Quantitative analysis of sponsorship bias in economic studies of antidepressants

C. BRUCE BAKER, MICHAEL T. JOHNSRUD, M. LYNN CRISMON, ROBERT A. ROSENHECK and SCOTT W. WOODS

Background Concern is widespread about potential sponsorship influence on research, especially in pharmacoeconomic studies. Quantitative analysis of possible bias in such studies is limited.

Aims To determine whether there is an association between sponsorship and quantitative outcomes in pharmacoeconomic studies of antidepressants.

Method Using all identifiable articles with original comparative quantitative cost or cost-effectiveness outcomes for antidepressants, we performed contingency table analyses of study sponsorship and design v. study outcome.

Results Studies sponsored by selective serotonin reuptake inhibitor (SSRI) manufacturers favoured SSRIs over tricyclic antidepressants more than non-industry-sponsored studies. Studies sponsored by manufacturers of newer antidepressants favoured these drugs more than did non-industry-sponsored studies. Among industry-sponsored studies, modelling studies favoured the sponsor's drug more than did administrative studies. Industry-sponsored modelling studies were more favourable to industry than were non-industry-sponsored ones.

Conclusions Pharmacoeconomic studies of antidepressants reveal clear associations of study sponsorship with quantitative outcome.

Declaration of interest Range of industry and non-industry funding received, detailed in Acknowledgements.

Long-standing concern exists about the potential influence of financial interests on medical decision-making (e.g. Hillman et al, 1990; Rennie & Flanagin, 1992). Especially vigorous discussion has centred on the conduct and reporting of pharmacoeconomic research (e.g. Hillman et al, 1991; Udrarhelyi et al, 1992; Gulati & Bitran, 1995; Siegel et al, 1996; Neumann, 1998; Hill et al, 2000; Jones & Cockrum, 2000; Neumann et al, 2000b). However, there has been little quantitative study of potential bias in pharmacoeconomic research throughout medicine. Reported studies have reached mixed conclusions (e.g. Sacristan et al, 1997; Azimi & Welch, 1998; Friedberg et al, 1999; Neumann et al, 2000a), perhaps in part because with one exception (Friedberg et al, 1999) they investigated several drugs and in some cases included medical devices. We are unaware of any study focused on psychiatric medication.

We studied associations between sponsorship and study design with quantitative outcome in pharmacoeconomic studies by examining the test case of antidepressants. We asked the following primary questions. First, is there an association between industry v. non-industry sponsorship of studies and quantitative conclusions? Second, among industry-sponsored studies and between industry-sponsored v. non-industry-sponsored studies, is there an association between study design and quantitative conclusions?

METHOD

We chose antidepressants licensed in the UK or the USA as our test case because of their large market share and the number of pharmacoeconomic studies. Antidepressants rank in the top three drug classes worldwide in terms of sales dollars. Their growth in sales ranks them among the top five drug classes worldwide (IMS Health, 2001). Additionally, these antidepressants are the

subject of multiple cost-outcome studies reporting quantitative results.

Study sample

To locate reports of pharmacoeconomic studies of antidepressant drugs we used the Cochrane Library, Medline and Health-STAR databases supplemented by manual searches based on the references cited in the studies located through the databases. We searched for all articles between 1987 - the year the first 'newer' antidepressant, fluoxetine, received US Food and Drug Administration (FDA) approval and April 2001. The search terms we used in Medline and HealthSTAR were COST-BENEFIT ANALYSIS or COST SAVINGS or DRUG COSTS or COST-EFFECTIVE-NESS (text word) and ANTIDEPRESSIVE AGENTS or ANTIDEPRESSANT (text word). The search term we used in the Cochrane Library was ANTIDEPRESSIVE AGENTS. We identified 46 articles (Jonsson & Bebbington, 1993, 1994; Hatziandreu et al, 1994; Le Pen et al, 1994; McFarland, 1994; Sclar et al, 1994, 1995, 1998, 1999; Stewart, 1994; Anton & Revicki, 1995; Einarson et al, 1995, 1997; Lapierre et al, 1995; Nuijten et al, 1995; Revicki et al, 1995, 1997; Skaer et al, 1995; Bentkover & Feighner, 1996; Forder et al, 1996; Hylan et al, 1996, 1998; Montgomery et al, 1996; Smith & Sherrill, 1996; Croghan et al, 1997, 2000; Melton et al, 1997; Obenchain et al, 1997; Woods & Rizzo, 1997; Boyer et al, 1998; Canadian Coordinating Office for Health Technology Assessment, 1998; Crown et al, 1998; Simon & Fishman, 1998; Thompson et al, 1998; Brown et al, 1999a,b; Griffiths et al, 1999; Nurnberg et al, 1999; Russell et al, 1999; Simon et al, 1999; Borghi & Guest, 2000; Sullivan et al, 2000; Casciano et al, 2001; Doyle et al, 2001; Poret et al, 2001; Wan et al, 2002). We excluded two studies (Boyer et al, 1998; Simon et al, 1999) because they were randomised trials, unlike all the other studies, which were modelling studies or analyses of administrative databases. The remaining articles represent 45 separate studies. Two articles report the results of one study (Jonsson & Bebbington, 1993, 1994). Two articles report two studies each, one in-patient, one out-patient (Einarson et al, 1995, 1997). Two articles reported slight variations on two studies, one in-patient and one out-patient (Casciano et al, 2001; Doyle et al, 2001).

Classification of studies

For the primary analysis we categorised each study according to whether it was industry-sponsored. The study was categorised as industry-sponsored if at least one author was listed as a pharmaceutical company employee, or an acknowledgement listed pharmaceutical company support; otherwise, it was categorised as non-industry-sponsored. For secondary analyses we categorised studies authored by industry employees separately from studies only listing financial support.

Study sponsors were categorised by product into those manufacturing selective serotonin reuptake inhibitors (SSRIs: fluoxetine, sertraline, paroxetine and citalopram) or 'atypical' antidepressant drugs (venlafaxine, bupropion and mirtazapine).

Operationalisation of outcomes

For either of the questions posed in our study no single means of operationalising the issue of which antidepressant was favoured could be applied to all studies. Therefore, we performed separate analyses alternative operationalisations. Specifically, for question one (the industry v. non-industry comparison), no single standard was applicable that allowed analysis of all 46 studies. Seemingly simple standards such as 'sponsor's antidepressant favoured' could not apply: in nonindustry-sponsored studies, there is no 'sponsor's antidepressant'. In our primary analysis of industry-sponsored ν . nonindustry-sponsored studies we examined whether the outcome favoured SSRIs or tricyclic antidepressants (TCAs), excluding studies sponsored by 'atypical' antidepressant manufacturers. To allow analysis of the latter studies, we performed an alternative analysis based on whether the outcome favoured the 'newest antidepressant' ('newness' was based upon date of FDA approval). In this analysis, studies in which the sponsor's drug was not the newest were excluded.

In addressing our second question, regarding the association of study design with bias on outcome, we examined the issue both within industry-sponsored trials and between industry-sponsored and non-industry-sponsored trials. Within the first group we looked at the association of modelling v. administrative study designs with outcome. We operationalised the outcomes and groups in two alternative ways: favouring the newest drug among all industry-sponsored studies, or favouring

the sponsored drug among all industrysponsored studies.

In examining the association of study design with outcome between industry v. non-industry sponsors, we compared the outcome patterns within modelling studies. We could not compare outcome patterns in administrative data studies given there was only one such non-industry-sponsored study. We operationalised outcomes in two alternative ways: favouring the newest drug, or favouring SSRIs v. TCAs.

Rating study outcomes

Initially two of the authors (C.B.B. and M.N.J.) independently categorised sponsorship and outcomes of each study. If their ratings were inconsistent, a third author (S.W.W.) rated the study. Initial ratings agreed in all cases but one.

Most studies contained several outcomes. However, we wished to rate a single outcome from each study and employed the following decision rules to select that outcome. First, we selected only quantitative outcomes. Second, among base case and variants, we selected the base case. Third, among outcomes adjusted for bias and unadjusted outcomes, we selected the adjusted outcome. Fourth, among outcomes for various time periods, we selected the longest period. Fifth, among multiple pharmacoeconomic indicators, we selected a single outcome on the basis of the following rules: if only cost outcomes were reported, we chose total costs over more limited costs; if cost and cost-effectiveness outcomes were reported, we chose costeffectiveness outcomes; and if more than one type of cost-effectiveness ratio was reported, we chose incremental over average ratios. Sixth, if results were reported separately for individual countries, we selected the results for the UK and the USA.

After selecting a single outcome for each study, the researchers rated each study as favourable, neutral or unfavourable for the drug of interest, depending on the particular analysis (e.g. SSRI in the SSRI v. TCA analysis, or newest antidepressant in the newest v. older antidepressant analysis): 'favourable' meant that a drug's quantitative cost-effectiveness results were unequalled by any of the other drugs in the study; 'neutral' meant that although other drugs' results might be equal to it, none surpassed the drug of interest; and 'unfavourable' meant that other drugs' results did surpass the drug of interest.

Raters used all available information to judge differences in outcomes among drugs. If the study reported statistical significance, raters based their judgements on statistically significant differences. If the study did not report statistical significance, raters based their judgements on the reported numerical differences. With quality-adjusted life-years (QALYs), raters judged a treatment superior if marginal cost-effectiveness was less than US\$20 000 per QALY, a common applied limit (Laupacis *et al*, 1992). Subsequently, we performed a sensitivity analysis by varying the marginal threshold between \$20 000 and \$100 000 per QALY.

The following is an example of how raters applied the rules noted above to designate a specific study as favourable, neutral or unfavourable. In the SSRI v. tricyclic or heterocyclic antidepressant analysis of the Hatziandreu study (Hatziandreu et al, 1994) the preceding rules led raters to judge that the study favoured the SSRI. The study reported the base case incremental cost-effectiveness ratio to be £2172 (\$3692) for each QALY gained by using the SSRI rather than the TCA. This cost per QALY gained is less than the \$20000 per QALY cut-off noted in the raters' decision rules; therefore, the study was rated as favourable for the SSRI.

In addition to the planned analyses described above, we performed two exploratory analyses: one was based on the number of industry authors and the second was based on the ordinal position of any industry authors. Neither of these analyses yielded a significant association.

We analysed the association between sponsorship and outcome using Fisher's exact test as generalised for 2×3 tables. We chose contingency table analysis rather than a meta-analytic technique because of the qualitative heterogeneity of the pharmacoeconomic outcome types across studies, which ranged from direct costs per patient, to direct costs per treatment success, to direct costs per symptomfree day, to lifetime direct costs per discounted QALY. We judged inappropriate to transform these qualitatively disparate types of outcomes into a common effect size. We selected the 0.05 α level, two-tailed.

RESULTS

Details of the studies are listed in Tables 1–3.

 Table I
 Economic outcome studies of antidepressant therapy sponsored by manufacturers of selective serotonin reuptake inhibitors

Study	Study design	Sponsor's antidepressant	Newest antidepressant	Control antidepressant	Study favours SSRI or	Study favours newest Study favours sponsor's	udy favours sponsor's
		(or class)	is sponsor's (newest)	(or class)	SSRI class v. TCA/HCA ¹ antidepressant or class ¹	antidepressant or class'	antidepressant
Jonsson & Bebbington (1993, 1994)	Simulation	Paroxetine	Yes	Imipramine	Favourable	Favourable	Favourable
Hatziandreu et al (1994) A	Simulation	Sertraline	Yes	Dothiepin	Favourable	Favourable	Favourable
Lapierre et al (1995) A	Simulation	Paroxetine	Yes	Imipramine	Favourable	Favourable	Favourable
Nuijten et al (1995)	Simulation	Citalopram	Yes	TCAs	Favourable	Favourable	Favourable
Bentkover & Feighner (1996)	Simulation	Paroxetine	Yes	Imi pramine	Favourable	Favourable	Favourable
Hylan et al (1996) A	Simulation	Fluoxetine	Yes	Imi pramine	Favourable	Favourable	Favourable
Sclar et <i>al</i> (1994) A	Administrative	Fluoxetine	Yes	Amitriptyline,	Favourable	Favourable	Favourable
	databases			nortriptyline,			
				desipramine			
Sclar et al (1995) A	Administrative	Fluoxetine	No (paroxetine)	Sertraline,	Not applicable	Not applicable	Favourable
	databases			paroxetine			
Skaer et <i>al</i> (1995)	Administrative	Sertraline	Yes	Amitriptyline,	Favourable	Favourable	Favourable
	databases			nortriptyline,			
				desipramine			
Forder et al (1996)	Administrative	Sertraline	Yes	TCAs	Favourable	Favourable	Favourable
	databases						
Smith & Sherrill (1996) A	Administrative	SSRI class	Yes	TCAs	Neutral	Neutral	Neutra
	databases						
Melton et al (1997)	Administrative	Paroxetine	Yes	TCAs	Neutra	Neutra	Neutral
	databases						
Obenchain et al (1997) A	Administrative	Fluoxetine	Yes	TCAs, trazodone	Favourable	Favourable	Favourable
	databases						
Croghan et al (1997) A	Administrative	Fluoxetine	Yes	TCAs, trazodone	Neutral	Neutra	Neutral
	databases						
Crown et al (1998) A	Administrative	SSRI class	Yes	TCAs	Favourable	Favourable	Favourable
	databases	i					:
Hylan et al (1998) A	Administrative	Huoxetine	No (paroxetine)	TCAs, paroxetine,	Not applicable	Not applicable	Favourable
	databases			sertraline			
Sclar et al (1998)	Administrative	Huoxetine	Yes	Amitriptyline,	Favourable	Favourable	Favourable
	databases			nortriptyline			
Thompson et al (1998) A	Administrative	SSRI class	Yes	TCAs	Neutra	Neutra	Neutral
	databases						
Russell et al (1999)	Administrative	Sertraline	No (paroxetine)	Huoxetine,	Not applicable	Not applicable	Neutral
	databases			paroxetine			
Sclar et al (1999)	Administrative	Citalopram	Yes	Amitriptyline,	Favourable	Favourable	Favourable
	databases			fluoxetine,			
				sertraline,			
				paroxetine			
Croghan et al (2000) A	Administrative	Fluoxetine	No (sertraline)	TCAs, sertraline	Not applicable	Not applicable	Favourable
	databases						

A, at least one author employed by manufacturer; HCA, heterocyclic antidepressant; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

1. Column headings are defined in the Method section.

 Table 2
 Economic outcome studies of antidepressant therapy sponsored by manufacturers of atypical antidepressants

Study	Study design	Sponsor's antidepressant	Newest antidepressant is sponsor's (newest)	Control antidepressant (or class)	Study favours SSRI or SSRI class v. TCA	Study favours newest antidepressant ¹	Study favours sponsor's antidepressant¹
Anton & Revicki (1995) A	Simulation	Nefazodone	Yes	Imipramine, fluoxetine	Favourable	Favourable	Favourable
Einarson et al (1995)	Simulation	Venlafaxine (in-patient	Yes	TCAs, SSRIs, HCAs	Unfavourable	Favourable	Favourable
		start)					
Einarson et <i>al</i> (1995)	Simulation	Venlafaxine (out-patient	Yes	TCAs, SSRIs, HCAs	Favourable	Unfavourable	Unfavourable
		start)					
Revicki et al (1995) A	Simulation	Nefazodone	Yes	Imipramine, fluoxetine	Favourable	Favourable	Favourable
Montgomery et al (1996)	Simulation	Nefazodone	Yes	Imipromine	Not applicable	Favourable	Favourable
Revicki et al (1997)	Simulation	Nefazodone	Yes	Fluoxetine, imipramine	Favourable	Favourable	Favourable
Einarson et <i>al</i> (1997)	Simulation	Venlafaxine (in-patient	Yes	TCAs, SSRIs	Unfavourable	Favourable	Favourable
		start)					
Einarson et al (1997)	Simulation	Venlafaxine (out-patient	Yes	TCAs, SSRIs	Favourable	Favourable	Favourable
		start)					
Brown et al (1999a) A	Simulation	Mirtazapine	Yes	Amitriptyline,	Not applicable	Favourable	Favourable
				fluoxetine			
Brown et al (1999b) A	Simulation	Mirtazapine	Yes	Amitriptyline	Not applicable	Favourable	Favourable
Borghi & Guest (2000)	Simulation	Mirtazapine	Yes	TCAs, fluoxetine	Not applicable	Favourable	Favourable
Doyle et al (2001)	Simulation	Venlafaxine (in-patient	Yes	SSRIs, TCAs	Favourable	Favourable	Favourable
Casciano et al (2001)		start)					
Doyle et al (200 l)	Simulation	Venlafaxine (out-patient	Yes	SSTIs, TCAs	Favourable	Favourable	Favourable
Casciano et al (2001)		start)					
Griffiths et al (1999)	Administrative	Venlafaxine	Yes	TCA	Not applicable	Neutral	Neutral
	databases						
Sullivan et al (2000)	Administrative	Venlafaxine	No (nefazodone)	TCAs, SSRIs,	Favourable	Not applicable	Neutral
	databases			nefazodone,			
				bupropion,			
				trazodone			
Poret et al (200 I) A	Administrative	Bupropion SR	Yes	Fluoxetine, paroxetine,	Not applicable	Neutral	Neutral
	databases			sertraline, TCAs,			
				venlafaxine			
Wan e <i>t al</i> (2002) A	Administrative	Venlafaxine	Yes	SSRIs	Not applicable	Neutral	Neutral
	Jacabases						

A, at least one author employed by manufacturer; HCAs, heterocyclic antidepressants maprotiline and trazodone; SR, sustained release; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

I. Column headings are defined in the Method section.

 Table 3
 Non-industry-sponsored economic outcome studies of antidepressant therapy

Study	Study design	Newest antidepressant or class	Control antidepressant	Study favours SSRI or SSRI class v. TCA ¹	Study favours newest antidepressant ¹
McFarland (1994)	Simulation	Paroxetine	Imipramine	Neutral	Neutral
Le Pen et al (1994)	Simulation	Fluoxetine	TCAs	Favourable	Favourable
Stewart (1994)	Simulation	Paroxetine	Amitriptyline, imipramine, sertraline	Unfavourable	Unfavourable
Woods & Rizzo (1997)	Simulation	Paroxetine	Imipramine	Unfavourable	Unfavourable
Canadian Coordinating Office for Health Technology Assessment (1998)	Simulation	SSRIs	TCAs alone TCAs/switch	Unfavourable	Unfavourable
Simon & Fishman (1998)	Administrative databases	Fluoxetine	Imipramine, desipramine	Neutral	Neutral
Nurnberg et al (1999)	Administrative	Paroxetine	Fluoxetine, sertraline	Not applicable	Unfavourable

SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

Association of sponsorship

SSRI v. TCA analysis

For the primary analysis of industry v. non-industry sponsorship of SSRI v. TCA studies, six of seven non-industry-sponsored studies were eligible for analysis (see Table 3). Seventeen industry studies were eligible (see Tables 1 and 2).

Distribution and results for Fisher's exact test are noted in Table 4. The association between industry sponsorship and outcome favouring SSRIs ν . TCAs was statistically significant. Each of the two secondary analyses contrasting studies with industry-employed authors ν . non-industry-sponsored studies and contrasting studies with industry funding alone ν . non-industry-sponsored studies demonstrated a statistically significant association between industry sponsorship and outcome favouring SSRIs ν . TCAs, with probability values of 0.0420 and 0.0163 respectively.

New v. old antidepressant analysis

All non-industry-sponsored studies were eligible (see Table 3). Thirty-three industry-sponsored studies (see Tables 1 and 2) were eligible. Distribution and results of Fisher's exact test are noted in Table 4. The association between industry sponsorship and outcome favouring the newest antidepressant was statistically significant. Each of the two secondary analyses contrasting studies

with industry-employed authors v. non-industry-sponsored studies and contrasting studies with industry funding alone v. non-industry-sponsored studies demonstrated a statistically significant association between industry sponsorship and outcome favouring the newest antidepressant, with probability values of 0.0047 and 0.0018 respectively.

Association between study design and sponsorship bias

Question I

Within industry-sponsored studies, is there a difference in tendency to favour the sponsor's drug over a competitor's drug or drug class, based on type of study design? For the principal analysis, 'favouring the

Table 4 Fisher's exact test: distributions and probabilities

		Outcome	
	Favourable	Neutral	Unfavourable
Sponsorship v. outcome favouring	SSRIs over TCAs: industry	v. non-industry stud	ies ¹
Industry sponsor	13	4	0
Non-industry sponsor	1	2	3
Sponsorship v. outcome favouring	newest antidepressant: ii	ndustry v. non-indust	ry studies²
Industry sponsor	25	7	1
Non-industry sponsor	1	2	4
Study design v. outcome favouring	newest antidepressant: v	within-industry studie	es³
Simulation design	18	0	1
Administrative data design	7	7	0
Sponsorship v. outcome favouring studies ⁴	newest antidepressant: ii	ndustry v. non-indust	ry modelling
Industry sponsor	18	0	1
Non-industry sponsor	1	ĺ	3

 ${\color{blue} {\sf SSRI}, selective \ serotonin \ reuptake \ inhibitor; TCA, tricyclic \ antidepressant.} \\$

- I. Fisher's exact test, P=0.0037.
- 2. Fisher's exact test, P=0.0005.
- Fisher's exact test, P=0.0008.
- 4. Fisher's exact test, P=0.0023.

I. Column headings are defined in the Method section.

sponsor's drug or drug class' was defined based on favouring the newest drug among all manufacturer-sponsored studies. Thirty-three industry-sponsored studies were eligible (see Tables 1 and 2). Distribution and results of the Fisher's exact test are noted in Table 4. The association between modelling ν . administrative study design and outcome favouring the newest drug was statistically significant.

For our alternative analysis based on whether the sponsor's drug or drug class won, regardless of whether it was newest, all 38 industry-sponsored studies were eligible (see Tables 1 and 2). This analysis yielded a probability value of 0.0011, consistent with the results in the primary analysis.

Question 2

Between industry-sponsored and nonindustry-sponsored modelling design studies, is there a difference in outcome patterns? For the principal analysis of this question we examined the patterns of favouring the newest drug. Nineteen industry studies (see Tables 1 and 2) and five non-industry studies (see Table 3) were eligible. The distribution and results of the Fisher's exact test are noted in Table 4. The association between industry ν . nonindustry sponsorship of modelling studies and outcome favouring the newest drug was statistically significant. Each of the two secondary analyses contrasting studies with industry-employed authors v. nonindustry-sponsored studies and contrasting studies with industry funding alone v. non-industry-sponsored studies demonstrated a statistically significant association between industry sponsorship and outcome favouring the newest antidepressant in modelling studies, with probability values of 0.0010 and 0.0100 respectively.

In an alternative analysis we examined the patterns of favouring SSRIs ν . favouring TCAs in modelling studies. We performed this analysis with the five eligible non-industry-sponsored studies (see Table 3) contrasted first with all twelve eligible industry-sponsored modelling studies (see Tables 1 and 2) that included SSRI ν . TCA comparisons, and then with the six eligible modelling studies sponsored by SSRI manufacturers (see Table 1) that included SSRI ν . TCA comparisons. The results of the Fisher's exact test in the two cases were 0.0139 and 0.0151 respectively, indicating that the tendency for industry-

sponsored simulations to favour SSRIs more often than non-industry-sponsored studies is unlikely to be due to chance.

The sensitivity analysis varying the marginal cost-effectiveness threshold from \$20 000 to \$100 000 per QALY did not change any of the results reported above.

DISCUSSION

Our analyses show that, regardless of how the question was operationalised, for each of our study questions there was greater than chance association between study sponsorship and outcome. Among industrysponsored v. non-industry-sponsored studies, industry-sponsored studies more frequently reported results favourable to the industry sponsor than did non-industry-sponsored studies. This was true whether industry sponsorship was defined as industry authorship, industry financial support alone, or both. Among industry studies, modelling studies were more likely to report results favourable to the sponsor than administrative data studies. Between industry-sponsored and non-industrysponsored modelling design studies, industry studies were more likely to report results favourable to industry.

Consistency with prior studies

Our overall finding of sponsorship bias is consistent with prior studies in that all three prior fully reported studies found some association between study sponsorship and outcomes (Azimi & Welch, 1998; Friedberg et al, 1999; Neumann et al, 2000a). However, any detailed comparison between our study and previous studies is necessarily limited given that the previous studies mixed drugs, devices and other health interventions, and mixed various classes of medicines (Azimi & Welch, 1998; Neumann et al, 2000a); focused on qualitative conclusions (Friedberg et al, 1999); and used various definitions to select the specific study outcomes to be analysed (Azimi & Welch, 1998; Friedberg et al, 1999; Neumann et al, 2000a). The study by Friedberg et al (1999) of oncology drugs is perhaps most comparable with our current study, given their focus on a single pharmaceutical class and their categorisation of study conclusions as favourable, neutral or unfavourable, although they focused on qualitative rather than quantitative conclusions. Like our study, that of Friedberg et al did find an association between study conclusion and funding source.

Support for concern about modelling studies

In addition to supporting the general concern about sponsorship bias in pharmacoeconomic studies, our findings support the more specific concerns that have been raised about the potential for bias in modelling studies (Luce, 1995; O'Brien, 1996; Sheldon, 1996; Maynard & Cookson, 1998; McCabe & Dixon, 2000). Such support stems from the combination of our two findings regarding study design: among industry studies, modelling studies are more favourable to the sponsor than administrative studies, and in a comparison of industry-sponsored and non-industrysponsored modelling studies, studies sponsored by industry are significantly more favourable to industry.

Limitations of our study

Our study has clear limitations. Randomised pharmacoeconomic trials could not be compared on the basis of sponsorship because there were only two such trials in this area. Relatively few non-industry-sponsored studies were available. We examined only one class of medications; analyses of other classes of medications should be conducted.

Bias v. accuracy

Although we have demonstrated several associations between study sponsorship and outcome, these associations do not suggest which (if either) side presents a more accurate estimate of relative pharmacoeconomic outcome. Both industrysupported and non-industry-supported researchers may be subject to forces that could potentially bias their work (Yee & Hillman, 1997; Drummond, 1998; Rennie & Luft, 2000). Additionally, journal editorial processes can result in a biased sample of studies being published. It has been observed that journals tend to publish studies with 'positive' rather than 'negative' results (Freemantle & Mason, 1997).

Causes of bias

Many ideas have been offered to explain how sponsorship could result in biased reported outcomes (Udrarhelyi *et al*, 1992; Freemantle & Mason, 1997; Drummond, 1998; Cook, 1999; Neumann *et al*, 2000a; Rennie & Luft, 2000). Industry, motivated to enhance sales of its products, might only pursue studies on products and select comparators that would yield favourable results. They might select biased populations within administrative data-sets, overtly or subtly influence analytical methods or models, or veto submission for publication of studies yielding unfavourable results. Non-industrysponsored researchers might bias the studies submitted for publication in similar ways, although perhaps from different motivations such as controlling formulary costs, personal or academic rivalries, or career promotion.

We are unable to pinpoint the causes of bias among the reports analysed here. Examination of the individual studies does not reveal a common element that differentiates industry-sponsored non-industry-sponsored studies; rather, the methodological limitations in the studies vary widely. These limitations have been discussed extensively elsewhere (Hotopf et al, 1996; Woods & Baker, 1997, 2002). However, at least two suggested causes seem unlikely. First, some commentators have noted the potential role of selection bias - i.e. the tendency of researchers not to submit and of journals not to publish small studies or studies with negative statistical outcomes (Freemantle & Mason, 1997; Neumann, 1998). This would help to explain how an overall preponderance of statistically positive studies could exist even if there were true uncertainty about alternative medications (Djulbegovic et al, 2000). The difference we have shown between industry-sponsored and nonindustry-sponsored studies suggests that submission or editorial selection bias based on statistical significance alone does not adequately explain the bias in the present case. Second, it has been suggested that a particular sponsor weeds out weak alternatives among its drugs in early preliminary processes; therefore, drugs that reach the stage of being marketed are strong competitors and likely to yield analyses that favour the sponsor's drug (Gagnon, 2000). However, these same strong competitors performed less well in non-industrysponsored studies, as shown clearly in our analysis of outcomes favouring either SSRIs or TCAs. Moreover, it should be noted that in the 18 studies with head-to-head comparisons among such strong competitors, the sponsor's drug lost only once (Einarson et al, 1995).

Bias in efficacy v. pharmacoeconomic studies

It is not possible to comment about whether the bias revealed in the current study of pharmacoeconomic reports of antidepressants is any greater or less than the sponsorship bias that may exist in efficacy studies of antidepressants. There is no published report on sponsorship bias in efficacy studies in any medication category within psychiatry. The only published report devoted to such quantitative analysis of psychiatric medications is a letter reviewing efficacy studies of any psychiatric medication in one journal over a 1-year period (Mandelkern, 1999). This author reported a tally for industry-supported studies of 16 favourable to the manufacturer's drug and none unfavourable, and for unsupported studies 10 favourable and 6 unfavourable, concluding that there was a correlation between source of support and efficacy outcome. In other areas of medicine, bias has been demonstrated repeatedly in efficacy studies (Davidson, 1986; Rochon et al, 1994; Stelfox et al, 1998; Djulbegovic et al, 2000). A study of sponsorship bias in efficacy trials of antidepressants would provide a useful comparison for our study.

It is important for pharmacoeconomic studies to attempt to give estimates that are as accurate and uninfluenced by bias as possible, given the large and growing number of health care dollars spent on medications. Pharmaceutical sales for North America were reported to be US\$153 billion in 2000, representing a 14% growth over the previous year (IMS Health, 2001). Owing to the importance of cost constraint in medicine the volume of pharmacoeconomic research has been growing (Detsky, 1994) and is linked to governmental purchasing decisions in some (Canadian jurisdictions Coordinating Office for Health Technology Assessment, 1994; Ontario Ministry of Health, 1994; Australian Government, 1995). However as we noted previously, financial and other incentives create strong motives for bias. Our results for antidepressants suggest that actual bias related to sponsorship appears to exist, although whether or how the bias and specific motives are related cannot be determined. Until the mechanisms producing the bias are better understood, interpretation of results from pharmacoeconomic studies should take sponsorship into account.

ACKNOWLEDGEMENTS

Consultancies, research grants or unrestricted grants were received from Pfizer, Bristol-Myers Squibb, Eli Lilly & Co. and Forest Laboratories. The work was supported in part by PHS KO8 MH-01718 (C.B.B.), NARSAD R03051 (C.B.B.), Robert Wood Johnson Foundation (M.L.C.), Substance Abuse and Mental Health Services Administration (M.L.C.), Meadows Foundation (M.L.C.), Houston Endowment (M.L.C.), Hogg Foundation (M.L.C.), Texas Department of Mental Health and Mental Retardation (M.L.C.) and PHS MH-54446 (SWW).

REFERENCES

Anton, S. F. & Revicki, D. A. (1995) The use of decision analysis in the pharmacoeconomic evaluation of an antidepressant: a cost-effectiveness study of nefazodone. *Psychopharmacology Bulletin,* **31**, 249–258.

Australian Government (1995) Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee Including Major Submissions Involving Economic Analysis. Canberra: Australian Government Publishing Service.

Azimi, N. A. & Welch, G. H. (1998) The effectiveness of cost-effectiveness analysis in containing costs. *Journal of General Internal Medicine*, **13**, 664–669.

Bentkover, J. D. & Feighner, J. P. (1996) Cost analysis of paroxetine vs imipramine in major depression. *Pharmacoeconomics*, **8**, 223–232.

Borghi, J. & Guest, J. F. (2000) Economic impact of using mirtazapine compared to amitriptyline and fluoxetine in the treatment of moderate and severe depression in the UK. *European Psychiatry*, **15**, 378–387.

Boyer, P., Danion, J. M., Bisserbe, J. C., et al (1998) Clinical and economic comparison of sertraline and fluoxetine in the treatment of depression. Pharmacoeconomics, 13, 157–169.

Brown, M. C., Nimmerrichter, A. A. & Guest, J. F. (1999a) Cost-effectiveness of mirtazapine compared to amitriptyline and fluoxetine in the treatment of moderate and severe depression in Austria. European Psychiatry: Journal of the Association of European Psychiatrists, 14, 230–244.

____, van Loon, J. M.T. & Guest, J. F. (1999b) Costeffectiveness of mirtazapine relative to amitriptyline in the treatment of moderate and severe depression in France. European Journal of Psychiatry, 13, 197–208.

Canadian Coordinating Office for Health Technology Assessment (1994) Guidelines for Economic Evaluation of Pharmaceuticals. Ottawa: CCOHTA.

_____(1998) Selective serotonin reuptake inhibitors (SSRIs) for major depression. Part II: The cost effectiveness of SSRIs in treatment of depression. August 1997. Report 4E. Evidence-Based Medicine, 3, 87.

Casciano, J., Doyle, J., Arikian, S., et al (2001) The health economic impact of antidepressant usage from a payer's perspective: a multinational study. International Journal of Clinical Practice, 55, 292–299.

Cook, J. R. (1999) Challenges in economic outcomes research. *Hepatology*, **29**, 23S–25S.

Croghan, T. W., Lair, T. J., Engelhart, L., et al (1997) Effect of antidepressant therapy on health care utilization and costs in primary care. *Psychiatric Services*, 48, 1420–1426

___, Kniesner, T. J., Melfi, C. A., et al (2000) Effect of antidepressant choice on the incidence and economic

intensity of hospitalization among depressed individuals. Administration and Policy in Mental Health, 27, 183–195.

Crown, W. H., Hylan, T. R. & Meneades, L. (1998) Antidepressant selection and use and healthcare expenditures — an empirical approach. *Pharmacoeconomics*, 13, 435–448.

Davidson, R. A. (1986) Source of funding and outcome of clinical trials. *Journal of General Internal Medicine*, **I**, 155–158.

Detsky, A. S. (1994) Using cost-effectiveness analysis for formulary decision making: from theory into practice. *Pharmacoeconomics*, **6**, 281–288.

Djulbegovic, B., Lacevic, M., Cantor, A., et al (2000) The uncertainty principle and industry-sponsored research. *Lancet*, **356**, 635–638.

Doyle, J. J., Casciano, J., Arikian, S., et al (2001) A multinational pharmacoeconomic evaluation of acute major depressive disorder (MDD); a comparison of cost-effectiveness between venlafaxine, SSRIs, and TCAs. *Value in Health, 4*, 16–31.

Drummond, M. F. (1998) A reappraisal of economic evaluation of pharmaceuticals: science or marketing? *Pharmacoeconomics*, **14**, 1–9.

Einarson, T. R., Arikian, S., Sweeney, S., et al (1995) A model to evaluate the cost-effectiveness of oral therapies in the management of patients with major depressive disorders. *Clinical Therapeutics*, **17**, 136–153.

___, Addis, A. & Iskedjian, M. (1997)

Pharmacoeconomic analysis of venlafaxine in the treatment of major depressive disorder. *Pharmacoeconomics*, **12**, 286–296.

Forder, J., Kavanagh, S. & Fenyo, A. (1996) A comparison of the cost-effectiveness of sertraline versus tricyclic antidepressants in primary care. *Journal of Affective Disorders*, **38**, 97–111.

Freemantle, N. & Mason, J. (1997) Publication bias in clinical trials and economic analyses. *Pharmacoeconomics*, 12, 10–16.

Friedberg, M., Saffran, B., Stinson, T. J., et al (1999) Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA*, **282**, 1453–1475.

Gagnon, J. P. (2000) Sources of bias in the economic analysis of new drugs (letter). *JAMA*, **283**, 1423.

Griffiths, R. I., Sullivan, E. M., Frank, R. G., et al (1999) Medical resource use and cost of venlafaxine or tricyclic antidepressant therapy. *Pharmacoeconomics*, 15, 495–505.

Gulati, S. C. & Bitran, J. D. (1995) Cost-effectiveness analysis: sleeping with an enemy or a friend? *Journal of Clinical Oncology,* **13**, 2152–2154.

Hatziandreu, E. J., Brown, R. E., Revicki, D. A., et al (1994) Cost utility of maintenance treatment of recurrent depression with sertraline versus episodic treatment with dothiepin. *Pharmacoeconomics*, 5, 249–264.

Hill, S. R., Mitchell, A. S. & Henry, D. A. (2000) Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. JAMA, 283, 2116–2121.

Hillman, A. L., Eisenberg, J. M., Pauly, M. V., et al (1991) Avoiding bias in the conduct and reporting of cost-effectiveness research sponsored by pharmaceutical companies. New England Journal of Medicine, 324, 1362–1365.

Hillman, B. J., Joseph, C. A., Magel, M. R., et al (1990) Frequency and costs of diagnostic imaging in office practice: a comparison of self referring and radiologist referring physicians. New England Journal of Medicine, 323, 1504–1508.

CLINICAL IMPLICATIONS

- Evaluation of pharmacoeconomic studies of antidepressants for treatment of major depression should take into account the source of study funding and of authorship.
- Evaluation of modelling studies in this area may warrant additional scrutiny.
- Evaluation of pharmacoeconomic studies should be informed by a thorough understanding of their methodology.

LIMITATIONS

- Randomised pharmacoeconomic trials could not be compared because of their scarcity.
- Few non-industry-sponsored trials were available.
- Only one class of medication was examined.

C. BRUCE BAKER, MD, JD, Department of Psychiatry, Yale School of Medicine, and Connecticut Mental Health Center, New Haven, Connecticut; MICHAELT. JOHNSRUD, PhD, Center for Pharmacoeconomic Studies, University of Texas, Austin, Texas; M. LYNN CRISMON, PharmD, Center for Pharmacoeconomic Studies, University of Texas, and Office of the Medical Director, Texas Department of Mental Health and Mental Retardation, Austin, Texas; ROBERT A. ROSENHECK, MD, Department of Psychiatry, Yale School of Medicine, and VA Northeast Program Evaluation Center, West Haven, Connecticut, and Yale School of Medicine Departments of Epidemiology and Public Health, New Haven, Connecticut; SCOTT W.WOODS, MD, Department of Psychiatry, Yale School of Medicine; Connecticut Mental Health Center, New Haven, Connecticut, USA

Correspondence: Dr C. Bruce Baker, Department of Psychiatry, Yale School of Medicine, CMHC Rm 38B, 34 Park St, New Haven, CT 06519, USA. Tel: (203) 974 7051; fax: (203) 974 7057; e-mail: Bruce.Baker@Yale.edu

(First received 3 January 2003, final revision IO July 2003, accepted 31 July 2003)

Hotopf, M., Lewis, G. & Normand, C. (1996) Are SSRIs a cost-effective alternative to tricyclics? *British Journal of Psychiatry*, **168**, 404–409.

Hylan, T. R., Kotsanos, J. G., Andersen, J. S., et al (1996) Comparison of a decision analytic model with results from a naturalistic economic clinical trial: an application of evaluating alternative antidepressants. *American Journal of Managed Care*, **2**, 1211–1223.

___, Crown, W. H., Meneades, L., et al (1998) Tricyclic antidepressant and selective serotonin reuptake inhibitors antidepressant selection and health care costs in the naturalistic setting: a multivariate analysis. Journal of Affective Disorders, 47, 71–79.

IMS Health (2001) The global pharmaceutical market in 2000. www.ims-global.com/insight/news_story/0103/news_story_010314.htm

Jones, M.T. & Cockrum, P. C. (2000) A critical review of published economic modelling studies in depression. *Pharmacoeconomics*, **17**, 555–583.

Jonsson, B. & Bebbington, P. (1993) Economic studies of the treatment of depressive illness. In *Health Economics of Depression* (eds B. Jonsson & J. Rosenbaum), pp. 35–48. Chichester: John Wiley.

— & — (1994) What price depression? The cost of depression and the cost-effectiveness of pharmacological treatment. *British Journal of Psychiatry*, **164**, 665–673.

Lapierre, Y., Bentkover, J., Schainbaum, S., et al (1995) Direct costs of depression: analysis of treatment costs of paroxetine versus imipramine in Canada. *Canadian Journal of Psychiatry*, **40**, 370–377.

Laupacis, A., Feeny, D., Detsky, A. S., et al (1992) How attractive does a new technology have to be to warrant adoption and utilization? *Canadian Medical* Association Journal, 146, 473–481.

Le Pen, C., Levy, E., Ravily, V., et al (1994) The cost of treatment dropout in depression: a cost—benefit analysis of fluoxetine vs. tricyclics. *Journal of Affective Disorders*, **31**, 1–18.

Luce, B. R. (1995) Policy implications of modelling the cost-effectiveness of health care technologies. *Drug Information Journal*, **29**, 1469–1475.

Mandelkern, M. (1999) Manufacturer support and outcome (letter). *Journal of Clinical Psychiatry*, **60**, 122.

Maynard, A. & Cookson, R. F. (1998) Computer modelling: the need for careful evaluation and public audit. *Pharmacoeconomics*, 14, 67–72.

- McCabe, C. & Dixon, S. (2000) Testing the validity of cost-effectiveness models. *Pharmacoeconomics*, 17, 501–513
- McFarland, B. H. (1994) Cost-effectiveness considerations for managed care systems: treating depression in primary care. *American Journal of Medicine*, 97 (suppl. 6A), 475–58S.
- Melton, S. T., Kirkwood, C. K., Farrar, T. W., et al (1997) Economic evaluation of paroxetine and imipramine in depressed outpatients. Psychopharmacology Bulletin, 33, 93–100.
- Montgomery, S. A., Brown, R. E. & Clark, M. (1996) Economic analysis of treating depression with nefazodone v. imipramine. *British Journal of Psychiatry*, 168, 768–771.
- **Neumann, P. J. (1998)** Paying the piper for pharmacoeconomic studies. *Medical Decision Making*, **18**, \$23–26.
- ___, Sandberg, E. A., Bell, C., et al (2000a) Are pharmaceuticals cost-effective? A review of the evidence. Health Affairs, 19, 92–109.
- ___, Stone, P.W., Chapman, R. H., et al (2000b) The quality of reporting in published cost—utility analyses, 1976–1997. Annals of Internal Medicine, 132, 964–972.
- Nuijten, M. J. C., Hardens, M. & Soutre, E. (1995) A Markov process analysis comparing the cost effectiveness of maintenance therapy with citalopram versus standard therapy in major depression. *Pharmacoeconomics*, **8**, 159–168.
- Nurnberg, G. H., Hensley, P. L., Thompson, P. M., et al (1999) Modeling the pharmacoeconomic cost of three selective serotonin reuptake inhibitors. *Psychiatric Services*, **50**, 1351–1353.
- Obenchain, R. L., Melfi, C. A., Croghan, T. W., et al (1997) Bootstrap analyses of cost effectiveness in antidepressant pharmacotherapy. *Pharmacoeconomics*, 11, 464–472.
- **O'Brien, B. (1996)** Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? *Medical Care*, **34**, DS99–I08.
- Ontario Ministry of Health (1994) Ontario Guidelines for Economic Analysis of Pharmaceutical Products.

 Toronto: Ontario Ministry of Health, Drug Programs

 Branch
- Poret, A. W., Neslusan, C., Ricci, J., et al (2001) Retrospective analysis of the health-care costs of bupropion sustained release in comparison with other antidepressants. Value in Health, 4, 362–369.
- **Rennie, D. & Flanagin, A. (1992)** Publication bias: the triumph of hope over experience. *JAMA*, **267**, 411–412.

- **Luft, H. S. (2000)** Pharmacoeconomic analyses: making them transparent, making them credible. *JAMA*, **283.** 2158–2160.
- Revicki, D. A., Brown, R. E., Palmer, W., et al (1995) Modelling the cost effectiveness of antidepressant treatment in primary care. *Pharmacoeconomics*, **8**, 524–540
- _____, ____, Keller, M. B., et al (1997) Cost-effectiveness of newer antidepressants compared with tricyclic antidepressants in managed care settings. Journal of Clinical Psychiatry, **58**, 47–58.
- Rochon, P. A., Gurwitz, J. H., Simms, R. W., et al (1994) A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. Archives of Internal Medicine, 154, 157–163.
- Russell, J. M., Berndt, E. R., Miceli, R., et al (1999) Course and cost of treatment for depression with fluoxetine, paroxetine, and sertraline. *American Journal* of Managed Care, **5**, 597–606.
- Sacristan, J. A., Bolanos, E., Soto, J., et al (1997) Publication bias in health economic studies (letter). Pharmacoeconomics, II, 289–290.
- Sclar, D. A., Robison, L. M., Skaer, T. L., et al (1994) Antidepressant pharmacotherapy: economic outcomes in a health maintenance organization. *Clinical Therapeutics*, 16, 715–730.
- ___, ___, et al (1995) Antidepressant pharmacotherapy: economic evaluation of fluoxetine, paroxetine, and sertraline in a health maintenance organization. Journal of International Medical Research, 23, 395–412.
- ____, Skaer, T. L., Robison, L. M., et al (1998) Economic outcomes with antidepressant pharmacotherapy: a retrospective intent-to-treat analysis. *Journal of Clinical Psychiatry*, **59**, 13–17.
- ___, ___, et al (1999) Economic appraisal of citalopram in the management of single-episode depression. Journal of Clinical Psychopharmacology, 19, 475–54S.
- **Sheldon, T. A. (1996)** Problems of using modelling in the economic evaluation of health care. *Health Economics*, **5**, I–II.
- Siegel, J. E., Weinstein, M. C., Russell, L. B., et al (1996) Recommendations for reporting cost-effectiveness analyses. *JAMA*, **276**, 1339–1341.
- **Simon, G. E. & Fishman, P. (1998)** Cost implications of initial antidepressant selection in primary care. *Pharmacoeconomics*, **13**, 61–70.
- ____, Heilegenstein, J. H., Revicki, D. A., et al (1999) Long-term outcomes of initial antidepressant drug

- choice in a 'real world' randomized trial. Archives of Family Medicine, **8**, 319–325.
- Skaer, T. L., Sclar, D. A., Robison, L. M., et al (1995) Economic evaluation of amitriptyline, desipramine, nortriptyline, and sertraline in the management of patients with depression. *Current Therapeutic Research*, **56**. 556–567.
- **Smith, W. & Sherrill, A. (1996)** A pharmacoeconomic study of the management of major depression: patients in a TennCare HMO. *Medical Interface*, **9**, 88–92.
- Stelfox, H. T., Chua, G., O'Rourke, K., et al (1998) Conflict of interest in the debate over calcium-channel antagonists. New England Journal of Medicine, 338, 101–106.
- **Stewart, A. (1994)** Anti-depressant pharmacotherapy: cost comparison of SSRIs and TCAs. *British Journal of Medical Economics*, **7**, 67–79.
- Sullivan, E. M., Griffiths, R. I., Frank, R. G., et al (2000) One-year costs of second-line therapies for depression. *Journal of Clinical Psychiatry*, 61, 290–298.
- **Thompson, D., Hylan, T. R., McMullen, W., et al** (1998) Predictors of a medical-offset effect among patients receiving antidepressant therapy. *American Journal of Psychiatry*, **155**, 824–827.
- Udrarhelyi, I. S., Graham, A., Rai, A., et al (1992) Cost-effectiveness and cost-benefit analyses in the medical literature: are the methods being used correctly? Annals of Internal Medicine, 116, 238–244.
- Wan, G. J., Crown, W. H., Berndt, E. R., et al (2002) Treatment costs of venlafaxine and selective serotonin-reuptake inhibitors for depression and anxiety. Managed Care Interface, 15, 24–30.
- **Woods, S.W. & Baker, C. B. (1997)** Cost-effectiveness of newer antidepressants. *Current Opinion in Psychiatry*, **10**, 95–101.
- __ & __ (2002) Cost effectiveness of the newer generation of antidepressants. In *Psychopharmacology:* The Fifth Generation of *Progress* (eds K. L. Davis, J. Coyle, D. S. Charney, et al), pp. III9–II38. New York: Raven Press.
- **& Rizzo, J. A. (1997)** Cost-effectiveness of antidepressant treatment reassessed. *British Journal of Psychiatry*, **170**, 257–263.
- Yee, G. C. & Hillman, A. L. (1997) Applied pharmacoeconomics: when can publication be legitimately withheld? *Pharmacoeconomics*, 12, 511–516.