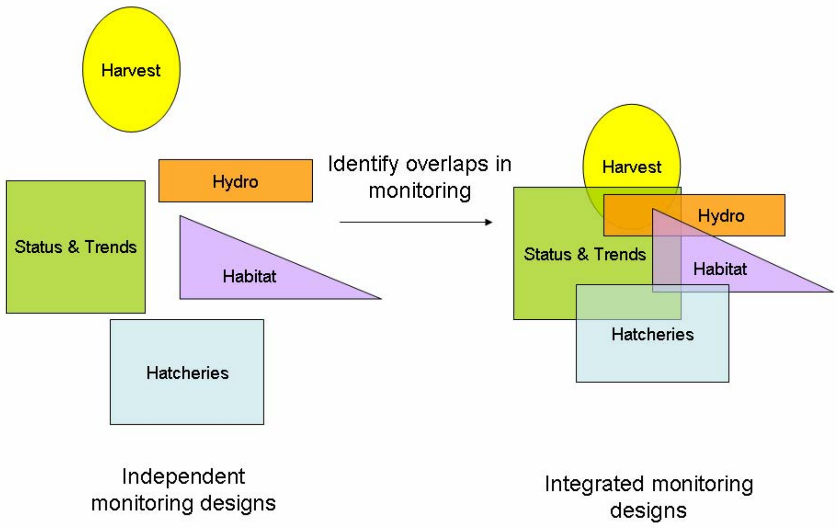

Figure. 2.6.1. Conceptual illustration of identification of opportunities and subsequent development of integrated monitoring designs across CSMEP subgroups.

Integration of M&E depends on the policy and management priorities of each domain and its constituent questions. Consequently, there is no "optimal" design that will exactly suit the preferences of all agencies. Therefore, program managers will need to iteratively review and collaboratively revise integrative strategies and designs. To this end CSMEP has been developing a suite of analytical tools and simulation models that will allow managers and scientists to jointly explore alternative M&E designs and associated trade-offs (i.e., statistical power, costs, sampling effort, etc.).

CSMEP has completed a preliminary analysis of the potential for an integrated PIT-tagging program to address a range of monitoring questions across M&E domains. The intent was to evaluate what intensities of basin-wide PIT-tagging would be required at which life stages and locations (Table 2.6.1) to provide reliable estimates of survival. CSMEP intends to extend this approach to assess statistical-cost tradeoffs; and evaluate other marking and monitoring techniques that have the potential for integration across domains. Figure 2.6.2 illustrates some of the linkages across M&E domains that are possible using PIT tags and other monitoring techniques.

Table 2.6.1. Abbreviated list of questions answerable in whole or in part with PIT-tagged fish.[15]

CSMEP Subgroup:	Question:	Indicator:	Tagging:	Detection:
Status & Trend	Straying of hatchery fish in to wild	Detections of tagged hatchery adults	Hatchery smolts	At tributary weirs or in carcass surveys
	Productivity (smolts per spawner)	Enumeration of smolt emigrants	Parr (for trap efficiency, early emigration), smolts	At smolt trap
	Productivity (adult recruits per spawner)	Age-at-return for adults	Parr or smolts	At LGR as adults or at weir
	SARs	Smolt-to-adult survival	Parr or smolts in tributary	At LGR as adults or at weir
	Hatchery-origin fish spawning in wild	hatchery-origin PIT tagged fish	As smolts in hatchery	At weir or carcass surveys
Habitat effectiveness	Parr abundance, treatment/control areas	Parr #'s	Parr in T/C areas	At traps, flat plate detectors
	Parr-to-smolt survival - treatment/control areas	Parr-to-smolt survival	Parr in treatment, control areas	At dams
	SAR - treatment/control areas	SAR	Parr or smolts in treatment, control areas	At dams
Harvest	Stock composition	Rates of adult tag recovery at dams, in harvest	Parr or smolts	At dam ladders, or in harvested fish
	Age composition of harvested fish	Age-at-return for adults	Parr or smolts	At dam ladders, or in harvested fish
	Harvest rates for listed stocks	Harvest rates	As parr or smolts in Snake	At netting or landing - must happen before fish are gutted
	Upstream survival rate	Upstream survival rate	As parr or smolts in Snake	At BON and LGR adult ladders
Supplementation Hatchery	In-season vs. pre-season adult return estimates	SAR, # of adults returning to supplementation hatchery	Parr or smolts at hatchery	At LGR as adults or at hatchery weir
	Harvest contribution of supplementation fish	Rates of adult tag recovery at dams, in harvest	Parr or smolts at hatchery	At dam ladders, or in harvested fish
	Life-stage survival rates, supplemented pops	Parr-to-smolt survival	Parr	At dams
	Upstream survival	SAR, survival BON to LGR	parr or smolts	At BON and LGR adult ladders
Hydro	Hydrosystem survival, inriver migrants	Smolt survival	Parr or smolts	At dams
	SAR, inriver migrants	SAR	Parr or smolts	At BON and LGR adult ladders
	SAR, transported fish	SAR	Parr or smolts	At BON and LGR adult ladders

[15] The full analysis can be found at www.cbfwa.org/csmep/web/documents/general/Documents/PITtagV4-12-14-05.pdf

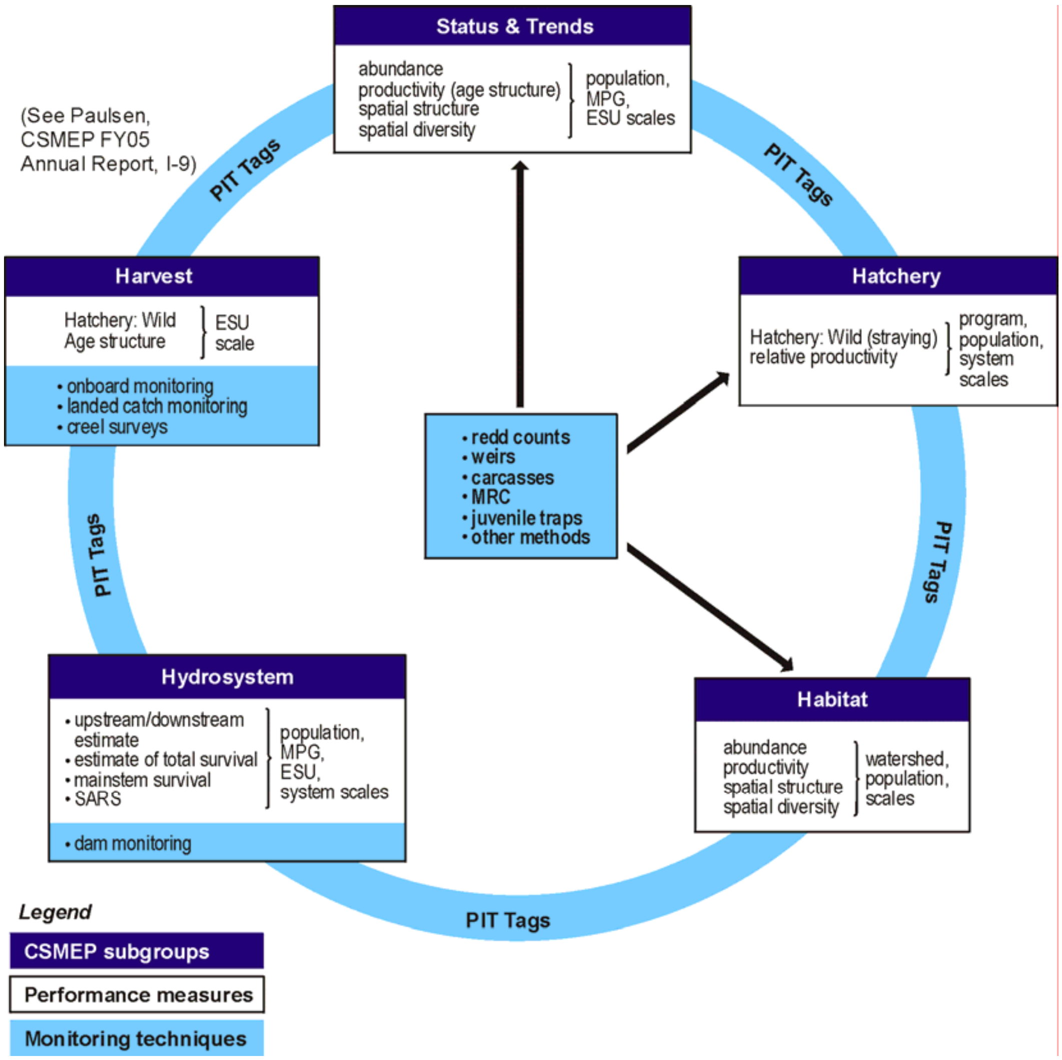

Figure 2.6.2. Monitoring techniques and potential linkages across status & trends and action effectiveness monitoring.

CSMEP is also developing an Integrated Costs Database Tool, a relational database that will assist evaluations of the cost and performance of integrated monitoring designs. The tool is able to combine the varied costs of equipment, personnel and analyses required for both stationary (weirs, smolt traps, etc.) and mobile techniques (redd counts, snorkeling, electroshocking, etc.) used for monitoring. The tool simulates deployment of field crews and specialized analysts working on component projects, and also incorporates the additional costs of different types of fish marking or processing required for analyses. The tool will also identify the full range of performance measures that can be captured across domains as proposed alternative monitoring components are built into an integrated M&E design. As individual domain-specific M&E designs are developed, the tool will help identify infrastructure redundancies and quantify the improved cost efficiencies of overlaying and integrating design components. This database tool and accompanying User Guide will be available shortly for download from the CSMEP public website. A screen capture of the front-end user interface for this developing database tool is shown in Figure 2.6.3.

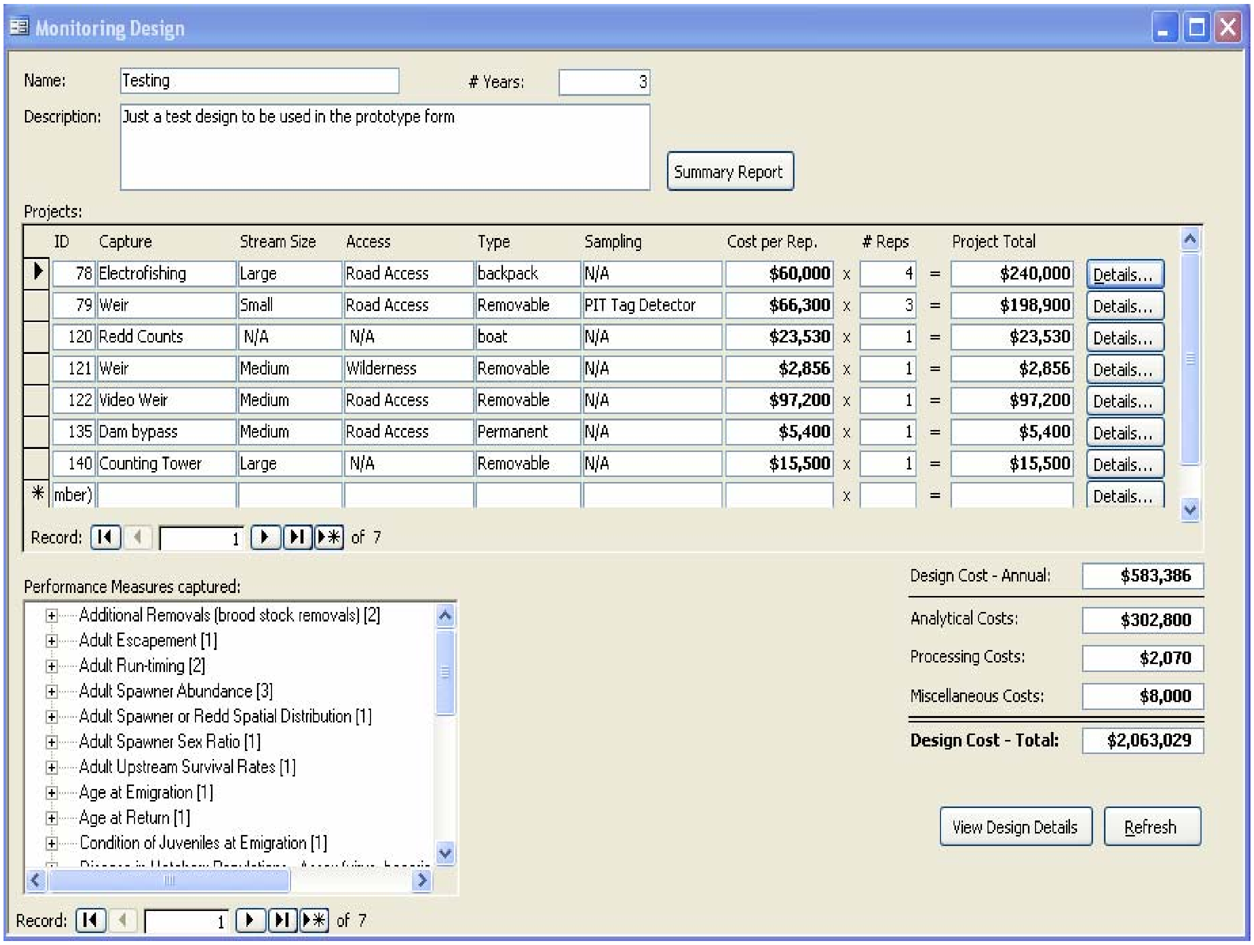

Figure 2.6.3 Front-end user interface for CSMEP's Cost Integration Database Tool.

2.7 Summary of general recommendations

Based upon analyses undertaken within its Snake River Basin Pilot study CSMEP suggests the following general recommendations for developing consistent, cost effective, coordinated, regional status & trends monitoring and action effectiveness monitoring within and among all the 'Hs' (Harvest, Hydro, Habitat, and Hatcheries). Recommendations specific to CSMEP designs for each M&E domain were identified in Sections 2.1–2.6.

Recommendation 1

Regional M&E for fish populations should be developed through a long term, systematic process that has the following attributes:

a. involves dialogue with Columbia River Basin fish managers and decision makers to identify the key management decisions, spatial and temporal scales of decisions, information needs, time frame for actions, and the level of acceptable risks when making the decisions;

b. conducts an inventory of existing M&E methods and evaluates their strengths and weaknesses for meeting information needs;

c. involves the long term participation of Columbia River Basin scientists with both field and statistical expertise, to ensure that M&E approaches meet information needs, are cost-effective, practical, statistically reliable, and have the support of state and tribal agencies;

d. recognizes that information needs, available funding, and scales of interest vary across agencies and it addresses the tradeoffs among design objectives and evaluation criteria; and

e. recognizes that M&E is an essential element of an adaptive management loop (Figure 2.7.1) to iteratively improve habitat, hydrosystem, and fisheries management actions, and that M&E approaches themselves need to be iteratively improved through the evaluation of projects.

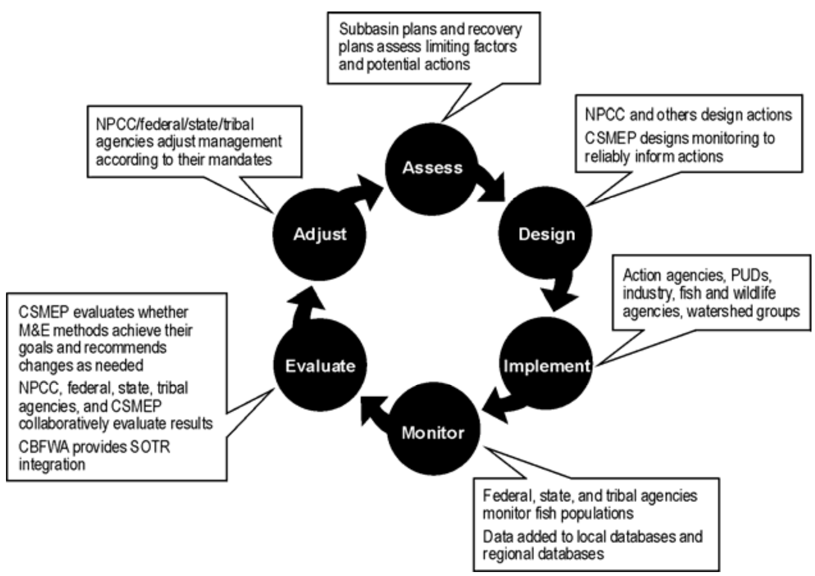

Figure 2.7.1. The adaptive management cycle, with example Columbia Basin entities included. The rigorous M&E designs being developed by CSMEP are essential for adaptive management.

Decisions on regional M&E designs need to be based on a quantitative evaluation of the costs and benefits of the Status Quo and alternative designs to answer management questions. The alternative designs should build on the strengths of each subbasin's existing monitoring infrastructure and data, remedy some of the major weaknesses, and adapt to regional variations that affect monitoring protocols. Without a formal quantitative evaluation of costs and benefits (e.g., statistical reliability, cost, ability to answer key questions, practicality), there is a risk that *ad hoc* M&E decisions will be made that are not cost-effective and preclude data aggregation for decisions and evaluations at greater spatial or temporal scales. Each region in the Columbia River Basin has invested considerable resources to develop a monitoring infrastructure that is primarily adapted to address local needs. It is much more cost-effective to build on the strengths of the existing monitoring infrastructure, rather than applying a uniform "cookie-cutter" approach throughout the Columbia River Basin. These improved designs can be developed to overcome weakness in the existing M&E programs to allow assessments at larger spatial and longer temporal scales.

Recommendation 2

The development and implementation of sound M&E designs must be accompanied by strong data management systems which facilitate the sharing, analysis and synthesis of data across agencies, spatial and temporal scales, and disciplines. Without a strong investment in data management, even the best monitoring designs will falter.

Recommendation 3

Status and trends monitoring should provide the foundation of a regional M&E program but it must be integrated with action effectiveness monitoring. An integrated M&E program provides economy of scale, prevents duplicative efforts, and is cost effective. Action effectiveness monitoring is more focused on specific questions that influence fish populations hence, it is typically of fixed duration and usually provides more precision. Action effectiveness M&E can respond to adaptive management needs by focusing its efforts to address the mechanistic causes of uncertainty in the relationship between management actions and fish population responses.

Recommendation 4

Status and trends monitoring of fish populations must satisfy the needs of population and ESU level assessments (for both listed and unlisted species) of viability, as well as assessments of overall trends in population abundance and productivity at larger spatial and longer temporal scales. It must also meet the needs of multiple agencies with different objectives, questions, and scales of interest. There are challenging tradeoffs to meet all M&E objectives but using the collaborative process CSMEP has adapted should result in cost effective designs to adequately address information needs.

Recommendation 5

M&E designs under development must also be integrated across species. CSMEP is currently working to incorporate steelhead into the Chinook salmon designs that have been developed for the Snake and mid-Columbia basins. CSMEP is working to integrate the use of PIT-tags and other techniques to answer multiple questions, improving the cost-effectiveness of Status & Trends, Habitat, Hydrosystem, Harvest, and Hatchery M&E designs.

Recommendation 6

Agencies should evaluate hybrid sampling designs to improve fish population monitoring that is based on fixed index sites. A hybrid sampling design would supplement the existing non-random, index monitoring sites with spatially representative sites. While index sites are not representative, sampling random sites throughout the range of a fish population is often not efficient (considerable time can be spent getting to each site). The hybrid approach takes advantage of the fact that index sites often efficiently sample a large fraction of the population and uses the supplementary random sampling to accurately determine just how big that fraction is. This approach would allow agencies to assess the bias in index sites, get reliable estimates of population abundance for viability assessments, permit aggregation to a variety of larger spatial scales (e.g., MPG, sub-basin), support the sharing of data collected by different agencies with different interests, and facilitate data analyses.

References

CSMEP (Collaborative Systemwide Monitoring and Evaluation Project) - Marmorek, D.R., M. Porter and D. Pickard (eds). 2006. Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) – Year 3, Project No. 2003-036-00, Annual Report for FY 2006. Prepared by ESSA Technologies Ltd., Vancouver, B.C. on behalf of the Columbia Basin Fish and Wildlife Authority, Portland, OR. www.cbfwa.org/csmep/web/documents/general/Project/ CSMEPFY06AnnualReportFinal.pdf).

CSMEP (Collaborative Systemwide Monitoring and Evaluation Project) – Marmorek, D.R., M. Porter, D. Pickard, and K. Wieckowski (eds.). 2007. (*in prep*). Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) Snake River Basin Pilot Study: Volume 2. Prepared by ESSA Technologies Ltd., Vancouver, B.C. on behalf of the Columbia Basin Fish and Wildlife Authority, Portland, OR

EPA (United States Environmental Protection Agency). 2000. Guidance for the Data Quality Objectives Process. EPA QA/G-4. Office of Environmental Information, Washington, DC. www.epa.gov/quality1/qs-docs/g4-final.pdf

Grant, S.W. (Editor). 1997. Genetic effects of straying of non-native hatchery fish into natural populations: Proceedings of the workshop. U.S. Department of Commerce. NOAA Technical Memo. NMFS-NWFSC-30, 130 p.

Hammond, J. S., R.L. Keeney, and H. Raiffa. 1998. Even swaps: a rationale method for making trade-offs. Harvard Business Review.

Hammond, J.S., R.L. Keeney and H. Raiffa. 1999. Smart Choices: a practical guide to making better decisions. Harvard Business School Press. Boston, Massachusetts.

Hillman, T.W. 2004. Monitoring strategy for the Upper Columbia basin. Draft Report. Prepared by Tracy W. Hillman, BioAnalysts, Inc. Eagle, Idaho. Prepared for Upper Columbia Regional Technical Team, Upper Columbia Salmon Recovery Board, Wenatchee, Washington. February 1, 2004.

HSRG (Hatchery Scientific Review Group). Lars Mobrand (chair), John Barr, Lee Blankenship, Don Campton, Trevor Evelyn, Tom Flagg, Conrad Mahnken, Robert Piper, Paul Seidel, Lisa Seeb and Bill Smoker. 2004. Hatchery Reform: Principles and Recommendations of the HSRG. Long Live the Kings, 1305 Fourth Avenue, Suite 810, Seattle, WA 98101 (www.hatcheryreform.org).

National Marine Fisheries Service. 2000. Recovery Planning Guidance for Technical Recovery Teams. www.nwfsc.noaa.gov/trt/guidanc9.pdf

National Parks Service (NPS). 2006. Vital Signs Monitoring: Guidance for designing an integrated monitoring program. National Parks Service, U.S. Department of Interior. science.nature.nps.gov/im/monitor/GoalsObjectives.cfm

PSC (Pacific Salmon Commission). 2005. Report of the expert panel of the coded wire tag recovery program for Pacific salmon. Vancouver, BC, Canada (www.psc.org/info_codedwiretagreview.htm)