# $A_N$ APPROACH TO EVALUATING HEALTH OUTCOMES

### Les Irwig, Associate Professor, Department of Public Health, University of Sydney

ealth services often seem to expend enormous resources without a clear picture of the outcomes achieved. With rapid improvements in computer hardware and software, it is enticing to use health service data to generate information about the outcomes of health care. Evaluating health outcomes, which also encompasses measuring the process of care, has enormous potential for improving health. To achieve maximum gains, we need a clear framework to guide this work and its interpretation. In this article I propose a framework based on four major questions<sup>1</sup>. By assessing which of the questions is being addressed and whether the methods are appropriate for that question, you may discover ways in which your own work or your interpretation of others' findings could be improved.

### QUESTION 1: WHAT IS THE RIGHT THING TO DO? (EFFECTIVENESS)

This is the most critical of the questions. It is concerned with deciding policy whereas the remaining questions are concerned with quality assurance of its implementation. Without reasonable certainty about the effectiveness of the policy, addressing later questions may do a disservice by encouraging uniform application of a policy which may be harmful. Which methods should be used to ensure Question 1 is answered adequately?

1 Choose the right study design: randomised trials Health policies or practice guidelines should be based on randomised trials and not on observational studies or opinion<sup>2</sup>. This is contrary to much of the health outcomes literature which suggests the efficacy of interventions can be ascertained from sophisticated analysis of outcome variability between different sources of care, often using routinely collected data3. I do not think such data can be used to test interventions3. The effects of therapy are often small and easily overshadowed by the selection biases inherent in observational studies. Selection biases are likely to be much greater in studies of interventions than in studies of environmental, lifestyle or occupational exposures. This occurs because both physicians and patients are deliberately choosing interventions on the basis of patient characteristics, not all of which will be explicit. Moreover, it will not be possible to adjust adequately for important predictors of outcome, such as case-mix, because they are imperfectly measured4.

## 2. Measure outcomes relevant to patients: quality of life and survival

Quality of life has been ignored for too long because it is difficult to measure. However, rapid advances are being made in the methodology which should enable its wider application<sup>5</sup>.

Apart from their quality of life, patients will be interested in their survival. We therefore should examine survival, or its proxy, all-cause mortality. Mortality from the disease at which the intervention is aimed may be reduced at the expense of an increase in mortality from other causes6. For decision making, we need to shift away from measures of relative improvement, such as relative risk, to measures of absolute improvement, such as risk difference, or its inverse the number of people who need to be detected/ treated to prevent one death or illness. For example, one needs to screen 23,000 women aged 35-44 for high blood cholesterol and treat 4,500 of them for five years to prevent one death7. Moreover, this calculation is based on the assumption of a 20 per cent reduction in mortality from cholesterol intervention in all age-sex groups, which seems far too optimistic an estimate8.

### 3. Meta-analyse all the trials

Meta-analysing all randomised trials will provide an estimate of the efficacy of an intervention with narrower confidence intervals than any of the individual trials and may give sufficient power to examine efficacy in different subgroups of patients<sup>668</sup>. Continuously updating meta-analyses and making results readily accessible in computer format will help health practitioners make rational decisions and should form the basis for clinical practice guidelines<sup>10,11,23</sup>.

This approach assumes the existence of randomised trials. I expect our hands will be full with addressing policy formulation and quality assurance for common conditions for which trials have been done. Doing randomised trials for important conditions where none has been done will be made easier if clinicians and patients become more accepting of the idea that there is uncertainty about the best treatment and more willing to enter trials. Trials should be made easier to incorporate into clinical practice by simplifying inclusion criteria and consent procedures, reducing the data requirements and changing funding arrangements<sup>11</sup>.

### QUESTION 2: DID WE DO THE RIGHT THING? (APPROPRIATENESS OF CARE)

Assessing appropriateness of care should be based on measuring the process of care (practice) against the 'gold standard' of the evidence-based policy generated in response to Question 1. It will have the greatest impact if the condition is common, actual practice varies considerably from the policy and clinical opinion leaders are keen to collaborate<sup>12</sup>. Commonly, appropriateness is measured by examining what proportion of people on whom a procedure was done should have had it<sup>13</sup>. This approach does not indicate how many patients got the procedure when they should have. Knowing how many people who needed the procedure did not get it is clearly as important as knowing how many got the procedure inappropriately. Methods for monitoring appropriateness of care should include both components and take steps to deal with error in the measurement of appropriateness<sup>14</sup>.

## QUESTION 3: DID WE DO THE RIGHT THING RIGHT? (PERFORMANCE)

The assessment of performance includes three main components: technical performance, patient satisfaction with the process of care, and efficiency. Technical performance often requires assessments which are more detailed than can be ascertained from systems set up to monitor appropriateness. Examples include the proportion of 'lumpectomies' in which the breast cancer has been removed with an adequate margin. Patient satisfaction includes several dimensions concerned with the process of care<sup>15</sup>. Efficiency is concerned with whether care was carried out in the most streamlined way, at least cost and without delays.

# QUESTION 4: DID IT HAVE THE RIGHT RESULT? (OUTCOME)

A flow diagram of the expected sequence of events of an intervention helps in deciding which, if any, outcomes to measure<sup>18</sup>. Long-term outcomes such as five-year survival are often not worth measuring as it may not be possible to disentangle the effects of intervention from other sources of variability in outcome. Furthermore, distant events will not provide timely feedback for monitoring quality. Surrogates (indicators) of outcome should be used only if there is good evidence that they are an important intermediate step

### **Evaluating health outcomes**

#### Continued from page 135

between the intervention and long-term outcome<sup>17,18,19</sup>. Surrogates may suggest mortality will be reduced when the opposite occurs<sup>20,21</sup>. Outcomes or their surrogates need to be unambiguously interpretable as an effect of intervention. This may occur if the effects are large, immediate, or rarely occur in the absence of the intervention, for example post-operative morbidity and mortality. Measuring immediate adverse events such as post-operative mortality may be useful for weighing up long-term benefits, estimated from randomised trials, against the risks in your patients<sup>22</sup>.

In instances where outcomes will not be unambiguously interpretable as an effect of the intervention, quality assurance should be based on measuring the appropriateness and performance of the intervention, and we will need to assume it will have the benefit demonstrated in randomised trials.

Berwick DM. Health services research and quality of care – Assignments for the 1990s. *Med Care* 1989; 27:763-71.
Sox HC Jr, Woolf SH. Evidence-based practice guidelines from the US

Sox HC JF, Wooli SH. Evidence-Jased practice guidelines from the US Preventive Services Task Force. JAMA 1993; 269:2678.
Cotton P. Doctors are asked: Does it work? Determining more good than harm is not easy. JAMA 1993; 270:153-8.
Thomas D, Stram D, Dwyer J. Exposure measurement error: Influence

on exposure-disease relationships and methods of correction. Ann Rev Publ Health 1993; 14:69-93.

5. Jenkins CD. Assessment of outcomes of health intervention. Soc Sci Med 1992; 35:367-75.

6. Barratt A, Irwig L. Is cholesterol testing/treatment really beneficial? Med J Aust 1993; 159:644-7.

7. Khaw KT, Rose G. Cholesterol screening programmes: How much potential benefit? *Brit Med J* 1989; 299:606-7. 8. Davey Smith G, Song F, Sheldon TA. Cholesterol lowering and

mortality: the importance of considering initial level of risk. Brit Med J 1993; 306:1367-73.

9. L'Abbe KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. Ann Intern Med 1987; 107:224-33.

Antriach Inder, Jane Joseph J. Let Gold and J. Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New Eng J Med* 1992; 327:248-54.

11. Chalmers TC. The need for early randomization in the development of new drugs for AIDS. J of Acquired Immune Deficiency Syndrome. 1990; 3(Supp):S10-S15. 12. Jencks SF, Wilensky GR. The health care quality improvement.

initiative - A new approach to quality assurance in medicare. JAMA 1992; 268:900-3

13. Leape LL, Hilborne LH, Park RE, Bernstein SJ, Kamberg CJ, Is beape LL, minorine LL, Fark RB, Bernstein SS, Rainberg CS, Sherwood M, Brook RH. The appropriateness of use of coronary artery bypass graft surgery in New York State. JAMA 1993; 269:753-60.
Phelps CE. The methodologic foundations of studies of the appropriateness of medical care. New Eng J Med 1998; 329:1241-5. 15. Cockburn J, Hill D, Irwig LM, DE Luise T, Turnbull D, Schofield P. Development & validation of an instrument to measure satisfaction of

participants at breast screening programme. European J of Cancer 1991; 27:827-30.

Iroig L. Epidemiology and health care evaluation. In Public Health Association – Health Care Evaluation in Australia 1989.
Freedman LS, Schatzkin A. Sample size for studying intermediate

endpoints within intervention trials or observational studies. Am J *Epidemiol* 1992; 136:1148-59. 18. Morrison AS. Intermediate determinants of mortality in the

evaluation of screening. Int J Epidemiol 1991; 20:642-50.

19. Day NE. Surrogate measures in the design of breast screening trials. In: Eds: Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PC. Cancer Screening. New York: Cambridge University Press, 1991; 391-

403. 20. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW and the CAST Investigators. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. *New Eng J Med* 1991; 324:781-8.

21. Teo KK, Yusuf S, Furberg CD. Effects of prophylactic antiarrhythmic drug therapy in acute myocardial infarction - An overview of results from randomized controlled trials. *JAMA* 1993; 270:1589-95. 22. Marwick C. Federal agency focuses on outcomes research. *JAMA* 1993; 270:164-5.

### What is the difference?

#### Continued from page 134

clinical indicators, clinical audit, peer review, utilisation review, best practice and managed care. The health outcomes approach and quality assurance initiatives have several elements in common.

First, they have a common purpose: the continual improvement of health services. They all involve a reiterative cycle of evaluation, adjustment of services (when necessary), and re-evaluation, leading to continual improvement.

Second, their evaluative processes are based on specified indicators - indicators of structure or process quality, or indicators of outcome.

Third, they are designed to be integrated into the work ethos and practices of all relevant personnel.

Fourth, they espouse an intention to promote improvements through positive measures rather than recrimination. They seek to respect the professional integrity of individual providers, especially clinicians, and they involve service providers in the evaluation and improvement of their own services.

Finally, implicitly or explicitly they advance the notion of customer focus. This involves identifying the customer, for whom any given service is undertaken, and seeking to provide optimal fulfilment of the customer's requirements. The customer may be external to the organisation or within it.

### QUALITY, OUTCOMES, AND COSTS

Traditionally, quality assurance initiatives have been concerned with the way services are provided without systematic consideration of costs. However, it is now recognised that quality of care cannot be improved without regard for cost<sup>13</sup>. Information on both outcomes and costs is needed to ensure optimal use of resources. The health outcomes approach emphasises that decisions must be based on the health outcomes of services as well as costs.

Frommer M, Rubin G, Lyle D. The NSW health outcomes program. NSW Pub Health Bull, 1992; 3:135,137.
Australian Council on Healthcare Standards. The ACHS accreditation

guide, Sydney: Australian Council on Healthcare Standards, 1993. 3. Duckett SJ, Coombs EM, Schmiede AM. Hospital accreditation in New South Wales, Kensington: School of Health Administration, University of New South Wales, 1980.
Best J. A review of the New South Wales Department of Health

Quality Assurance Program. Sydney: Diagnosis, 1989. 5. Scarfe GC, Weaver CJ, Duckett SJ, Schmiede AM. Quality assurance

Jamison JH. Report of the commission of inquiry into the efficiency and administration of hospitals. Carberra: AGPS, 1981.

7. Commonwealth Department of Health. Health care and the consumer. Canberra: AGPS, 1985.

8. Renwick M, Harvey R. Quality assurance in hospitals. Canberra: Australian Institute of Health, 1989.

9. Berwick DM. Controlling variation in health care: A consultation from

Walter Shewhart. Med Care, 1991; 29:1212-1225. 10. Kritchevsky SB, Simmons BP. Continuous quality improvement. Concepts and applications for physician care. JAMA, 1991; 266:1817-1823

11. Kaluzny AD, McLaughlin CP, Simpson K. Applying total quality management concepts to public health organisations. *Pub Health Reports*, 1992; 107:257-264.

 Donabedian A. Evaluating the quality of medical care. Milbank Memorial Fund Quarterly: Health and Society, 1966; 44:166-206.
Braithwaite J, Westbrook JI. For richer for poorer, in sickness and in health: international issues of quality and cost. Australian Hospital Association, Canberra, 1992.