Psychological Medicine (2010), **40**, 181–186. © Cambridge University Press 2009
doi:10.1017/S0033291709006102

EDITORIAL

Is the antidepressive effect of second-generation antidepressants a myth?

P. Bech*

Psychiatric Research Unit, Frederiksborg General Hospital, Hillerød, Denmark

Two recent meta-analyses on second-generation antidepressants *versus* placebo in mild to moderate forms of major depression, based on data on all randomized clinical trials using the Hamilton Depression Scale (HAMD) submitted to FDA, have shown an effect size of ~0.30 in favour of antidepressants in the acute therapy of major depression. The clinical significance of an effect size at this level was found to be so poor that these meta-analyses have subscribed to the myth of an exclusively placebo-like effect of second-generation antidepressants. A re-allocation of HAMD items focusing on those items measuring severity of clinical depression, the $HAMD_6$, has identified effect sizes of ⩾0.40 for second-generation antidepressants in placebo-controlled trials for which even a dose–response relationship can be demonstrated. In the relapse-prevention phase during continuation therapy of patients with major depression, the advantage of second-generation antidepressants over placebo was as significant as in the acute therapy phase. To explore a myth is not to deny the facts but rather to re-allocate them.

Received 16 September 2008; Revised 7 April 2009; Accepted 29 April 2009; First published online 3 June 2009

Key words: Antidepressive effect, dose–response relationship, effect size, $HAMD_6$, relapse prevention.

Introduction

In two recent meta-analyses of second-generation antidepressants *versus* placebo in mild to moderate forms of major depression, the data on all trials submitted to the US Food and Drug Administration (FDA) were used (Kirsch *et al*. 2008; Turner *et al*. 2008). In these meta-analyses the full Hamilton Depression Scale (HAMD; Hamilton, 1967) was employed as outcome measure and effect-size statistics were used to demonstrate the clinical response. Both meta-analyses concluded that the advantage of second-generation antidepressants over placebo in acute 6–8 weeks' therapy of patients with major depressive episodes in terms of effect-size statistics was minimal, thereby maintaining the myth of a merely placebo-like effect of antidepressant medication in mild to moderate depression.

First-generation antidepressant medication was introduced by Kuhn when, 50 years ago, he demonstrated that over a treatment period of 4 weeks imipramine significantly improved the depressive symptoms of patients hospitalized for depressive illness (Kuhn, 1958). In the 1960s, randomized, placebo-controlled clinical trials in hospitalized patients confirmed by use of the full HAMD that imipramine was superior to placebo (Beck, 1973).

When developing first-generation cognitive psychotherapy of depression, Beck and his group (Rush *et al*. 1977) used imipramine as an active comparator to assess the antidepressive effect of cognitive psychotherapy in depressed patients because placebo cognitive therapy was too difficult to perform. The clinical evaluations in the Rush *et al*. (1977) study were obviously not blind regarding treatment assignment. As outcome scale the clinicians used the full HAMD. The baseline score on the HAMD was ~18, corresponding to mild depression. The results showed that both cognitive therapy and imipramine were effective, although no placebo–imipramine arm was included. One of the first placebo-controlled imipramine trials in a family doctor setting failed, however, to discriminate between active drug and placebo (Porter, 1970). On the other hand, a consequence of the Rush *et al*. (1977) study was the gradual acceptance of imipramine for use even in patients with mild depression.

With the advent in the 1980s of second-generation antidepressants [specific serotonin reuptake inhibitors (SSRIs)], the use of this type of antidepressive medication in the milder forms of depression became popular due to the very favourable SSRI side-effect profile compared to that of imipramine. In the 1970s the FDA requirement for achieving approval of an experimental antidepressant was a substantial amount of evidence based on randomized, placebo-controlled

* Address for correspondence: P. Bech, MDSci., Professor of Psychiatry and Applied Psychometrics, Psychiatric Research Unit, Frederiksborg General Hospital, 3400 Hillerød, Denmark.
(Email: pebe@noh.regionh.dk)

trials including fixed graded dose of the experimental drug (Leber, 1996). In such trials the side-effects of the experimental drug were to be systematically evaluated as well. In the placebo-controlled clinical trials with SSRIs, imipramine was often used as an active comparator, and it was shown that the advantage of SSRIs was their favourable side-effect profile compared to imipramine rather than an advantage regarding their antidepressive effect (Øhrberg *et al*. 1992). The imipramine side-effects such as sedation and increased appetite might be considered as desired effects in the initial acute therapy of moderately to severely depressed patients but are generally regarded as serious side-effects in long-term therapy (e.g. Mayer, 1975).

In the mild to moderate degree of depressive illness, the use of placebo is still important for regulatory issues, and the Kirsch *et al*. (2008) and Turner *et al*. (2008) meta-analyses using the full HAMD as outcome measure have certainly subscribed to the myth of the placebo-like effect of SSRIs in patients with mild to moderate depression.

To explore a myth is, as stated by Ryle (1949), not to deny the facts but rather to re-allocate them. In the following, the facts emerging from the Kirsch *et al*. (2008) and Turner *et al*. (2008) meta-analyses will not only be re-allocated but also itemized within the universe of items in the full HAMD, while the effect-size statistics will be maintained as relevant for the measuring of clinical response in the acute therapy of major depressive episodes.

Effect size as descriptive outcome statistics in the acute therapy of depression

Both the Kirsch *et al*. (2008) and the Turner *et al*. (2008) meta-analyses have used effect size to indicate antidepressive activity of an experimental drug *versus* placebo. In these meta-analyses effect size is defined as the difference in mean change from baseline to the respective time-points in HAMD scores between patients treated with the experimental drug and patients treated with placebo divided by the pooled standard deviation for the two groups of patients (Hedges & Olkin, 1985). According to Kirsch *et al*. (2008) the FDA seems to accept a mean drug–placebo difference in HAMD change score of ~ 2 as evidence of an antidepressive effect whereas the British National Institute for Clinical Excellence (NICE, 2004) recommends a drug–placebo difference in HAMD improvement score of 3.

When estimating the sample size for a randomized, placebo-controlled trial with duloxetine in patients with major depression, the mean drug–placebo improvement score on HAMD was assumed to be 3.25 and the pooled standard deviation 7.0, resulting in an effect size of 0.46 (Detke *et al*. 2004). However, in the placebo-controlled trial with desvenlafaxine, the latest antidepressant approved by FDA (Young & Plosker, 2008), the mean drug–placebo improvement score on HAMD was assumed to be 3.0 and the pooled standard deviation 8.0, resulting in an expected effect size of 0.38 (Boyer *et al*. 2008).

The pooled standard deviation in placebo-controlled trials of antidepressants in patients with mild to moderate major depression seems thus to be in the interval between 7 and 8 when using $HAMD_{17}$ as outcome measure. With a pooled standard deviation of 7.5, the drug–placebo improvement score on HAMD seems to have been 2.4 to obtain the average effect size of 0.32 as reported by Kirsch *et al*. (2008). This meta-analysis focused on fluvoxetine, paroxetine, nefazodone and venlafaxine. In the meta-analysis of Turner *et al*. (2008) the average effect size was 0.31, and this analysis had identified 12 different second-generation antidepressants, including the drugs covered by Kirsch *et al*. (2008).

Turner (2008) correctly states that when Cohen (1976) introduced the effect size as descriptive statistics in clinical trials, he recommended the value of 0.50, as adopted by NICE (2004), for a clinically significant improvement. However, this was made on Cohen's own subjective intuition, not with any reference to trials of antidepressants.

In both meta-analyses (Kirsch *et al*. 2008; Turner *et al*. 2008) it seems that the unbiased effect-size formula as recommended by Hedges & Olkin (1985) has been used. In this formula (published elsewhere, see Bech, 2007) there is a correction for trials with less than 100 patients in each treatment arm. An effect size of 0.40 corresponds to an antidepressant advantage of 15–20% over placebo in the acute therapy of depression when using the response criterion of $\geqslant 50\%$ reduction on HAMD from baseline to endpoint (Bech *et al*. 2000). This 15–20% advantage was also demonstrated in the most comprehensive review of first-generation antidepressants (Smith *et al*. 1969).

Re-allocating the HAMD items to measure response in the acute therapy of depression

The full HAMD has been used in all the trials analysed by Kirsch *et al*. (2008) and almost all trials analysed by Turner *et al*. (2008). However, no attempt has been made in their analyses to identify the various versions of HAMD used in regard to, e.g. the number of items ($HAMD_{17}$, $HAMD_{21}$ or $HAMD_{24}$; Bech, 2009). Thus, the $HAMD_{24}$ was recommended by Beck and his group (Riskind *et al*. 1987) to cover the cognitive triad of depression with the extra items of worthlessness, helplessness and hopelessness. However, these three

https://doi.org/10.1017/S0033291709006102 Published online by Cambridge University Press

items are already included in the $HAMD_{17}$ as observed by Hamilton (1986) as 'guilt' covers worthlessness, 'depressed mood' covers hopelessness and 'work and interests' covers helplessness.

In the meta-analysis Turner *et al*. (2008) the calculated mean effect size for each drug is a combination of both fixed-graded experimental-drug dose trials and flexible dose trials. Therefore, a drug such as mirtazapine, in which no dose–response relationship has been demonstrated (Pinder & Zivkov, 1998), has been favoured by Turner *et al*. (2008) compared, for example, with duloxetine or venlafaxine in which a dose–response relationship has been demonstrated (Bech, 2009).

In our original study on the clinical validity of the $HAMD_{17}$ to measure severity of depressive states we identified six of the 17 items as corresponding symptomatologically to experienced psychiatrists' global perception of depressive states (Bech *et al*. 1975). In our next studies (Bech *et al*. 1981, 1984) we used item response theory models to investigate to what extent the total score of these six symptoms [depressed mood, guilt feelings, work and interests, psychomotor retardation, psychic anxiety, and general somatic symptom (tiredness)] was a sufficient statistic, i.e. a profile of the individual items is not necessary. In contrast to factor analysis which focuses on conceptual knowledge, item response theory models, e.g. the Rasch model (Rasch, 1980), focus on perceptual knowledge by considering the individual items to be placed on the graduated line on which the severity of the depressive state is listed (Bech, 2009).

Whereas the total score of these six HAMD items ($HAMD_6$) is a sufficient statistic for measurement of the pure antidepressive effect, many of the other HAMD items (insomnia, agitation, somatic anxiety, gastrointestinal symptoms, weight loss and sexual problems) might be considered to be antidepressant medication side-effects. The total score of these items is not a sufficient statistic as side-effects must be analysed individually (Lingjærde *et al*. 1987). Adverse events increase significantly with antidepressant dose (Bollini *et al*. 1999).

Table 1 shows the effect-size statistics in placebo-controlled trials of second-generation antidepressants in the acute therapy (6–8 weeks) of patients with mild to moderate major depression.

In our fluoxetine study on placebo-controlled trials in patients with DSM-III major depression we obtained an effect size of 0.38 using the $HAMD_6$ and an effect size of 0.30 on $HAMD_{17}$ (Table 1). When fluoxetine was used as an active comparator in placebo-controlled venlafaxine trials (Table 1) an effect size of 0.40 was obtained with $HAMD_6$ while the full $HAMD_{17}$ only obtained an effect size of 0.24. In both fluoxetine analyses (Bech *et al*. 2000; Entsuah *et al*. 2002) the fluoxetine dose range was 20–60 mg/day and no dose–response relationship was found.

Table 1. *Effect-size statistics in placebo-controlled trials of second-generation antidepressants in acute therapy (6–8 weeks) in patients with mild to moderate depression*

Study	$HAMD_{17}$	$HAMD_6$
Bech *et al*. (2000)		
Fluoxetine 20–60 mg	0.30	0.38
Entsuah *et al*. (2002)		
Fluoxetine 20–60 mg	0.24	0.40
Bech *et al*. (2002)		
Citalopram 20 mg	0.09	0.21
Citalopram 40 mg	0.39	0.51
Bech *et al*. (2004)		
Escitalopram 10 mg	N.A.	0.38
Escitalopram 20 mg	N.A	0.61
Bech (2001)		
Mirtazapine 15–60 mg	0.49	0.42
Bech *et al*. (2006)		
Duloxetine 60 mg	0.46	0.51
Duloxetine 120 mg	0.49	0.57

HAMD, Hamilton Depression Scale; N.A., not applicable.

Regarding citalopram, a dose–response relationship was identified (Bech *et al*. 2002), as 40 mg was found to be the optimal dose. In most placebo-controlled escitalopram trials the Montgomery–Åsberg Depression Rating Scale (Montgomery & Åsberg, 1979) has been used as outcome measure but in our analysis (Bech *et al*. 2004) the $HAMD_6$ has also been included, and we showed that 20 mg was the optimal dose with an effect size of 0.61. For duloxetine (Bech *et al*. 2006) both $HAMD_{17}$ and $HAMD_6$ showed effect sizes >0.40, even for 60 mg/day.

For SSRIs as well as for duloxetine $HAMD_6$ effect sizes have all been higher than $HAMD_{17}$ effect sizes, while for mirtazapine (which has many different actions, e.g. antihistamine, adrenergic, and serotonin receptor 2A blocking) the full $HAMD_{17}$ scale has an effect size higher than that of the $HAMD_6$ (Bech, 2001).

Relapse prevention in the continuation treatment of major depressive episode

According to the European Guidelines for Clinical Investigation (European Union, 1994) a substance is accepted as an antidepressant only if the advantage over placebo in clinical effect can be demonstrated both in the acute phase and in the continuation phase. As discussed elsewhere, electroconvulsive therapy (ECT), as conventionally administered with about 12

https://doi.org/10.1017/S0033291709006102 Published online by Cambridge University Press

sessions over 4 weeks, is effective only within 2–3 months after the last session (Lauritzen *et al*. 1996). However, when compared to placebo in this controlled clinical trial lasting up to 6 months after the last ECT session, paroxetine was found to be relapse preventive, as 65% relapsed on placebo and 12% on paroxetine ($p \leqslant 0.05$) (Lauritzen *et al*. 1996). A systematic review on relapse prevention with antidepressant drug treatment in depressive disorders showed an approximate relapse of 41% for placebo continuation therapy while 18% relapsed on antidepressants, without any difference between SSRIs and tricyclic antidepressants (Geddes *et al*. 2003). This is similar to the results of Bent-Hansen *et al*. (2003). Here the recurrence of depression in maintenance therapy with placebo was 43% but with citalopram was 13% ($p \leqslant 0.01$).

In order to cover the pure antidepressive effect as demonstrated by Kuhn (1958, 1970), the six items in HAMD covering clinical depression symptoms (depressed mood, guilt, work and interests, tiredness, anxiety, and psychomotor retardation) should be used as outcome measure. These six items ($HAMD_6$) have been found not only to reflect experienced psychiatrists' global assessment of depression (Bech *et al*. 1975) but also to fulfil the item response theory model when the total score is used as a sufficient statistic during a trial of antidepressants (Bech *et al*. 1984). These items constitute the dimension of manifest depression (Overall, 1962) and were identified by Steinmeyer & Möller (1992) using facet theory analysis of HAMD during treatment with paroxetine or amitriptyline. When testing the sensitivity of the individual items in $HAMD_{17}$ to response to paroxetine in all placebo-controlled trials with this drug, Santen *et al*. (2008) identified the $HAMD_6$ items as superior to the other items.

Many of the other HAMD items (insomnia, agitation, somatic anxiety, gastrointestinal symptoms, weight loss and sexual problems) might be considered to be side-effects of antidepressants. Together these items are not a sufficient statistic, as side-effects have to be analysed individually (Lingjærde *et al*. 1987).

Discussion

Parker (2009) has discussed the meta-analysis of Kirsch *et al*. (2008) but not that of Turner *et al*. (2008) and concludes that if we: 'wish to reject the imputation that antidepressant drugs are little better than placebo, we need first to recognize limitations of current RCT (randomized clinical trial) procedures'. However, Parker (2009) seems not to reject the necessity for randomized clinical placebo-controlled trials but rather to call for the use of better outcome scales as well as better diagnostic classification of depression.

By focusing on the $HAMD_6$ rather than the $HAMD_{17}$ we need not reject the Hamilton scale but rather use it much more appropriately. Antidepressants act, as discussed by Angst (2007), on the target dimension of depression across disorders.

When Kuhn demonstrated the antidepressive effect of imipramine he had no access to the Hamilton Depression Scale but he did confess that depressed patients 'recount absolutely nothing spontaneously about their depressive experience, and these come to light only on questioning' (Kuhn, 1970). Hamilton developed his scale (Hamilton, 1987) to enable clinicians to focus on the current symptomatology of patients, implying that the total scores on $HAMD_{17}$ should give a global impression of the burden of the illness (Bech, 2009). However, the dimension on which the antidepressive drugs act clinically is that of the $HAMD_6$ symptoms; this is also in accord with Kuhn (1958, 1970).

In a review on item response theory models and health outcome measurements in the 21st century it was shown that these models, especially the Rasch model, have a potential advantage over the classical models, e.g. factor analysis, when improving existing rating scales (Hays *et al*. 2000). Responsiveness to change during treatment is not a separate dimension of validity (Hays & Hadorn 1992). In other words, the $HAMD_6$ is really a unidimensional depression scale.

Effect-size statistics are of especial importance when comparing the responsiveness of two scales such as the $HAMD_{17}$ *versus* $HAMD_6$ in the acute therapy of depression. In continuation therapy it is the standardization of the scales that is important when defining relapse. In the meta-analysis trials a $HAMD_{17}$ score of $\geqslant 16$ and a $HAMD_6$ score of $\geqslant 9$ are often used to indicate a relapse (Bent-Hansen *et al*. 2003; Ruhé *et al*. 2005). In post-stroke prevention of depression when comparing the SSRI sertraline with placebo, we demonstrated that as early as after 6 weeks of therapy that sertraline was statistically superior to placebo on $HAMD_6$ ($p \leqslant 0.05$) but on the $HAMD_{17}$ the advantage of sertraline over placebo only appeared after ~ 20 weeks of therapy (Rasmussen *et al*. 2003).

The prevention of post-stroke depression by SSRIs is still the best example of the prevention of depression in patients who had never previously had an episode of depression but due to their physical illness belonged in a high-risk group for the development of depression (Robinson & Jorge, 2009).

Conclusion

The antidepressant effect of second-generation antidepressants does seem to be a myth when using $HAMD_{17}$ as outcome scale in the acute therapy of

https://doi.org/10.1017/S0033291709006102 Published online by Cambridge University Press

depressed patients. However, when using the core items of depression ($HAMD_6$), to measure antidepressive activity, no such myth of mere placebo activity is in operation for second-generation antidepressants, for which even a dose–response relationship can be demonstrated in the acute phase. In relapse prevention during continuation therapy the advantage of second-generation antidepressants over placebo is even more pronounced than in the acute phase treatment.

Declaration of Interest

Over the past 3 years until August 2008 Professor Bech has occasionally received funding from and been a speaker or member of advisory boards for pharmaceutical companies with an interest in the drug treatment of affective disorders (AstraZeneca, Lilly, Lundbeck A/S, Lundbeck Foundation, Organon).

References

Angst J (2007). Psychiatric diagnoses: the weak component of modern research. *World Psychiatry* **6**, 94–95.

Bech P (2001). Meta-analysis of placebo-controlled trials with mirtazapine using the core items of the Hamilton Depression Scale as evidence of a pure antidepressive effect in the short-term treatment of major depression. *International Journal of Neuropsychopharmacology* **4**, 337–345.

Bech P (2007). Dose-response relationship of Pregabalin in patients with generalized anxiety disorder. A pooled analysis of four placebo-controlled trials. *Pharmacopsychiatry* **40**, 163–168.

Bech P (2009). Fifty years with the Hamilton scales for anxiety and depression. *Psychotherapy and Psychosomatics* **78**, 202–211.

Bech P, Allerup P, Gram LF, Reisby N, Rosenberg R, Jacobsen O, Nagy A (1981). The Hamilton Depression Scale. Evaluation of objectivity using logistic models. *Acta Psychiatrica Scandinavica* **63**, 290–299.

Bech P, Allerup P, Reisby N, Gram LF (1984). Assessment of symptom change from improvement curves on the Hamilton Depression Scale in trials with antidepressants. *Psychopharmacology* **84**, 276–281.

Bech P, Cialdella P, Haugh M, Birkett MA, Hours A, Boissel JP, Tollefson GD (2000). A meta-analysis of randomised controlled trials of fluoxetine versus placebo and tricyclic antidepressants in the short-term treatment of major depression. *British Journal of Psychiatry* **176**, 421–428.

Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG (1975). Quantitative rating of depressive states. *Acta Psychiatrica Scandinavica* **51**, 161–170.

Bech P, Kajdasz DK, Porsdal V (2006). Dose-response relationship of duloxetine in placebo-controlled clinical trials in patients with major depressive disorder. *Psychopharmacology* **188**, 273–280.

Bech P, Tanghøj P, Andersen HF, Overø K (2002). Citalopram dose-response revisited using an alternative psychometric approach to evaluate clinical effects of four fixed citalopram doses compared to placebo in patients with major depression. *Psychopharmacology* **163**, 20–25.

Bech P, Tanghøj P, Cialdella P, Friis Andersen H, Pedersen AG (2004). Escitalopram dosee–response revisited: an alternative psychometric approach to evaluate clinical effects of escitalopram compared to citalopram and placebo in patients with major depression. *International Journal of Neuropsychopharmacology* **7**, 283–290.

Beck AT (1973). *The Diagnosis and Management of Depression*. University of Pennsylvania Press: Philadelphia.

Bent-Hansen J, Lunde M, Klysner R, Andersen M, Tanghøj P, Solstad K, Bech P (2003). The validity of the depression rating scales in discriminating between citalopram and placebo in depression recurrence in the maintenance therapy of elderly unipolar patients with major depression. *Pharmacopsychiatry* **6**, 313–316.

Bollini P, Pampallona S, Tbaldi G, Kupelnick B, Munizza C (1999). Effectiveness of antidepressants. Meta-analysis of dose-effect relationships in randomised clinical trials. *British Journal of Psychiatry* **174**, 297–303.

Boyer P, Montgomery S, Lepola U (2008). Efficacy, safety and tolerability of fixed-dose desvenlafaxine 50 mg and 100 mg/day for major depression in a placebo-controlled trial. *International Clinical Psychopharmacology* **23**, 248–253.

Cohen J (1976). *Statistical Power Analysis for the Behavioural Sciences*. Lawrence Erlbaum, New York.

Detke MJ, Wiltse CG, Mallinckrodt CH, Mc Namara RK, Demitrack MA, Bitter I (2004). Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *European Neuropsychopharmacology* **14**, 457–470.

Entsuah R, Shaffer M, Zhang J (2002). A critical examination of the sensitivity of unidimensional subscales derived from the Hamilton Depression Rating Scale to antidepressant drug effects. *Journal of Psychiatric Research* **36**, 437–448.

European Union (1994). Guidelines on psychotropic drugs: antidepressant medical products. *European Neuropsychopharmacology* **4**, 62–65.

Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, Goodwin GM (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet* **61**, 653–661.

Hamilton M (1967). Development of a rating scale for primary depressive illness. *British Journal of social and Clinical Psychology* **6**, 278–296.

Hamilton M (1986). The Hamilton Depression Scale. In *Assessment of Depression* (ed. N. Sartorius and T. A. Ban), pp. 143–152. Springer: Berlin.

Hamilton M (1987). Assessment of psychopathology. In *Human Psychopharmacology. Measures and Methods* (ed. I. Hindmarch and P. D. Stonier), pp. 1–18. Wiley: Chichester.

Hays RD, Hadorn D (1992). Responsiveness to change: an aspect of validity, not a separate dimension. *Quality of Life Research* **1**, 73–75.

Hays RD, Morales LS, Reise SP (2000). Item response theory and health outcome measurement in the 21st century. *Medical Care* **38** (Suppl. 2), 28–42.

Hedges LV, Olkin I (1985). *Statistical Methods for Meta-analysis*. Academic Press: Orlando.

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *Public Library of Science Medicine* **5**, 260–268.

Kuhn R (1958). The treatment of depressive states with G 22355 (imipramine hydrochloride). *American Journal of Psychiatry* **115**, 459–464.

Kuhn R (1970). The imipramine story. In *Discoveries in biological psychiatry* (ed. F. J. Ayd and B. Blackwell), pp. 205–217. Lippincott: Philadelphia.

Lauritzen L, Odgaard K, Clemmesen L, Lunde M, Ôhrström J, Black C, Bech P (1996). Relapse prevention by means of paroxetine in ECT-treated patients with major depression: a comparison with imipramine and placebo in medium-term continuation therapy. *Acta Psychiatrica Scandinavica* **94**, 241–251.

Leber P (1996). The role of the regulator in the evaluation of the acceptability of new drug products. In *Psychotropic Drug Development* (ed. D. Healy and D. P. Doogan), pp. 69–77. Chapman & Hall Medical: London.

Lingjærde O, Ahlfors UG, Bech P, Dencker SJ, Elgren K (1987). The UKU side effect rating scale. A new comprehensive rating scale for psychotropic drugs and a cross-sectional study of side effects in neuroleptic-treated patients. *Acta Psychiatrica Scandinavica* (Suppl.) **334**, 1–100.

Mayer DY (1975). Psychotropic drugs and the 'anti-depressed' personality. *British Journal of Medical Psychology* **48**, 349–357.

Montgomery SA, Åsberg M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* **134**, 382–389.

NICE (2004). (http://www.nice.org.uk/nicemedia/pdf/CG23fullguideline.pdf), p. 41. National Institute for Health and Clinical Excellence.

Øhrberg S, Christiansen PE, Severin B (1992). Paroxetine and imipramine in the treatment of depressive patients in psychiatric practice. *Acta Psychiatrica Scandinavica* **86**, 437–444.

Overall JE (1962). The dimensions of manifest depression. *Psychiatric Research* **1**, 239–245.

Parker G (2009). Antidepressants on trial: how valid is the evidence. *British Journal of Psychiatry* **194**, 1–3.

Porter AMW (1970). Depressive illness in a general practice. A demographic study and a controlled trial of imipramine. *British Medical Journal* **1**, 773–778.

Pinder RM, Zivkov M (1998). On demonstrating the correct dose for new antidepressants. In *Clinical Pharmacology in Psychiatry: Finding the Right Dose of Psychotropic Drugs* (ed. L. P. Balant, J. Benitez, S. G. Dahl, L. F. Gram, R. M. Pinder and W. Z. Potter), pp. 31–42. European Commission: Brussels.

Rasch G (1980). *Probalistic Models for Some Intelligence and Attainment Tests*. University of Chicago Press: Chicago.

Rasmussen A, Lunde M, Poulsen DL, Sørensen K, Qvitzau S, Bech P (2003). A double-blind, placebo-controlled study of sertraline in the prevention of depression in stroke patients. *Psychosomatics* **44**, 216–221.

Riskind JH, Beck AT, Brown G, Steer RA (1987). Taking the measure of anxiety and depression. Validity of the reconstructed Hamilton Scales. *Journal of Nervous and Mental Disorders* **175**, 474–479.

Robinson RG, Jorge RE (2009). Prevention of first-episode depression: progress and potential. *British Journal of Psychiatry* **194**, 296–297.

Ruhé HG, Dekker JJ, Peen J, Holmen R, de Jonghe F (2005). Clinical use of the Hamilton Depression Rating scale: is increased efficiency possible? A post hoc comparison of Hamilton Depression Rating Scale, Maier and Bech subscales, Clinical Global Impression, and Symptom Checklist-90 scores. *Comprehensive Psychiatry* **46**, 417–427.

Rush AJ, Beck AT, Kovacs M, Hollon S (1977). Comparative efficacy of cognitive therapy and imipramine in the treatment of depressed outpatients. *Cognitive Therapy and Research* **1**, 17–37.

Ryle G (1949). *The Concept of Mind*. Hutchinson: London.

Santen G, Gomeni R, Danhof M, Pasqua OD (2008). Sensitivity of the Hamilton Depression Scale to response and its consequence for the assessment of efficacy. *Psychiatric Research* **42**, 100–109.

Smith A, Tranganzo E, Harrison G (1969). Studies on the effectiveness of antidepressant drugs. *Psychopharmacology Bulletin* (Suppl.), 1–53.

Steinmeyer EM, Möller H-J (1992). Facet theoretic analysis of the Hamilton-D scale. *Journal of Affective Disorders* **25**, 53–62.

Turner EH (2008). Efficacy of antidepressants. *British Medical Journal* **336**, 516–517.

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal B (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* **358**, 252–60.

Young LPH, Plosker GL (2008). Desvenlafaxine extended release. *CNS Drugs* **22**, 1061–1069.