

Research report

Effectiveness of pindolol plus serotonin uptake inhibitors in depression: a meta-analysis of early and late outcomes from randomised controlled trials

Javier Ballesteros^{a,b,*}, Luis F. Callado^c

^aDepartment of Neuroscience, Medical Psychology and Psychiatry, University of the Basque Country, Barrio Sarriena S/N, E-48940 Leioa, Spain

^bInstitute of Psychiatric Research, María Josefa Recio Foundation, Bilbao, Spain

^cDepartment of Pharmacology, University of the Basque Country, Leioa, Spain

Received 10 May 2002; received in revised form 21 October 2002; accepted 8 November 2002

Abstract

Background: Contradictory results on the efficacy of pindolol associated with selective serotonin reuptake inhibitors (SSRIs) in depressive illness have been published and no former review has produced an overall figure of its efficacy. This study aims to review the efficacy and tolerability of pindolol plus SSRIs in depressive illness. **Methods:** A meta-analysis of randomised controlled trials (RCTs) comparing pindolol plus SSRIs with placebo plus SSRIs. **Results:** Nine RCTs met inclusion criteria. Outcome favoured pindolol at 2 weeks time ($N=5$; $OR=2.8$; 95% CI 1.4–5.7), but not at four to 6 weeks ($N=7$; $OR=1.4$; 95% CI 0.8–2.7). Results for early outcome studies were robust to sensitivity analysis. Nineteen more studies, averaging null results, would be needed to change the overall probability ($P=0.0001$) to a non-significant figure. **Conclusions:** Pindolol seems to hasten the response to SSRIs in depression with a timing window circumscribed to the first weeks of treatment.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Depression; Effectiveness; Meta-analysis; Outcome; Pindolol; SSRI

1. Background

Blockade of serotonin (5-HT) reuptake is a common property of many antidepressant drugs, but although selective 5-HT reuptake inhibitors (SSRIs) block monoamine uptake within hours of administra-

tion to patients, its full clinical effect does not appear until 2–3 weeks after treatment onset (Sugrue, 1983; Leonard, 1984). This unwanted delay has brought attention towards treatment strategies that might shorten that time lag (Artigas et al., 1996; Blier et al., 1997; Zanardi et al., 1997; Montgomery, 1999). Experimental research has shown that SSRIs, such as fluvoxamine or citalopram, increase the extracellular concentration of 5-HT not only at synaptic terminals but also in the proximity of cell bodies and dendrites

* Corresponding author. Tel.: +34-94-601-2799; fax: +34-94-464-9266.

E-mail address: onpbaroj@lg.ehu.es (J. Ballesteros).

of 5-HT neurones in the raphe nucleus (Bel and Artigas, 1992; Invernizzi et al., 1992; Blier and Bergeron, 1998). The extracellular increase of 5-HT would trigger the action of inhibitory somatodendritic 5-HT_{1A} autoreceptors. These, in turn, would limit the extra-cellular rise of 5-HT by inhibiting the firing of 5-HT neurons, at the same time as the 5-HT transporter at synaptic endings is blocked by the SSRIs (Artigas et al., 1996). On such grounds, current hypothesis postulate that the delay of clinical antidepressant effects might be due to the time 5-HT_{1A} autoreceptors need to desensitise and, therefore, decrease the feedback inhibition of monoamine release (Blier and de Montigny, 1994). If this were true pharmacological blockade of 5-HT_{1A} autoreceptors should mimic the effect of autoreceptor desensitisation and hasten the antidepressant effects of SSRIs. Open clinical studies have backed this hypothesis by showing a faster onset of antidepressant effects by SSRIs when associated with the 5HT_{1A}/β-adrenoreceptor antagonist pindolol (Artigas et al., 1994; Blier and Bergeron, 1995). Nevertheless, the expectations raised by open clinical studies and further supported by randomised controlled clinical trials (RCTs) (Pérez et al., 1997, 1999; Tome et al., 1997; Zanardi et al., 1997) have been challenged by others (Berman et al., 1997; Moreno et al., 1997). As a result, heterogeneous findings have been reported on the efficacy of SSRIs plus pindolol either for early outcomes—outcomes assessed up to 2 weeks of treatment—(Bordet et al., 1998; Pérez et al., 1999) or late outcomes—at 4–6 weeks since randomisation to treatment arms (Berman et al., 1997; Pérez et al., 1997; Zanardi et al., 1998). Because of the uncertainty of the potential benefit of pindolol associated with SSRIs in the treatment of depression we undertook a systematic review of published RCTs to investigate its efficacy. We were interested in the assessment of both early and late outcomes. None of the formerly published reviews on this issue have reported an overall figure of efficacy (Blier and Bergeron, 1998; McAskil et al., 1998; Montgomery, 1999; Puzantian and Kawase, 1999; Olver et al., 2000; Artigas et al., 2001). This review also focuses on both the rate of drop-outs not explicitly due to adverse events—as a surrogate indicator for tolerability—and the safety of the SSRIs plus pindolol association as assessed by the rate of reported adverse events.

2. Methods

2.1. Inclusion and exclusion criteria

Peer reviewed published RCTs using a parallel design for comparing pindolol plus SSRIs against placebo plus SSRIs in depressive illness. A minimum follow-up of 2 weeks and a maximum of 6 weeks was required to assess early and/or late outcomes, respectively. Adequate concealment of treatment to trial arms and frequency data on clinical outcomes either directly reported or extractable from the methods and/or results section of the original paper. Crossover studies were excluded as were those using any non SSRI antidepressant as main or unique treatment.

2.2. Search strategy

A computerised search was done by using a combination of the following keywords—either as free text or mapping them to appropriate thesaurus terms: SSRI* or FLUOXETINE or FLUVOXAMINE or SERTRALINE or PAROXETINE or CITALOPRAM or SEROTONIN-UP-TAKE-INHIBITORS) and (PINDOLOL or SEROTONIN-ANTAGONISTS). No language restriction was imposed. The searched databases were: MEDLINE (1966 to 12/2001), and EMBASE (1980 to 12/2001) through Ovid; PSYCINFO (1967 to 12/2001) and CURRENT CONTENTS (1995 to 12/2001) through WinSPIRS. A secondary search was done by hand searching the reference lists from primary studies and former reviews. All retrieved references were managed with the Reference Manager v.9.5 program (ISI Research-Soft, Berkeley, CA).

2.3. Selected outcomes

Efficacy was assessed by the number of patients who responded to treatment (a decrease of $\geq 50\%$ —or similar criterion—in depression rating scores since random allocation) out of the total number of randomised patients. The Hamilton Depression Rating Scale (HDRS) (Hamilton, 1960, 1967) and the Montgomery-Åsberg Depression Rating Scale (MADRS) (Montgomery and Åsberg, 1979) were selected as outcome measures since both scales have shown

similar changes through time in psychopharmacological studies and therefore provide comparable data (Lauge et al., 1998).

Tolerability was assessed by the number of patients failing to complete the study out of the total number of randomised patients. Safety was assessed by the number of subjects who complained of side effects out of the total number of randomised patients.

2.4. Data extraction

Data on outcome frequency and ancillary information were extracted by one reviewer (JB) and checked for accuracy and completeness by the other (LFC). For each study, frequencies for the main outcome were extracted to conform to a two-by-two table.

2.5. Statistical analysis

Results are expressed as odds-ratios (OR) to assess the strength of the effect (Fleiss, 1993) and as the number of patients needed-to-treat to benefit (NNTB) or to harm (NNTH) to evaluate its clinical importance (McQuay and Moore, 1997; Altman, 1998). The studies were individually reanalysed to get an estimate of the OR and its 95% confidence interval (95% CI). Fisher's test was used to get exact probabilities from two-by-two tables.

Because of the different clinical conditions and settings represented in the studies we expected that data sets on efficacy would be heterogeneous. Therefore, the statistical combination of results was done by a random effects model (DerSimonian and Laird, 1986) and complemented with a full Bayesian random effects model (Smith et al., 1995). Post hoc influence or sensitivity analysis was done by leaving out studies and checking the consistency of the overall effect estimate. The influence of publication bias was indirectly assessed by the fail safe N method (Rosenthal, 1979) as the few expected studies retrieved would make it unreliable to run more formal evaluations based on the asymmetry of funnel plots. To do that, an overall probability value was obtained by combining one-side *P*-values by the method of Fisher (Hedges and Olkin, 1985). For tolerability and safety analysis a fixed effects model by the Mantel-

Haenszel approach followed by a *Q*-test of heterogeneity, was used to combine the individual OR's (Fleiss, 1993).

Statistical analysis for fixed effect models and DerSimonian and Laird random effect models were done with specific macros written for Stata (Bradburn et al., 1998; StataCorp, 2001). The full Bayesian random effects models were run with WinBUGS (Spiegelhalter et al., 2000) assuming a normal distribution for the log OR with non-informative priors for the unknown parameters. The Bayesian estimates (medians and 95% credible intervals) were obtained through 5000 iterations after a burn-in of 5000 iterations.

3. Results

3.1. Characteristics of included studies

Nine studies out of eleven selected papers met the inclusion criteria (Berman et al., 1997, 1999; Pérez et al., 1997, 1999; Tome et al., 1997; Zanardi et al., 1997, 1998; Bordet et al., 1998; Maes et al., 1999). Of the excluded papers, one was a cross-over trial (Moreno et al., 1997) and another did not provide the necessary information on outcome frequency (Smeraldi et al., 1998). All studies reported in former reviews were tracked by the computerized search and included in the meta-analysis except Moreno et al. (1997) for the reason stated above and Maes et al. (1996) because of using an heterocyclic antidepressant (trazodone) at subtherapeutic doses as main treatment. Overall, the studies provided data from 594 patients (298 in the pindolol plus SSRI's group, 296 in the placebo group). Seven studies contributed data on late clinical outcomes (4–6 weeks of treatment) (Berman et al., 1997, 1999; Pérez et al., 1997; Tome et al., 1997; Zanardi et al., 1997, 1998; Maes et al., 1999) whereas five presented appropriate data to assess early clinical response—10 days to 2 weeks—(Tome et al., 1997; Zanardi et al., 1997, 1998; Bordet et al., 1998; Pérez et al., 1999).

The primary studies were carefully checked to avoid including multiple publications based on the same results. Two serial papers (Berman et al., 1997, 1999) report data from the same design but by accumulating the number of participants. We came

to the decision to include these studies as two independent pieces of data by extracting the full results as appeared on the first paper but only data corresponding to the second wave from the other report. This was done by retrieving information through two-by-two tables and removing from the second study the frequencies formerly reported. In this way, both studies were included in the meta-analysis set without biasing the overall effect estimate either because of multiple publications or by giving inappropriate more weight to the accumulated numbers of the second report.

The study by Zanardi et al. (1997) included a comparative arm with paroxetine plus pindolol during 1 week which was discarded because it did not conform with inclusion criteria. Bordet et al. (1998) used pindolol for 3 weeks but full frequency data is only reported for day 10 of the trial. Only at that time paroxetine plus pindolol outperformed paroxetine plus placebo. To assess the bias this study could have on the overall results it was selected for sensitivity analysis. Maes et al. (1999) trial arm with fluoxetine plus mianserine was discarded because it did not conform with inclusion criteria. Also the authors randomised 34 patients but analysed only 31 since three individuals withdrew in the first week of treatment before any post-baseline measure were available. Following the intention-to-treat principle our estimates were made over the number of initially randomised patients.

The study by Pérez et al. (1999) randomised either pindolol or placebo to patients with treatment-refractory depression who were treated with 5-HT reuptake inhibitors, including SSRIs (fluoxetine, fluvoxamine, paroxetine) and clomipramine, at the time of randomization. As it was not possible from the published data to keep apart clomipramine users from the rest, all patients on 5-HT reuptake inhibitors were included in the meta-analysis. Because of these features this study was also selected for sensitivity analysis.

Table 1 shows the main characteristics of the nine studies reviewed. Four studies used fluoxetine, three paroxetine, one fluvoxamine and another several SSRIs, in addition to pindolol or placebo. Female patients outnumbered males in all studies except two (Berman et al., 1997; Maes et al., 1999). All studies used the intention-to-treat principle for analysing results. Definition of clinical response varied between

studies—with some reporting more than one criterion—therefore, we extracted data from comparable criteria (Table 1). Overall, data from four studies reach conventional statistical significance ($P \leq 0.05$) supporting the pindolol plus SSRIs association, whereas the other five do not favour such an association (Table 1).

3.2. Assessment of efficacy for early clinical outcome

Table 2 shows the main results of the five studies which reported data to assess the early response to SSRIs plus pindolol. Fig. 1 displays the individual study results and the combined estimate. Two studies favoured the efficacy of SSRIs plus pindolol, another two are inconclusive with OR greater than unity but with 95% CI including the null hypothesis value for the OR, and one study has an OR equal to one. There was no significant heterogeneity between studies (Q -test=6.2, 4 df; $P=0.18$) and the combined estimate clearly favours the efficacy of SSRIs plus pindolol over SSRIs plus placebo (DerSimonian and Laird random effects model OR=2.8, 95% CI 1.4–5.7; Z -test=2.96, $P=0.003$; full Bayesian random effects model OR=2.9, 95% credible interval 1.5–7.3; NNTB=6, 95% CI 4–20).

This result was consistent after sensitivity analysis was done by deleting two problematic studies or one study at a time. The deletion of Bordet et al. (1998) study barely changed the overall results (DerSimonian and Laird model OR=3.2, 95% CI 1.1–9.0; Z -test=2.18, $P=0.03$; full Bayesian model OR=3.2, 95% credible interval 1.2–14.9; NNTB=6, 95% CI 3–100). Similar result was found when Pérez et al. (1999) study was deleted (DerSimonian and Laird model OR=3.4, 95% CI 1.8–6.6; Z -test=3.66, $P<0.001$; full Bayesian model OR=3.7, 95% credible interval 1.8–11.4; NNTB=5, 95% CI 3–10). The deletion of both studies increased somewhat the pooled effect estimate at the cost of increasing its imprecision because of combining only three studies (DerSimonian and Laird model OR=4.7, 95% CI 1.6–14.2; Z -test=2.75, $P=0.006$; full Bayesian model OR=4.9, 95% credible interval 1.3–38.7; NNTB=4, 95% CI 3–17). Deleting one study at a time, the most conservative effect estimate was obtained by the deletion of Zanardi et al. (1998) study (DerSimonian and Laird model OR=2.6, 95%

Table 1
Characteristics of studies included in the systematic review^a

Authors (Year)	Sample and setting	Design and intervention	Response criterion selected for the systematic review	Main results ^b
Berman et al. (1997)	43 MD outpatients (DSM-IV criteria) with HDRS-25 score ≥ 18 at baseline Age range: 18–70 years Sex ratio: 44.2% females	Two arms RCT: FLX (20 mg/day)+POL (7.5 mg/day) vs. FLX (20 mg/day)+PBO Length of trial: 6 weeks	A decrease $\geq 50\%$ in total HDRS-25 score from baseline to endpoint with a maximum post-treatment score of 15 (ITT analysis) Reported as partial remission	Rates of partial remission did not differ between POL and PBO groups (Fisher's exact $P=0.34$ as reported by authors)
Pérez et al. (1997)	111 unipolar MD outpatients (DSM-IV criteria) with HDRS-17 score ≥ 18 at baseline Age: 18 years or older Sex ratio: 71.2% females	Two arms RCT after 1 week of placebo run-in period: FLX (20 mg/day)+POL (7.5 mg/day) vs. FLX (20 mg/day)+PBO Length of trial: 6 weeks	A decrease $\geq 50\%$ in total HDRS-17 score from baseline to endpoint (ITT analysis) Reported as clinical response	The proportion of responders was greater in the POL group than in the PBO group ($P=0.04$ as reported by authors)
Tome et al. (1997)	80 outpatients with a diagnosis of mild, moderate or severe unipolar depression without psychotic symptoms (ICD-10 criteria) with MADRS score ≥ 18 at baseline Age range: 18–65 years Sex ratio: 60% females	Two arms RCT: PRX (20 mg/day)+POL (7.5 mg/day) vs. PRX (20 mg/day)+PBO Length of trial: 6 weeks	A decrease $\geq 50\%$ in total MADRS score from baseline to endpoint (ITT analysis)	Rates of improvement did not differ between the POL and PBO group at the end of trial (Fisher's exact $P=1.0$ by secondary analysis)
Zanardi et al. (1997)	42 inpatients with recurrent MD (DSM-IV criteria) with HDRS ≥ 18 at baseline Age range: 18–65 years Sex ratio: 69% females	Two arms RCT after 1 week of placebo run-in period: PRX (20 mg/day)+POL (7.5 mg/day) vs. PRX (20 mg/day)+PBO Length of trial: 4 weeks (The design included another arm with PRX+POL for 1 week not appropriate for inclusion)	A total score in the HDRS ≤ 8 at endpoint (ITT analysis)	Rates of improvement were marginally not significant between the POL and PBO group at the end of trial (Fisher's exact $P=0.05$ by secondary analysis)
Bordet et al. (1998)	100 inpatients and outpatients with unipolar MD nonpsychotic subtype (DSM-IV criteria) with HDRS-17 ≥ 18 at baseline Age range: 18–65 years Sex ratio: 70% females	Two arms RCT: PRX (20 mg/day)+POL (15 mg/day) vs. PRX (20 mg/day)+PBO Length of trial: 3 weeks	A total score in the HDRS-17 ≤ 10 at endpoint (ITT analysis)	Rates of improvement favored the POL group at 10 days when full data is only reported (Fisher's exact $P=0.04$ by secondary analysis)
Zanardi et al. (1998)	72 inpatients with MD with psychotic features, MD single or recurrent episode and bipolar disorder (DSM-III-R criteria) Age range: 18–65 years Sex ratio: 77.8% females	Two arms RCT after 1 week of placebo run-in period: FVX (150 mg/day)+POL (7.5 mg/day) vs. FVX (150 mg/day)+PBO Length of trial: 6 weeks	A total score in the HDRS-21 ≤ 8 at endpoint (ITT analysis)	Rates of improvement did not differ between the POL and the PBO group at endpoint (Fisher's exact $P=1.0$ by secondary analysis)
Berman et al. (1999)	43 MD outpatients (DSM-IV criteria) with HDRS-25 score ≥ 18 at baseline Age range: 18–70 years Sex ratio: 75% females	Two arms RCT: FLX (20 mg/day)+POL (7.5 mg/day) vs. FLX (20 mg/day)+PBO Length of trial: 6 weeks	A decrease $\geq 50\%$ in total HDRS-25 score from baseline to endpoint with a maximum post-treatment score of 15 (ITT analysis) Reported as partial remission	Rates of partial remission did not differ between POL and PBO groups (Fisher's exact, $P=1.0$ by secondary analysis)
Maes et al. (1999)	21 MD inpatients—mainly with melancholia and TRD-(DSM-III-R criteria) with HDRS-17 score ≥ 16 at baseline Age range: 25–70 years Sex ratio: 38.1% females	Two arms RCT: FLX (20 mg/day)+POL (7.5 mg/day) vs. FLX (20 mg/day)+PBO Length of trial: 5 weeks (The design included another arm with FLX+mianserine not appropriate for inclusion)	A decrease $\geq 50\%$ in total HDRS-17 score from baseline to endpoint (ITT analysis)	Rates of improvement favored the POL group at endpoint (Fisher's exact $P=0.03$ by secondary analysis)

(continued on next page)

Table 1 (continued)

Authors (Year)	Sample and setting	Design and intervention	Response criterion selected for the systematic review	Main results ^b
Pérez et al. (1999)	80 outpatients diagnosed of MD single or recurrent (DSM-IV criteria) with a current TRD episode and with HDRS-17 total score >16 Age range: 18–65 years Sex ratio: 71.3% females	Two arms RCT after a 5 days placebo run-in period: SSRIs and 5-HT reuptake inhibitors + POL (7.5 mg/day) vs SSRIs and 5-HT reuptake inhibitors + PBO Length of trial: 10 days	A decrease $\geq 50\%$ in total HDRS-17 score from baseline to endpoint (ITT analysis)	The percentage of responder patients was the same in the two treatment arms (Fisher's exact $P=1.0$ by secondary analysis)

^a MD, indicates major depression; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders 4th edition; HDRS, Hamilton Depression Rating Scale; RCT, randomized controlled trial; FLX, fluoxetine; POL, pindolol; PBO, placebo; ITT, intention-to-treat; PRX, paroxetine; ICD-10, International Classification of Mental and Behavioural Disorders 10th revision; MADRS, Montgomery and Åsberg Depression Rating Scale; DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders 3rd edition revised; FVX, fluvoxamine; TRD, treatment-resistant depression; SSRIs, selective serotonin reuptake inhibitors; 5-HT, serotonin.

^b P -values reported are either extracted from the original papers or calculated by secondary analysis.

CI 1.2–5.6; Z -test = 2.51, $P=0.01$; full Bayesian model OR = 2.8, 95% credible interval 1.2–7.3; NNTB = 5, 95% CI 3–59).

With the overall result of the five data sets, 19 extra studies—confirming the null hypothesis in average—would be needed to raise the overall probability value obtained by the method of Fisher ($P=0.0001$) above the classical cut-off value of 0.05.

3.3. Assessment of efficacy for late clinical outcome

Table 2 shows the main results of the seven studies which contributed data to assess the efficacy of SSRIs plus pindolol at the end of RCTs. Fig. 2 displays the individual study results and the combined estimate. Two studies favoured the efficacy of SSRIs plus pindolol, another two are inconclusive with OR great-

Table 2

Outcomes from the studies included in the systematic review^a

Authors (Year)	Early clinical response		P -value	Clinical response at the end of trial		P -value
	SSRIs + pindolol No. improved/total no. (Event rate, %)	SRIs + placebo No. improved/total no. (Event rate, %)		SSRIs + pindolol No. improved/total no. (Event rate, %)	SRIs + placebo No. improved/total no. (Event rate, %)	
Berman et al. (1997)	NA	NA	NA	13/23 (56.5)	15/20 (75.0)	0.34
Pérez et al. (1997)	NA	NA	NA	41/55 (74.5)	33/56 (58.9)	0.11
Tome et al. (1997)	19/40 (47.5)	11/40 (27.5)	0.10	21/40 (52.5)	22/40 (55.0)	1.0
Zanardi et al. (1997)	12/21 (57.1)	2/21 (9.5)	0.002	17/21 (81.0)	10/21 (47.6)	0.05
Bordet et al. (1998)	24/50 (48.0)	13/50 (26.0)	0.04	NA	NA	NA
Zanardi et al. (1998)	6/36 (16.7)	1/36 (2.8)	0.11	29/36 (80.6)	28/36 (77.8)	1.0
Berman et al. (1999)	NA	NA	NA	15/22 (68.2)	15/21 (71.4)	1.0
Maes et al. (1999)	NA	NA	NA	6/11 (54.5)	1/12 (8.3)	0.03
Pérez et al. (1999)	5/40 (12.5)	5/40 (12.5)	1.0	NA	NA	NA

^a SSRIs, indicates selective serotonin reuptake inhibitors; NA, data not available. All P -values report 2-sided Fisher's exact probabilities calculated by secondary analysis.

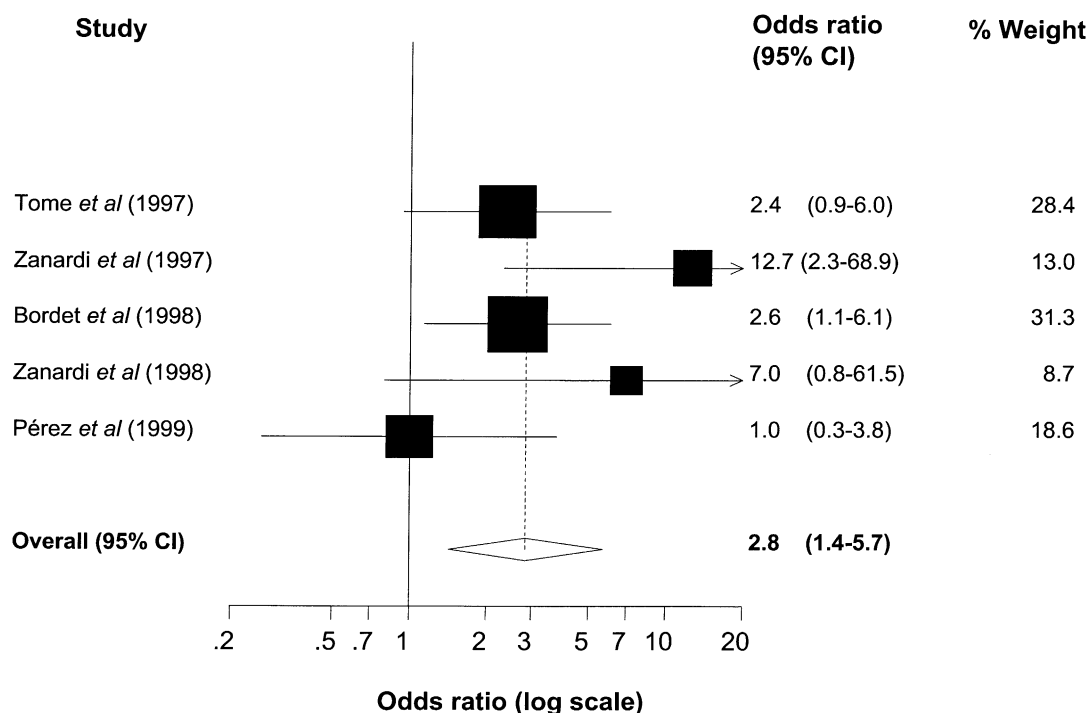


Fig. 1. Odds ratios of randomised trials and summary overall estimate for early response studies (10 days to 2 weeks). DerSimonian and Laird random effects model. Solid squares are proportional to individual study weights.

er than unity but wide 95% CI, and three studies have an OR lower than one. There was evidence of heterogeneity (Q -test = 11.9, 6 df, $P = 0.06$), which was singularly accounted for by the results from Maes *et al.* (1999) study. Its deletion reduced unambiguously the heterogeneity (Q -test = 8.3, 5 df, $P = 0.14$). The combined estimate does not favour the efficacy of SSRIs plus pindolol over SSRIs plus placebo (DerSimonian and Laird random effects model OR = 1.4, 95% CI 0.8–2.7; Z -test = 1.12, $P = 0.26$; full Bayesian random effects model OR = 1.5, 95% credible interval 0.8–2.9; NNTB = 12, [NNTH 20 to ∞ to NNTB 5]). Since the pooled results did not support here the efficacy of pindolol no further sensitivity analysis or fail safe N were performed.

3.4. Efficacy as compared at both early and late clinical outcome

Three studies (Tome *et al.*, 1997; Zanardi *et al.*, 1997, 1998) provided data to analyse the efficacy of

pindolol at two selected time points over the clinical trial: at 2 weeks and at the end of trial. Their results matched the pattern previously found. At 2 weeks the pooled outcome favoured pindolol (DerSimonian and Laird model OR = 4.7, 95% CI 1.6–14.2; Z -test = 2.75, $P = 0.006$; full Bayesian model OR = 4.9, 95% credible interval 1.3–38.7) whereas the difference between arms vanished by the end of trial (DerSimonian and Laird model OR = 1.5, 95% CI 0.6–3.7; Z -test = 0.90, $P = 0.37$; full Bayesian model OR = 1.4, 95% credible interval 0.4–6.4).

3.5. Tolerability and adverse events

Overall, most of the randomized patients finished the study. There were not differences between pindolol plus SSRIs and placebo plus SSRIs on tolerability (Mantel-Haenszel pooled OR = 1.3, 95% CI 0.8–2.3; Z -test = 1.12, $P = 0.26$; Q -test = 1.2, df = 6, $P = 0.98$) or adverse events (Mantel-Haenszel pooled OR = 1.3, 95% CI 0.7–2.1; Z -test = 0.91, $P = 0.36$; Q -test = 4.3,

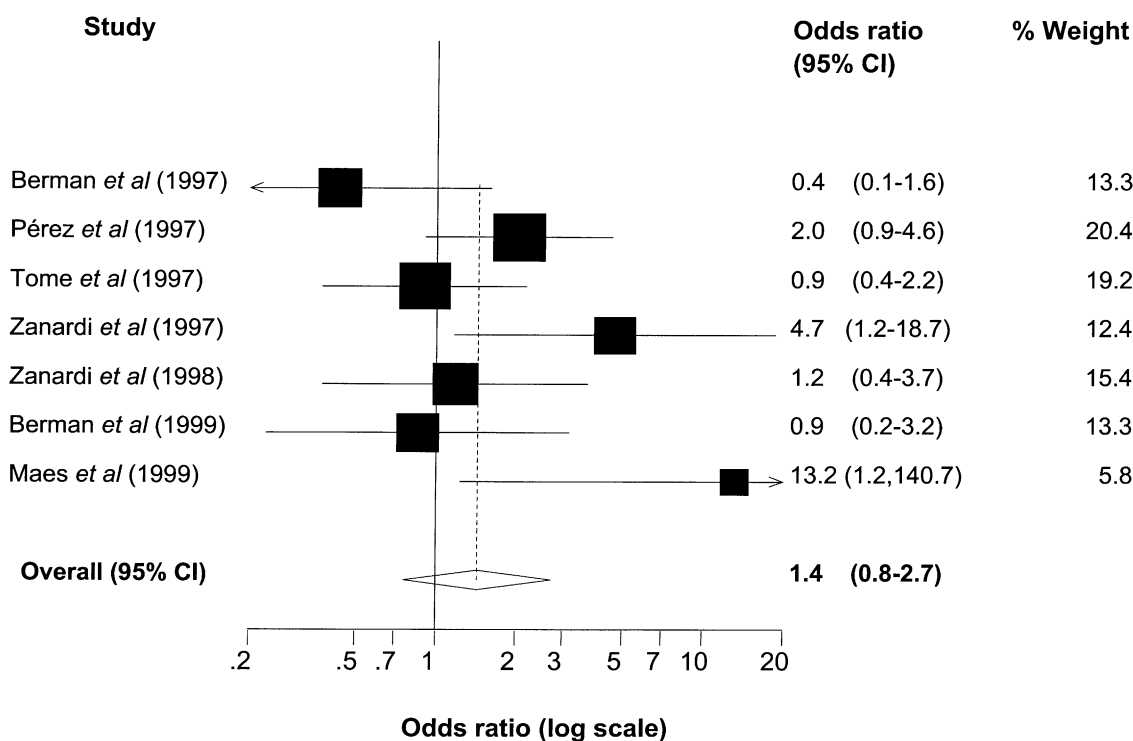


Fig. 2. Odds ratios of randomised trials and summary overall estimate for late response studies (3–6 weeks). DerSimonian and Laird random effects model. Solid squares are proportional to individual study weights.

df=7, $P=0.75$). The recorded complaints were mild; mainly nausea, vomiting, headache, somnolence, sedation and postural hypotension.

4. Discussion

Our results suggest the efficacy of pindolol plus SSRIs in depression is restricted to the first 2 weeks of treatment. We have not found evidence to support the efficacy of pindolol beyond that timing window. Also, pindolol plus SSRIs does not present a worse profile than placebo plus SSRIs regarding tolerability and adverse events. Nevertheless, we must be cautious with the evidence on adverse events so far gathered since the total number of subjects exposed to pindolol plus SSRIs is far from being large enough to discard the possibility of more severe adverse events as the induction of a serotonin syndrome (Corkeron, 1995) or other psychiatric problems (Yatham *et al.*, 1999). Pharmacoepidemiological studies with large samples

of patients are needed to substantiate the safety of pindolol plus SSRIs for regular clinical use. There is also evidence, worth pursuing in meta-analysis with individual patient data, that pindolol might not work in treatment-resistant depression (Moreno *et al.*, 1997; Pérez *et al.*, 1999). Our results seem robust as the sensitivity analysis showed. The effect estimates hardly changed when one study at a time or the two most problematic studies were deleted from the early response data set. The bias, theoretically against the null hypothesis, likely induced by the selective data from the study of Bordet *et al.* (1998) did not act as expected. Its deletion barely increased the effect estimate (OR from 2.8 to 3.2). Conversely, the deletion of Pérez *et al.* (1999) study had the expected effect of increasing the OR from 2.8 to 3.4.

Several aspects of this study—choice of effect size, heterogeneity among studies and publication bias—must be considered in depth to adequately interpret its findings. Our decision to pool the data by using ORs instead of, or jointly with, standardised mean differ-

ences (SMDs) was made by several reasons. Primarily because most studies in the area of depressive disorders use either percentage change of depression scores or a defined cut-off score as outcome criteria for improvement or remission (Prien et al., 1991; BALLENGER, 1999; Thase, 2001). In fact, these were the main effect measures reported in the studies we have analysed. These outcome criteria lead directly to dichotomous categories of outcome and hence to ORs as the favoured metric to estimate and combine effect sizes (Fleiss, 1993). Additional reasons were that ORs seem to be more easily grasped by clinicians than SMDs and that extraction of categorical results from individual studies is a sensible way for checking or carrying out intention-to-treat analysis when the quantitative results reported are not informative.

We also made the a priori decision to combine ORs by a random effects model to account not only for the statistical heterogeneity we expected but for the clinical heterogeneity formerly found and reported (Tome et al., 1997). In fact, our analysis did not show compelling evidence for heterogeneity regarding early response and in the case of late response the heterogeneity was restricted to the effect of a singular study (Maes et al., 1999). Therefore, the estimation of a pooled effect made sense even in the constrained circumstances of a fixed effects model. Nevertheless, since our results account for more variation than randomly, they tend to be more conservative, less prone to reject the null hypothesis of no effect, than those obtained by a fixed effects model.

Our results rely on conventional published evidence since former efforts to unveil the 'grey literature' were unrewarding. It means that we cannot rule out the presence of publication bias but only assess the robustness of the reported results. Since the number of retrieved studies were not large enough for a reliable assessment of funnel plots asymmetry, we relied on the Rosenthal approach to the 'file-drawer problem' to calculate the number of studies which were necessary to change the overall results (Rosenthal, 1979). We are aware of the criticisms of such approach (Oakes, 1993) but consider it at least a rough indicator of robustness against publication bias. It was calculated that 19 extra-studies, averaging null results, would be needed to change the combined probability, regarding early efficacy of pindolol plus SSRIs, to a non-significant figure. We think it is

unlikely to find so many unpublished or untracked studies since the subject we deal with continues to be a topical issue (Kinney et al., 2000; Pérez et al., 2001; Artigas, 2001). Nevertheless, these 19 extra-studies are less than the number needed to attain the tolerance level ($n=35$) which could support a robust result (Rosenthal, 1979). In summary, our results are not robust enough against publication bias in case it was present in our study. Unfortunately, publication bias seem to be more problematic and difficult to deal with in psychiatry than in other medical areas where steps have been taken to make 'grey literature' and 'data-on-file' readily available to researchers (Gilbody et al., 2000).

The timing window found in our meta-analysis agrees with experimental results about the desensitisation of somatic 5-HT_{1A} autoreceptors. It has been described (de Montigny et al., 1990) that a 2-day treatment with SSRIs decreased the firing of 5-HT neurons, but if the treatment was continued during 2 weeks these neurons progressively regained their normal firing activity. Electrophysiological studies in the raphe neurons have shown that the repeated administration of SSRIs leads to a functional desensitisation of the somatodendritic 5-HT_{1A} autoreceptors that no longer had the capability to inhibit firing and therefore enables the neurons to recover (Chaput et al., 1986). As the 5-HT reuptake inhibition by SSRIs is maintained over the weeks, the antidepressant effects will be observed only after the time 5-HT_{1A} autoreceptors require for their desensitisation, and this time seems to correlate well with the onset delay of action described for the SSRIs (Blier and de Montigny, 1994). In this way, it is sound to assume that pindolol would have its main clinical action associated to SSRIs during the approximately 2 weeks 5-HT_{1A} autoreceptors need for their desensitisation after SSRIs intake. Beyond that timing window pindolol could no longer be needed to hasten the clinical response to SSRIs and therefore the results of placebo or pindolol associated with SSRIs would be similar.

This research must be seen as a preliminary but required step to endorse the efficacy of the combination of pindolol plus SSRIs for early outcomes in depression. To settle the issue far beyond any reasonable doubt more research is needed, either by analysing weekly outcomes throughout trials length or time-to-event data. Both approaches heavily rely on

the possibility to obtain accurate data from independent trials at the level required for performing individual patient data meta-analysis.

Acknowledgements

Dr Callado holds a post-doctoral Fellowship from the Basque Government.

References

- Altman, D.G., 1998. Confidence intervals for the number needed to treat. *Br. Med. J.* 317, 1309–1312.
- Artigas, F., 2001. Limitations to enhancing the speed of onset of antidepressants. Are rapid action antidepressants possible? *Hum. Psychopharmacol. Clin. Exp.* 16, 29–36.
- Artigas, F., Celada, P., Laruelle, M., Adell, A., 2001. How does pindolol improve antidepressant action? *Trends Pharmacol. Sci.* 22, 224–228.
- Artigas, F., Pérez, V., Alvarez, E., 1994. Pindolol induces a rapid improvement of depressed patients treated with serotonin reuptake inhibitors. *Arch. Gen. Psychiatry* 51, 248–251.
- Artigas, F., Romero, L., de Montigny, C., Blier, P., 1996. Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT_{1A} antagonists. *Trends Neurosci.* 19, 378–383.
- Ballenger, J.C., 1999. Clinical guidelines for establishing remission in patients with depression and anxiety. *J. Clin. Psychiatry* 60 (Suppl. 22), 29–34.
- Bel, N., Artigas, F., 1992. Fluvoxamine preferentially increases extracellular 5-hydroxytryptamine in the raphe nuclei: an in vivo microdialysis study. *Eur. J. Pharmacol.* 229, 101–103.
- Berman, R.M., Anand, A., Cappiello, A., Miller, H.L., Hu, X.S., Oren, D.A., Charney, D.S., 1999. The use of pindolol with fluoxetine in the treatment of major depression: final results from a double-blind, placebo-controlled trial. *Biol. Psychiatry* 45, 1170–1177.
- Berman, R.M., Darnell, A.M., Miller, H.L., Anand, A., Charney, D.S., 1997. Effect of pindolol in hastening response to fluoxetine in the treatment of major depression: a double-blind, placebo-controlled trial. *Am. J. Psychiatry* 154, 37–43.
- Blier, P., Bergeron, R., 1995. Effectiveness of pindolol with selected antidepressant drugs in the treatment of major depression. *J. Clin. Psychopharmacol.* 15, 217–222.
- Blier, P., Bergeron, R., 1998. The use of pindolol to potentiate antidepressant medication. *J. Clin. Psychiatry* 59 (Suppl. 5), 16–23.
- Blier, P., Bergeron, R., de Montigny, C., 1997. Selective activation of postsynaptic 5-HT_{1A} receptors induces rapid antidepressant response. *Neuropsychopharmacology* 16, 333–338.
- Blier, P., de Montigny, C., 1994. Current advances and trends in the treatment of depression. *Trends Pharmacol. Sci.* 15, 220–226.
- Bordet, R., Thomas, P., Dupuis, B., 1998. Effect of pindolol on onset of action of paroxetine in the treatment of major depression: intermediate analysis of a double-blind, placebo-controlled trial. *Reseau de Recherche et d'Experimentation Psychopharmacologique. Am. J. Psychiatry* 155, 1346–1351.
- Bradburn, M.J., Deeks, J.J., Altman, D.G., 1998. Meta—an alternative meta-analysis command. *Stata Tech. Bull.* 44, 4–15.
- Chaput, Y., de Montigny, C., Blier, P., 1986. Effects of a selective 5-HT reuptake blocker, citalopram, on the sensitivity of 5-HT autoreceptors: electrophysiological studies in the rat brain. *Naunyn Schmiedeberg's Arch. Pharmacol.* 333, 342–348.
- Corkeron, M.A., 1995. Serotonin syndrome—a potentially fatal complication of antidepressant therapy. *Med. J. Aust.* 163, 481–482.
- de Montigny, C., Chaput, Y., Blier, P., 1990. Modification of serotonergic neuron properties by long-term treatment with serotonin reuptake blockers. *J. Clin. Psychiatry* 51 (Suppl. B), 4–8.
- DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. *Control. Clin. Trials* 7, 177–188.
- Fleiss, J.L., 1993. The statistical basis of meta-analysis. *Stat. Methods Med. Res.* 2, 121–145.
- Gilbody, S.M., Song, F., Eastwood, A.J., Sutton, A., 2000. The causes, consequences and detection of publication bias in psychiatry. *Acta Psychiatr. Scand.* 102, 241–249.
- Hamilton, M., 1960. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 23, 56–62.
- Hamilton, M., 1967. Development of a rating scale for primary depressive illness. *Br. J. Soc. Clin. Psychol.* 6, 278–296.
- Hedges, L.V., Olkin, I., 1985. *Statistical Methods for Meta-Analysis*. Academic Press, Inc, San Diego, CA.
- Invernizzi, R., Belli, S., Samanin, R., 1992. Citalopram's ability to increase the extracellular concentrations of serotonin in the dorsal raphe prevents the drug's effect in the frontal cortex. *Brain Res.* 584, 322–324.
- Kinney, G.G., Taber, M.T., Gribkoff, V.K., 2000. The augmentation hypothesis for improvement of antidepressant therapy: is pindolol a suitable candidate for testing the ability of 5HT_{1A} receptor antagonists to enhance SSRI efficacy and onset latency? *Mol. Neurobiol.* 21, 137–152.
- Lauge, N., Behnke, K., Søgaard, J., Bahr, B., Bech, P., 1998. Responsiveness of observer rating scales by analysis of number of days until improvement in patients with major depression. *Eur. Psychiatry* 13, 143–145.
- Leonard, B.E., 1984. Pharmacology of new antidepressants. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 8, 97–108.
- Maes, M., Libbrecht, I., van Hunsel, F., Campens, D., Meltzer, H.Y., 1999. Pindolol and mianserin augment the antidepressant activity of fluoxetine in hospitalized major depressed patients, including those with treatment resistance. *J. Clin. Psychopharmacol.* 19, 177–182.
- Maes, M., Vandoelaeghe, E., Desnyder, R., 1996. Efficacy of treatment with trazodone in combination with pindolol or fluoxetine in major depression. *J. Affect. Disord.* 41, 201–210.
- McAskill, R., Mir, S., Taylor, D., 1998. Pindolol augmentation of antidepressant therapy. *Br. J. Psychiatry* 173, 203–208.
- McQuay, H.J., Moore, A., 1997. Using numerical results from systematic reviews in clinical practice. *Ann. Intern. Med.* 126, 712–720.

- Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. *Br. J. Psychiatry* 134, 382–389.
- Montgomery, S.A., 1999. New developments in the treatment of depression. *J. Clin. Psychiatry* 60, 10–15.
- Moreno, F.A., Gelenberg, A.J., Bachar, K., Delgado, P.L., 1997. Pindolol augmentation of treatment-resistant depressed patients. *J. Clin. Psychiatry* 58, 437–439.
- Oakes, M., 1993. The logic and role of meta-analysis in clinical research. *Stat. Method. Med. Res.* 2, 147–160.
- Olver, J.S., Cryan, J.F., Burrows, G.D., Norman, T.R., 2000. Pindolol augmentation of antidepressants: a review and rationale. *Aust. N.Z. J. Psychiatry* 34, 71–79.
- Pérez, V., Gilaberte, I., Faries, D., Alvarez, E., Artigas, F., 1997. Randomised, double-blind, placebo-controlled trial of pindolol in combination with fluoxetine antidepressant treatment. *Lancet* 349, 1594–1597.
- Pérez, V., Puigdemont, D., Gilaberte, I., Alvarez, E., Artigas, F., 2001. Augmentation of fluoxetine antidepressant action by pindolol: analysis of clinical, pharmacokinetic, and methodologic factors. *J. Clin. Psychopharmacol.* 21, 36–45.
- Pérez, V., Soler, J., Puigdemont, D., Alvarez, E., Artigas, F., 1999. A double-blind, randomized, placebo-controlled trial of pindolol augmentation in depressive patients resistant to serotonin reuptake inhibitors. *Grup de Recerca en Trastorns Afectius. Arch. Gen. Psychiatry* 56, 375–379.
- Prien, R.F., Carpenter, L.L., Kupfer, D.J., 1991. The definition and operational criteria for treatment outcome of major depressive disorder. A review of the current research literature. *Arch. Gen. Psychiatry* 48, 796–800.
- Puzantian, T., Kawase, K., 1999. Does the addition of pindolol accelerate or enhance the response to selective serotonin reuptake inhibitor antidepressants? *Pharmacotherapy* 19, 205–212.
- Rosenthal, R., 1979. The ‘file drawer problem’ and tolerance for null results. *Psychol. Bull.* 86, 638–641.
- Smeraldi, E., Zanardi, R., Benedetti, F., Di Bella, D., Perez, J., Catalano, M., 1998. Polymorphism within the promoter of the serotonin transporter gene and antidepressant efficacy of fluvoxamine. *Mol. Psychiatry* 3, 508–511.
- Smith, T.C., Spiegelhalter, D.J., Thomas, A., 1995. Bayesian approaches to random-effects meta-analysis: a comparative study. *Stat. Med.* 14, 2685–2699.
- Spiegelhalter, D.J., Thomas, A., Best, N.G., 2000. WinBUGS Version 1.2. User Manual. Cambridge, MRC Biostatistics Unit.
- StataCorp, 2001. Stata Statistical Software: Release 7.0. Stata Corporation, College Station, TX.
- Sugrue, M.F., 1983. Chronic antidepressant therapy and associated changes in central monoaminergic receptor functioning. *Pharmacol. Ther.* 21, 1–33.
- Tome, M.B., Isaac, M.T., Harte, R., Holland, C., 1997. Paroxetine and pindolol: a randomized trial of serotonergic autoreceptor blockade in the reduction of antidepressant latency. *Int. Clin. Psychopharmacol.* 12, 81–89.
- Thase, M.E., 2001. Methodology to measure onset of action. *J. Clin. Psychiatry* 62 (Suppl. 15), 18–21.
- Yatham, L.N., Lint, D., Lam, R.W., Zis, A.P., 1999. Adverse effects of pindolol augmentation in patients with bipolar depression. *J. Clin. Psychopharmacol.* 19, 383–384.
- Zanardi, R., Artigas, F., Franchini, L., Sforzini, L., Gasperini, M., Smeraldi, E., Perez, J., 1997. How long should pindolol be associated with paroxetine to improve the antidepressant response? *J. Clin. Psychopharmacol.* 17, 446–450.
- Zanardi, R., Franchini, L., Gasperini, M., Lucca, A., Smeraldi, E., Perez, J., 1998. Faster onset of action of fluvoxamine in combination with pindolol in the treatment of delusional depression: a controlled study. *J. Clin. Psychopharmacol.* 18, 441–446.