The Placebo Effect: is achieving trial success all in our minds?

Sameer Saxena[1] and Chathika Weerasuriya[2]

[1]Atlantia Food Trials, Ireland
ssaxena@atlantiafoodtrials.com
[2]Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
c.weerasuriya@nhs.net

Introduction

The placebo response, and other closely related phenomena, have been a constant source of fascination since they were first described and although precise, well demarcated definitions of the effects may not exist[1–3] the broad related areas of research have never been starved of attention. Over the last decade we have observed a great increase in the amount of dedicated experimental work performed on these processes[4,5], and ultimately we have arrived at a consensus that they constitute real psychobiological phenomena with fundamental implications for the pathophysiology and management of medical conditions[6–9].

In tandem with our evolving understanding of the effect, we have become increasingly reliant on a single experimental design model. The randomised, double-blind, placebo-controlled randomised clinical trial (RDBPCT) has been the gold standard instrument for establishing and comparing the efficacy (active comparators or non-placebo controlled RCTs) of treatments for decades now[10]. This single study design model has become central to the regulatory approval process for pharmaceuticals and supplements alike. This has led to a massive increase in the number of RDBPCTs conducted globally – expected to be at least 50,000 per year by 2020. Parallel to this trend is the now well documented increase in the magnitude of placebo responses observed in RDBPCTs, especially in trials using relatively subjective outcome measures. This has been a topic of considerable interest in recent years, primarily because adequately differentiating active and placebo groups (i.e. achieving adequate "assay sensitivity") has become more difficult. In the current paper we describe the trends in more detail, as well as approaches to mitigate the placebo response and investigate whether newer study design models can improve the chances of achieving adequate separation between treatment groups.

Mechanistic insights

From its very first descriptions[11], research into the placebo effect has been focused inside areas of medicine that surround conditions intuitively thought to be more likely to be remedied by "inactive" modalities. These conditions share common characteristics: the severity is assessed by self-reported, highly subjective symptoms; the condition may have a relapsing-remitting course (and thus a high probability of regression to the mean); and there is often a paucity of objective clinical signs or biomarkers to aid therapy. As such, pain, psychiatric conditions and a diverse collection of "functional" disorders (e.g. gastrointestinal and rheumatological) are over-represented.

Considerable progress has been made in recent years clarifying the exact physiological mechanisms involved in the effect, and the field has now established itself as an important area of original research.

There does not appear to be a single neurobio-

Placebo	The word placebo is the Latin word for "I shall please". It is used to indicate sham treatments or inert substances such as sugar pills or saline infusions.
Nocebo	The term nocebo ("I shall harm") was introduced in analogy to "placebo" to distinguish the positive from the noxious effects of placebos, when an inert substance is given in a negative context inducing negative expectations about the outcome, for example adverse events in placebo-controlled trials, or experimental hyperalgesia and nausea in the laboratory.
Placebo effect	The placebo effect is defined as average improvement of a symptom or physiological condition following a placebo intervention in a RCT. It includes methodological factors (regression to the mean, response bias), the natural course of the disease, and context factors (expectation, learning). The "true" placebo effect is the placebo effect cleared for other contributing (or in fact confounding) factors such as the natural course of the disease or spontaneous symptom fluctuations. However, RCTs usually do not control for the natural course because this would require a "no treatment" control group; instead, it is assumed that the natural course is equal in the drug and the placebo arm, which might not be the case.
Placebo response	The placebo response refers to the outcome caused by a placebo manipulation. It reflects the neurobiological and psychophysiological response of an individual to an inert substance or sham treatment and is mediated by various factors that make up the treatment context.
Natural course of a disease	This term describes the course and outcome of an illness in the absence of any treatment or intervention; the disease is left 'to run its natural course' that includes spontaneous variation in symptom severity, which is an immanent phenomenon in most chronic clinical conditions.
Regression to the mean	A statistical phenomenon; individuals tend to have extreme values in symptom severity or physiological parameters when enrolled into a clinical trial. These values tend to be lower and closer to the average at subsequent assessments, because they are more likely to change in the direction of the mean score, instead of developing even more extreme scores. This phenomenon in part explains the improvement observed in placebo groups in clinical trials.
Assay sensitivity	The ability of a clinical trial to differentiate between an effective treatment (for example, a drug) and a less effective or ineffective treatment (for example, placebo).
CER trial	A comparative effectiveness research (CER) trial is performed to analyse the efficacy of a novel pharmacological agent or treatment in comparison with standard treatments or approved drugs. Patients are therefore randomly allocated to receive the treatment under investigation or one or more standard treatments.

Figure 1: Terms and definitions

logical basis which underpins the placebo effect. Multiple mechanisms have been suggested – acting in concert or alone – in differing disease and experimental contexts. The psychological triggers which activate these mechanisms include expectation and (Classical/Pavlovian) conditioning; they are best exemplified by key experiments in the fields of pain, Parkinson's disease and neuroendocrinology.

The role of expectation is highlighted by studies of placebo analgesia (a reported reduction in pain in response to administration of an inert substance). Neuroimaging studies have demonstrated that both placebo mediated- and μ-receptor mediated opiate (remifentanil) analgesia leads to activation of neural networks in the rostral anterior cingulate cortex (rACC) and orbitofrontal cortex[12]. These areas form part of a postulated endogenous opiate system, which is believed to modulate pain transmission at the brainstem and spinal levels via descending opioid based signals from the cortex. μ-receptor selective radiotracer studies have demonstrated endogenous opiate release in the rACC in response to placebo analgesia[13]; rACC activation in turn correlates with activation of established antinociceptive systems in subcortical areas including the periaqueductal gray and amygdale. Finally, placebo mediated analgesic effects can be abolished by opiate antagonist (naloxone) administration[14,15]; indeed, naloxone can suppress placebo induced opiate specific side effects such as respiratory depression.

In addition to the endogenous opiate system, cholecystokinin (CCK) has been implicated in placebo analgesia. CCK is postulated to have pronociceptive effects via antagonism of opiate pathways; as such, proglumide, a CCK antagonist, has been demonstrated to disinhibit placebo analgesia. CCK signalling may also underlie nocebo hyperalgesia as evidenced by proglumide mediated inhibition of the nocebo response[16].

Parkinson's disease (PD) provides further insight into the relationship between expectation and placebo effect, but as mediated by dopaminergic neurotransmission. Radiotracer studies utilising ^{11}C-raclopride (a dopamine D2-D3 receptor antagonist) have demonstrated increased striatal dopamine release in response to placebo anti-Parkinsonian therapy. Further, in awake Parkinson's disease patients undergoing implantation of deep brain stimulation electrodes, subcutaneous administration of placebo apomorphine (i.e. saline) has been shown to elicit a reduction in the firing rate of single neurons in the subthalamic nucleus (STN)[17]. STN hyperactivity is a key component of the current pathophysiological model of PD and this reduction in activity may be mediated by placebo induced dopamine release (although direct evidence for this is lacking).

Expectation of reward (which in the trial contexts above would be analgesia or motor improvement) is then a likely general mediator of the placebo response. In support of this, Scott et al[18] have demonstrated established dopaminergic reward pathways to be implicated in the placebo response. Utilising PET-CT and a standardised pain protocol, placebo responders were found to have co-incident activation of opioid receptor and dopamine receptor pathways in the nucleus accumbens, a brain region postulated to mediate reward. High placebo responses were associated with greater receptor activation.

Finally, classical (Pavlovian) conditioning is also implicated in the placebo effect. In Pavlovian conditioning, the Unconditioned Stimulus (US – typically the active agent or drug) and the Conditioned Stimulus (CS – an inert cue, i.e. the placebo) are repeatedly and administered simultaneously to establish an association between them. The CS is subsequently able to elicit a Conditioned Response (CR) as would be expected with the US. Sumatriptan is a 5-HT$_{1B/D}$ receptor agonist which stimulates growth hormone secretion and inhibits cortisol secretion. In a series of experiments[19], administration of placebo (saline) to patients who had been conditioned for two days prior with sumatriptan was able to elicit similar hormonal responses to sumatriptan itself. In this context, the CS was the process of drug administration, whereas the US was sumatriptan in the initial conditioning procedure. Critically, verbal suggestion alone, in the absence of conditioning, of the expected hormonal effect was inadequate to elicit the effect.

Certain key conclusions can be derived from the above. Firstly, it is evident that there are tangible neurobiological correlates to the placebo phenomenon – it is more than a patient report bias. Secondly, these mechanisms are heterogeneous and may be condition or intervention (e.g. drug) specific.

The context dependent nature of the placebo effect also has implications for clinical trial design and interpretation. The modern double blind placebo controlled randomised clinical trial is built on an ”additive model” (Figure 2), wherein ”non-specific” effects (including the placebo effect) are assumed to be equal, qualitatively and quantitatively, in the drug and placebo trial arms. The drug specific effects are expected to be additive to these non-specific effects in the drug arm. Therefore, subtracting the effect size of the placebo arm from the drug arm results an estimation of the drug effect. Experimental evidence now calls this assumption into question.

In meta-analyses of placebo controlled trials which report high placebo response rates, those factors which correlate with the magnitude of the response in the placebo arm have been found to

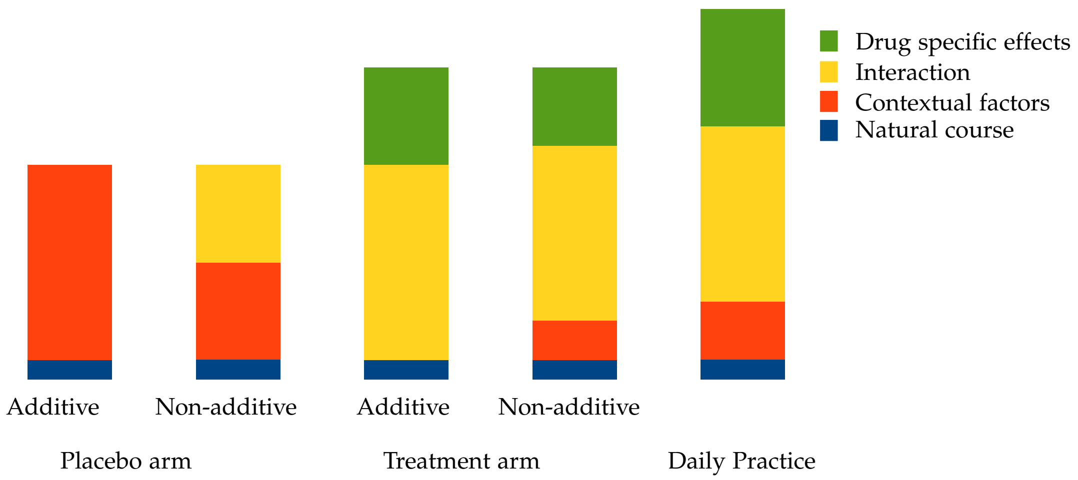

Figure 2: Additive vs Non-additive models of the Placebo response

be uncorrelated to the magnitude of the response in the drug arm. This suggests that different factors may modulate the placebo component of the drug arm compared to the placebo arm.

There is emerging neurobiological evidence to support this claim. In a double blind study involving opiate vs placebo administration in a thermal pain protocol, Petrovic et al[20] have demonstrated, via PET-CT imaging, that although bilateral activation of rACC occurs under both experimental conditions (as above), differences exist to the degree and in the simultaneous activation of other brain regions. Rostral anterior cingulate cortex activation is greater in the opiate condition, whereas orbitofrontal cortex and ventrolateral prefrontal cortex activation is greater in placebo. As such, different mechanisms may mediate the response under the two experimental conditions. Critically, regional activation under the placebo condition is not a simple subset of the opiate condition – the activation seen in the placebo condition is absent or diminished in the opiate condition. In the context of clinical trials, this may be reflected as a reduction in the placebo component of the response in the drug arm, as such violating the additivity assumption.

Genomic variation may also contribute to differential placebo response in patients, conceptually encapsulated within a framework referred to as the "placebome"[21]. Unlike the myriad other study design, patient and healthcare provider factors which contribute to the response (see discussion below), the genome of a given trial participant is stable. Identification of genomic biomarkers of the placebo response may enable both further characterisation of the physiologic basis of the response and, perhaps more pragmatically, identification of placebo responders prior to enrolment in trials. Rs4680, a single nucleotide polymorphism in the gene encoding catechol-O-methyltransferase (COMT), an enzyme involved in dopamine catabolism, leads to a val-to-met change at codon 158. The met allele is less active (i.e. reduced dopamine catabolism). In an experimental pain protocol examining resting state fMRI measures of brain reward regions, COMT genetic variability and personality traits on placebo analgesia, the number of COMT met alleles was found to be linearly correlated to placebo pain suppression[22]; increased extracellular dopamine levels in brain reward regions due to the lower activity alloenzyme is hypothesized to mediate this effect. Similarly, genetic variation in serotonergic, endocannabinoid and opiate pathways have been implicated in placebo responses. The significance of these findings are wide. Firstly, confounding: unbalanced randomisation of genetically predisposed placebo responders or non-responders to drug or placebo trial arms will invalidate estimates of drug efficacy. Secondly, some placebo gene products or pathways themselves may be involved in the drug mechanism of action, e.g. dopaminergic pathways in anti-Parkinsonian drugs. Potential drug-placebo-gene interactions may affect response in the drug arm, violating the additivity assumption.

Recent trends

The current understanding of the phenomenon suggests that the overall placebo response rate in any given RDBPCT is influenced by multiple contributing factors (Figure 2). For clarity, this specific discussion will be limited to considering variables which influence the overall placebo effect rate in any given study, which are expected to be distributed amongst all arms of the study (including the treatment arms) and will exclude statistical contributions (regression to the mean, interactive effects, fluctuating natural course of disease, etc).

Careful analysis of vast amounts of readily available RCT data demonstrate that the placebo response rates vary considerably across different studies. Despite this, few notable factors have been identified that reliably predict the levels of the response in a given trial. Due to the (safely presumed) minute nature of the effect sizes that would be at play in such situations, it would not be commercially viable[23] to undertake adequately powered studies comparing the effects of varying study design or inclusion/exclusion criteria on drug-placebo differences. Similarly individualised patient or experimenter data (which may influence placebo response rates) is not routinely reported in the literature and analysis would require access to raw data, which is currently not feasible. Thus, most of the insights gained into factors influencing the placebo effect are derived from retrospective meta-analyses of previously published trials using summary group-level data and are hence limited to study design factors. Discussions of the varied experimental study designs used to investigate the placebo response itself in prospective studies is outside the scope of this paper, and so we will not explore the vast area.

A clear rise in observed industry-wide placebo response rates has been observed over the last few decades[24]. This effect has been most pronounced in the areas of psychopharmacology[5,25–27] and analgesia[28–30], however the broad effects of the phenomenon have been felt in other fields as well.

In a comprehensive analysis of RDBPCTs conducted between 1990 – 2013 for analgesic agents for chronic neuropathic pain, Tuttle et al[31] have delineated specific trends and associations with placebo responses. In this analysis, drug response (percentage change in pain at end-point compared to baseline) remained stable over the observed 23 year interval, but placebo response magnitude demonstrated a statistically significant increase. Consequently, the treatment advantage of drugs over placebo fell from 27.3% more analgesia in 1996 to 8.9% in 2013. Trial size (and therefore placebo group size) and length are correlated to placebo response magnitude. Moreover, the demonstrated increase in placebo magnitude is restricted to trials undertaken in the United States, where trials have become larger and longer (during the analysed period) compared to other geographic regions. Patient factors – such as baseline severity – were not reliable predictors of placebo response magnitude.

Similar trends have been identified in trials of antipsychotic medication for schizophrenia[23,32–34] and major adult depression[35–39]. In a meta-analysis of trials for antipsychotic treatment for schizophrenia, Rutherford et al[40] have also demonstrated increasing placebo response magnitude with publication year. As for neuropathic pain trials, treatment advantage of drug over placebo is negatively correlated with trial duration, but in contrast to neuropathic pain, is positively correlated with baseline symptom severity.

Meta-analyses of trials for GI disorders, including inflammatory bowel disease, irritable bowel syndrome, functional dyspepsia, gastro-oesophageal reflux disease and duodenal ulceration, have also demonstrated large and persistent placebo response magnitudes of 20-40%[34,41]; these results are comparable to psychiatric conditions noted above. However, trends and associations of the response are more heterogeneous: placebo response magnitude has decreased in magnitude over 1975 – 2015 in IBS, and trial duration is both positively and negatively correlated in different meta-analyses[42].

There is clearly heterogeneity in the trends of placebo responses, which may be disease or therapy specific[4,43]. Various explanations have been proposed for the results described above. Increasing trial length may confer greater education for trial participants or greater contact with health care professionals and the clinical environment with its incident greater opportunities for support. Conversely, increasing trial size – e.g. through the use of contract research organisations (CROs) – may be mediated through differential enrolled patient characteristics as different recruiting institutions (academic centres vs CROs) face differential pressures. Changing definitions of disease, responders and end-points (e.g. Rome criteria for irritable bowel syndrome) may also contribute to response rate measurement, as might coincident regulatory framework changes which affect trial parameters such as length, size or design. Finally, publication bias is evident – trials which fail to demonstrate drug over placebo benefit are likely to be under-reported, potentially further underestimating the placebo response magnitude overall.

Because the placebo response is determined by the entire psychosocial context of the patient – including previous experience of treatment and expectations of future treatment – it is difficult to isolate instances where the patient is not subject to placebo influences. For example, patients wait listed for future treatment – regarded as a retrospective "no treatment comparator" – are themselves potentially subject to a placebo by virtue of being waitlisted.

The overall trends have far reaching consequences for evidenced based medicine in general. It has become increasingly difficult to distinguish between the specific effects of an active treatment compared with placebo, leading to a crisis in drug development. This is most clearly illustrated by the fact that the percentage of CNS drugs entering phase I testing that receive regulatory approval is lower than all other therapeutic areas except oncology – in fact, around half of all drug development failures in the area are due to the inability to demonstrate efficacy in phase II studies (up 15% from 1990-2000), and now failure rates in phase III studies is over 50% as well[44]. Needless to say, the impact on drug development is extremely significant, particularly when one considers the simple fact that most negative findings in early Proof-of-Concept studies are never re-evaluated. Instead of attempting to reverse a potentially false negative, current business practices dictate that most companies abandon further investigation of new compounds when early experiments fail to show a benefit. This emphasises the importance of "getting it right the first time" in terms of selecting a study design that maximises the chances of signal detection. This has led to rapid new developments in RCT design to delineate or mitigate the placebo effect more clearly.

Mitigating the placebo response in RCTs

As noted above, RCTs are central to the drug development and approval process. As such, research interest into methods which improve assay sensitivity (i.e. optimise drug-placebo differences) has increased substantially, particularly in psychopharmacology. In this section we elaborate on a few such methods, focusing on aspects relating to the placebo phenomenon. Additionally, any such discussion would be incomplete were inadequate emphasis given to other study design factors which influence overall trial outcome, in particular the choice of primary endpoint variable. All such factors should ideally undergo review by experts in study design, on a case-by-case basis, prior to finalising trial protocols.

As far as the standard, randomised, double-blind and placebo controlled trial (RDBPCT – sometimes referred to as a conventional parallel group study) goes, there are a few salient study design characteristics that have come to light as important contributors to placebo response rates. These can vary depending on the condition that is being investigated, however some aspects have now been replicated enough in contrasting areas to be generalisable to most RDBPCTs. The methods outlined in the current paper, as discussed above, are inherently limited to broad study design factors, as data reported in published RDBPCTs typically does not include any individual patient or experimenter characteristics. Moreover, since the overall placebo response in any given study is the summation of statistical reasons for improvement and the broad mechanisms of the placebo effect, as pointed out above, one could logically expect that the statistical reasons be equally distributed across all arms of a study, assuming adequate sample size necessary for effective randomisation. Thus, the following discussion will focus on the placebo effect, unless otherwise stated, and will attempt to distil the recommendations to general guiding principles. Similarly, the discussion will also be limited to those study designs geared towards modulating the placebo response rates in treatment trials, as opposed to those used to explore the nature of the placebo response itself.

Allocation ratios and randomisation schedules interact with the placebo response through two mechanisms: firstly, expectation bias and secondly, statistical power.

Unbalanced randomisation (i.e. any deviation from a 1:1 allocation to between treatment and placebo groups) has been strongly associated with RDBPCTs which demonstrate above average placebo response rates[4,45]. This implies that employing balanced randomisation schedules, even in the context of multiple treatment arms, would enable optimal signal detection. A review of the literature suggests that a higher frequency of allocation to the intervention arm (for example, a 3:1 randomisation schedule with three active treatment arms) is especially associated with above-average placebo response rates. Additionally, the studies above are typically characterised by the use of patient-reported outcome (PRO) measures as primary or secondary endpoints, especially in diseases areas such as psychopharmacology, pain, and functional GI disorder interventions. Together, this provides credence to the proposed expectation based mechanism likely to underlie the placebo effect. Analyses of response rates in active drug comparator trials when compared to placebo controlled trials provides further evidence to this theory. Investigators[40] have demonstrated that efficacy of experimental drugs is higher in trials where they compared to existing effective active comparators as opposed to placebo. This implies that the certainty of receiving active treatment increases, via expectation, the magnitude of the placebo component of the response.

The proportion of patients allocated to placebo has direct implications for the statistical power of the trial, that is the ability to detect placebo:treatment differences. In a situation with one contrast, i.e. a placebo arm and a single treatment arm, optimal power is derived from a 1:1 allocation. As the number of treatment arms increases, for example in studies with two, three or four treatment arms, optimal power is derived from 4:3:3, 3:2:2:2 and 2:1:1:1:1 ratios respectively[46–48]. In simple terms, the addition of an extra patient to the placebo group increases the power of all placebo:drug contrasts, whereas the addition of an extra patient to any given treatment arm increases the power of the contrast between that arm and the placebo group only. Therefore, it seems prudent to make this recommendation for most RCTs in these disease categories, especially confirmatory (phase III) studies where the phenomenon is more pronounced.

Even more evidence for the benefits of balanced randomisation comes from the classic enrichment/multidosing trials carried out in migraine, depression, and schizophrenia, which exhibited higher placebo response rates as compared to standard RDBPCTs[49]. Also of note is the fact that this phenomenon does not seem to extend into other areas of research like functional GI disorders. The evidence from IBS studies would suggest that randomisation ratios have no overall effect on observed placebo responses[4]. Current expert opinion suggests that more work has to be done to confirm whether it extends to trials in areas of medicine that rely more on objective outcome measures such as biochemical or physiological markers, as opposed to those areas outlined above that rely predominantly on PROs. Nevertheless, one can also deduce that balanced randomisation in these trials did not have a deleterious effect on the overall observed placebo response rates, which strengthens the case for balanced randomisation as a generalisable recommendation to maximise drug-placebo differences.

Other strategies designed to maximise assay sensitivity have not been as replicable, either within a given field of study or between fields; as such, they cannot form general recommendations to follow for all RDBPCTs. This is illustrated by the heterogeneity in the overall trends of placebo response when trials in functional GI disorders are compared with those in pscyhopharmacology or pain, as noted above.

Effective blinding has also been sought as a solution for improving signal detection in RDBPCTs. This is based on the premise that active treatments with physiological side effects are more likely to exhibit symptomatic improvement. Support for this hypothesis comes from meta-analyses[50] which demonstrate that drug benefit is positively and significantly correlated to the number of adverse events reported in the respective drug arm of the trial, thus indicating a potential un-blinding effect of the adverse events occurring during a trial that co-determines overall drug efficacy. Recent reviews of published trials which analyse how blinding status is reported in the literature reveal that less than 1 in 40 trials reported tests for the overall effectiveness of blinding strategies, with a worrying downward trend in reporting over time[51]. Further, blinding was successful in less than half of the studies that reported presumed blinding status. Somehow, this has not been the focus of many research efforts despite evidence suggesting the importance of presumed allocation on overall symptomatic improvement. We would like to support the proposal put forward by other commentators that at least presumed treatment allocation should be evaluated after the study and this should become standard practice for all RDBPCTs. Exactly how this affects data analysis and interpretation of results following study completion is yet to be ascertained, however, but that can only be expected to change once adoption of methods to assess effectiveness of blinding improves. We can see a role for this additional data collection as a tool to explain, for example, the heterogeneity of results obtained in RDBPCTs of the same compound. In parallel, the use of active placebos has arisen as a potential solution to the problem of effective blinding. However, documented attempts have inauspiciously been confined to areas where such suitable compounds were already available (e.g. atropine in antidepressant studies), as these compounds are notoriously challenging to develop and they invariably result in detection of lower effect sizes[52–54]. As a result the use of active placebos is not a viable option for most early phase studies.

A higher frequency and duration of visits, higher frequency of IP dosing, and longer trial duration have also been (to a lesser extent) shown to be associated with uncharacteristically high placebo response rates. This might be explained by the fact that all of these factors increase interaction with health care professionals. Moreoever, and although based more on expert opinion than replicated empirical evidence, it is recommended that the number of investigators should be kept to minimum and centralised raters should be used whenever possible, in order to minimise variability and magnitude of placebo response rates,.

Crossover studies

The use of approaches intended to identify and exclude early placebo responders is another method that has been tried to achieve maximum assay sensitivity. Since any exclusion of subjects from trial evaluation in a posthoc manner would inevitably be heavily investigator-biased, a number of study designs (that would be implemented prior to the commencement of recruitment) have been proposed to counteract this.

Crossover designs had been of interest since the rise in popularity of RDBPCTs mid-century, which revealed that intra-participant variability of responses is lower than inter-participant variability under most clinical conditions. This led to the idea of each participant providing his/her own control data, by utilising a design that exposed participants to both IP and placebo in a blinded fashion with wash-out periods in between (the crossover design). The Classic two-treatment, two-period crossover (AB|BA) study is shown in Figure 3, and the following discussion will primarily focus on this.

The differences between treatment effects can be assessed by means of statistical tests for independent samples using the intra-participant differences between the outcomes in both periods as the raw data. This design thus avoids problems of similitude of study and control groups with respect to confounding variables (e.g., age and sex), and has a few other advantages compared to the standard RDBPCT. Firstly, since all patients are guaranteed to be exposed to active treatment, there are obvious ethical and recruitment benefits compared with placebo-controlled studies. In addition, the crossover design is advantageous regarding the power of the statistical test carried out to confirm the existence of a treatment effect: crossover trials require lower sample sizes than parallel-group trials to meet the same criteria in terms of type I and type II error rates. In a situation where the between-subject variance

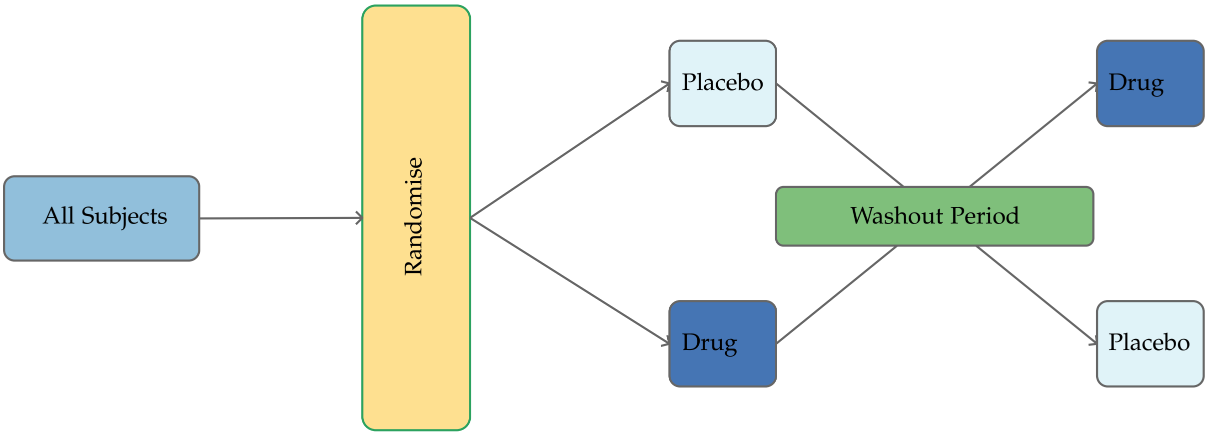

Figure 3: Cross-over design

is twice as large as that due to measurement error, for instance, six times as many patients are required to achieve the same power in a parallel-group study as in a crossover trial.

These benefits, however come with a few caveats. Many authors have commented on the prevalence of statistical errors in the analysis of crossover studies and the discussion shows no sign of abating. For example, researchers analyzing the data of crossover trials commonly proceed as though they were conducting a simple pre/post comparison. This does not allow for systematic differences in outcome between patients who receive treatment A in period 1 and treatment B in period 2 (or vice versa). This is too important to ignore, as differences can be detected even when A and B have identical effects (e.g., when the same drug is given each time), because time effects (or period effects) may be at play. In other words, the failure to accommodate stratification by sequence group, where the investigators proceed as would be appropriate when analysing a study with fixed order of treatments (e.g. by utilising a paired t-test or a Wilcoxon signed-rank test). Performing statistical analysis in this way risks depreciating the results of a crossover: In an extreme case, a significant result will only indicate that a pronounced period effect could be established, while the efficacy of the treatments themselves was equal. As a consequence, researchers planning and analysing a crossover trial have to take precautions to avoid any confounding of treatment effects and period effects. A simple example of a period effect is familiarization with the study situation, the so-called "learning effect". Problems relating to such phenomena can be minimized or eliminated when one is able to include more than two periods or more than two treatments, as in those cases intra-participant comparisons among treatments can be obtained, e.g. in an ABA | BAB study.

Additionally, the two trial periods in which the patient receives the different treatments whose effects are being compared must be separated by a washout phase that is *sufficiently long* (usually at least 5 half lives of the IP) to rule out any (first-order) carryover effect. In other words, the effect of the first treatment must have disappeared completely before the commencement of the second period. The importance of getting this right at the planning stage of a study cannot be understated. This is often difficult to ensure, and there are many cases where *a priori* knowledge of the pharmacodynamics and pharmacokinetics of test compounds is limited.

The preliminary test for differential carryover effects is usually performed as an initial step of the confirmatory analysis of the study data. However, even the primary literature on applied statistics provides no conclusive answer to the question of how one should proceed when the pre-test yields a significant result. Ordinarily, the established biometric practice in presence of a significant carryover effect in a two-period crossover

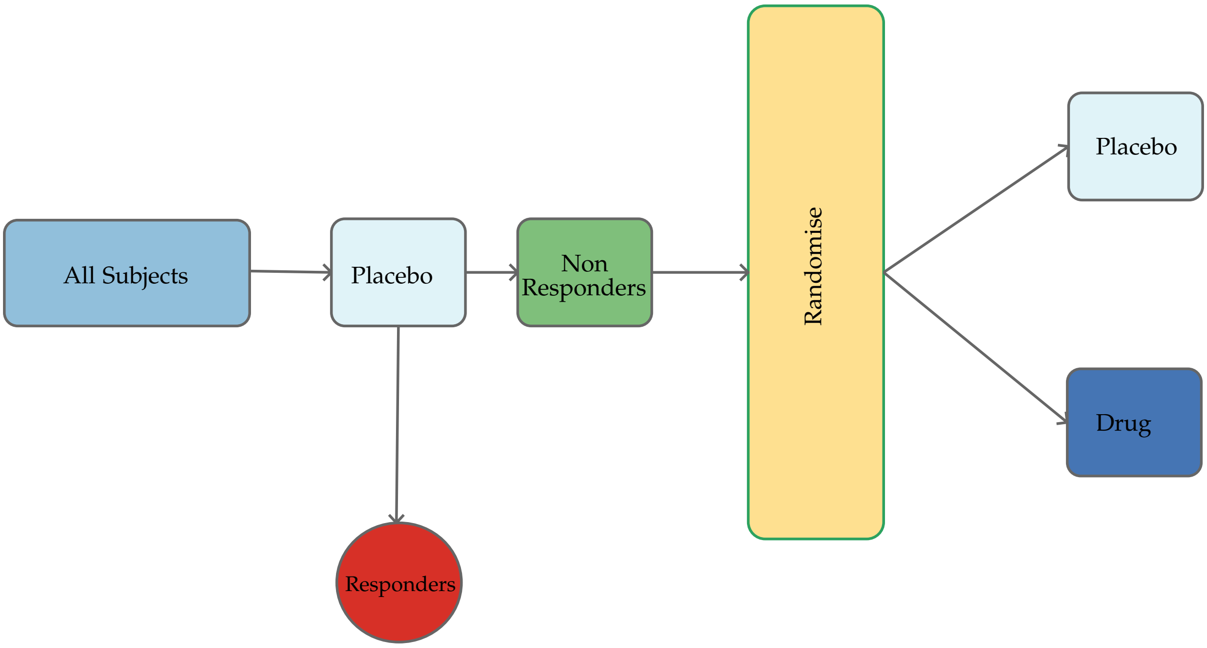

Figure 4: Placebo lead-in design

trial was to analyze the data from the first study period just as if it had been obtained from a conventional parallel-group study. However this essentially akin to throwing out half of your data! Furthermore, the test for carryover in the AB|BA design is a between-experimental unit comparison, and consequently, there may not be sufficient power in the AB/BA design to reject carryover when carryover exists. In fact, recent work has shown that the unpaired t-test, used as part of such a two-stage procedure, no longer exhibits its basic properties and may, under certain circumstances, markedly overestimate the target significance level. That is, the two-stage analysis performs poorly because the unconditional Type I error rate operates at a much higher level than desired. Part of the reason for this is that the test for differential carryover and the test for treatment differences in the first period are highly correlated and do not act independently. If differential carryover effects are of concern, then a better approach would be to use a study design that can account for them.

Crossover designs are also prone to un-blinding of the study, as differences in the side effect profiles of drug versus placebo groups can vary considerably. Thus in cases where active treatment side effects are likely to be pronounced, there is limited scope for applying a crossover design unless an adequate active placebo is available. Unfortunately, this just does not seem to be a viable option for the vast majority of compounds being tested as discussed earlier, although this might change in the future.

Placebo lead-in and other related designs

Placebo lead-in periods to standard RDCPCTs is another one of such methods used initially by large industry-sponsored psychopharmacology trials in the 1980s, that was a further step in identification and elimination of placebo responders from primary analysis. This method (Figure 4) uses a (usually open label, sometimes single blinded) placebo run-in phase of variable duration prior to randomisation, in which all participants receive placebo. Those that demonstrate significant symptom improvement are then excluded from the trial, with the remainder of participants randomised to either drug or placebo in classic RDBPC fashion. This design has important limitations, however, due to the underlying assumptions that the model makes. Firstly, being a placebo responder or a placebo non-responder is assumed to be stable individual attribute that prevents the placebo responses from occurring in non-excluded patients subsequently treated by placebo. This has been shown to be false[55] as trials using repeated treatment period designs (discussed below) have demonstrated this effect. The use of placebo run-in also risks introducing selection bias by systematically eliminating a subgroup of participants with certain characteristics[25,39,56], for example those with lower symptom severity that are prone to respond to placebo. Such a bias needs to be controlled for, otherwise drug approval authorities may be inclined to limit the indication for the drug under investigation. Additionally, this design feature is usually unblinded for the investigator (and maybe for some patients if they read the patient information carefully) and thus generates a bias in clinical assessment, although double blind versions have been developed to counteract this.

The placebo lead-in design has been the subject of many meta-analyses, and initially the consensus expert opinion more or less dismissed it as a useful tool for improving assay sensitivity, especially in psychopharmacology[57]. It has also been shown to be ineffective for that purpose in trials of functional GI disorders. More recent analyses, that importantly included FDA data on antidepressant and antipsychotic studies[9], however, have reached different conclusions. This may be just be a case of publication bias confounding the results, although there seems to exist some utility in excluding placebo responders in select cases.

Randomised Run-in/Withdrawal

Currently favoured by European (EMA, EFSA) and North American (FDA) regulatory agencies, randomised run-in/withdrawal is an unbiased method used to test whether the transition from placebo to active treatment (run-in) and from active treatment to placebo (withdrawal) creates strong placebo/nocebo effects. In this design, the run-in and withdrawal phases of the study are double-blinded, allowing temporal separation of symptom improvements/worsening from the initiation/discontinuation of treatment, which in turn aids in distinguishing "true" drug responses from drug+placebo compound effects. The logic behind the design is that a patient who has shown symptomatic improvement to an active drug is more at risk to lose that benefit when switched to placebo as opposed to remaining on drug. There are insufficient published data to recommend this method simply as a tool for maximising signal detection in early phase studies, since most have been phase III studies, where more information about toxicity and dosing are available. Additionally when using the design the total study duration is inevitably longer, as not all patients exposed to the experimental drug will be randomised. However, the conceptual advantages it affords investigative efforts combined with its current favourable standing with regulatory bodies, mean that it deserves a mention. This is of note particularly in heterogeneous diseases with classically waxing/waning natural courses, and those in which only subsets of the patient population are postulated to derive benefit from new treatments.

Novel Study Design

A number of even newer study designs that evolved from the ideas described above have also been slowly gaining traction recently. As a response to the growing problem of high placebo response rates specifically in Psychiatry RCTs, a few groups have devised original study designs intended primarily to mitigate the magnitude of placebo responses and the waxing/waning course of certain conditions. What must be stressed prior to the discourse, however, is that although the conceptual and statistical aspects of these study designs have been demonstrated (at least theoretically) to be satisfactory, there has been relatively modest experimental validation of these models using empirical data to compare them prospectively to standard RDBPCTs. This is again, in part due to the underutilisation of these novel designs in current clinical trial design practices, and also the likely costly and tedious nature of performing such validating studies. In our opinion (as with randomised run-in/withdrawal), the lack of vali-

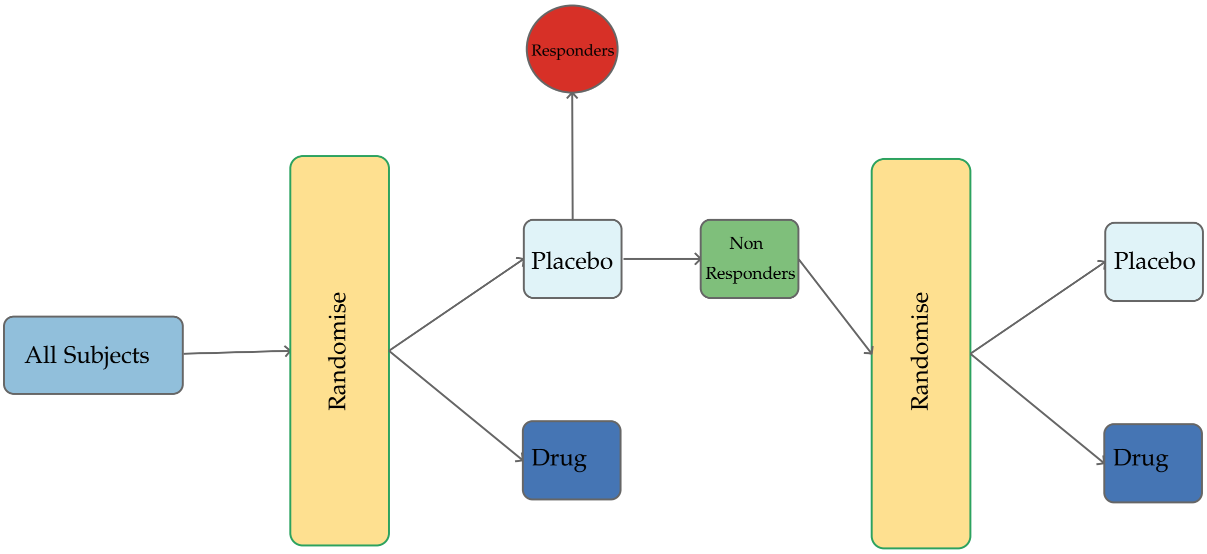

Figure 5: Sequential Parallel Design comparison

dation is very much a chicken-and-egg situation, which can and should be rectified by performing more studies with innovative study design. The statistical basis for these designs afford a number of theoretical advantages compared to standard RCTs, and these can only be realised by increasing the usage of these theoretically sound methods in practice. Again, the fact that regulatory agencies around the world are currently in favour of such designs should encourage investigators to apply them to situations where they would be useful.

Fava, Ivanova, Tamura and colleagues have described statistical derivations for novel study designs that go beyond crossover studies and attempt to combine the advantages of placebo lead-in with randomised withdrawal. The Sequential Parallel Comparison Design (SPCD; Figure 5) consists two sequential trial periods with parallel arms and two randomisation events. In Phase 1, patients are randomized to receive either drug or placebo in a conventional manner (RDBPCT), but eventually with more patients randomized to placebo[27]. In Phase 2, placebo non-responders from phase 1 are re-randomized to receive either drug or placebo. Statistical methods[58] allow the experimenter to either appraise both phases individually or, given equal treatment duration in both phases, to merge data for a common evaluation. The intellectual property rights relating to this particular study design have now been acquired by a private company, and therefore may not be a readily available option for every trial, however results of published studies invariably showed lower placebo response rates in phase 2 compared to phase 1[59–61].

The Two way enriched Design (TED; Figure 6) is another related design that was proposed as an evolution of SPCD. The idea behind the TED design is that a drug which is significantly superior to placebo in achieving short-term efficacy will also be superior to placebo in the maintenance of efficacy in drug responders[62]. This assumption, of course, needs to be examined on a case-by-case basis and may be influenced by considerations such as chronicity of the disease and time period associated with achieving response. A TED trial is also conducted in two stages: Participants are randomized to one of four sequences placebo–placebo, placebo–drug, drug–placebo and drug–drug. In the TED, first-stage placebo responders and first-stage drug non-responders are not included in the efficacy analysis in Stage 2 (although patients might be included in the trial for blinding purposes) because it is unlikely to observe a treatment effect in these patients. In other words, it re-randomizes not only placebo nonresponders

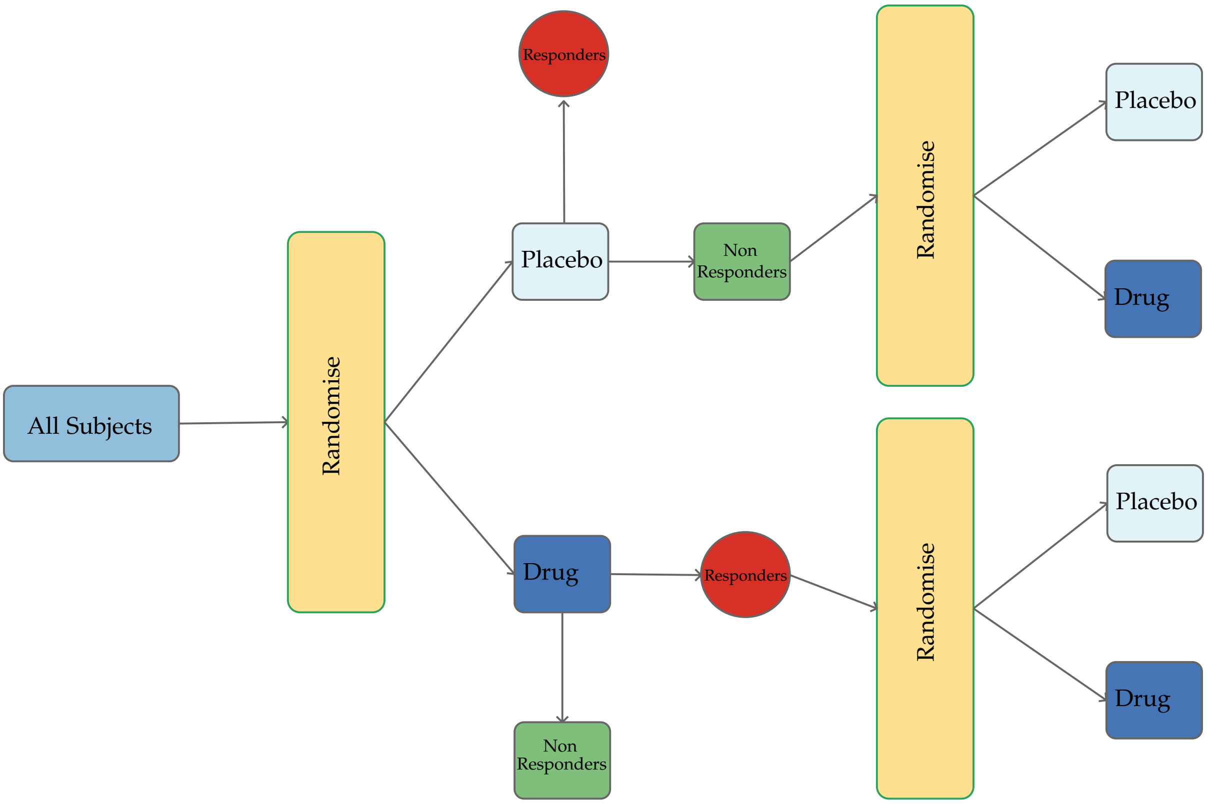

Figure 6: Two way Enriched Design

but also drug-responders to drug or placebo in Phase 2, this way proposing to enhance the drug response and decrease the placebo response of the complete trial.

A Sequential Enriched Design (SED; Figure 7) is an even newer model[63] that is designed to exclude patients who do not respond to any treatment from the study (so called "always non-responders") in addition to placebo responders. The SED commences with a double-blind placebo lead-in phase followed by a conventional parallel design in the first stage. Only patients who respond to the drug in the first stage are re-randomized to the drug or placebo at the second stage. Simulated trial data plugged into the model revealed a tendency towards lower bias in estimating treatment effects and beneficial characteristics related to statistical power. Real-world validation at this point, however is still lacking.

Since SPCD, TED and SED (similar to randomised run-in/withdrawal studies) trials are still relatively new, they are also lacking in published studies proving their utility in maximising assay sensitivity in drug development trials. However the designs have a number of built-in advantages. In the first place, no eligible patients are recruited and then not used and more responses are observed compared to parallel design or placebo lead-in. Furthermore, there is an increase in statistical power for any given sample size – both from potentially detecting a larger effect size in placebo non-responders and from the reuse of participants. Finally, although the trial is longer for each individual subject, for any given power, the overall trial duration is typically shorter because the total sample size is smaller. These trial designs are therefore suitable mainly for trials in chronic illnesses that are not yet curable, such as asthma, arthritis, hypertension, epilepsy, migraine headache and pain; that is, illnesses for which

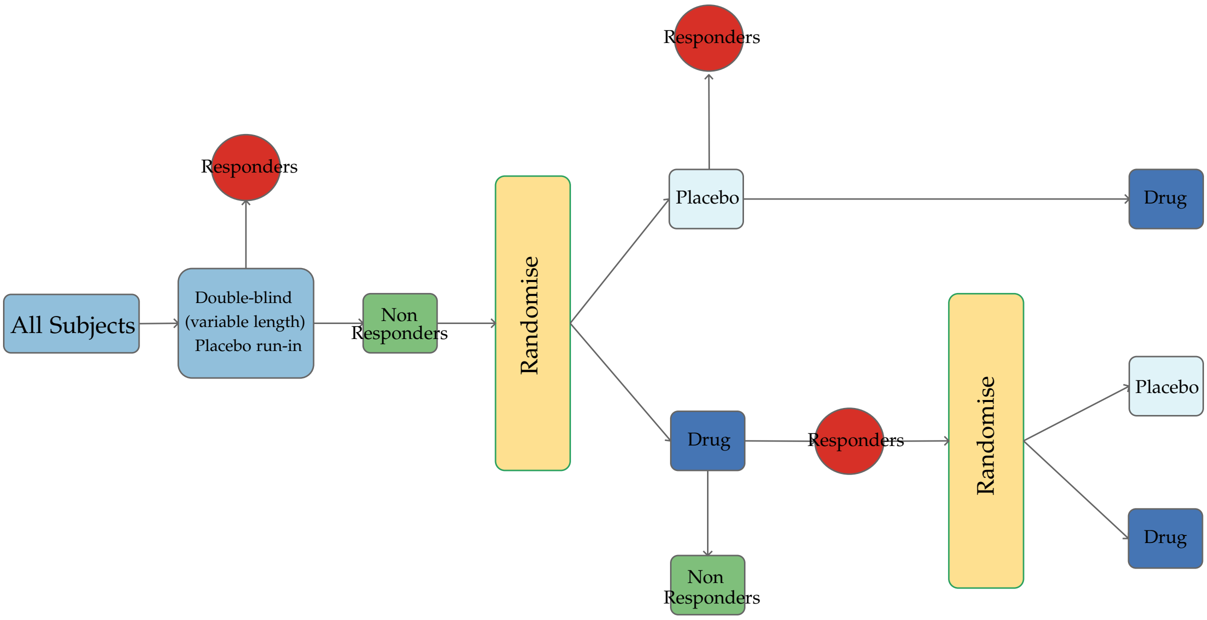

Figure 7: Sequential Enriched Design

crossover designs are sometimes considered, but frequently investigators do not use crossover design because of a concern as to the risk of a 'carry over' effect. They generally increase the efficiency of randomised clinical trials, including proof-of-concept/ Phase II and registration trials, particularly in the presence of high placebo response, but also in cases of low placebo response.

Conclusion

In the present paper we have attempted to summarise the salient aspects of the placebo phenomena in contemporary clinical research, particularly its role in influencing the probability of detecting clinically meaningful results in the classic RDBPCT design. Some interesting insights have been gained recently into the mechanisms of the phenomena, firmly establishing the neurophysiological pathways at play. Alongside this, the RDBPCT has become the de facto standard in clinical efficacy inquiry, and obtaining positive results is now the cornerstone of obtaining regulatory approval, subsequently leading to an explosion in the number of RDBPCTs conducted. The rates of new compounds successfully exiting all successive phases of drug development (especially psychopharmacological) has also declined over the same period, the cost of bringing a new compound to market has risen, and trends suggesting a rise in the "background" placebo response in trials have been put forward as an explanation. Currently it remains unclear if this effect is prevalent in all disease categories, but the trend in psychopharmacology and analgesia trials is more robust. In response, novel study designs have been proposed in the literature to mitigate the effect of these tendencies, and studies utilising them have been successfully completed lending some validation. Crossover studies and placebo lead-in periods led conceptually to the development of more statistically sophisticated models like two-way enrichment design (TED) and sequential enrichment design (SED).

In our opinion, the clinical trial landscape today is not a monotonous sea of standardised RDBPCTs, but instead an exciting area at one of the frontiers of health science with many possibilities. The choice of study design available to the investigator within this framework, it seems, has never been more open. Furthermore, we anticipate that concepts in drug development and the regulatory approval process will, in accordance with recent developments, continue to filter into the areas of nutritional and preventive research. For example the general level of scrutiny that EFSA are placing on food trials in recent years has risen – typified by the continual revision of satisfactory endpoints in health claim validation research. Similarly, in functional disease areas with a few validated alternatives beginning to surface (e.g. probiotics in functional constipation), we foresee the importance of comparative effectiveness research (CER) trials increasing, as they have done in classical drug development.

The future of placebo research seems bright, and new developments with important consequences are likely to be uncovered in the near future. The placebome and its intricacies are increasingly being delineated[21]. This is likely to be an area of intense research, especially with the continued decrease in cost, and increase in access to the technologies required. Similarly, we see "mundane" placebo metrics becoming more important in routine clinical trial work. Measures such as the effectiveness of blinding should, logically become a part of standard RDBPCTs.

Therefore (outside of the few general recommendations outlined in the main text) it seems prudent to recommend that every case should be examined in an a priori fashion with experts in trial design and the clinical disease area to determine the best possible model for achieving success. This varies depending on the investigational product in question, the disease area under investigation, the specific current regulatory framework, and many other factors.

References

1. Wolf, S. Effects of suggestion and conditioning on the action of chemical agents in human subjects; the pharmacology of placebos. J Clin Invest 1950;29:100–109.
2. Wolf, S. The pharmacology of placebos. Pharmacol Rev 1959;11:689–704.
3. Beecher, HK. The powerful placebo. J Am Med Assoc 1955;159:1602–1606.
4. Enck, P, Klosterhalfen, S, Weimer, K, Horing, B, and Zipfel, S. The placebo response in clinical trials: more questions than answers. Philos Trans R Soc Lond B Biol Sci 2011;366:1889–1895.
5. Weimer, K, Colloca, L, and Enck, P. Placebo eff ects in psychiatry: mediators and moderators. Lancet Psychiatry 2015;2:246–257.
6. Benedetti, F and Amanzio, M. Mechanisms of the placebo response. Pulm Pharmacol Ther 2013;26:520–523.
7. Hróbjartsson, A and Gøtzsche, PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344:1594–1602.
8. Hróbjartsson, A and Gøtzsche, PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev 2010:CD003974.
9. Hróbjartsson, A, Kaptchuk, TJ, and Miller, FG. Placebo effect studies are susceptible to response bias and to other types of biases. J Clin Epidemiol 2011;64:1223–1229.
10. Stolberg, M. Inventing the randomized double-blind trial: the Nuremberg salt test of 1835. J R Soc Med 2006;99:642–643.
11. Schedlowski, M, Enck, P, Rief, W, and Bingel, U. Neuro-Bio-Behavioral Mechanisms of Placebo and Nocebo Responses: Implications for Clinical Trials and Clinical Practice. Pharmacol Rev 2015;67:697–730.
12. Petrovic, P. Placebo and Opioid Analgesia– Imaging a Shared Neuronal Network. Science 2002;295:1737–1740.
13. Zubieta, JK. Placebo Effects Mediated by Endogenous Opioid Activity on -Opioid Receptors. J Neurosci 2005;25:7754–7762.
14. Levine, JD, Gordon, NC, and Fields, HL. The mechanism of placebo analgesia. Lancet 1978;2:654–657.
15. Benedetti, F. The opposite effects of the opiate antagonist naloxone and the cholecystokinin antagonist proglumide on placebo analgesia. Pain 1996;64:535–543.
16. Benedetti, F, Amanzio, M, and Maggi, G. Potentiation of placebo analgesia by proglumide. Lancet 1995;346:1231.
17. Benedetti, F, Colloca, L, Torre, E, et al. Placebo-responsive Parkinson patients show decreased activity in single neurons of subthalamic nucleus. Nat Neurosci 2004;7:587–588.
18. Scott, DJ, Stohler, CS, Egnatuk, CM, Wang, H, Koeppe, RA, and Zubieta, JK. Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. Arch Gen Psychiatry 2008;65:220–231.
19. Benedetti, F, Pollo, A, Lopiano, L, Lanotte, M, Vighetti, S, and Rainero, I. Conscious expectation and unconscious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. J Neurosci 2003;23:4315–4323.
20. Petrovic, P, Kalso, E, Petersson, KM, Andersson, J, Fransson, P, and Ingvar, M. A prefrontal non-opioid mechanism in placebo analgesia: Pain 2010;150:59–65.
21. Hall, KT, Loscalzo, J, and Kaptchuk, TJ. Genetics and the placebo effect: the placebome. Trends Mol Med 2015;21:285–294.

22. Yu, R, Gollub, RL, Vangel, M, Kaptchuk, T, Smoller, JW, and Kong, J. Placebo analgesia and reward processing: Integrating genetics, personality, and intrinsic brain activity: Resting State, Gene, Personality, and Conditioning Placebo Effect. Hum Brain Mapp 2014;35:4583–4593.

23. Agid, O, Siu, CO, Potkin, SG, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970-2010. Am J Psychiatry 2013;170:1335–1344.

24. Yang, H, Cusin, C, and Fava, M. Is there a placebo problem in antidepressant trials? Curr Top Med Chem 2005;5:1077–1086.

25. Bridge, JA, Birmaher, B, Iyengar, S, Barbe, RP, and Brent, DA. Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. Am J Psychiatry 2009;166:42–49.

26. Dunlop, BW, Thase, ME, Wun, CC, et al. A meta-analysis of factors impacting detection of antidepressant efficacy in clinical trials: the importance of academic sites. Neuropsychopharmacology 2012;37:2830–2836.

27. Fava, M, Evins, AE, Dorer, DJ, and Schoenfeld, DA. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. Psychother Psychosom 2003;72:115–127.

28. Vase, L, Riley, JL, and Price, DD. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. Pain 2002;99:443–452.

29. Arakawa, A, Kaneko, M, and Narukawa, M. An investigation of factors contributing to higher levels of placebo response in clinical trials in neuropathic pain: a systematic review and meta-analysis. Clin Drug Investig 2015;35:67–81.

30. Häuser, W, Bartram-Wunn, E, Bartram, C, Reinecke, H, and Tölle, T. Systematic review: Placebo response in drug trials of fibromyalgia syndrome and painful peripheral diabetic neuropathy-magnitude and patient-related predictors. Pain 2011;152:1709–1717.

31. Tuttle, AH, Tohyama, S, Ramsay, T, et al. Increasing placebo responses over time in U.S. clinical trials of neuropathic pain: Pain 2015;156:2616–2626.

32. Kemp, AS, Schooler, NR, Kalali, AH, et al. What Is Causing the Reduced Drug-Placebo Difference in Recent Schizophrenia Clinical Trials and What Can be Done About It? Schizophr Bull 2010;36:504–509.

33. Alphs, L, Benedetti, F, Fleischhacker, WW, and Kane, JM. Placebo-related effects in clinical trials in schizophrenia: what is driving this phenomenon and what can be done to minimize it? Int J Neuropsychopharmacol 2012;15:1003–1014.

34. Leucht, S, Heres, S, and Davis, JM. Increasing placebo response in antipsychotic drug trials: let's stop the vicious circle. Am J Psychiatry 2013;170:1232–1234.

35. Walsh, BT, Seidman, SN, Sysko, R, and Gould, M. Placebo response in studies of major depression: variable, substantial, and growing. JAMA 2002;287:1840–1847.

36. Stolk, P, Ten Berg, MJ, Hemels, MEH, and Einarson, TR. Meta-analysis of placebo rates in major depressive disorder trials. Ann Pharmacother 2003;37:1891–1899.

37. Rief, W, Nestoriuc, Y, Weiss, S, Welzel, E, Barsky, AJ, and Hofmann, SG. Meta-analysis of the placebo response in antidepressant trials. J Affect Disord 2009;118:1–8.

38. Fournier, JC, DeRubeis, RJ, Hollon, SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA 2010;303:47–53.

39. Kirsch, I, Deacon, BJ, Huedo-Medina, TB, Scoboria, A, Moore, TJ, and Johnson, BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med. 2008;5:e45.

40. Rutherford, BR, Pott, E, Tandler, JM, Wall, MM, Roose, SP, and Lieberman, JA. Placebo Response in Antipsychotic Clinical Trials: A Meta-analysis. JAMA Psychiatry 2014;71:1409.

41. Pitz, M, Cheang, M, and Bernstein, CN. Defining the predictors of the placebo response in irritable bowel syndrome. Clin Gastroenterol Hepatol 2005;3:237–247.

42. Elsenbruch, S and Enck, P. Placebo effects and their determinants in gastrointestinal disorders. Nat Rev Gastroenterol Hepatol 2015;12:472–485.

43. Quessy, SN and Rowbotham, MC. Placebo response in neuropathic pain trials. Pain 2008;138:479–483.

44. Hurko, O and Ryan, JL. Translational research in central nervous system drug discovery. NeuroRx 2005;2:671–682.

45. Enck, P, Junne, F, Klosterhalfen, S, Zipfel, S, and Martens, U. Therapy options in irritable bowel syndrome. Eur J Gastroenterol Hepatol 2010;22:1402–1411.

46. Mallinckrodt, CH, Tamura, RN, and Tanaka, Y. Recent developments in improving signal detection and reducing placebo response in psychiatric clinical trials. J Psychiatr Res 2011;45:1202–1207.

47. Mallinckrodt, CH, Zhang, L, Prucka, WR, and Millen, BA. Signal detection and placebo response in schizophrenia: parallels with depression. Psychopharmacol Bull 2010;43:53–72.

48. Potter, WZ, Mallinckrodt, CH, and Detke, MJ. Controlling Placebo Response in Drug Development: Lessons Learned from Psychopharmacology. Pharmaceut Med 2014;28:53–65.

49. Enck, P, Bingel, U, Schedlowski, M, and Rief, W. The placebo response in medicine: minimize, maximize or personalize? Nat Rev Drug Discov 2013;12:191–204.

50. Shah, E, Triantafyllou, K, Hana, AA, and Pimentel, M. Adverse events appear to unblind clinical trials in irritable bowel syndrome. Neurogastroenterol Motil 2014;26:482–488.

51. Hróbjartsson, A, Forfang, E, Haahr, MT, Als-Nielsen, B, and Brorson, S. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. Int J Epidemiol 2007;36:654–663.

52. Moncrieff, J, Wessely, S, and Hardy, R. Antidepressants using active placebos. Cochrane Database Syst Rev 2001:CD003012.
53. Moncrieff, J, Wessely, S, and Hardy, R. Active placebos versus antidepressants for depression. Cochrane Database Syst Rev 2004:CD003012.
54. Edward, SJL, Stevens, AJ, Braunholtz, DA, Lilford, RJ, and Swift, T. The ethics of placebo-controlled trials: a comparison of inert and active placebo controls. World J Surg 2005;29:610–614.
55. Lee, S, Walker, JR, Jakul, L, and Sexton, K. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. Depress Anxiety 2004;19:10–19.
56. Enck, P, Vinson, B, Malfertheiner, P, Zipfel, S, and Klosterhalfen, S. The placebo response in functional dyspepsia–reanalysis of trial data. Neurogastroenterol Motil 2009;21:370–377.
57. Trivedi, MH and Rush, H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? Neuropsychopharmacology 1994;11:33–43.
58. Ivanova, A, Qaqish, B, and Schoenfeld, DA. Optimality, sample size, and power calculations for the sequential parallel comparison design. Stat Med 2011;30:2793–2803.
59. Gewandter, JS, Dworkin, RH, Turk, DC, et al. Research designs for proof-of-concept chronic pain clinical trials: IMMPACT recommendations. Pain 2014;155:1683–1695.
60. Jeon, HJ, Fava, M, Mischoulon, D, et al. Psychomotor symptoms and treatment outcomes of ziprasidone monotherapy in patients with major depressive disorder: a 12-week, randomized, double-blind, placebo-controlled, sequential parallel comparison trial. Int Clin Psychopharmacol 2014;29:332–338.
61. Papakostas, GI, Vitolo, OV, Ishak, WW, et al. A 12-week, randomized, double-blind, placebo-controlled, sequential parallel comparison trial of ziprasidone as monotherapy for major depressive disorder. J Clin Psychiatry 2012;73:1541–1547.
62. Ivanova, A and Tamura, RN. A two-way enriched clinical trial design: combining advantages of placebo lead-in and randomized withdrawal. Stat Methods Med Res 2015;24:871–890.
63. Chen, YF, Zhang, X, Tamura, RN, and Chen, CM. A sequential enriched design for target patient population in psychiatric clinical trials. Stat Med 2014;33:2953–2967.