International Journal of Risk & Safety in Medicine 20 (2008) 73–81
DOI 10.3233/JRS-2008-0426
IOS Press

Clinical trials and drug promotion: Selective reporting of study 329

Jon N. Jureidini [a,*], Leemon B. McHenry [b] and Peter R. Mansfield [c]

[a] *Discipline of Psychiatry, University of Adelaide, Adelaide, South Australia*
[b] *Department of Philosophy, California State University, Northridge, CA, USA*
[c] *Discipline of General Practice, University of Adelaide, Adelaide, South Australia*

Abstract. Selective reporting is prevalent in the medical literature, particularly in industry-sponsored research. In this paper, we expose selective reporting that is not evident without access to internal company documents. The published report of study 329 of paroxetine in adolescents sponsored by GlaxoSmithKline claims that "paroxetine is generally well tolerated and effective for major depression in adolescents". By contrast, documents obtained during litigation reveal that study 329 was negative for efficacy on all 8 protocol specified outcomes and positive for harm.

Keywords: Selective reporting, SSRI, litigation, industry sponsorship

1. Introduction

Selective reporting is a significant problem in drug trials [1]. One study found a discrepancy between the primary outcomes specified in the trial protocols, and those listed in the published paper in 62% of 112 trials. Only 50% of efficacy outcomes and 35% of harm outcomes were reported. Efficacy outcomes were more than twice as likely to be reported if they were statistically significant [2]. Selective reporting has been especially problematic in antidepressant research [3,4]. We use internal documents made available during a class action lawsuit (*Beverly Smith vs. SmithKline Beecham*) to illuminate selective reporting of study 329 of paroxetine (Paxil/Seroxat) in adolescent depression. This study was funded by SmithKline Beecham (SKB), now GlaxoSmithKline (GSK) after merging with Glaxo Wellcome in 2000. A report of study 329 was published by the *Journal of the American Academy of Child and Adolescent Psychiatry* (*JAACAP*) in July 2001 [5]. Its authors claimed that paroxetine was "generally well tolerated and effective for major depression in adolescents". The paper became one of the most cited in the medical literature in supporting the use of antidepressants in child and adolescent depression [6] and GSK claimed it demonstrated "REMARKABLE Efficacy and Safety" [7]. In 2007 systematic reviewers were still describing study 329 as a positive trial [8]. Yet this study was negative on all protocol-defined outcomes and demonstrated important safety problems [9].

The aim of this paper is to expose selective reporting that would not be apparent without access to documents that only emerged through litigation [10]. In June 2004 Californian law firm, Baum Hedlund, alleged that GSK misrepresented the safety and efficacy of paroxetine in the pediatric population. In the

* Address for correspondence: Jon N. Jureidini, Department of Psychological Medicine, Women's and Children's Hospital, North Adelaide, 5006 Australia. Tel.: +61881617226; E-mail: jon.jureidini@cywhs.sa.gov.au.

course of that (now largely settled) litigation, GSK were required to provide all relevant documents. Approximately 10,000 pages of documents were made available to J.N.J. by Baum Hedlund, who had contracted him to provide independent psychiatric review of the data. All documents were initially deemed confidential, but after Baum Hedlund made challenges that certain documents did not reveal trade secrets to competitors, some were released into the public domain. To ensure that this paper is based only on the publicly available documents P.R.M., who has not had access to the confidential documents, was given responsibility for quality control of a draft prepared by J.N.J. and L.B.M. All documents referred to in this paper are available to the reader (www.healthyskepticism.org/documents/PaxilStudy329.php), allowing verification of all claims. J.N.J. and L.B.M. assert that no document withheld from the public domain by confidentiality constraints imposed by GSK contradicts any of the documents cited here.

2. Changes in outcome measures

2.1. Initial study design

In 1992 Martin Keller, MD, Chairman of Psychiatry at Brown University, Rhode Island and colleagues successfully proposed to SKB a multi-site study of a selective serotonin reuptake inhibitor and a tricyclic antidepressant in adolescent major depression [11]. The 1993 protocol for the study (and its subsequent amendments) specified two primary outcome measures: change in total Hamilton Rating Scale (HAM-D) score; and proportion of responders (HAM-D $\leqslant 8$ or reduced by $\geqslant$50%) [12]. The protocol also specified six secondary outcome measures (see Table 1). Patients were enrolled between April 1994 and March 1997, with 275 patients completing the acute phase of the study by May 1997. The blind was broken in October, 1997 [13, p. 891]. *There was no significant difference between the paroxetine and placebo groups on any of the eight pre-specified outcome measures* [14].

2.2. Introduction of new outcome measures

However, by the time the data were analysed, many other new measures had been added to the list of secondary outcomes. There was a statistically significant difference between the paroxetine and placebo groups for only two of these additional secondary outcomes: remission (defined as HAM-D $\leqslant 8$); and the HAM-D depression item. Only these two of the extra measures introduced before analysing the data were reported when study 329 was first written up for submission to *JAMA* in 1999 [15]. By that

Table 1
Outcome measures (significant results in **bold**); ordering of outcome measures is from originals

Protocol (1993, 1996) [12]	p	Final paper (2001) [5]	p
*Change in HAM-D total score	0.13	**HAM-D $\leqslant 8$**	**0.02**
*Responders (HAM-D $\leqslant 8$ or reduced by $\geqslant$50%)	0.11	*Responders (HAM-D $\leqslant 8$ or reduced by $\geqslant$50%)	0.11
Depression scale of K-SADS-L	0.07	**HAM-D depressed mood item**	**0.001**
Mean Clinical Global Improvement (CGI) score	0.09	**K-SADS-L depressed mood item**	**0.05**
Autonomous function checklist	0.15	**CGI 1 or 2**	**0.02**
Self-perception profile	0.54	Depression scale of K-SADS-L	0.07
Sickness impact scale	0.46	Mean CGI	0.09
Relapse during maintenance	0.24**	*HAM-D total score	0.13

*Protocol specified primary outcomes. **Not published, calculated by us, trend favours placebo.

Box 1
History of the four positive 'depression related variables' unspecified in the trial protocol

HAM-D $\leqslant 8$	
1992 December	Part of the complex definition of 'responder' in Keller's proposal to SKB [11].
1996 October	Not specified as an outcome measure in the acute-phase protocol [14].
1997 April	First labelled as 'remission', a second "definition of 'response' during the acute phase" [16].
1999 February	Listed as an outcome variable in early drafts of the paper [15].
2001 July	By publication, 'remission' disappears altogether as a label, and 'HAM-D $\leqslant 8$' is conflated with 'HAM-D $\leqslant 8$ or reduced by $\geqslant 50\%$' – see Box 2 [5].
HAM-D depression item	
1997 August	Not mentioned before the official unblinding.
CGI 1 or 2	
1997 April	Mentioned as possible outcome [16].
1998 January	Not mentioned in 'Top Line Results' [17] three months after the blind was broken. Study 329 co-author Ryan noted at the time by hand on his copy of these 'Top Line Results' the percentage of subjects fitting into each of the CGI categories but there is no indication of any decision as to how to make use of this data [18, p. 450].
K-SADS-L depressed mood item	
1998 November	First documented as an outcome variable [14, p. 44].

time, four of the six negative protocol-specified secondary measures had also been removed from the list of secondary outcomes, and two further additional new positive outcome measures had been added (changes in K-SADS depression item and Clinical Global Improvement (CGI) scale of 1 or 2) (see Box 1). No document prior to eight months *after* breaking the blind mentions the K-SADS depression item as an outcome measure. The introduction of these additional outcome measures so long after initial data analysis is consistent with a statement by GSK's senior scientist James McCafferty, that analysis had revealed 'a strong statistical trend and we were looking for corroborative evidence' [13, p. 375].

Overall four of the eight negative outcome measures specified in the protocol were replaced with four positive ones, many other negative measures having been tested and rejected along the way (see Table 1). The rationale given for the extra measures was that they were added according to "an analytical plan developed prior to opening of the blind" [14, p. 15]. No written evidence of this plan has been produced, raising uncertainty about Keller et al.'s claim that their "depression-related variables were declared *a priori*" [5, p. 764].

2.3. Conflation of primary and secondary outcomes

Many drafts of a report of the study were written before submission, initially to *JAMA*. In the first draft, the distinction between primary and secondary outcome variables in the protocol was removed so that all 8 outcomes were described as 'primary' in the results section [15]. However, in later drafts the term 'primary' was replaced with 'depression-related' [19]. These 1999 drafts reported that paroxetine was effective on the grounds that four out of eight of these measures were positive, without disclosing that there was no significant difference on either pre-specified *primary* outcome measure (see Table 1). *JAMA* rejected the paper in October, 1999, and it was revised for submission to *JAACAP*. One of the *JAMA* reviewers had noted that "the definition of remission and response overlaps in this manuscript" [20]. From the first draft for *JAACAP* in April 2000 'remission' (HAM-D $\leqslant 8$) was eliminated from the "depression-related variables … declared a priori" listed in the Methods section, thus reducing these

Box 2

Response (HAM-D $\leqslant$ 8)/remission (HAM-D $\leqslant$ 8 or reduced by $\geqslant$50%) conflation in the published paper [5]

Abstract

Under Method, the first main outcome measure is listed as "endpoint response" (defined as HAM-D $\leqslant$ 8 or reduced by $\geqslant$50%) but in the Results "HAM-D total score $\leqslant$ 8" appears for the first time and response is not mentioned (p. 762).

Method

Response is listed as an outcome and defined as "HAM-D $\leqslant$ 8 or reduced by $\geqslant$50%".

HAM-D $\leqslant$ 8 is not listed amongst the outcomes (p. 764).

Results

Efficacy Results (p. 765) begins with an explicit false claim:

> Of the depression-related variables, paroxetine separated statistically from placebo at end point among 4 of the parameters: [including] response (i.e. primary outcome measure).

In the following paragraph, where we would expect to find the response data, we instead see the data for "HAM-D total score $\leqslant$ 8 at end point".

All but the most careful readers would conclude that the HAM-D $\leqslant$ 8 figures being quoted show that response was a positive outcome.

Figures

The reader might easily assume that the two figures (p. 767) illustrate the two primary outcomes, but one of them is for HAM-D $\leqslant$ 8.

Discussion

Response is absent from the list of those items that did not separate statistically from placebo (p. 769). This change from earlier drafts [20] takes away a cue that might otherwise alert the reader to the repeated substitution of 'remission' for 'response'.

variables from eight to seven [21]. However the positive HAM-D $\leqslant$ 8 results were still reported in the Efficacy Results section, just where the reader would expect to find 'response' scores.

In July 2000, a *JAACAP* reviewer called for primary outcomes to be stated [22], forcing the authors to re-introduce them. But the lack of significant advantage for paroxetine on the two pre-specified primary outcomes was still not declared. Instead the authors continued to claim efficacy for paroxetine based on the conflation between response and HAM-D $\leqslant$ 8 that had begun in April 2000 with the elimination of 'remission' from the Method. This conflation now extended throughout the published paper, obscuring the negative primary outcome results by reporting positive HAM-D $\leqslant$ 8 results where negative 'response' (HAM-D $\leqslant$ 8 or reduced by $\geqslant$50%) results would be expected (see Box 2). Although HAM-D $\leqslant$ 8 is not listed as an outcome in the Method, its positive result is prominent in the text, appears at the head of the main results table and is the sole focus of Keller et al.'s Fig. 1. GSK subsequently sought to justify the conflation on the grounds that both 'response' and 'remission' were different ways of defining 'responders' [13, p. 287; 16]. But both the acute-phase protocol [14] and the published paper contained just one definition of responder: HAM-D $\leqslant$ 8 or reduced by $\geqslant$50%.

2.4. Presentation of other outcomes

The results of the other negative primary outcome measure, change from baseline HAM-D score, are also omitted from the text of the Results in the final paper. This outcome is graphically represented in a figure (Keller et al.'s Fig. 2), but without clear indication that the difference was non-significant. Only the main results table (Keller et al.'s Table 2) reports all eight outcomes accurately. The *JAACAP* paper

has been defended as follows: 'it clearly tells the reader in that table two all these variables are exactly described along with the exact key values so they could make their own decision on that' [13, p. 573].

3. Reporting of adverse effects

The abstract of the *JAACAP* paper states that: "Paroxetine was generally well tolerated in this adolescent population, and most adverse effects were not serious". Yet SKB's final report on the acute phase (completed in November 1998) documented many serious and severe adverse effects in the paroxetine group, several of them significantly more frequent than for placebo – see Table 2. Although suicidal thoughts and behaviour were grouped under the euphemism of 'emotional lability', Table 48 of SKB's internal final report [14, p. 109] clearly shows that five of the six occurrences of emotional lability were rated 'severe' and that all five had self-harmed or reported emergent suicidal ideas. Just a few minutes' reading of the serious adverse events narratives in this final report (pp. 276–307) would have revealed three more cases of suicidal ideas or self-harm that had not been classified as emotional lability. So the authors should have known that at least eight adolescents in the paroxetine group had self-harmed or reported emergent suicidal ideas compared to only one in the placebo group.

Relatively small numbers and brief follow up in RCTs lessen the likelihood of detecting serious adverse events (SAEs), so any signal should be highlighted. Yet early drafts of the paper prepared for *JAMA* did not discuss SAEs at all [15]. Subsequently SKB senior scientist McCafferty composed a paragraph on SAEs that appeared for the first time in the draft of July, 1999. It disclosed that 11 patients on paroxetine, compared to two on placebo, had SAEs, but did not mention the statistical significance of these figures. Subsequently McCafferty's disclosures of overdose and mania were edited out, and SAEs on paroxetine were attributed to other causes. Where McCafferty's draft reads:

> worsening depression, emotional lability, headache, and hostility were considered related or possibly related to treatment [20],

the published *JAACAP* paper states:

> only headache (1 patient) was considered by the treating investigator to be related to paroxetine treatment.

Table 2
Adverse events documented in SKB's final report of study 329 [14]

Type of adverse event	Paroxetine ($N = 93$)	Placebo ($N = 87$)	$p^{\wedge}$	Source table
Serious[#]	11 (12%)	2 (2.3%)	0.01	48, p. 109
Severe[##]	27 (29%)	15 (17%)	0.06	14.3.1, pp. 231–238
Hospitalisation	6[*] (6.5%)	0	0.004	48, p. 109
Nervous system				
Any	56 (60%)	29 (33%)	0.001	14.2.1, p. 227
Severe[**]	17 (18%)	4 (4.6%)	0.003	14.3.1, pp. 231–238
Requiring withdrawal	8 (8.6%)	2 (2.3%)	0.056	49, p. 111
Leading to dose reductions	8 (8.6%)	2 (2.3%)	0.056	46, p. 105

[^]Calculated by us; [#]'resulted in hospitalisation, was associated with suicidal gestures, or was described by the treating physician as serious' [5]; [##]'incapacitating and prevents normal everyday activities' [14, p. 565]; [*]stated as 7 in published paper; [**]stated as 16 for paroxetine and 3 for placebo in Table 44, p. 101.

4. Reporting study 329 to health professionals

From the late 1990s paroxetine was promoted to SKB/GSK Neuroscience sales representatives [23]. In August, 2001, a memorandum from Paxil Product Management to "all sales representatives selling Paxil" stated: "This 'cutting edge,' landmark study is the first to compare efficacy of an SSRI and a TCA with placebo in the treatment of major depression in adolescents. *Paxil* demonstrates REMARKABLE Efficacy and Safety in the treatment of adolescent depression" [7]. The memorandum mentioned only positive outcomes. By contrast, the cardiac adverse effects of imipramine were emphasised.

SKB/GSK produced a series of *Med Query Letters* to doctors who requested information about paroxetine for childhood depression via sales representatives. There is no publicly available information about whether or not sales representatives actively prompted doctors to request this information. *Letters* characteristically started and ended with disclaimers like "*Paxil* is not FDA-approved for use in children or adolescents; therefore, we may not offer any recommendations regarding the use of *Paxil* in these patients" [24], but nonetheless provided selected information about study 329. For example, *Letters* omitted primary outcome results (1998 [25], 1999 [26], 2001 [24]) and serious adverse event results (1998, 1999, 2000 [27], 2001), or failed to mention other negative childhood depression studies when these results became available (2000, 2001). Other academic publications and presentations frequently did not disclose the results for the primary outcomes and serious adverse events [19,28–33].

5. Discussion

5.1. Were the results for study 329 positive or negative?

There was no significant efficacy difference between paroxetine and placebo on the two primary outcomes or six secondary outcomes in the original protocol. At least 19 additional outcomes were tested. Study 329 was positive on 4 of 27 known outcomes (15%). There was a significantly higher rate of SAEs with paroxetine than with placebo. Consequently, study 329 was negative for efficacy and positive for harm.

5.2. Did selective reporting occur?

Claims that paroxetine was "generally well tolerated and effective" [5] arose from selective reporting of the 15% of outcomes that were positive and selective under reporting of the other efficacy and SAE findings. The *JAACAP* paper has been defended on the grounds that readers could read in the results table that the two outcomes described as primary elsewhere (but not in that table) were negative [13]. However readers are more likely to be influenced by the abstract than the tables of a clinical trial report, as evidenced by the continued retransmitting of the false impression that study 329 found "significant efficacy on one of the two primary endpoints" [8]. A likely cause of this misunderstanding is the conflation of 'remission' and 'responder' and especially the false statement that "paroxetine separated statistically from placebo at end point among 4 of the parameters: [including] response (i.e. primary outcome measure) . . ." [5].

5.3. How did selective reporting happen?

In response to criticism in *JAACAP* in 2003, Keller et al. [34] indicated that they believed that paroxetine was effective and therefore viewed the efficacy results as a false negative arising from their mistake of using the HAM-D as their depression measure. They then searched for other outcomes that matched their beliefs about efficacy. Such searching has been described as "data torturing" [35], a form of confirmation bias in which information is sought to support pre-conceived beliefs. Confirmation bias could also lead authors who were unconcerned about adverse events to look less closely at that data and to attribute adverse events in the paroxetine group to non-drug causes such as "arguments with boyfriends" [36]. Confirmation bias could be well-intentioned, so that investigators might believe that what they had done was entirely appropriate. However it does not explain the conflation of 'remission' and 'responder', the changes to the descriptions of SAEs, or flaws that were detected by peer reviewers but were not corrected.

6. Conclusions

Since the publication of the results of study 329 in 2001, suspicions have emerged about its selective reporting [37,38]. Our detailed case study of proprietary documents from GSK regarding this study adds to the evidence that flaws in industry-funded research can be severe, and difficult to detect. The documents reveal that the published conclusions of study 329 and information provided by GSK to health professionals understated adverse effect rates and emphasised post-hoc measures that were not consistent with the unpublished, protocol-defined primary and secondary outcomes.

Funding: none.

Acknowledgements

We wish to thank Skip Murgatroyd, Michael Baum and Cindy Hall of Baum Hedlund, Peter Baghurst for statistical advice, and Shelley Jofre of BBC Scotland.

References

[1] H. Melander, J. Ahlqvist-Rastad, G. Meijer and B. Beermann, Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications, *BMJ* **326** (2003), 1171–1173.
[2] A.-W. Chan, A. Hróbjartsson, M.T. Haahr, P.C. Gøtzsche and D.G. Altman, Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles, *JAMA* **291** (2004), 2457–2465.
[3] E.J. Garland, Facing the evidence: Antidepressant treatment in children and adolescents, *CMAJ* **170** (2004), 489–491.
[4] J.N. Jureidini, C.J. Doecke, P.R. Mansfield, M.M. Haby, D.B. Menkes and A.L. Tonkin, Efficacy and safety of antidepressants for children and adolescents, *BMJ* **328** (2004), 879–883.
[5] M.B. Keller, N.D. Ryan, M. Strober, R.G. Klein, S.P. Kutcher, B. Birmaher et al. Efficacy of paroxetine in the treatment of adolescent major depression: a randomized, controlled trial, *J. Am. Acad. Child Adolesc. Psychiatr.* **40** (2001), 762–772.
[6] Journal Citation Reports, http://thomsonscientific.com/products/jcr/.
[7] Hawkins to all sales representatives selling Paxil, Aug 16, 2001, PAR000651577; http://www.healthyskepticism.org/documents/documents/20010816Hawkinstoreps.pdf (accessed 10 March, 2007).
[8] C. Moreno, C. Arango, M. Parellada, D. Shaffer and H. Bird, Antidepressants in child and adolescent depression: where are the bugs?, *Acta Psychiatr. Scand.* **115** (2007), 184–195.
[9] T.A. Hammad, T. Laughren and J. Racoosin, Suicidality in pediatric patients treated with antidepressant drugs, *Arch. Gen. Psychiatry* **63** (2006), 332–339.

[10] A.S. Kesselheim and J. Avorn, The role of litigation in defining drug risks, *JAMA* **297** (2007), 308–311.
[11] Adolescent unipolar major depression: Multisite psychopharmacology study, draft of 5 Dec, 1992, PAR000754956; http://www.healthyskepticism.org/documents/documents/19921205KellerProposal.pdf (accessed 10 March, 2007).
[12] SmithKline Beecham, A multi-center, double-blind, placebo controlled study of paroxetine and imipramine in adolescents with unipolar major depression 1993/amended 1996, RYAN002807-45; www.healthyskepticism.org/documents/Protocol329.pdf.
[13] McCafferty Deposition, 24–26 Aug, 2006; http://www.healthyskepticism.org/documents/documents/McCaffertyDeposition.pdf (accessed 10 March, 2007).
[14] SmithKline Beecham, A multi-center, double-blind, placebo controlled study of paroxetine and imipramine in adolescents with unipolar major depression – acute phase, Final clinical report, SB Document Number: BRL-029060/RSD-100TW9/1/CPMS-329, 24 November, 1998; http://www.gsk.com/media/paroxetine/depression_329_full.pdf (accessed March 2007).
[15] Draft I. 18 Dec, 1998, PAR004368103-136; http://www.healthyskepticism.org/documents/documents/DraftI.pdf (accessed 10 March, 2007).
[16] Draft minutes of teleconference re paroxetine study 329 efficacy analysis, 22 Apr, 1997, PAR000753425-9; http://www.healthyskepticism.org/documents/documents/970422teleconference.pdf (accessed 10 March, 2007).
[17] Top Line Results 1998/01/21 RYAN002664; http://www.healthyskepticism.org/documents/documents/TopLineResultsRyannotes.pdf (accessed 10 March, 2007).
[18] Ryan Deposition 5 Oct 2006; http://www.healthyskepticism.org/documents/documents/RyanDeposition.pdf (accessed 10 March, 2007).
[19] Draft submitted to JAMA, 30 July, 1999, PAR000212704-39; http://www.healthyskepticism.org/documents/documents/19990726JAMAsubmission.pdf (accessed 10 March, 2007).
[20] S. Laden, Response to JAMA review, 10 Dec, 1999, PAR000757046-58; http://www.healthyskepticism.org/documents/documents/ResponsestoJAMA.pdf (accessed 10 March, 2007).
[21] Draft VI. 25 Apr, 2000, PAR000756969-7009; http://www.healthyskepticism.org/documents/documents/20000425draftVI.pdf (accessed 10 March, 2007).
[22] Response to *JAACAP* reviewers, November 3, 2000, PAR008947745-8; http://www.healthyskepticism.org/documents/documents/responsetoJAACAPReviews.pdf (accessed 10 March, 2007).
[23] SmithKline Beecham, *Nulli Secundus* 8 Dec, 1999, PAR000094170; http://www.healthyskepticism.org/documents/documents/NulliSecundus_000.pdf (accessed 10 March, 2007).
[24] GlaxoSmithKline, Med Query Letter, 11 Sep, 2001, PAR001986848-57; http://www.healthyskepticism.org/documents/documents/2001MedQuery.pdf (accessed 10 March, 2007).
[25] SmithKline Beecham, Med Query Letter, 1 Jul, 1998, PAR001986766-74; http://www.healthyskepticism.org/documents/documents/1998MedQuery.pdf (accessed 10 March, 2007).
[26] SmithKline Beecham, Med Query Letter, 15 Jul, 1999, PAR001986804-14; http://www.healthyskepticism.org/documents/documents/1999MedQuery.pdf (accessed 10 March, 2007).
[27] SmithKline Beecham, Med Query Letter, 4 Jan, 2000, PAR001986827-37; http://www.healthyskepticism.org/documents/documents/2000MedQuery.pdf (accessed 10 March, 2007).
[28] M.B. Keller, N.D. Ryan, B. Birmaher, R.G. Klein, M. Strober, K.D. Wagner and E.B. Weller, Paroxetine and imipramine in the treatment of adolescent depression (Abstract NR562:209), Presentation to American Psychiatric Association, Toronto, 2 Jun, 1998, PAR000599842; http://www.healthyskepticism.org/documents/documents/KellerPoster.pdf (accessed 10 March, 2007).
[29] K. Wagner, B. Birmaher, G. Carlson, G. Clarke, G. Emslie, B. Geller et al., Safety of paroxetine and imipramine in the treatment of adolescent depression, Poster No. 69 at New Clinical Drug Evaluation Unit Annual Meeting, Boca Raton, FL, 11 Jun, 1998, PAR001335408; http://www.healthyskepticism.org/documents/documents/WagnerPoster.pdf (accessed 10 March, 2007).
[30] R. Berard and N. Ryan, Adolescent depression: Efficacy of paroxetine, Poster at European College of Neuropsychopharmacology, Paris, Oct 1998, PAR001986606; http://www.healthyskepticism.org/documents/documents/BerardPoster.pdf (accessed 10 March, 2007).
[31] C. Gagiono, Paroxetine in adolescent depression, Poster at World Congress of Psychiatry, Hamburg, Oct 1999, PAR001982238; http://www.healthyskepticism.org/documents/documents/GagianoPoster.pdf (accessed 10 March, 2007).
[32] K.D. Wagner, Paroxetine treatment of mood and anxiety disorders in children and adolescents, *Psychopharm. Bull.* **37**(S1) (2003), 167–175.
[33] A. Benbow, Depressing misrepresentation?, *Lancet* **363** (2004), 1732–1733.
[34] M.B. Keller, N.D. Ryan, M. Strober, E.B. Weller, J.P. McCafferty, O.R. Hagino et al., Paroxetine in major depression (reply), *J. Am. Acad. Child Adolesc. Psychiatr.* **42** (2003), 514–515.
[35] J.L. Mills, Data torturing, *NEJM* **329** (1993), 1196–1199.

[36] M.B. Keller, N.D. Ryan and K.D. Wagner, Paroxetine in adolescent major depression (reply), *J. Am. Acad. Child Adolesc. Psychiatr.* **41** (2002), 364.
[37] SmithKline Beecham, Seroxat/Paxil adolescent depression position piece on the Phase III clinical studies, October 1998, PAR003019178; http://www.healthyskepticism.org/documents/documents/19981014PositionPiece.pdf (accessed July 2007).
[38] J. Jureidini and A. Tonkin, Paroxetine in major depression, *J. Am. Acad. Child Adolesc. Psychiatr.* **42** (2003), 514.
[39] J. Leo, The SSRI trials in children: disturbing implications for academic medicine, *Ethic. Hum. Psychol. Psychiatr.* **8** (2006), 33–34.