Psychological Medicine, 1994, **24**, 307–316.

Antipsychotic drugs: is more worse? A meta-analysis of the published randomized control trials

P. BOLLINI,[1] S. PAMPALLONA, M. J. ORZA, M. E. ADAMS AND T. C. CHALMERS

From the Technology Assessment Group, Harvard School of Public Health, Boston, MA., USA

SYNOPSIS Effectiveness and side-effects of high- *versus* low-dose neuroleptic treatment of chronic psychosis have been assessed through a meta-analysis of 22 published randomized control trials comparing different neuroleptic doses. No incremental clinical improvement was found at doses above 375 mg equivalent of chlorpromazine, while a significant increase in adverse reactions was observed.

INTRODUCTION

Neuroleptic (NL) medications, also called antipsychotics, are widely used in clinical practice to treat the symptoms of psychosis and other severe mental disorders. Their primary beneficial effect is the reduction of positive symptoms of psychosis (e.g. delusions, hallucinations, bizarre behaviour, etc.). However, their effectiveness in controlling negative symptoms (e.g. withdrawal, blunted affect, psychomotor retardation, etc.) has been questioned. Because of the occurrence of adverse reactions, which can be severe and impair patients' social functioning, the balance between the benefits and risks of neuroleptic treatment is extremely delicate (Baldessarini, 1985).

Concern about adverse reactions, in particular tardive dyskinesia (involuntary movements of oral and limb muscles, often irreversible, for which no cure has been found), coupled with the apparent diminished effectiveness for negative symptoms characteristic of chronic psychosis, have raised questions about the appropriateness of long-term maintenance with NL drugs. Furthermore, the most appropriate dose of NL drugs administered as maintenance treatment has been debated for several years (see for example Gardos & Cole, 1973; Anderson & Kuehnle, 1976; Baldessarini *et al.* 1976; Eriksen *et al.* 1976; Aubree & Lader, 1980; Baldessarini & Davis, 1980; Kane *et al.* 1986; Baldessarini *et al.* 1988).

[1] Address for correspondence: Dr Paola Bollini, I.O.M., 17 Route des Morillons, 1211 Geneva, Switzerland.

A few years after the discovery of the usefulness of the first NL (chlorpromazine) in psychiatric patients, a wide range of doses were explored in clinical practice, from 200 mg daily in the first trial (Delay & Deniker, 1952) up to 4000 mg in subsequent trials (Aubree, 1979). The justification for high doses was ill-defined, but, nonetheless, treatment with high doses became widespread. Largely based on uncontrolled studies, high-dose NL treatment was purported to suppress symptoms faster than conventional doses, leading to decreased length of hospitalization and to few, if any, adverse reactions (Oldham & Bott, 1971; Sangiovanni *et al.* 1973; Ayd, 1977).

Although most of the subsequent randomized control trials (RCTs) have not supported the greater advantage of high-dose regimens over conventional ones, high-dose regimens are still widely used in clinical practice (Baldessarini *et al.* 1984; Bollini *et al.* 1984; Magno Zito *et al.* 1987; Holloway, 1988; Reardon *et al.* 1989). Unfortunately, the level of exposure to and duration of treatment with NL drugs appear related to the prevalence of some types of adverse reactions, suggesting that many patients are exposed to unnecessarily high and potentially dangerous treatments.

The optimal maintenance dose of NL drugs has been tested in a number of clinical studies with reasonably similar treatment protocols. Until now, no attempt has been made to combine the results of clinical trials with random patient assignment to different dose groups, or to assess the differing quality of the trials, although

thoughtful reviews of the field have been published (Baldessarini *et al.* 1988).

Meta-analysis is a technique for combining the results of collections of research papers to answer specific questions, often in a quantitative manner (Glass, 1976; Louis *et al.* 1985; Sacks *et al.* 1987). Prompted by the need to synthesize effectively the large body of research findings that has progressively accumulated, and that no longer seems adequately summarized by qualitative reviews, meta-analysis has been used in several fields of medicine and public health. Recently, the impact of NL on tardive dyskinesia and the control of agitated demented patients has been the object of two meta-analyses (Morgenstern *et al.* 1987; Schneider *et al.* 1990). We believe that a quantitative estimate of the benefits and risks of different NL doses in chronic schizophrenic patients is both timely and relevant because of the widespread use of high dose regimens, whose effectiveness and safety have not been adequately assessed.

Therefore, we have designed the present study to answer two questions: (1) are higher doses of NL drugs for maintenance treatment of psychosis any more effective than lower doses?; (2) are higher doses safer?

METHOD

We searched the MEDLINE database, from 1966 to June 1989, and other sources (review articles, reference lists from other papers), to find all the published RCTs that compared different doses of the same NL drug (when two or more papers reported the same study, the most recent version was chosen). To be selected for the meta-analysis, papers had to satisfy the following six criteria:

1. maintenance treatment of psychotic disorders (excluding treatment of the acute phase);
2. adult patients (older than 16 years);
3. random assignment of patients to treatment;
4. availability of conversion factor of studied treatment to chlorpromazine;
5. duration of trial of at least 4 weeks;
6. English language.

A trained meta-analysis technician blinded all articles selected by the initial screening process, to diminish the bias in the selection of the final papers (Sacks *et al.* 1987). Authors, date of publication, journal of publication, date of study, site of study, and all the study findings were obliterated, and the articles re-photocopied so that the reader could not identify them. Two readers (M.J.O. and M.E.A.) scored the blinded version of the articles according to criteria for quality previously described, and determined a quantitative estimate of quality by consensus (Sacks *et al.* 1987).

We used tables devised from extensive review of comparative RCTs and from clinical experience (Baldessarini *et al.* 1980, 1988) to convert doses of the various NL used in the selected articles to the equivalent amount in mg of chlorpromazine. As an approximate equivalence, 25 mg of Fluphenazine enanthate or decanoate every 2–3 weeks was converted to a daily oral dose of 500 mg of chlorpromazine. For Flupenthixol decanoate, 18 mg every 2 weeks was converted to a daily oral dose of 225 mg of chlorpromazine. This procedure enabled us to obtain comparable doses across studies. All but three studies considered two doses of a chosen NL; when more than two doses of the same NL were used, we considered only the highest and the lowest in the analysis. We examined clinical effectiveness and occurrence of side effects as the main endpoints.

In an attempt to provide a *post-hoc* analysis by intention to treat, we identified three broad categories of drop-outs from the papers: (1) lack of symptomatic control; (2) presence of severe adverse reactions; and (3) other reasons (e.g. intercurrent illnesses, transfer to other wards, lack of compliance, and discharge from the hospital). For each of the outcomes of interest, patients who dropped out for lack of symptomatic control or adverse reactions have been reconsidered for the present meta-analysis as described in the following sections.

Clinical effectiveness

Each study defined improvement either as a rating on a clinical global impression scale, or as no psychotic relapse, or as no rehospitalization. The three measures point in the same direction and are consistent expressions of positive outcome. With either of these definitions, each study expressed the outcome of the study group as a proportion of improved patients. For the purpose of the meta-analysis, we considered

patients who had dropped out of a study because of lack of symptomatic control as not improved.

Side effects

Side effects were the second main endpoint that we considered in the analysis. Seven studies did not mention side-effects, and five reported them in such a way that their data could not be combined because only the average scores of specific scales were reported. Therefore, we extracted data on side-effects from only 10 studies. We have analysed adverse reactions in three different ways: (1) any; (2) neurological (i.e. acute dystonia, akathisia, tardive dyskinesia, and Parkinsonism); and (3) Parkinsonism alone. For the purpose of the meta-analysis, we counted patients dropping out because of adverse reactions among those with side-effects.

Statistical analysis

For each treatment arm we computed the following summary statistics: (1) proportion of improvement, defined as the number of patients who had improved divided by the number of patients who completed the study plus those who dropped out because of lack of symptomatic control; (2) average number of adverse reactions per patient, defined as the total number of adverse reactions (including those experienced by patients who dropped out because of side effects) divided by the number of patients who completed the study plus those who dropped out because of side effects; (3) average number of neurological adverse reactions (i.e. acute dystonia, akathisia, tardive dyskinesia, and Parkinsonism), defined as the total number of neurological adverse reactions divided by the total number of patients (as defined above); and (4) proportion of Parkinsonism, defined as the number of patients who had experienced Parkinsonism divided by the total number of patients (as defined above).

In an attempt to describe the dose–response relationship, we conducted an analysis in which each treatment arm constituted the sampling unit. For this analysis, we considered each of the above quantities in turn as the dependent variable and assessed its relationship with the independent variable, NL dose used in each arm of each study, by means of a covariance model.

In order to avoid the assumption of a linear dose–response model, which is not appropriate in modeling the dose–effect relationship for NL, we considered four dose levels, dividing the treatment arms into four groups of equal size, according to the observed quartiles of administered dose: 25–165 mg, 166–375 mg, 376–830 mg, > 830 mg. We treated dose level and study as categorical variables. We have defined a dummy variable for the first three levels of dose, thus considering the highest level as reference. The model also included a random effect for study, which permits adjustment for between study variation and also allows for the correlation between observations pertaining to the same study. In order to account for the different precision of the summary statistics provided by each treatment arm, we weighted the analysis by the corresponding sample size. Appropriateness of the linear models was assessed by analysis of residuals, and statistical significance was tested by appropriate F tests. No adjustments of P values for multiple comparisons have been considered.

RESULTS

From the above sources, we identified and blinded a total of 30 potentially acceptable studies. Of these, we excluded 8 studies for two reasons: (1) four because they were cross-over studies, with data not separable by arm (Faleni, 1970; de Buck, 1972; Deneker *et al.* 1978; Wiles, 1980); and (2) four because their data on effectiveness could not be extracted (Brotman *et al.* 1969; Clark *et al.* 1970; Chouinard & Annable, 1976; McCreadie *et al.* 1979). Thus, we included in the analysis a total of 22 RCTs comparing two doses of the same NL drug as maintenance treatment of chronic psychosis. Two of these studies were originally designed to evaluate two doses of each of three different NL drugs: thus, for the purpose of further analyses, we have considered a total of $20+6=26$ sampling units.

The main characteristics of the selected studies are shown in Table 1. The studies with one or more pairs of treatment arms, one at a lower and one at a higher dose, as defined by the investigators, generated a collection of 52 treatment groups. A total of 1638 patients were considered, 833 in the lower dose category and 805 in the higher dose category. Almost invariably, patients had a diagnosis of schizo-

Table 1. *Randomized Control Trials comparing high and low neuroleptic treatment as maintenance in chronic psychosis*

Study	Setting (I/O)*	Length (weeks)	Type of NL	CPZ/eq. daily dose (mg)	No. of patients	
					Analysed	Improved
Caffey *et al.* 1964	I	16	Phenothiazines	160	89	76
				375	88	84
Prien and Cole, 1968	I	24	Chlorpromazine	300	215	34
				1888	201	52
Simpson *et al.* 1968	I	16	Butaperazine	263	6	1
				1450	6	0
Williams *et al.* 1969	I	24	Fluphenazine HCl	250	6	6
				500	4	4
Williams *et al.* 1969	I	24	Trifluopromazine	165	6	2
				330	4	1
Williams *et al.* 1969	I	24	Trifluopromazine	415	7	1
			+ Fluphenazine	830	9	8
Prien *et al.* 1969	I	24	Trifluoperazine	375	105	26
				2000	113	26
Gardos *et al.* 1974	I	16	Thiotixene	250	18	4
				1000	19	4
Quitkin *et al.* 1975	I	6	Fluphenazine HCl	950	13	9
				39000	18	4
Chien, 1975	O	52	Fluphenazine	258	16	10
			enanthate	795	16	14
McClelland *et al.* 1976	I	24	Fluphenazine	500	25	6
			decanoate	10000	24	9
Clark *et al.* 1977	I	12	Loxapine	472	11	5
				931	12	8
Goldstein *et al.* 1978	O	6	Fluphenazine	125	45	38
			enanthate	500	51	48
Bjorndal *et al.* 1980	I	12	Haloperidol	552	11	6
				6210	12	5
Branchey *et al.* 1981	I	42	Loxapine	160	22	6
				664	11	9
Kane *et al.* 1983	O	52	Fluphenazine	50	62	36
			decanoate	500	64	61
Nishikawa *et al.* 1984	O	52	Haloperidol	45	12	3
				270	11	6
Nishikawa *et al.* 1985	O	52	Thioridazine	25	11	0
				75	7	1
Nishikawa *et al.* 1985	O	52	Pimozide	100	10	0
				300	8	3
Nishikawa *et al.* 1985	O	52	Pimozide +	125	9	6
			Thioridazine	375	5	5
Marder *et al.* 1987	O	104	Fluphenazine	100	33	25
			decanoate	500	25	22
Carpenter *et al.* 1987†	O	104	Various	196	14	7
				720	12	5
Johnson *et al.* 1987	O	52	Flupenthixol	150	28	19
			decanoate	225	31	28
Huang *et al.* 1987	I	9	Thiotixene	1500	20	4
				6781	20	13
Cookson, 1987	I	44	Flupenthixol	1475	9	6
			decanoate	4163	9	8
Hogarty *et al.* 1988	O	104	Fluphenazine	76	30	21
			decanoate	500	25	19

* I = In-patients; O = Out-patients.
† Involved continuous *versus* targeted neuroleptic treatment.

phrenia with a chronic clinical course. Studies conducted before 1980 often dealt with in-patients (10 out of 12), while after 1980 most studies involved out-patients (7 out of 10), mirroring the shift of care from the mental hospital to community settings. Sixteen studies were conducted in the US, three in the UK, two in Japan and one in Denmark. Study length varied from 6 to 104 weeks, with an average of 38·5 weeks (±S.D. 31·3 weeks). In five studies

Table 2. *Results of the regression of clinical effectiveness (proportion of patients improved) on neuroleptic dose levels, corrected for study effect*

Clinical effectiveness	B*	S.E. (B)	*P*
Dose level (mg)			
⩽ 165	−0·226	0·0797	0·009
166–375	−0·058	0·0501	0·256
376–830	−0·004	0·0839	0·961
⩾ 831	—	—	—

F test for regression with dose and study: $F = 10{\cdot}93$; df = 28,23; $P < 0{\cdot}0001$.
F test for dose, after inclusion of study: $F = 5{\cdot}08$; df = 3,23; $P = 0{\cdot}008$.

* Estimated regression coefficient.

drug-resistant patients (i.e. patients who previously did not respond to rather high NL regimens administered from 6 weeks to more than 2 years) were recruited. Finally, eight studies used fixed doses (Simpson *et al.* 1968; Prien *et al.* 1969; Williams *et al.* 1969; Gardos *et al.* 1974; Quitkin *et al.* 1975; McClelland *et al.* 1976; Nishikawa *et al.* 1985, 1985; Huang *et al.* 1987), and the remaining flexible doses.

Lower doses, as defined by the investigators, ranged from 25–1500 mg chlorpromazine equivalent, with an average of 348 mg (±S.D. 391 mg). Higher doses, similarly defined by the investigators, spread from 75–39000 mg, with an average of 3110 mg (±S.D. 7712 mg). Although the dose of 39000 mg was a definite outlier, out of 52 treatment groups daily doses of 1000 mg and above were tested eleven times, four of which above 6000 mg. The year of publication of the selected studies went from 1964 to 1987. The doses of neuroleptics tested did not show any appreciable change over time. The average low dose administered in trials conducted before 1980 was 345 mg, and after 1980 350 mg. Similarly, the average high dose tested before 1980 was 1717 mg (excluding the study of Quitkin *et al.* 1975, having a high dose of 39000 mg), and 1637 mg thereafter.

We explored the association between NL dose and clinical effectiveness and side-effects, respectively, by means of a linear regression model. Possible differences among studies (in terms of patient sex, age, type of setting, etc.) were taken into account by introducing a term for each study in the regression, and we weighted the analysis by sample size. Thereafter, coefficients for dose given in subsequent sections have to be considered as adjusted for possible differences of baselines among studies.

Regression for clinical effectiveness

We regressed the proportion of improvement on dose groups (25–165 mg, 166–375 mg, 376–830 mg, > 830 mg) and study as described in the

Table 3. *Number and type of adverse reactions reported in* 10 *studies for low and high chlorpromazine dose*

Study	CPZ dose (mg)	Acute dyst.	Parkinsonism	Akathisia	Tardive dyskin.	Hypotension	Seizures	Oversedation	Skin reactions	Other
Gardos *et al.*	250	0	4	5	0	0	0	0	0	6
1974	1000	0	9	10	0	0	0	0	0	4
Prien & Cole	300	2	17	19	0	2	2	31	6	66
1968	1888	19	39	19	0	4	12	79	60	215
Quitkin *et al.*	950	0	0	0	0	0	0	0	0	0
1975	39000	0	7	0	0	0	0	0	0	7
McClelland *et al.*	500	0	5	2	2	0	2	0	0	0
1976	10000	1	8	5	5	0	0	0	0	1
Prien *et al.*	375	11	12	14	0	21	0	6	6	36
1969	2000	29	56	39	0	15	1	14	6	42
Hogarty *et al.*	76	0	0	0	4	0	0	0	0	0
1988	500	0	0	0	7	0	0	0	0	0
Simpson *et al.*	263	1	4	0	0	0	0	0	0	0
1968	1450	5	6	0	0	0	0	0	0	0
Bjorndal *et al.*	552	0	1	1	2	0	0	0	0	0
1980	6210	0	3	3	0	0	1	5	0	0
Clark *et al.*	472	0	5	1	0	3	0	6	0	3
1977	931	1	5	2	1	1	0	9	0	9
Huang *et al.*	1500	9	4	3	0	3	0	7	0	1
1987	6781	12	6	0	0	4	0	12	0	8

Method section. The value of the overall test for the significance of the regression was 10·93, 28 and 23 df, $P < 0·0001$. The partial F test for the terms for dose, when the terms for study were in the model, was $F = 5·08$, 3 and 23 df, $P < 0·01$. The estimated coefficients for the four dose levels are shown in Table 2. This Table must be interpreted considering the coding system we have adopted for dose category. For instance, the coefficient for the lowest dose level tells how the prediction is affected by changing the dose from the highest to the lowest level. The coefficients for the other two intermediate dose levels should be interpreted in a similar way. Thus, the highest dose level is always taken as the reference, and its coefficient is set to zero by default.

The only statistically significant coefficient is the one for the lowest dose group ($b = -0·226$, S.E. $(b) = 0·080$, $P < 0·01$). Because the dependent variable indicates the percentage of improved patients, the interpretation of this coefficient is that on average the percentage of improved patients in the lowest dose group was 22·6 percentage points lower than in the highest dose group, and significantly so. For the other 2 dose groups (166–375 mg and 376–830 mg), the associated reduction in percentage of improved patients was 5·8 ($P = 0·256$) and 0·4 ($P = 0·961$) percentage points respectively as compared to the percentage of improved patients in the highest dose category. In other words, for doses beyond the range 166–375 mg, no significant increases in effect could be detected as compared to the highest dose group. On average, the same percentage of improved patients was found at doses in the ranges 166–375 mg, 376–830 mg, and > 830 mg.

Regression for side effects

Ten studies reported adverse reactions in a way suitable for data extraction, as shown in Table 3. Parkinsonism was the single most reported adverse reaction (191 cases in both dose groups, corresponding to 20% of the cases), followed by oversedation (18%), akathisia (13%) and acute dystonia (9%). A total of 21 cases of tardive dyskinesia were reported, corresponding to 2% of the patients considered. With the exception of hypotension, adverse reactions were more frequent in the higher dose group, as defined by the investigators.

Table 4. *Results of the regression of side-effects (average number per patient) on neuroleptic dose levels, corrected for study effect*

	B*	S.E. (B)	P
Parkinsonism			
Dose level (mg)			
⩽ 375	−0·181	0·0564	0·0126
376–830	−0·124	0·1285	0·3635
⩾ 831	—	—	—
F test for regression with dose and study: $F = 3·20$; df = 11,8; $P = 0·055$.			
Partial F test for dose level, after inclusion of study: $F = 5·24$; df = 2,8; $P = 0·035$			
Neurological side-effects†			
Dose level (mg)			
⩽ 375	−0·359	0·1040	0·0086
376–830	−0·262	0·2372	0·3013
⩾ 831	—	—	—
F test for regression with dose and study: $F = 3·83$; df = 11,8; $P = 0·033$.			
Partial F test for dose level, after inclusion of study: $F = 6·12$; df = 2,8; $P = 0·0024$.			
All side-effects			
Dose level (mg)			
⩽ 375	−1·009	0·1038	0·0001
376–830	−0·739	0·2365	0·0142
⩾ 831	—	—	—
F test for regression with dose and study: $F = 21·30$; df = 11,8; $P < 0·0001$.			
Partial F test for dose level, after inclusion of study; $F = 48·58$; df = 2,8; $P < 0·0001$.			

* Estimated regression coefficient.
† Including acute dystonia, akathisia, tardive dyskinesia, Parkinsonism.

The association between NL dose and adverse reactions was explored by means of a linear regression model described in the Method section, having as dependent variable at a time the proportion of Parkinsonism, the average number of neurological adverse reactions (i.e. acute dystonia, akathisia, tardive dyskinesia, and Parkinsonism), and the average number of adverse reactions of any kind. For consistency, the same categories of dose level were considered in this regression as in the analysis of clinical effectiveness. However, because only 10 studies reported side effects in a way suitable for data extraction, the sparseness of the data has required lumping together the first two categories of dose. Table 4 summarizes the results of the analysis. For the three groups of side effects, the F tests for the regression were always significant, as were the F tests for the effect of dose after inclusion of the terms for study. On

Table 5. *Some key criteria for quality assessment*

Criteria	Maximum score possible	% Achieving maximum score ($N = 22$)
Selection description	3	64
Therapeutic regimen definition	3	36
Blinding of observers	8	27
Blinding of randomization process	10	5
Prior estimate of sample size	3	0
Reporting of withdrawals	4	36
Handling of withdrawals	4	9
Reporting of side effects	3	27

average, after adjustment for study, the percentage of patients with Parkinsonism, as predicted by the model, was 18 percentage points ($P = 0{\cdot}0126$) less in the lowest dose category ($\leqslant 375$ mg) than in the highest dose category ($\geqslant 831$ mg). The same patients experienced on average 0·359 ($P = 0{\cdot}0086$) fewer neurological adverse reactions and 1·009 fewer adverse reactions of any kind ($P = 0{\cdot}0001$). For the latter endpoint, a significant reduction ($b = -0{\cdot}739, P = 0{\cdot}0142$) was observed also in the intermediate dose group (376–830 mg).

Quality scoring

The mean quality score of the 22 selected papers was 0·32 ($\pm$S.D. 0·11). Table 5 shows the percentage of articles that achieved the maximum score possible on eight criteria we judged to be particularly important for studies of high- and low-dose NL. Overall, only the reporting of the selection of the study patients was satisfactory, while the reporting of the calculation of the sample size, the blinding of randomization process, the handling of withdrawals in the analysis and the reporting of side-effects received the lowest quality score.

DISCUSSION

Meta-analysis is a quantitative approach to the review and synthesis of a number of research papers addressing the same research issue. Because clinical studies are often conceived differently, and often report information differently, meta-analysis summarizes simple data on treatment and major endpoints. For this reason, it usually answers research questions more sharply defined than the ones of the original collection of studies (Louis *et al.* 1985). Meta-analysis is helpful in the following ways: (1) increase statistical power for major endpoints and subgroup analyses; (2) help resolve uncertainty when reports disagree; (3) improve estimates of effect size; and (4) answer questions not posed at the start of individual trials (Sacks *et al.* 1987).

In the present study, we used meta-analysis to approach in a quantitative way the issue of the best maintenance NL dose in chronic psychosis. Drug utilization studies using mean chlorpromazine equivalent doses showed that high doses are still frequently used (Baldessarini *et al.* 1984; Bollini *et al.* 1984; Magno Zito *et al.* 1987; Holloway, 1988). One report indicated that the average dose administered has increased over the years (Reardon *et al.* 1989). Randomized control trials comparing different NL doses provided controversial results. With these premises, we felt that a quantitative overview could provide an objective synthesis of the available evidence.

In particular, using multiple regression we were able to evaluate clinical effectiveness and adverse reactions at different doses. Introducing a term for each study in the model allowed us to take into account different study characteristics (for instance, patient's age and sex, drug resistance status, therapeutic setting, etc.). Moreover, we weighted the regression according to study size, in order to account for the different variability of the estimates of the effect.

The present meta-analysis of higher *versus* lower dose maintenance treatment of chronic psychosis gave the following answers to our original questions: (1) there is no therapeutic advantage to be gained beyond the dose range between 166 and 375 mg equivalent of chlorpromazine; and (2) adverse reactions increase significantly above these doses.

Our results suggest that, on average, a dose up to 165 mg equivalent is significantly less effective as compared with the reference category of more than 830 mg. However, no significant therapeutic advantage is found above 375 mg equivalent of chlorpromazine. Of course, dose equivalence is an imperfect measure of NL dose, but the range of lowest effective dose that we have obtained is very similar to the figure of 300 mg equivalent of chlorpromazine identified

in the analysis of 23 clinical trials as the threshold dose for effectiveness, beyond which clinical advantages may not be found regularly (Baldessarini & Davis, 1980).

The apparent lack of incremental benefit with increase in NL dose might have been caused by an increase in frequency and severity of adverse reactions that affect behaviour. It is well known in fact that some dose-dependent side effects of NL drugs, such as Parkinsonism, over-sedation and akathisia, could easily be confused with clinical manifestations of psychosis (Quitkin *et al.* 1975; Bollini *et al.* 1984). Accordingly, a patient with Parkinsonism or oversedation may be considered withdrawn or depressed, or one with akathisia may be seen as agitated. Consistently with this hypothesis, both neurological and overall side effects were significantly increased with higher doses.

About one-third (7 out of 22) of the trials did not mention side effects, although they were probably present in all studies, and only 27% of the papers that reported side effects did so in a complete and accurate way. Side effects are often inadequately reported in clinical studies. Venulet and colleagues (1982), in a review of quality and completeness of articles on adverse drug reactions published between 1972 and 1979, observed that only 19% reported them adequately. Moreover, RCTs usually select patients in order to maximize the treatment difference they want to show. This selection, which is totally legitimate, often leads to the exclusion of patients more prone to side effects. These two reasons, inadequate reporting of side effects and selection of patients, probably led to incidence figures lower than the ones usually found in epidemiological studies.

The small number of studies included in the meta-analysis reporting adverse reactions unfortunately reduced the sample size for these analyses, and did not allow the study of each side effect separately, except for Parkinsonism. Neurological side effects, as expected, were the most common problem reported, in particular Parkinsonism and akathisia. Few incident cases of tardive dyskinesia were reported, a proportion of 0·02 and 0·03 in the lower and higher dose group respectively, as defined by the investigators. However, most of the patients had already been exposed to chronic NL treatment, and thus a number of them had already acquired tardive dyskinesia. Overall, patients receiving more than 375 mg equivalent of chlorpromazine showed a significantly higher occurrence of both Parkinsonism and neurological adverse reactions. It is worth noting that neurological adverse effects, and Parkinsonism in particular, might severely limit patients' social functioning.

The studies examined achieved an average quality score of 0·32. This compares with a mean score of 0·42 for a sample of 376 trials published between 1949 and 1986 in 18 subject areas examined by Reitman and colleagues (1987). However, none of the trials scored was from the psychiatric literature, making a quality comparison difficult. None of the trials included in the present meta-analysis reported a calculation of sample size before start of the study. Furthermore, only 8% of the studies handled drop-outs satisfactorily, although early interruption of treatment due to both worsening of clinical state and adverse reactions are likely to differ by treatment. Our decision to include the first in the category with known outcome, and the latter among adverse reactions is the best approximation to an adequate statistical handling of dropouts. In his review of design and analysis of psychopharmacology research, Goldberg (1987) acknowledged handling of dropouts and sample size requirements among the issues still in need of improvement.

Finally, we should mention that the present meta-analysis dealt only with published RCTs in English. The scope of the meta-analytical effort was initially to cover all literature on the topic, irrespective of language. During the initial phase of literature search, the authors realized that invariably publications in languages different from English did not meet the methodological standards to be eventually included in the overview. This is not to be interpreted as poor quality of research in non-English speaking countries, but rather as a reflection of the more stringent peer review system in journals published in English. At that stage, no further evaluation of non-English literature was performed.

Previous researchers have argued that unpublished results might importantly affect meta-analysis conclusions, insofar as studies with positive results are more likely to be published than studies with negative results (Begg & Berlin, 1989), although publication bias seems greater

in observational and laboratory-based experimental studies than in randomized clinical trials (Easterbrook *et al.* 1991). This point is of course very difficult to prove, because the results of unpublished studies are usually unavailable. In our case, we can only speculate that, since higher doses constituted the 'innovative' treatment, it is more likely that trials reaching results unfavourable to higher doses may not have been published. If this were so, then published studies would over-represent studies favouring 'effective' high doses: the results of the present meta-analysis would thus be biased downwards.

CONCLUSION

The results of the present meta-analysis support the view that moderate doses of NL drugs (roughly between 165 and 375 mg equivalent of chlorpromazine) should be preferred in the maintenance treatment of chronic psychosis. Higher doses fail to produce incremental improvement but significantly increase the occurrence of adverse reactions. Neuroleptic drugs have proven to be a crucial therapeutic tool in the treatment of psychosis. This does not imply that all patients should receive NL drugs, or drug treatment alone (Karon, 1989). Quantitative evidence from meta-analysis suggests that, when pharmacological treatment is deemed necessary for chronic schizophrenic patients, the optimal balance between clinical effectiveness and adverse reactions is found in the low to moderate dose range.

This project was supported by grant no. HS 05936 from the National Center for Health Services Research and Health Care Technology Assessment.

We wish to thank Frederick Mosteller for the guidance offered throughout the project, also Leon Eisenberg, Michael R. Reich and Thomas McLaughlin for their useful comments on the manuscript. We are indebted to Bruce Kupelnick and Robert Boyle for their contribution to the retrieval and blinding of the papers.

REFERENCES

Anderson, W. H. & Kuehnle, J. C. (1976). Dosage of antipsychotic drugs. *New England Journal of Medicine* **294**, 670.

Aubrée, J. C. (1979). Le problème des posologies des neuroleptiques et leurs variations géographiques. Mémoire de Psychiatrie, Université René Descartes (Paris V).

Aubrée, J. C. & Lader, M. H. (1980). High and very high dosage antipsychotics: a critical review. *Journal of Clinical Psychiatry* **41**, 341–350.

Ayd, F. J. Jr. (1977). Guidelines for using short-acting intramuscular Neuroleptics for rapid neuroleptisation. *International Drug Therapy Newsletter* **12**, 5–12.

Baldessarini, R. J. (1985). *Chemotherapy in Psychiatry. Principles and Practice*. Harvard University Press: Cambridge, MA.

Baldessarini, R. J. & Davis, J. M. (1980). What is the best maintenance dose of Neuroleptics in schizophrenia? *Psychiatry Research* **3**, 115–122.

Baldessarini, R. J., Gelenberg, A. J. & Lipinski, J. F. (1976). Grams of antipsychotics? *New England Journal of Medicine* **294**, 113–114.

Baldessarini, R. J., Cole, J. O., Davis, J. M., Gardos, G., Simpson, G. & Tarsy, D. (1980). *Tardive Dyskinesia. Task Force Report no. 18*. American Psychiatric Association: Washington DC.

Baldessarini, R. J., Katz, B. & Cotton, P. (1984). Dissimilar dosing with high-potency and low-potency neuroleptics. American Journal of Psychiatry **141**, 748–752.

Baldessarini, R. J., Cohen, B. M. & Teicher, M. (1988). Significance of neuroleptic dose and plasma level in the pharmacological treatment of psychoses. *Archives of General Psychiatry* **45**, 79–91.

Begg, C. G. & Berlin, J. A. (1989). Publication bias and dissemination of clinical research. *Journal of National Cancer Institute* **81**, 107–115.

Bjorndal, N., Bjerre, M., Gerlach, J., Kristjansen, P., Magelund, G., Oestrich, I. H. & Waehrens, J. (1980). High dosage haloperidol therapy in chronic schizophrenic patients: a double blind study of clinical response, side effects, serum haloperidol, and serum prolactin. *Psychopharmacology* **67**, 17–23.

Bollini, P., Andreani, A., Colombo, F., Bellantuono, C. Beretta, P., Arduini, A., Galli, T. & Tognoni, G. (1984). High-dose neuroleptics: uncontrolled clinical practice confirms controlled clinical trials. *British Journal of Psychiatry* **144**, 25–27.

Branchey, M. H., Branchey, L. B. & Richardson, M. A. (1981). Effects of neuroleptic adjustment on clinical condition and tardive dyskinesia in schizophrenic patients. *American Journal of Psychiatry* **138**, 608–612.

Brotman, R. K., Muzekari, L. H. & Shanken, P. M. (1969). Butaperazine in chronic schizophrenic patients: a double blind study. *Current Therapeutic Research* **11**, 5–8.

Caffey, E. M., Diamond, L. S., Frank, T. V., Grasberger, J. C., Herman, L., Klett, C. J. & Rothstein, C. (1964). Discontinuation or reduction of chemotherapy in chronic schizophrenics. *Journal of Chronic Diseases* **17**, 347–358.

Carpenter, W. T., Heinrichs, D. W. & Hanlon, T. E. (1987). A comparative trial of pharmacologic strategies in schizophrenia. *American Journal of Psychiatry* **144**, 1466–1470.

Chien, C. P. (1975). Drugs and rehabilitation in schizophrenia. In *Drugs in Combination with Other Therapies* (ed. M. Greenblatt), pp. 13–34. Grune and Stratton: New York.

Chouinard, G. & Annable, L. (1976). Alpha-methyldopa-chlorpromazine combination in schizophrenic patients. *Neuropsychobiology* **2**, 118–126.

Clark, M. L., Ramsey, H. R., Ragland, R. E., Rahhal, D. K., Serafetinides, E. A. & Costiloe J. P. (1970). Chlorpromazine in chronic schizophrenia: behavioral dose–response relationships. *Psychopharmacologia* (Berl.) **18**, 260–270.

Clark, M. L., Paredes, A., Costiloe, J. P., Fulkerson, F. G. & Wood, F. (1977). Evaluation of two dose levels of loxapine succinate in chronic schizophrenia. *Diseases of the Nervous System* **38**, 7–10.

Cookson, I. B. (1987). The effects of a 50% reduction of Ciz(z)-Flupenthixol Decanoate in chronic schizophrenic patients maintained on a high dose regime. *International Clinical Psychopharmacology* **2**, 141–149.

De Buck, R. P. (1972). Relative safety and efficacy of high and low dose administration of fluphenazine HCl to psychotic patients. In *Proceedings of the Eighth International Congress of Neuropsychopharmacology, Excerpta Medica, International Congress Series No. 148*, pp. 265–272. Princeton, NJ.

Delay, J. & Deniker, P. (1952). Trente-huit cases de psychoses traitées par la cure prolongée et continue de 4560RP. Le Congrès des Al. et Neurol. de Langue Française. In *Compte Rendue du Congrès*, Masson et Cie: Paris.

Deneker, S. J., Johansson, R., Lundin, L. & Malm, U. (1978). High doses of fluphenazine enanthate in schizophrenia: a controlled study. *Acta Psychiatrica Scandinavica* **57**, 405–414.

Easterbrook, P. J., Berlin, J. A., Gapalan, R. & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet* **337**, 867–872.

Eriksen, S. E., Hurt, S. W. & Davis, J. M. (1976). Dosage of antipsychotic drugs. *New England Journal of Medicine*, **294**, 1296–1297.

Faleni, R. A. (1970). The use of high doses of fluphenazine in the treatment of psychotic patients. *Psychosomatics* **11**, 496–499.

Gardos, G. & Cole, J. O. (1973). The importance of dosage in antipsychotic drug administration. A review of dose-response studies. *Psychopharmacologia* (Berlin) **29**, 221–230.

Gardos, G., Hecht Orzack, M., Finn, G. & Cole, J. (1974). High and low dose thiothixene treatment in chronic schizophrenia. *Diseases of the Nervous System* **35**, 53–58.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Research* **5**, 3–8.

Goldberg, S. C. (1987). Persistent flaws in the design and analysis of psychopharmacology research. In *Psychopharmacology: the Third Generation of Progress* (ed. H. Y. Meltzer), pp. 1005–1012. Raven Press: New York.

Goldstein, M. J., Rodnick, E. H., Evans, J. R., May, P. R. A. & Steinberg, M. R. (1978). Drug and family therapy in the aftercare of acute schizophrenics. *Archives of General Psychiatry* **35**, 1169–1177.

Hogarty, G. E., McEvoy, J. P., Munetz, M., DiBarry, A. L., Bartone, P., Cather, R., Cooley, S. J., Ulrich, R. F., Carter, M. & Madonia, M. J. (1988). Dose of Fluphenazine, familial expressed emotion, and outcome in schizophrenia. Results of a two-year controlled study. *Archives of General Psychiatry* **45**, 797–805.

Holloway, F. (1988). Prescribing for the long-term mentally ill. A study of treatment practices. *British Journal of Psychiatry* **152**, 511–515.

Huang, C. C., Gerhardstein, R. P., Kim, D. Y. & Hollister, L. (1987). Treatment-resistant schizoprenia: controlled study of moderate- and high-dose thiothixene. *International Clinical Psychopharmacology* **2**, 69–75.

Johnson, D. A. W., Ludlow, J. M., Street, K. & Taylor, R. D. W. (1987). Double blind comparison of half-dose and standard-dose Flupenthixol Decanoate in the maintenance treatment of stabilised out-patients with schizophrenia. *British Journal of Psychiatry* **151**, 634–638.

Kane, J. M., Rifkin, A., Woerner, M., Reardon, G., Sarantakos, S., Schiebel, D. & Ramos-Lorenzi, J. (1983). Low-dose neuroleptic treatment of outpatient schizophrenics. *Archives of General Psychiatry* **40**, 893–896.

Kane, J. M., Woerner, M. & Sarantakos, S. (1986). Depot neuroleptics: a comparative review of standard, intermediate, and low-dose regimens. *Journal of Clinical Psychiatry* **47** (5, Suppl), 30–33.

Karon, B. P. (1989). Psychotherapy versus medication for schizophrenia: empirical comparisons. In *The Limits of Biological Treatments for Psychological Distress. Comparisons with Psychotherapy and Placebo* (ed. S. Fisher and R. P. Greenberg), pp. 105–150. Lawrence Erlbaum Associates Publishers: Hillsdale, NJ.

Louis, T., Fineberg, H. V. & Mosteller, F. (1985). Findings for public health from meta-analyses. *Annals Review of Public Health* **6**, 1–20.

McClelland, H. A., Farquharson, R. G., Leyburn, P., Furness, J. A. & Schiff, A. A. (1976). Very high dose fluphenazine decanoate. A controlled trial in chronic schizophrenia. *Archives of General Psychiatry* **33**, 1435–1439.

McCreadie, R. G., Flanagan, W. L., McKnight, J. & Jorgensen, A. (1979). High dose flupenthixol decanoate in chronic schizophrenia. *British Journal of Psychiatry* **135**, 175–179.

Magno Zito, J., Craig, T., Wanderling, J. & Siegel, C. (1987). Pharmaco-epidemiology in 136 hospitalized schizophrenic patients. *American Journal of Psychiatry* **144**, 778–782.

Marder, S. R., Van Putten, T., Mintz, J., Lebell, M., McKenzie, J. & May, P. R. A. (1987). Low- and conventional-dose maintenance therapy with Fluphenazine Decanoate. *Archives of General Psychiatry* **44**, 518–521.

Morgenstern, H., Glazer, W. M., Niedzwiecki, D. & Nourjah, P. (1987). The impact of neuroleptic medication on tardive dyskinesia: a meta-analysis of published studies. *American Journal of Public Health* **77**, 717–724.

Nishikawa, T., Tsuda, A., Tanaka, M., Hoaki, I., Koga, I. & Uchida, Y. (1984). Prophylactic effect of neuroleptics in symptom-free schizophrenics: a comparative dose-response study of haloperidol and propericiazine. *Psychopharmacology* **82**, 153–156.

Nishikawa, T., Tsuda, A., Tanaka, M., Koga, I. & Uchida, Y. (1985). Prophylactic effects of neuroleptics in symptom-free schizophrenics: roles of dopaminergic and noradrenergic blockers. *Biological Psychiatry* **20**, 1161–1166.

Oldham, A. J. & Bott, M. (1971). The management of excitation in a general hospital psychiatric ward by high dosage haloperidol. *Acta Psychiatrica Scandinavica* **47**, 369–376.

Overall, J. E. & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports* **10**, 799–812.

Prien, R. F. & Cole, J. O. (1968). High dose chlorpromazine therapy in chronic schizophrenia. *Archives of General Psychiatry* **18**, 482–495.

Prien, R. F., Levine, J. & Cole, J. O. (1969). High dose trifluoperazine therapy in chronic schizophrenia. *American Journal of Psychiatry* **126**, 305–313.

Quitkin, F., Rifkin, A. & Klein, D. F. (1975). Very high dosage vs standard dosage fluphenazine in schizophrenia. A double-blind study of nonchronic treatment-refractory patients. *Archives of General Psychiatry* **32**, 1276–1281.

Reardon, G. T., Rifkin, A., Schwartz, A., Myerson, A. & Siris, S. (1989). Changing patterns of neuroleptic dosage over a decade. *American Journal of Psychiatry* **146**, 726–728.

Reitman, D., Sacks, H. S. & Chalmers, T. C. (1987). Technical quality assessment of randomized control trials (RCTs). *Controlled Clinical Trials* **8**, 232.

Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A. & Chalmers, T. C. (1987). Meta-analysis of randomized controlled trials. *New England Journal of Medicine* **316**, 450–455.

Sangiovanni, F., Taylor, M. A., Abrams, R. & Gaztanaga, P. (1973). Rapid control of psychotic excitement states with intramuscular haloperidol. *American Journal of Psychiatry* **130**, 1155–1156.

Schneider, L. S., Pollock, V. E. & Lyness, S. A. (1990). A metaanalysis of controlled trials of neuroleptic treatment in dementia. *Journal of the American Geriatric Society* **38**, 553–563.

Simpson, G. M., Amin, M., Kunz-Bartholini, E., Watts, T. P. S. & Laska, E. (1968). Problems in the evaluation of the optimal dose of a phenothiazine (butaperazine). *Diseases of the Nervous System* **29**, 478–484.

Venulet, J., Blattner, R., von Bulow, J. & Berneker, G. C. (1982). How good are articles on adverse drug reactions? *British Medical Journal* **284**, 1792–1794.

Wiles, D., Franklin, M., Dencker, S. J., Johansson, R., Lundin, L. & Malm, U. (1980). Plasma Fluphenazine and prolactin levels in schizophrenic patients during treatment with low and high doses of fluphenazine enanthate. *Psychopharmacology* **71**, 131–136.

Williams, J. R., Solecki, R. T. & Putkammer, S. (1969). Effects of single, combined and non-drug treatments on chronic mental patients. (A preliminary study.) *Diseases of the Nervous System* **30**, 696–701.