Departament de Psiquiatria i Psicobiologia Clínica
Facultat de Medicina
Universitat de Barcelona

COST-EFECTIVITAT DEL TRACTAMENT ANTIDEPRESSIU

Avaluació del cost-efectivitat i cost-utilitat dels antidepressius en el tractament de persones amb trastorns depressius en Atenció Primària de Salut a Catalunya

Tesi doctoral presentada per Antoni Serrano Blanco
per optar al Grau de Doctor en Medicina

Castelldefels, juliol de 2006

ELSEVIER

Journal of Affective Disorders 91 (2006) 153–163

JOURNAL OF AFFECTIVE DISORDERS

www.elsevier.com/locate/jad

Research report

Effectiveness and cost-effectiveness of antidepressant treatment in primary health care: A six-month randomised study comparing fluoxetine to imipramine☆

A. Serrano-Blanco [a,*], E. Gabarron [a], I. Garcia-Bayo [b], M. Soler-Vila [b], E. Caramés [b], M.T. Peñarrubia-Maria [b], A. Pinto-Meza [a], J.M. Haro [a]
Depressió en Atenció Primària de Gavà group (D. A. P. G. A.)[1]

[a] *Sant Joan de Déu, Serveis de Salut Mental, Fundació Sant Joan de Déu, St. Boi de Llobregat, Barcelona, Spain*
[b] *CAP Gavà II, Unitat Docent Costa de Ponent, Institut Català de Salut, Barcelona, Spain*

Received 28 April 2005; received in revised form 25 November 2005; accepted 25 November 2005
Available online 3 February 2006

Abstract

Background: Over the past decade, studies of the effectiveness of pharmacological treatment for depression have often been based on research designs intended to measure efficacy, and for this reason the results are of limited generalizability. Research is needed comparing the clinical and economic outcomes of antidepressants in day-to-day clinical practice.
Methods: A six-month randomised prospective naturalistic study comparing fluoxetine to imipramine carried out in three primary care health centres. Outcome measures were the Montgomery Asberg Depression Rating Scale (MADRS), direct costs, indirect costs and total costs. Subjects were evaluated at the beginning of treatment and at one, three and six months thereafter.
Results: Of the 103 patients, 38.8% ($n=40$) were diagnosed with major depressive disorder, 14.6% ($n=15$) with dysthymic disorder, and 46.6% ($n=48$) with depressive disorder not otherwise specified. Patients with major depressive disorder or dysthymic disorder achieved similar clinical improvement in both treatment groups (mean MADRS ratings decrease in major depressive disorder from baseline to 6 months of 18.3 for imipramine and 18.8 for fluoxetine). For patients with major depressive disorder and dysthymic disorder, the imipramine group had fewer treatment-associated costs (imipramine € 469.66 versus fluoxetine € 1585.93 in major depressive disorder, $p<0.05$; imipramine € 175.39 versus fluoxetine € 2929.36 in dysthymic disorder, $p<0.05$). The group with depressive disorder *not otherwise specified* did not experience statistically significant differences in clinical and costs outcomes between treatment groups.
Limitations: Exclusion criteria, participating physicians may not represent GPs.

☆ This research was supported by the Catalan Agency for Health Technology Assessment and Research (AATRM 15/47/98) and forms part of the IRYSS Network (Investigación en Resultados de Salud y Servicios Sanitarios) (FIS GG03/202).

* Corresponding author. Sant Joan de Déu–SSM, Fundació Sant Joan de Déu, Antoni Pujadas 42, 08830 Sant Boi de Llobregat (Barcelona), Spain. Tel.: +34 93 640 63 50; fax: +34 93 471 44 51.
E-mail address: aserrano@sjd-ssm.com (A. Serrano-Blanco).

[1] The DAPGA group is a multidisciplinary group of researchers studying depression in primary care. Members of the DAPGA group are: Emilia Caramés Duran (MD), Ángel Espín Martínez (MD), Inmaculada García Bayo (MD), M. Carmen García Pulido (MD), Irma Miranda Chueca (MD), Maria Teresa Peñarrubia María (MD), Dolores Ruiz Blanes (MD), Ignacio Sáenz (MD), Maria Soler Vila (MD), Rosa Villafáfila (MD), Pere Simonet (MD), Camino Verduras Rubio (MD), Rosa Aranzana (MD), Amelia Prieto (MD), José Luis Caballé (MD), Fernando Teba Escabías (MD), Susana Araya LaRibera (MD), Elia Gabarron Hortal, Antoni Serrano-Blanco (MD) and Josep Maria Haro Abad (PhD).

0165-0327/$ - see front matter
doi:10.1016/j.jad.2005.11.014

Conclusions: In a primary care context, imipramine may represent a more cost-effective treatment option than fluoxetine for treating major depressive disorder or dysthymic disorder. There were no differences in cost-effectiveness in the treatment of depressive disorder *not otherwise specified*.

Keywords: Depression; Cost-effectiveness; Antidepressants; Primary care; Major depressive disorder

1. Introduction

Clinicians have a wide range of pharmacological treatment alternatives for depression, which have been usually found to be of similar efficacy (Song et al., 1993; Nielsen et al., 1993; MacGillivray et al., 2003). The choice of antidepressant drug for a given individual can be made according to a variety of criteria. In clinical practice, this choice is usually based on clinical outcomes, which may vary depending on clinical subtypes, or on the side effect profile. However, current understanding of treatment outcomes is mostly based on clinical trials that measure efficacy, and are usually conducted under experimental conditions bearing little resemblance to day-to-day clinical practice. These studies, therefore, are of limited use in clinical decision-making (Donoghue and Hylan, 2001). According to Wells (1999), Zimmerman et al. (2002) and Concato et al. (2000), clinical trials are based on samples selected by stringent criteria and measure highly specific results in a time frame usually more limited than naturalistic follow-up studies. Also, clinicians are blind to the treatment the patient receives, and the experimental protocol for therapy may not be representative of clinical practice. As a result, the information obtained cannot easily be generalized. Clinical decisions should be based on the effectiveness of treatments, which has been defined as the degree to which a particular intervention, when used habitually in clinical conditions, achieves the result it is said to achieve.

In the last decade, the increase in the number of articles that highlight the relevance of burden and cost-evaluation of mental disorders has led to the development of methods for comparing the cost-effectiveness of treatments (Badia, 2003; Greenberg et al., 1993; Rice and Miller, 1995; Stewart, 1997; Murray and López, 1996). Published studies of the efficacy and effectiveness of antidepressant treatments focus on comparisons between selective serotonin reuptake inhibitors (SSRIs) and tricyclic antidepressants (TCAs), and addressed primarily whether the former were better because despite their higher price, they produced fewer side effects (Le Pen et al., 1994; Mulrow et al., 2000). A meta-analysis by Anderson and Tomenson (1995), which included 62 double-blind randomised trials, concluded that TCAs were slightly more cost-effective per patient treated than SSRIs, but the latter were more cost-effective as measured in terms of "successful treatment". However, these results are of limited generalizability because they are based on randomised controlled trials of limited external validity (Crown, 2000).

Various authors (Crown, 2001; Anderson and Tomenson, 1995; Garattini et al., 1998; Mulrow et al., 2000) have proposed randomised prospective studies with naturalistic follow-up and direct measures of economic variables as the best method of investigating pharmacological treatment outcomes in depression. In a study based on these principles conducted in a primary care context, Simon et al. (1996) found that direct costs for patients receiving fluoxetine (FLU) were no different from those for patients receiving imipramine (IMI) or desipramine. It is not known yet whether SSRIs are more cost-effective than TCAs.

This randomised naturalistic study compared the effectiveness and cost-effectiveness of IMI, a TCA, relative to FLU, a SSRI, for persons suffering from depressive disorders treated in primary care. The results of the project provide information not only on clinical effectiveness and direct costs, but also on indirect costs, which have been studied very little thus far.

2. Methods

The study was conducted in three Primary Health Care Centres (PHCC) within the metropolitan area of Barcelona, Spain. The PHCC are located in two municipalities, Gavà and Viladecans, and are situated around 10 km from the centre of Barcelona. Each centre serves a population of approximately 25,000. A staff of 7 to 10 general practitioners (GP) and 7 to 10 nurses work in each centre. Each centre treats 20,000 people per year, forming a mean of 90,000 yearly office visits, each lasting approximately 5 min per visit, although each GP is allowed two extended visits per day. A separate service provider, a Mental Health Care Centre (MHCC) located in Gavà provides specialized treatment, mostly for people with treatment-resistant anxiety and depressive disorders, and severe mental disorders.

A referral from a GP or a psychiatrist is required for uptake in the MHCC. The MHCC supports 5 PHCC serving a mostly urban population of approximately 125,000 inhabitants, and treats patients over 18 years of age.

In Spain, the National Healthcare Service (NHS) is financed by the general taxes levied by the state, and in the municipalities where the study was carried out, administered by the autonomous government of Catalonia. Office visits and hospital admissions are completely covered by the NHS, and prescription medications are covered completely for retired persons, and partially for those still employed. Sick leave requires a physician's authorization, and patients unable to work continue to receive most of their salary.

2.1. Patients

Patients eligible for participation were those aged 18 to 65 years who were seen by their GP between May 1999 and May 2001, and were evaluated by the GP to receive antidepressant treatment for a depressive disorder. Inclusion criteria were defined in order to select a population of patients similar to those for whom GPs usually prescribe antidepressant medication for the treatment of a depressive disorder. Exclusion criteria were: pharmacological antidepressant treatment in the previous 60 days; a history of alcohol or drug abuse; psychotic symptoms or a history of bipolar disorder; antipsychotic drug use, lithium or antiepileptic medication in the previous six months; pregnancy, lactation, or women intending to become pregnant; a general medical condition that counterindicated the use of the antidepressants used in the study (a history of epilepsy, renal failure, heart failure, a history of myocardial infarction, severe liver or prostatic disorders, mental disorder due to a general medical condition, glaucoma, allergy to imipramine, fluoxetine or lactose); need of hospital admission for depression in the judgment of the GP; and requirement of intensive psychotherapy or electroconvulsive therapy.

Patients requiring antidepressant treatment for a depressive disorder were invited to participate, provided they did not fill any of the exclusion criteria. The enrolment visit was within the normal course of care. At the end of the enrolment visit, the GP informed the patient about the study. If the patient agreed to participate, the GP filled out a short recruitment form with basic information about the patient and the initial dosage of medication (for both imipramine and fluoxetine) that he or she would prescribe. The first evaluation visit with the blind interviewer (a senior psychiatric resident) was scheduled as soon as possible (always within the first week after the enrolment visit).

2.2. First evaluation and randomisation

During the first evaluation, the blind interviewer administered the Structured Clinical Interview for DSM-IV (SCID-I) (First et al., 2001), the Montgomery Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979; Lobo et al., 2002), the Clinical Global Impression (CGI) rating scale and a questionnaire with basic sociodemographic, clinical and health service use data. Following SCID-I assessment, only patients fulfilling a depressive disorder diagnosis (major depressive disorder [MDD], dysthymic disorder [DD], or depressive disorder not otherwise specified [DDNOS]) were included in the study.

Before the study began, the investigators created the random group sequence with a random number sheet. Randomisation was stratified by PHCC in 10-patient blocks (5 to fluoxetine and 5 to imipramine) derived from the group sequence. With the list, the investigators created a series of numbered envelopes that contained the group assignment for each patient.

All patients had a recruitment number. At the end of the visit the evaluator opened the envelope with the patient number that contained the medication that had been randomly selected. This medication was then prescribed to the patient at the dosage that had been recommended by the GP. The GP did not participate in the random allocation of patients.

2.3. Patient care

The GP was responsible for all patient care following his or her usual clinical practice. There were no recommendations in the protocol regarding medication dosage, use of concomitant medications, and duration of treatment or referral to a MHCC. GPs were only instructed to note in the patient's chart any change in antidepressant medication.

2.4. Follow-up evaluations

At one, three and six-month intervals after the initial evaluation (one week before or after the evaluation date was allowed), the patients were evaluated by a clinical psychologist blind to their treatment status who administered the assessment instruments (MADRS and a questionnaire recording the patient's use of health care services). To increase reliability between the assessment of the psychiatrist and that of the clinical

psychologist, both interviewers received identical training in the use of the instruments. After the patients completed the study, their clinical charts were reviewed and data concerning medication use was recorded.

Sample size was calculated for a type I error of 0.05 and a power of 0.80. In order to detect a statistically significant difference of € 800 in costs between the groups, the required sample size was 57 patients per group (total 114). This sample size calculation took into account that both treatments were equally effective, and that costs for depression in primary care in Spain were about € 1000 (SD 1512) after 6 months (Sacristan et al., 1999).

The study protocol was approved by the ethics committee of Sant Joan de Déu Mental Health Services. All participating patients gave their written consent.

2.5. Statistical analysis

The primary outcome measure for effectiveness was measured by the MADRS. The primary analysis was conducted using the entire sample. Secondary analysis was based on a stratification of the data based on the SCID-I diagnosis (MDD, DD, DDNOS). A linear regression model for each of the outcome measures including baseline rating of the outcome scale and treatment was used to calculate the statistical significances of the difference between treatment groups.

Societal cost perspective was used for the calculation of costs. The economic variables used in this study were total, direct and indirect costs. Total costs were obtained by adding direct and indirect costs. Direct costs were calculated by adding the costs of psychotropic drug treatment (antidepressants and benzodiazepines), office visits, nursing care and social work, the cost of visits to specialist practitioners, psychiatric and general medicine emergency-room care, and psychiatric and general medicine hospital admissions. The cost of the drugs was calculated by determining the price per milligram of the drug most frequently prescribed during the study according to the International Vademecum (Red Book) of 2001, and included value-added tax (VAT) (Annex 1). The total cost of drug treatment was calculated by multiplying the price per milligram by the daily dose in milligrams and the number of treatment days. For other antidepressants or anxiety suppressants, the price of the drug brand and presentation prescribed was used. The costs of patients' use of healthcare services were obtained from the Base de Datos de Costes Unitarios de Soikos (2001). The indi-

Annex 1
Drug acquisition costs in euros, 2001 and 2003

Drug	Brand name	Dose (mg)/tablet	Tablets	2001 price ()	2003 price ()
Antidepressants					
Amitriptiline	Tryptizol	50	30	2.88	2.88
Citalopram	Prisdal	20	28	30.07	30.08
Clomipramine	Anafranil	25	40	3.62	3.62
Fluoxetine	Prozac	20	28	25.83	23.25
Fluvoxamine	Dumirox	50	30	10.58	9.51
Imipramine	Tofranil	25	50	2.16	2.15
Maprotiline	Ludiomil	10	30	2.24	2.23
Mianserine	Lantanon	10	50	4.36	4.36
Mianserine	Lantanon	30	30	7.65	7.65
Mirtazapine	Rexer	15	30	27.49	27.49
Moclobemide	Manerix	150	30	14.56	14.55
Nefazodone	Dutonin	50	14	4.66	
Paroxetine	Seroxat	20	28	33.19	33.19
Reboxetine	Norebox	4	20	13.95	13.95
Sertraline	Besitran	50	30	31.99	31.99
Anxiolytic/sleep inducers					
Alprazolam	Trankimazin	0.25	30	2.68	2.40
Bromazepam	Lexatin	1.5	30	1.59	1.60
Clorazepato	Tranxilium	5	30	1.89	1.88
Clotiazepam	Distensan	5	30	2.64	2.63
Diazepam	Valium	5	30	1.53	1.53
Halazepam	Alapryl	40	30	3.61	3.61
Lorazepam	Orfidal	1	25	2.01	2.01
Lormetazepam	Loramet	1	30	3.17	2.92
Zolpidem	Stilnox	10	30	6.05	6.05
Zopiclone	Limovan	7.5	30	5.23	5.23

rect costs were calculated by multiplying authorized sick leave days by the minimum daily wage in Spain (€ 24.04) for 2001.

The SPSS 10.0 was used for data analyses. It was compared to each intervention group based on original treatment assignment, regardless of treatment received (intent-to-treat, hereafter ITT). Comparisons of clinical outcomes by treatment group were analysed using multivariate linear regression, including the baseline rating of the rating scale as a covariate. The dependent variable of the models was the outcome of the rating scale. For patients who dropped out of the study but had a post-baseline measure, the Last Observation Carried Forward (LOCF) imputation method was used.

Multivariate economic data analysis was performed using a logarithmic transformation of costs to reduce the skewness of the distribution (Manning and Mullahy, 2001). Treatment and total costs in the two months before baseline were included as covariates. The transformed variable fulfilled the homocesdasticity requirement. Means and confidence intervals presented in the tables were based on untransformed data.

Post hoc hypotheses were tested as to possible differences between those with major depression, those with dysthymic disorder and those with depressive disorders not otherwise specified.

Finally, given the decrease in cost for fluoxetine after it became available as a generic drug, drug acquisition costs during the study period may over-estimate current costs. The sensitivity analysis was conducted using current drug acquisition costs.

3. Results

3.1. Patient characteristics

Out of the 110 patients initially selected by the GPs, 103 were included. Two patients declined to participate in the evaluation and five could not be assessed in the week after the initial GP visit. Patients were randomised either to FLU (53 patients) or to IMI (50 patients) treatment. Table 1 describes the sociodemographic and clinical characteristics of the patients included. There were no significant differences in gender, age, professional occupation, diagnosis or clinical severity between the two treatment groups.

Of the 103 patients, 94 (91.3%) completed at least one of the follow-up visits and were included in the statistical analyses (Fig. 1). Of the 9 patients lost to follow-up, 4 were from the FLU group and 5 from the IMI group. Patients lost to follow-up were younger, with a mean age of 33.8 (SD=15.5) versus 44.4 (SD=11.3) ($p<0.05$), and more severely ill, with a mean MADRS score of 30.2 (SD=7.8) versus 21.7 (SD=9.1), p value<0.05. After the study ended, the investigators reviewed the patients' clinical charts and noted that none of the patients lost to follow-up utilized mental health care services nor were hospitalized. The mean starting dosage for FLU was 19.4 mg (SD=2.3, median 20, minimum 10, maximum 20), and for IMI it was 33.5 mg (SD=16.8, median 25, minimum 25, maximum 100). The mean daily dosage for FLU at day 30 was 20.7 mg (SD=4.8, median 20, minimum 10, maximum 40) and for IMI 58.3 mg (SD=30.3, median 50, minimum 25, maximum 125). There was no titration in the FLU group. Patients treated with IMI increased their mean dosage from 33.5 to 58.3 mg/day during the first month of treatment (Table 2). Doses remained relatively stable after that.

Table 1
Baseline sociodemographic and clinical characteristics of the patients included in the study

	Fluoxetine	Imipramine
Number of patients	53	50
	N (%)	
Women	40 (75.5)	35 (70)
Level of professional occupation		
High[a]	8 (15.1)	4 (8)
Medium[b]	22 (41.5)	18 (36)
Low[c]	23 (43.4)	28 (56)
Diagnosis		
Major depressive disorder	20 (37.7)	20 (40.0)
Dysthymic disorder	8 (15.1)	7 (14.0)
Depressive disorder not otherwise specified	25 (47.2)	23 (46.0)
	Mean (SD)	
Age (years)	43.1 (12.2)	43.9 (12.0)
MADRS	23.15 (9.7)	21.66 (8.9)
CGI	3.50 (1.09)	3.54 (1.15)

MADRS: Montgomery Asberg Depression Rating Scale; CGI: Clinical Global Impression scale.
[a] High: executive, business person or high ranking professions.
[b] Medium: secretary, civil workers or manual labour supervisor.
[c] Low: blue collar job, i.e. manual labourer or a job with minimal qualification.

3.2. Effectiveness

Patients in both treatment groups improved during the follow-up and there were no differences in effectiveness between them. However, the analysis of treatment response revealed differences in response among diagnostic groups (Table 3). Patients with MDD improved more on IMI than on FLU at 30 days as measured by the MADRS ($p<0.05$) (Table 3). The adjusted mean difference in clinical response between the FLU and IMI groups at 30 days using MADRS was

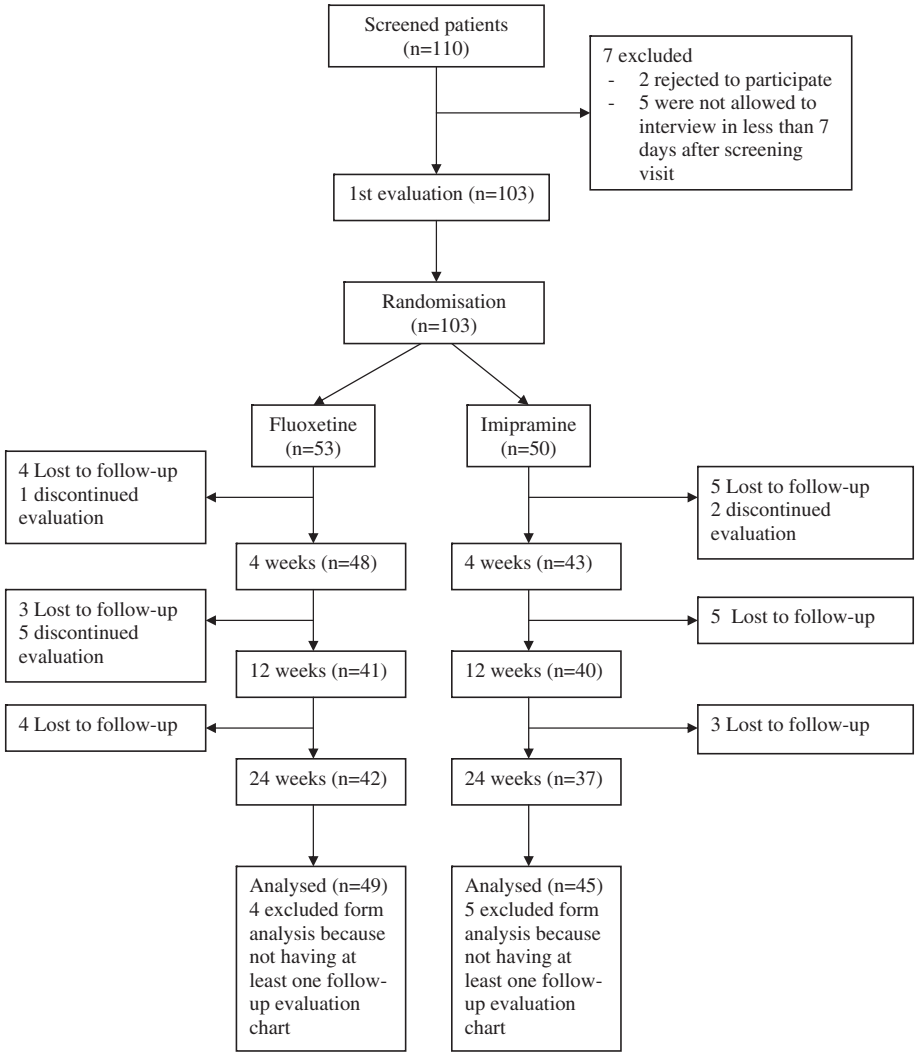

Fig. 1. Flow-chart of patient distribution.

−6.50 (95% CI −12.40, −0.65). There were no statistically significant differences in outcome in later evaluations. For patients with DD, IMI achieved better outcomes than FLU at three months follow-up ($p<0.05$). The adjusted mean difference in clinical response between the FLU and IMI groups at 3 months

Table 2
Pharmacological treatment patterns of patients included in the study

	Baseline	30 days	90 days	180 days
Fluoxetine				
Mean daily dose in milligrams (SD)[a]	19.4 (2.3)	20.7 (4.8)	20.7 (3.7)	20.9 (4.2)
Patients taking initial prescription (%)	53	39 (73.6)	30 (56.6)	23 (43.4)
Patients taking other antidepressant treatments (%)		4 (7.5)	8 (15.1)	6 (11.3)
Patients without pharmacological treatment (%)		7 (13.2)	10 (20.8)	19 (35.8)
Imipramine				
Mean daily dose in milligrams (SD)[a]	33.5 (16.8)	56.6 (29.7)	64.2 (35.8)	52.8 (34.2)
Patients taking initial prescription (%)	50	34 (64)	30 (60)	18 (36)
Patients taking other antidepressant treatments (%)		4 (8)	7 (14)	8 (16)
Patients without pharmacological treatment (%)		9 (18)	7 (14)	18 (36)

[a] Pharmacological doses calculated in patients maintaining initial treatment at current and previous evaluation points.

Table 3
MADRS and CGI mean (SD) scores at baseline and at 30, 90 and 180 days by diagnostic and treatment groups

		MADRS				CGI			
		Baseline	Day 30	Day 90	Day 180	Baseline	Day 30	Day 90	Day 180
Major depressive disorder ($n=37$)	Fluoxetine ($n=18$)	29.7 (9.3)	22.7* (11.4)	17.1 (9.9)	10.9 (9.2)	4.3 (1.1)	3.7 (1.6)	3.0 (1.4)	2.3 (1.5)
	Imipramine ($n=19$)	25.3 (8.7)	14.1* (8.0)	11.3 (7.9)	7.0 (6.4)	4.3 (1.1)	2.9 (1.2)	2.5 (1.4)	1.6 (1.3)
Dysthymic disorder ($n=14$)	Fluoxetine ($n=7$)	18.6 (5.3)	18.3 (8.6)	18.2* (7.7)	13.4 (7.9)	3.7 (0.5)	3.6* (0.8)	3.6 (1.1)	3.1 (0.7)
	Imipramine ($n=7$)	17.3 (12.2)	9.9 (9.0)	8.4* (8.0)	7.3 (6.9)	3.0 (0.9)	2.1* (1.4)	1.7 (1.5)	2.0 (1.8)
Depressive disorder not otherwise specified ($n=43$)	Fluoxetine ($n=24$)	17.7 (5.8)	10.4* (6.5)	8.0 (8.5)	7.2 (7.1)	2.9 (0.7)	2.3 (1.0)	1.6 (1.4)	1.4 (1.4)
	Imipramine ($n=19$)	18.2 (7.0)	14.7* (7.5)	11.3 (8.6)	11.2 (9.9)	2.8 (0.8)	2.7 (1.5)	2.3 (1.5)	2.0 (1.7)
All patients ($n=94$)	Fluoxetine ($n=49$)	22.2 (9.1)	16.1 (10.4)	12.8 (10.0)	9.5 (8.2)	3.5 (1.1)	3.0 (1.4)	2.4 (1.6)	2.0 (1.5)
	Imipramine ($n=45$)	21.7 (9.2)	13.7 (7.9)	10.9 (8.1)	8.8 (8.2)	3.5 (1.2)	2.7 (1.4)	2.3 (1.4)	1.8 (1.5)

* $p<0.05$, linear regression using LOCF and baseline rating and treatment as covariates.

using MADRS was −10.06 (95% CI −18.51, −1.62). Patients with DDNOS showed more improvement on FLU, as measured by MADRS during the first month of treatment ($p<0.05$). The adjusted mean difference in clinical response between the FLU and IMI groups at one month using MADRS was 4.12 (95% CI 0.15, 8.10).

3.3. Treatment costs

Table 4 shows direct, indirect and total cumulative costs of the patients by diagnostic and treatment group. A comparison of the economic data for all patients in the study showed significant differences in total costs between treatment groups after 90 and 180 days of follow-up. The FLU group showed higher total costs after 90 days of follow-up (€ 784.26; SD=869.91) when compared to the IMI group [€ 489.11; SD=693.54; $p<0.05$]. The FLU group showed higher total costs after 180 days of follow-up (€ 1,330.38; SD=1499.07) when compared to the IMI group [€ 691.81; SD=1071.62; $p<0.001$]. A separate comparison of direct and indirect costs showed that the FLU group had higher indirect costs after 180 days of follow-up (€ 999.21; SD=1409.91) when compared to the IMI group [€ 431.44; SD=982.02; $p<0.05$]. No statistically significant differences were observed between FLU and IMI direct costs.

Comparisons among diagnostic groups revealed that in MDD, the FLU group showed significantly higher cumulative total costs after 30 ($p<0.01$), 90 ($p<0.05$) and 180 days ($p<0.001$) days. Separate comparison of direct and indirect costs showed that the FLU group had higher direct costs after 30, 90 and 180 days of follow-up ($p<0.05$ in each measure). No statistically significant differences were observed between FLU and IMI indirect costs. Patients with DD in the FLU group showed significantly higher total costs after 30 ($p<0.05$), 90 ($p<0.01$) and 180 ($p<0.05$) days than patients in the IMI group. Separate comparison of direct and indirect costs showed that the FLU group had higher direct costs after 90 days of follow-up ($p<0.05$), and higher indirect costs after 90 and 180 days of follow-up ($p<0.05$ and $p<0.01$, respectively). Patients with DDNOS did not show differences in total costs between the treatment groups.

3.4. Sensitivity analysis

In the past few years fluoxetine has become one of the most economical antidepressants. A sensitivity analysis of price was conducted using drug acquisition costs according to the International Vademecum (Red Book) of 2003, with value-added tax (VAT) included. This analysis showed that the statistical significance of total costs remains the same, though that of some direct costs has changed. Thus, for all of the patients, there were still statistically significant differences in total costs at 90 and 180 days ($p<0.05$ and $p<0.001$, respectively). In patients with a diagnosis of MDD, statistically significant differences between treatment groups remained at 30, 90 and 180 days, as they also did in patients diagnosed with DD. The patients diagnosed with DDNOS continued as before, without statistically significant differences between treatment groups.

4. Discussion

This randomised study has shown that, during the first six months of treatment in primary care for a depressive disorder, IMI may be a more cost-effective treatment than FLU since the similar effectiveness was balanced by lower cost. Disaggregating the results by disorder, in major depressive disorder and dysthymic disorder, patients starting IMI had lower costs than patients initiating FLU and at some assessment points had also better outcomes in clinical severity. In contrast, for depressive disorders that do not fulfil major depres-

Table 4
Mean (SD) cumulative direct, indirect and total costs by treatment and diagnostic groups in euros

	Previous			30 days			90 days			180 days		
	DC	IC	TC	DC	IC	TC	DC	IC	TC	DC	IC	TC
Fluoxetine												
MDD ($n=18$)	231.6	320.5	552.1	120.0[a]	360.6	480.6[c]	241.14[a]	761.28	1002.42[a]	395.36[a]	1190.58	1585.93[c]
	(279.1)	(368.8)	(561.6)	(97.2)	(371.1)	(414.1)	(126.24)	(941.07)	(1010.69)	(157.44)	(1411.01)	(1497.95)
DD ($n=7$)	233.3	412.1	645.4	105.54	412.12	517.7[a]	252.24[a]	1281.60[b]	1533.84[b]	411.39	2517.97[b]	2929.36[a]
	(116.0)	(385.5)	(467.6)	(50.39)	(385.51)	(354.4)	(79.18)	(1105.15)	(1150.10)	(134.15)	(2259.08)	(2350.86)
DDNOS ($n=24$)	88.4	210.4	298.8	62.15	161.52	223.67	148.15	253.87	402.02	259.64	412.71	672.35
	(92.0)	(334.9)	(368.7)	(45.51)	(298.38)	(299.85)	(85.31)	(315.23)	(344.24)	(128.41)	(505.97)	(555.52)
Total ($n=49$)	161.7	279.7	441.4	89.59	270.46	360.04	197.18	587.09	784.26[a]	331.18	999.21[a]	1330.38[c]
	(196.5)	(355.1)	(474.3)	(73.61)	(349.16)	(371.34)	(110.83)	(806.95)	(869.91)	(154.82)	(1409.91)	(1499.07)
Imipramine												
MDD ($n=19$)	141.9	113.9	255.8	74.7[a]	52.1	126.8[c]	159.97[a]	222.90	382.87[a]	228.30[a]	241.36	469.66[c]
	(159.3)	(270.2)	(343.8)	(113.5)	(167.4)	(194.2)	(144.48)	(461.78)	(525.90)	(175.76)	(461.42)	(551.99)
DD ($n=7$)	89.8	103.0	192.8	49.23	0.0	49.2[a]	87.07[a]	0.0[b]	87.07[b]	175.39	0.0[b]	175.39[a]
	(109.7)	(272.6)	(378.1)	(43.65)		(43.7)	(59.41)		(59.41)	(90.42)		(90.42)
DDNOS ($n=19$)	118.3	265.7	384.0	81.44	265.71	347.15	221.55	521.93	743.48	323.77	780.46	1104.23
	(83.9)	(357.4)	(366.8)	(71.51)	(357.43)	(368.76)	(219.51)	(801.61)	(868.89)	(270.90)	(1380.33)	(1471.17)
Total ($n=45$)	123.8	176.3	300.1	73.59	134.18	207.77	174.63	314.48	489.11[a]	260.38	431.44[a]	691.81[c]
	(123.5)	(313.5)	(358.6)	(87.97)	(277.45)	(294.16)	(176.06)	(622.94)	(693.54)	(217.03)	(982.02)	(1071.62)
Total												
MDD (37)	185.5	214.4	399.9	96.7	202.2	298.9	199.45	484.81	684.27	309.57	703.14	1012.71
	(227.0)	(334.2)	(480.2)	(106.9)	(321.7)	(363.3)	(140.19)	(774.11)	(848.06)	(185.25)	(1130.47)	(1237.69)
DD ($n=14$)	161.5	257.6	419.1	77.4	206.1	283.4	169.66	640.80	810.46	293.39	1258.98	1552.37
	(131.6)	(358.6)	(471.3)	(53.9)	(338.1)	(343.4)	(108.94)	(1002.95)	(1084.28)	(164.55)	(2015.54)	(2143.93)
DDNOS ($n=43$)	101.6	234.8	336.4	70.1	207.6	278.2	180.58	372.32	552.90	287.98	575.21	863.18
	(88.8)	(342.0)	(366.0)	(58.5)	(326.0)	(333.7)	(161.24)	(589.88)	(646.44)	(203.76)	(995.44)	(1069.42)
Total ($n=94$)	143.6	230.2	373.7	81.9	205.2	287.1	186.38	456.59	642.97	297.28	727.40	1024.68
	(165.8)	(338.0)	(426.6)	(80.8)	(322.5)	(343.5)	(145.38)	(733.78)	(800.08)	(189.53)	(1250.42)	(1343.89)

MDD: major depressive disorder; DD: dysthymic disorder; DDNOS: depressive disorder not otherwise specified; DC: direct costs; IC: indirect costs; TC: total costs.
[a] $p<0.05$, [b] $p<0.01$, [c] $p<0.005$. Statistical significance was calculated using linear regression models with log-transformed costs as outcome variable.

sive or dysthymic disorder criteria we found no differences in costs between both antidepressants, and better outcomes for FLU at one of the assessment points.

Most studies comparing the cost-effectiveness of TCA with SSRI in primary care have been meta-analyses or mathematical models (Jönsson and Bebbington, 1994; Doyle et al., 2001; Revicki et al., 1995). Mathematical models are very sensitive to the assumptions of the model, which has led sometimes to discordant conclusions. For example, after the Jönsson and Bebbington (1994) study that recommended the use of paroxetine in preference to imipramine, Woods and Rizzo (1997) published a new analysis stating that if the follow-up period were increased and the proportion of patients who changed drugs when the initial treatment proved ineffective were different, imipramine would be more cost-effective than paroxetine. Published reviews also show contradictory results. While Hotopf et al. (1995) suggested that imipramine should be the first-line treatment for depression, Crott and Gilis (1998) argued that it should be fluoxetine. Contradictions may have emerged because these reviews included studies based on mathematical models, which have lower external validity (Hotopf et al., 1995; Crott and Gilis, 1998; Barbui et al., 2003).

This is a randomised prospective study with naturalistic follow-up and direct measures of economic variables. The methodology recommended by Crown (2000) and Simon et al. (1995) was used for the evaluation of the cost-effectiveness of antidepressants. Also, as suggested by Mendlewicz (2001), the direct total costs were evaluated in addition to medication costs, given that the price of the drug, normally a SSRI versus a TCA, tends to be compensated for by fewer office visits. It is also necessary to analyse the indirect costs (lost work time) associated with depression (Berto et al., 2000; Panzarino, 1998; Stewart et al., 2003).

A previous six-month randomised naturalistic follow-up study comparing FLU with two different TCAs found that direct costs for the FLU group were lower than those in either TCA group (imipramine and desipramine), but not significantly so (Simon et al., 1996). The design of the present study is very similar to that study, but the patients in the Simon et al. study were less severely ill, and indirect costs were not measured.

In order to maximize external validity, in our study the inclusion criteria were purposely broad in order to enroll a representative sample of patients who started antidepressant treatment in primary care. There were no restrictions in treatment patterns after the initial randomisation. Participating primary care physicians used their clinical judgment in treating patients, changing drug prescriptions, authorizing sick leave, making referrals to specialized mental health services, or admitting the patient. They were given no treatment recommendations or restrictions, and were asked only to opt first for the other antidepressant used in the study if they considered changing the patient's medication.

One of the interesting findings of the study is that the mean dosage of IMI at day 30 is lower than the dosage recommended in clinical guidelines (1993 APA treatment guidelines for MDD). Some authors think that the high dosage level established for randomised controlled trials may not be necessary in primary care (Simon et al., 1996; Blakey, 1999). A systematic review found that TCA in low doses results in more reduction in depression than placebo, and there was no strong evidence to show that standard-dosage tricyclic antidepressants obtain more response than low dosage (Furukawa et al., 2004).

We have defined decrease in MADRS rating as the main outcome measure. A more pragmatic approach could have been to analyse the proportion of patients still receiving the antidepressant assigned by the random allocation at the end of the study. If that outcome is analysed, 23 (43.4%) patients in fluoxetine group were still receiving the same antidepressant at the end of the study, while the figure for imipramine was 18 (36%). Differences were not statistically significant. Sixty four percent of patients in the fluoxetine and imipramine groups were on antidepressant therapy at the end of the study. In contrast, Simon et al. (1999) found that patients taking fluoxetine therapy were significantly more likely to continue taking the initial antidepressant compared to imipramine or desipramine at 24 months, but were equally likely to still continue any antidepressant therapy. These differences can be attributed to the study duration.

Some considerations need to be made in the interpretation of the results. First, physicians participated in the study on a voluntary basis. It is possible that only the more motivated physicians formed a self-selected group that is less representative of all GPs. Second, the requirement of informed consent may have influenced the selection of patients. However, this is true of all studies using human subjects. Third, the use of exclusion criteria, though it is both necessary and less restrictive than in randomised controlled trials, limits the generalizability of the results. Fourth, randomisation affects the physician's choice of treatment and thus distances the setting from normal clinical practice. Fifth, the measurement of indirect costs associated with sick leave represents another limitation. The same cost per sick day was used for all patients, but a

more direct measure of this variable is needed in order to arrive at a more accurate calculation of the real costs involved. Nevertheless, there were no differences in the level of professional occupation between treatment groups. Sixth, sick leave days may reflect physical as well as mental disorders. Seventh, the small sample size of our study limited our ability to detect differences by diagnostic group. Finally, this trial followed patients for six months, which is longer in comparison to typical 6–8 week trials. However, in everyday clinical practice patients receive treatment for longer periods of time (Simon et al., 1999).

A further limitation of the study is that no data was recorded on the incidence of side effects, since GPs poorly recorded them in the clinical charts. An indirect measure of tolerability is the rate of treatment change. In a comparison of SSRI versus TCA, Anderson and Tomenson (1995) found that the most frequent reason for treatment dropout was the presence of side effects. In our study, dropout rates are similar for both groups. A meta-analysis performed by Song et al. (1993) also revealed no significant differences in dropout rates between patients treated with SSRI and those treated with TCA.

The data for this study and the unit costs were recorded during the period of 1999–2001. Drug prices have varied since then, and it is possible that direct costs do not reflect exact drug prices of today. However, in our study drug prices only represent 6% of total costs.

In conclusion, this study demonstrates no differences in the effectiveness of antidepressants. Longer studies are needed to confirm these results and their applicability to clinical practice.

Acknowledgments

The authors gratefully acknowledge the contribution of Maria Mallen and Juan Vicente Torres Martín to the statistical analysis presented here. We also wish to acknowledge Susan M. DiGiacomo, Ph.D., of the Fundació Sant Joan de Déu, for editorial assistance in the preparation of successive English-language versions of the manuscript.

References

American Psychiatric Association, M., 1993. Practice guideline for major depressive disorders in adults. Am. J. Psychiatry 150 (4 suppl), 1–22.

Anderson, I.M., Tomenson, B., 1995. Treatment discontinuation with selective serotonin reuptake inhibitors compared with tricyclic antidepressants: a meta-analysis. BMJ, 1433–1438.

Badia, X., 2003. La investigación de resultados en salud y el impacto de los costes de la salud mental. In: López-Ibor Aliño, J.J., Gómez Pérez, J.C., Gutiérrez Fuentes, J.A. (Eds.), Retos para la Psiquiatría y la Salud Mental en España. Ars Medica, Barcelona, pp. 25–37.

Barbui, C., Percudani, M., Hotopf, M., 2003. Economic evaluation of antidepressive agents: a systematic critique of experimental and observational studies. J. Clin. Psychopharmacol. 23, 145–154.

Base de Datos de Costes Unitarios de Soikos, 2001.

Berto, P., D'Ilario, D., Ruffo, P., Di Virgilio, R., Rizzo, F., 2000. Depression: cost-of-illness studies in the international literature, a review. J. Mental Health Policy Econ. 3, 3–10.

Blakey, A.F., 1999. Use of tricyclic antidepressants in primary care. Acta Psychiatr. Scand. 100, 317.

Concato, J., Shah, N., Horwit, R.I., 2000. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N. Engl. J. Med. 342, 1887–1892.

Crott, R., Gilis, P., 1998. Economic comparisons of the pharmacotherapy of depression: an overview. Acta Psychiatr. Scand. 97, 241–252.

Crown, W.H., 2000. Economic outcomes associated with tricyclic antidepressant and selective serotonin reuptake inhibitor treatments for depression. Acta Psychiatr. Scand. 101 (suppl 403), 62–66.

Crown, W.H., 2001. Antidepressant selection and economic outcome: a review of methods and studies from clinical practice. Br. J. Psychiatry 179 (suppl 42), s18–s22.

Donoghue, J., Hylan, T.R., 2001. Antidepressant use in clinical practice: efficacy v. effectiveness. Br. J. Psychiatry 179 (suppl 42), s9–s17.

Doyle, J.J., Casciano, J., Arikian, S., Tarride, J., Gonzalez, M.A., Casciano, R., 2001. A multinational pharmacoeconomic evaluation of acute major depressive disorder (MDD): a comparison of cost-effectiveness between venlafaxine, SSRI and TCAs. Value Health 4 (1), 16–31.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2001. Entrevista Clínica Estructurada Para Los Trastornos Del Eje I del DSM-IV. Masson, Barcelona.

Furukawa, T., McGuire, H., Barbui, C., 2004. Low dosage tricyclic antidepressants for depression. The Cochrane Library, Issue 2. John Wiley & Sons, Ltd., Chichester, UK, pp. 1–36.

Garattini, S., Barbui, C., Saraceno, B., 1998. Antidepressant agents: from tricyclics to serotonin uptake inhibitors. Psychol. Med. 28, 1169–1178.

Greenberg, P.E., Stiglin, L.E., Finkelstein, S.N., Berndt, E.R., 1993. The economic burden of depression in 1990. J. Clin. Psychiatry 54 (11), 405–418.

Hotopf, M., Lewis, G., Normand, C., 1995. Are SSRIs a cost-effective alternative to tricyclics. Br. J. Psychiatry 168, 404–408.

Jönsson, B., Bebbington, P.E., 1994. What price depression? The cost of depression and the cost-effectiveness of pharmacological treatment. Br. J. Psychiatry 164, 665–673.

Le Pen, C., Ley, E., Ravily, V., Beuzen, J.N., Meurgey, F., 1994. The cost of treatment dropout in depression. A cost–benefit analysis of fluoxetine vs. tricyclics. J. Affect. Disord. 31, 1–18.

Lobo, A., Chamorro, L., Luque, A., Dal-Re, R., Badia, X., Baro, E., 2002. Validación de las versiones en español de la Montgomery–Asberg Depression Rating Scale y la Hamilton Anxiety Rating Scale para la evaluación de la depresión y de la ansiedad. Med. Clin. (Barc) 118 (13), 493–499.

MacGillivray, S., Arroll, B., Hatcher, S., Ogston, S., Reid, I., Sullivan, F., Williams, B., Crombie, I., 2003. Efficacy and tolerability of selective serotonine reuptake inhibitors compared with tricyclic

antidepressants in depression treated in primary care: systematic review and meta-analysis. BMJ 326, 1014–1018.

Manning, W.G., Mullahy, J., 2001. Estimating log models: to transform or not to transform? J. Health Econ. 20, 461–494.

Mendlewicz, J., 2001. Optimising antidepressant use in clinical practice: towards criteria for antidepressant selection. Br. J. Psychiatry 179 (suppl 42), s1–s3.

Montgomery, S.A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. Br. J. Psychiatry 134, 382–389.

Mulrow, C.D., Williams, J.W., Ciquette, E., Aguilar, C., Hitchcock Noël, P., Lee, S., Cornell, J., Stamm, K., 2000. Efficacy of newer medications for treating depression in Primary Care patients. Am. J. Med. 108, 54–64.

Murray, C., López, A.D., 1996. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Global of burden of disease of injuries series. The Harvard School of Public Health, on behalf of the World Heald Organization and The World Bank, Boston.

Nielsen, B., Behnke, K., Christiansen, P., Geisler, A., Ipsen, E., Maach-Moller, B., Ohrberg, S., 1993. A comparison of fluoxetine and imipramine in the treatment of outpatients with major depressive disorder. Acta Psychiatr. Scand. 87, 269–272.

Panzarino, P.J., 1998. The costs of depression: direct and indirect; treatment versus nontreatment. J. Clin. Psychiatry 59 (suppl 20), 11–14.

Revicki, D.A., Brow, R.E., Palmer, W., Bakish, D., Rosser, W., Anton, S.F., Feeny, D., 1995. Modelling the cost effectiveness of antidepressant treatment in primary care. PharmacoEconomics 8 (6), 524–540.

Rice, D.P., Miller, P., 1995. The economic burden of affective disorders. Br. J. Psychiatry 166, 34–42 (suppl).

Sacristan, J.A., Gilaberte Asin, I., Hylan, T.R., Crown, W.H., Bueno, G., Garzón, M.A., Montejo, A.L., 1999. Costes del tratamiento con nuevos antidepresivos en la práctica clínica. Aten. Prim. 23, 15–25.

Simon, G.E., Wagner, E., Vonkorff, M., 1995. Cost-effectiveness comparisons using "real world" randomized trials: the case of new antidepressant drugs. J. Clin. Epidemiol. 48 (3), 363–373.

Simon, G.E., Vonkorff, M., Heilingenstein, J.H., Revicki, D.A., Grothaus, L., Katon, W., Wagner, E., 1996. Initial antidepressant choice in Primary Care. Effectiveness and cost of fluoxetine vs tricyclic antidepressants. JAMA 275, 1897–1902.

Simon, G.E., Heilingenstein, J.H., Revicki, D.A., Vonkorff, M., Katon, W., Ludman, E., et al., 1999. Long-term outcomes of initial antidepressant drug choice in a "real world" randomized trial. Arch. Fam. Med. 8 (4), 319–325.

Song, F., Freemantl, N., Sheldon, T.A., House, A., Watson, P., Long, A., Mason, J., 1993. Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. BMJ 306, 683–687.

Stewart, A., 1997. Choosing an antidepressant: effectiveness based pharmacoeconomics. J. Affect. Disord. 48, 125–133.

Stewart, W.F., Ricci, J.A., Chee, E., Hahn, S.R., Morganstein, D., 2003. Cost of lost productive work time among US workers with depression. JAMA 289, 3135–3144.

Wells, K.B., 1999. Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. Am. J. Psychiatry 156, 5–10.

Woods, S.W., Rizzo, J.A., 1997. Cost-effectiveness of antidepressant treatment. Br. J. Psychiatry 170, 257–263.

Zimmerman, M., Mattia, J.I., Posternak, M.A., 2002. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? Am. J. Psychiatry 159 (3), 469–473.