
Lessons to be Learned from Placebo Arms in Psychopharmacology Trials

Bettina K. Doering, Winfried Rief, and Keith J. Petrie

Contents

1	Introduction	274
2	The Placebo Effect in Psychopharmacology Trials	275
2.1	The Relevance of Spontaneous Remission	275
2.2	The Increasing Power of Placebo	277
2.3	The Impact of Trial Design on Placebo Responses	278
2.4	Open-Label Placebo Application	281
3	Side Effects in Psychopharmacology Trials	283
3.1	The Nocebo Effect	283
3.2	Onset Sensations	284
4	Implications for Drug Trials: Possible Interaction Effects	285
	References	287

Abstract

Large placebo effects are typically reported in clinical drug trials and evidence suggests placebo effects have increased over time. The diminishing drug–placebo difference calls into question the effectiveness of pharmacological treatments and provides a challenge to prove the effectiveness of new medications. This chapter discusses explanations for the increasing placebo effect. It highlights the contribution of spontaneous remission to the improvement in placebo groups, but focuses particularly on the role of patient and clinician expectations. Certain characteristics of the trial design can influence the formation of patient expectations and, subsequently, true placebo responses. Side effects in clinical trials may also contribute inadvertently to placebo

B.K. Doering (✉) • W. Rief
University of Marburg, Marburg, Germany
e-mail: doeringb@staff.uni-marburg.de

K.J. Petrie
University of Auckland, Auckland, New Zealand

responses. Side effects after starting medication can inform participants about their allocation to an active treatment group. Thus, they may enhance expectations of improvement and contribute to nonspecific effects in clinical trials. It is argued that specific and nonspecific effects interact in drug groups of clinical trials. This interaction influences drug–placebo differences in clinical trials (i.e., trial sensitivity). Future research should aim to identify which patients will respond best to drugs and those who may be better treated with placebos.

Keywords

Clinical trials • Placebo • Psychopharmacology • Additive model • Onset sensation • Nocebo

1 Introduction

Double-blind randomized placebo-controlled trials (RCT) have been the standard research design to investigate the effect of a new pharmacological substance on a medical condition since the 1950s (Hill 1990). Placebo interventions may consist of pharmacologically inactive pills or other sham treatments. In RCTs, patients are randomized to either an active drug arm or a placebo arm, and patient outcomes in both study arms are contrasted. Thus, RCTs seek to disentangle the specific effect of the pharmacological substance under investigation from nonspecific effects of the treatment. Nonspecific effects manifest themselves as an improvement in the placebo arm. This improvement is partly due to phenomena such as symptom fluctuation or statistical artifacts. According to Enck and colleagues (2013), the term “placebo effect” will be used in this chapter to denote all symptom changes in the placebo group, irrespective of their origin. There are different mechanisms underlying this phenomenon, including spontaneous remission, regression to the mean, natural course of a disease, biases, and true placebo responses. The term placebo response, therefore, will be reserved for the neurobiological and psychophysiological response of an individual to an inert substance or sham treatment that is mediated by factors of the treatment context.

The double-blind RCT design makes several basic assumptions (Enck et al. 2013). First, nonspecific effects should be identical in placebo and drug arms. True placebo responses due to expectancy and learning mechanisms should therefore be equally present in placebo and active drug arms. Secondly, the nonspecific effect in the placebo group should be independent of the drug. Thirdly, the improvement in the placebo group should be constant, i.e., it should not change over the course of the trial, or, at least, changes of nonspecific effects over the course of treatment should be parallel in drug and placebo groups. Lastly, the outcome in the active drug group is thought to indicate clinical relevance, i.e., to mirror the drug’s effectiveness in clinical practice. Specific and nonspecific effects must be additive in order to identify the drug-specific effect by means of comparing the symptom change in the active drug group with the change in the placebo group.

This chapter discusses empirical evidence for placebo and nocebo phenomena that challenges these assumptions. It leads to the question of whether we are drawing the right conclusions from the placebo groups of clinical trials. It focuses on placebo arms in psychopharmacology trials, since notably strong placebo effects have been observed in clinical trials involving psychiatric disorders (Kirsch et al. 2008; Price et al. 2008; Rief et al. 2011b). The discussion is therefore of particular relevance to psychopharmacology trials.

2 The Placebo Effect in Psychopharmacology Trials

Psychopharmacological drug trials often report significant symptom improvement in their respective placebo arms. This is especially true for antidepressant pharmacological treatment. Based on the results of published antidepressant drug trials, 30 % of patients in the placebo group respond to treatment, compared to 50 % of patients in the active medication arm (Walsh et al. 2002). This may still underestimate the prevalence of the placebo effect, since serious concerns about a publication bias in the antidepressant trial literature have been raised. Among 74 antidepressant trials registered with the Food and Drug Administration (FDA), 31 % of the trials, accounting for 3,449 study participants, were not published (Turner et al. 2008). Publication was associated with study outcome, with lower probability of publication for studies that were viewed by the FDA as having unfavorable results for the investigational treatment.

Kirsch and colleagues analyzed both published and unpublished data from the Food and Drug Administration (FDA) for a subgroup of four new-generation antidepressant drugs (Kirsch et al. 2008). They reported a strong placebo effect that questioned the clinical effectiveness of antidepressant treatment. When changes in the Hamilton Depression Rating Scale (Hamilton 1960) were considered as primary outcome, patients in active drug groups demonstrated a weighted mean improvement of 9.6 points on the scale, while patients assigned to placebo groups reported 7.8 points improvement. The mean drug–placebo difference therefore amounts to only 1.8 points. This has led researchers to claim that up to 75 % of the positive effect of antidepressant medication is accounted for by placebo effects (Kirsch et al. 2008). Reanalyses of the FDA data (Fountoulakis et al. 2013; Horder et al. 2010) have questioned the statistical approach of Kirsch and colleagues and have argued in favor of drug-specific antidepressant effects. However, these new analyses have not reached substantially different conclusions and have inadvertently corroborated the substantial magnitude of nonspecific effects.

2.1 The Relevance of Spontaneous Remission

The improvement seen in placebo arms of clinical trials only partially represents a true placebo response. A portion of change over the study course is likely to be caused by symptom fluctuation, i.e., spontaneous improvement or worsening in a

patient's disease. Epidemiologic surveys report high spontaneous remission rates for depression (Rhebergen et al. 2009, 2011; Wells et al. 1992). However, data from these naturalistic study designs cannot evaluate the effect of treatment on the observed course of depressive symptoms and the proportion of treated and untreated depressed study participants in the sample. In order to assess the true placebo response, a comparison of placebo groups with untreated control groups that demonstrate the natural course of the disease is needed. Unfortunately, data from no-treatment control groups from antidepressant treatment trials are scarce: in psychopharmacological trials, no-treatment control groups are not considered a valid and necessary control condition (Laughren 2001). Additionally, the inclusion of a no-treatment control group raises ethical concerns, since patients are left without treatment. The scarcity of natural course data from psychopharmacological trials is illustrated by a recent meta-analysis (Krogsboll et al. 2009). The meta-analysis attempted to quantify the spontaneous improvement in RCTs, based on a Cochrane review of the effect of placebo interventions across different medical conditions (Hrobjartsson and Gotzsche 2004). Only three-armed trials with no-treatment groups, placebo groups, and active treatment groups were included. Across all medical conditions, only 5 of 37 trials with this design employed a pharmacological treatment. For antidepressant treatment, only three 3-armed trials could be identified, and two of these trials used non-pharmacological interventions. Based on the paucity of data, it is very difficult to draw definite conclusions about the contribution of spontaneous remission to the observed symptom changes in placebo groups of antidepressant pharmacological trials.

Data concerning spontaneous improvement come primarily from trials of psychotherapeutic interventions for depression. Waitlist-controlled trials offer the treatment under investigation to patients of the control group only after a fixed waiting time, and observe the natural course of the disease during the wait. However, change in a waitlist group may be caused by various factors. While symptom fluctuation, spontaneous remission, and regression to the mean are obvious factors of influence, it is also necessary to consider other explanations (Arrindell 2001). Patients randomized to a waitlist may be disappointed about the wait, potentially resulting in an exacerbation of symptoms and an underestimation of spontaneous remission. On the other hand, diagnostic assessments during the wait may exert therapeutic benefit, and a guaranteed treatment option may induce hope and thus lead to additional improvement above natural course. It is also unclear whether patients who are enrolled in psychotherapy trials and patients enrolling in psychopharmacology trials are comparable regarding for example symptom severity or other disease-specific characteristics: if not, their spontaneous symptom change may not be identical. In spite of the limited explanatory power of waitlist control groups, they provide the best estimate of spontaneous remission effects. Therefore, focusing on psychotherapy trials, a recent meta-analytic review has attempted to investigate the contribution of spontaneous improvement to the symptom changes in placebo groups of antidepressant trials (Rutherford et al. 2012a). The authors report a medium effect size for the change in depression scores in the waitlist group. This translates to a mean improvement of four points on

the Hamilton Rating Scale for Depression (Hamilton 1960) in waitlist control groups. In placebo groups of antidepressant drug trials, an average improvement of 8 points on the scale has been reported (Kirsch et al. 2008). Since these results are based on different data sets, they cannot be compared directly. However, the estimated improvement in waitlist control groups is unlikely to account for the full magnitude of the placebo effect seen in antidepressant trials. This highlights the relevance of true placebo responses. To summarize, preliminary evidence argues that placebo effects in antidepressant clinical trials are substantially more than only spontaneous remission. Due to the paucity of data, however, additional explanatory factors need to be taken into account when interpreting the symptom change in placebo arms.

2.2 The Increasing Power of Placebo

An increasing number of clinical trials in psychopharmacology fail to demonstrate the superiority of active medication over placebo. Substantial improvement in their respective placebo arms is considered an important explanatory factor for the high failure rate in clinical trials, including antidepressant medication in both adult populations (Khin et al. 2011) and pediatric populations (Bridge et al. 2009), and antipsychotic drugs (Kemp et al. 2010). A recent meta-analysis investigated the effectiveness of second-generation antipsychotic drugs in placebo-controlled RCTs (Leucht et al. 2009). Thirty-seven RCTs representing data from over seven thousand patients diagnosed with schizophrenia were included and analyzed concerning 13 different outcome measures. Forty-one percent of patients responded to the drug compared with 24 % of patients who responded to placebo. Effect sizes varied across the treatment outcome, but they were all of only moderate size (standardized mean difference -0.51 for “overall symptoms” as predefined primary outcome). Meta-regression showed a decline in drug–placebo differences over time and a funnel plot suggested the possibility of publication bias. This bias indicates a selective publication of trials that demonstrate a significant superiority of the active medication and possibly report larger drug–placebo differences. This could mean that the already substantial placebo effect observed may be only a conservative estimate.

A decline in drug–placebo differences has already been reported in other meta-analyses of antipsychotic trials (Chen et al. 2010; Kemp et al. 2010; Potkin et al. 2011), and it is mirrored in antidepressant trials (Khin et al. 2011; Rief et al. 2009b; Walsh et al. 2002). The so-called “publication year effect” describes that the reported magnitude of placebo effects over the years has grown steadily, while experimental design has fundamentally stayed the same. Various explanations have been proposed for this effect: changes in the populations included in the trials or decreasing quality in the implementation of recent trials could contribute to this finding. To cite one example, the number of trials conducted outside the United States of America has increased. Region of data acquisition (U.S. trials versus non-U.S. trials) has been implicated as a factor of influence for

diminished drug–placebo differences in antipsychotic drug trials (Chen et al. 2010), but not in antidepressant trials (Khin et al. 2011). From a statistical point of view, larger effect sizes may originate from increased sample homogeneity in clinical trials. More homogeneous samples artificially inflate the effect size since effect sizes are calculated by dividing the mean difference by the pooled standard deviation. Indeed, evidence for increased sample homogeneity over time has been reported, for example, in a moderate association of the standard deviation of baseline depression scores with publication year (Mora et al. 2011). Nevertheless this finding can only partially explain the magnitude of the publication year effect.

Therefore, our meta-analysis investigated alternative explanations focusing on methods of assessing treatment outcome and their potential role in the publication year effect. Like Walsh and colleagues (2002) we found that effect sizes based on observer ratings in antidepressant trials correlate significantly and substantially with publication year (Rief et al. 2009b). If, however, effect sizes in placebo groups based on patient self-ratings were considered, these ratings demonstrated no significant association with publication year. Thus, while observer ratings demonstrate an increasing placebo effect, this trend is not apparent in the patients' self-ratings. To explain this surprising finding it is helpful to consider not only the role of patient expectation for placebo responses but also the role of clinician expectation about the trial. Trials of an investigational treatment that is likely to be perceived as ineffective by the study personnel report extremely low placebo effects (Shelton et al. 2001). In line with Fava and colleagues (2003), we would argue that clinician expectations about the effectiveness of antidepressant medication have probably increased over time, for example, through positive clinical experience with antidepressant pharmacotherapy. Clinician expectation may therefore be more positive than patient expectation and thus contribute to the increase of the placebo effect. However, this hypothesis awaits further investigation. Nevertheless, the diverging pattern of effect sizes in placebo groups based on observer ratings and patient self-ratings certainly questions the exclusive role of observer ratings as the gold standard of outcome assessment.

2.3 The Impact of Trial Design on Placebo Responses

While the impact of different assessment methods has already been discussed in the previous section, there are other additional characteristics of the trial design that influence the magnitude of the placebo response. Among these, characteristics that have been investigated in psychopharmacology trials (Alphs et al. 2012; Enck et al. 2013; Papakostas and Fava 2009) are:

- The duration of the clinical trial
- The number of active treatment groups or presence of a placebo group
- The number of study visits
- The use of placebo run-in phases
- Crossover design

- Flexible or fixed dosing regimes

The next section will discuss two trial characteristics that may result in changes in patient expectations based on examples from psychopharmacology trials: the number of treatment arms and effects of blinding/concealment.

An important factor of the trial design is the blinding. Double-blind design involves the blinding of study personnel and raters who evaluate the outcome. Additionally, it pertains to the blinding of patients since absent or deficient patient blinding may confound the trial outcome. While most psychopharmacological trials are designed as double-blind randomized controlled trials, a minority of trials are conducted with an open-label design, i.e., both study participants and study personnel are informed about the individual allocation to treatment arms. Additionally, some trials may be conceptualized as double blind but blinding may be broken inadvertently (cf. onset sensations). In a double-blind comparison of alprazolam, imipramine, and placebo for panic disorder, the majority of both patients and physicians were able to correctly guess the assignment to active treatment and placebo arm, respectively. Additionally, physicians were also able to accurately guess the type of active treatment that a patient had been assigned to (Margraf et al. 1991).

The influence of blinding has been demonstrated impressively in a meta-analysis of antipsychotic drugs versus placebo for relapse prevention in schizophrenia (Leucht et al. 2012). The analysis included randomized trials of patients with schizophrenia who were continued or withdrawn from antipsychotic medication after an initial stabilization period. Relapse between 7 and 12 months was defined as primary outcome and assessed by clinical judgment, e.g., need for medication or rating scales. As anticipated, all antipsychotic drugs were more successful at preventing relapse than placebo. Additionally, however, a significant difference emerged between blinded and unblinded studies. The proportion of patients in the drug groups of unblinded trials who relapsed was only 17 % compared to 28 % in blinded trials, while the proportion of patients who relapsed in the respective placebo groups was practically identical in blinded and unblinded trials (64 and 65 %, respectively). This translates to a significantly reduced risk ratio of relapse in the drug groups of unblinded trials ($RR=0.26$) compared to blinded trials ($RR=0.42$). Thus, antipsychotic drugs are apparently more effective in open-label trials. This finding is important, because open-label conditions mimic clinical practice more closely than double-blind trials.

This result also leads to the question of whether the increase in effectiveness with open-label use is caused by nonspecific treatment factors, i.e., a placebo mechanism such as expectancy. Patients who knew that they were certainly receiving the active medication may have developed more positive expectations that in turn may have resulted in a better treatment outcome. However, patient expectations are not routinely assessed in clinical trials; therefore, this explanation remains hypothetical. Another explanation could focus on clinician expectation. Since the study personnel also knew about individual allocation, this may have biased their rating of symptom severity and stability in the open-label trials and led

to an overestimation of the effectiveness of the antipsychotic drugs. However, both open-label studies used the criterion “hospital admission” in addition to more subjective data like rating scales in order to define the occurrence of “relapse.” Nevertheless, without further data, both explanations are possible. Again, they call our attention to the need for refined assessment methods on a multimodal level.

Another important characteristic of clinical trial designs is the number of active treatment arms and the definition of the control group. Adequate and well-controlled trials are needed to provide evidence for a drug’s effectiveness. The use of both placebo control groups and active medication control groups is considered to meet this requirement. Comparative effectiveness research conducts trials that employ active medication control groups: the investigational product is tested against an established standard treatment, so that all patients receive active therapy. Active comparators can also be used in combination with an additional placebo control group to result in a three-armed clinical trial design (investigational treatment, active comparator, and placebo). Obviously, these designs vary with regard to the likelihood of receiving active medication or placebo.

A recent meta-analysis of atypical antipsychotic trials in schizophrenia examined whether the investigational active treatments performed equally well in active-controlled or low-dose controlled trials compared to placebo-controlled trials (Woods et al. 2005). Based on published and unpublished data, it demonstrated that the effectiveness of investigational treatments depended on trial design: all investigational treatments were associated with greater symptom improvement in active-controlled designs. The same drugs and doses were almost twice as effective when employed in an active-controlled design compared to placebo-controlled studies. Similar results have been reported for antidepressant trials. The response rate to antidepressants is higher in trials that do not include a placebo arm (65.4 %) than in placebo-controlled trials (57.7 %) (Sinyor et al. 2010). However, active-controlled trials and placebo-controlled trials may vary with regard to additional characteristics, e.g., different completion rates and study sample selection. These differences could partially account for the design-specific placebo effect. In antidepressant trials, however, different dropout rates in active-controlled designs and placebo-controlled design do not seem to add to this effect (Rutherford et al. 2012b). A convincing explanation for this differential improvement is an expectancy effect: patients who are enrolled in an active-controlled trial know that they will definitively receive active medication after the informed consent procedure. This knowledge engenders positive treatment expectations. These expectations in turn act as nonspecific treatment factors (i.e., a placebo mechanism) that contribute to the symptom improvement observed in both active treatment groups. In addition to patient expectations, expectations of study personnel will probably also differ in active-controlled and placebo-controlled trials for the same reason. The clinician expectations may also impact ratings of improvement in the placebo groups. Since the definition of treatment response in the meta-analysis was based on observer ratings (Sinyor et al. 2010), concurrent influences of patient and clinician expectations cannot be quantified.

In either case, the improvement in active drug arms varies as a function of control group. This finding is complemented by varying response rates in placebo groups in antidepressant trials with one or more active medication arms (Sinyor et al. 2010). Trials that include only the investigational treatment as active medication and a placebo treatment as control group yield lower response rates in placebo groups (34 %) than trials that include at least a second active treatment arm (46 %). Thus, depressed patients respond better to placebo in trials that offer a higher likelihood of receiving active medication than in trials that offer only a 50 % chance of active treatment. In consequence, the trial design can lead to increased placebo responses (i.e., nonspecific treatment factors) that may not only impact the improvement in the placebo group but also in the active medication group. In a meta-regression of antidepressant trials a greater probability of receiving placebo predicted a better efficacy separation of drug and placebo (Papakostas and Fava 2009). This association remained significant independent of a simultaneous consideration of publication year and baseline depression severity as additional predictors.

The meta-analytical evidence that the number of treatment arms can impact the placebo response is corroborated by preliminary evidence from a pilot study. In this trial, assignment to placebo-controlled or active-controlled trial, respectively, directly influenced treatment expectation (Rutherford et al. 2013). Depressed patients were randomly allocated to either a placebo-controlled trial or a comparative effectiveness trial of two active antidepressant treatments. Expectancy of improvement was assessed once before randomization and at beginning of the trial. Group assignment led to the hypothesized changes in expectancy: patients in the active-controlled trial reported significantly greater expectancy of improvement than patients in the placebo-controlled trial. Importantly, baseline depression, which may be a source of more negative expectations, was not associated with this expectancy score. Additionally, higher expectancy scores were associated longitudinally with lower depression scores at the end of the study and a greater improvement in depressive symptoms over time. The mean difference between active medication groups in the placebo-controlled and active controlled trials, however, was not statistically significant. While these results should certainly be interpreted with caution, due to the limited sample size and minor methodological concerns, they illustrate the importance of accounting for patient expectation when assessing clinical trial outcome. The study also offers preliminary evidence that trial design may exert its influence on trial outcome in placebo and active medication groups through changes in patient expectation.

2.4 Open-Label Placebo Application

A special case of placebo use in clinical trials is open-label placebo application. Open label in this context means that patients are correctly informed that the pill they are receiving contains no pharmacologically active ingredient. However, positive treatment expectancies are formed through additional information, e.g.,

referring to large effects that placebo pills have demonstrated in other clinical trials. This is a novel approach to the research of placebo effects since deception has long been regarded a prerequisite for placebo responses by both healthcare professionals and laypeople. Early proof-of-principle experiments employed methodologically weak research designs (Aulas and Rosner 2003; Park and Covi 1965) and are therefore of limited internal validity. Recently, open placebo application has also been investigated in pilot RCTs. A groundbreaking study in the treatment of Irritable Bowel Syndrome (Kaptchuk et al. 2010) contrasted open-label application of a placebo pill with a natural history control group. The open-label placebo condition introduced the pill truthfully as pharmacologically inactive but also as known to result in significant improvement in Irritable Bowel Syndrome through mind–body self-healing processes. Results demonstrated clinically meaningful improvements: participants of the open-label placebo application reported significantly greater global improvement, reduced symptom severity, and increased relief.

Based on the substantial improvements in the placebo arms of psychopharmacological trials that have been reported in previous sections of this chapter, the identical rationale has been applied to the treatment of Major Depressive Disorder. Kelley and colleagues (2012) conducted a pilot waitlist-controlled RCT in 20 patients. Placebo pills were correctly introduced as pharmacologically inactive but also with regard to their substantial positive effects in clinical trials of depression and with additional explanations for their use. Patients were assessed at baseline before treatment and again after 2 weeks with the Hamilton Rating Scale for Depression (Hamilton 1960). The experimental group and the control group demonstrated no significant differences. However, preliminary data show an interesting trend: the improvement in Hamilton Rating scores was of medium effect size in the open-label placebo group ($d = 0.53$). Notably, this trend emerged in spite of a minimal sample size ($n = 11$) and a very limited observation period. A replication investigating a larger sample over a longer period of time is desirable, before any conclusions about the efficiency of open-label placebo application in the treatment of depression can be drawn.

A different approach to an open-label placebo application has been investigated in the treatment of Attention Deficit Hyperactivity Disorder (Sandler and Bodfish 2005; Sandler et al. 2010). The design combined pharmacologically active drugs and open-label placebo application in a classical conditioning paradigm using two control conditions. In the experimental group, mixed amphetamine salts were paired in the acquisition period with a visually distinct placebo capsule that was truthfully specified as a placebo. Additionally, the placebo pill was also referred to as a “dose extender” that could generate positive effects on ADHD based on mind–body interactions and the placebo mechanisms of learning and expectancy. After 1 month of acquisition, the dose of amphetamines was reduced for 1 month to 50 % of the original amount and again paired with the placebo. Outcomes were contrasted with two control groups. The first control group received their original dose continuously. The second control group received only 50 % of the original dose (similar to the experimental group) but without the placebo application. Compared to the simple dose reduction group, the open-label placebo group

demonstrated better outcomes such as lower side effect rates and maintained ADHD symptom control.

Evidence for the effectiveness of open-label placebo application is still sparse. The few studies suffer from weaknesses such as small sample sizes or the inherently impossible double-blind masking in open-label applications. Special attention must also be paid to the role of the patient–provider relationship and the instructions about the placebo pill in the respective contexts (Kaptchuk et al. 2010). Nevertheless, this innovative approach has yielded first encouraging results. Open-label placebo application may be of special interest to medical conditions that demonstrate substantial placebo effects in clinical trials and that involve a treatment that is associated with severe side effects.

3 Side Effects in Psychopharmacology Trials

Adverse events that occur in the placebo group of a clinical trial have been termed nocebo effects (Barsky et al. 2002). Like placebo responses, nocebo responses are induced by patient’s response expectations about the treatment outcome and the medication under investigation. The nocebo phenomenon is of great relevance to clinical practice (Doering and Rief 2013) but also to clinical trials. The next section discusses evidence that nocebo effects may lead to an increase in the symptom burden and may distress the patient. Nocebo-induced side effects significantly influence a patient’s decision to adhere to a prescribed treatment and may ultimately lead to the decision to discontinue participation in a clinical trial. The second section elaborates how sensations or minor symptoms that patients associate with study medication intake may inadvertently contribute to placebo responses.

3.1 The Nocebo Effect

Nocebo research requires systematic assessment of adverse events in clinical trials, preferably both on objective and on subjective level (Rief et al. 2011a). Unfortunately, this issue is not routinely addressed in psychopharmacological trials: in clinical studies of antipsychotic medication, only a minority of studies investigated subjectively experienced side effects and standardized, systematic assessment methods were rarely used (Pope et al. 2010). Our knowledge about nocebo responses in clinical trials is therefore limited and has to be interpreted in the context of differing and mostly unsystematic assessment methods.

A convincing example for the relevance of the nocebo phenomenon comes from a review of statin drug trials (Rief et al. 2006): in these trials, a comparable number of patients from both active treatment and placebo groups discontinued trial participation, with dropout rates varying from 10 to 28 %. Of note, a considerable number of patients from the placebo group discontinued treatment specifically because of side effects that they had experienced (4–26 %). A meta-analysis of antidepressant drug trials (including only tricyclic antidepressants and selective

serotonin reuptake inhibitors) reports comparable results: discontinuation rates were nearly identical for placebo groups and corresponding drug groups, 24.7 and 24.8 %, respectively (Rief et al. 2009a). Similar results have been reported for clinical trials in the pharmacological treatment of fibromyalgia, investigating drugs including the antidepressants duloxetine and milnacipran and the anticonvulsant gabapentine (Mitsikostas et al. 2012). Thus, adverse events in placebo groups of psychopharmacology trials are relatively frequent, can even lead to trial discontinuation, and must certainly be taken into account when interpreting clinical trial data.

Recently, research has focused on the comparison of the side effect profile that is reported in the placebo group with the side effect profile that is reported in the respective active drug group. Adverse events are assumed to originate from the pharmacological profile of the drug. In the case of antidepressant drug trials for example, tricyclic antidepressants (TCA) would be expected to produce more adverse events than serotonin reuptake inhibitors (SSRI) due to their differential pharmacological mode of action. Interestingly, the placebo groups mirror this expectation: placebo groups from TCA trials report significantly more side effects than placebo groups from SSRI trials (Rief et al. 2009a). In a similar vein, adverse events in placebo groups of clinical trials of drug treatment for fibromyalgia mirrored quantitatively and qualitatively the side effects of the respective active drug arm (Mitsikostas et al. 2012). A meta-analysis of clinical trials of various anti-migraine medications (nonsteroidal anti-inflammatory drugs, triptans, anticonvulsants) reports the same drug-specific nocebo effects in the placebo arm (Amanzio et al. 2009): only placebo groups of anticonvulsant trials report anticonvulsant-specific side effects, e.g., memory difficulties and anorexia, while patients in placebo groups of nonsteroidal anti-inflammatory drug trials report more gastrointestinal symptoms. This is an important finding that again demonstrates how symptom changes of patients in placebo groups mirror those of patients in the respective active drug arm. Both the improvement of symptoms and the development of side effects in placebo groups can only be understood within the context of the individual study. This illustrates that a pooling of placebo groups derived from different clinical trials may lead to false conclusions.

3.2 Onset Sensations

Minor bodily symptoms associated with medication intake are not necessarily only considered in the context of adverse events, but also conceptualized as “onset sensations.” In clinical trials these onset sensations may occasionally be experienced in placebo groups as nocebo phenomena, but they occur primarily in the drug group. Onset sensations have been discussed as a confounding influence that can unblind trial participants and raters to the randomization, and thus endanger the internal validity of clinical trials (Fava et al. 2003; Margraf et al. 1991; Rief et al. 2011b). Therefore, active placebos have been proposed as an alternative; these placebos induce minor side effects that mimic those of the active drug.

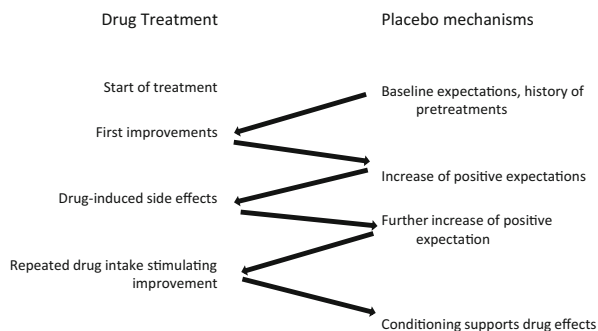
However, the placebo contains no active ingredient with specific therapeutic benefit to the medical condition under investigation. In antidepressant research atropine has been used as an active placebo in several clinical trials of TCA. A review of these studies (Moncrieff et al. 2004) concludes that the drug–placebo difference for trials using active placebos is reduced below any clinical relevance: the pooled effect size for antidepressants over placebo was 0.17, and the 95 % confidence interval ranged from 0.00 to 0.37. The review can be criticized since it only included a limited number of relatively old studies, focusing only on TCAs. However, the findings suggest that drug–placebo differences become less evident when active placebos are used as a control condition. This could be caused by a rather unlikely decrease in drug effectiveness, though it seems more likely that active placebos may be more powerful than “inert” placebos.

This hypothesis was tested empirically in the domain of placebo analgesia in healthy volunteers (Rief and Glombiewski 2012). In an experimental study inert placebos were compared with active placebos in combination with different instructions about group allocation (probability of receiving drug: 0, 50, 100 %). Participants were informed that they either had a 50 % chance of receiving the active drug (to mirror a clinical trial) or that they had a 100 % chance of receiving active medication (to mirror clinical practice). In reality, all volunteers received only placebo. Pain thresholds were assessed before and after placebo treatment. In inert placebo conditions, the well-known expectancy effect of placebo analgesia was replicated: participants who believed they had received an active drug reported the highest pain thresholds. Pain thresholds in the active placebo group differed substantially from the inert placebo group in the 50 % chance condition. Compared to participants who noted no bodily symptoms after “medication” intake, participants with minor onset sensations from active placebo intake demonstrated a greater placebo effect. It can be hypothesized that these onset sensations convinced participants that they were receiving the active medication. Increased placebo analgesia was then triggered by this expectancy effect. Since the 50 % condition most closely resembles clinical trial design, the results argue that minor onset conditions serve to strengthen nonspecific effects in clinical studies. The placebo effect observed in experimentally induced pain in healthy participants is not necessarily identical to placebo effects observed in patients who suffer from a chronic disease. In combination with data from clinical trials using active placebos (Moncrieff et al. 2004), however, these results question the relevance of drug–placebo differences stemming from inert placebos.

4 Implications for Drug Trials: Possible Interaction Effects

The evidence presented in this chapter argues strongly for the consideration of interactions between drug-specific and nonspecific effects in clinical trials, as illustrated in Fig. 1. Before a trial starts, patients will form outcome expectancies, based for example on their individual chance of receiving the active treatment in the respective trial design. Moreover, patients will probably hold expectations about

Fig. 1 Complex interaction of placebo mechanisms with specific treatment effects. Therefore, in this example, nonspecific effects in placebo and drug groups can differ



the respective treatment in general or have previous experience with the treatment in the case of chronic medical conditions. This may also influence their response to placebo and medication, and possibly to a varying degree: in depression, previous treatment experience has been reported to have a negative impact on symptom change in placebo groups, but not in active treatment groups (Hunter et al. 2010). Furthermore, the magnitude of nonspecific effects varies not only with patient expectation but also with clinician expectation, as the publication year effect suggests.

During a clinical trial, onset sensations may unblind patients to their treatment allocation and trigger expectancy effects that in turn lead to more positive treatment outcomes. However, these nonspecific effects are more likely to occur in the active treatment arm, since most clinical trials employ only inert placebos. Additionally, associative learning processes (i.e., conditioning) that link the ritual of medication intake with the experience of symptom alleviation may occur and support the drug effect.

These considerations challenge the basic assumptions of the additive model in RCTs. If nonspecific effects interact with specific effects and are strengthened by onset sensations, then nonspecific effects are not identical in placebo and drug groups. If the nonspecific effects in the placebo group vary with regard to the treatment under investigation, as demonstrated by the drug specificity of nocebo effects, then nonspecific effects can no longer be considered independent of the drug. If drug-specific effects and nonspecific effects interact and reinforce each other, true placebo responses (as a portion of the improvement observed in the drug group) will not remain constant but change over the course of the trial. Thus, an interactive model of RCTs is proposed (Enck et al. 2013) that accounts for these interaction effects in the drug group of clinical trials. This new model should guide our interpretation of clinical trial results.

Conclusion: Lessons to Be Learned

The accumulated evidence demonstrates that placebo effects are substantial, even when accounting for methodological bias and spontaneous remission. Large placebo effects challenge the development of new drugs due to diminished drug–placebo differences. Various explanations have been proposed for this

phenomenon, both pointing to methodological biases and increasing expectancy effects. Placebo effects and, in consequence, drug–placebo differences in clinical trials must be interpreted within the context of the RCT design. For example, placebo-controlled clinical trials with a second active comparator (three-armed RCTs) may yield different drug–placebo differences for a given drug than a two-armed, placebo-controlled trial of the same drug. The large placebo response in psychopharmacological trials needs to be investigated in more detail and with more suitable assessment methods. Patient and clinician expectation should be considered, and side effects assessed more carefully, in order to advance our understanding of placebo and nocebo responses in clinical trials. In the context of clinical research, alternative trial designs that are better suited to evaluate the true efficacy of the investigational drug should be employed.

Nevertheless, the large placebo response in psychopharmacology should also be a warning: for at least 25 % of depressed patients receiving antidepressants, placebos may be better options (Gueorguieva et al. 2011). For some patients, no treatment could be the recommendation of choice (i.e., spontaneous remission in natural course), especially when considering potential unwanted consequences of antidepressant treatment. At present, physicians have no empirically founded guidelines to help them to determine which depressed patient should receive no treatment, placebo treatment, or active drug treatment. Thus, refined treatment guidelines for the use of psychopharmacological medication are clearly needed, both to reduce overtreatment and to prevent under-treatment. Special attention must be paid to ethical concerns in informed consent procedures, both when using placebo interventions in clinical trials and when using verum treatments with a considerable placebo or nocebo component in clinical practice (Blease 2010; Miller and Colloca 2009; Wells and Kaptchuk 2012). The findings should also be a motivation to harness nonspecific effects and maximize them in clinical practice.

References

- Alphs L, Benedetti F, Fleischhacker WW, Kane JM (2012) Placebo-related effects in clinical trials in schizophrenia: what is driving this phenomenon and what can be done to minimize it? *Int J Neuropsychopharmacol* 15:1003–1014
- Amanzio M, Corazzini LL, Vase L, Benedetti F (2009) A systematic review of adverse events in placebo groups of anti-migraine clinical trials. *Pain* 146:261–269
- Arrindell WA (2001) Changes in waiting-list patients over time: data on some commonly-used measures. Beware! *Behav Res Ther* 39:1227–1247
- Aulas JJ, Rosner I (2003) Efficacy of a non blind placebo prescription. *Encéphale* 29:68–71
- Barsky AJ, Saintfort R, Rogers MP, Borus JF (2002) Nonspecific medication side effects and the nocebo phenomenon. *JAMA* 287:622–627
- Blease C (2010) Deception as treatment: the case of depression. *J Med Ethics* 37:13
- Bridge JA, Birmaher B, Iyengar S, Barbe RP, Brent DA (2009) Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. *Am J Psychiatry* 166:42–49

- Chen Y-F, Wang S-J, Khin NA, Hung HMJ, Laughren TP (2010) Trial design issues and treatment effect modeling in multi-regional schizophrenia trials. *Pharm Stat* 9:217–229
- Doering BK, Rief W (2013) Nocebos in daily clinical practice: the potential side effects of the treatment context and the patient–doctor interaction on pain in clinical populations. In: Colloca L, Flaten MA, Meissner K (eds) *Placebo and pain*. Academic, San Diego, pp 257–266
- Enck P, Bingel U, Schedlowski M, Rief W (2013) The placebo response in medicine: minimize, maximize or personalize? *Nat Rev Drug Discov* 12:191–204
- Fava M, Evins AE, Dorer DJ, Schoenfeld DA (2003) The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom* 72:115–127
- Fountoulakis K, Veroniki A, Siamouli M, Moller H-J (2013) No role for initial severity on the efficacy of antidepressants: results of a multi-meta-analysis. *Ann Gen Psychiatry* 12:26
- Georgieva R, Mallinckrodt C, Krystal JH (2011) Trajectories of depression severity in clinical trials of duloxetine: insights into antidepressant and placebo responses. *Arch Gen Psychiatry* 68:1227–1237
- Hamilton M (1960) A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56–62
- Hill AB (1990) Suspended judgment. Memories of the British streptomycin trial in tuberculosis. The first randomized clinical trial. *Control Clin Trials* 11:77–79
- Horder J, Matthews P, Waldmann R (2010) Placebo, Prozac and PLoS: significant lessons for psychopharmacology. *J Psychopharmacol* 25:1277–1288
- Hrobjartsson A, Gotzsche PC (2004) Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev*:Cd003974
- Hunter AM, Cook IA, Leuchter AF (2010) Impact of antidepressant treatment history on clinical outcomes in placebo and medication treatment of major depression. *J Clin Psychopharmacol* 30:748–751
- Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, Singer JP, Kowalczykowski M, Miller FG, Kirsch I, Lembo AJ (2010) Placebos without deception: a randomized controlled trial in irritable bowel syndrome. *PLoS ONE* 5:e15591
- Kelley JM, Kaptchuk TJ, Cusin C, Lipkin S, Fava M (2012) Open-label placebo for major depressive disorder: a pilot randomized controlled trial. *Psychother Psychosom* 81:312–314
- Kemp AS, Schooler NR, Kalali AH, Alphas L, Anand R, Awad G, Davidson M, Dube S, Ereshefsky L, Gharabawi G, Leon AC, Lepine JP, Potkin SG, Vermeulen A (2010) What is causing the reduced drug-placebo difference in recent schizophrenia clinical trials and what can be done about it? *Schizophr Bull* 36:504–509
- Khin NA, Chen YF, Yang Y, Yang P, Laughren TP (2011) Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. *J Clin Psychiatry* 72:464–472
- 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. *PLoS Med* 5:e45
- Krogsboll LT, Hrobjartsson A, Gotzsche PC (2009) Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Med Res Methodol* 9:1
- Laughren TP (2001) The scientific and ethical basis for placebo-controlled trials in depression and schizophrenia: an FDA perspective. *Eur Psychiatry* 16:418–423
- Leucht S, Arber D, Engel RR, Kissling W, Davis JM (2009) How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Mol Psychiatry* 14:429–447
- Leucht S, Tardy M, Komossa K, Heres S, Kissling W, Salanti G, Davis JM (2012) Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet* 379:2063–2071
- Margraf J, Ehlers A, Roth WT, Clark DB, Sheikh J, Agras WS, Taylor CB (1991) How “blind” are double-blind studies? *J Consult Clin Psychol* 59:184–187

- Miller FG, Colloca L (2009) The legitimacy of placebo treatments in clinical practice: evidence and ethics. *Am J Bioeth* 9:39–47
- Mitsikostas DD, Chalarakis NG, Mantonakis LI, Delicha E-M, Sfikakis PP (2012) Nocebo in fibromyalgia: meta-analysis of placebo-controlled clinical trials and implications for clinical practice. *Eur J Neurol* 19:672–680
- Moncrieff J, Wessely S, Hardy R (2004) Active placebos versus antidepressants for depression. *Cochrane Database Syst Rev*:CD003012
- Mora MS, Nestoriuc Y, Rief W (2011) Lessons learned from placebo groups in antidepressant trials. *Philos Trans R Soc Lond B Biol Sci* 366:1879–1888
- Papakostas GI, Fava M (2009) Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol* 19:34–40
- Park LC, Covi L (1965) Nonblind placebo trial: an exploration of neurotic patients' responses to placebo when its inert content is disclosed. *Arch Gen Psychiatry* 12:36–45
- Pope A, Adams C, Paton C, Weaver T, Barnes TR (2010) Assessment of adverse effects in clinical studies of antipsychotic medication: survey of methods used. *Br J Psychiatry* 197:67–72
- Potkin S, Agid O, Siu C, Watsky E, Vanderburg D, Remington G (2011) Placebo response trajectories in short-term and long-term antipsychotic trials in schizophrenia. *Schizophr Res* 132:108–113
- Price DD, Finnis DG, Benedetti F (2008) A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol* 59:565–590
- Rhebergen D, Beekman ATF, Rd G, Nolen WA, Spijker J, Hoogendijk WJ, Penninx BWJH (2009) The three-year naturalistic course of major depressive disorder, dysthymic disorder and double depression. *J Affect Disord* 115:450–459
- Rhebergen D, Batelaan NM, de Graaf R, Nolen WA, Spijker J, Beekman ATF, Penninx BWJH (2011) The 7-year course of depression and anxiety in the general population. *Acta Psychiatr Scand* 123:297–306
- Rief W, Glombiewski JA (2012) The hidden effects of blinded, placebo-controlled randomized trials: an experimental investigation. *Pain* 153:2473–2477
- Rief W, Avorn J, Barsky AJ (2006) Medication-attributed adverse effects in placebo groups. *Arch Intern Med* 166:155–160
- Rief W, Nestoriuc Y, von Lilienfeld-Toal A, Hofmann SG, Barsky AJ, Avorn J (2009a) Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: a systematic review and meta-analysis. *Drug Saf* 32:1041–1056
- Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG (2009b) Meta-analysis of the placebo response in antidepressant trials. *J Affect Disord* 118:1–8
- Rief W, Barsky AJ, Glombiewski JA, Nestoriuc Y, Glaesmer H, Braehler E (2011a) Assessing general side effects in clinical trials: reference data from the general population. *Pharmacoepidem Drug Saf* 20:405–415
- Rief W, Bingel U, Schedlowski M, Enck P (2011b) Mechanisms involved in placebo and nocebo responses and implications for drug trials. *Clin Pharmacol Ther* 90:722–726
- Rutherford BR, Mori S, Sneed JR, Pimontel MA, Roose SP (2012a) Contribution of spontaneous improvement to placebo response in depression: a meta-analytic review. *J Psychiatr Res* 46:697–702
- Rutherford BR, Sneed JR, Roose SP (2012b) Does differential drop-out explain the influence of study design on antidepressant response? A meta-analysis. *J Affect Disord* 140:57–65
- Rutherford BR, Marcus SM, Wang P, Sneed JR, Pelton G, Devanand D, Duan N, Roose SP (2013) A randomized, prospective pilot study of patient expectancy and antidepressant outcome. *Psychol Med* 43:975–982
- Sandler AD, Bodfish JW (2005) Open-label use of placebos in the treatment of ADHD: a pilot study. *Child Care Hlth Dev* 34:104–110
- Sandler AD, Glesne CE, Bodfish JW (2010) Conditioned placebo dose reduction: a new treatment in attention-deficit hyperactivity disorder? *J Dev Behav Pediatr* 31:369–375

- Shelton RC, Keller MB, Gelenberg A, Dunner DL, Hirschfeld R, Thase ME, Russell J, Lydiard RB, Crits-Christoph P, Gallop R, Todd L, Hellerstein D, Goodnick P, Keitner G, Stahl SM, Halbreich U (2001) Effectiveness of St. John's worth in major depression: a randomized controlled trial. *JAMA* 285:1978–1986
- Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lancot KL (2010) Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. *J Clin Psychiatry* 71:270–279
- Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *New Engl J Med* 358:252–260
- Walsh BT, Seidman SN, Sysko R, Gould M (2002) Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* 287:1840–1847
- Wells R, Kaptchuk TJ (2012) To tell the truth, the whole truth, may do patients harm: the problem of the nocebo effect for informed consent. *Am J Bioeth* 12:22–29
- Wells KB, Burnam MA, Rogers W, Hays R, Camp P (1992) The course of depression in adult outpatients. Results from the medical outcomes study. *Arch Gen Psychiatry* 49:788–794
- Woods SW, Gueorguieva RV, Baker CB, Makuch RW (2005) Control group bias in randomized atypical antipsychotic medication trials for schizophrenia. *Arch Gen Psychiatry* 62:961–970