

20. Scannapieco M, Connell-Carrick K. Focus on the first years: correlates of substantiation of child maltreatment for families with children 0 to 4. *Child Youth Serv Rev* 2005; **27**: 1307–1323.
21. Flaherty EG, Thompson R, Litrownik AJ, et al. Effect of early childhood adversity on child health. *Arch Pediatr Adolesc Med* 2006; **160**: 1232–1238.
22. Royal College of Paediatrics and Child Health. *Responsibilities of doctors in child protection cases with regard to confidentiality*. London: RCPCH, 2004.
23. Fairclough N. *Discourse and social change*. Cambridge: Polity Press, 1992.
24. Royal College of General Practitioners. *Grasping the nettle*. www.rcgp.org.uk/PDF/Corp_grasping_the_nettle.pdf (accessed 1 Dec 2009).
25. Blackwell PL. The Solihull approach resource pack. *Infant Ment Health J* 2004; **25**: 74–76.
26. Bakermans-Kranenburg MJ. Less is more: meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bull* 2003; **129**: 195–215.
27. Polansky NA, Chalmers MA, Buitenen EW, Williams DP. *Damaged parents; an anatomy of child neglect*. Chicago, IL: University of Chicago Press, 1981.
28. Kim J. Type-specific intergenerational transmission of neglectful and physically abusive parenting behaviors among young parents. *Child Youth Serv Rev* 2009; **31**: 761–767.
29. The Scottish Government. A guide to getting it right for every child. Edinburgh: The Scottish Government, 2008. <http://www.scotland.gov.uk/Topics/People/Young-People/childrenservices/girfec/publications>
30. Wolock I, Horowitz B. Child maltreatment as a social problem: the neglect of neglect. *Am J Orthopsychiatry* 1984; **54**: 530–543.

DOI: 10.3399/bjgp10X482031

ADDRESS FOR CORRESPONDENCE

Philip Wilson

Section of General Practice and Primary Care, Faculty of Medicine, University of Glasgow, 1 Horselethill Road, Glasgow, G12 9LX.

E-mail: p.wilson@clinmed.gla.ac.uk

NICE work: how NICE decides what we should pay for

The National Institute for Health and Clinical Excellence (NICE) is responsible for providing guidance on the promotion of good health and the prevention and treatment of ill health in the UK. Fundamental to the decision-making process is the need to make recommendations based on the best available evidence with input from all stakeholders in a transparent and collaborative manner.¹ Health technologies considered by NICE include pharmaceuticals, medical devices, diagnostic techniques, surgical procedures, other therapeutic technologies, and health promotion activities.

The budget for the NHS is fixed by a political process and decisions about which health technologies to recommend are based on a combination of clinical effectiveness and cost-effectiveness, taking into account the opportunity cost of technologies displaced by new, generally more expensive, technologies. Evidence regarding clinical effectiveness often comes from randomised controlled trials (RCTs) because they have high internal validity. However, RCTs generally estimate efficacy in a much narrower population than the target population and may be conducted over much shorter time periods relative to how long the health technology will be applied in clinical practice. Consequently, when assessing the relative cost-effectiveness of two or more health technologies it is usual to make decisions based on an economic model to capture the expected lifetime costs and benefits.

Drug regulatory authorities define a

confirmatory trial as an adequately controlled trial in which the hypotheses are stated in advance.² Such trials are predominantly designed and analysed using a frequentist (or classical) approach to statistics in which a hypothesis to be tested is specified (that is, the null hypothesis), the sample size necessary to generate sufficient information to reject the null hypothesis if it is false is determined, and the strength of evidence against the null hypothesis is calculated (that is, the *P*-value).

Trials are typically designed either as superiority or non-inferiority trials. In superiority trials, a minimum effect that has clinical relevance is specified. In non-inferiority trials, a non-inferiority margin is specified such that if the effect of the new intervention were no worse than this, then the conclusion would be that it was clinically non inferior to the standard. At the design stage, the sponsor's risk, or power of the test, is the probability that we will reject the null hypothesis if the true treatment effect equals the effect size of interest. Power is conventionally set at 80% or 90%, which means that the sponsor is prepared to accept probabilities of 0.20 and 0.10 of not rejecting the null hypothesis respectively.

At the analysis stage the *P*-value, or significance level, is the regulator's risk of wrongly approving a drug as efficacious, usually set at 0.05 (5%) or less, and represents the probability of obtaining a result at least as extreme as the one observed on the assumption that there is no true difference. Any effect can be shown to

be statistically significant given enough information, although such differences may not be clinically relevant.

The RESPECT trial team³ designed the RESPECT trial to estimate the effect of pharmaceutical care for older people, shared between GPs and community pharmacists in the UK, relative to usual care. The primary outcome measure was the UK Medication Appropriateness Index (UK-MAI). Following conventional trial design considerations for a superiority trial, they defined the treatment effect to be detected as a difference of 0.4 of a standard deviation in UK-MAI, a significance level of 5%, and a power of the test of 80%. At the analysis stage, the *P*-value can be regarded as a decision rule and, because the *P*-value for the effect of the intervention was estimated as 0.402, the authors concluded that 12 months of pharmaceutical care delivered by community pharmacists to older people did not affect the appropriateness of repeat medication as assessed by the UK-MAI.

However, absence of evidence is not the same as evidence of absence and it is helpful to supplement *P*-values with 95% confidence intervals which provide a range of plausible values for the true treatment effect. Unfortunately, the authors found it necessary to transform the data prior to their analysis and, as a consequence, an appraisal committee would find it difficult deciding whether important treatment effects are plausible on the original scale.

Most submissions to NICE will usually include strong evidence for the clinical

effectiveness of the new health technology, such as a statistically significant hazard ratio for progression-free or overall survival, before providing evidence to support a claim of cost-effectiveness. RESPECT is unusual in this regard because it failed to detect a statistically significant benefit of pharmaceutical care. Nevertheless, in accordance with the appraisal process, the RESPECT trial team⁴ present an analysis of the cost-effectiveness of pharmaceutical care.

The decision rule for cost-effectiveness is based on the incremental cost-effectiveness ratio (ICER), which presents the ratio of the mean differences in costs to the mean difference in benefit. When deciding whether to adopt a health technology for the target population, it is the mean cost and mean effectiveness over the whole population that is relevant because the decision applies to the whole population. The NHS will have to pay a cost equal to the total of all the costs for individual patients under the chosen health technology and, when expressed on a per-patient basis, this is the population mean cost. For a similar reason, the per-patient mean effectiveness measures the benefit that the NHS obtains for that cost in terms of improved health.

NICE has a preference for expressing health gain in terms of QALYs (quality adjusted life years) because the health technology is expected to have an effect on survival as well as health-related quality of life, and this was the outcome measure used by the RESPECT trial team in their cost-effectiveness analysis. The UK-MAI is unlikely to be an appropriate outcome measure for a cost-effectiveness analysis because it is doubtful that it is linear in the sense that a decision maker would be prepared to pay twice as much for two units of benefit as it would for one unit of benefit. Thus, the RESPECT trial team's studies^{3,4} come to different conclusions regarding clinical effectiveness and cost-effectiveness, not least because they are addressing different objectives based on different outcome measures.

Appraisal committees have to make decisions based on the available evidence because not doing so is equivalent to accepting the current health technology. They do not use a precise ICER threshold, although above an ICER of £30 000 per

QALY an appraisal committee will consider other factors, including the decision uncertainty. Decision uncertainty is represented by the cost-effectiveness acceptability curve (CEAC), which plots the probability of cost-effectiveness at different values of the ICER given the available evidence.⁵

The CEAC is a Bayesian concept in which inferences are based on a combination of sample data and prior information. This is more useful than a *P*-value because we are no longer interested in testing a simple hypothesis but want to make inferences based on all available information. A willingness-to-pay of zero for a QALY corresponds to making inferences about which is the cheaper health technology; a willingness-to-pay tending to infinity corresponds to making inferences about which is the most effective health technology; values in between correspond to making inferences based on a trade-off between costs and effectiveness.

In their trial-based analysis, the RESPECT trial team showed that there is good evidence to suggest that pharmaceutical care is more expensive but more effective than usual care. At their estimated ICER of £10 000, an appraisal committee would normally accept the health technology as being cost-effective because it is smaller than £30 000 per QALY and would be regarded as an efficient use of resource, subject to the extent of the decision uncertainty and any concerns about the quality of the analysis and sensitivity of the results to particular assumptions. The authors did not consider lifetime costs, although they suggest that these are unlikely to be a major concern because most costs were incurred early. In addition, the authors did not consider lifetime benefits, which may be of particular interest given the relatively small incremental QALY estimated in the trial.

The authors presented subgroup analyses which suggest that the ICER is greater for older patients with more repeat prescriptions. For example, in 90-year-old patients with 15 repeat prescriptions, the ICER is estimated at £35 195, which would raise some doubt as to whether pharmaceutical care is a cost-effectiveness strategy for this population. Furthermore, pharmaceutical care would probably not be approved in this population even when

referring the estimated ICER to the criteria for end of life, because the procedure does not offer an extension to life of at least 3 months.

The authors concluded that their results are uncertain and that further research into the long-term benefits of pharmaceutical care might be worthwhile. Of course, future research will itself incur a cost and the value of conducting further research to resolve the uncertainties should be evaluated using expected value of information analysis. The prospectively designed RESPECT trial suggests that a statistically non-significant effect of pharmaceutical care on UK-MAI translates into a cost-effective increase in QALYs. Although an appraisal committee would normally accept health technologies with similar ICERs as being cost-effective, some members may raise questions concerning the 'mechanism of action' of pharmaceutical care given that differences in UK-MAI smaller than judged to be clinically relevant by the authors appear to translate into cost-effective improvements in QALYs gained.

John W Stevens,

Director, Centre for Bayesian Statistics in Health Economics, SchARR, University of Sheffield.

Provenance

Commissioned; not peer reviewed.

REFERENCES

1. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal*. NICE, 2004. http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf (accessed 27 Nov 2009).
2. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. *Statistical principles for clinical trials*: E9. ICH 1998. <http://www.ich.org/LOB/media/MEDIA485.pdf> (accessed 27 Nov 2009).
3. RESPECT Trial Team. Effectiveness of shared pharmaceutical care for older patients: RESPECT trial findings. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp09X473295.
4. RESPECT Trial Team. Cost-effectiveness of shared pharmaceutical care for older patients: RESPECT trial findings. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp09X482312.
5. O'Hagan A, Stevens JW, Montmartin J. Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio. *Pharmacoeconomics* 2000 17: 339–349.

DOI: 10.3399/bjgp10X482040

ADDRESS FOR CORRESPONDENCE

John W Stevens

Centre for Bayesian Statistics in Health Economics, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
E-mail: J.W.Stevens@sheffield.ac.uk