

What is already known on this topic

Studies have suggested that migraine may be a risk factor for stroke, but results from these studies have been inconsistent

What this study adds

Migraine both with and without aura may be an independent risk factor for ischaemic stroke

This risk is higher among oral contraceptive users and younger adults (<45 years)

define ischaemic stroke, including duration of symptoms of at least 24 hours, as well as confirmation of diagnosis by brain imaging or autopsy. Finally, we could not infer from the studies a temporal relation between the onset of migraine and the diagnosis of stroke.

Conclusion

Data from observational studies suggest that migraine may be a risk factor in developing stroke. More studies are needed to explore the mechanism of this potential association. The risk of migraine among users of oral contraceptives must be further investigated.

Contributors: See bmj.com

Funding: No specific funding. ME is funded by a Canadian Institutes of Health Research postdoctoral fellowship award. AS is supported by the National Institutes of Health and the Department of Veterans Affairs through the Parkinson Disease Research Education and Clinical Center grant.

Competing interests: None declared.

Ethical approval: Not needed.

- Silberstein SD. Migraine. *Lancet* 2004;363:381-91.
- Tzourio C, Tehindranarivelo A, Iglesias S, Alperovitch A, Chedru F, d'Anglejan-Chatillon J, et al. Case-control study of migraine and risk of ischaemic stroke in young women. *BMJ* 1995;310:830-3.
- Carolei A, Marini C, De Matteis G. History of migraine and risk of cerebral ischaemia in young adults. *Lancet* 1996;347:1503-6.

- Woods RP, Iacoboni M, Mazziotta JC. Brief report: bilateral spreading cerebral hypoperfusion during spontaneous migraine headache. *N Engl J Med* 1994;331:1689-92.
- Lidegaard O, Kreiner S. Contraceptives and cerebral thrombosis: a five-year national case-control study. *Contraception* 2002;65:197-205.
- Bhutta A, Cleves M, Casey P, Cradock M, Anand KJS. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;288:728-37.
- Collaborative Group for the Study of Stroke in Young Women. Oral contraceptives and stroke in young women. *JAMA* 1975;231:718-22.
- Henrich JB, Horwitz RI. A controlled study of ischemic stroke risk in migraine patients. *J Clin Epidemiol* 1989;42:773-80.
- Marini C, Carolei A, Roberts RS, Prencipe M, Gandolfo C, Inzitari D, et al. Focal cerebral ischemia in young adults: a collaborative case-control study. *Neuroepidemiology* 1993;12:70-81.
- Tzourio C, Iglesias S, Hubert JB, Visy JM, Alperovitch A, Tehindranarivelo A, et al. Migraine and risk of ischemic stroke: a case-control study. *BMJ* 1993;307:289-92.
- Lidegaard O. Oral contraceptives, pregnancy and the risk of cerebral thromboembolism: the influence of diabetes, hypertension, migraine and previous thrombotic disease. *Br J Obstet Gynaecol* 1995;102:153-9.
- Haapaniemi H, Hillbom M, Juvela S. Life-style associated risk factors for acute brain infarction among persons of working age. *Stroke* 1997;28:26-30.
- Chang CL, Donaghy M, Poulter N. Migraine and stroke in young women: case-control study. *BMJ* 1999;318:13-8.
- Donaghy M, Chang CL, Poulter N. Duration, frequency, recency, and type of migraine and the risk of ischaemic stroke in women of childbearing age. *J Neurol Neurosurg Psychiatry* 2002;73:747-50.
- Schwaag S, Nabavi DG, Frese A, Husstedt IW, Evers S. The association between migraine and juvenile stroke: a case-control study. *Headache* 2003;43:90-5.
- Buring JE, Hebert P, Romero J, Kittross A, Cook N, Manson J, et al. Migraine and subsequent risk of stroke in the physicians' health study. *Arch Neurol* 1995;52:129-34.
- Merikangas KR, Fenton B, Cheng SH, Stolar MJ, Rish N. Association between migraine and stroke in a large-scale epidemiological study of the United States. *Arch Neurol* 1997;54:362-8.
- Nightingale AL, Farmer RD. Ischemic stroke in young women: a nested case-control study using the UK general practice research database. *Stroke* 2004;35:1574-8.
- Kruit MC, Buchem MA, Hofman P, Bakkers JT, Terwindt GM, Ferrari MD, et al. Migraine as a risk factor for subclinical brain lesions. *JAMA* 2004;291:427-34.
- Schwartz SM, Pettiti DB, Siscovick DS, Longstreth WT Jr, Sidney S, Raghunathan TE, et al. Stroke and use of low-dose oral contraceptives in young women: a pooled analysis of two US studies. *Stroke* 1998;29:2277-84.
- Tietjen GE, Day M, Norris L, Aurora S, Halvorsen A, Shultz LR, et al. Role of anticardiolipin antibodies in young persons with migraine and transient focal neurologic events: a prospective study. *Neurology* 1998;50:1433-40.

(Accepted 9 November 2004)

doi 10.1136/bmj.38302.504063.8F

Comparing estimates of cost effectiveness submitted to the National Institute for Clinical Excellence (NICE) by different organisations: retrospective study

A H Miners, Martina Garau, Dogan Fidan, A J Fischer

Abstract

Objective To assess the association between different types of organisation and the results from economic evaluations.

Design Retrospective pairwise comparison of evidence submitted to the technology appraisal programme of the National Institute for Clinical Excellence (NICE) by manufacturers of the relevant healthcare technologies and by contracted university based assessment groups.

Data sources Data from the first 62 appraisals.

Main outcome measure Incremental cost effectiveness ratios.

Results Data from 27 of the 62 appraisals could be compared. The analysis of 54 pairwise comparisons showed that manufacturer's estimates of incremental cost effectiveness ratios were lower (suggesting a more cost effective use of resources) than those produced by the assessment groups (25 were lower, 29 were the same, none were higher, $P < 0.01$). Restriction of this dataset to include only one pairwise comparison per appraisal (27 pairs) produced a similar result (21 were lower, two were the same, four were higher, $P < 0.001$).

Health Economics Research Group, Brunel University, Uxbridge UB8 3PH

A H Miners
visiting research fellow

Economics Department, University of York, Heslington, York YO1 5DD

Martina Garau
MSc student

continued over

BMJ 2005;330:65-8



This is the abridged version of an article that was posted on bmj.com on 15 December 2004: <http://bmj.com/cgi/doi/10.1136/bmj.38285.482350.82>

Health Policy Unit,
Department of
Public Health and
Policy, London
School of Health
and Tropical
Medicine, London
WC1E 7HT
Dogan Fidan
*honorary research
fellow*

Department of
Community Health
Sciences, St
George's Hospital
Medical School,
London SW17 0RE
A J Fischer
senior lecturer

Correspondence to:
A H Miners
Alec.Miners@
nice.nhs.uk

Conclusions The estimated incremental cost effectiveness ratios submitted by manufacturers were on average significantly lower than those submitted by the assessment groups. These results show that an important role of NICE's appraisal committee, and of decision makers in general, is to determine which economic evaluations, or parts of evaluations, should be given more credence.

Introduction

Evidence suggests that profit making organisations are more likely to report favourable results and conclusions from clinical studies than non-profit organisations. Three studies have formally assessed this association with economic evaluations. One showed that evaluations of products sponsored by drug companies were less likely to report unfavourable qualitative conclusions (5% *v* 38%; $P=0.04$) compared with studies sponsored by non-profit organisations.¹ The other two studies reported similar findings.^{2 3}

The technology appraisals programme of the National Institute for Clinical Excellence (NICE) provides guidance to the NHS in England and Wales on the use of new and existing health technologies. Manufacturers (or relevant UK agents) of the relevant technology, professional groups, and national groups representing patients submit evidence. Collectively, these three groups are known as consultees. An academic centre commissioned through the NHS's health technology assessment programme, called an assessment group, also assesses the evidence. Guidance to the NHS is formulated with regard to this evidence, which includes information on cost effectiveness.

The criterion traditionally used to assess cost effectiveness is the magnitude of a statistic known as the incremental cost effectiveness ratio, defined as the difference in costs between two technologies divided by the difference in their benefits.⁴ The lower the incremental cost effectiveness ratio, the more cost effective a technology.

We evaluated the association between source of funding and the results from economic evaluations submitted to NICE's technology appraisals programme.

Methods

Detailed information regarding the technology appraisals programme is available elsewhere (www.nice.org.uk).

We extracted estimated incremental cost effectiveness ratios for the first 62 technology appraisals, issued between March 2000 and May 2003. Some contained guidance on more than one aspect of a technology, so each recommendation was treated as a separate observation for which cost effectiveness estimates could exist. We included pairings of estimates in the final dataset only if the manufacturer and the assessment group had each submitted at least one incremental cost effectiveness ratio for a technology by using the same health outcome measure. We assumed that results from cost minimisation analyses indicated clinical equivalence, together with an increase or decrease in cost, and included them in the analysis accordingly.

We abstracted estimates from published guidance, assessment reports, and manufacturers' submissions, and these were double checked by a second reviewer. Where the two reviewers' estimates did not match, we sought opinion from the relevant technology analyst at NICE.

Statistical analysis

We used two different approaches to analysis. The first categorised incremental cost effectiveness ratio estimates according to a five point scale. We chose the value of £30 000 (€42 790, \$55 770) per unit of health outcome as one upper limit because of published comments on NICE's decisions to date.^{5 6} By comparing categories of estimates (as opposed to continuous variables) we avoided some of the problems posed by negative estimates.

Our null hypothesis was that the incremental cost effectiveness ratios produced by the assessment groups would not differ systematically from those produced by manufacturers.

Our second analysis made only one comparison between the estimates of the assessment groups and those of the manufacturers involved in each technology appraisal, because more than one pairwise comparison of incremental cost effectiveness ratio estimates could occur (as in the first method). For example, there might have been more than one manufacturer per appraisal or separate results for different subgroups of patients. Such clustered estimates will almost certainly be correlated and, if all such estimates are used, will overestimate the accuracy of the difference being investigated (see bmj.com for details). This analysis was based on point ICER estimates.

Results

Of the 62 appraisals, we excluded 35, leaving 27 appraisals containing 54 pairwise comparisons. Over 80% of the manufacturers' incremental cost effectiveness ratios were between dominant (meaning that the appraised technology was considered to be less costly and more effective than the comparator technology) and £15 000 per unit of outcome (categories 0 and 1 combined), whereas the assessment groups' estimates were more uniformly distributed over the five categories (fig 1). Only four (7%) estimates from manufacturers were above £30 000 per unit of outcome compared with 19 (35%) estimates from the assessment groups.

There were 25 cases of negative rank (where the manufacturers' estimates were lower than those of the corresponding assessment group), 29 ties (where the estimates were in the same category), and no positive ranks ($P<0.01$). Figure 2 plots the log of these pairwise comparisons; the 45° line indicates identical incremental cost effectiveness ratios. Points above this line indicate that a manufacturer's estimate was higher (less favourable) than the estimate produced by the assessment group, whereas points below this line indicate the manufacturer submitted a lower (more favourable) estimate than that of the assessment group. The analysis conducted with the second method reduced the 54 pairs of incremental cost effectiveness ratios to 27 pairs, of which two were a tie. The analysis showed a similar result. In 21 instances the manufacturer

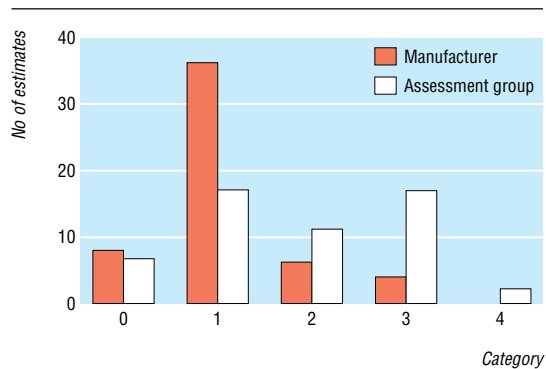


Fig 1 Distribution of estimates of incremental cost effectiveness ratios from assessment groups and manufacturers according to category of cost effectiveness (0=negative incremental cost effectiveness ratios favouring use of technology or, in case of cost minimisation analysis, cheaper than alternative technology; 1=£1-£15 000; 2=£15 001-£30 000; 3=>£30 000; 4=negative incremental cost effectiveness ratios not favouring use of technology or, in case of cost minimisation analysis, more expensive than alternative technology)

submitted a lower estimate than that of the assessment group and in the four remaining it was higher ($P < 0.001$).

Discussion

Our analysis showed that the estimates of incremental cost effectiveness ratio submitted by manufacturers were on average significantly lower than those provided by the assessment groups. This finding is similar to those reported by others.¹⁻³ Unlike previous studies, however, ours is based on specific pairwise comparisons of estimates that have been put forward for the same purposes and constructed with the same terms of reference.

Whether differences between competing cost effectiveness estimates ultimately matter in terms of decision making depends on several factors, including the level of uncertainty surrounding the incremental cost effectiveness ratios⁷ and their absolute value.

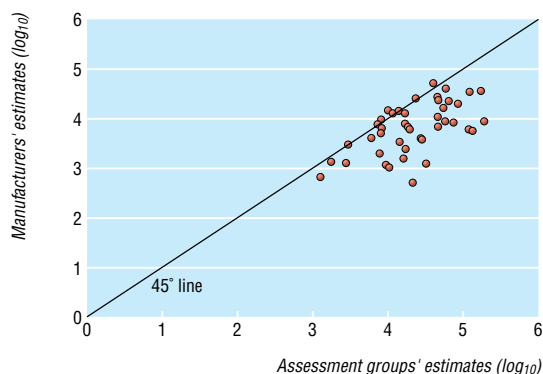


Fig 2 Logged pairwise comparison of incremental cost effectiveness ratios. Only 45 of the 54 pairs of recommendations have been plotted because the remaining nine pairs contained at least one negative estimate and could not be logged. However, in six of nine pairs the estimates from the two sources were broadly similar and in three pairs the estimates reported by manufacturers were much more favourable towards the technology than those reported by the assessment groups

What is already known on this topic

Economic evaluations are used to produce estimates of the cost effectiveness of healthcare technologies

Profit making organisations are more likely to report favourable outcomes from clinical and economic studies than non-profit organisations

One of the key roles of NICE's appraisal committee is to judge the appropriateness of each evaluation and to determine which estimate is the most reasonable

What this study adds

Economic evaluations submitted by manufacturers to NICE's technology appraisals programme were significantly more favourable than evaluations produced by academic research groups

The findings pose questions as to the appropriate methods and processes used by NICE (and reimbursement agencies in general) in the determination of cost effectiveness

NICE will soon be adopting a "reference case" approach designed to be the most appropriate for the NHS to improve consistency across all submitted economic evaluations.⁸ It is therefore feasible that differences between estimated incremental cost effectiveness ratios submitted will be smaller in the future.

We used two different analyses because, though the first method was more informative, we were concerned about the possibility of overstating the accuracy of the difference being investigated. Both approaches, however, produced similar results.

Limitations of study

The evaluations submitted by a manufacturer and the corresponding assessment group are not totally independent, as the assessment group has usually had the opportunity to review the manufacturer's evaluation before completing its own. In contrast, manufacturers do not have the opportunity to review the assessment group's evaluation until it has been submitted to the institute. It is difficult to gauge, however, whether this is likely to lead to systematic differences in the paired ratios or to influence the extent or direction of this difference.

It is also feasible that differences in estimates are not uniform across all categories of cost effectiveness. Logic suggests that because technologies with high incremental cost effectiveness ratios are less likely to be recommended for use within the NHS, groups with vested interests in a technology would be more likely to underestimate incremental cost effectiveness when the "true" estimate is high than when it is low. Insufficient data existed to test this hypothesis.

The reasons that incremental cost effectiveness ratios varied within each appraisal have not yet been investigated. For the moment, we can conclude only that estimates from manufacturers and assessment

groups were significantly different from each other, not that biased results were associated with either type of organisation.

Conclusions

Incremental cost effectiveness ratios submitted to NICE's technology appraisals programme by different types of organisation were significantly different from each other. This questions the appropriate methods and processes used by the institute when determining cost effectiveness, and highlights the need for decision makers to have unhindered access to the methods used to produce cost effectiveness estimates.

We thank all the members of the appraisals team at NICE who helped collect the data, Caroline Sabin (professor of epidemiology and medical statistics, University College London) for her advice on the statistical analysis, Martin Buxton (professor of health economics, Brunel University) for his comments on earlier drafts of the manuscript, and the reviewers for their constructive comments.

Contributors: See bmj.com

Funding: None.

Competing interests: AHM and AJF are members of NICE's appraisals team. DF was a member of the appraisals team when this study was undertaken.

Ethical approval: Not required.

- 1 Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 1999;15:1453-7.
- 2 Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326:1167-70.
- 3 Azimi NA, Welch HG. The effectiveness of cost-effectiveness analysis in containing costs. *J Gen Intern Med* 1998;10:664-9.
- 4 Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 1997.
- 5 Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ* 2004;13:437-52.
- 6 Raftery J. NICE: faster access to modern treatments? Analysis of guidance on health technologies. *BMJ* 2001;323:1300-3.
- 7 Van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E ratios alongside clinical trials. *Health Econ* 1994;3:309-19.
- 8 Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.

(Accepted 18 October 2004)

doi 10.1136/bmj.38285.482350.82

Optimal search strategies for retrieving systematic reviews from Medline: analytical survey

Victor M Montori, Nancy L Wilczynski, Douglas Morgan, R Brian Haynes, for the Hedges Team

Abstract

Objective To develop optimal search strategies in Medline for retrieving systematic reviews.

Design Analytical survey.

Data sources 161 journals published in 2000 indexed in Medline.

Main outcome measures The sensitivity, specificity, and precision of retrieval of systematic reviews of 4862 unique terms in 782 485 combinations of one to five terms were determined by comparison with a hand search of all articles (the criterion standard) in 161 journals published during 2000 (49 028 articles).

Results Only 753 (1.5%) of the 49 028 articles were systematic reviews. The most sensitive strategy included five terms and had a sensitivity of 99.9% (95% confidence interval 99.6% to 100%) and a specificity of 52% (51.6% to 52.5%). The strategy that best minimised the difference between sensitivity and specificity had a sensitivity of 98% (97% to 99%) and specificity of 90.8% (90.5% to 91.1%). Highest precision for multiterm strategies, 57% (54% to 60%), was achieved at a sensitivity of 71% (68% to 74%). The term "cochrane database of systematic reviews.jn." was the most precise single term search strategy (sensitivity of 56% (52% to 60%) and precision of 96% (94% to 98%)). These strategies are available through the "limit" screen of Ovid's search interface for Medline.


Conclusions Systematic reviews can be retrieved from Medline with close to perfect sensitivity or specificity, or with high precision, by using empirical search strategies.

Introduction

Finding systematic reviews in Medline poses two challenges. Firstly, only a tiny proportion of citations in Medline are for literature reviews, and only a fraction of these are systematic reviews. Secondly, the National Library of Medicine's Medlars indexing procedures do not include "systematic review" as a "publication type." Rather, the indexing terms and publication types include a number of variants for reviews, including "meta-analysis" (whether or not from a systematic review); "review, academic"; "review, tutorial"; "review literature"; as well as separate terms for articles that often include reviews, such as "consensus development conference", "guideline", and "practice guideline". The need for special search strategies (hedges) for systematic reviews could be substantially reduced if such reviews were indexed by a separate publication type, but indexers need to be able to dependably distinguish systematic reviews from other reviews. Pending this innovation, there is need for validated search strategies for systematic reviews that optimise retrieval for clinical users and researchers.

In this paper we report on the generation, validation, and performance characteristics of new search strategies to identify systematic reviews in Medline, and compare them with previously published strategies.

 A table showing PubMed translations of Ovid search strategies is on bmj.com

 This is the abridged version of an article that was posted on bmj.com on 24 December 2004: <http://bmj.com/cgi/doi/10.1136/bmj.38336.804167.47>

Department of Medicine, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

Victor M Montori
assistant professor

Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada L8N 3J5

Nancy L Wilczynski
doctoral student

Douglas Morgan
data analyst

R Brian Haynes
professor

Correspondence to: R B Haynes, Department of Clinical Epidemiology and Biostatistics, Room 2C10B Health Sciences Center, McMaster University Faculty of Health Sciences, 1200 Main Street West, Hamilton, ON, Canada L8N 3J5
bhaynes@mcmaster.ca