Brit. J. Psychiat. (1982), **140**, 64–68

Side Effects and Placebo Amplification

RICHARD THOMSON

Summary: A sample of placebo-controlled tricyclic antidepressant studies was examined retrospectively to determine whether there was any difference in the relative efficacy of the tricyclic when it was compared against an inert placebo as against an active (atropine) placebo with anticholinergic side effects. Fewer studies showed a significant difference between atropine placebo and drug than between inert placebo and drug. The possibilities that atropine has a specific antidepressant effect, or that side effects amplify placebo responses are considered.

Reviews by Shapiro (1971) and Shapiro and Morris (1978) have clearly demonstrated the importance of the placebo effect in drug therapy, yet of the three broad areas which might be expected to modify the placebo response, only those pertaining to patient personality variables, and the characteristics of the treatment (doctors, setting, etc.) have been widely investigated. The third, and equally important factor, the intervention strategy itself, has been conspicuously ignored, although it may somehow suggest increased efficacy to the patient, and hence a greater, expectancy induced, effect.

The side effects produced by most psychotropic drugs may play a part in producing or amplifying placebo response, especially in double blind evaluations where both subjects and investigators are to remain ignorant of who is actually receiving the experimental drug. Several workers have commented upon the fact that side effects may reveal to the experimenters who is on what, thus allowing bias to interfere with clinical rating of improvement (e.g. Nash, 1962; Stallone *et al*, 1975; Letemendia and Harris, 1959).

Equally importantly, however, Marini *et al* (1976) note that subjects themselves may break the double blind because of this factor. They found that subjects could determine, through side effects, whether they were on lithium or not.

If this is a general phenomenon, it has important implications for the clinical evaluation of drugs. If subjects in a drug trial are warned to expect certain side effects, and then do not experience them, they may conclude they have been given a placebo, or that the drug is ineffective for them. Conversely, those that do experience side effects may feel that the drug is working and will not think they have received inert or ineffective medication. Thus not only may side effects tell the staff who *should* be getting better, they may strengthen the power of the placebo effect in the patient. If swallowing an inert placebo can give significant improvement to about 30 per cent of those suffering from depression (Morris and Beck, 1974) then it seems intuitively logical that swallowing a tablet which can be felt and/or seen to be doing something by the patient would be even more likely to increase his expectations of therapeutic efficacy.

There has been little direct testing of this hypothesis. However, Lyerly *et al* (1964) found that the effects of amphetamine and chloral hydrate were dependent on the instructional set given to the subjects, but that these instructions affected mood only when the drug was present and not when an inert placebo was taken. They suggested that the drugs gave rise to non-specific internal stimuli which were then interpreted in relation to the expectancies produced by the instructions. Penick and Hinkle (1964), Penick and Fisher (1965) and Dinnerstein and Halm (1970) note similar findings. In a more direct, but methodologically unsound, study, Brune *et al* (1962) found that mild side effects enhanced drug success whilst zero or severe side effects lessened success. Thomson and Chute (in preparation) found that an active placebo (nicotinic acid) induced significantly greater depression in normal volunteers than did an inert (lactose) placebo. Kast and Loesch (1961) used atropine sulphate to produce a dry mouth which was then either positively or negatively interpreted to patients. In keeping with expectations, a positive interpretation led to increased clinical improvement whereas a negative interpretation decreased therapeutic efficacy. Contradictory results are provided by Lipman *et al* (1966).

Dinnerstein and Halm (1970) talk of drugs as being possible placebo amplifiers, i.e. non specific physio-

https://doi.org/10.1192/bjp.140.1.64 Published online by Cambridge University Press

logical effects produced by the pharmacological action of the drugs may act as magnifiers of any placebo effect by convincing subjects that the drug has potency. One test of this hypothesis is to take a well studied drug class (in this case the tricyclic antidepressants) and examine, by literature review, whether there is any difference in the relative efficacy of the experimental drug when it is compared against an inert placebo as opposed to when it is compared against an active (atropine) placebo, which mimics closely the anticholinergic side effects of the tricyclics.

My experimental hypothesis then was that significantly less studies would show a significant difference between atropine placebo and drug, than would be evident in those studies where an inert placebo was used.

Method

Sample

Morris and Beck (1974) have reviewed all double blind placebo controlled studies of tricyclic antidepressants carried out between 1958 and 1972. They list 93 group comparisons of which 61 showed tricyclic antidepressants to be significantly ($P < .05$) more effective than a placebo. These 93 group comparisons were taken as the sample to be studied and were divided into three categories:

(1) Those where an inert or presumably inert (i.e. not stated) placebo is used (n = 68).

(2) Those where an atropine placebo is used (n = 7).

(3) Those where the comparison was an invalid one for methodological reasons (n = 6); where no comparison was possible due to non-presentation of data (n = 9); where the article was in a foreign language (n = 3); or where a non atropine (barbiturate and mannitol) active placebo was used (n = 2).

The studies in category three were then dropped from the sample leaving a total of 75 group comparisons for analysis.

Design

All 75 studies were analysed by the author using set criteria for evaluation, and then allocated to either a 'drug superior' or an 'equivocal' category. In only one comparison was the placebo shown to be superior and this was included in the equivocal category. The difficulty in comparing results across all manner of methodologies, samples, time spans and evaluations, necessitated the drawing up of strict criteria for evaluation. The criteria by which a decision was made were:

(1) In all studies the drug was only considered superior to the placebo if it could be shown to achieve statistically significant improvements ($P < .05$) over the placebo. Where non-parametric methods were used, evaluation was via a 2×2 chi square table. i.e.

	Drug	Placebo
all improved		
all not improved		

(2) Where more than one treatment was compared against the placebo only the tricyclic was included in the analysis. Where more than one tricyclic was included in the analysis, then each one was evaluated against the placebo as a separate analysis.

(3) Where more than one clinical population was used then each sample was evaluated separately. However, where insufficient data was provided to allow an analysis of each sub-population then only the main drug effect across all samples was used.

N.B. Separate populations were just that (e.g. two different hospitals, outpatients *vs* inpatients, etc.). This did not apply if a population was subdivided retrospectively into different types of subject (e.g. high income *vs* low income, mild *vs* severe, etc.). Retrospective breakdown was analysed through the main effect only.

(4) Where no raw data was provided but the authors stated that there was a statistically significant difference (and quoted the test used) then the authors' statement was presumed to be accurate.

(5) All evaluative comparisons used were the last formal comparison of the study (follow-up comparisons were not used). In a few studies patients had been discharged because of significant improvement before the final formal assessment. In these cases the last comparison before any patients were discharged was used.

N.B. Discharges did not include dropouts.

(6) In some studies the following design was used: (a) Tricyclic+additive *vs* (b) additive alone *vs* (c) placebo.

These studies were included as valid comparisons only if the a *vs* b comparison was not significantly different. If this requirement was met then an a *vs* c comparison was carried out.

(7) Most studies used more than one measure in their evaluations and some studies used very many. Where results on these measures conflicted then a value judgement had to be made as to which measures to give the most weighting to. In general, psychometric measures were weighted

https://doi.org/10.1192/bjp.140.1.64 Published online by Cambridge University Press

most highly. However this was only a general guide as, for example, more weight would have been put on a patient self-report than a WAIS block design. Ultimately the decision in most of these studies became a matter of professional judgement. Where one or two significant results were scattered amongst many non-significant results the tendency was to rate the study as equivocal. Points which were borne in mind, but for which no formal criteria could be provided were:

(a) how methodologically sound was the study?

(b) how equivalent were the pre-treatment groups?

(c) were the statistically significant changes in level of depression, or some less relevant area (e.g. in sedative side effects)?

In order to provide a check on my use of these criteria a second person also evaluated the seven comparisons using an atropine placebo, plus three studies using an inert placebo. The second evaluator was not blind to the study hypothesis but he was blind to the author's conclusion for each of the studies. Although not providing a formal check against experimenter bias it did allow a check of the reliability of the author's criteria and the conclusions drawn from them. In fact the second evaluation was exactly the same as the first.

Results

In total, 59 per cent of the studies showed the drug to be superior to placebo, which compares favourably with Morris and Beck's (1974) results of 65 per cent. More importantly there was a significant difference ($P < .05$) where atropine placebo was used, the advantage of tricyclic over placebo being significantly less than where the placebo was inert.

To allow others to make their own comparisons the atropine studies designated equivocal (on the criteria above) were: Friedman *et al* (1966); Hollister *et al* (1964); Hussain (1970); Uhlenhuth and Park (1963) and Wilson *et al* (1963). (The Hollister *et al*, 1964, study had two comparison groups). The atropine study which demonstrated a significant superiority for the drug over the placebo was Daneman (1964).

TABLE

Outcome of double blind trial of tricyclic therapy in 75 studies

Placebo	Equivocal	Drug superior
Atropine	6	1
Inert	25	43

$\chi^2 = 4.42$ (Yates correction).
$P < .05$.

Discussion

These results provide strong empirical support for the experimental hypothesis and suggest that side effects may act as a placebo amplifier, contributing some of the variance commonly attributed to the specific chemical effect of the drug. Several explanations of this can be considered.

The anticholinergic effects of atropine may be antidepressant. Several workers (e.g. Janowsky *et al*, 1972; Snyder and Yamamura, 1977) have suggested that central cholinergic factors may play a role in depression and that anticholinergic drugs may improve a depressive condition. Van der Kolk *et al* (1978) note that such hypotheses are not entirely consistent with clinical observations of anticholinergic activity in depression, but such an explanation cannot be ruled out.

A second explanation is that those authors careful enough to incorporate an atropine placebo in their study may be more rigorous workers generally. Smith *et al* (1969) showed in a review that the apparent efficacy of an antidepressant was inversely related to the methodological strength of its study. Perhaps, then, the authors of the atropine studies were more conservative and stringent in their experimental work and hence less likely to get a significant difference between drug and placebo.

Evaluator bias cannot be completely ruled out but several factors argue against it. The author's overall result was very similar to that of Morris and Beck (1974), but slightly smaller, which would tend to make a significant difference between studies with and without atropine harder to obtain rather than easier. The exactly similar but independent results of the second evaluator would seem to indicate that the criteria were being followed in an appropriate manner.

Whether the criteria themselves are appropriate is a matter for the reader to decide. Clearly a large number of possible criteria might have been used, provided they were the same for all studies. The difficulty of designing criteria for such a large number of very disparate studies is great. However, with the exception of one study (Hollister *et al*, 1964) the results of the atropine comparisons are relatively clear cut. Within the Hollister *et al* (1964) study the amitriptyline *vs* atropine comparison, while equivocal with regard to the main effect, did show some significant differences

https://doi.org/10.1192/bjp.140.1.64 Published online by Cambridge University Press

retrospectively with regard to those subjects diagnosed severely depressed. If the criteria had allowed it then there would have been a second atropine comparison where the tricyclic was significantly superior. By the same token, there would have resulted two further comparisons that were equivocal (the mild and moderate samples) and the other 73 comparisons would have had to be re-examined in the same light with probable further changes occurring. In short, while other criteria could be argued, the author believes that the ones used are methodologically sound enough to illustrate the point made.

The 'placebo amplification' phenomenon, where the physiological effects (such as side effects) produced by drugs convince subjects already influenced favourably by the pill-taking ritual that the drug has potency, seems to account for a considerable proportion of the supposed clinical value of antidepressants. This is not to say, of course, that there is no specific effect due to the tricyclic. Indeed there is evidence from several studies (e.g. Hollister *et al*, 1964) that the use of a heterogeneous group of depressed subjects obscures drug effects whereas a comparison using only severely depressed subjects may show a highly significant difference between placebo and drug.

The present results support Dinnerstein and Halm's (1970) placebo amplification hypothesis. To strengthen this support would require the elimination of the other explanations outlined above, something which could be done experimentally using a double blind, tricyclic *vs* inert placebo *vs* atropine placebo study. To the best of the author's knowledge such a study has not yet been carried out.

The paucity of study in this area should be a matter of some concern to anyone interested in the methodology of drug evaluation in particular, and therapeutic intervention in general. Frequently now, workers compare new treatments against old without the inclusion of a placebo control. They argue that the old, or standard, treatment has already been compared against a placebo control and therefore the only need for comparison is to establish whether the new drug is better than the old. However, the old may contain a large component of placebo amplification, and therefore simply give an active placebo result for comparison with the new.

The present study has considered only tricyclics, but similar results may be possible with many other drugs. Therefore wherever possible drug trials should use a placebo control which mimics as closely as possible all non-specific effects.

Acknowledgement

The author wishes to thank Dr David Marks for his advice, and assistance as a second evaluator of the sample.

References

BRUNE, G., MORPUGO, C., BIELKUS, A., KOBAYASHI, T., TOWLENTES, T. & HIMWICH, H. (1962) Relevance of drug induced extrapyramidal reactions to behavioral changes during neuroleptic treatment. Treatment with trifluoperazine singly and in combination with trihexyphenidyl. *Comprehensive Psychiatry*, **3**, 227–34.

DANEMAN, E. (1964) Imipramine in office management of depressive reactions (a double blind clinical study). *Diseases of the Nervous System*, **22**, 213–17.

DINNERSTEIN, A. & HALM, J. (1970) Modification of placebo effects by means of drugs: Effects of aspirin and placebos on self-rated moods. *Journal of Abnormal Psychology*, **75**, 308–14.

FRIEDMAN, A., GRANICK, S., COHEN, H. & COWITZ, B. (1966) Imipramine (Tofranil) vs placebo in hospitalized psychotic depressives. (A comparison of patients self ratings, psychiatrists ratings and psychological test scores). *Journal of Psychiatric Research*, **4**, 13–36.

HOLLISTER, L., OVERALL, J., JOHNSON, M., PENNINGTON, V., KATZ, G. & SHELTON, J. (1964) Controlled comparison of amitriptyline, imipramine, and placebo in hospitalized depressed patients. *Journal of Nervous and Mental Disease*, **139**, 370–5.

HUSSAIN, Z. (1970) Drugs in Depressive Illness. *British Medical Journal*, *ii*, 482.

JANOWSKY, D., EL-YOUSEF, M., DAVIS, J. & SEKERKE, H. (1972) A cholinergic-adrenergic hypothesis of mania and depression. *Lancet*, *ii*, 632–5.

KAST, E. & LOESCH, J. (1961) Influence of the doctor patient relationship on drug action. *Illinois Medical Journal*, **119**, 390–3.

LETEMENDIA, F. & HARRIS, A. (1959) The influence of side effects on the reporting of symptoms. *Psychopharmacologia*, **1**, 39–47.

LIPMAN, R., PARK, L., RICKELS, K. & CHASE, C. (1966) Paradoxical influence of a therapeutic side-effect interpretation. *Archives of General Psychiatry*, **15**, 462–74.

LYERLY, S., ROSS, S., KRUGMAN, A. & CLYDE, D. (1964) Drugs and placebos: The effects of instructions upon performance and mood under amphetamine sulphate and chloral hydrate. *Journal of Abnormal and Social Psychology*, **68**, 321–7.

MARINI, J., SHEARD, M., BRIDGES, C. & WAGNER, E. (1976) An evaluation of the double blind design in a study comparing lithium carbonate with placebo. *Acta Psychiatrica Scandinavica*, **53**, 343–54.

MORRIS, J. & BECK, A. (1974) The efficacy of antidepressant drugs. A review of research (1958–1972). *Archives of General Psychiatry*, **30**, 667–74.

NASH, M. (1962) The double blind procedure: Rationale and empirical evaluation. *Journal of Nervous and Mental Disease*, **134**, 34–7.

PENICK, S. & FISHER, S. (1965) Drug-set interaction: psychological and physiological effects of epinephrine under differential expectations. *Psychosomatic Medicine*, **27**, 177–82.

—— & HINKLE, L. (1964) The effect of expectation on response to phenmetrazine. *Psychosomatic Medicine*, **26**, 369–73.

https://doi.org/10.1192/bjp.140.1.64 Published online by Cambridge University Press

SHAPIRO, A. (1971) Placebo effects in Medicine, Psychotherapy and Psychoanalysis. In *Handbook of Psychotherapy and Behavior Change*, (eds. A. Bergin and S. Garfield). New York: Wiley.

—— & MORRIS, L. (1978) The placebo effect in medical and psychological therapies. In *Handbook of Psychotherapy and Behavior Change*, (eds. S. Garfield and A. Bergin). New York: Wiley.

SMITH, A., TRAGANZA, E. & HARRISON, G. (1969) Studies on the effectiveness of antidepressant drugs. *Psychopharmacology Bulletin*, Supplement, March, 1–53.

SNYDER, S. & YAMAMURA, H. (1977) Antidepressants and the muscarinic acetylcholine receptor. *Archives of General Psychiatry*, **2,** 236–9.

STALLONE, F., MENDLEWICZ, J. & FIEVE, R. (1975) Double blind procedure: an assessment in a study of lithium prophylaxis. *Psychological Medicine*, **5,** 78–82.

UHLENHUTH, E. & PARK, L. (1963) The influence of medication (imipramine) and doctor in relieving depressed psychoneurotic outpatients. *Journal of Psychiatric Research*, **2,** 101–22.

VAN DER KOLK, B., SHADER, R. & GREENBLATT, D. (1978) Autonomic effects of psychotropic drugs. In *Psychopharmacology: A Generation of Progress*, (eds. M. Lipton, A. Di Mascio and K. Killam). New York: Raven Press.

WILSON, I., VERNON, J. & SANDILER, M. JR. (1963) A controlled study of treatment of depression. *Journal of Neuropsychiatry*, **4,** 331–7.

Richard J. Thomson, M.A., Dip.Clin.Psych., *Lecturer in Clinical Psychology, Department of Psychological Medicine, University of Otago Medical School, Dunedin, New Zealand*

(*Received 27 May 1981*)

https://doi.org/10.1192/bjp.140.1.64 Published online by Cambridge University Press