Practical application of meta-analysis results: Avoiding the double use of data

Article *in* Canadian Journal of Fisheries and Aquatic Sciences · April 2005

DOI: 10.1139/f04-245

Carolina Viviana Minte-Vera
Inter-American Tropical Tuna Commission

Trevor A. Branch
University of Washington Seattle

Ian J. Stewart
International Pacific Halibut Commission

Martin W. Dorn
National Oceanic and Atmospheric Administration

Practical application of meta-analysis results: avoiding the double use of data

C.V. Minte-Vera, T.A. Branch, I.J. Stewart, and M.W. Dorn

Abstract: Meta-analysis is an important new tool for synthesizing scientific knowledge from many previous studies. In fisheries, meta-analyses can be used to obtain prior distributions or penalty functions for parameters used in stock assessment models. Two types of results are generally published in a meta-analysis: Type A, the updated results for each stock used in the meta-analysis, and Type B, the results that would best describe a new stock. Including these results in assessments for the individual stocks would result in double use of the data if the assessments include the input data used in the meta-analyses, which they typically would. To solve this problem, we recommend that an additional form of results should be reported in meta-analyses: Type C, the results for a new stock obtained by sequentially excluding each stock's data set and repeating the meta-analysis. Type C results should be used whenever the assessment input data overlap with the meta-analysis input data, avoiding the double use of data. We illustrate the impact of this reporting change on the results of a recent meta-analysis.

Résumé : La méta-analyse est un nouvel outil précieux pour faire la synthèse des informations scientifiques contenues dans un grand nombre d'études antérieures. Dans le cas des pêches, les méta-analyses peuvent servir à déterminer des distributions a priori et des fonctions de pénalité pour les paramètres dans les modèles d'évaluation des stocks. On publie généralement deux types de résultats dans une méta-analyse, les premiers, de type A, les données mises à jour de chacun des stocks utilisés dans la méta-analyse et les seconds, de type B, les résultats qui décriraient le mieux un nouveau stock. Inclure ces résultats dans des évaluations de stocks individuels reviendrait à utiliser deux fois les même données, si les évaluations comprennent les données originales d'entrée de la méta-analyse, ce qui arrive normalement. Pour éviter ce problème, nous recommandons d'inclure un troisième type de résultats dans les méta-analyses, le type C, soit les résultats pour un nouveau stock en excluant tour à tour les données de chaque stock et en répétant la méta-analyse. Les résultats de type C devraient être utilisés lorsque les données d'entrée de l'évaluation chevauchent celles de la méta-analyse, ce qui permettrait d'éviter une double utilisation des données. Nous illustrons les effets de ce changement de présentation sur les résultats d'une méta-analyse récente.

[Traduit par la Rédaction]

Introduction

Meta-analyses have become a powerful way of synthesizing the results of many different studies (Gurevitch and Hedges 1993; Arnqvist and Wooster 1995; Osenberg et al. 1999). Meta-analyses have been performed mainly on the results of multiple experiments (Brett and Goldman 1996; Gurevitch et al. 2000), but in applied fields of ecology, such as fisheries, time series of observations are also included. A growing number of papers have used meta-analysis to address problems that have long thwarted fisheries scientists, including the relationship between biomass and catch-per-unit-effort (Harley et al. 2001), computing survey catchabilities (Millar and Methot 2002), estimating the steepness of stock–recruitment curves (Harley and Myers 2001; Dorn 2002), and determining whether depensation occurs at small stock sizes (Myers et al. 1995; Liermann and Hilborn 1997). Meta-analyses in fisheries have been conducted using either mixed-effects models (Myers et al. 1995, 1999, 2001) or Bayesian hierarchical models (Liermann and Hilborn 1997; Harley and Myers 2001; Millar and Methot 2002). Both methods assume that there is an underlying true

Received 26 February 2004. Accepted 9 October 2004. Published on the NRC Research Press Web site at http://cjfas.nrc.ca on 11 May 2005.
J18007

C.V. Minte-Vera,[1,2] T.A. Branch,[3] and I.J. Stewart. School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, USA.
M.W. Dorn. Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA 98115-6349, USA.

[1]Corresponding author (e-mail: cminte@u.washington.edu).
[2]Present address: Núcleo de Pesquisa em Limnologia Ictiologia e Aqüicultura, Universidade Estadual de Maringá, Maringá, PR, Brazil.
[3]Present address: Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, Western Cape, South Africa.

 doi: 10.1139/F04-245

distribution of the parameter as well as some real interstock differences, which "pull" results from individual stocks away from the average of this true distribution.

Meta-analyses hold special promise in the field of fisheries because they allow information from other stocks or species ("plausible parameter values") to be formally included in stock assessments (Hilborn and Liermann 1998). This process is most easily defined in Bayesian stock assessments, which are becoming increasingly popular (Punt and Hilborn 1997). In Bayesian models, all the parameters are assumed to be random variables with a probability distribution based on previous knowledge (a prior) that will be updated given the available data to produce a resultant probability distribution (a posterior). This process is formalized by Bayes Theorem:

$$(1)\qquad p(\boldsymbol{\theta}|\mathbf{y}) \propto p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})$$

The probability of the parameters ($\boldsymbol{\theta}$) given the data ($\mathbf{y}$) (the posterior) is proportional to the probability of the parameters (the prior) multiplied by the probability of the data given the parameters (the likelihood, or sampling distribution). If we assume that we know nothing at all about the parameter, then an uninformative prior is typically used (see Gelman et al. (1995) for a discussion of how to select an uninformative prior). For many parameters, such as natural mortality, depensation, and those describing stock–recruitment functions, there is often little information in a single data set. Informative priors, based on meta-analyses, are a formal way of combining information from other species or stocks for these poorly estimated parameters. Assessments based on likelihood methods can also incorporate the results of meta-analyses in the form of penalty functions that constrain the estimation process to plausible parameter values (Bard 1974). For both Bayesian and likelihood methods, the prior (or penalty function) and the data used to calculate the likelihood must be independent.

Meta-analyses are typically performed using hierarchical models. In these models, the stock-specific parameters are considered exchangeable components of a hyperdistribution, with hyperparameters assumed to be random variables (Gelman et al. 1995; Efron 1996). A hierarchical Bayesian model can be written as

$$(2)\qquad p(\boldsymbol{\theta},\boldsymbol{\varphi}|\mathbf{y}) \propto p(\boldsymbol{\varphi})p(\boldsymbol{\theta}|\boldsymbol{\varphi})p(\mathbf{y}|\boldsymbol{\theta})$$

The joint posterior probability of the hyperparameters ($\boldsymbol{\varphi}$) and the parameters ($\boldsymbol{\theta}$) given the data ($\mathbf{y}$) is proportional to the probability of the hyperparameters (the hyperprior) multiplied by the probability of the parameters given the hyperparameters and the probability of the data given the parameters.

Standard meta-analysis outputs

We are firm believers in both meta-analysis and Bayesian stock assessments, both in the methods used and in their applications. However, we have a growing sense of unease about the manner in which all meta-analysis results are currently presented because they encourage the double use of data. Two types of results are typically presented for the parameter of interest (Fig. 1). The first is Type A, the result for each individual stock included in the meta-analysis, allowing the data for that particular stock to be "influenced" by data from the $K - 1$ other stocks (the "posterior distribution" in Bayesian meta-analysis). Following the Bayesian notation above, this distribution is

$$(3)\qquad p(\theta_i|\mathbf{y},\boldsymbol{\varphi},\tilde{\boldsymbol{\theta}}) \propto \int_{\varphi}\int_{\tilde{\theta}} p(\boldsymbol{\theta},\boldsymbol{\varphi}|\mathbf{y})\mathrm{d}\tilde{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{\varphi}$$

The marginal posterior probability distribution for the stock i given the data ($\mathbf{y}$), the parameters of the other stocks excluding stock i ($\tilde{\boldsymbol{\theta}}$), and the hyperparameters ($\boldsymbol{\varphi}$) is proportional to the integral of the joint posterior distribution. These results should be used for further work on one of the stocks in the meta-analysis, provided that none of the data in the meta-analysis will be included. The second is Type B, the result for an "unobserved stock" (a synthesis of the data from all the stocks in the meta-analysis, the "posterior predictive distribution" in Bayesian meta-analysis). In Bayesian terms, the distribution of the parameters for an unobserved stock (θ_{K+1}) given the information from the K observed stock is

$$(4)\qquad p(\theta_{K+1}|\mathbf{y}) \propto \int_{\varphi}\int_{\theta} p(\theta_{K+1}|\boldsymbol{\varphi})p(\boldsymbol{\theta},\boldsymbol{\varphi}|\mathbf{y})\mathrm{d}\boldsymbol{\theta}\mathrm{d}\boldsymbol{\varphi}$$

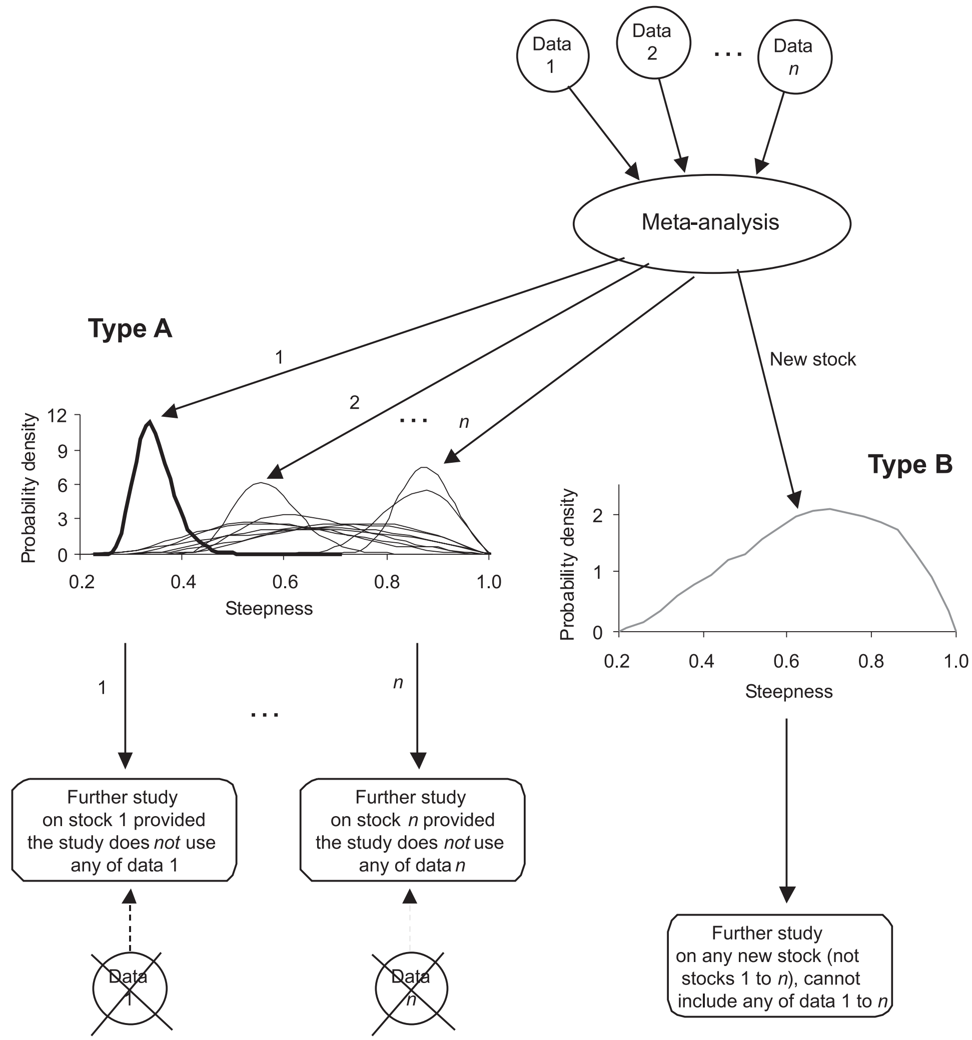
This distribution can answer questions about the evidence for a particular hypothesis as well as the effect size of a given parameter. Type B results should be used as assessment inputs for stocks that were not included in the meta-analysis.

In standard Bayesian analyses, informative priors are obtained using knowledge available before the new experiment commences (Gelman et al. 1995). The collected data are then used to update the priors and obtain posteriors. In contrast, stock assessments are not independent discrete experiments but are based on time series data. Stock assessments usually incorporate many sources of data, including survey indices, yearly total catches, proportions of fish of each age, and commercial catch-per-unit-effort indices. These data sources are simultaneously fit to model predictions based on assumed underlying population dynamics equations (see, for example, Methot 1990). Those model predictions include stock biomasses and recruitment used as inputs in fisheries meta-analyses (e.g., Myers et al. 1999; Dorn 2002). Because these time series of data are increased each year by one or more data points, past data are nearly always retained as part of the updated assessment. For this reason, any informative prior based on past knowledge, i.e., on meta-analyses that use any of the input data used in past stock assessments, is going to include data that are still used in the current assessment. This double use of data will generally result in a posterior distribution for the parameter that is biased and has an incorrect associated level of uncertainty (Gelman et al. 1995, p. 122). Both Type A and Type B results are afflicted by this double use of data if they are used as informative priors (in Bayesian analyses) or as penalty functions (in likelihood analyses) in future analyses.

Solution for the double use of data

We suggest that to discourage the inadvertent double use of data and to broaden the applicability of meta-analytical results, authors should also report a third form of results in their papers (Fig. 2): Type C, the result for an "unobserved stock" obtained by excluding the stock being assessed from the meta-analysis. This could also be called an "independent

Fig. 1. Standard reporting of results from meta-analysis showing the flow from input data to the meta-analysis, to Type A and Type B results from the meta-analysis, and finally to future studies or stock assessments that incorporate Type A and Type B results. The distributions for steepness are those obtained from the meta-analysis conducted by Dorn (2002). Under Type A, stock 1 (in bold) is the west coast Pacific ocean perch (*Sebastes alutus*) stock.

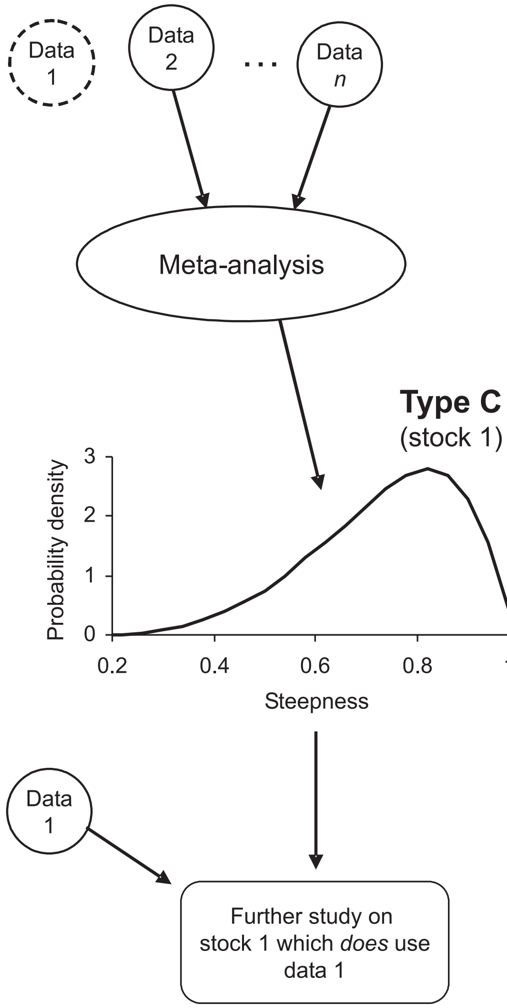


predictive posterior” in Bayesian meta-analysis. A separate Type C result will be required for each of the stocks included in the meta-analysis. These results should be used whenever the same input data are used for both the meta-analysis and the subsequent application, avoiding the double use of data.

Illustration

We illustrate our point by examining the recently published Bayesian meta-analysis on the steepness of rockfish stock–recruitment relationships (Dorn 2002) with the aim of producing results appropriate for use in an updated US west coast Pacific ocean perch (*Sebastes alutus*) stock assessment. Steepness is the fraction of preexploitation recruitment produced when spawning biomass is at 20% of preexploitation levels (Fig. 3). Dorn’s (2002) meta-analysis produced posterior distributions for steepness based on fitting a hierarchical model to spawning biomass and recruitment time series generated by previous stock assessments. This framework allowed sharing of information between stocks, strengthening the inference for those stocks with short time series (Dorn 2002). However, because the biomass and recruitment series were assessment model predictions, they were based on time series data that could not be discarded in future assessments. In the case of west coast Pacific ocean perch, when producing an updated assessment model, only a few years of additional data would be added to the 40-year time series used in the meta-analysis. The Type A estimate of steepness (Beverton–Holt model) for the west coast Pacific ocean perch stock had a mean of 0.35 (SD = 0.039) and a relatively narrow distribution. The Type B estimate of steepness for an unobserved stock had a mean of 0.66 (SD = 0.17). The estimate of steepness for west coast Pacific ocean perch was markedly lower than estimates for the other stocks (Dorn 2002), and therefore, the Type B estimate was influenced in two ways: the overall mean was decreased and the estimated variability among stocks was increased. However, neither the Type A nor the Type B results should be included in the stock assessment for west coast Pacific ocean perch because of the double use of data. The process leading to double use of data is outlined in Fig. 1, where data from west coast Pacific ocean perch (stock 1) flow into

Fig. 2. Proposed additional reporting of results from meta-analysis showing how data from stock 1 (west coast Pacific ocean perch (*Sebastes alutus*) in this example) need to be excluded from the meta-analysis to obtain Type C results for this stock. We reran the meta-analysis of Dorn (2002) omitting data for west coast Pacific ocean perch to obtain the illustrated Type C curve. Note that future stock assessments can use the Type C posterior from the meta-analysis distribution as a prior on steepness in the assessment and still include the data for this stock without double use of data.



the meta-analysis and are incorporated into the Type A and Type B results. Future assessments that include Type A or Type B results and the data used in the meta-analysis will be therefore be guilty of double use of data.

To produce results appropriate for use in an updated stock assessment, we therefore reran the analysis in Dorn (2002), excluding the west coast Pacific ocean perch data, to obtain the Type C distribution of steepness (as illustrated in Fig. 2). The Type C distribution has a higher mean of 0.74 and less uncertainty (SD = 0.14) than the Type B distribution (mean = 0.66, SD = 0.17). This Type C distribution (the independent predictive posterior for the meta-analysis) can be safely used as the prior in a Bayesian stock assessment. We compared the effect of using the incorrect Type B distribution and the correct Type C distribution as priors for steepness in the current stock assessment model for west coast Pacific ocean

Fig. 3. Illustration of the steepness parameter for two stock–recruitment functions with steepness values of 0.8 (broken curve) and 0.4 (solid curve). Steepness is the fraction of preexploitation recruitment produced when spawning biomass is at 20% of preexploitation levels.

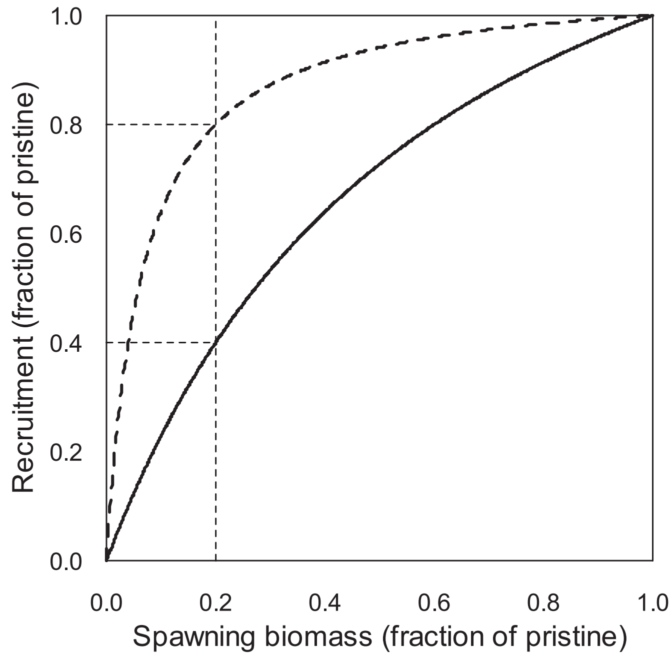
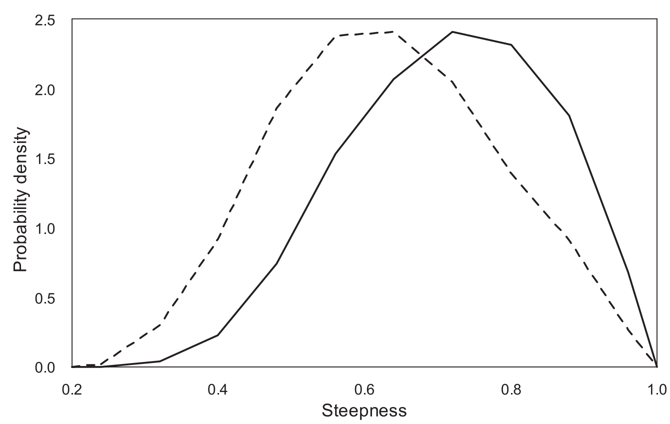


Fig. 4. Posterior distributions for steepness obtained from a stock assessment of west coast Pacific ocean perch (*Sebastes alutus*) that either incorrectly used the Type B distribution (broken line) or correctly used the Type C distribution (solid line) as a prior on steepness. The meta-analysis distributions (Types B and C) were incorporated as informative priors into the current assessment model for this stock, and updated posterior distributions for steepness were obtained from the assessment.



perch (Ianelli 2002). These priors were then updated using the assessment model, producing posteriors for steepness that were markedly different for the Type B prior (mean = 0.63, SD = 0.15) versus the Type C prior (mean 0.72, SD = 0.14) (Fig. 4). This difference in the posterior distribution for steepness from the assessment model does not result in a large change in the mean estimate of current depletion (which increases from mean = 0.30, SD = 0.10 to mean = 0.33, SD = 0.10) but will translate into substantial differences in future recruitment predictions. For example, median predicted recruitment (1 year into the future) when using the Type B prior is 7.2% higher than when using the Type C prior. This difference would increase if predictions were made farther into the future, as higher predicted recruitments will be followed by larger spawning biomass, which in turn will produce higher subsequent recruitments.

Discussion

The approach we recommend will entail an additional burden in both analysis time and reporting space. It may be possible to show through simulation that, for a large number of data sets, there is little effect of removing the single stock of interest. However this “contamination” of the prior with the data would still be incorrect. Because the fish stocks used in meta-analyses are generally those with better data, and those for which assessments are done on a regular basis (i.e., the most commercially important), these are the very stocks most likely to be assessed with the informative priors created. Therefore, it is imperative that reported meta-analysis results be practically applicable to all future studies and assessments. Current reporting methods require stock assessment scientists to contact the authors of meta-analyses to request that they rerun their analyses to obtain proper Type C results.

In summary, we suggest that in order for scientists to use the results of meta-analysis correctly, analysts should report all three types of results stated above. It is then up to the users to employ Type A results if their stock was included in the meta-analysis but they are not reusing any of the data that were used in the meta-analysis, Type B results if their stock was not included in the meta-analysis, and Type C results if their model shares any of the data that were used in the meta-analysis. The development and use of Type C results may also provide an additional benefit as a tool for comparing the information content of data sets used in meta-analysis.

In this paper, we have focused on the problems of applying meta-analyses to fisheries stock assessments (our specialty). However, we are well aware that this issue, and our proposed solution, are applicable to any scientific field that presents or uses the results of meta-analyses based on time series data, like ecological risk assessments and climate modeling.

Acknowledgements

We would like to thank Dr. A.E. Punt for suggestions on the preparation of this manuscript and Dr. J.N. Ianelli for providing his code to rerun the west coast Pacific ocean perch assessment. Dr. K. Newman, Dr. N. Barrowman, and anonymous reviewers provided valuable comments. Funding for C.V.M.-V. was provided by CAPES/Brazil. I.J.S. was supported by a National Sea Grant/National Marine Fisheries Service fellowship. T.A.B. received funding from the South African National Research Foundation and NMFS grant NA07FE0473.

References

Arnqvist, G., and Wooster, D. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends Ecol. Evol. **10**: 236–240.

Bard, Y. 1974. Nonlinear parameter estimation. Academic Press, New York.

Brett, M.T., and Goldman, C.R. 1996. A meta-analysis of the freshwater trophic cascade. Proc. Natl. Acad. Sci. U.S.A. **93**: 7723–7726.

Dorn, M.W. 2002. Advice on West Coast rockfish harvest rates from Bayesian meta-analysis of stock–recruit relationships. N. Am. J. Fish. Manag. **22**: 280–300.

Efron, B. 1996. Empirical Bayes methods for combining likelihoods. J. Am. Stat. Assoc. **91**: 538–550.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 1995. Bayesian data analysis. Chapman and Hall/CRC, Boca Raton, Fla.

Gurevitch, J., and Hedges, L.V. 1993. Meta-analysis: combining the results of independent experiments. *In* Design and analysis of ecological experiments. *Edited by* S.M. Scheiner and J. Gurevitch. Chapman and Hall, New York. pp. 378–398.

Gurevitch, J., Morrison, J.A., and Hedges, L.V. 2000. The interaction between competition and predation: a meta-analysis of field experiments. Am. Nat. **155**: 435–453.

Harley, S.J., and Myers, R.A. 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. Can. J. Fish. Aquat. Sci. **58**: 1569–1584.

Harley, S.J., Myers, R.A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. **58**: 1760–1772.

Hilborn, R., and Liermann, M. 1998. Standing on the shoulders of giants: learning from experience in fisheries. Rev. Fish Biol. Fish. **8**: 273–283.

Ianelli, J.N. 2002. Simulation analyses testing the robustness of productivity determinations from West Coast Pacific ocean perch stock assessment data. N. Am. J. Fish. Manag. **22**: 301–310.

Liermann, M., and Hilborn, R. 1997. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. Can. J. Fish. Aquat. Sci. **54**: 1976–1984.

Methot, R.D. 1990. Synthesis model: an adaptive framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. **50**: 259–277.

Millar, R.B., and Methot, R.D. 2002. Age-structured meta-analysis of U.S. West Coast rockfish (Scorpaenidae) populations and hierarchical modelling of trawl survey catchabilities. Can. J. Fish. Aquat. Sci. **59**: 383–392.

Myers, R.A., Barrowman, N.J., Hutchings, J.A., and Rosenberg, A.A. 1995. Population dynamics of exploited fish stocks at low population levels. Science (Wash., D.C.), **269**: 1106–1108.

Myers, R.A., Bowen, K.G., and Barrowman, N.J. 1999. Maximum reproductive rate of fish at low population sizes. Rev. Fish Biol. Fish. **56**: 2404–2419.

Myers, R.A., MacKenzie, B.R., Bowen, K.G., and Barrowman, N.J. 2001. What is the carrying capacity for fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod. Rev. Fish Biol. Fish. **58**: 1464–1476.

Osenberg, C.W., Sarnelle, O., Cooper, S.D., and Holt, R.D. 1999. Resolving ecological questions through meta-analysis: goals, metrics, and models. Ecology, **80**: 1105–1117.

Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. **7**: 35–63.