If only they knew! A non-inferiority randomized controlled trial comparing deceptive and open-label placebo in healthy individuals

Druart, L ; Graham Longsworth, S E ; Terrisse, H ; Locher, Cosima ; Blease, C ; Rolland, C ; Pinsault, N

DOI: https://doi.org/10.1002/ejp.2204

Received: 8 July 2023 | Revised: 26 September 2023 | Accepted: 3 November 2023

DOI: 10.1002/ejp.2204

EJP European Journal of Pain

ORIGINAL ARTICLE

If only they knew! A non-inferiority randomized controlled trial comparing deceptive and open-label placebo in healthy individuals

L. Druart[1,2] | S. E. Graham Longsworth[1,2] | H. Terrisse[1] | C. Locher[3] | C. Blease[4] | C. Rolland[1] | N. Pinsault[1,2]

[1]Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, Grenoble, France

[2]Department of Physiotherapy, University Grenoble Alpes, Grenoble, France

[3]Department of Consultation-Liaison Psychiatry and Psychosomatic Medicine, University Hospital Zurich, Zurich, Switzerland

[4]Department of Psychiatry, Digital Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA

Correspondence
L. Druart, Department of Physiotherapy, University Grenoble Alpes, 175 Av Centrale St-Martin d'Hères, Grenoble 38400, France.
Email: leo.druart@univ-grenoble-alpes.fr

Abstract

Background: Placebo use is widespread in clinical practice. However, they are most often administered deceptively rather than openly. It is often suggested that open-label placebos (OLP) are less effective than deceptive placebos (DP). This study aimed to compare the use of DP and OLP treatments to reduce pain in healthy volunteers.

Methods: We conducted a non-inferiority, parallel, randomized, controlled trial, which also included a nested cross-over no-treatment condition. This study was conducted at a university clinic in France.

Results: We included 60 subjects and the main result shows that the OLP was not inferior to the DP by a margin of 10 mm. The mean difference between both groups regarding intensity of pain was 0.7 mm with a 95% compatibility interval (95% CI) of]−∞; 5.4], and 97.5% CI of]−∞; 6.3]. Secondary outcomes require cautious interpretation of the effect of placebo versus no treatment due to a time–treatment interaction.

Conclusion: The study indicates that OLP may perform just as well as DP and could provide support for the use of OLP as an ethical alternative to DP when they are to be used in a clinical setting. If only patients knew about the placebo nature of some treatments they are receiving, unnecessary lies could be avoided while maintaining similar placebo effects.

Significance: This study is the first to show non-inferiority of placebos administered honestly, also called OLP, compared to DP in reducing pain. This suggests that OLP could be as effective as their deceptive counterparts while having the ethical advantage of not being required to lie. If deception is not a necessary condition for efficacy, OLP should be preferred over DP.

1 | INTRODUCTION

Vigorous debate surrounds the clinical use of deceptive placebo (DP) treatments. Although their use is believed to be widespread (Linde et al., 2018), a major ethical pitfall they bring is the need to deceive patients in administering such treatments (Annoni & Miller, 2014; Barnhill, 2011). In response to this, honestly prescribed, so-called "open label placebos" (hereafter, open-label placebos [OLP]) have been suggested as a more ethical solution to the use of placebo treatments in a clinical setting (Kaptchuk & Miller, 2018). As such, the efficacy of honestly prescribed placebo treatments has been studied (Aulas & Rosner, 2003). Two recent meta-analyses showed a moderate effect size when compared to no treatment (Charlesworth et al., 2017; von Wernsdorff et al., 2021): more precisely, the most recent meta-analysis of the two, found an effect size of 0.79 with a 95% compatibility interval (95% CI) of 0.38 to 1.20 (we choose here to use the term compatibility interval instead of confidence intervals as suggested by Rafi & Greenland (2020)). However, although OLPs have a supposed ethical advantage over DPs, some patients consider effectiveness over autonomy when deciding whether a placebo treatment is acceptable or when choosing their preferred placebo administration (Bishop et al., 2014; Druart et al., 2023). As pointed out by Charlesworth et al. (2017): "it is often suggested that OLP are likely to be less effective than placebos delivered deceptively". To date, several trials have compared the effectiveness of DP and OLP on pain (Disley et al., 2021; Kube et al., 2020; Locher et al., 2017; Mundt et al., 2017). Mundt et al. (2017), Locher et al. (2017), Kube et al. (2020) and Disley et al. (2021) all found no statistical difference between DP and OLP when testing for superiority. However, no studies have tested OLP and DP for non-inferiority.

Interestingly, Locher et al. (2017) also found no difference between an OLP administered without a rationale and no treatment, showing the necessity of the rationale when administering OLPs. Indeed, one of the reasons for OLP's effectiveness could be the suggested benefit through information (Blease et al., 2019). Similarly, when administering DPs, information about treatment mechanisms boosts the placebo response (Tang & Colagiuri, 2013). However, in published trials on OLPs, there are important variations in the rationale given to patients before treatment administration (Heiss et al., 2021; von Wernsdorff et al., 2021). These variations mean clinical applications might depend on the rationale given by the therapist. To this end, using a standardized rationale could be interesting to ease replication and clinical transferability.

Since no superiority is not an indication of non-inferiority or equivalence, and adding the need to replicate previous studies, we need better comparisons of DP and OLP. Therefore, as there is no superiority of DP compared to OLP and there is a supposed ethical benefit of OLP compared to DP, non-inferiority (or equivalence) design and analysis are indicated. Therefore, this study aimed to compare OLP and DP through a non-inferiority analysis. Our hypothesis is that OLP provided with standardized information upon administration will be non-inferior to DP.

2 | METHODS

This trial and analysis plan have been approved by the French National Ethical Committee (no. 2017-A01643-50) and registered on ClinicalTrials (NCT03934138). The protocol along with the analysis plan has been published in a separate article (Druart et al., 2020).

2.1 | Trial design

We conducted a non-inferiority, randomized, controlled trial comparing the use of DP and OLP. Due to both the supposed ethical superiority of the OLP over the DP and previous studies showing no superiority of DP over OLP, a non-inferiority trial is the appropriate design. The trial was a parallel study comparing a group receiving a DP and one receiving an OLP. Within the parallel design was nested a cross-over where each participant also received a no-treatment condition. This allows comparing OLP and DP (parallel design) while comparing both placebos to no treatment as a secondary outcome (nested cross-over). This design limits the impact of this secondary outcome on the number of subjects. It is also appropriate when expecting high inter-individual variability. Both the allocation to the group as well as the order of the placebo and no-treatment condition were randomized. This is made apparent in the flowchart in Figure 1. At the end of the study, some participants were also invited to participate in a qualitative study regarding the acceptability of placebo treatments (Druart et al., 2023).

2.2 | Participants

We recruited healthy participants aged between 18 and 40 to participate in a study on pain. Participants were informed before the study that it involved administration of three painful stimulations via cold water. Written consent was given by all participants. Participation was compensated 20€. Participants were recruited via advertisements for a study of a painkiller cream, on the Grenoble University campus and social media. Inclusion criteria were being aged between 18 and 40 years old,

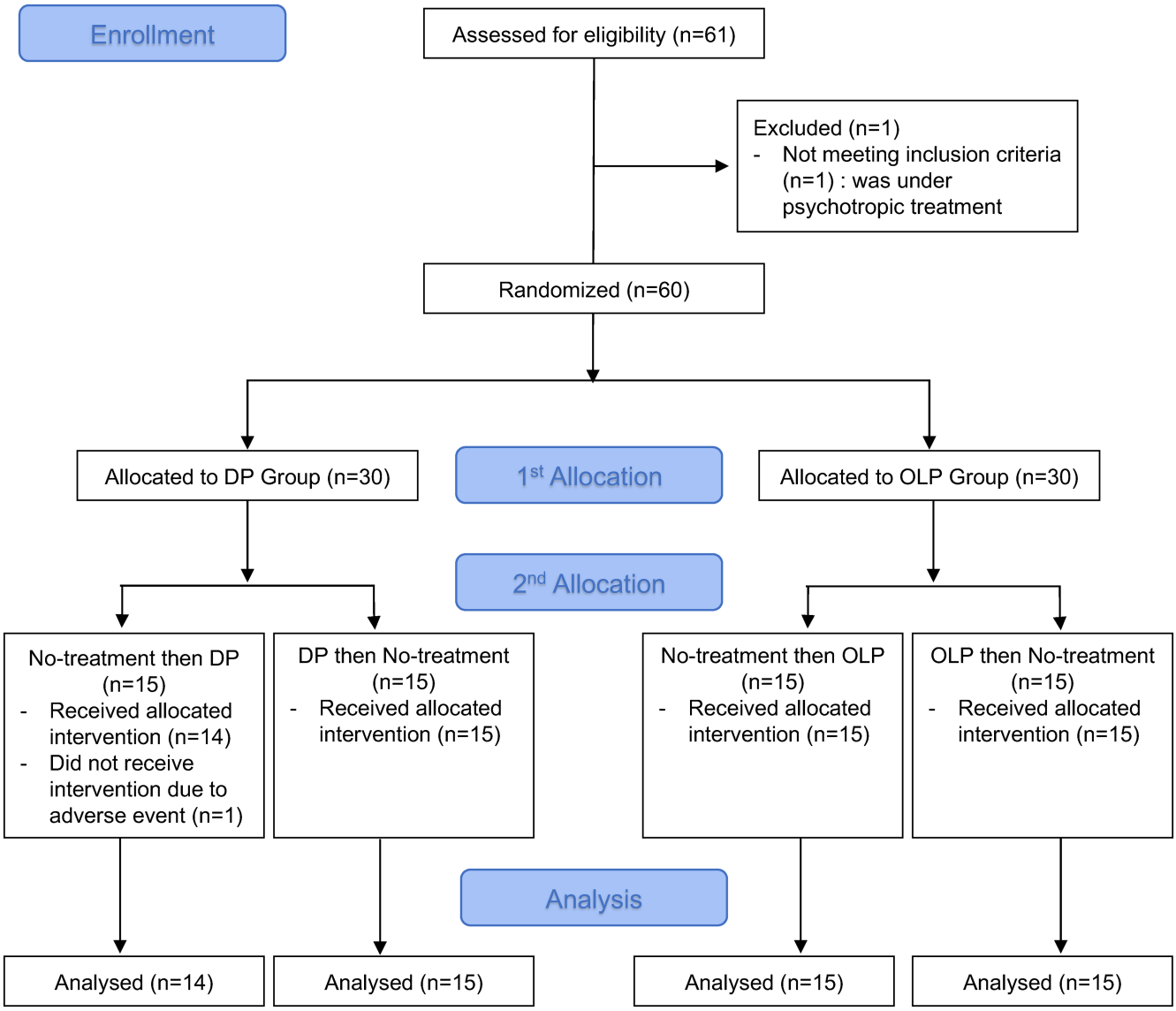

FIGURE 1 Flowchart of participants in the clinical trial.

registered to the national healthcare system and having understood and signed the written consent. Non-inclusion criteria were the following: legal impossibility to participate in the protocol (i.e. pregnant women and people deprived of their liberty) or affections modifying the perception of cold-related pain (any known pathology affecting the venous, arterial or lymphatic system, diabetes, known cardiac ailments, asthma, frostbite on the hand, epilepsy, hand arthritis, lupus erythematosus, allergic reactions to the cream or being under psychotropic or pain medication).

2.3 | Interventions

Each participant visited the research facility during a single face-to-face individual visit lasting 2 h with two physiotherapists well-versed in placebo effects conducting the study. This visit was divided into three phases: preparation, experimentation and debriefing.

The preparation phase was identical for both groups. During this part of the study, participants were allowed to ask any questions they had regarding the study procedure. The narrative of the study, at that moment, was that the aim of the study was to study the effect of a cream on pain. Participants were informed once again of the fact they would undergo three painful stimulations during the trial. Inclusion and exclusion criteria were checked, and a written consent was signed. The first cold pressor test (CPT) stimulation was the calibration CPT. During this CPT, participants immersed their hand and the distal third of their forearm in water at 1°C. A 20-min break was respected between each of the following CPTs to ensure a correct wash-out period for pain.

They filled out a Visual Analog Scale (VAS) rating for pain intensity every 5 s and were told to take their hands out once the VAS reached 7 of 10 at least. This time was recorded and set as the duration of the following experimental CPTs for this individual.

The experimental phase started after the calibration CPT. After completing a survey on knowledge regarding placebo effects, the OLP group watched a video revealing the cream studied in this trial was inert and explaining mechanisms behind the placebo effect as well as a brief explanation of the mechanisms of pain (http://bit.ly/Placethic-Video). To ensure structural equivalence (Blease et al., 2019; Locher et al., 2018), the DP group watched a control video on the history of washing hands under the pretence that they had to take their mind off the calibration before the experimental part of the study (https://www.youtube.com/watch?v=WQVYWUsrfbk). During the experimental phase, each participant underwent two more CPT procedures: one with no treatment and one with a placebo treatment either deceptive or open. The order between the no-treatment condition and the placebo condition (whether deceptive or open) was randomized. For both CPTs, participants immersed their arms in the CPT for the duration recorded during calibration and evaluated their pain intensity on the VAS once the time was up. In one case, they immersed their arm with no additional treatment and in the placebo condition an inert cream was applied before immersion. Investigators wore a white coat and carried a stethoscope during the trial. The cream was conditioned in a small 2-mL plastic syringe and administered using vinyl gloves and a short 1-min massage saying it was to help the skin "absorb" the cream. The OLP group was administered the inert cream along with the sentence: "[I will now apply a placebo cream that does not contain any active substance. It will make it easier to bear the pain during the next test through the placebo mechanisms seen in the video.]" The DP group was administered the inert cream along with the sentence: "(I will now apply an effective cream to combat cold-related pain. It will make it easier to bear the pain of the next test.)" Before immersing their arm in the placebo cream, participants filled out a questionnaire measuring treatment credibility and expectancy (Coste et al., 2020; Devilly & Borkovec, 2000; Mertens et al., 2017).

During the debriefing of the study, both groups filled out several questionnaires. The first measured perceived knowledge of the research hypothesis (Rubin, 2016), the second measured the perception of the investigators ("During the study, I trusted the investigator" on a 5-point Likert scale) and the last one measured knowledge regarding placebo effects. The questionnaire on placebo effect knowledge was designed for this study. It was inspired by the main misconceptions surrounding placebo effects (Hughes et al., 2017). It was pre-tested via cognitive interviews to check understandability and reading difficulty with 15 volunteers drawn from a convenience sample from the authors' network sharing the same characteristics as our study sample.

After study participation, all recruits were offered the 20€ compensation and investigators debriefed participants. During this discussion, we disclosed the purpose of the trial and also answered honestly any questions they had (Bishop et al., 2012).

2.4 | Blinding

During the study, only the analyst was blinded. Participants in the DP group were blind to the inert nature of their treatment and participants in the OLP group were aware they were receiving an inert treatment. Investigators were not blind to the treatment they were administering.

2.5 | Sample size

In accordance with the Food and Drug Administration and supported by the minimal clinically interesting difference recommended by Myles et al. (2017), we set the non-inferiority margin at 10 mm. Streff et al. (2010) had a similar utilization of experimental pain in their study using a CPT on healthy subjects and a VAS to estimate pain. In the dataset they present, a standard deviation of 21.9 mm is observed (Streff et al., 2010). Taking this into account and assuming a unilateral alpha of 0.05, the minimum sample size required to achieve a power of 80% to reject the inferiority null hypothesis was 60 per group. Adding to that a 5% margin of estimated non-usable data, the necessary number of subjects was set at 126.

2.6 | Randomization

Subjects were randomly allocated to a group determining which placebo they received (either DP or OLP) and to the order in which they received their placebo treatment (either placebo then no-treatment or no-treatment then placebo). Both the group and order of treatment randomization were blocked with random block sizes between 2 and 4 participants. Group allocation for each participant was stored in a sealed envelope, its content unbeknownst to investigators. A participant's envelope was opened only once he or she had signed the consent form and the experimentation had started.

2.7 | Outcomes and statistical methods

We performed the statistical analysis with alpha = 0.05. Non-inferiority was unilateral, and all other tests were bilateral. We used an intent-to-treat population approach: all participants were analysed in the group in which they were initially randomized. For the primary endpoint, we planned to impute missing data by multiple imputations if between 5% and 20% of the measurements were missing. If less than 5% were missing, we planned complete case analysis. If more than 20% of the data were missing, the results would be interpreted with caution. We planned complete case analysis on the other endpoints.

The primary endpoint was the difference in pain intensity, on a 100-point VAS, between the OLP and the DP treatments at the end of the placebo condition CPT. Pain intensity was chosen as it seems to involve less motivational and cognitive components than pain threshold (Handwerker & Kobal, 1993). We tested the non-inferiority of the OLP condition compared to the DP condition using the unilateral 95% CI of the difference in pain intensity. The non-inferiority margin was set to 10 mm. We estimated the compatibility interval by linear regression. The linear regression accounted for pain intensity during no-treatment conditions and the CPT's order. As an intent-to-treat analysis can be biased when evaluating non-inferiority (Wiens & Zhao, 2007), to check the consistency of our results, we also tested non-inferiority with a per-protocol analysis: subjects were included in the group only if there were no deviations from the protocol.

The main secondary endpoint was the difference in pain intensity between each placebo condition and no treatment. The superiority of the DP condition compared to no treatment was tested by a cross-over ANOVA including an interaction effect between time and treatment. Similarly, we also tested the superiority of the OLP condition compared to no treatment by a cross-over ANOVA.

Other secondary outcomes include the questionnaire measuring knowledge developed for this study, ranging from 0 to 17, which was compared between both groups as well as before and after watching the video for the OLP group. This questionnaire is available in the Data S1. In addition, all subjects completed a questionnaire about their perception of the investigators and the research hypothesis (Rubin, 2016). We tested these between-group differences with Student's *t*-test for two independent samples. To compare the knowledge about placebo before and after the educational video in the OLP group, we used Student's *t*-test for paired samples. Lastly, before each placebo condition CPT, participants filled out treatment credibility and expectancy questionnaire (Coste et al., 2020; Devilly & Borkovec, 2000; Mertens et al., 2017). Credibility and expectancy were each scored out of 100.

Blood pressure and heart rate values were obtained before, during and after each CPT to ensure the CPT is well tolerated.

Analysis was conducted on Stata software, version 17.0 (StataCorp LLC, College Station, TX, USA).

3 | RESULTS

3.1 | Recruitment, participant flow and harms

Recruitment spread over the 3 May 2019 to the 15 October 2021 with periods of interruption due to the COVID pandemic and lockdowns. During the end of 2021, due to the sanitary situation's impact on the trial, recruitment had to be halted.

In this timeframe, we recruited 60 volunteers of whom, one person did not receive the intervention due to an adverse event (fainting after calibration CPT). Figure 1 shows the flow of participants in the study.

3.2 | Population description

Our analysed population included 59 subjects. Their characteristics are presented in Table 1.

3.3 | Outcomes and estimation

Table 2 shows the results for each outcome depending on group. For the primary outcome, results are shown also depending on the order of treatment received.

The main result of our study is the non-inferiority comparison of the intensity of pain of the DP and the OLP groups measured by VAS. This was calculated through an

TABLE 1 Description of the participants.

Group	DP	OLP
Number of subjects	29	30
Women[a]	20 (69.0)	19 (63.3)
Age in years[b]	21.0 [19.0; 27.0]	22.0 [21.0; 22.0]
Time in seconds to reach 7/10 during calibration CPT[b]	31.0 [21.0; 46.0]	30.5 [24.0; 53.0]

[a] *n* (%).

[b] Median [Q1–Q3].

TABLE 2 Descriptive results of primary and secondary outcomes.

Group	DP	OLP
Number of participants	29	30
Treatment order: NT then placebo	14 (48.3)	15 (50.0)
VAS_{NT} (mm)	60.5 [59.0; 66.0]	66.0 [55.0; 73.0]
$VAS_{placebo}$ (mm)	55.5 [51.0; 69.0]	60.0 [48.0; 73.0]
Treatment order: placebo then NT	15 (51.7)	15 (50.0)
$VAS_{placebo}$ (mm)	64.0 [62.0; 70.0]	66.0 [60.0; 81.0]
VAS_{NT} (mm)	68.0 [60.0; 79.0]	71.0 [66.0; 88.0]
Knowledge before educative video	–	13.0 [11.0; 14.0]
Knowledge at end of trial	13.0 [11.0; 14.0]	15.5 [14.0; 16.0]
Treatment credibility (%)	66.7 [58.3; 83.3]	56.3 [33.3; 66.7]
Treatment expectancy (%)	60.8 [46.7; 71.7]	47.5 [35.0; 60.8]

ANOVA adjusted for group, treatment order and pain intensity during calibration. The results show that the mean difference (VAS_{OLP}-VAS_{DP}) was 0.7 mm with a unilateral 95% CI of]−∞; 5.4] and a bilateral 95% CI (equivalent to unilateral 97.5% CI for non-inferiority studies) of [−4.9; 6.3]. In both cases, the upper bound of the CI is below the non-inferiority margin of 10 mm, allowing to accept H1. When smaller values are desirable, non-inferiority can be claimed when the upper limit of the estimated CI is below the non-inferiority margin. Thus, we draw the following conclusion: in our study, the OLP condition was not inferior to the DP condition by a margin of 10 mm. There does not appear to be any significant deviation from the assumptions of the linear model.

The first secondary outcome that was of interest was the difference between the placebo conditions and no treatment. There was a significant interaction between time and treatment in the analysis of this outcome. During the second CPT, both placebo conditions showed no difference with the no-treatment condition. In the DP group, the mean difference (VAS_{DP}-VAS_{NT}) was 2.7 with a 95% CI of [−5.5; 10.9]. In the OLP group, the mean difference (VAS_{OLP}-VAS_{NT}) was 6.9 with a 95% CI of [−2.2; 15.9]. During the third CPT, both groups showed a statistically significant difference to no treatment. In the DP group, the mean difference (VAS_{DP}-VAS_{NT}) was −9.3 with a 95% CI of [−17.5; −1.1]. In the OLP group, the mean difference (VAS_{OLP}-VAS_{NT}) was −15.4 with a 95% CI of [−24.5; −6.3]. In both groups, one outlier was abnormally low. Sensitivity analysis was conducted to see if excluding this data point would change interpretation of results. This was not the case indicating the outlier had little impact on interpretation. In summary for this outcome, these results indicate there is no difference when comparing placebo conditions and no treatment during the second CPT and a statistically significant difference during the third CPT. Figure 2 represents the findings for the previously mentioned endpoints.

Another secondary outcome was our participant's knowledge regarding placebo. This was measured thanks to a questionnaire ranging from 0 to 17. Results were compared with a Satterthwaite test due to a difference in variability between groups. When comparing both groups (OLP_{after}-DP), the difference was 2.4 with a 95% CI of [1.4; 3.4]. This can be interpreted as the OLP group scoring significantly higher than the DP group. We also compared the scores of the OLP group before and after watching the educational video. The difference (OLP_{after}-OLP_{before}) was 2.4 with a 95% CI of [1.6; 3.3]. Figure 3 represents these findings.

Just after treatment administration and while the treatment was "taking effect on the skin", we asked participants to fill out scores regarding their expected effects and the credibility of the treatment. These questionnaires were scored out of 100. When comparing treatment credibility (OLP-DP), the difference was −16.6% with a 95% CI of [−27.3; −5.9]. This indicates there was a significantly lower credibility of the OLP treatment compared to the DP. Similarly, the difference in treatment expectancy was −11.2% with a 95% CI of [−20.9; −1.4], indicating there also was a significantly lower expectancy for the OLP than for the DP treatment. These are illustrated in Figure 4.

Lastly, there was no statistically significant difference in how both groups trusted the investigators and it seems that the participants in the OLP group were more confident they knew the research hypotheses although the questionnaire does not measure whether or not they are correct (results available in Data S2).

4 | DISCUSSION

4.1 | Findings and interpretations

In this study, we showed that an OLP with a convincing rationale was non-inferior to a DP with a 10 mm margin on the VAS and a 5% alpha risk. This study contributes

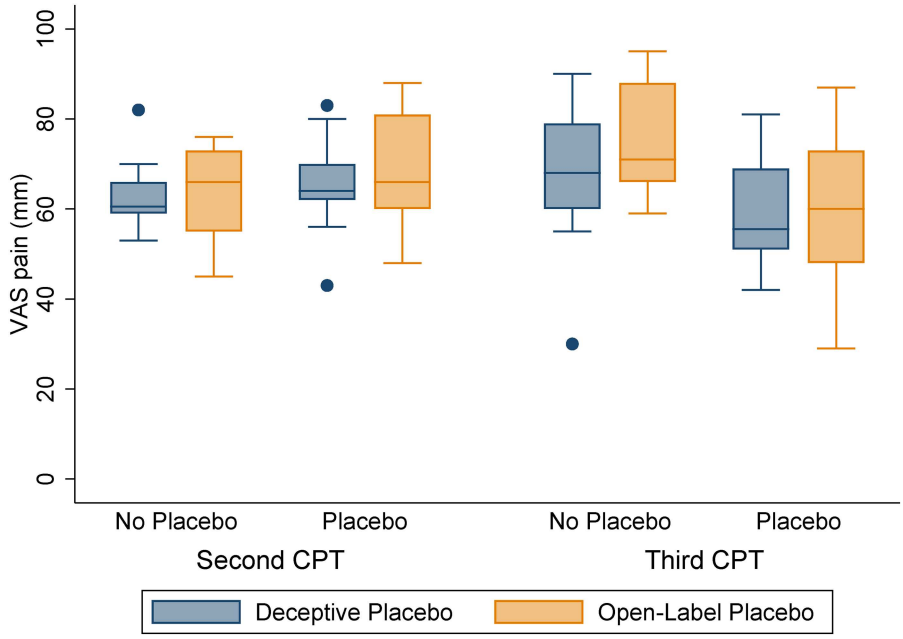

FIGURE 2 Main endpoint graphical representation measuring pain intensity with a visual analogue scale (VAS) ranging from 0 to 100 mm. Participants receiving no treatment during the second cold pressor test (CPT) are those receiving placebo in the third CPT and vice versa.

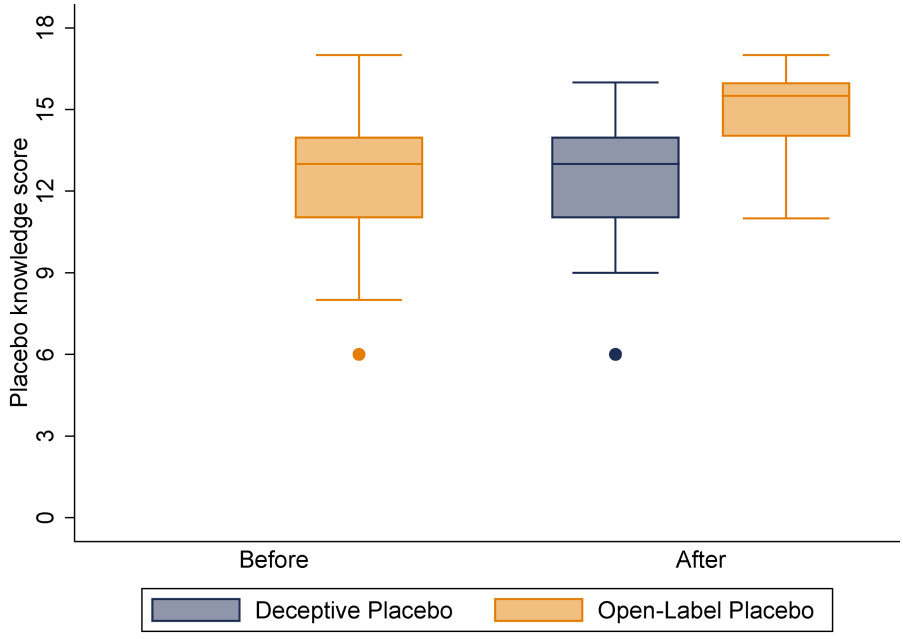

FIGURE 3 Graphical evolution of scores on the placebo knowledge questionnaire. For the OLP group, knowledge was evaluated before and after watching the educational video.

to replication and confirms the findings from previous trials comparing OLP and DP (Disley et al., 2021; Kube et al., 2020; Locher et al., 2017; Mundt et al., 2017). It also allows further interpretation than previous studies finding no superiority as it concludes with a non-inferiority.

Secondary outcomes are interesting to interpret, especially the comparison with the no-treatment conditions. Surprisingly, during the second CPT, there were no differences between placebo conditions and no-treatment. However, during the third CPT, there were significant differences. This differs from our initial hypothesis of showing superiority of placebo treatments compared to no treatment on both CPTs as shown in some published comparisons of OLP with no treatment (von Wernsdorff et al., 2021). However, in reality, our results show that time had an impact on the result of our study. This is

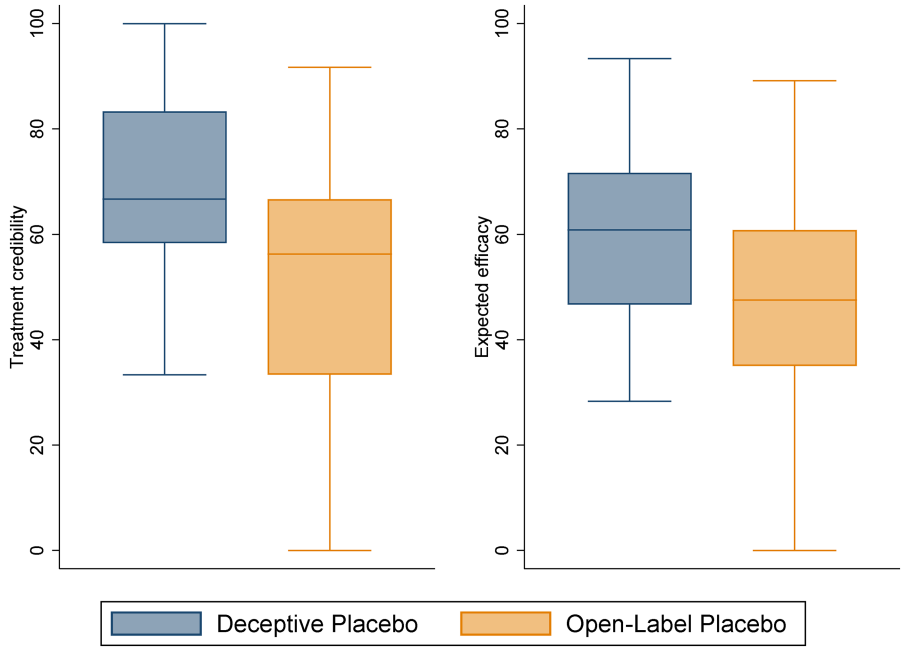

FIGURE 4 Credibility and expectancy of treatment effect out of 100 just after application of the placebo cream.

different from initial hypothesis of generating placebo effects (superiority to no treatment) on both occasions. One way this could be explained is that participants may have been influenced by their a priori experiences. As such, during the second CPT, they may have re-assessed their pain intensity due to the difference between the first and second CPT (filling a VAS every 5 s or waiting for the time to be up). Once this was done, they noted a change during the third CPT: participants to whom we took away the treatment worsened, and participants to whom we added a placebo treatment improved. This could be in line with findings from Colloca et al. (2020) showing that a priori experiences modulate response to DP more than expectancy (Colloca et al., 2020; Colloca & Benedetti, 2006). The role of expectation has also been questioned for OLPs (Schaefer et al., 2018). This would also be consistent with the credibility and expectancy scores we measured for each treatment. Indeed, although DP and OLP were non-inferior, OLP showed significantly lower expectancy and credibility. Haas et al. (2020) also showed lower credibility and expectancy scores for OLP rather than DP using the same questionnaire (Haas et al., 2020).

Another point of discussion in our findings pertains to the rational content and format we choose to use. This is an important point to discuss when administering OLPs as Blease et al. (2019) propose it is unclear what explains the potential effect of OLPs (e.g. whether it the rationale, the pill or the doctor–patient relationship, or some combination thereof) (Blease et al., 2019). OLPs in clinical trials have been administered with highly variable rationales (von Wernsdorff et al., 2021). In contrast, some authors have suggested key information to include in the OLP rationale (Heiss et al., 2021). In our study, we chose a video format with mostly informative content. Other studies have used appeals to other patient's experience, boosting hopefulness or increasing expectations (Schaefer et al., 2018). Our results show that our participants may indeed have benefited from the educational video to improve their knowledge about placebo. However, a video rationale may have also taken away part of the doctor–patient relationship during administration. The use of this format allowed for a better replicability of our findings, to limit variability in future research and, if future results call for it, an easier clinical application.

4.2 | Strengths and limitations

Several methodological considerations should be discussed in this study. Some have already been reviewed in a separate article (Druart et al., 2020). First, the use of a cross-over nested in the parallel trial allowed us to look at inter-individual comparison as well as consider how a priori experiences changed pain experience. This also allowed us to have a no-treatment condition in addition to our two placebo conditions without increasing the number of subjects needed allowing for additional power. However, in hindsight, we planned for a cross-over because we hypothesized there would only be a small interaction of time on the treatment effect. This was not the case as the interaction of time in our linear regression was important. Therefore, in this study, the interpretation

of the difference, or lack thereof, between placebo conditions and no-treatment requires caution because of the discordance. If we had planned for three groups (OLP vs. DP vs. NT), this would have facilitated the interpretation. However, other studies have already shown the effect of OLPs versus no treatment (von Wernsdorff et al., 2021). Another major point of discussion, common to most OLP studies to date, is the lack of patient and therapist blinding. Indeed, due to the nature of the treatment, patients are aware of the inertness of what they are receiving. However, we could have improved the blinding in the administration of the OLP and DP (Blease et al., 2019). Our study did not do better than other studies in this regard. Due to this, results must always be interpreted with caution as they are difficult to distinguish from reporting bias from patients or investigators. Relatedly, we cannot rule out whether the effectiveness of DP or OLP was owed to placebo effects proper, or participant responder biases. As has been argued, conflating placebo responses with placebo effects means researchers often tend to inflate the size of placebo effects (Blease et al., 2019, 2023; Kirsch, 2013).

Underpowered studies are an issue in medical research and placebo research is no exception (Blease et al., 2023). In our study, we had initially planned for 126 subjects considering our 10 mm non-inferiority margin and an estimated standard deviation of 21.6 mm. Due to the pandemic, we were unable to recruit many participants. However, we hypothesize this did not affect the results for two main reasons. First, the observed standard deviation was significantly lower in our study than anticipated. This is probably thanks to our first CPT functioning as calibration. This made our population more homogeneous in pain ratings and increased our power. Second, as our results show H0 is unlikely on the primary outcome, there would be no type II risk. We accept H1 with a type I risk. Thus, although our study did not recruit the initially planned number of subjects, we believe this had no impact on our power or our interpretation of the primary outcome. Regarding the secondary outcomes, in which the testing would have indicated H0 to be more likely, the study could be underpowered. For these reasons and those linked to the study design, the results comparing placebo conditions to no treatments should be met with caution.

Lastly, for the knowledge questionnaire, we could add that in our study design one outcome was not balanced for both groups because structural equivalence was not perfect (Blease et al., 2019; Locher et al., 2018). The knowledge questionnaire was completed twice by our OLP group and only once by our DP group. Although there was at least an hour between both completions for the OLP group, there is no guarantee that this did not bias their responses as we have not checked the test–retest reliability of our questionnaire.

4.3 | Implications

Nevertheless, these findings have two major implications. First, amidst a replication crisis in medical and particularly in placebo research, replication and confirmation of findings are important contributions to scientific knowledge (Blease et al., 2023). Second, our findings bring serious doubts to the pertinence and justification of the currently widespread clinical uses of DPs (Linde et al., 2018). Indeed, several studies have found no superiority between a DP and a well-explained OLP and we add to this by demonstrating non-inferiority. If placebo treatments are to be used, OLPs should be favoured over DPs. However, looking at how OLPs should best be administered is still to be determined. For example, in our nested qualitative study (Druart et al., 2023), participants suggested the following conditions to administer OLPs in the clinical setting: a convincing rationale, time to discuss this treatment option with their healthcare provider, proven effectiveness compared to DP, appropriateness regarding the clinical situation and being included in the decision to take an OLP.

4.4 | Future studies

Moving further, research comparing the effect of DPs and OLPs with patients instead of healthy subjects is needed before any clinical applications are suggested. It is reasonable to assume that findings with healthy subjects will be similar among patients as placebo effects seem to have a smaller effect size for healthy subjects rather than patients (Forsberg et al., 2017). Another important area of research that needs to be addressed is to better understand what we are measuring in no-treatment conditions. As such, in our study, it seems that reverting to a no-treatment condition after having been given a placebo treatment could have worsened the pain similar to a nocebo effect. Furukawa et al. (2014) suggest considering some no-treatment conditions such as waiting lists as nocebo effects and thus poor tools to distinguish the placebo effect from the placebo response (Furukawa et al., 2014). Finally, further research is needed to explore the acceptability of placebos, including OLPs, among patients (Blease, 2019; Druart et al., 2023).

AUTHOR CONTRIBUTIONS

L. Druart: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, visualization, writing original draft, writing review and editing. **S. E. Graham Longsworth**: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, writing original draft, writing review and editing.

H. Terrisse: Conceptualization, data curation, formal analysis, methodology, software, visualization, writing review and editing. **C. Locher**: Methodology, supervision, validation, writing review and editing. **C. Blease**: Methodology, supervision, validation, writing review and editing. **C. Rolland**: Methodology, project administration, resources, writing review and editing. **N. Pinsault**: Conceptualization, formal analysis, investigation, methodology, project administration, resources, supervision, validation, writing review and editing.

ACKNOWLEDGEMENTS

The authors would like to thank: Pascale Calabrese and the PRETA team for their invaluable help in the technical setup of the physiological measurements in this trial as well as their material support. SCREEN for their material assistance in the realization of this trial. Maia Dolgopoloff for her advice regarding methodology and statistical analysis in the first steps of this project. Jean-Luc Bosson for his medical advice during this project. Matthias Bieber, Cécile Foltzer and Emilie Bailly-Bassin for their assistance during their respective internships.

FUNDING INFORMATION

This trial was funded by the APICIL foundation (grant no. 1251.18) which had no role in the design of the study; in the collection, analysis or interpretation of data; or in the writing of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing or conflicts of interest.

DATA AVAILABILITY STATEMENT

Data and code for analysis are available upon request to the corresponding author.

TRIAL REGISTRATION

French National Ethical Committee no. 2017-A01643-50 & ClinicalTrials no. NCT03934138.

ORCID

L. Druart https://orcid.org/0000-0002-5060-6620
H. Terrisse https://orcid.org/0000-0001-8239-1903
C. Locher https://orcid.org/0000-0002-9660-0590
C. Blease https://orcid.org/0000-0002-0205-1165
C. Rolland https://orcid.org/0000-0002-8311-2475
N. Pinsault https://orcid.org/0000-0001-5034-0802

REFERENCES

Annoni, M., & Miller, F. G. (2014). Placebos in clinical practice: An ethical overview. *Douleur et Analgésie*, *27*, 215–220.

Aulas, J. J., & Rosner, I. (2003). Efficacy of a non-blind placebo prescription. *Encephale*, *29*, 68–71.

Barnhill, A. (2011). What it takes to defend deceptive placebo use. *Kennedy Institute of Ethics Journal*, *21*, 219–250.

Bishop, F. L., Aizlewood, L., & Adams, A. E. M. (2014). When and why placebo-prescribing is acceptable and unacceptable: A focus group study of patients' views. *PLoS One*, *9*, e101822.

Bishop, F. L., Jacobson, E. E., Shaw, J., & Kaptchuk, T. J. (2012). Participants' experiences of being debriefed to placebo allocation in a clinical trial. *Qualitative Health Research*, *22*, 1138–1149.

Blease, C., Colagiuri, B., & Locher, C. (2023). Replication crisis and placebo studies: Rebooting the bioethical debate. *Journal of Medical Ethics*, *49*, 663–669.

Blease, C. R. (2019). The role of placebos in family medicine: Implications of evidence and ethics for general practitioners. *Australian Journal of General Practice*, *48*, 700–705.

Blease, C. R., Bernstein, M. H., & Locher, C. (2019). Open-label placebo clinical trials: Is it the rationale, the interaction or the pill? *BMJ Evidence-Based Medicine*, *25*(5), 159–165.

Charlesworth, J. E. G., Petkovic, G., Kelley, J. M., Hunter, M., Onakpoya, I., Roberts, N., Miller, F. G., & Howick, J. (2017). Effects of placebos without deception compared with no treatment: A systematic review and meta-analysis. *Journal of Evidence-Based Medicine*, *10*, 97–107.

Colloca, L., Akintola, T., Haycock, N. R., Blasini, M., Thomas, S., Phillips, J., Corsi, N., Schenk, L. A., & Wang, Y. (2020). Prior therapeutic experiences, not expectation ratings, predict placebo effects: An experimental study in chronic pain and healthy participants. *Psychotherapy and Psychosomatics*, *89*, 1–8.

Colloca, L., & Benedetti, F. (2006). How prior experience shapes placebo analgesia. *Pain*, *124*, 126–133.

Coste, J., Tarquinio, C., Rouquette, A., Montel, S., & Pouchot, J. (2020). Cross-cultural adaptation and validation of the French version of the credibility/expectancy questionnaire. Further insights into the measured concepts and their relationships. *Psychologie Française*, *65*, 173–184.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*, 73–86.

Disley, N., Kola-Palmer, S., & Retzler, C. (2021). A comparison of open-label and deceptive placebo analgesia in a healthy sample. *Journal of Psychosomatic Research*, *140*, 110298.

Druart, L., Graham Longsworth, S., Rolland, C., Dolgopoloff, M., Terrisse, H., Bosson, J. L., & Pinsault, N. (2020). Can an open-label placebo be as effective as a deceptive placebo? Methodological considerations of a study protocol. *Medicines (Basel)*, *7*, 3.

Druart, L., Vauthrin, O., Pinsault, N., Locher, C., & Blease, C. (2023). 'It's not my greengrocer, it's someone from the medical profession': A qualitative study regarding acceptability of deceptive and open-label placebo prescribing in France. *British Journal of Health Psychology*, *28*, 273–290.

Forsberg, J. T., Martinussen, M., & Flaten, M. A. (2017). The placebo analgesic effect in healthy individuals and patients: A meta-analysis. *Psychosomatic Medicine*, *79*, 388–394.

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, *130*, 181–192.

Haas, J. W., Rief, W., & Doering, B. K. (2020). Open-label placebo treatment: Outcome expectations and general acceptance in the lay population. *International Journal of Behavioral Medicine*, *28*, 444–454.

Handwerker, H. O., & Kobal, G. (1993). Psychophysiology of experimentally induced pain. *Physiological Reviews*, *73*, 639–671.

Heiss, U., Rosenfield, M., & Bernstein, M. (2021). Can the open-label placebo rationale be optimized? *Frontiers in Pain Research*, *2*, 734882.

Hughes, J., Greville-Harris, M., Graham, C. A., Lewith, G., White, P., & Bishop, F. L. (2017). What trial participants need to be told about placebo effects to give informed consent: A survey to establish existing knowledge among patients with back pain. *Journal of Medical Ethics*, *43*, 867–870.

Kaptchuk, T. J., & Miller, F. G. (2018). Open label placebo: Can honestly prescribed placebos evoke meaningful therapeutic benefits? *BMJ*, *363*, k3889.

Kirsch, I. (2013). The placebo effect revisited: Lessons learned to date. *Complementary Therapies in Medicine*, *21*, 102–104.

Kube, T., Rief, W., Vivell, M. B., Schäfer, N. L., Vermillion, T., Körfer, K., & Glombiewski, J. A. (2020). Deceptive and nondeceptive placebos to reduce pain: An experimental study in healthy individuals. *The Clinical Journal of Pain*, *36*, 68–79.

Linde, K., Atmann, O., Meissner, K., Schneider, A., Meister, R., Kriston, L., & Werner, C. (2018). How often do general practitioners use placebos and non-specific interventions? Systematic review and meta-analysis of surveys. *PLoS One*, *13*, e0202211.

Locher, C., Frey Nascimento, A., Kirsch, I., Kossowsky, J., Meyer, A., & Gaab, J. (2017). Is the rationale more important than deception? A randomized controlled trial of open-label placebo analgesia. *Pain*, *158*, 2320–2328.

Locher, C., Gaab, J., & Blease, C. (2018). When a placebo is not a placebo: Problems and solutions to the gold standard in psychotherapy research. *Frontiers in Psychology*, *9*, 2317.

Mertens, V. C., Moser, A., Verbunt, J., Smeets, R., & Goossens, M. (2017). Content validity of the credibility and expectancy questionnaire in a pain rehabilitation setting. *Pain Practice*, *17*, 902–913.

Mundt, J. M., Roditi, D., & Robinson, M. E. (2017). A comparison of deceptive and non-deceptive placebo analgesia: Efficacy and ethical consequences. *Annals of Behavioral Medicine*, *51*, 307–315.

Myles, P. S., Myles, D. B., Galagher, W., Boyd, D., Chew, C., MacDonald, N., & Dennis, A. (2017). Measuring acute postoperative pain using the visual analog scale: The minimal clinically important difference and patient acceptable symptom state. *British Journal of Anaesthesia*, *118*, 424–429.

Rafi, Z., & Greenland, S. (2020). Semantic and cognitive tools to aid statistical science: Replace confidence and significance by compatibility and surprise. *BMC Medical Research Methodology*, *20*, 244.

Rubin, M. (2016). The perceived awareness of the research hypothesis scale: Assessing the influence of demand characteristics.

Schaefer, M., Sahin, T., & Berstecher, B. (2018). Why do open-label placebos work? A randomized controlled trial of an open-label placebo induction with and without extended information about the placebo effect in allergic rhinitis. *PLoS One*, *13*, e0192758.

Streff, A., Kuehl, L. K., Michaux, G., & Anton, F. (2010). Differential physiological effects during tonic painful hand immersion tests using hot and ice water. *European Journal of Pain*, *14*, 266–272.

Tang, C. W., & Colagiuri, B. (2013). Can an educational handout enhance placebo analgesia for experimentally-induced pain? *PLoS One*, *8*, e77544.

von Wernsdorff, M., Loef, M., Tuschen-Caffier, B., & Schmidt, S. (2021). Effects of open-label placebos in clinical trials: A systematic review and meta-analysis. *Scientific Reports*, *11*, 3855.

Wiens, B. L., & Zhao, W. (2007). The role of intention to treat in analysis of noninferiority studies. *Clinical Trials*, *4*, 286–291.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Druart, L., Graham Longsworth, S. E., Terrisse, H., Locher, C., Blease, C., Rolland, C., & Pinsault, N. (2023). If only they knew! A non-inferiority randomized controlled trial comparing deceptive and open-label placebo in healthy individuals. *European Journal of Pain*, *00*, 1–11. https://doi.org/10.1002/ejp.2204