TRAINING AND CONDUCTING ECONOMIC EVALUATION IN PUBLIC HEALTH

Phaedra S. Corso
College of Public Health, University of Georgia, United States of America
pcorso@uga.edu

Economic evaluation (EE) refers to applied analytic methods used to identify, measure, value, and compare the costs and consequences of prevention and treatment strategies. Economic evaluation provides important information to assist policy makers who are faced with making funding decisions with scarce public health resources. However, the application and understanding of the statistical tools used to conduct economic evaluation has been limited in the context of evaluation training. This paper provides an introduction to three methods for conducting economic evaluation of public health programs: cost-effectiveness analysis, cost-utility analysis, and benefit-cost analysis. Each method is discussed in turn, with a special emphasis on the statistical tools used in each, and the uses and misuses of the methods in the policy arena.

INTRODUCTION TO ECONOMIC EVALUATION METHODS

Incorporating principles of economics is a natural extension of evaluation training and research, where the scarcity of societal resources requires policy makers to contemplate the economic principles of opportunity costs and efficiency. Opportunity costs refer to the foregone benefits of the next best alternative use of resources. Policy makers invariably have to make difficult tradeoffs, such as those between interventions that focus on prevention versus treatment, the young versus old, and urban versus rural populations. Economic efficiency is achieved when benefits of an intervention are maximized, while at the same time minimizing opportunity costs (Donaldson et al., 2002). How to define and measure benefits and costs in this context depends upon the perspective, as in who pays and who benefits. As such, information on efficiency provides answers to the following normative questions: What interventions should be used to maximize outcomes of interest? What is the proper mix of interventions to provide to a population?, and What is the best way to allocate public resources across the population?

Health economics, an evolving sub-discipline of economics, focuses on the supply and demand of health care in general and the impact of health care as a market commodity on a population (Santerre & Neun, 2000; Culyer & Newhouse, 2000; Maynard & Kanavos, 2000; and Williams, 1987). The principles of health economics can be applied to public health more specifically (Kenkel, 2000), by focusing on economic determinants of health-related behaviors, health status, and health disparities; the valuation and assessment of costs associated with disease, injury, disability, and premature mortality; and the economic evaluation of healthcare, health promotion, and prevention technologies and strategies.

Sometimes referred to as the non-traditional application of health economics (Hoch et al., 2002), economic evaluation is characterized by experimental and quasi-experimental methods and is used to produce information that compares intervention outcomes to costs. The main analytic tools used to conduct economic evaluations include cost-effectiveness analysis (CEA), cost-utility analysis (CUA) as a special variant of CEA, and benefit-cost analysis (BCA). Primers on how to use these methods specifically for public health interventions have been described in detail elsewhere (Gold et al., 1996; Haddix et al., 2003; Drummond et al., 2005).

In a CEA, an intervention's costs are compared to its health benefits as measured in natural units of health, such as cost per case prevented. In a CUA, an intervention's costs are compared to the impact of the intervention on health outcomes that are expressed as a combined measure of length of life (survival) and health-related quality of life (HRQoL). Quality-adjusted life years (QALYs) and disability-adjusted life years (DALYs) are two commonly used indices of health used in a CUA. Either the QALY or the DALY can be described as a measure of an intervention's impact on life expectancy adjusted for a person's quality of life for that remaining life expectancy. For example, if the quality of life, on a scale where 0 is death and 1.0 is perfect health, of a person being blind in both eyes is equal to 50% of perfect health, then a person with 35 years of remaining life expectancy can expect only 17.5 years of quality-adjusted life expectancy. Assessing quality of life weights from populations and discerning the tradeoffs they would make between improving

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

HRQoL or increasing life expectancy, has at its foundation core aspects of economic theory, where people are asked to make tradeoffs between health and life.

In a BCA, an intervention's costs are compared to its health benefits as measured in monetary terms, such that the summary measure is presented as net benefits (total benefits minus total costs). A monetary assessment of burden is typically referred to as a cost-of-illness (COI) analysis, which estimates the costs of health outcomes associated with disease or injury. Most costs incorporated into a COI analysis include direct medical costs (e.g., hospitalizations, MD visits), direct nonmedical costs (e.g., travel costs, childcare expenses), and productivity losses associated with morbidity and premature mortality (e.g., days missed from work). Contingent valuation is another technique for monetizing benefits in a BCA, which includes one approach that measures society's willingness to pay for reductions in morbidity or mortality risks (Mitchell & Carson, 1989). For example, when people in a national survey conducted in the United States were asked to state their willingness to pay for reductions in gun violence, their answers suggest that the benefits of reducing gun violence are worth approximately $100 billion per year (Cook & Ludwig, 2002).

STATISTICAL ISSUES IN ECONOMIC EVALUATION METHODS

There are several statistical challenges in the application of economic evaluation methods in public health. The first issue relates to costs and their characteristics that make them difficult to model in a CEA. Medical costs present an analytic challenge because of their distribution, which is right (i.e., positively) skewed with a high frequency of zero costs (Briggs & Gray, 1999). Since incurred costs cannot be negative, they are bounded by zero but have no upper bound; thus, there is a high frequency of those with zero claims costs and a low frequency of those with high claims costs. The few with high claims costs represent a small portion of the population or sample with severe adverse events with costly treatments.

Many statistical solutions have been proposed over the history of economic evaluation to deal with the challenges of dealing with medical costs. Specifically, utilizing the median as a measure of central tendency, log-transforming cost data, and various parametric and non-parametric methods. Utilizing the median is one option since the median is less affected by outliers; however, economic evaluation requires a per person cost that can only be derived from the mean (Zhou & Tu, 1999). Another proposed solution is transforming cost data to a logarithmic scale, with the goal that the log-transformed data will be normally distributed, and then conducting a one-part regression model. Unfortunately, this method also presents challenges. First, transformed data is difficult to interpret, particularly for the purposes of economic evaluation, and the inferential statistics estimate the log costs, a statistical relationship that is not the same for the original cost data if the assumption of heteroskedasticity is violated. Secondly, cost data is often so severely skewed that the data remains non-normally distributed even after the log-transformation (Briggs & Gray, 1999; Zhou & Liang, 2006).

Parametric methods, such as analysis of variance (ANOVA), analysis of covariance (ANCOVA), Tobit model, and Heckman's selection model have been proposed. However, each of these methods has been deemed inappropriate for analyzing medical claims data (Duan et al., 1983). The best-practice for analyzing cost data with regards to the distribution concerns was first proposed by Duan et al. (1983). This 2-part regression model includes two stages in which the first stage uses a probit equation for the dichotomous event of having zero or positive medical claims costs, and the second stage uses a linear regression model for non-zero values on the log-scale. Finally, the predicted cost of an individual is the product of the expected cost and the probability of having costs. As noted previously, there are concerns with using costs on a log-scale. Thus, a smearing factor must be used to accurately estimate mean costs from the estimated mean log-costs. Per the distribution of the errors, either a parametric or non-parametric smearing factor should be used (Cooper, Sutton, Mugford, & Abrams, 2003). Although the two-part model presents many advantages, the four-part model is arguably an even better method. The four-part model uses the above method to separate inpatient users from non-impatient users. Thus the first two equations separate the population into three groups: 1) nonusers, 2) inpatient users, and 3) ambulatory-only users.

Another statistical challenge in economic evaluation, in CEA specifically, is the need to establish a confidence interval for the summary measure of the CEA, which is the cost-

effectiveness ratio. The ratio includes costs in the numerator and the effects in the numerator, cases prevented or quality-adjusted life years (QALYs) saved, for example. Because the summary measure is a ratio, establishing a confidence interval is statistically challenging. For example, assuming person-level data on an intervention's costs and effects, one can calculate a 95% confidence interval (CI) separately for each. However, combining the 95% CI for costs and the 95% CI for effects gives a 90% confidence box rather than the 95% CI because $(1-\alpha)2 = 90\%$. But this may overestimate the actual size of the confidence limits. Of course, there is evidence that in many situations, there is a strong correlation between costs and effects, and that the direction of the correlation may not always be obvious. It could be that people who live a long time utilize more healthcare resulting in higher costs, or that people who have terrible health outcomes have higher medical costs, or that people who die immediately have lower costs. If costs and effects follow a joint normal distribution, the joint cost and effect density function might be elliptical in shape. But adjusting the confidence limits to include correlation requires knowledge of the covariance between the costs and effects. Although there is no exact formula for estimating the variance of costs/effects, one can approximate it through Taylor's expansion, which is a function of the variance of each variable, the covariance of the variables, and the rate of change of each variable with respect to the other. Although Taylor's expansion creates a cost and effect confidence limits that are symmetric in cost effectiveness space (a confidence ellipse), it assumes that the ratio itself is normally distributed (which we know often it is not). Further, it does not work well when there is a high probability of an effect or cost being close to zero. And if you observe that the distribution of the cost-effectiveness ratio in the population sample is far from normal, Taylor's expansion will be unlikely to provide accurate estimates (van Hout et al., 1994). To correct for this, researchers began to use Feiller's Theorem as a way to calculate a more accurate confidence interval (Willan & O'Brien, 1996).

Although initial efforts were concentrated on providing confidence intervals for cost-effectiveness ratios, recent efforts have focused on the concept of incremental net benefits, NB, (Willan, 2001). NB is defined as $(\Delta\text{effects} - \Delta\text{costs})/\lambda$, where λ is the pre-determined amount society is willing to pay per unit of effectiveness. In this way, the statistical properties of the NB measure become more manageable because the function is linear and regression diagnostics can be used to determine robustness of estimates. From this method arose the application of the cost-effectiveness acceptability curve, CEAC (Fenwick et al., 2004). A CEAC is a visual representation of the probability that the intervention is cost-effective, for all potential values of the pre-determined funding ceiling or willingness to pay per unit of effectiveness, λ.

CONCLUSION

Economic evaluation is a valuable tool for policy makers faced with making difficult policy decisions with scarce public resources. But the complexities of the data used in economic evaluations and the summary measures used for comparisons across interventions pose a number of statistical challenges. Important strides have been made in the last few decades to deal with these statistical challenges. Training in these statistical techniques is therefore an important part of any evaluation curriculum that intends to focus on economic evaluation as a priority area.

REFERENCES

Briggs, A. H., & Gray, A. M. (1999). Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment, 3*(2), 1-134.

Cook, P. J., & Ludwig, J. (2002). The costs of gun violence against children. *The Future of Children, 12*(2), 87-99.

Cooper, N. J., Sutton, A. J., Mugford, M., & Abrams, K. R. (2003). Use of Bayesian Markov Chain Monte Carlo methods to model cost-of-illness data. *Medical Decision Making, 23*(1), 38-53.

Culyer, A. J., & Newhouse, J. P. (Eds.) (2000). *Handbook of Health Economics, volume 1b.* Amsterdam: Elsevier Science B.V.

Donaldson, C., Mugford, M., & Vale, L. (Eds.) (2002). *Evidence-based Health Economics*. London: BMJ Books.

Drummond, M. E., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the Evaluation of Health Care Programmes, third edition*. Oxford: Oxford Medical Publications.

Duan, N. (1983). Smearing estimate: a non-parametric retransformation method. *J Am Stat Assoc, 78*, 605–610.

Fenwick, E., Obrien, B., & Briggs, A. (2004). Cost-effectiveness acceptability curves – facts, fallacies, and frequently asked questions. *Health Economics, 13*(5), 405–415.

Gold, M. R., Siegel, J. E., Russel, L. B., & Weinstein, M. C., eds. (1996). *Cost-effectiveness in Health and Medicine*. New York: Oxford University Press.

Haddix, A. C., Teutsch, S. M., & Corso, P. S., (Eds.) (2003). *Prevention Effectiveness: An Introduction to Decision Analysis and Economic Evaluation, second edition.* New York: Oxford University Press.

Hoch, J. S., Briggs, A. H., & Willan, A. R. (2002). Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economic, 11*, 415–30.

Kenkel, D. S. (2000). Chapter 31: Prevention. In A. J. Culyer & J. P. Newhouse (Eds.) *Handbook of Health Economics, volume 1b*. Amsterdam: Elsevier Science B.V.

Maynard, A., & Kanavos, P. (2000). Health economics: an evolving paradigm. *Health Economics, 9*, 183-90.

Mitchell, R. C., & Carson, R. T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington D.C.: Resources for the Future.

Santerre, R. E., & Neun, S. P. (2000). *Health Economics: Theories, Insights, and Industry Studies*. Orlando: Harcourt Brace & Company.

Van Haut, B. **A.,** Al, M. J., Gordon, G. S., & Rutten, F. F. H. (1994). Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics, 3*, 309-319.

Willan, A. R., & O'Brien, B. J. (1996). Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Economics, 5*, 297–305.

Willan, A.R. (2001). Analysis, sample size and power for estimating incremental net health benefit from clinical trial data. *Controlled Clinical Trials, 22*, 228 –237.

Williams, A. (1987). Health economics: the cheerful face of a dismal science. In A. Williams (Ed.) *Health and Economics*. London: Macmillan.

Zhou, X.-H., & Liang, H. (2006). Semi-parametric single-index two-part regression models. *Computational Statistics & Data Analysis, 50*(5), 1378-1390.

Zhou, X. H., & Tu, W. (1999). Comparison of several independent population means when their samples contain log-normal and possibly zero observations. *Biometrics, 55*(2), 645-651.