International Journal of Law and Psychiatry 31 (2008) 453–462

Contents lists available at ScienceDirect

International Journal of Law and Psychiatry

The antidepressant debate and the balanced placebo trial design: An ethical analysis

Duff R. Waring *

York University, Toronto, Canada

ARTICLE INFO

Keywords:
Active placebos
Antidepressant debate
Deception
Drug effects
Expectancy
Experimental subordination
Placebo effects
Trial design

ABSTRACT

There is ongoing debate about whether randomized, placebo-controlled trials under a double-blind have reliably established the pharmacological efficacy of antidepressants. Numerous meta-analyses of antidepressant efficacy trials, e.g., Kirsch et al. [Kirsch, I., Moore, T. J., Scoboria, A., & Nicholls, S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. food and drug administration. *Prevention and Treatment, 5*, Article 23. (Retrieved July 19, 2007 from http://journals.apa.org/prevention/volume5)], have shown a modest drug-placebo difference but methodological problems with standard trial design preclude a definitive conclusion that this difference results from specific biological effects of antidepressants or the nonspecific factors that have not been adequately excluded. Standard trial design assumes the additivity thesis of pharmacological efficacy, being the assumption that the specific or "true" magnitude of the pharmacological effect is limited to the difference between the drug and placebo responses in a standard trial. If the drug effects are as small as these meta-analyses suggest, then their clinical effectiveness is questionable. If the drug effects are actually larger but masked by placebo effects, then the additivity thesis is not valid and we risk false negative results with standard trial design. Kirsch et al. propose an alternative, four arm balanced placebo trial design (BPTD) that can accurately test the additivity thesis. The BPTD uses antidepressants, active placebos and the intentional deception of research subjects. My focal question is whether the BPTD is ethically defensible. I will explore two objections that can be raised against it: 1) lying to BPTD research subjects violates their autonomy and exploits their illness and 2) the BPTD may not enable us to test the additivity thesis with accuracy, i.e., it may contribute to the masking of drug effects that it aims to avoid. I argue that these objections support the conclusion that the BPTD is ethically indefensible.

1. Introduction

A basic tenet of biological psychiatry is that antidepressants are efficacious in two ways: they exert both pharmacological efficacy and clinical effectiveness. The results of some placebo-controlled, randomized drug trials under a double-blind (hereinafter standard trials) supposedly confirm their pharmacological efficacy, i.e., that antidepressants exert specific and measurable effects on research subjects (Zimmerman, Mattia, & Posternak, 2002: 469–473). The principal aim of standard trials is to determine whether there is a difference between specific drug and nonspecific placebo effects. The results of subsequent therapeutic practice may confirm the clinical effectiveness of antidepressants, i.e., that they are beneficial treatments for depressed patients (Parker, 2001: 95–96). What are promoted as major advances in psychopharmacology have ushered in the antidepressant era (Healy, 1997). But there is heated debate about the extent to which antidepressants are efficacious in either of these two ways (Fisher & Greenberg, 1997; Healy, 1997; Moncrieff, 2002: 193–194).

* Room 625 Atkinson Bldg., York University, 4700 Keele St., Toronto, ON, Canada M3J 1P3. Tel.: +1 416 736 2100x33522.
E-mail address: dwaring@yorku.ca.

0160-2527/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijlp.2008.09.001

The work of Kirsch, Scoboria, Moore, and Nicholls (2002) has further enkindled this debate. They challenged both the clinical effectiveness of antidepressants and the standard trial design that purports to confirm their efficacy. Kirsch et al. analyzed all the efficacy data submitted to the U.S. Food and Drug Administration for approval of the six most widely used antidepressants that were licensed for use between 1987 and 1999.[1] This data was derived from 47 short-term efficacy trials using standard trial design (i.e., a verum or drug arm and a placebo control arm, both under a double-blind). All but one of the trials involved patients diagnosed as "moderately to severely depressed." On their meta-analysis of these trials, Kirsch et al. found that approximately 82% of the drug response was duplicated in the placebo control arms. This indicated that about 18% of the drug response is due to the pharmacological effects of the antidepressants. The rest is supposedly placebo effect. The mean difference between drug and placebo response was roughly two points on the Hamilton Depression Scale. There was no difference in improvement at either the highest or lowest doses of medication (Kirsch et al., 2002: 1, 3–7).

Kirsch et al. suggested three possible explanations for the small drug-placebo difference. First, the difference could be an enhanced placebo effect in the verum arm due to the breaking of the blind. Second, the difference could be a "very small" but statistically significant pharmacological effect. Third, the difference might actually be a larger pharmacological effect that is masked by placebo effects. If the difference is a small pharmacological effect, then the clinical effectiveness of antidepressants is questionable. Indeed, we may lack an adequate justification for exposing patients to their side effects. If the pharmacological effect is actually larger but masked by placebo effects, then we need an alternative trial design that could show this. This design would also test the additivity thesis of pharmacological efficacy, being the assumption that the specific or "true" magnitude of the pharmacological effect is limited to the difference between the drug and placebo responses in a standard trial (Kirsch, 2000: 733; Kirsch et al., 2002: 8).

Kirsch et al. suggested that we test this thesis with a four arm "balanced placebo trial design" (hereinafter BPTD) using antidepressants, active placebo controls and the intentional deception of research subjects (Kirsch et al., 2002: 8; Marlatt & Rosenhow, 1980). As I understand it, the goal would be to test the six antidepressants reviewed by Kirsch et al. with the BPTD and then compare the drug and placebo responses across the standard and alternative trial designs. Antidepressants could thus be represented as drugs that have been shown to elicit greater responses than placebo controls and that have been designed, licensed and approved for use as active treatments. While some subjects in the BPTD will know they are taking them, antidepressants will also be used to make some subjects think they are receiving active placebo controls. Active placebo controls could be represented as compounds that have been designed to merely simulate the side effects of antidepressants. While some subjects in the BPTD will know they are taking them, active placebo controls will also be used to make some subjects think they are receiving antidepressants. My focal question is whether the BPTD is ethically defensible. I will explore two objections that can be raised against it: 1) lying to BPTD research subjects violates their autonomy and exploits their illness and 2) the BPTD may not enable us to test the additivity thesis with accuracy because it can contribute to the masking of drug effects that it aims to avoid. I argue that these objections support the conclusion that the BPTD is not ethically defensible.

2. Alternative trial design

Let us consider the three possible explanations for the small drug-placebo difference found by Kirsch et al. One possibility is that it is an enhanced placebo effect in the verum arm due to the breaking of the blind. Knowing that one has been randomized to the verum arm will likely enhance the expectation or belief that one will improve. Knowledge of assignment to the placebo control arm ought to decrease this belief (Fisher & Greenberg, 1993: 345–350; Kirsch, 2000: 733; Kirsch & Moore, 2002: 8; Kirsch et al., 2002: 7). Studies indicate that the ability of researchers and subjects to deduce whether they have been randomized to provide or receive drugs or placebos exceeds levels of chance (Fisher & Greenberg, 1993: 345–350, 1997: 23–24,133–135; Rabkin et al., 1986: 75–86). We also know that post-trial checks on whether the blind was maintained are rare (Fergusson, Glass, Waring, & Shapiro, 2004: 432–434).

The second possibility is that this small difference reflects the additivity thesis. If we parse the additivity thesis we see that it contains at least three assumptions. First, that there is an equal level of placebo response in the verum and placebo control arms of a standard trial. Second, that the magnitude of a drug's "true" pharmacological effect is limited to the difference between the responses in the verum and placebo control arms and third, that we can gauge the true pharmacological effect by subtracting the total response in the placebo control arm from the total response in the verum arm. Standard trials are based on this thesis. The drug demonstrates efficacy only if the response to it is significantly greater than the placebo response (Kirsch, 2000: 733–734; Kirsch et al., 2002: 8–9).

While additivity is the foundational thesis of standard trials, its applicability to antidepressants has not been tested. Kirsch et al. argued that it should be (Kirsch, 2000: 733–734; Kirsch et al., 2002: 9). Research indicates that some drug effects are not additive. Alcohol and caffeine have been shown to have some effects that are additive and some effects that are non-additive (Hull & Bond, 1986; Kirsch & Rosadino, 1993). Additive effects are those for which the differences between the verum and placebo responses under a double-blind are preserved in other settings where people either knowingly take the same drug and expect an effect or where people unknowingly take the same drug and do not expect an effect. Additive drug effects are not influenced by expectancy, i.e., by the subjects' expectation or belief that they will improve as a result of what they are, or think they are, taking. Thus Kirsch and Rosadino (1993) found that caffeine-induced alertness is additive under a double-blind and in a trial arm in which subjects

[1] The breakdown by drug and efficacy trial was: fluoxetine (5), paroxetine (16), sertraline (7), venaflaxine (6), nefadozone (8) and citalopram (5).