International Journal of Technology Assessment in Health Care, **19:4** (2003), 682–684.

REPLY TO COYLE'S COMMENTS ON 'UNCERTAINTY IN COST-EFFECTIVENESS ANALYSIS: PROBABILISTIC UNCERTAINTY ANALYSIS AND STOCHASTIC LEAGUE TABLES'

Rob M. P. M. Baltussen
Institute for Medical Technology Assessment, World Health Organization

Raymond C. W. Hutubessy
David B. Evans
Christopher J. M. Murray
World Health Organization

Stochastic league tables (SLT) have been introduced as a tool for communicating uncertainty around cost-effectiveness estimates to decision makers (1;9). Coyle (7) criticizes this approach on three specific points. We will discuss them in turn.

ESTIMATING ICER: "RATIO OF MEANS" VERSUS "MEAN RATIO"

The first comment relates to the correct approach to estimate incremental cost-effectiveness ratio (ICER). Coyle correctly argues that the "ratio of means" approach has strong theoretical foundations based on both constrained optimization and individual utility maximization, whereas the "mean ratio" approach lacks this basis and may lead to incorrect estimates of the ICER. We fully agree with Coyle in this matter. However, Coyle argues that the SLT approach applies the "mean ratio" approach in its calculations, and that, therefore, its results would be misleading. This is not correct.

Nonparametric bootstrapping techniques typically involve three levels of analysis. The first level is that of a trail in which data are collected from individual patients receiving alternative therapies. On the basis of these data, the ICER can be estimated as follows:

$$\mathrm{ICER} = \frac{\sum_{\mathrm{I}=1..\mathrm{R}} \mathrm{C}_{\mathrm{Ai}} - \mathrm{C}_{\mathrm{Bi}}}{\sum_{\mathrm{I}=1..\mathrm{R}} \mathrm{E}_{\mathrm{Ai}} - \mathrm{E}_{\mathrm{Bi}}} \tag{1}$$

where C_{Ai} and C_{Bi} and E_{Ai} and E_{Bi} are the costs and effectiveness of individual patients following therapies A and B. A single estimate of ICER is calculated on the basis of the "ratio of means" approach that Coyle correctly favors.

The second level of analysis is the bootstrap procedure for the sake of uncertainty analysis. Repeated samples of cost/effects pairs are taken from the groups of patients who received therapy A and who receive therapy B. The bootstrap replicate of the ICER is again calculated using formula 1. By repeating this procedure a large number of times, the empirical sampling distribution of the ICER is obtained.

The third level of analysis involves the determination of the confidence interval on the basis of the empirical sampling distribution. The simple percentile method can be used to estimate the upper and lower confidence limits for the ICER by taking the $100(\alpha/2)$ and the $100(1-(\alpha/2))$ percentile values of the empirical sampling distribution of the ICER. The mean ICER is estimated as the mean of all the individual replicated ICER and, therefore, could indeed be labeled as a "mean ratio" approach.

The numerical example to illustrate SLT, as presented in Baltussen et al. (2), used parametric bootstrapping rather than nonparametric bootstrapping, but the above principles are maintained. Parametric bootstrapping makes distributional assumption concerning the statistic in question and samples from that distribution, whereas nonparametric bootstrapping samples from an empirical data set. The data as reported in our example, and where was sampled from, represent mean costs and mean effect at the population level. Rather than taking many samples at the individual patient level, the second level of analysis, therefore, involves taking a single sample of mean population costs and effects for therapies A and B to determine an ICER replicate.

The use of "mean ratios" in the third level of analysis is where Coyle's critique seems to focus. However, we believe that this approach is appropriate for the derivation of confidence interval; many other authors seem to agree with this approach, as they have used the same technique in bootstrapping procedures (3–6). The use of "mean ratios" to determine confidence intervals is conceptually very different from the situation that Stinnett and Paltiel (13) describe, in which they are concerned about the consistency and validity of the calculation of single ICER estimates on the basis of trial data. Coyle's critique of SLT, therefore, is not valid.

CONFIDENCE INTERVALS AROUND THE ICER

Coyle argues that the percentile method to estimate confidence intervals may lead to ambiguous results. As an example, he states that two replicates A and B can have the same negative ICER but one may estimate therapy A as both cost saving and more effective and the other may estimate therapy B as more costly but less effective. In our opinion, this can be solved by a simple algorithm in which the negative ICER of therapy A is listed in the top percentiles of ICERs and the negative ICER of therapy B is listed in the bottom percentile of ICERs. In the context of generalized cost-effectiveness analysis (CEA), this option would not even be necessary, as this approach evaluates interventions compared with "doing nothing" and costs, by definition, are positive (10).

Furthermore, in our study, we propose a combination of parametric bootstrapping (also called probabilistic uncertainty analysis) and nonparametric bootstrapping, as first developed by Lord and Asante (11). Coyle argues that we confuse methods of analysis: whereas nonparametric bootstrapping is designed for economic evaluation alongside clinical trials, parametric bootstrapping is designed for economic evaluation based on decision analysis that includes a set of estimates at the population level. We argue that model approaches in CEA often use data from a variety of sources and that a combination of parametric and nonparametric bootstrapping allows the assessment of uncertainty around clinical and population level data in the same analysis.

Coyle proposes an alternative approach by expressing the net monetary benefit of interventions by weighing the expected health gains by a shadow price λ less the expected costs. Cost-effectiveness acceptability curves are based on this approach and inform decision-makers, for a given λ, the probability that an intervention is cost-effective. Recently, cost-acceptability effectiveness curves have been applied to a multiple intervention environment (8), which makes this approach very similar to SLT. Because λ represents the shadow price of a budget constraint, the rank order of interventions following either the cost-effectiveness acceptability curves or SLT approach would be the same. However, the use of a budget constraint in our optimization procedure seems more relevant from the policy perspective than the application of a shadow price, which, if estimated correctly, still requires reference to the budget constraint. Moreover, the use of a fixed shadow price ignores that, as new interventions are funded, other interventions have to be canceled to avoid an ever-growing budget (12).

STOCHASTIC LEAGUE TABLES

The two "fundamental problems" of SLT as mentioned by Coyle have been debated in detail elsewhere (1), with the conclusion that SLT do have a strong theoretical background. We believe that multi-intervention cost-effectiveness acceptability curves as well as SLT are an important contribution to uncertainty analysis and are an apt approach for communicating uncertainty around cost-effectiveness estimates to policy makers.

REFERENCES

1. Baltussen RMPM, Hutubessy RCW, Barendregt JJ, Evans DB, Murray CJL. Formal response to "Determining the optimal combinations of mutually exclusive interventions: A response to Hutubessy and colleagues." *Health Econ*. 2003;12:163-164.
2. Baltussen RMPM, Hutubessy RCW, Evans DB, Murray CJL. Uncertainty in cost-effectiveness analysis: Probabilistic uncertainty analysis and stochastic league tables. *Int J Technol Assess Health Care*. 2002;18:112-119.
3. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation. *Health Econ*. 1997;6:327-340.
4. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. An illustration and application to blood pressure control in Type 2 Diabetes. *Int J Technol Assess Health Care*. 2002;17:69-82.
5. Campbell MK, Torgerson DJ. Bootstrapping: Estimating confidence intervals for cost-effectiveness ratios. *Q J Med*. 1999;92:177-182.
6. Chaudhary MA, Stearn SC. Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trail. *Stat Med*. 1996;15:1447-1458.
7. Coyle
8. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: The role of cost-effectiveness acceptability curves. *Health Econ*. 2001;10:779-787.
9. Hutubessy RCW, Baltussen RMPM, Barendregt JJ, Evans DB, Murray CJ. Stochastic league tables: Communicating cost-effectiveness results to decision makers. *Health Econ*. 2001;10:473-477.
10. Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic uncertainty analysis. *Health Econ*. 1999;8:323-333.
11. Murray CJ, Evans DB, Acharya A, Baltussen RMPM. Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Econ*. 2000;9:235-251.
12. Sendi P, Gafni A, Birch S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Econ*. 2002;11:23-31.
13. Stinnett AA, Paltiel AD. Estimating CE ratios under second-order uncertainty – the mean ratio versus the ratio of means. *Med Decis Making*. 1997;17:483-489.