Universität Basel
Fakultät für Psychologie

The Therapeutic Oxymoron: Exploring the Mechanisms of Open-Label Placebo

Inauguraldissertation zur Erlangung der Würde einer Doktorin der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

Sarah Bürgler

aus Meggen, Luzern

Basel, 2023

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. Jens Gaab

Prof. Dr. Claudia Carvalho

Datum des Doktoratsexamen: 04.07.2023

Prof. Dr. Jens Gaab

Dekan der Fakultät für Psychologie

Universität Basel

Fakultät für Psychologie

Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Des Weiteren versichere ich, sämtliche Textpassagen, die unter Zuhilfenahme KI-gestützter Programme verfasst wurden, entsprechend unter «auxiliary means» gekennzeichnet sowie mit einem Hinweis auf das verwendete KI-gestützte Programm versehen zu haben. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

- **Buergler, S**.*, Sezer, D.*, Gaab, J., & Locher, C. (2023). The role of population, expectation, modality, and comparator on open-label placebo effects: A network meta-analysis. *Scientific Reports (in Review)*
  * shared first authorship
- **Buergler, S**., Sezer, D., Bagge, N., Kirsch, I. Locher, C., Carvalho, C., & Gaab, J. (2023). Imaginary pills and open-label placebos can reduce test anxiety by means of placebo mechanisms. *Scientific Reports, 13*(1), 2624. https://doi.org/10.1038/s41598-023-29624-7
- **Buergler, S**., Sezer, D., Busch, A., Enzmann, M., Bakis, B., Locher, C., Bagge, N., Kirsch, I., Carvalho, C., & Gaab, J. (2023). A qualitative study on imaginary pills and open-label placebos in test anxiety. *PLOS ONE (in Review)*

Universität Basel
Fakultät für Psychologie

Spezifizierung des eigenen Forschungsbeitrags zu den Manuskripten:

1. Eigener Beitrag nach CRediT[1]:

☒ Conceptualization	☒ Data curation	☒ Formal Analysis
☐ Funding acquisition	☒ Investigation	☒ Methodology
☒ Project administration	☐ Resources	☒ Software
☒ Supervision	☒ Validation	☒ Visualization
☒ Writing – original draft		
☒ Writing – review & editing		

Das Manuskript wurde bisher für keine anderen Qualifikationsarbeiten eingereicht.

2. Eigener Beitrag nach CRediT[1]:

☒ Conceptualization	☒ Data curation	☒ Formal Analysis
☐ Funding acquisition	☒ Investigation	☒ Methodology
☒ Project administration	☐ Resources	☒ Software
☒ Supervision	☒ Validation	☒ Visualization
☒ Writing – original draft		
☒ Writing – review & editing		

Das Manuskript wurde bisher für keine anderen Qualifikationsarbeiten eingereicht.

3. Eigener Beitrag nach CRediT[1]:

☒ Conceptualization	☒ Data curation	☒ Formal Analysis
☐ Funding acquisition	☒ Investigation	☒ Methodology
☒ Project administration	☐ Resources	☐ Software
☒ Supervision	☒ Validation	☒ Visualization
☒ Writing – original draft		
☒ Writing – review & editing		

Das Manuskript wurde bisher für keine anderen Qualifikationsarbeiten eingereicht.

[1] https://casrai.org/credit/

Universität Basel

Fakultät für Psychologie

Open-Science Aspekte der Manuskripte:

1. Preregistration: ☒ ja ☐ nein
   Open-Access-Publikation: ☒ ja ☐ nein
   Open-Access-Data/Analyse: ☐ ja ☒ nein
   Ort/URL der Daten und Analysen: Meta-Analyse, Daten in einzelnen Studien

2. Preregistration: ☒ ja ☐ nein
   Open-Access-Publikation: ☒ ja ☐ nein
   Open-Access-Data/Analyse: ☐ ja ☒ nein

3. Preregistration: ☒ ja ☐ nein
   Open-Access-Publikation: ☒ ja ☐ nein
   Open-Access-Data/Analyse: ☒ ja ☐ nein
   Ort/URL der Daten und Analysen: tbd

Ort, Datum Luzern, 27. April 2023

Signatur 

Vorname Nachname Sarah Bürgler

Acknowledgments

To all those who have supported, accompanied, and inspired me on this journey toward completing my dissertation – a heartfelt thank you!

My first thanks go to Jens Gaab for his unwavering support, encouragement, and confidence in my abilities. His inspiring and energetic as well as personal, honest, and collaborative approach to working together has been extraordinary, and I will always cherish the countless delicious Tuesday lunches we shared. I am also very grateful to Claudia Carvalho for taking on the important role of second supervisor in my dissertation.

I also want to thank the entire team for the honor and pleasure of working alongside you. The inspiration, support, advice, and enjoyment that I experienced while collaborating with all of you were priceless. My special thanks go to Cosima Locher for her extraordinary and inspirational support and guidance in this endeavor.

My sincere thanks also go to Niels Bagge and Irving Kirsch for their invaluable contributions to this project. Their ideas and inputs have been incredibly helpful and enriching, and their assistance and guidance have been instrumental in the direction and success of this venture.

My family and friends also deserve a big thank you for their support, attentive listening and belief in me. Their encouragement has been a constant source of strength and energy that has propelled me forward.

Importantly, I would like to express my heartfelt appreciation to Philipp Keller for the countless ways he has brightened and lightened my life and work. I am grateful for his steadfast companionship, genuine interest and curiosity in my work, his thought-provoking questions, and his humorous and encouraging manner, which motivates and inspires me.

Last but certainly not least, I want to express my profound gratitude to Dilan Sezer, my faithful companion, fellow student, close friend, and co-worker. Without her boundless energy, sharp intellect, and inspiring authenticity, this doctoral journey would have been a much rockier one. She has not only accompanied and supported me throughout the process, but also motivated, encouraged and contributed to both the content and the emotional landscape of this journey.

Table of contents

Abstract

Placebos have emancipated themselves from earlier, often unethical or merely research-based practices, toward an open and thereby ethical administration. Despite a plethora of efficacy studies and nearly two decades of empirical research on open-label placebos (OLPs), little to nothing is known about the circumstances under which they realize their potential. While the underlying processes of deceptive placebos are fairly well understood, these findings cannot simply be transferred to the field of OLPs, given that the two areas most probably operate through distinct mechanisms. To address this research gap, this thesis aims to identify potential explanatory forces behind OLPs.

To do the unexploredness justice, a multi-method approach was adopted to shed light on the topic: A network meta-analysis of OLPs (study 1) supports the notion that OLPs can be an efficacious intervention for nonclinical and clinical populations. In particular, this study highlights the importance of inducing positive treatment expectancies, for example, by providing a treatment rationale. The mere administration of a pill without embedding it in a narrative apparently is insufficient, implying that the pill itself might not be a necessary component for OLP effects. This finding was confirmed in a three-week randomized controlled trial (study 2) in which OLPs and the sheer imagination of a pill intake both led to significantly lower test anxiety in students compared to a control group. The presence of a placebo effect solely through means of imagination implies that the daily therapeutic ritual itself evoked beneficial effects. Although these two studies provide valuable insights into OLP mechanisms, they leave the question unanswered as to which participant-related factors play a role in distinct OLP effects. Therefore, a qualitative analysis (study 3) was conducted, in which a central finding was the great extent to what an individual's desire for relief determines the magnitude of the OLP effect.

OLP effects can therefore be understood as a relief of symptoms that occurs within the framework of a plausible narrative and a ritual that corresponds to the individual's desire for relief when seeking treatment. Taken together, these three studies suggest, that the words and rituals, as well as the patient's desire for relief, are key determinants responsible for the effect of OLP. A better understanding of the relative weight of each of these components can inform research regarding underlying mechanisms. This, in turn, can assist clinicians in considering or utilizing these mechanisms more frequently in their clinical practice and, where appropriate, communicate the therapeutic value of each element openly.

Introduction

Placebos have moved beyond being suggestive procedures of earlier centuries, methodological research tools, or unethically used treatments for difficult patients. Over the past years, a branch of research has started to flourish in which placebos are administered openly and, in this sense, ethically (e.g., Carvalho et al., 2016; Kaptchuk et al., 2010). This contradictory or paradoxical treatment, in fact, an oxymoron (open placebo), allows the often large and clinically meaningful effects – in some cases equal or nearly equal to active treatments (e.g., Kirsch, 2019) – to be exploited in a wide range of conditions. This, while preserving the patient's autonomy, the physician's honesty, and the trust in the therapeutic relationship (Annoni, 2018; Gaab et al., 2016). Indeed, the potential of these so-called open-label placebos (OLPs) has been well demonstrated by a number of studies: Meta-analyses found medium to large-sized effects in comparison to control conditions in clinical populations (Charlesworth et al., 2017; von Wernsdorff et al., 2021) and medium-sized effects in self-reported outcomes in experimental nonclinical samples (Spille et al., 2023). However, despite their efficacy in various disorders and populations and their 16-year empirical research tenure with up to 50 experimental and clinical studies, little to nothing is known about why, how and under what circumstances OLPs are efficacious.

Deceptive placebo mechanisms

Toward a better understanding of potential OLP determinants, it is worthwhile to turn the arc back to deceptive placebos. Here, various complex and multifaceted mechanisms, involving both psychological and physiological processes, are discussed. The most popular and best-studied mechanism is treatment-related expectancy (e.g., Kirsch, 2018; Meissner & Linde, 2018; Schedlowski et al., 2015): When patients believe they are receiving an effective treatment, they may experience a sense of hope and optimism, which can, in turn, reduce anxiety and stress (Meissner & Linde, 2018). Thereby, not only verbal information but also nonverbal cues or contextual factors (such as dosage, administration route, therapeutic encounter, and costs) can influence the extent of expectancies surrounding medical treatments and, consequently, the magnitude of the placebo effect (de Craen et al., 2000; Finniss et al., 2010; Kam-Hansen et al., 2014; Meissner & Linde, 2018). Besides that, more non-conscious forms of expectancy, such as classical, observational, operant conditioning (Colloca & Benedetti, 2009; Colloca & Miller, 2011; Montgomery & Kirsch, 1997), and related to this social learning (Chen et al., 2019), establish another key mechanism of deceptive placebo effects. Aside from these

two fundamental determinants, other patient and treatment-related factors play non-negligible roles: personal treatment choice (Tang et al., 2022), gender and personality (Kelley et al., 2009) as well as adherence (Simpson et al., 2006). These processes are accompanied by several neurobiological underpinnings, including specific brain regions and pathways that modulate the body's stress response and activate the release of endogenous opioids as well as other neurotransmitters (Benedetti et al., 2022). Nonetheless, it is crucial to bear in mind that these patient and treatment-related factors can equally influence expectancies and thus realize their effects based on this mechanism.

Differences in open-label and deceptive placebo mechanisms

While the underlying processes and mechanisms in deceptive placebos are well-researched and broadly understood, the findings found therein cannot unquestioningly be carried over to the field of OLPs. Ted Kaptchuk, the researcher who kickstarted this branch of research with his OLP study on irritable bowel syndrome (Kaptchuk et al., 2010; apart from an early but uncontrolled study by Park & Covi, 1965 and a study using OLP as a dose extender by Sandler & Bodfish, 2008) argues that in the absence of deception, OLPs most likely have their distinct dynamics (Kaptchuk, 2018). In other words, as the long-believed main effect-driving element of a placebos, i.e., "deception," and the associated expectancies of receiving a verum are no longer present, unsurprisingly, other psychophysiological processes might be at work. While the prevailing research no longer considers deception to be a necessary element for placebos effects, the inner workings of OLP yet remain unclear.

That different predictor are involved in deceptive and OLPs is supported by a few studies which directly compare the two placebo intervention types: Dispositional optimism has been shown to be associated with deceptive but not OLP effects (Locher et al., 2019), and symptom-specific anxiety (i.e., greater self-efficacy regarding the potential to influence or manage one's symptoms) as well as pain catastrophizing (i.e., qualities of hopelessness and helplessness) were significant predictors of response for OLPs but not for deceptive placebos (Ballou et al., 2022). The hypothesis put forward by the authors suggests that the ability to think flexibly or conditionally (i.e., "If I do X, my symptoms will improve") may be a reliable indicator of the likelihood of improvement with the OLP treatment. In contrast, rigid thinking and a sense of helplessness when managing symptoms ("No matter what I do, I will feel bad") could hinder response to the OLP treatment. Similarly, qualitative research found that OLP elicited more self-examination, ambivalent feelings, and active

engagement than deceptive placebos, whereas OLP participants were more hesitant to attribute symptom improvement to their treatment (Haas, Ongaro, et al., 2022). Also, unlike deceptive placebos, OLPs appear not to be dose-dependent and influenced by adherence (El Brihi et al., 2019; Simpson et al., 2006). Genetic evidence supports the notion that there are variations in clinical and psychological characteristics between individuals who react to open-label placebos and those who react to deceptive placebos. This is demonstrated by the fact that a distinct polymorphism predicts the response to open-label placebos, as opposed to deceptive placebos (Hoenemeyer et al., 2021; Zhou et al., 2019). Similarly, when it comes to neural underpinnings, differences concerning a missing engagement of prefrontal brain regions in OLPs suggest that treatment expectancies may play a less prominent role than in deceptive placebos (Schaefer et al., 2023). In summary, there is compelling evidence reaching from genetics to personality and neural involvement indicating that different mechanisms are at play.

Potential open-label placebo mechanisms

Apart from the comparison with deceptive placebo studies, a handful of investigations of clinical and nonclinical nature have examined individual determinants and contextual factors that may bear influences on the likelihood and strength of OLP effects: However, personality traits did not appear to be related to outcome (Zhou et al., 2019), nor did different lengths of placebo rationales (Schneider et al., 2020), rationale styles (i.e., personal vs. scientific in Friehs et al., 2022; expectancy vs. hope in Kube et al., 2020) or different rationale contents (Olliges et al., 2022). Accordingly, participants' suggestibility (i.e., their ability to imagine a situation) or further personality variables (e.g., neuroticism, extraversion, openness, or positive attitude towards complementary or alternative/conventional medicine) were not significantly associated with outcomes (Bräscher et al., 2022). Conversely, the effect of OLP appeared to be more pronounced in females and participants with more severe baseline pain, which in turn was influenced by age, i.e., younger patients achieving better treatment outcomes (Flowers et al., 2021). Similarly, only depressed patients under 65 whose first onset occurred before age 50 showed OLP effects (Nitzan et al., 2020). This finding, however, conflicts with others showing that older patients also appear to profit from OLP treatments (Olliges et al., 2022). Further, while a stronger general belief in the benefit of medication (Kleine-Borgmann et al., 2021) and a firmer belief in the power of placebos improved OLP effects in several studies (Leibowitz et al., 2019; Schaefer et al., 2021), another study did not find a significant relationship between beliefs and outcome measures of emotional distress

(Guevarra et al., 2020). Besides the influence of person-related factors, the patient-provider relationship is associated with different magnitudes of OLP effects (Leibowitz et al., 2019): however, OLP studies commonly control for this, ensuring that participants across conditions have the same quality and quantity of interaction with a caring physician. Specific designs with different ways to encounter the participant (e.g., Kube et al., 2021; Leibowitz et al., 2019; Rathschlag & Klatt, 2021) provide insights into the role of this factor: Even in the absence of an overall treatment effect, Kube and colleagues (2021), for instance, found that participants receiving OLP benefitted more from the treatment if a warm, empathic provider delivered it. Similarly, the practitioners' characteristics (Lee et al., 2022) influenced treatment outcomes.

Importantly, and as part of the patient-provider interaction, the OLP administration is usually accompanied by a treatment rationale. The most commonly used treatment protocol for OLPs is modeled after the one used in the first OLP study by Kaptchuk et al. (2010) in order to exploit or maintain, through the means of verbal suggestion, a person's treatment expectancy and belief that the open administration of "nothing" may have lowered. And indeed, there is evidence of the importance of the rationale in OLP effects (Klinger et al., 2017; Locher et al., 2017; Rathschlag & Klatt, 2021). Yet, meta-analytical systematic investigations of the role of the treatment rationale remain limited to the experimental context (Spille et al., 2023). Interestingly, most authors of OLP studies equate verbal suggestions with the induction of treatment expectancies. Accordingly, in about half of the existing OLP studies, expectancies were assessed (e.g., as a manipulation check). However, the limited available evidence to date is inconsistent: In fact, of a total of 49 OLP studies screened, 24 assessed expectancies, and only 6 reported a positive correlation with the outcome. Even in these studies, however, the results do not always appear to be entirely clear-cut but somewhat dependent on the type of expectancy measurement, the intervention, or the outcomes (Bräscher et al., 2022: correlation only with expectancies measured shortly before the placebo phase, not with those at baseline; El Brihi et al., 2019: not for sleep quality; Klinger et al., 2017: only for conditioned OLP; Lee et al., 2022: only for sham acupuncture and not for OLP pills; Meeuwis et al., 2021; Meeuwis et al., 2019: not for all outcomes). In many other OLP studies, no association was detected, or only expectancies in the deceptive placebo group correlated with symptom improvement (Lembo et al., 2021), differed from the control group (Locher et al., 2017), or differences in baseline expectancies between groups were found, but not associated with outcomes (e.g., Kube et al., 2020; Pan et al., 2020).

When reopening the discussion to all placebo types (i.e., including deceptive placebos), the discrepancies between treatment expectancies and outcomes may be due to the well-known finding that people often do not have direct access to their internal states, complicating an accurate representation of their expectancies (Nisbett & Wilson, 1977). Along these lines, research has indicated that (deceptive) placebos have been shown to work through unconscious aspects of expectancy. This implies that participants might encounter a type of expectancy that operates beyond their conscious awareness and that cannot be reflected in self-reported measures (Jensen et al., 2012): If the placebo effect can be produced without conscious expectancy, it is plausible that classical conditioning is involved, i.e., a process in which a stimulus (the active drug) is paired with an initially neutral stimulus (the act of taking a pill), and over time this ritual of pill-taking alone may trigger effects. In this sense, the pill serves as a symbol of healing, generating symptomatic improvement on its own, even in the absence of active ingredients (Blease et al., 2020). However, while several studies have successfully implemented conditioned OLP paradigms (e.g., Flowers et al., 2021; Klinger et al., 2017; Morales-Quezada et al., 2020; Mundt et al., 2017; Sandler et al., 2010), both the conditioning and the expectancy theories are criticized as insufficient in explaining OLP effects. Critical voices emphasize that, especially in clinical samples, patients enter a study with a history of treatment failure. In other words, participants often have low or even negative expectancies toward new treatments (Ballou et al., 2017; Kaptchuk, 2018).

Instead, hope has found its way into the discussion about potential OLP mechanisms. Qualitative studies showed that a number of patients – rather than having positive treatment expectancy – were hopeful for a reduction in symptoms and had an attitude of "let’s see what happens” (Haas, Ongaro, et al., 2022; Kaptchuk et al., 2009; Pan et al., 2022). As a result, it has been postulated that hope might be a critical element contributing to the response to OLPs beyond mere expectancy (Ballou et al., 2017). Kaptchuk defines hope as a “life jacket” against despair, protecting patients from potential harm or disappointment through inappropriately strong expectancies (Kaptchuk, 2018). Connected to the construct of hope is novelty as well as uncertainty or skepticism (Ballou et al., 2017), the latter being a common theme in qualitative OLP research (e.g., in the form of participants’ ambivalence, being both open and skeptical; Bernstein et al., 2021; Haas, Ongaro, et al., 2022; Hruschak et al., 2022; Locher et al., 2021).

In line with this, more sophisticated explanatory approaches within the framework of Bayesian models have been proposed to understand the positive effects of OLP, namely

prediction error processing (PEP) and embodied cognition (Kaptchuk, 2018). Put simply, PEP is a cognitive process occurring when the predicted input (cognitively expected events) and actual input (perceived events) mismatch. In OLPs, such ambiguous messages are nested ("This is an empty pill, so it will not help me" vs. "This inert pill may still be helpful as there is a healing setting and someone prescribing me it"). The medical ritual involved causes the brain to interpret even small bodily changes as a result of healing and to experience relief accordingly to fulfill the prediction and minimize the error (Ongaro & Kaptchuk, 2019). In essence, based on PEP, the placebo effect is predominantly triggered by actions and, to a lesser extent, if at all, by thoughts (Kaptchuk, 2018). This kind of prediction process is related to embodied cognition, i.e., the idea that bodily experiences such as the act of pill intake and the sensory experience of ingesting a substance can itself influence our expectancies and perceptions of symptoms (Ballou et al., 2017).

Aim of the thesis

Even though a plethora of studies sought to unpack the mechanisms underlying the treatment efficacy of OLPs, the research is yet to move in a more fruitful direction. Taken together, the currently available data indicate that the effects of OLPs can be affected by contextual and individual factors, although not always in the same way as in deceptive placebos. Also, the factors considered important for OLPs efficacy show inconclusive results. Whereas these findings offer some intriguing initial insights, the relatively limited amount of research on OLPs compared to deceptive placebos and the inconsistent results leave a substantial gap in our understanding of OLP effects and their mechanisms. From a scientific point of view, unpacking the most critical components of OLP effects can lead to a better understanding of the intervention and thereby contribute considerably to the quantification of the combined effect of different plausible OLP mechanisms. From a clinical point of view, improved knowledge would help educate clinicians about the necessary components of OLPs so that they might leverage those forces in their clinical practice and – where necessary – communicate the distinct components' therapeutic value openly to patients. Lacking such mechanistic knowledge, clinicians might adopt a mere "medical model" and thus, for instance, operate under the assumption that the prescription of a pill is sufficient to produce placebo effects (Blease et al., 2020).

To better grasp the fundamental principles of OLPs, this cumulative thesis strives to address this research gap by exploring potential mechanisms at play. Therefore, a multi-method approach was used to illuminate the topic from different angles. First, a network

meta-analyses (study 1) will provide insight into the role of different moderators (i.e., population, expectation, modality, and comparator) on OLP effects. Second, a randomized controlled trial (study 2) takes the OLP concept one step further by examining the necessity of a physical pill to induce beneficial effects. Finally, a qualitative analysis (study 3) includes patients' idiosyncratic perspectives to shed light on their treatment experience.

Methods and Results

Each of the three studies – individually and in combination – attempts to help elaborate key components of the OLP treatment regime. To do so, different methodological approaches were employed to explore the topic at different levels of abstraction, from large-scale (i.e., network) to small-scale (i.e., qualitative analysis). In the following, no detailed description of the methods and results for the individual studies is provided (full descriptions in the main publications to be found in the appendices); instead, the importance and advantage of the particular methods used are described. Furthermore, key findings relevant to this thesis are presented.

Network meta-analyses (study 1)

Although three meta-analyses have already been conducted on OLPs, several important aspects could not be adequately addressed (Charlesworth et al., 2017; Spille et al., 2023; von Wernsdorff et al., 2021). First, these analyses only considered one type of population (either clinical or nonclinical), which may limit the comparability of effect sizes across these two samples. Second, due to methodological limits, these pair-wise meta-analyses cannot comprehensively examine the effects of different routes of administration; namely, it remains unclear whether different placebo administration modalities result in distinct effects. Third, these current meta-analyses combined all control groups into one arm, hindering the investigation of OLPs compared to different control groups, risking either over or underestimation of effects. Last, a systematic investigation exploring the importance of treatment expectancies in OLPs (induced through verbal suggestions or conditioning) is missing. To examine these open questions, we have applied network-metanalytic procedures and conducted the first network meta-analyses in OLPs (study 1), which methodologically allows us to address the previously mentioned challenges.

Our analyses revealed that (1) OLPs can be beneficial in comparison to no treatment (NT) in nonclinical (12 trials; 1’015 participants) and clinical populations (25 trials; 2’006 participants). Overall, higher effects were found in the clinical sample (e.g., standardized mean difference, SMD, for OLP pills vs. NT = 0.46) compared to the nonclinical (e.g., SMD for OLP pills vs. NT = 0.10, *n.s.*). (2) The kind of modality (i.e., the route of administration) had no substantial impact on OLP effects. However, (3) OLP effects can vary depending on the comparator used, and (4) the induction of positive treatment expectancies was found to be important in order for OLPs to work. This last finding was especially pronounced within the nonclinical network, where OLPs delivered without the evocation

of at least minimal treatment expectancies (herein referred to as OLP-) were less efficacious as compared to all other groups within the network, even to NT (SMD = -0.60). In this sense, OLP conditions that do not include any expectancy-building component could, at best, serve as controls, controlling for the component of the pill. The pill itself might, therefore, not be necessary to produce positive treatment effects in OLP studies.

Randomized controlled trial (study 2)

In order to test this hypothesis, the second study aimed to take the OLP concept one step further and investigate the necessity of a physical pill to produce positive effects (study 2). Such extended or different placebo paradigms can aid in understanding the mechanisms of OLP by systematically manipulating the treatment setting and application. Hence, we tested placebo effects without the use of a physical placebo by having participants imagine taking a pill in an online randomized controlled trial (RCT) with healthy students self-reporting test anxiety. A novel imaginary pill (IP; n = 55) intervention was compared to OLP (n = 59) and a control group (CG; n = 59). Three weeks before the exam, all three groups received a rationale according to their group allocation in an online treatment session. Both intervention groups were instructed to take two (imaginary) pills daily for three weeks, and the outcomes were assessed weekly (e.g., test anxiety) or after the exam (i.e., test performance).

Groups' test anxiety differed at study endpoint, $F(2,169) = 11.50$, $p < .001$. Test anxiety was lower in the intervention groups compared to the CG, $t(169) = -4.44$, $p < .001$, $d = -0.71$. The two intervention groups, however, did not differ significantly, i.e., both were similarly efficacious. In terms of our secondary outcome, test performance, the intervention groups (OLP/IP) had higher test scores compared to the CG, $t(117) = 1.98$, $p = .050$, $d = 0.38$, whereas the intervention groups did again not differ. Furthermore, participants receiving an intervention (IP/OLP) expected fewer symptoms compared to the CG, $t(169) = -5.76$, $p < .001$, $d = -0.92$, with scores of the two intervention groups being comparable. Mean expectancy significantly correlated with endpoint test anxiety ($r = 0.56$, $p < .001$). When including expectancy as an additional covariate in the overall model, expectancy was significantly associated with test anxiety, $F(1, 168) = 21.14$, $p < .001$, but the treatment group remained significant, $F(2, 168) = 12.87$, $p < .001$.

Qualitative analysis (study 3)

Whereas (network) meta-analyses and RCTs can answer questions about phenomena or interventions as a whole, their aggregated data does not reflect individual perspectives.

Especially for new treatments, the idiosyncratic views of participants are of significant value to better understand the acceptance and the individual treatment experience. This allows for a better understanding of what factors were perceived as helpful, critical, or hindering the success of the treatment. To clarify this endeavor, a qualitative study (study 3) was nested in the before-mentioned RCT (study 2). A reflexive thematic analysis of semi-structured interviews (*N* = 20) was conducted, and open-ended questions from the RCT were qualitatively evaluated (*N* = 114).

Four key themes were identified: (1) attitude towards the intervention, (2) applicability of the intervention, (3) experience of effects, and (4) characteristics of the imagination. The IP intervention was well-accepted, easily applicable, and various effects, pill characteristics, and appearances were described. Many participants did not desire a physical pill, either due to the absence of the imagination component or aversion to pills. Still, the IP approach was considered to be cognitively and time-demanding, which in turn had the positively experienced effect of establishing the pill intake as a therapeutic ritual that protected against the increase in test anxiety during the preparation phase. The importance of a treatment rationale was stressed, counteracting an often-existing initial ambivalent attitude. Participants' desire for relief further influenced motivation, effect, and application of the two interventions due to differences in levels of suffering or proximity to the upcoming exam. Treatment expectancies were met by the majority of interviewees; however, this finding is potentially due to low initial expectancies. The open-ended questions of the RCT corroborated the reflexive thematic analysis findings.

Discussion

This cumulative work sought to improve the understanding of basic OLP mechanisms by identifying agents that contribute to the success of OLP practices. The network meta-analyses (study 1) support the notion that OLPs can be an efficacious intervention for nonclinical and clinical populations, with clinical samples benefiting to a greater extent from the treatment altogether. These sample-distinct effect sizes can offer valuable information about potential factors that contribute to greater or lower OLP effects (discussed below). Notably, a salient and intriguing finding was that embedding pill administration in a treatment *rationale* is an essential component for OLPs' success. On the other hand, simply delivering a pill without embedding it in an explanatory story seems insufficient or even produces negative effects (as was the case when OLP- was compared to NT in nonclinical samples). Therefore, OLP conditions lacking expectancy-building components may, at best, be suitable as controls for the component of the pill, which, on the other hand, may not be a necessary determinant in producing OLP effects.

This finding was confirmed in an RCT with healthy students (study 2), where OLPs and the sheer imagination of a pill both led to significantly lower test anxiety scores as compared to a control group with a moderate-to-large effect (d = 0.71). Also, test performance (i.e., students' grades) was significantly better in the intervention groups compared to the CG with a small effect (d = 0.38). Thus psychological components, for their part, may be sufficient on their own to exploit placebo effects. Although the observed effects were associated with expectancy, as this measure was positively correlated with outcome (r = 0.56); our analyses indicated that not only expectancy but also other factors must be responsible for the group-specific improvement in test anxiety. Importantly, the study shows that a crucial element for the OLP effects seems to be the daily therapeutic *ritual* of (imaginary) pill-taking that both intervention groups engaged in.

To gain a deeper understanding of the experiences at an individual participant level the qualitative analysis (study 3) may prove useful: The challenging implementation, especially of the IP intervention, promoted the establishment of a therapeutic ritual, supporting the previous finding. Also, the importance of a treatment rationale was highlighted. Besides that, common themes were connected to participants' *desire for relief*: Motivation, application, and effects of the two interventions fluctuated over time due to different levels of suffering or the proximity of the upcoming exam. Treatment

expectancies, on the other hand, seemed to be low, particularly at the start of the interventions.

In the following, the empirical results of the three studies are integrated into existing literature, and one central mechanism is derived from each study and explored in more depth. The mechanisms are discussed in relation to one another, incorporating the results of all three studies, to broaden our understanding of the OLP phenomena.

Rationale

The most clear-cut and intriguing finding of our network meta-analyses (study 1) relates to the importance of the *treatment rationale*. Receiving an OLP intervention without an expectancy-building component seems not beneficial and can even result in negative effects. This implies that participants who receive an "empty" pill without any explanation may feel not being taken seriously, for reasons such as potential underestimation of the power of (open-label) placebos or since placebos are often still regarded as negative or connoted with deception (Smits et al., 2021). Alternatively, they may be disappointed to find themselves in a control condition as placebos are often associated with “control.” Hence, many OLP studies strengthen the rationale's vital importance (Klinger et al., 2017; Locher et al., 2017; Rathschlag & Klatt, 2021). As such, OLP treatments are cost-effective but not as time-saving as over-the-counter medications. In other words, the pill has to be filled with a meaning (which casts doubt on whether placebo pills are really "empty pills"), which requires a (human) counterpart and time.

The open-ended questions from the RCT (study 2) supported the importance of the rationale by illustrating that the explanations led to a better understanding, reinforced prior knowledge, made sense, and created confidence in the intervention. Notably, and along these lines, the qualitative analysis (study 3) showed that the treatment rationale addressed and mitigated initial skepticism. Especially the *discussion points* about underlying mechanisms (e.g., conditioning) and the information that an open attitude helps but is not necessary were most remembered and convincing. In line, a recent survey found that the mention of brain mechanisms was most plausible to explain placebo effects (Smits et al., 2021). However, these findings contrast the ones of Locher and colleagues, where participants seldom emphasized brain-related mechanisms, such as classical conditioning, as a mechanism of placebo effects (Locher et al., 2021). Other mindsets of approaching the intervention (such as “let’s give it a try”), however, were more often used, as was also often the case in our qualitative analysis. The participants' perceptions of what

is essential seem, therefore, to be rather unique and certainly also dependent on the nature of their illness, cultural background, and personal attitudes. Perkins and Repper (2021) aptly describe that within our multicultural society, there are large *individual and cultural differences* in understanding mental health challenges and expectancies of mental health services, and this diversity of meaning is the cause of the large individual and cultural differences in placebo effects or contextual healing (Miller & Kaptchuk, 2008). Specifically, for someone who holds a biological model of mental disorders, the meanings, and narratives fitting that model (i.e., neurobiological explanations) may lead to greater placebo effects, whereas for someone else, a more spiritual explanatory model may be meaningful. Therefore, the compatibility of mental health beliefs between the individual and the clinician with his or her corresponding intervention is essential (Perkins & Repper, 2021). Hence, it is arguably of great significance that we present not just one but several discussion points, offering a variety of options for consideration. Most participants in our qualitative analysis (study 3) regarded the biological discussion point as important. The biological explanation might have been most meaningful and credible as our sample was academic (100% students and 88% of them bachelor psychology students) and thus displayed a mindset rooted in natural science. Similarly, the survey sample by Smits and colleagues was highly educated and of relatively young age (Smits et al., 2021). Thus, future research in populations with other characteristics will illuminate the diversity of plausible and meaningful OLP treatment explanations, thereby enriching the range of discussion points from which OLP-treated individuals can draw their personal meaning.

Zooming out again from distinct discussion points to the treatment rationale as a whole, the question of the ideal conditions that make a treatment explanation “effective” apart from its content remains. In the spirit of provider transparency, we were completely honest with our participants by offering an evidence-based explanation that may be somewhat modest or demystifying but is as close to the facts as can be (it’s just a placebo, after all). Likewise, with our placebo treatments, especially the IP intervention, we provide a maximally *adaptive and customizable context*. Patients can fill this context with their own desires, meanings, and preferences (regarding the characteristics of the intervention and its effects). This, in turn, maximizes *autonomy and self-efficacy*, another frequently discussed theme in OLPs (Ballou et al., 2022), as the effect of the (physical and even more the imaginary) pill can consciously only be attributed to the person themselves and not to an external entity such as an (active) pill.

From a more contextual understanding of placebo effects (Miller & Kaptchuk, 2008), the importance of the *context of the therapeutic encounter* in which a treatment rationale is delivered is noteworthy, as it is often considered the primary vehicle of therapeutic benefit (Benedetti et al., 2011). Within OLP literature, the potential significance of the patient-clinician interaction is accordingly discussed (Blease et al., 2020). Studies have already confirmed the considerable impact of the patient-clinician relationship in deceptive placebo application (Gaab et al., 2019; Kaptchuk et al., 2008), whereas the results of OLP studies are still too under-researched to draw clear conclusions (e.g., Leibowitz et al., 2019). The database on distinct OLP effects for suggestions in a written form (e.g., Friehs et al., 2022), via video (e.g., Bräscher et al., 2022; Carvalho et al., 2016; Haas, Winkler, et al., 2022; Kleine-Borgmann et al., 2019; Schneider et al., 2020) or virtually (e.g., Kube et al., 2021; study 2) is inconclusive. Yet, contact with a treatment provider either in-person or remotely seems to produce larger OLP effects and thus is supposedly essential. In summary, obtaining a treatment explanation in the context of a patient-provider interaction is of vital importance for the effect of OLP to be realized.

Ritual

The findings of our network meta-analyses (study 1) question the ability of inert remedies as a sole source to produce beneficial effects. Therefore, we concluded that the pills' physical presence might not be essential for yielding favorable treatment effects, a finding confirmed by our RCT (study 2). In there, both the OLP and IP (i.e., the imagination of a pill intake) significantly reduced test anxiety compared to the control group. The finding that simply imagining a pill intake can lead to similar results as ingesting placebo pills highlights the power of the psychological component of OLP mechanisms. That psychological components for their part may be sufficient on their own to leverage placebo effects is also evidenced by research demonstrating the feasibility of eliciting placebo effects without any physical treatment component (Crum & Langer, 2007; Gaab et al., 2019; Kong et al., 2018; K. j. Peerdeman et al., 2017; Wai-Lan Yeung et al., 2020). Thus, the question is why placebos, purely psychological in nature, can still produce beneficial effects.

Despite the absence of a physical pill, participants in the IP group of the RCT still engaged in a *therapeutic ritual* of taking a pill twice daily (study 2). The qualitative interviews (study 3) reveal that, in addition to the importance of the rationale, participants most frequently mentioned that they ritualized pill-taking and/or embedded it in a preexisting ritual. This therapeutic ritual in itself, as well as the creation of a small break in daily life, were

perceived as supportive and protective against the increase in test anxiety during the three-week preparation phase. The importance of establishing a ritual is supported by the non-significant OLP finding in an RCT of insomnia by Haas, Winkler, et al., (2022), who argued that the dosing regimen of placebo application warrant careful consideration: Namely, in successful OLP studies, placebos are taken regularly, usually over a period of two or three weeks, whereas a single OLP application, as performed in their and other studies, has been shown to be efficacious only in healthy samples (Kube et al., 2020; Locher et al., 2017; Saunders et al., 2019). Findings of our network meta-analyses (study 1) offer additional support for this hypothesis, as the effects in the clinical network, where pills were usually ingested over a longer period of time, were greater as compared to the ones in the nonclinical network. Furthermore, in the clinical sample, no significantly worse effects of OLP participants without expectancy induction (i.e., OLP-) compared to NT were found, as was the case in the nonclinical network: A potential explanation for this finding might be that the disappointment of being assigned to a placebo control condition was buffered by at least performing a ritual of, e.g., taking pills over a period of time. No such ritual could have been established in a single administration, which was often the case in the nonclinical studies.

Apart from ritualization, the *symbol of the pill itself* is meaningful. Our qualitative study (study 3) shows that the symbolism of a pill is associated with a certain expectancy (i.e., getting better) and creates a sense of familiarity among participants. This finding highlights the cultural significance of pill-taking, as it is deeply embedded in participants' customs and beliefs. The Western cultural understanding of a pill itself has a therapeutic meaning; the association of the pill and its beneficial effects can be viewed as a product of culturalization. Hence, a culturally specific symbol such as a pill can be a significant healing aspect in the therapeutic context and, in turn, lead to placebo effects (Perkins & Repper, 2021). Here, the appearance of a pill can serve an important role: Interestingly, the IPs described most often by participants (study 2/study 3) corresponded with the conventional pill type and packaging in Switzerland (small, white, and round pills, packaged in a blister). A realistic appearance imagined could, in turn, underpin the credibility and familiarity of the intervention and thus enhance expectancies. Interestingly, in a survey about expectancies associated with pharmaceutical pill color and shape, white pills were perceived as significantly more effective in treating headaches than any other colors – regardless of the country in which the survey was conducted (i.e., USA, China, Colombia). The authors speculate that the preference for white pills in treating headaches

may be attributed to the global popularity of the headache medication "aspirin," which happens to be white in color (Wan et al., 2015). Google searches (March, 2023) on the most common pharmaceutical drugs against test anxiety (i.e., beta-blocker, benzodiazepine) showed that these pills are very often small, white, and packaged in a blister. This finding underscores the significant influence of contextual factors, such as prior experience with a particular drug or advertising, on the relationship between drug appearance and perceived efficacy (Meissner & Linde, 2018). Also, in a survey of laypeople, 82.7% agreed that the packaging of a placebo (i.e., the color of the pill) could influence its effect (Smits et al., 2021).

Yet, the question of how a ritual or a symbol of a pill can lead to significant changes still remains. A series of psychophysiological studies by Benedetti demonstrated how different social stimuli, including words and rituals of the therapeutic act, can modify the chemistry and circuitry of a patient's brain. Such processes can lead to a placebo effect in which the prefrontal cortex is supposed to be involved (Benedetti et al., 2011). The absence of prefrontal activation in OLPs, however, suggests that the *expectancy* of effects may play a less prominent role (Schaefer et al., 2023). Even though we found a significant relationship between expectancy and outcome in our RCT (study 2), statistical analyses suggest that expectancy alone does not account for group-specific improvement. The effects can, for example, be discussed in the context of *conditioning*: In fact, our network meta-analyses (study 1) demonstrated that clinical OLP groups in which a conditioning paradigm with pills was applied had larger effect sizes when compared to NT (SMD = 0.89) as compared to unconditioned OLP pills (SMD = 0.46). When it comes to conditioning effects that are rooted in our learning histories (e.g., pill means healing), on the other hand, criticism (for instance, about negative learning histories in clinical samples) has been raised.

Here, more sophisticated paradigms, such as *Prediction Error Processing* (PEP), are proposed as more adequate explanatory models for the effects of rituals and meaning carriers such as pills (e.g., Kaptchuk, 2018). In the case of OLPs' ambiguous messages, a discrepancy is likely to occur between prediction and actual experience (“This is an empty pill, so it won't help me” vs. “I take this pill every day, received it within a medical setting”). To reduce this mismatch, the daily therapeutic ritual of (imaginary) pill-taking can lead to interpretations of even small physical changes as results of treatment (e.g., being more relaxed due to small “pill-breaks”), which in turn can result in self-fulfilling healing (Kaptchuk, 2018; Ongaro & Kaptchuk, 2019). In a similar vein, the model of *embodied*

cognition states that our experiences are not only consciously stored as memories but also directly imprinted in our bodies without any cognitive process being involved (Thompson et al., 2009): In the case of an IP, the mere act of imagining taking a pill can lead to changes in bodily sensations and physiological responses, as cognitive representations of pill-taking are associated with, e.g., a feeling of relaxation in our body (Kaptchuk, 2018). In support of both of these theories are themes that emerged in our qualitative analyses (study 3): The majority of IP participants were, for instance, positively surprised, astonished, or impressed by the effects of the pill, many of which were (accompanying) bodily symptoms such as being more relaxed. Similarly, the argument of ambiguity central to both theories was reflected by the majority of participants in their skepticism towards the intervention, particularly at the beginning, but also open-mindedness. Thus, this initial ambivalent attitude might have been consciously countered by providing an explanation, but also unconsciously due to embodied pill-taking behavior and the daily ritual performed. Consequently, a person's attention may be focused on positive physical changes as a result of these rituals, e.g., taking a small break for the pill intake leaves time to imagine the desired state and thereby may cause becoming somewhat calmer. Crucially, predictive processing demonstrates that the therapeutic ritual and the active ingredients of an intervention act, albeit in different ways, on the same process through which we experience symptom relief (Ongaro & Kaptchuk, 2019). In conclusion, the three studies in this thesis contribute to the view that performing a culturally plausible healing ritual is critical to initiating a variety of healing processes within the body.

Desire for relief

To gain a deeper understanding of the intrapersonal factors that drive OLP effects, the following sections attempt to integrate several concepts discussed in the context of internal states that promote the occurrence of OLP effects. First, despite the above-mentioned more recently discussed approaches, the most commonly mentioned underlying mechanism of OLP in literature and often considered to be the very essence of OLP efficacy, still are concepts related to treatment *expectancy*. This assumption might be derived from the prominent role of response expectancy in deceptive placebo mechanisms (Kirsch, 2018). To investigate its role in OLP mechanisms, a majority of OLP studies assessed treatment expectancy, and a few report positive associations with outcomes (see introduction for a summary). In line with these observations, expectancy for relief was significantly associated with test anxiety in our RCT (study 2). Also, the fact

that different routes of treatment administration are linked with different treatment expectancies and thus also with varying (placebo) effects supports this notion (Meissner & Linde, 2018): More invasive routes of placebo delivery are, for instance, associated with more substantial effects than oral or nasal administration. Also, even if counter-intuitive at first, the finding that OLP administration routes did not differ in our network (study 1) may be attributed to the influence of expectancy on the intervention. SMDs varying by as much as 0.50 between modalities suggest that (particularly within the clinical sample) potential differences between OLP modalities may be masked, given that the present analyses have examined the OLP effects for a variety of different somatic and mental health conditions. Peerdeman et al., (2018) found that expectancies of the efficacy of different routes of administration differed for pain and itch; for example, injected medications were rated as most effective for relieving pain, and topical medications for relieving itch. These results may reflect the influence of knowledge, medical culture, and prior experience on treatment expectancies (Kirsch, 2018).

As mentioned, however, neural investigations indicate that prefrontal brain regions do not exhibit activation in the case of OLPs, and relatively few OLP studies actually found positive associations with treatment outcomes. This suggests that expectancies may play a relatively minor role. In line, expectancy as a mechanism has been criticized in the context of OLPs, specifically because patients with a history of negative treatment outcomes encounter an intervention with no or negative expectancy (Ballou et al., 2017). Our qualitative study (study 3) revealed that participants' treatment expectancies were largely met. On closer examination, however, the reason for this was often a very low initial treatment expectancy at baseline (“I [initially] underestimated that it can really work”). The possibility that the participants had previous negative experiences with treatments cannot be ruled out, but the nonclinical nature of the sample tends to argue against this. Yet, reported mindsets such as “I have nothing to lose,” “let’s give it a try,” or “it won’t harm me if it doesn’t work” (a Swiss German saying) are related to open-mindedness and the idea of *hope*. This less cognitive and more affective factor of hope has been suggested to be important in OLP effects (Haas, Ongaro, et al., 2022; Kaptchuk et al., 2009; Pan et al., 2022), as patients, who had no previous success with medication for their symptoms, could adopt a try-out-attitude of ”what if it helps?” (Haas, Ongaro, et al., 2022). This attitude might allow for a more adequate assessment of help and thus protects against inflated treatment expectancies (Kaptchuk, 2018).

Connected with hope is *novelty* (Ballou et al., 2017), namely, the assumption of being able to try something new and possibly opening the door to an effective treatment. In line with this is the finding of our network (study 1) of smaller effect sizes for OLP pills in the clinical sample (SMD = 0.46) as compared to previous investigations (SMD = 0.88 in Charlesworth et al., 2017 and 0.72 in von Wernsdorff et al., 2021). This trend towards smaller effects over time may, on the one hand, indicate research-related idiosyncrasies (i.e., reporting bias, time lag bias) but, on the other hand, also a loss of novelty on the part of the study participants (especially since they are often psychology students with knowledge of (open-label) placebo effects). This is supported by half of the OLP interviewees in our qualitative analysis (study 3) having pre-existing familiarity with (open-label) placebos. Thus, the newer a treatment is, the greater the hope for its efficacy can be, and the larger the effects.

An attempt to combine the aspects of hope and expectancy mentioned above can potentially be offered by the superordinate construct of *desire for relief*. Patients in clinical OLP studies often reported feeling despair and a strong desire for relief (Ballou et al., 2017; Kaptchuk, 2018), and in a similar vein, studies investigating deceptive placebo analgesia have found larger effect sizes in patients as compared to healthy individuals (Forsberg et al., 2017). Also, findings of larger OLP effects in the clinical compared to the nonclinical sample in our network (study 1) fit well into this: Clinical populations, and thus individuals who are suffering, enter a study with a stronger desire for relief as compared to healthy individuals. Patients have a specific desire, namely, to get better. In case they do not receive treatment and are, for instance, assigned to a control group, this can lead to disappointment or bring them in a waiting position (not doing anything to improve, but awaiting treatment; Blease et al., 2020). Healthy participants receiving OLPs after a symptom provocation, on the other hand, are assumed not to be hopeful, yet may have certain expectancies for symptom relief after treatment application. In line with this differentiation and despite his strong criticism of expectancy theory in the context of OLP mechanisms, Kaptchuk (2018) acknowledges that expectancy is involved with many placebo effects, especially in acute experiments with nonclinical samples, in short-term interventions for patients, or for new patient conditions without negative treatment experiences. Indeed, a large meta-analysis investigating patients' pain with expectation interventions (i.e., verbal suggestion, conditioning, and mental imagery), found greater effects in acute pain as compared to chronic pain (Peerdeman et al., 2016). As such, it seems that different mechanisms come into play as a function of participants' desire for

relief: Expectancies might play a more central role in nonclinical samples, whereas in clinical populations the mechanism of hope is more prominent. In the case of our RCT (study 2), all baseline text-anxiety values were within a normal range of scores, indicating that our sample was overall healthy. Still, our sample can be categorized as preclinical as participants wished to experience less test anxiety and therefore both of the before-mentioned mechanisms could be at work. In our qualitative analysis (study 3), for instance, motivation, effects, and application of the two interventions fluctuated over time due to different levels of despair. Illustratively, a participant shared that motivation for pill intake dropped after some time due to a lower level of suffering. Along these lines, participants reported that the exam proximity changed the experience of effects. With the increase in pressure and nervousness due to the upcoming exam, respondents showed enhanced motivation for the pill intake and increased effects ("As the exam was approaching, you became more nervous, and you needed something that would help you [...], the need increased and the effect also went up").

Conclusion

Towards a more conclusive answer as to why such a paradoxical and contradictory treatment as openly administered placebos can unfold an effect, this paper attempted to integrate empirical results on potential key components from three studies into the current state of research. To this end, three major mechanisms were identified that might be unique in bringing the OLP treatment regime into flourishing: (1) Incorporating the OLP administration into a narrative, such as providing a *treatment rationale* within a healing (interpersonal) context, is key. As such, giving meaning and credibility to the treatment through a story (e.g., the OLP discussion points) resonating with the individuals' preferences is vital to successful OLP treatment. (2) Essential for OLP effects is likewise the *therapeutic ritual* involved, such as the (imaginary) pill intake. The pill symbolizes healing in our participants' culture and can thus activate conscious expectancies and unconscious processes, such as conditioning, predictive processing, or embodied cognition. (3) Other than these treatment-specific aspects, an important factor is inherent to the treated individual that greatly influences the magnitude of the OLP effect: Namely, the *desire for relief*, which may vary depending on a person's level of suffering and the corresponding expectancies or hopes for the treatment.

In short, the effects of the paradoxical OLP treatment may be interpreted as a phenomenon that occurs within the framework of a therapeutic interaction that offers a plausible treatment narrative, involves various healing rituals, symbols, and meanings that

affect both conscious and unconscious expectancies depending on the individual's particular desire for relief when seeking treatment. From a research perspective, knowledge of these mechanisms can provide insight into the relative weight of each OLP component. Recognizing the significance of the treatment rationale, the associated therapeutic rituals, and the patient's desires upon entering treatment, can, in turn, educate clinicians so that they might consider or use these mechanisms in their clinical practice and, where necessary and for the sake of transparency, overtly communicate the therapeutic value of each component.

References

Annoni, M. (2018). The ethics of placebo effects in clinical practice and research. *International Review of Neurobiology*, *139*, 463–484.

Ballou, S., Haas, J. W., Iturrino, J., Nee, J., Kirsch, I., Rangan, V., Cheng, V., Lembo, A., Kaptchuk, T. J., & Kelley, J. M. (2022). Psychological Predictors of Response to Open-label versus Double-blind Placebo in a Randomized Controlled Trial in Irritable Bowel Syndrome. *Psychosomatic Medicine*, *84*(6), 738–746.

Ballou, S., Kaptchuk, T. J., Hirsch, W., Nee, J., Iturrino, J., Hall, K. T., Kelley, J. M., Cheng, V., Kirsch, I., & Jacobson, E. (2017). Open-label versus double-blind placebo treatment in irritable bowel syndrome: Study protocol for a randomized controlled trial. *Trials*, *18*(1), 1–15. https://doi.org/10.1186/s13063-017-1964-x

Benedetti, F., Carlino, E., & Pollo, A. (2011). How placebos change the patient's brain. *Neuropsychopharmacology*, *36*(1), 339–354.

Benedetti, F., Frisaldi, E., & Shaibani, A. (2022). Thirty years of neuroscientific investigation of placebo and nocebo: The interesting, the good, and the bad. *Annual review of pharmacology and toxicology*, *62*, 323–340.

Bernstein, M. H., Fuchs, N., Rosenfield, M., Weiss, A.-P., Blease, C., Locher, C., Magill, M., Rich, J., & Beaudoin, F. L. (2021). Treating pain with open-label placebos: A qualitative study with post-surgical pain patients. *The Journal of Pain*, *22*(11), 1518–1529.

Blease, C. R., Bernstein, M. H., & Locher, C. (2020). Open-label placebo clinical trials: Is it the rationale, the interaction or the pill? *BMJ Evidence-Based Medicine*, *25*(5), Article 5. https://doi.org/10.1136/bmjebm-2019-111209

Bräscher, A.-K., Ferti, I.-E., & Witthöft, M. (2022). Open-label placebo effects on psychological and physical well-being: A conceptual replication study. *Clinical Psychology in Europe*, *4*(4), 1–16.

Carvalho, C., Caetano, J. M., Cunha, L., Rebouta, P., Kaptchuk, T. J., & Kirsch, I. (2016). Open-label placebo treatment in chronic low back pain: A randomized controlled trial. *Pain*, *157*(12), 2766–2772. PubMed. https://doi.org/10.1097/j.pain.0000000000000700

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*(2), 97–107. https://doi.org/10.1111/jebm.12251

Chen, P.-H. A., Cheong, J. H., Jolly, E., Elhence, H., Wager, T. D., & Chang, L. J. (2019). Socially transmitted placebo effects. *Nature human behaviour*, *3*(12), 1295–1305.

Colloca, L., & Benedetti, F. (2009). Placebo analgesia induced by social observational learning. *PAIN®*, *144*(1–2), 28–34.

Colloca, L., & Miller, F. G. (2011). How placebo responses are formed: A learning perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1572), 1859–1869.

Crum, A. J., & Langer, E. J. (2007). Mindset matters: Exercise and the placebo effect. *Psychological Science*, *18*(2), 165–171. https://doi.org/10.1111/j.1467-9280.2007.01867.x

de Craen, A. J., Tijssen, J. G., de Gans, J., & Kleijnen, J. (2000). Placebo effect in the acute treatment of migraine: Subcutaneous placebos are better than oral placebos. *Journal of Neurology*, *247*(3), 183–188. https://doi.org/10.1007/s004150050560

El Brihi, J., Horne, R., & Faasse, K. (2019). Prescribing placebos: An experimental examination of the role of dose, expectancies, and adherence in open-label placebo effects. *Annals of Behavioral Medicine*, *53*(1), 16–28. https://doi.org/10.1093/abm/kay011

Finniss, D. G., Kaptchuk, T. J., Miller, F., & Benedetti, F. (2010). Placebo effects: Biological, clinical and ethical advances. *Lancet*, *375*(9715), 686.

Flowers, K. M., Patton, M. E., Hruschak, V. J., Fields, K. G., Schwartz, E., Zeballos, J., Kang, J. D., Edwards, R. R., Kaptchuk, T. J., & Schreiber, K. L. (2021). Conditioned open-label placebo for opioid reduction after spine surgery: A randomized controlled trial. *Pain*, *162*(6), 1828. https://doi.org/10.1097/j.pain.0000000000002185

Forsberg, J. T., Martinussen, M., & Flaten, M. A. (2017). The placebo analgesic effect in healthy individuals and patients: A meta-analysis. *Psychosomatic Medicine*, *79*(4), 388–394. https://doi.org/10.1097/psy.0000000000000432

Friehs, T., Rief, W., Glombiewski, J. A., Haas, J., & Kube, T. (2022). Deceptive and non-deceptive placebos to reduce sadness: A five-armed experimental study. *Journal of*

Affective Disorders Reports, *9*, 100349. https://doi.org/10.1016/j.jadr.2022.100349

Gaab, J., Blease, C., Locher, C., & Gerger, H. (2016). Go open: A plea for transparency in psychotherapy. *Psychology of Consciousness: Theory, Research, and Practice*, *3*(2), 175.

Gaab, J., Kossowsky, J., Ehlert, U., & Locher, C. (2019). Effects and components of placebos with a psychological treatment rationale–three randomized-controlled studies. *Scientific reports*, *9*(1), 1421.

Guevarra, D. A., Moser, J. S., Wager, T. D., & Kross, E. (2020). Placebos without deception reduce self-report and neural measures of emotional distress. *Nature communications*, *11*(1), 3785.

Haas, J. W., Ongaro, G., Jacobson, E., Conboy, L. A., Nee, J., Iturrino, J., Rangan, V., Lembo, A., Kaptchuk, T. J., & Ballou, S. (2022). Patients' experiences treated with open-label placebo versus double-blind placebo: A mixed methods qualitative study. *BMC psychology*, *10*(1), 1–12.

Haas, J. W., Winkler, A., Rheker, J., Doering, B. K., & Rief, W. (2022). No open-label placebo effect in insomnia? Lessons learned from an experimental trial. *Journal of Psychosomatic Research*, *158*, 110923. https://doi.org/10.1016/j.jpsychores.2022.110923

Hoenemeyer, T. W., Baidwan, N. K., Hall, K., Kaptchuk, T. J., Fontaine, K. R., & Mehta, T. S. (2021). An exploratory analysis of the association between Catechol-O-methyltransferase and response to a randomized open-label placebo treatment for cancer-related fatigue. *Frontiers in Psychiatry*, *12*, 684556.

Hruschak, V., Flowers, K. M., Patton, M., Merchantz, V., Schwartz, E., Edwards, R., Kaptchuk, T., Kang, J., Dossett, M., & Schreiber, K. (2022). Experiences of Patients Taking Conditioned Open-Label Placebos for Reduction of Postoperative Pain and Opioid Exposure After Spine Surgery. *International Journal of Behavioral Medicine*, 1–13.

Jensen, K. B., Kaptchuk, T. J., Kirsch, I., Raicek, J., Lindstrom, K. M., Berna, C., Gollub, R. L., Ingvar, M., & Kong, J. (2012). Nonconscious activation of placebo and nocebo pain responses. *Proceedings of the National Academy of Sciences*, *109*(39), 15959–15964.

Kam-Hansen, S., Jakubowski, M., Kelley, J. M., Kirsch, I., Hoaglin, D. C., Kaptchuk, T. J.,

& Burstein, R. (2014). Altered placebo and drug labeling changes the outcome of episodic migraine attacks. *Science translational medicine*, *6*(218), 218ra5-218ra5.

Kaptchuk, T. J. (2018). Open-Label Placebo: Reflections on a Research Agenda. *Perspectives in Biology and Medicine*, *61*(3), Article 3. https://doi.org/10.1353/pbm.2018.0045

Kaptchuk, T. J., Friedlander, E., Kelley, J. M., Sanchez, M. N., Kokkotou, E., Singer, J. P., Kowalczykowski, M., Miller, F. G., Kirsch, I., & Lembo, A. J. (2010). Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome. *PLOS ONE*, *5*(12), Article 12. https://doi.org/10.1371/journal.pone.0015591

Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., Kirsch, I., Schyner, R. N., Nam, B. H., & Nguyen, L. T. (2008). Components of placebo effect: Randomised controlled trial in patients with irritable bowel syndrome. *Bmj*, *336*(7651), 999–1003.

Kaptchuk, T. J., Shaw, J., Kerr, C. E., Conboy, L. A., Kelley, J. M., Csordas, T. J., Lembo, A. J., & Jacobson, E. E. (2009). "Maybe I made up the whole thing": Placebos and patients' experiences in a randomized controlled trial. *Culture, Medicine, and Psychiatry*, *33*, 382–411.

Kelley, J. M., Lembo, A. J., Ablon, J. S., Villanueva, J. J., Conboy, L. A., Levy, R., Marci, C. D., Kerr, C., Kirsch, I., & Jacobson, E. E. (2009). Patient and practitioner influences on the placebo effect in irritable bowel syndrome. *Psychosomatic medicine*, *71*(7), 789.

Kirsch, I. (2018). Response expectancy and the placebo effect. *International review of neurobiology*, *138*, 81–93.

Kirsch, I. (2019). Placebo effect in the treatment of depression and anxiety. *Frontiers in Psychiatry*, 407. https://doi.org/10.3389/fpsyt.2019.00407

Kleine-Borgmann, J., Schmidt, K., Billinger, M., Forkmann, K., Wiech, K., & Bingel, U. (2021). Effects of open-label placebos on test performance and psychological well-being in healthy medical students: A randomized controlled trial. *Scientific reports*, *11*(1), 1–11.

Kleine-Borgmann, J., Schmidt, K., Hellmann, A., & Bingel, U. (2019). Effects of open-label placebo on pain, functional disability, and spine mobility in patients with chronic back pain: A randomized controlled trial. *Pain*, *160*(12), 2891–2897.

Klinger, R., Kothe, R., Schmitz, J., Kamping, S., & Flor, H. (2017). Placebo effects of a sham opioid solution: A randomized controlled study in patients with chronic low back pain. *Pain*, *158*(10), 1893–1902.

Kong, J., Wang, Z., Leiser, J., Minicucci, D., Edwards, R., Kirsch, I., Wasan, A. D., Lang, C., Gerber, J., Yu, S., Napadow, V., Kaptchuk, T. J., & Gollub, R. L. (2018). Enhancing treatment of osteoarthritis knee pain by boosting expectancy: A functional neuroimaging study. *NeuroImage: Clinical*, *18*, 325–334. https://doi.org/10.1016/j.nicl.2018.01.021

Kube, T., Hofmann, V. E., Glombiewski, J. A., & Kirsch, I. (2021). Providing open-label placebos remotely—A randomized controlled trial in allergic rhinitis. *PLoS One*, *16*(3), e0248367.

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*(2), 68–79.

Lee, S., Choi, D.-H., Hong, M., Lee, I.-S., & Chae, Y. (2022). Open-label placebo treatment for experimental pain: A randomized-controlled trial with placebo acupuncture and placebo pills. *Journal of Integrative and Complementary Medicine*, *28*(2), 136–145.

Leibowitz, K. A., Hardebeck, E. J., Goyer, J. P., & Crum, A. J. (2019). The role of patient beliefs in open-label placebo effects. *Health Psychology*, *38*(7), 613.

Lembo, A., Kelley, J. M., Nee, J., Ballou, S., Iturrino, J., Cheng, V., Rangan, V., Katon, J., Hirsch, W., & Kirsch, I. (2021). Open-label placebo vs double-blind placebo for irritable bowel syndrome: A randomized clinical trial. *Pain*, *162*(9), 2428–2435.

Locher, C., Buergler, S., Nascimento, A. F., Kost, L., Blease, C., & Gaab, J. (2021). Lay perspectives of the open-label placebo rationale: A qualitative study of participants in an experimental trial. *BMJ open*, *11*(8), e053346.

Locher, C., Nascimento, A. F., 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*(12), 2320–2328.

Locher, C., Nascimento, A. F., Kossowsky, J., Meyer, A., & Gaab, J. (2019). Open-label placebo response–Does optimism matter? A secondary-analysis of a randomized controlled trial. *Journal of psychosomatic research*, *116*, 25–30.

Meeuwis, S. H., van Middendorp, H., Lavrijsen, A. P., Veldhuijzen, D. S., & Evers, A. W. (2021). Open-and closed-label placebo and nocebo suggestions about a sham transdermal patch. *Psychosomatic Medicine*, *83*(1), 33.

Meeuwis, S. H., Van Middendorp, H., Van Laarhoven, A. I., Veldhuijzen, D. S., Lavrijsen, A. P., & Evers, A. W. (2019). Effects of open-and closed-label nocebo and placebo suggestions on itch and itch expectations. *Frontiers in Psychiatry*, *10*, 436.

Meissner, K., & Linde, K. (2018). Are blue pills better than green? How treatment features modulate placebo effects. *International review of neurobiology*, *139*, 357–378. https://doi.org/10.1016/bs.irn.2018.07.014

Miller, F. G., & Kaptchuk, T. J. (2008). The power of context: Reconceptualizing the placebo effect. *Journal of the Royal Society of Medicine*, *101*(5), 222–225.

Montgomery, G. H., & Kirsch, I. (1997). Classical conditioning and the placebo effect. *Pain*, *72*(1–2), 107–113.

Morales-Quezada, L., Mesia-Toledo, I., Estudillo-Guerra, A., O'Connor, K. C., Schneider, J. C., Sohn, D. J., Crandell, D. M., Kaptchuk, T., & Zafonte, R. (2020). Conditioning open-label placebo: A pilot pharmacobehavioral approach for opioid dose reduction and pain control. *Pain Reports*, *5*(4). https://doi.org/10.1097/PR9.0000000000000828

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*(2), 307–315.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231. https://doi.org/10.1037/0033-295X.84.3.231

Nitzan, U., Carmeli, G., Chalamish, Y., Braw, Y., Kirsch, I., Shefet, D., Krieger, I., Mendlovic, S., Bloch, Y., & Lichtenberg, P. (2020). Open-Label placebo for the treatment of unipolar depression: Results from a randomized controlled trial. *Journal of Affective Disorders*, *276*, 707–710. https://doi.org/10.1016/j.jad.2020.07.077

Olliges, E., Stroppe, S., Haile, A., Reiß, F., Malhis, M., Funke, S. A., & Meissner, K. (2022). Open-label placebo administration decreases pain in elderly patients with symptomatic knee osteoarthritis–A randomized controlled trial. *Frontiers in Psychiatry*, *13*.

Ongaro, G., & Kaptchuk, T. J. (2019). Symptom perception, placebo effects, and the Bayesian brain. *Pain*, *160*(1), 1.

Pan, Y., Frank, M. L., Kaptchuk, T. J., & Nestoriuc, Y. (2022). “Let’s see what happens:”—Women’s experiences of open-label placebo treatment for menopausal hot flushes in a randomized controlled trial. *Plos one*, *17*(11), e0276499.

Pan, Y., Meister, R., Löwe, B., Kaptchuk, T. J., Buhling, K. J., & Nestoriuc, Y. (2020). Open-label placebos for menopausal hot flushes: A randomized controlled trial. *Scientific Reports*, *10*(1), 20090.

Park, L. C., & Covi, L. (1965). Nonblind placebo trial: An exploration of neurotic patients’ responses to placebo when its inert content is disclosed. *Archives of general psychiatry*, *12*(4), 336–345.

Peerdeman, K. J., Tekampe, J., Van Laarhoven, A. I. M., van Middendorp, H., Rippe, R. C. A., Peters, M. L., & Evers, A. W. M. (2018). Expectations about the effectiveness of pain-and itch-relieving medication administered via different routes. *European Journal of Pain*, *22*(4), 774–783. https://doi.org/10.1002/ejp.1163

Peerdeman, K. J., van Laarhoven, A. I., Keij, S. M., Vase, L., Rovers, M. M., Peters, M. L., & Evers, A. W. (2016). Relieving patients’ pain with expectation interventions: A meta-analysis. *Pain*, *157*(6), 1179–1191.

Peerdeman, K. j., van Laarhoven, A. i. m., Bartels, D. j. p., Peters, M. l., & Evers, A. w. m. (2017). Placebo-like analgesia via response imagery. *European Journal of Pain*, *21*(8), 1366–1377. https://doi.org/10.1002/ejp.1035

Perkins, R., & Repper, J. (2021). Reframing the placebo effect. The importance of meaning. *Mental Health and Social Inclusion*, *25*(3), 213–217.

Rathschlag, M., & Klatt, S. (2021). Open-label placebo interventions with drinking water and their influence on perceived physical and mental well-being. *Frontiers in Psychology*, *12*, 658275.

Sandler, A. D., & Bodfish, J. W. (2008). Open-label use of placebos in the treatment of ADHD: A pilot study. *Child: care, health and development*, *34*(1), 104–110.

Sandler, A. D., Glesne, C. E., & Bodfish, J. W. (2010). Conditioned placebo dose reduction: A new treatment in ADHD? *Journal of developmental and behavioral pediatrics: JDBP*, *31*(5), 369.

Saunders, B., Saito, T., Klosterhoff, R., de Oliveira, L. F., Barreto, G., Perim, P., Pinto, A. J., Lima, F., de Sá Pinto, A. L., & Gualano, B. (2019). “I put it in my head that the supplement would help me”: Open-placebo improves exercise performance in female cyclists. *PLoS One*, *14*(9), e0222982.

Schaefer, M., Hellmann-Regen, J., & Enge, S. (2021). Effects of open-label placebos on state anxiety and glucocorticoid stress responses. *Brain Sciences*, *11*(4), 508.

Schaefer, M., Kühnel, A., Schweitzer, F., Enge, S., & Gärtner, M. (2023). Neural underpinnings of open-label placebo effects in emotional distress. *Neuropsychopharmacology*, *48*(3), 560–566.

Schedlowski, M., Enck, P., Rief, W., & Bingel, U. (2015). Neuro-bio-behavioral mechanisms of placebo and nocebo responses: Implications for clinical trials and clinical practice. *Pharmacological reviews*, *67*(3), 697–730.

Schneider, T., Luethi, J., Mauermann, E., Bandschapp, O., & Ruppen, W. (2020). Pain response to open label placebo in induced acute pain in healthy adult males. *Anesthesiology*, *132*(3), 571–580.

Simpson, S. H., Eurich, D. T., Majumdar, S. R., Padwal, R. S., Tsuyuki, R. T., Varney, J., & Johnson, J. A. (2006). A meta-analysis of the association between adherence to drug therapy and mortality. *Bmj*, *333*(7557), 15.

Smits, R. M., Veldhuijzen, D. S., Olde Hartman, T., Peerdeman, K. J., Van Vliet, L. M., Van Middendorp, H., Rippe, R. C., Wulffraat, N. M., & Evers, A. W. (2021). Explaining placebo effects in an online survey study: Does ‘Pavlov’ring a bell? *Plos one*, *16*(3), e0247103.

Spille, L., Fendel, J. C., Seuling, P. D., Göritz, A. S., & Schmidt, S. (2023). Open-label placebos—A systematic review and meta-analysis of experimental studies with non-clinical samples. *Scientific Reports*, *13*(1), 3640.

Tang, B., Barnes, K., Geers, A., Livesey, E., & Colagiuri, B. (2022). Choice and the placebo effect: A meta-analysis. *Annals of Behavioral Medicine*, *56*(10), 977–988. https://doi.org/10.1093/abm/kaab111

Thompson, J. J., Ritenbaugh, C., & Nichter, M. (2009). Reconsidering the placebo response from a broad anthropological perspective. *Culture, Medicine, and Psychiatry*, *33*(1), 112–152. https://doi.org/10.1007/s11013-008-9122-2

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*(1), 3855. https://doi.org/10.1038/s41598-021-83148-6

Wai-Lan Yeung, V., Geers, A. L., & Colloca, L. (2020). Merely possessing a placebo analgesic improves analgesia similar to using the placebo analgesic. *Annals of Behavioral Medicine*, *54*(9), 637–652.

Wan, X., Woods, A. T., Salgado-Montejo, A., Velasco, C., & Spence, C. (2015). Assessing the expectations associated with pharmaceutical pill colour and shape. *Food Quality and Preference*, *45*, 171–182. https://doi.org/10.1016/j.foodqual.2015.06.009

Zhou, E. S., Hall, K. T., Michaud, A. L., Blackmon, J. E., Partridge, A. H., & Recklitis, C. J. (2019). Open-label placebo reduces fatigue in cancer survivors: A randomized trial. *Supportive Care in Cancer*, *27*, 2179–2187. https://doi.org/10.1007/s00520-018-4477-6

Auxiliary means

ChatGPT version March 23, OpenAI: https://chat.openai.com/ Assistance with linguistic improvement of part of the work in terms of clarity of phrases and sentence structure.

DeepL Translate, DeepL SE: https://www.deepl.com/translator Translation of sentences and text passages independently formulated in German throughout parts of the work.

Appendices

A. Study 1

B. Study 2

C. Study 3

Universität Basel

Fakultät für Psychologie

A. Study 1

Buergler, S.*, Sezer, D.*, Gaab, J., & Locher, C. (2023). The role of population, expectation, modality, and comparator on open-label placebo effects: A network meta-analysis. *Scientific Reports (in Review)* * shared first authorship

The role of population, expectation, modality and comparator on open-label placebo effects: A network meta-analysis

Sarah Buergler[1*°], Dilan Sezer[1*], Jens Gaab[1], & Cosima Locher[2, 3]

[1] Division of Clinical Psychology and Psychotherapy, Faculty of Psychology, University of Basel, Switzerland

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

[3] Faculty of Health, University of Plymouth, Plymouth, UK

*Shared first authorship

°Corresponding author: Sarah Buergler, MSc, Division of Clinical Psychology and Psychotherapy, Faculty of Psychology, University of Basel, Switzerland. Emails: sarah.buergler@unibas.ch

Keywords: **open-label placebo, network meta-analysis**

Short title: **Open label-placebos – network meta-analysis**

Abstract

Three meta-analyses demonstrate the clinical potential of open-label placebos (OLPs). However, there is a need to synthesize the existing evidence through more complex analyses that allow to answer questions beyond mere efficacy. This serves to better understand why and under what circumstances OLPs work (e.g., in what populations or through which routes of administration). To answer these questions, we conducted the first network meta-analysis in the field of OLPs. Our analyses revealed that OLPs can be beneficial in comparison to NT in nonclinical (12 trials; 1'015 participants) and clinical populations (25 trials; 2'006 participants). The kind of modality had no substantial impact on OLP effects. However, positive treatment expectations were found to be important in order for OLPs to work. Further, OLP effects can vary depending on the comparator used. Thus, the population, modality, expectation and comparator should be considered when designing and interpreting OLP studies.

Introduction

Placebos have been found to have clinically significant effects in a variety of clinical conditions[1,2], but their use in clinical practice is denied as it violates ethical obligations. In this regard, open-label placebos (OLP) administered under full disclosure and transparency can be considered both ethical and feasible[3]. Several studies show medium sized to large clinically relevant effects of OLPs[4–6] that can be comparable in magnitude to deceptively administered placebos (DP)[7–12]. However, given that this field of research is still in its infancy with the first controlled study published in 2008[13], there are still many questions that need to be addressed.

In OLP, no meta-analysis has so far explored the differential effects across distinct populations. Whereas in clinical conditions OLPs were significantly more efficacious compared to NT with moderate to large effects (SMD = 0.88[4] and 0.72[5]), in nonclinical experimental conditions a moderate effect was found in OLPs for self-reported outcomes (SMD = 0.43) and no significant effect was observed for objective outcomes (SMD = -0.02[6]). Thus, it appears that clinical populations may benefit more from OLP treatments than nonclinical populations, a finding known in deceptive placebos where placebo analgesia tends to be higher in patients compared to healthy subjects[14,15].

OLP effects may not only vary across populations but also across treatment modalities. For example, more invasive placebo procedures, such as injections and sham procedures, have been shown to increase expectations towards a treatment's efficacy – and in turn enhance placebo effects[16–18]. However, while placebo effects in itch seem not to differ between oral and injective placebo administration[19], in osteoarthritis intra-articular and topical placebo were more efficacious than orally administered placebo[20]. It is argued that more invasive administrations of placebos have stronger effects than less invasive administration (oral or nasal) in the case of pain, whereas in nonpain conditions such as itch, this might not be the case[21]. Nonetheless, placebo experts strongly agree that clinicians should not prescribe more invasive treatments merely to obtain stronger placebo effects, due to practical and ethical restrictions, higher costs, and higher risk of undesirable side effects[1]. This is especially true for OLPs as it is unclear to date whether the findings on deceptive placebos that more invasive treatments are more beneficial can be applied to the field of OLPs.

Not only the route of administration, but also associative learning (i.e., conditioning) and verbal suggestions that accompany a treatment play a key role in the expected and actual placebo effects[22]. In the majority of OLP studies the administration of the placebo is accompanied with a rationale consisting of four discussion points in order to induce positive treatment expectations (see e.g.,[23]). So far, the impact of positive expectation on OLP effects has been explored in studies comparing OLPs with expectation induction (i.e., through verbal suggestions or conditioning) to OLPs without such expectation-inducing procedures (hereafter, OLP-[8,24–27]). Some authors have concluded that the treatment rationale is crucial when it comes to the efficacy of OLP (e.g.,[8]), however, systematic investigations are limited to the experimental context[6].

Effect sizes may also depend on different control conditions used in trials. For example, it has been found that waitlist (WL) control groups lead to larger effects than no treatment (NT) controls[28]. This result could be due to the fact that subjects who are assigned to a WL group are not actively looking for improvement opportunities during the waiting phase, as might be the case with the NT group. Blease and colleagues (2019)[29] compared this phenomenon with the induction of nocebo effects in the context of OLPs, especially in the case when the experimenter mentions the potential advantages of the OLP intervention before the assignment to the WL. Further, the use of treatment as usual (TAU) controls can be considered problematic as the “treatment as usual” is typically not monitored or sufficiently reported, which may lead to structural inequivalence across studies which apply TAU[30,31]. Thus, it is warranted to take a closer look at the different comparators that are used across OLP studies.

As illustrated above, currently existing meta-analyses on OLP effects did not address several important aspects: (1) In each of the three analyses only one type of population (clinical or nonclinical) was considered. The comparability of effect sizes across different individual meta-analyses, however, might be limited, as these studies used different definitions for eligible OLP interventions and for conditions that qualify for the nonclinical and clinical population. This especially holds true for the question, whether subclinical conditions (e.g., menopausal hot-flushes, self-reported test-anxiety or general well-being) are to be considered nonclinical or clinical. Therefore, there is a need for a clear definition of these samples and for meta-analytic analyses that apply the same inclusion criteria in both areas. (2) These meta-analyses, as well as other meta-analyses of different placebo administration modes, cannot comprehensively examine different routes of administration, in part because interpretation of results from multiple meta-analyses is compromised by indirect comparison via subgroup analyses[21]. Hence, it remains unclear whether different placebo administrations result in different effects to justify the choice of one route of administration over another. (3) Furthermore, no review study has to date systematically, and on the basis of a relatively large database, examined whether the effects of OLPs with positive expectations either through a rationale or other expectation-inducing measures (e.g., conditioning) differ from those without expectation induction. (4) Finally, the current OLP meta-analyses lumped all control groups into one arm and thus did not differentiate between the different control conditions. However, it is of great importance to investigate OLP efficacy in comparison to different kinds of control groups, thereby ensuring that OLP effect sizes are neither over- nor underestimated.

To examine these open questions, a network meta-analysis (NMA) is the method of choice. NMAs allow the comparison of multiple treatment and comparator groups. Further, an NMA produces more accurate effect sizes than a traditional meta-analysis by including both direct and indirect evidence. To the best of our knowledge, this is the first NMA on OLP treatments. On the basis of the above discussed challenges and open questions in OLP research, we derived the following research questions (RQ) that can be answered in a network meta-analytic framework: Is the magnitude of the OLP effects different across (RQ1) clinical vs. nonclinical populations, (RQ2) OLP treatment modalities, (RQ3) treatment expectation, and (RQ4) comparator groups.

Methods

Search Strategy

A systematic review and NMA was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Study population analysis (PRISMA) statement[32,33] (eAppendix 5 in the supplement). The search strategies were conducted in Medline, Embase, and PsycINFO via Ovid, the Cumulative Index of Nursing and Allied Health Literature (CINAHL), clinicaltrials.gov, Open-Trials, and Cochrane Register of Controlled Trials were developed in close collaboration with an information specialist. The four databases and the three registries were searched using text word synonyms and database-specific subject headings for open-label placebos in February 2nd, 2021 (eAppendix 1 in the supplement) and updated on June 8th, 2022 (eAppendix 2 in the supplement). No language restrictions were applied. For Medline and Embase, randomized controlled trial (RCT) filters were applied, and conference abstracts and conference reviews were excluded from Embase. References were exported to Endnote X9 and deduplicated using the Bramer method. Furthermore, additional trials were identified from an existing systematic review on OLPs[6] and a newsletter on placebo studies (https://jips.online/). If data was not available, the corresponding authors of the respective publication were contacted via email. Several reviewers, in pairs of two, independently screened the references based on their titles and abstracts using https://covidence.org. Selected references were retrieved in full-text and independently assessed for eligibility by two reviewers. Any disagreements over eligibility were resolved by consensus or, if necessary, by consultation of a third reviewer. This study was registered with Prospero (CRD42020161696).

Study Selection

We included RCTs comparing OLPs compared to a control group in clinical (e.g., chronic low-back pain, depression, irritable bowel syndrome, allergic rhinitis), subclinical (e.g., menopausal hot flushes or test anxiety), as well as nonclinical (e.g., experimental induced pain or allergic reactions) populations. There were no age restrictions. Our definition of OLP was as follows: (1) The placebo must have been given openly, i.e., the receiver was 100% aware of getting the placebo when applied. Studies needed to state explicitly that the placebo was delivered with the full awareness of the receiver, i.e., solely using the term "open-label" as description of the study was not enough, as this term was used inconsistently sometimes referring to treatment provider being unblind. Also, balanced placebo design studies (with

e.g., a 50% chance of receiving a placebo) were excluded. (2) The placebo had to consist of a "pharmacological" property, i.e., was defined as everything that can be swallowed (e.g., pills, capsules, sirups, etc.), applied on the skin or other body parts (such as a cream or eye drops) or injected. Studies testing devices (e.g., deep brain stimulation) as well as placebo exercises, and diets were excluded. Also, studies testing procedures such as placebo massage or acupuncture without including an additional treatment arm fulfilling our placebo definition were not eligible. (3) At least minimal positive expectation needed to be induced alongside the placebo administration (e.g., either through a rationale (i.e., positive suggestions) or conditioning). (4) The placebo needed to be applied with the intention of a positive effect (i.e., therapeutic or well-being enhancing, no nocebo effects). Based on these criteria, none of the open-label drug trials using OLP as a comparator, which we aimed to also include in these analyses, met our definition.

Crossover studies were only included if we were able to extract the results of the first period of the trial (i.e., before the first cross) separately. This is because data from crossover studies should not be treated as if data stems from parallel-trials[34]. If this data was not reported, authors were contacted. In case of no response, these studies were excluded from the analysis. In order to be included, studies needed to report a baseline and a post measure or alternatively report change scores from baseline to post. Studies reporting only post values or where we were not able to retrieve means and standard deviations (SDs) were excluded. For studies published more than once (i.e., secondary analysis), we included only the entry with the most relevant data to our analysis.

Data Extraction

All relevant data were extracted independently in pairs of two using a standardized excel template. Disagreements were clarified through consensus and by consultation with a third reviewer, if required. Means and standard deviations (SDs) were extracted and in case SDs were not reported, we calculated them from standard errors (SE), confidence intervals (CIs), or interquartile ranges (IQR) and medians were converted to means[34]. If the sample size used for the analysis was not reported, we used the sample size of the baseline data (i.e., participants randomized). If it was not possible to impute appropriate measures for the calculation of effect sizes or if data was missing, we contacted the authors to obtain them. If authors did not provide the respective information, studies were excluded from further analyses.

Primary Outcomes

We applied a hierarchy for the choice of outcomes: (1) As a first choice, we extracted the primary outcome as defined by the study authors. In the presence of two or more primary outcomes, we checked trial registries for additional information and/or contacted authors. If no information could be obtained, the outcome for the present analysis was selected based on (2) the most frequently reported outcome across our data pool (i.e., pain was preferred over medication use) in order to reduce heterogeneity, and if this was not applicable (3) the most informative outcome (e.g., a symptom-related scale preferred over a general quality of life assessment). In the absence of a baseline assessment for the primary

outcome, another outcome was chosen according to these rules, avoiding the exclusion of this study (see eTable 1 in the supplement for the rationale of choice for the outcomes). If more than one baseline measure was collected, we chose the timepoint closest to the start of the intervention[35]. If more than one post measurement was reported, we extracted the first assessment after intervention end (i.e., measured at the time point closest to the end of treatment), if no other explanation for the most clinically relevant time point was given in the publication (i.e., a definition of the primary endpoint measurement). In studies including a WL control group, participants additionally received the OLP treatment after study completion. Outcomes for these individuals were not included in the analyses because they lacked a control group for comparison and in order to avoid enrolling participants multiple times.

Sample building

We allocated each study to either the nonclinical or the clinical study pool. Nonclinical studies were defined as studies that: experimentally induced states (i.e., experimentally induced pain, itch, sadness), whereas clinical studies investigated the effects of OLPs in naturally occurring states (e.g., clinical: irritable bowel syndrome, chronic lower back pain; subclinical: test anxiety, well-being, relaxation). One study[25] experimentally induced pain in an IBS patient sample. This study was rubricated as clinical.

Node building

In order to be able to test the effects of different OLP modalities in comparison to different control groups, each group in a study was clustered together with similar other study groups. Our strategy to create the nodes was data-based and with the aim to restrain from a high number of nodes. This lumping approach has the methodological advantage to increase power and to allow for more accurate estimates of the effect sizes[36,37]. The following rules were applied: (1) Nodes were built according to the OLP administration route, i.e., nasal (vapor, spray), dermal (cream, patch), or injection. In the case of oral application, we differentiated between pills (capsules, tablets) and suspensions (drops, solutions). (2) Groups testing different treatment rationales or intervention components alongside with the placebo administration (i.e.,[12,26,38–40]) or different amounts of placebos per day (i.e.,[41]) were merged. However, study groups testing the effect of OLP without the application of any expectation induction (herein referred to as OLP-) were separately entered into the analyses. (3) If there were different comparator groups that fell within one category (e.g., several DP groups), we merged them into one node (i.e.,[25,27,38,40]). (4) To assess the differences of expectation induction (e.g., through verbal suggestion or conditioning paradigms), these nodes were defined separately (e.g., OLP vs. cOLP). (5) In all cases where participants could receive the intervention upon study conclusion, we used the node WL control group. When data of study groups were merged, we used different formulas[34].

Risk of Bias

We assessed the risk of bias of the included studies using the Cochrane risk of bias tool 2[42]. Each study was assessed by two reviewers, with conflicts resolved by consensus. To account for the special nature of included studies in this NMA (i.e., all of them not being blind), we employed some special rules: (1) If

we received a “high” risk of bias rating in domain 4 only due to signaling question 4.5 (“Is it likely that assessment of the outcome was influenced by knowledge of intervention received?”), we overwrote the suggestion of the algorithm for this domain to “some concerns”. The rationale for this decision is based on the fact that a single “high” judgment in one of the four domains leads to an overall high risk of bias. Thus, all of our included studies would have received a high overall risk rating and consequently we would have lost all variance in our assessments. (2) When answering signaling questions 2.1 (“Were participants aware of their assigned intervention during the trial?”) and 4.3.question (“Were outcome assessors aware of the intervention received by study participants?”) for the comparison OLP (i.e., being aware of receiving the intervention) and DP group (i.e., being not aware), we judged as if both groups were unblinded as suggested by the authors of the risk of bias tool 2[43]. (3) Because the risk of bias tool 2 requires an assessment of the level of study group comparisons within a study, multiple assessments were performed per study. However, all multiple assessments within a study were identical and thus reported in a single column (see eTable 1 in the supplement).

Statistical analysis

In order to answer our research questions (RQ) we proceeded as follows: (RQ1) Two different networks were conducted separately, one for the clinical and one for the nonclinical population. These networks were then compared qualitatively. (RQ2) OLP treatment modalities were compared directly using head to head comparisons, excluding OLP-. (RQ3) The effect of treatment expectations was assessed using head to head comparisons with OLP- to all other OLP modality groups. (RQ4) To assess the effects of different comparator groups (i.e., NT, TAU, WL) we compared all OLP modalities that were significantly better than NT with the other comparator groups.

Effect sizes of the interventions applying the standardized mean difference (SMD) were calculated, with their magnitude interpreted as small, moderate or large, with 0.20, 0.50, and 0.80 SD units[44]. We decided to employ random-effects models rather than fixed-effects models because the included studies were expected to be heterogeneous. Network meta-analytic methods were applied within a frequentist framework using the package “netmeta” in R[45,46]. Results are presented as SMDs with corresponding 95% confidence intervals.

NMA relies on the assumption of transitivity to estimate indirect treatment effects. This assumption implies that any study participant that meets all inclusion criteria in each network is likely, in principle, to be randomized to any of the interventions in the corresponding network. We addressed the assumption of transitivity[47], by first conducting two separate networks (i.e., nonclinical and clinical) in order for the distribution of potential modifiers (e.g. population) to be more balanced across comparisons and by second checking whether the direct and indirect treatment effects are in statistical agreement (via an assessment for inconsistency). We conducted a statistical evaluation of consistency, i.e., the agreement between direct and indirect evidence, using local (separating direct from indirect evidence[48]) as well as global (design- by-treatment interaction test[49]) approaches.

The various effects of the groups were ranked using *P* scores. *P* scores are values between 0 and 1 and have an interpretation analogous to the surface under the cumulative ranking curve values[50] and measure the extent of certainty that a treatment is better than another treatment, averaged over all competing treatments. The *P* scores result in a ranking of all treatments that essentially follows the ranking of the point estimates but takes precision into account[50].

For all treatment comparisons in a NMA, we assumed a common between-study heterogeneity. Different statistics were used to quantify heterogeneity: the (within design) Q statistic[45], the between-study variance $\tau 2$, and the heterogeneity statistic I^2 [50]. The I^2 value can be interpreted as follows: 0 to 40% might not be important; 30 to 60% may represent moderate heterogeneity; 50 to 90% may represent substantial heterogeneity; 75 to 100% represents considerable heterogeneity[51].

The certainty of evidence for the network estimates of the efficacy outcomes was evaluated by using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) ratings[47], which were conducted in CINeMA (Confidence in Network Meta-Analysis[50]). In GRADE, the quality of a body of evidence is defined as the study limitations, imprecision, inconsistency, indirectness and reporting bias[47]. To assess across-study bias (reporting bias), a comparison-adjusted funnel plot and the Egger test for funnel plot asymmetry were computed[48]. In case of asymmetry the trim and fill method was used to adjust for small-study effects with NT as reference[52,53]. Due to too few comparisons we were not able to use the tool for assessing risk of bias due to missing evidence in a synthesis (ROB-MEN[54]) as initially planned.

We conducted sensitivity analyses excluding studies in which the risk of bias was high. We decided to choose this criterion, as all studies had at least a moderate risk due to the fact that blinding was not given and that most outcomes were patient-reported. We also conducted sensitivity analyses to investigate if results differed within the clinical network when the subclinical studies were excluded. Furthermore, owing to the great variance of included conditions within each of the two networks and due to considerable heterogeneity in the nonclinical network, we performed subgroup analysis for two broad areas – pain and psychological conditions.

Results

A total of 12’991 records were retrieved by bibliographic database and registries searching. After removing duplicates, 6’811 remained and their title and abstracts were screened together with 21 additionally identified records. Subsequently, 731 full-texts were screened. Thirty-seven RCTs (comprising 3’021 participants) conducted between 2010 and 2022 comparing 12 interventions and 3 control groups met all of the eligibility criteria and were included into our analyses. A flow chart detailing the process of study identification and selection is shown in eFigure 1 in the supplement. All studies were reported in English and included an adult sample with a mean (SD) age of 36 (15.3) years (range 19-70 years). All selected outcomes were of continuous nature (see eTable 1 in the supplement for details on selected outcomes). The individual characteristics of the 37 studies included in the analysis are given in eTable 1 in the supplement.

Nonclinical sample

Twelve studies yielded sufficient data to be included in the analysis of the nonclinical sample (comprising 1'015 participants). The sample sizes of individual studies ranged from 21 to 151. The mean (SD) age of this sample was 23.6 (2.1) years (range: 20–28 years), and 67.7% of the sample population were female. Four studies examined experimentally induced pain, three itch, two sadness, one acute stress, one nausea, and one tested muscle strength. All studies were single-center studies except one[38]. Eight trials recruited participants from Europe (Germany, Netherlands, Switzerland and UK), three from North America (USA) and one from Australia. The studies had different routes of placebo administration such as nasal (4 studies), dermal (6 studies) and oral (2 studies). Ten studies used a NT control, nine included a DP and two an OLP- condition. One study used conditioning in order to evoke positive treatment expectations, all others used verbal suggestions.

Figure 1A shows the network of eligible comparisons and figure 2A shows the forest plot of the NMA including all treatments and control groups using NT as a reference. In this network, only nasal OLPs were significantly better than NT (SMD = 0.43, [0.02–0.84]). Dermally applied conditioned and unconditioned OLPs as well as OLP pills were not significantly better compared to NT (SMDs ranging from 0.10, [-0.60–0.80] to 0.47, [-0.33–1.28]). OLP- was worse than NT (SMD = -0.60, [-1.15– -0.05]). (RQ2) The investigation of head to head comparisons (see eTable 2 in the supplement) of different OLP modalities revealed no significant differences with SMDs ranging from 0.04, [-0.83–0.92] to 0.38, [-0.68–1.43]. (RQ3) OLPs without the induction of treatment expectation were statistically worse compared to all other OLP modalities (SMDs ranging from -0.86, [-1.41– -0.31] to -1.07, [-2.02– -0.12]) except for the comparison with OLP pills (SMD = -0.69, [-1.57– 0.19]). (RQ4) Within this network there was only one comparator (i.e., NT). Therefore, differential effects depending on the comparison groups used could not be investigated.

Clinical sample

The analysis of the clinical sample included 25 studies with 2'006 participants and sample sizes of individual studies ranging from 19 to 211. The mean (SD) age of this sample was 43.7 (14.9) years (range: 19–70 years), and 70.7% were female. The different populations used in the 25 included studies were the following: chronic low back pain (4 studies), allergic rhinitis (3 studies), cancer-related fatigue (3 studies), irritable bowel syndrome (2 studies), knee osteoarthritis (2 studies), major depressive disorder (2 studies), acute pain (following spine surgery; 1 study), acute pain (spinal cord injury and polytrauma; 1 study), chronic low back pain + experimental pain (1 study), menopausal hot flushes (1 study), primary insomnia (1 study), relaxation test (1 study), test anxiety (1 study), well-being (1 study), and well-being + cognitive enhancement (1 study). The mean duration of the treatment phase was three weeks (range: 1 day to 12 weeks). No study was multicentered. Thirteen trials recruited patients from Europe (Germany, Austria, Denmark, Portugal), eight from North America (USA), three from Asia (Japan and Israel), and one from Australia. Various routes of placebo administration were used such as nasal (4 studies) and dermal (5 studies) applications as well as injections (2 studies). Furthermore, oral applications included pills (21 studies) and suspensions (2 studies). Nine studies used a NT control condition, five TAU and eight a WL. Furthermore, two studies included a DP, two an OLP- group, one a psychological intervention and one a treatment program (exercise and education intervention) as a

comparator group. Overall, three studies used a conditioning paradigm to induce positive treatment expectation.

Figure 1B depicts the clinical network with eligible comparisons and figure 2A shows the Forest plot of the NMA including all treatments and control groups using NT as reference. In the clinical network, conditioned and unconditioned OLP pills outperformed NT (0.89, [0.01–1.76] to 0.46, [0.28–0.65], respectively). Injected OLPs and conditioned and unconditioned OLP suspensions were not statistically better than NT (SMDs ranging from 0.23, [-0.54–1.01] to 0.70, [-0.14–1.54]). (RQ2) The investigation of head to head comparisons (see eTable 2 in the supplement) of different OLP modalities revealed no significant differences with SMDs ranging from -0.08, [-1.18–1.01] to 0.65, [-0.50–1.81]. (RQ3) OLPs without the induction of treatment expectation were not statistically different from any other OLP modality (SMDs ranging from -0.26, [-1.03–0.51] to -0.92, [-1.87–0.04]) except for the comparison with OLP pills, here OLP- was significantly worse (SMD = -0.49, [-0.92– -0.07]). (RQ4) The investigation of the effects of treatment comparators showed that OLP pills was in addition to NT also significantly better than WL (SMD = 0.43, [0.22 – 0.64]) but not TAU (SMD = 0.16, [-0.48–0.80]). In addition, cOLP pills was significantly better than TAU (SMD = 0.58, [0.02–1.15]), and marginally not significant compared with WL, yet the effect was high (SMD = 0.86, [-0.02–1.74]).

Results of sensitivity analyses, adverse events and certainty of evidence assessment can be found in the supplement (see eAppendix 6).

Discussion

This systematic review and NMA of RCTs with 3’021 individuals assessed the efficacy of various OLP interventions in comparison to different types of control groups both in a nonclinical and clinical sample. The aim was to examine whether the size of the OLP effect is different across (RQ1) nonclinical vs. clinical populations, (RQ2) treatment modalities, (RQ3) treatment expectation, and (RQ4) comparator groups. Across both networks, a wide range of conditions was studied with pain and diverse psychological conditions being the most frequent.

Within the nonclinical sample the NMA revealed a significant effect of OLP administered as a spray or vapor (i.e, OLP nasal) compared with NT (SMD = 0.43). All other OLP interventions showed small to medium but insignificant SMDs compared with NT. Similar results were found for the clinical sample, where only OLP pills outperformed NT (SMD conditioned = 0.89; unconditioned= 0.46), with again all other modalities showing insignificant but small to medium effects. Even though only some OLP modalities were significantly better than NT, the comparison of the different employed OLP modalities in both networks showed no significant differences. However, OLPs without the induction of treatment expectation were statistically worse compared to the majority of OLP modalities within the nonclinical network (SMD ranging from -0.86 to -1.07) and compared to OLP pills in the clinical network (SMD = -0.49). Finally, the comparison of treatment comparator groups in the clinical network showed that OLP pills were better than WL (SMD = 0.43) but not better than TAU (SMD = 0.16).

In the following, the observed effects will be discussed with regard to the four distinct research questions. In order to investigate differential effect sizes across the nonclinical and clinical sample (RQ1), we compared the findings of both networks qualitatively. We found that the effect sizes for the comparison of OLP pills to NT yielded smaller and nonsignificant effects within the nonclinical sample (i.e., SMD nonclinical = 0.10; clinical = 0.46). This trend was also exemplified by the comparison of DP against NT (SMD nonclinical = 0.50; clinical = 0.76). Similar observations have previously been reported for both somatic and psychological conditions: For example, studies investigating placebo analgesia have found an average effect size of 1.24 in healthy individuals and an effect size of 1.49 in patients[14]. This finding not only suggests that DPs employed in OLP studies tend to yield smaller effects as compared to studies investigating DPs only, but also sheds light on the difference between the effect sizes of placebo effects across nonclinical and clinical samples. In this regard, our two networks may support the notion that placebo effects tend to be of greater magnitude in clinical as opposed to nonclinical populations. This trend was supported by our sensitivity analysis, where effect sizes were slightly bigger when excluding subclinical studies. A potential explanation could be the more pronounced desire of relief in patients as opposed to healthy individuals[55]. In summary, this finding suggests that clinical and subclinical populations might benefit from OLP treatments to a greater degree than healthy individuals and that experimental studies on healthy individuals may underestimate the magnitude of the OLP effect in patients. However, this comparison is only qualitative in nature and therefore could be further explored as part of a single study.

In terms of OLP modalities (RQ2), none of the direct comparisons were statistically significant, indicating that there might not be a difference in the effect across OLP intervention modalities in either sample. This finding stands in contrast to the roam of DP, where it is known that more invasive routes of administration can yield bigger effects compared to less invasive procedures[19,21]. This discrepancy suggests that findings from DP research might not be valid for the field of OLP. However, SMDs varying up to 0.50 across modalities suggest (especially within the clinical sample) that the current analyses might be underpowered in order to observe statistically significant differences. However, there is also reason to assume that potential differences in OLP modalities may be obscured given that the present analyses investigated the efficacy of OLP treatments across a variety of different somatic and psychological conditions. Supporting this line of reasoning, Peerdeman et al. (2017)[56] found that expectations towards the efficacy of different routes of administration differed for pain and itch, e.g., injected medications were expected to be most effective for relieving pain and topical medications for alleviating itching. These results might reflect the impact of knowledge and prior experience on treatment expectations. Regardless, placebo experts advise against prescribing more invasive treatments to yield stronger effects, as this entails practical and ethical limitation[1]. Especially the yet small database for OLPs calls for a cautious consideration regarding the use of more invasive procedures.

Regarding our research question on the impact of expectation (RQ3), evidence from both networks suggests that OLP interventions delivered without the evocation of at least minimal treatment expectations are less efficacious as compared to OLP interventions with the induction of treatment expectation. This finding was especially pronounced within the nonclinical network, where OLP- was less efficacious as compared to all other groups within the network, even to NT (SMD = -0.60). However,

the efficacy of OLP- within the nonclinical network was solely evaluated by two trials investigating dermal placebo applications[8,26]. Nevertheless, it appears that expectancy building is an important component of OLP interventions and that simply prescribing an inert treatment is not sufficient. Hence, OLP treatments might be cost-efficient but not as time-efficient as over the counter medicine. Possible explanations for this observation could be that participants do not feel taken seriously when they are simply told that they are receiving a placebo treatment, or that they are disappointed because they may not know about the power of placebo effects. The effects in the clinical setting (where no differences between OLP- and NT were observed) might potentially be buffered by at least performing a ritual of e.g. taking pills over a period of time. In a single administration, which was often the case in the nonclinical studies, no such ritual could be established. Therefore, the rationale seems to be an essential and potentially indispensable component for the efficacy of OLP[57]. In this sense, OLP conditions that do not include any expectation building component could at best serve as control groups, controlling for the component of the pill. The pill itself might therefore not be necessary to produce positive treatment effects in OLP studies. This finding is supported by a recently published RCT on OLPs and imaginary pills (Buergler et al. 2023).

With respect to the potential impact of different comparators (RQ4), our systematic search showed that due to the experimental setting all nonclinical studies used a NT control group. Differences across control groups could thus only be investigated within the clinical sample. There, we identified three different comparison groups, namely NT, WL and TAU. Comparison of effect sizes across different comparator groups showed that OLP pills was in addition to NT also significantly better than WL (SMD = 0.43) but not TAU (SMD = 0.16). In other words, this finding could imply that OLP pills are better than "nothing", but not better than "something". Thus, the efficacy of both of these interventions seems to depend on the kind of control group used, a finding in line with psychotherapy research[28]. However, whereas there WL was notably inferior to NT, in the present study both comparison groups yielded comparable effects. Conditioned OLP pills, on the other hand, were significantly better than NT as well as TAU (SMD = 0.58) and tended to be better than WL (SMD = 0.86; *n.s.*). However, these findings are based on two studies only, indicating that the obtained conclusions are not entirely conclusive and should be further explored. Nevertheless, these findings suggest that comparator groups within OLP studies should be chosen carefully as the effects might differ according to the chosen comparator.

Overall, the present analyses confirm the results of previous meta-analyses investigating the efficacy of OLP in clinical populations, which found moderate to high effect sizes[4,5]. In contrast, the results of the nonclinical sample contradict in part the findings by Spille et al. (2022)[6], which found a medium sized effect for subjective outcomes for OLPs in comparison to NT. This difference in findings might be explained by different inclusion criteria and thus another body of studies that contributed to the results (e.g., the inclusion of subclinical studies in their analysis) and might further be fostered by their differentiation between objective and subjective outcomes. Remarkably, the herein found effect sizes for OLP pills in the clinical sample were smaller as compared to previous investigations, which included only OLP pills in comparison to different control conditions (SMD = 0.46 vs. 0.88[4] and 0.72[5]). This trend towards smaller effects across the timespan suggests that in an early state of research, "positive" studies

are more likely to be published (reporting bias – which was also present within the clinical sample of this NMA) and with time insignificant results are more likely to be published (time lag bias).

This study has several strengths. First, the direct comparison of different placebo intervention modalities and comparators is of great importance to inform the young research field of OLPs about the comparative efficacy in order to better design future studies. The network meta-analytic approach uniquely allows investigating the effects of different modalities and comparators. Second, this methodology allows combining direct and indirect evidence to get the most precise estimate of the intervention differences. Third, we were able to include 13 more studies than the newest existing meta-analysis on OLP in clinical conditions[5], which strengthens the body of evidence. Fifth, the clear definition of OLPs is a strength of this analysis as well as the several sensitivity analyses that were conducted, which showed comparable results that further supported the trends of the overall analyses. Finally, the application of the same inclusion criteria for the nonclinical and the clinical sample allows to more reliably compare effect sizes across both populations. However, this study has several limitations that should be taken into account when interpreting the results. First, although the network meta-analytic approach allowed to include 12 studies within the nonclinical network and 25 within the clinical, which represents a considerably broad range of studies as compared to previous analysis, the relatively small number of studies in each node and the resulting small power might have led to a lack of significance (large confidence intervals). Second, a major limitation of our NMAs is associated with the fact that most interventions have been tested in less than 100 participants. It is therefore possible that the effect of some of these interventions is owing to a so-called small-study effect: smaller trials show different, often larger, treatment effects than bigger ones[58,59]. Third, substantial heterogeneity was found in our NMAs. The variety of the studied conditions, the format of the interventions (e.g., duration), and the reported outcomes differed widely, which may have contributed to the statistical heterogeneity and certainly to the clinical heterogeneity. However, we tried to reduce heterogeneity by applying a very precise and strict definition of OLPs, by conducting two separate networks and by choosing the most frequent outcome, in case of the presence of several outcomes. Furthermore, sensitivity analyses suggest that the results remain unchanged when looking at more homogeneous subgroups within the network as for example pain. Fourth, although NMAs have the advantage of making use of all available data, the indirect evidence does not directly stem from randomized comparisons[60]. Fifth, according to the GRADE framework, the within-study bias of many comparisons was assessed as “some concerns”, which can be attributed in part by methodological difficulties which arise through the nature of OLPs (i.e., participants being unblinded) and the nature of most outcomes being self-reported. Sixth, funnel plots and accompanying Egger’s tests indicated a risk for reporting bias for the clinical network because of the lack of small studies comparing NT versus OLP pills with negative effects. Seventh, we excluded cross-over studies due to analytical concerns regarding comparability with parallel-trials[34], which reduced the body of evidence to parallel trials with accompanying loss in power. Eighth, the clinical sample also includes undiagnosed subclinical conditions, which limits the comparability to other studies, which only included studies with diagnosed samples (e.g.,[5]). However, this was accounted for by conducting a sensitivity analysis which yielded comparable results. Tenth, the relatively early stage of OLP research did not allow to investigate the efficacy of OLPs within distinct conditions. Therefore, the

present NMA examined the interventions on meta-level lumping studies with different conditions, which might impair the requirement for NMAs of included populations being in theory jointly randomizable. Finally, the results on the comparable efficacy of OLP modalities might be explained by population specific choices of treatment modalities, obscuring potential differences within each domain.

With this NMA, we were able to identify several research gaps: First, larger studies should be conducted, as sample sizes are often relatively small (range: 19 - 211). Second, the population should be more representative: Currently, the majority of the study population is female (70%) and especially in the nonclinical sample very young (mean age: 23.6 years). This complicates, among other things, the transfer of nonclinical findings to the clinical population, which was on average older (mean age: 43.7 years). Third, adverse events should be reported more structured and consistently. Because of not or inconsistently reported adverse events, we were not able not analyze them in the present study. Fourth, it would be crucial to conduct future studies by more independent research teams with less allegiance to OLP research. Fifth, in future studies, the control group used should be chosen deliberately, because depending on the type of control group – as our study shows – different sizes of effects result. Also, in further meta-analysis, control groups should not be lumped together, as this can obscure possible treatment effects. Sixth, in a further (network) meta-analyses, it would be informative to distinguish between active and non-active OLP, which was not considered in these analyses. Seventh, further experimental studies should be designed more according to the needs of clinical populations: For instance, the OLP modalities OLP nasal (e.g., spays) and OLP dermal (e.g., creams) were only studied in nonclinical populations and not in clinical, possibly indicating that this route of administration is not suitable for clinical conditions. Eighth, future (network) meta-analyses should take into account that potential differences between OLP modalities may be masked, as their effects may differ depending on the type of disease. Finally, in order to reduce within study bias, future research should include objective outcomes and behavioral markers.

To conclude, OLPs can be beneficial compared to control conditions in nonclinical and clinical conditions. However, the magnitude of effects appears to be smaller compared to previous meta-analyses and further depend on several aspects that we have considered in our NMAs. (1) We identified a trend for greater effect sizes within the clinical network. Hence, research in nonclinical samples may underestimate the magnitude of OLP effects in patients. (2) There were no differences in the effect across OLP modalities in either sample. This finding calls for a cautious consideration regarding the use of more invasive OLP procedures. (3) Inducing positive treatment expectation is of great importance for the efficacy of OLPs. Simply prescribing an OLP seems not to be enough and might even hold the risk of being worse than receiving nothing. (4) Finally, we found that OLP effects can vary depending on the comparator used. In other words, some interventions facilitate relief when compared to “nothing” but their effect appears to vanish when compared to other treatments. With this NMA, we hope to expand the knowledge in the emerging research field of OLPs and inform future studies aimed at exploring ethical ways to use placebo effects for the good of patients.

Additional statements

Acknowledgments: The authors would like to thank all masters students: Laura Tincknell, Sarah Marie Hanig, Jana Hochreuter, Jennifer Lopez and interns: Sebastian Penz, Lea Masanti, Cyrill Hübscher as well as our research assistants: Yannick Schulz for aiding with data screening, extraction, eTable1 and references and Berfin Bakis with her help double-checking the results. Furthermore, the authors want to thank Hannah Ewald and Christian Appenzeller for their help with developing the search strategy and conducting the search and Guido Schwarzer for his help regarding analyzing the data in R.

Contributors: Authors contributed to the work as follows: Conceptualization: SB, DS, CL, JG; Methodology: SB, DS, CL; Formal Analyses: SB, DS, CL; Writing—original draft: SB, DS; Writing—review & editing: SB, DS, CL, JG; Visualization: SB, DS; Supervision: CL, JG; Project Administration: SB, DS

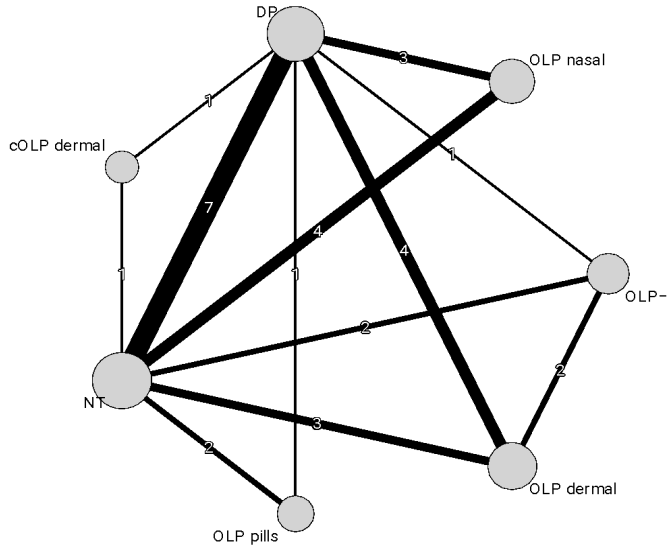
Competing interests' statement: The authors have no conflicts of interest to disclose.

Data availability statement: Data will be shared upon reasonable request addressed to sarah.buergler@unibas.ch.

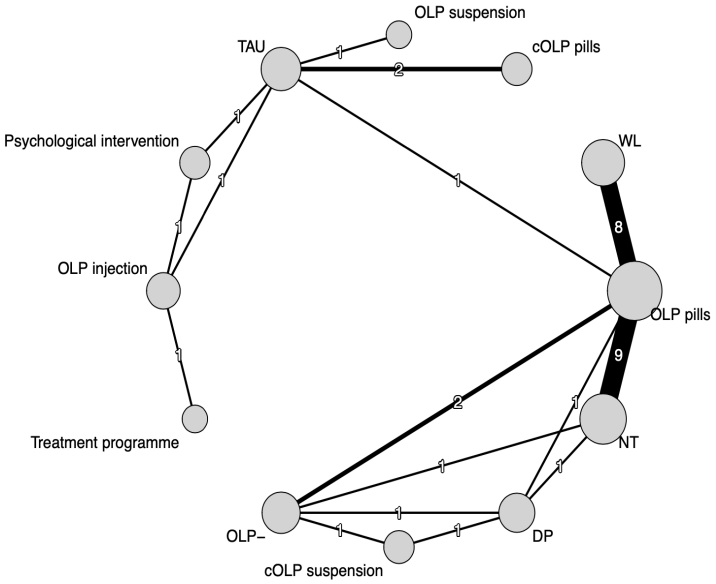
Funding: CL received funding from the Swiss National Science Foundation (SNSF): PZ00P1_201972**.**

Figures

Figure 1. Network Meta-analysis of eligible comparisons



A. Network meta-analysis of eligible comparisons for the nonclinical sample

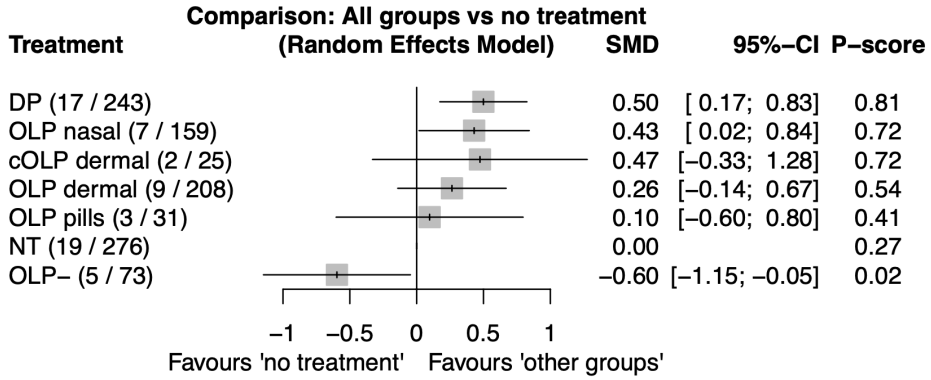


B. Network meta-analysis of eligible comparisons for the clinical sample

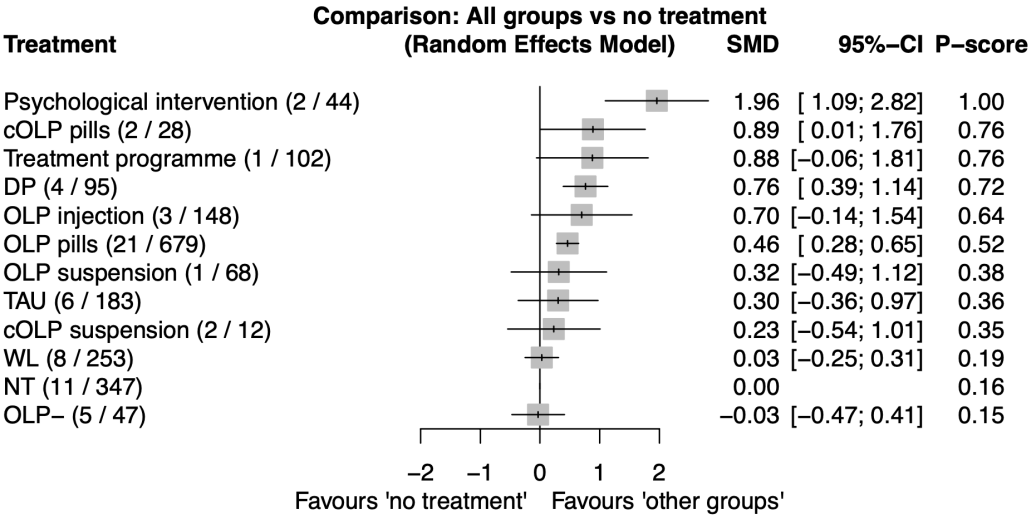
A, nonclinical. B, clinical. Width of the lines is proportional to the number of trials comparing every pair of treatments/groups.

Note. cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; TAU, Treatment as Usual; WL, Wait List.

Figure 2. Forest plot of network meta-analysis of all trials



A. Forest plot of network meta-analysis of all trials for the nonclinical sample



B. Forest plot of network meta-analysis of all trials for the clinical sample

All groups were compared with no treatment (NT), which was the reference group. The brackets behind the group names indicate the following: number of direct comparisons with this group/number of patients in which the intervention/control was examined. SMD indicates standardized mean difference.

Note. cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; TAU, Treatment as Usual; WL, Wait List.

References

1. Evers, A. W. M. *et al.* Implications of Placebo and Nocebo Effects for Clinical Practice: Expert Consensus. *Psychother Psychosom* **87**, 204–210 (2018).
2. Evers, A. W. M. *et al.* What Should Clinicians Tell Patients about Placebo and Nocebo Effects? Practical Considerations Based on Expert Consensus. *Psychother Psychosom* **90**, 49–56 (2021).
3. Blease, C., Colloca, L. & Kaptchuk, T. J. Are open-Label Placebos Ethical? Informed Consent and Ethical Equivocations. *Bioethics* **30**, 407–414 (2016).
4. Charlesworth, J. E. G. *et al.* Effects of placebos without deception compared with no treatment: A systematic review and meta-analysis. *J Evid Based Med* **10**, 97–107 (2017).
5. von Wernsdorff, M., Loef, M., Tuschen-Caffier, B. & Schmidt, S. Effects of open-label placebos in clinical trials: a systematic review and meta-analysis. *Sci. Rep.* **11**, 3855 (2021).
6. Spille, L., Fendel, J. C., Seuling, P., Göritz, A. S. & Schmidt, S. Open-label placebos: A systematic review and meta-analysis of experimental studies with non-clinical samples. (2022) doi:https://doi.org/10.21203/rs.3.rs-2093533/v1.
7. Sezer, D., Locher, C. & Gaab, J. Deceptive and open-label placebo effects in experimentally induced guilt: a randomized controlled trial in healthy subjects. *Sci. Rep.* **12**, 21219 (2022).
8. Locher, C. *et al.* Is the rationale more important than deception? A randomized controlled trial of open-label placebo analgesia. *Pain* **158**, 2320–2328 (2017).
9. Lembo, A. *et al.* Open-label placebo vs double-blind placebo for irritable bowel syndrome: a randomized clinical trial. *Pain* **162**, 2428–2435 (2021).
10. Mundt, J. M., Roditi, D. & Robinson, M. E. A comparison of deceptive and non-deceptive placebo analgesia: efficacy and ethical consequences. *Ann. Behav. Med.* **51**, 307–315 (2017).
11. Disley, N., Kola-Palmer, S. & Retzler, C. A comparison of open-label and deceptive placebo analgesia in a healthy sample. *J. Psychosom. Res.* **140**, 110298 (2021).
12. Kube, T. *et al.* Deceptive and Nondeceptive Placebos to Reduce Pain: An Experimental Study in Healthy Individuals. *Clin J Pain* **36**, 68–79 (2020).
13. Sandler, A. D. & Bodfish, J. W. Open-label use of placebos in the treatment of ADHD: a pilot study. *Child Care Health Dev* **34**, 104–10 (2008).
14. Forsberg, J. T., Martinussen, M. & Flaten, M. A. The Placebo Analgesic Effect in Healthy Individuals and Patients: A Meta-Analysis. *Psychosom Med* **79**, 388–394 (2017).
15. Charron, J., Rainville, P. & Marchand, S. Direct comparison of placebo effects on clinical and experimental pain. *Clin J Pain* **22**, 204–11 (2006).
16. Kaptchuk, T. J. *et al.* Sham device v inert pill: randomised controlled trial of two placebo treatments. *Bmj* **332**, 391–7 (2006).
17. de Craen, A. J., Tijssen, J. G., de Gans, J. & Kleijnen, J. Placebo effect in the acute treatment of migraine: subcutaneous placebos are better than oral placebos. *J Neurol* **247**, 183–8 (2000).
18. Macedo, A., Farré, M. & Baños, J. E. A meta-analysis of the placebo response in acute migraine and how this response may be influenced by some of the characteristics of clinical trials. *Eur J Clin Pharmacol* **62**, 161–72 (2006).
19. van Laarhoven, A. I. M. *et al.* Placebo effects on itch: a meta-analysis of clinical trials of patients with dermatological conditions. *J Invest Dermatol* **135**, 1234–1243 (2015).
20. Bannuru, R. R. *et al.* Effectiveness and Implications of Alternative Placebo Treatments: A Systematic Review and Network Meta-analysis of Osteoarthritis Trials. *Ann Intern Med* **163**, 365–72 (2015).
21. Meissner, K. *et al.* Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. *JAMA Intern Med* **173**, 1941–51 (2013).
22. Colloca, L. & Miller, F. G. How placebo responses are formed: a learning perspective. *Philos Trans R Soc Lond B Biol Sci* **366**, 1859–69 (2011).
23. Kaptchuk, T. J. *et al.* Placebos without deception: a randomized controlled trial in irritable bowel

syndrome. *PLoS One* **5**, e15591 (2010).

24. Haas, J. W., Winkler, A., Rheker, J., Doering, B. K. & Rief, W. No open-label placebo effect in insomnia? Lessons learned from an experimental trial. *J Psychosom Res* **158**, 110923 (2022).
25. Klinger, R., Kothe, R., Schmitz, J., Kamping, S. & Flor, H. Placebo effects of a sham opioid solution: a randomized controlled study in patients with chronic low back pain. *Pain* **158**, 1893–1902 (2017).
26. Leibowitz, K. A., Hardebeck, E. J., Goyer, J. P. & Crum, A. J. The role of patient beliefs in open-label placebo effects. *Health Psychol* **38**, 613–622 (2019).
27. Schaefer, M., Sahin, T. & Berstecher, B. 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 (2018).
28. Michopoulos, I. *et al.* Different control conditions can produce different effect estimates in psychotherapy trials for depression. *J. Clin. Epidemiol.* **132**, 59–70 (2021).
29. Blease, C. R., Bernstein, M. H. & Locher, C. Open-label placebo clinical trials: is it the rationale, the interaction or the pill? *BMJ Evid Based Med* **25**, 159–165 (2020).
30. Watts, S. E., Turnell, A., Kladnitski, N., Newby, J. M. & Andrews, G. Treatment-as-usual (TAU) is anything but usual: a meta-analysis of CBT versus TAU for anxiety and depression. *J Affect Disord* **175**, 152–67 (2015).
31. Cuijpers, P., Quero, S., Papola, D., Cristea, I. A. & Karyotaki, E. Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: a meta-analysis. *Psychol. Med.* **51**, 634–644 (2021).
32. Liberati, A. *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* **6**, e1000100 (2009).
33. Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L. & Bekhuis, T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc* **104**, 240–3 (2016).
34. Higgins, J. P. *et al. Cochrane handbook for systematic reviews of interventions*. (John Wiley & Sons, 2019).
35. Gervasio, J., Zheng, S., Skrotzki, C. & Pachete, A. The effect of oral contraceptive use on cortisol reactivity to the Trier Social Stress Test: A meta-analysis. *Psychoneuroendocrinology* **136**, 105626 (2022).
36. Koechlin, H. *et al.* Nonpharmacological interventions for pediatric migraine: a network meta-analysis. *Pediatrics* **147**, (2021).
37. Shi, C., Westby, M., Norman, G., Dumville, J. C. & Cullum, N. Node-making processes in network meta-analysis of nonpharmacological interventions should be well planned and reported. *J. Clin. Epidemiol.* **101**, 124–125 (2018).
38. Friehs, T., Rief, W., Glombiewski, J. A., Haas, J. & Kube, T. Deceptive and non-deceptive placebos to reduce sadness: A five-armed experimental study. *J. Affect. Disord. Rep.* **9**, 100349 (2022).
39. Barnes, K., Yu, A., Josupeit, J. & Colagiuri, B. Deceptive but not open label placebos attenuate motion-induced nausea. *J. Psychosom. Res.* **125**, 109808 (2019).
40. Kube, T., Hofmann, V. E., Glombiewski, J. A. & Kirsch, I. Providing open-label placebos remotely—A randomized controlled trial in allergic rhinitis. *PLoS One* **16**, e0248367 (2021).
41. El Brihi, J., Horne, R. & Faasse, K. Prescribing Placebos: An Experimental Examination of the Role of Dose, Expectancies, and Adherence in Open-Label Placebo Effects. *Ann Behav Med* **53**, 16–28 (2019).
42. Sterne, J. A. C. *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj* **366**, l4898 (2019).
43. Jelena Savovic. Personal communication. (2022).
44. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. vol. 2nd ed. (Routledge., 1988).
45. Rücker, G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* **3**, 312–24 (2012).

46. Rücker, G., Krahn, U., König, J., Efthimiou, O. & Schwarzer, G. Netmeta: network meta-analysis using frequentist methods. *R Package Version 09–0* (2016).

47. Efthimiou, O. *et al.* GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* **7**, 236–63 (2016).

48. Dias, S., Welton, N. J., Caldwell, D. M. & Ades, A. E. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* **29**, 932–44 (2010).

49. Higgins, J. P. *et al.* Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* **3**, 98–110 (2012).

50. Salanti, G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* **3**, 80–97 (2012).

51. Rücker, G. & Schwarzer, G. Resolve conflicting rankings of outcomes in network meta-analysis: Partial ordering of treatments. *Res Synth Methods* **8**, 526–536 (2017).

52. Duval, S. & Tweedie, R. A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *J. Am. Stat. Assoc.* **95**, 89–98 (2000).

53. Duval, S. & Tweedie, R. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463 (2000).

54. Chiocchia, V. *et al.* ROB-MEN: a tool to assess risk of bias due to missing evidence in network meta-analysis. *BMC Med.* **19**, 1–13 (2021).

55. Hyland, M. E. Motivation and placebos: do different mechanisms occur in different contexts? *Philos. Trans. R. Soc. B Biol. Sci.* **366**, 1828–1837 (2011).

56. Peerdeman, K. J. *et al.* Expectations about the effectiveness of pain- and itch-relieving medication administered via different routes. *Eur J Pain* **22**, 774–783 (2018).

57. Locher, C. *et al.* Lay perspectives of the open-label placebo rationale: a qualitative study of participants in an experimental trial. *BMJ Open* **11**, e053346 (2021).

58. Sterne, J. A., Gavaghan, D. & Egger, M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* **53**, 1119–29 (2000).

59. Sterling, T. D., Rosenbaum, W. L. & Weinkam, J. J. Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa. *Am. Stat.* **49**, 108–112 (1995).

60. Chaimani, A., Salanti, G., Leucht, S., Geddes, J. R. & Cipriani, A. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. *Evid Based Ment Health* **20**, 88–94 (2017).

Supplementary online content

The role of population, expectation, modality and comparator on open-label placebo effects: A network meta-analysis

Sarah Buergler*, Dilan Sezer*, Jens Gaab & Cosima Locher
*shared first authorship

eAppendix 1. Search strategies and hits

Medline Ovid

(20210202; 956 hits)

(((placebo* or sham) adj2 (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*)).ti,ab,kw,kf. or open placebo*.ti,ab,kw,kf. or ((placebos/ or Placebo Effect/) and (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*).ti,ab.)) and (exp Random Allocation/ or exp Randomized Controlled Trial/ or exp Randomized Controlled Trials as Topic/ or RCT or (randomiz* or randomis*).ti,ab. or ((controlled clinical or non-inferiority or noninferiority or superiority or equivalence or pragmatic) ADJ2 trial$).ti,ab.)

Embase Ovid

(20210202; 5,487 hits)

(((placebo* or sham) adj2 (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*)).ti,ab,kw. or open placebo*.ti,ab,kw. or ((placebo/ or Placebo Effect/ or sham procedure/) and (Open study/ or (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*).ti,ab.))) and (randomization/ or exp randomized controlled trial/ or randomized controlled trial topic/ or "randomized controlled trial (topic)"/ or RCT or (randomiz* or randomis*).ti,ab. or ((controlled clinical or non-inferiority or noninferiority or superiority or equivalence or pragmatic) ADJ2 trial$).ti,ab.)

NOT (conference abstract or conference review).pt

CINAHL Ebsco

(20210202; 589 hits)

((((TI placebo* OR AB placebo*) OR (TI sham OR AB sham)) N2 ((TI open-label* OR AB open-label*) OR (TI told OR AB told) OR (TI nondecept* OR AB nondecept*) OR (TI non-decept* OR AB non-decept*) OR (TI nonconceal* OR AB nonconceal*) OR (TI "non conceal*" OR AB "non conceal*") OR (TI unconceal* OR AB unconceal*) OR (TI unblind* OR AB unblind*) OR (TI nonblind* OR AB nonblind*) OR (TI "non blind*" OR AB "non blind*") OR (TI "without decept*" OR AB "without decept*") OR (TI "without conceal*" OR AB "without conceal*") OR (TI "without blind*" OR AB "without blind*"))) OR (TI "open placebo*" OR AB "open placebo*") OR (((MH "placebos") OR (MH "Placebo Effect")) AND ((TI open-label* OR AB open-label*) OR (TI told OR AB told) OR (TI nondecept* OR AB nondecept*) OR (TI non-decept* OR AB non-decept*) OR (TI nonconceal* OR AB nonconceal*) OR (TI "non conceal*" OR AB "non conceal*") OR (TI unconceal* OR AB unconceal*) OR (TI unblind* OR AB unblind*) OR (TI nonblind* OR AB nonblind*) OR (TI "non blind*" OR AB "non blind*") OR (TI "without decept*" OR AB "without decept*") OR (TI "without conceal*" OR AB "without conceal*") OR (TI "without blind*" OR AB "without blind*"))))

PsycINFO Ovid

(20210202; 406 hits)

(((placebo* or sham) adj2 (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*)).ti,ab. or open placebo*.ti,ab. or (placebo/ and (open-label* or told or nondecept* or non-decept* or nonconceal* or non conceal* or unconceal* or unblind* or nonblind* or non blind* or without decept* or without conceal* or without blind*).ti,ab.))

eAppendix 2. Hits update

Medline Ovid

(20210201 bis 20220608; 66 hits)

limit *SEARCH* to dt=20210201-20220608

Embase Ovid

(20210201 bis 20220608; 640 hits)

limit *SEARCH* to dc=20210201-20220608

CINAHL Ebsco

(20210201 bis 20220608; 38 hits)

PsycINFO Ovid

(20210201 bis 20220608; 43 hits)

limit *SEARCH* to up=20210201-20220608

eAppendix 3. GRADE Ratings for each network

We used the Grading of Recommendations Assessment, Development, and Evaluation ratings (GRADE[1]) and the corresponding web application to apply this framework[2,3]. The certainty of evidence for each network estimate was assessed according to the following criteria:

Study limitations (Within study bias): The overall risk of bias of each study was categorized. According to the Cochrane Risk of Bias tool 2[4], we rated five risk of bias domains. We then used the contribution matrix to calculate the percentage of contribution from each study, and finally assessed the study limitation for each network estimate based on the weighted average risk of bias of the contributing studies. We selected the rule "Average Risk of Bias" in order to calculate the within study bias.

Reporting bias (Across studies bias): Since each of our comparisons had less than 10 comparisons, we could not use the ROB-MEN[5] tool to assess reporting bias. Therefore, a comparison-adjusted funnel plot with accompanying Egger test for asymmetry was conducted and used as a basis for the judgment.

Indirectness: We judged that there was no concern in this domain as the included studies matched our inclusion criteria and study questions.

Imprecision: In line with previous analyses[6], we considered a clinically meaningful threshold for standardized mean difference (SMD) to be 0.20.

Heterogeneity: We evaluated the degree of concerns through comparing the clinical inference based on the 95% confidence intervals (CI), the latter reflecting the degree of heterogeneity. Appling the same clinical inference framework as for imprecision, we saw no concerns in heterogeneity when the two judgements matched (e.g. no concern based on 95% CI and no concern based on 95% PI), some concerns when they differed by one degree (e.g. no concern based on 95% CI but some concerns based on 95% PI), and major concerns when they differed by two degrees (e.g. no concern based on 95% CI but major concerns based on 95% PI).

Incoherence (Inconsistency): For inconsistency, we looked at the results of side splitting and we saw major concerns when $p<0.05$ but no concern otherwise.

Nonclinical network

We found some concerns for *within-study bias* (i.e., study limitations) for all pairwise comparisons, due to the nature of the studies being unblind and most outcomes being self-reported. In terms of the *across-study bias* (i.e., reporting bias), the Egger test for funnel plot asymmetry was non-significant (p = .666) indicating that selection bias is not a big threat to the network meta-analysis. There was no concern for *indirectness*, since the included studies all matched our study questions. Evaluating *imprecision*, we found that all statistically significant comparisons revealed a clinically significant effect size. Furthermore, we examined *heterogeneity*, which is represented by the 95% prediction interval for each individual comparison. For all statistically significant comparisons there were at least some concerns regarding heterogeneity, indicating that there is a high variability of effects. Furthermore, we found no evidence for substantial and statistically significant heterogeneity in the network as a whole (within design Q = 2.27, p = .811, tau2 = 0.13; $I2$ = 66%). Finally, there was no evidence of incoherence between the direct and indirect evidence, i.e., all p-values were above 5%. For those comparisons where only indirect evidence was available incoherence was set to major concerns. Also, we identified evidence of inconsistency in the NMA when calculating the global design-by-treatment interaction test (between designs Q = 41.43, p < .001).

Comparison	Number of Studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence
			Mixed evidence				
DP vs NT	7	Some concerns	Low risk	No concerns	No concerns	Major concerns	No concerns
DP vs OLP dermal	4	Some concerns	Low risk	No concerns	Some concerns	Some concerns	No concerns
DP vs OLP nasal	3	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns
DP vs OLP pills	1	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns
DP vs OLP-	1	Some concerns	Low risk	No concerns	No concerns	No concerns	No concerns
DP vs cOLP dermal	1	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns
NT vs OLP dermal	3	Some concerns	Low risk	No concerns	Some concerns	Some concerns	No concerns
NT vs OLP nasal	4	Some concerns	Low risk	No concerns	No concerns	Major concerns	No concerns
NT vs OLP pills	2	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns
NT vs OLP-	2	Some concerns	Low risk	No concerns	No concerns	Major concerns	No concerns
NT vs cOLP dermal	1	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns
OLP dermal vs OLP-	2	Some concerns	Low risk	No concerns	No concerns	Some concerns	No concerns

		Indirect evidence					
OLP dermal vs OLP nasal	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP dermal vs OLP pills	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP dermal vs cOLP dermal	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP nasal vs OLP pills	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP nasal vs OLP-	--	Some concerns ☐	Low risk	No concerns	No concerns	Some concerns ☐	Major concerns ☐
OLP nasal vs cOLP dermal	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP pills vs OLP-	--	Some concerns ☐	Low risk	No concerns	Some concerns ☐	Some concerns ☐	Major concerns ☐
OLP pills vs cOLP dermal	--	Some concerns ☐	Low risk	No concerns	Major concerns ☐	No concerns	Major concerns ☐
OLP- vs cOLP dermal	--	Some concerns ☐	Low risk	No concerns	No concerns	Major concerns ☐	Major concerns ☐

Clinical network

We found some concerns for *within-study bias* (i.e., study limitations) for most pairwise comparisons, due to the nature of the studies being unblind and most outcomes being self-reported. In terms of the *across-study bias* (i.e., reporting bias), the Egger test for funnel plot asymmetry was significant (p = .036) indicating that reporting bias is a threat to the network meta-analysis. There was no concern for *indirectness*, since the included studies all matched our study questions. Evaluating *imprecision*, we found that all statistically significant comparisons revealed a clinically significant effect size, except for two comparisons (cOLP suspension vs. DP, cOLP suspension vs. OLP-) where we found major concerns regarding the clinical significance of observed effects. Furthermore, we examine *heterogeneity*, which is represented by the 95% prediction interval for each individual comparison. For three statistically significant comparisons (TAU vs. cOLP pills, NT vs. cOLP pills, OLP- vs. OLP pills) there were some concerns regarding heterogeneity, indicating that there is some variability of effects. All other significant comparisons revealed no concerns. Furthermore, we found no evidence for substantial and statistically significant heterogeneity in the network as a whole (within design Q = 12.62, p = .557, tau2 = 0.024; $I2$ = 26.5%). Finally, there was evidence of incoherence between the direct and indirect evidence in three comparisons, i.e., cOLP suspension vs. OLP-, cOLP suspension vs. DP, DP vs. OLP-. For those comparisons where only indirect evidence was available incoherence was set to major concerns. Also, we identified evidence of inconsistency in the NMA when calculating the global design-by-treatment interaction test (between designs Q = 11.86, p = .018).

Comparison	Number of Studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence
			Mixed evidence				
DP vs NT	1	Some concerns	Some concerns	No concerns	No concerns	No concerns	Some concerns
DP vs OLP pills	1	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	No concerns
DP vs OLP-	1	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
DP vs cOLP suspension	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
NT vs OLP pills	9	Some concerns	Some concerns	No concerns	No concerns	No concerns	No concerns
NT vs OLP-	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	No concerns
OLP injection vs Psychological intervention	1	No concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
OLP injection vs TAU	1	No concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
OLP injection vs Treatment programme	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs OLP-	2	Some concerns	Some concerns	No concerns	No concerns	Some concerns	Major concerns
OLP pills vs TAU	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs WL	8	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
OLP suspension vs TAU	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP- vs cOLP suspension	1	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
Psychological intervention vs TAU	1	No concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
TAU vs cOLP pills	2	Some concerns	Some concerns	No concerns	No concerns	Some concerns	Major concerns

		Indirect evidence					
DP vs OLP injection	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
DP vs OLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
DP vs Psychological intervention	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
DP vs TAU	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
DP vs Treatment programme	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
DP vs WL	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
DP vs cOLP pills	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
NT vs OLP injection	--	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
NT vs OLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
NT vs Psychological intervention	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
NT vs TAU	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
NT vs Treatment programme	--	Some concerns	Some concerns	No concerns	Some concerns	No concerns	Major concerns
NT vs WL	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
NT vs cOLP pills	--	Some concerns	Some concerns	No concerns	No concerns	Some concerns	Major concerns
NT vs cOLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP injection vs OLP pills	--	No concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP injection vs OLP suspension	--	No concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP injection vs OLP-	--	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
OLP injection vs WL	--	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
OLP injection vs cOLP pills	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns

OLP injection vs cOLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs OLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs Psychological intervention	--	No concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
OLP pills vs Treatment programme	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs cOLP pills	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP pills vs cOLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP suspension vs OLP-	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP suspension vs Psychological intervention	--	No concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
OLP suspension vs Treatment programme	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP suspension vs WL	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP suspension vs cOLP pills	--	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
OLP suspension vs cOLP suspension	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP- vs Psychological intervention	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
OLP- vs TAU	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP- vs Treatment programme	--	Some concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns
OLP- vs WL	--	Some concerns	Some concerns	No concerns	Major concerns	No concerns	Major concerns
OLP- vs cOLP pills	--	Some concerns	Some concerns	No concerns	Some concerns	No concerns	Major concerns
Psychological intervention vs Treatment programme	--	No concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
Psychological intervention vs WL	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
Psychological intervention vs cOLP pills	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
Psychological intervention vs cOLP suspension	--	Some concerns	Some concerns	No concerns	No concerns	No concerns	Major concerns
TAU vs Treatment programme	--	No concerns	Some concerns	No concerns	Some concerns	Some concerns	Major concerns

TAU vs WL	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐
TAU vs cOLP suspension	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐
Treatment programme vs WL	--	Some concerns ☐	Some concerns ☐	No concerns	Some concerns ☐	Some concerns ☐	Major concerns ☐
Treatment programme vs cOLP pills	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐
Treatment programme vs cOLP suspension	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐
WL vs cOLP pills	--	Some concerns ☐	Some concerns ☐	No concerns	Some concerns ☐	No concerns	Major concerns ☐
WL vs cOLP suspension	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐
cOLP pills vs cOLP suspension	--	Some concerns ☐	Some concerns ☐	No concerns	Major concerns ☐	No concerns	Major concerns ☐

eAppendix 4. Details on inconsistency

Nonclinical network – local approach

```
Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

             comparison k prop   nma direct indir.  Diff     z p-value
         cOLP dermal:DP 1 0.78 -0.03  -0.09   0.20 -0.29 -0.29  0.7720
         cOLP dermal:NT 1 0.78  0.47   0.54   0.25  0.29  0.29  0.7720
 cOLP dermal:OLP dermal 0    0  0.21      .   0.21     .     .       .
  cOLP dermal:OLP nasal 0    0  0.04      .   0.04     .     .       .
  cOLP dermal:OLP pills 0    0  0.38      .   0.38     .     .       .
       cOLP dermal:OLP- 0    0  1.07      .   1.07     .     .       .
                 DP:NT 7 0.85  0.50   0.47   0.66 -0.19 -0.42  0.6773
          DP:OLP dermal 4 0.77  0.24   0.10   0.70 -0.61 -1.28  0.2011
           DP:OLP nasal 3 0.74  0.07   0.21  -0.33  0.55  1.09  0.2751
           DP:OLP pills 1 0.34  0.40  -0.01   0.62 -0.63 -0.79  0.4306
                DP:OLP- 1 0.43  1.10   1.44   0.84  0.60  1.02  0.3078
          OLP dermal:NT 3 0.70  0.26   0.20   0.41 -0.21 -0.47  0.6404
           OLP nasal:NT 4 0.87  0.43   0.50  -0.03  0.53  0.86  0.3890
           OLP pills:NT 2 0.91  0.10  -0.00   1.09 -1.09 -0.89  0.3747
                OLP-:NT 2 0.83 -0.60  -0.70  -0.12 -0.57 -0.77  0.4386
   OLP dermal:OLP nasal 0    0 -0.17      .  -0.17     .     .       .
   OLP dermal:OLP pills 0    0  0.17      .   0.17     .     .       .
        OLP dermal:OLP- 2 0.85  0.86   0.86   0.85  0.01  0.01  0.9882
    OLP nasal:OLP pills 0    0  0.33      .   0.33     .     .       .
         OLP nasal:OLP- 0    0  1.03      .   1.03     .     .       .
         OLP pills:OLP- 0    0  0.69      .   0.69     .     .       .

Legend:
 comparison - Treatment comparison
 k          - Number of studies providing direct evidence
 prop       - Direct evidence proportion
 nma        - Estimated treatment effect (SMD) in network meta-analysis
 direct     - Estimated treatment effect (SMD) derived from direct evidence
 indir.     - Estimated treatment effect (SMD) derived from indirect evidence
 Diff       - Difference between direct and indirect treatment estimates
 z          - z-value of test for disagreement (direct versus indirect)
 p-value    - p-value of test for disagreement (direct versus indirect)
```

Nonclinical network – global approach

```
Q statistics to assess homogeneity / consistency

                    Q df  p-value
Total           43.69 15   0.0001
Within designs   2.27  5   0.8112
Between designs 41.43 10 < 0.0001

Design-specific decomposition of within-designs Q statistic

          Design    Q df p-value
   DP:OLP dermal 0.12  1  0.7324
 NT:DP:OLP nasal 2.15  4  0.7083

Between-designs Q statistic after detaching of single designs

       Detached design     Q df  p-value
         DP:OLP dermal 40.30  9 < 0.0001
          NT:OLP nasal 37.82  9 < 0.0001
          NT:OLP pills 41.14  9 < 0.0001
      NT:DP:OLP dermal 40.17  8 < 0.0001
 NT:DP:OLP dermal:OLP-  6.98  7   0.4306
       NT:DP:OLP nasal 35.92  8 < 0.0001
       NT:DP:OLP pills 40.19  8 < 0.0001
    NT:OLP dermal:OLP- 14.51  8   0.0693

Q statistic to assess consistency under the assumption of
a full design-by-treatment interaction random effects model

                    Q df  p-value tau.within tau2.within
Between designs 41.43 10 < 0.0001          0           0
```

Clinical network – local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison	k	prop	nma	direct	indir.	Diff	z	p-value
cOLP pills:cOLP suspension	0	0	0.65	.	0.65	.	.	.
cOLP pills:DP	0	0	0.12	.	0.12	.	.	.
cOLP pills:NT	0	0	0.89	.	0.89	.	.	.
cOLP pills:OLP injection	0	0	0.19	.	0.19	.	.	.
cOLP pills:OLP pills	0	0	0.42	.	0.42	.	.	.
cOLP pills:OLP suspension	0	0	0.57	.	0.57	.	.	.
cOLP pills:OLP-	0	0	0.92	.	0.92	.	.	.
cOLP pills:Psychological intervention	0	0	-1.07	.	-1.07	.	.	.
cOLP pills:TAU	2	1.00	0.58	0.58	.	.	.	.
cOLP pills:Treatment programme	0	0	0.01	.	0.01	.	.	.
cOLP pills:WL	0	0	0.86	.	0.86	.	.	.
cOLP suspension:DP	1	0.89	-0.53	-0.93	2.58	-3.51	-2.98	0.0029
cOLP suspension:NT	0	0	0.23	.	0.23	.	.	.
cOLP suspension:OLP injection	0	0	-0.47	.	-0.47	.	.	.
cOLP suspension:OLP pills	0	0	-0.23	.	-0.23	.	.	.
cOLP suspension:OLP suspension	0	0	-0.08	.	-0.08	.	.	.
cOLP suspension:OLP-	1	0.78	0.26	0.89	-1.91	2.80	2.98	0.0029
cOLP suspension:Psychological intervention	0	0	-1.72	.	-1.72	.	.	.
cOLP suspension:TAU	0	0	-0.07	.	-0.07	.	.	.
cOLP suspension:Treatment programme	0	0	-0.65	.	-0.65	.	.	.
cOLP suspension:WL	0	0	0.20	.	0.20	.	.	.
DP:NT	1	0.69	0.76	0.52	1.31	-0.79	-1.92	0.0546
DP:OLP injection	0	0	0.06	.	0.06	.	.	.
DP:OLP pills	1	0.68	0.30	0.11	0.72	-0.61	-1.51	0.1298
DP:OLP suspension	0	0	0.45	.	0.45	.	.	.
DP:OLP-	1	0.35	0.79	1.82	0.24	1.58	2.98	0.0029
DP:Psychological intervention	0	0	-1.19	.	-1.19	.	.	.
DP:TAU	0	0	0.46	.	0.46	.	.	.
DP:Treatment programme	0	0	-0.11	.	-0.11	.	.	.
DP:WL	0	0	0.73	.	0.73	.	.	.
OLP injection:NT	0	0	0.70	.	0.70	.	.	.
OLP pills:NT	9	0.99	0.46	0.47	-0.39	0.86	1.09	0.2749
OLP suspension:NT	0	0	0.32	.	0.32	.	.	.
OLP-:NT	1	0.33	-0.03	0.42	-0.26	0.68	1.42	0.1564
Psychological intervention:NT	0	0	1.96	.	1.96	.	.	.
TAU:NT	0	0	0.30	.	0.30	.	.	.
Treatment programme:NT	0	0	0.88	.	0.88	.	.	.
WL:NT	0	0	0.03	.	0.03	.	.	.
OLP injection:OLP pills	0	0	0.24	.	0.24	.	.	.
OLP injection:OLP suspension	0	0	0.38	.	0.38	.	.	.
OLP injection:OLP-	0	0	0.73	.	0.73	.	.	.
OLP injection:Psychological intervention	1	1.00	-1.26	-1.26	.	.	.	.
OLP injection:TAU	1	1.00	0.39	0.39	.	.	.	.
OLP injection:Treatment programme	1	1.00	-0.18	-0.18	.	.	.	.
OLP injection:WL	0	0	0.67	.	0.67	.	.	.
OLP pills:OLP suspension	0	0	0.15	.	0.15	.	.	.
OLP pills:OLP-	2	0.69	0.49	0.23	1.07	-0.84	-1.77	0.0770
OLP pills:Psychological intervention	0	0	-1.49	.	-1.49	.	.	.
OLP pills:TAU	1	1.00	0.16	0.16	.	.	.	.
OLP pills:Treatment programme	0	0	-0.41	.	-0.41	.	.	.
OLP pills:WL	8	1.00	0.43	0.43	.	.	.	.
OLP suspension:OLP-	0	0	0.35	.	0.35	.	.	.
OLP suspension:Psychological intervention	0	0	-1.64	.	-1.64	.	.	.
OLP suspension:TAU	1	1.00	0.01	0.01	.	.	.	.
OLP suspension:Treatment programme	0	0	-0.56	.	-0.56	.	.	.
OLP suspension:WL	0	0	0.28	.	0.28	.	.	.
OLP-:Psychological intervention	0	0	-1.99	.	-1.99	.	.	.
OLP-:TAU	0	0	-0.34	.	-0.34	.	.	.
OLP-:Treatment programme	0	0	-0.91	.	-0.91	.	.	.
OLP-:WL	0	0	-0.06	.	-0.06	.	.	.
Psychological intervention:TAU	1	1.00	1.65	1.65	.	.	.	.
Psychological intervention:Treatment programme	0	0	1.08	.	1.08	.	.	.
Psychological intervention:WL	0	0	1.92	.	1.92	.	.	.
TAU:Treatment programme	0	0	-0.57	.	-0.57	.	.	.
TAU:WL	0	0	0.27	.	0.27	.	.	.
Treatment programme:WL	0	0	0.85	.	0.85	.	.	.

Legend:

comparison	Treatment comparison
k	Number of studies providing direct evidence
prop	Direct evidence proportion
nma	Estimated treatment effect (SMD) in network meta-analysis
direct	Estimated treatment effect (SMD) derived from direct evidence
indir.	Estimated treatment effect (SMD) derived from indirect evidence
Diff	Difference between direct and indirect treatment estimates
z	z-value of test for disagreement (direct versus indirect)
p-value	p-value of test for disagreement (direct versus indirect)

Clinical network – global approach

```
Q statistics to assess homogeneity / consistency

                    Q df p-value
Total           24.48 18  0.1400
Within designs  12.62 14  0.5569
Between designs 11.86  4  0.0184

Design-specific decomposition of within-designs Q statistic

         Design    Q df p-value
 cOLP pills:TAU 0.27  1  0.6041
   NT:OLP pills 4.40  6  0.6224
   OLP pills:WL 7.94  7  0.3375

Between-designs Q statistic after detaching of single designs

   Detached design     Q df p-value
      NT:OLP pills 11.86  3  0.0079
    OLP pills:OLP-  9.45  3  0.0239
   NT:DP:OLP pills  0.96  2  0.6199
 NT:OLP pills:OLP-  9.30  2  0.0096

Q statistic to assess consistency under the assumption of
a full design-by-treatment interaction random effects model

                    Q df p-value tau.within tau2.within
Between designs 11.86  4  0.0184          0           0
```

eAppendix 5. PRISMA checklist

Section/Topic	Item	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review *incorporating a network meta-analysis (or related form of meta-analysis).*	p.1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: **Background:** main objectives **Methods:** data sources; study eligibility criteria, participants, and interventions; study appraisal; and *synthesis methods, such as network meta-analysis.* **Results:** number of studies and participants identified; summary estimates with corresponding confidence/credible intervals; *treatment rankings may also be discussed. Authors may choose to summarize pairwise comparisons against a chosen treatment included in their analyses for brevity.* **Discussion/Conclusions:** limitations; conclusions and implications of findings. **Other:** primary source of funding; systematic review registration number with registry name.	p.2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known*, including mention of why a network meta-analysis has been conducted.*	p.3-4
Objectives	4	Provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	p.4
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists and if and where it can be accessed (e.g., Web address); and, if available, provide registration information, including registration number.	p.6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Clearly describe eligible treatments included in the treatment network, and note whether any have been clustered or merged into the same node (with justification).*	p.5-6

Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	p.5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	p.5 eAppendix 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	p.5-6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	p.6-7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	p.5-7
Geometry of the network	**S1**	Describe methods used to explore the geometry of the treatment network under study and potential biases related to it. This should include how the evidence base has been graphically summarized for presentation, and what characteristics were compiled and used to describe the evidence base to readers.	p.7-9
Risk of bias within individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	p.7-8 eAppendix 3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Also describe the use of additional summary measures assessed, such as treatment rankings and surface under the cumulative ranking curve (SUCRA) values, as well as modified approaches used to present summary findings from meta-analyses.*	p.8-9
Planned methods of analysis	14	Describe the methods of handling data and combining results of studies for each network meta-analysis. This should include, but not be limited to: ● *Handling of multi-arm trials;* ● *Selection of variance structure;* ● *Selection of prior distributions in Bayesian analyses; and* ● *Assessment of model fit.*	p.8-9
Assessment of Inconsistency	**S2**	Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network(s) studied. Describe efforts taken to address its presence when found.	p.8-9 eAppendix 4

Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	p.9 eAppendix 3
Additional analyses	16	Describe methods of additional analyses if done, indicating which were pre-specified. This may include, but not be limited to, the following: ● Sensitivity or subgroup analyses; ● Meta-regression analyses; ● *Alternative formulations of the treatment network; and* ● *Use of alternative prior distributions for Bayesian analyses (if applicable).*	p.9 eAppendix 6
RESULTS†			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	p.10 eFigure 1
Presentation of network structure	**S3**	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	Figures 1A, 1B
Summary of network geometry	**S4**	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	eTable 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	p.10-11 eTable 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	eTable 1 eAppendix 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	Figures 2A, 2B eTable 2, 3
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g. placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	p.10-11 Figures 2A, 2B eTable 2, 3

Exploration for inconsistency	**S5**	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	eAppendix 4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	eTable 1 eAppendix 3
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses*, alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses,* and so forth).	eAppendix 6
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy-makers).	p.12-14
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network geometry (e.g., avoidance of certain comparisons).*	p.14-15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	p. 16
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	p.17

PICOS = population, intervention, comparators, outcomes, study design.

* Text in italics indicate wording specific to reporting of network meta-analyses that has been added to guidance from the PRISMA statement.

† Authors may wish to plan for use of appendices to present all relevant information in full detail for items in this section.

eAppendix 6. Additional Results

Adverse events

Regarding adverse events, it is remarkable that few studies reported adverse events systematically or at all. In total, 15 of the 37 studies made a statement regarding adverse events. From these reports, it is apparent that relatively few adverse events occur in the context of OLP treatment. This suggests that OLP is a safe and mostly side effect free treatment. However, due to inconsistent or unreported adverse events, it is difficult to draw a conclusion.

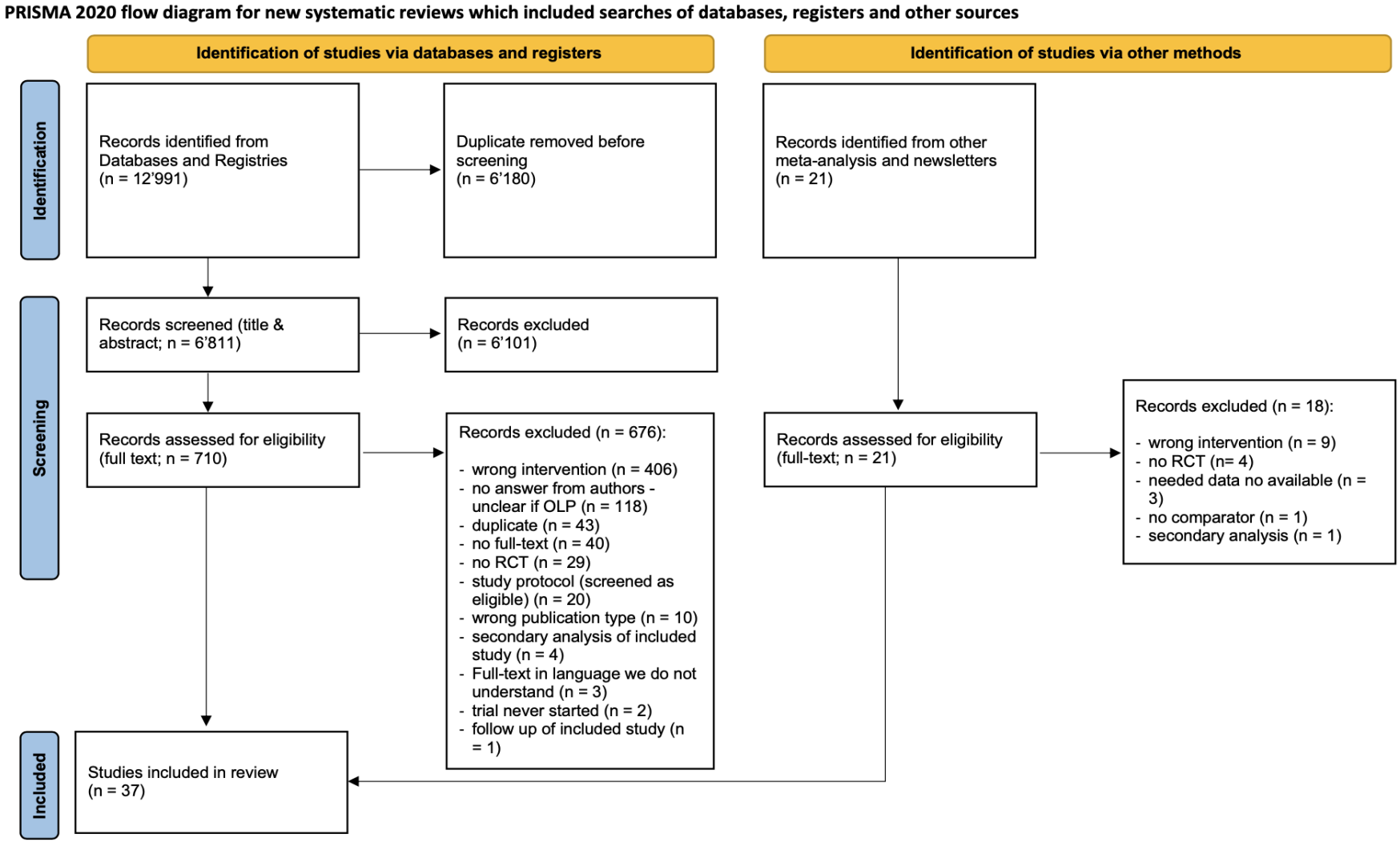
Certainty of the evidence

The certainty of evidence for the network estimates of both samples was examined by using GRADE. The results for study limitations (within study bias), reporting bias (across-studies bias), indirectness, imprecision, heterogeneity, and incoherence can be found in the supplement (eAppendix 3-4, eFigure 2).

Sensitivity analysis

To investigate the impact of high risk studies, we conducted the analyses including only studies in which the risk of bias was low or moderate. In each sample, one study was high risk of bias and thus excluded and compared to the whole sample. The results in the nonclinical network remained unchanged in principal, solely OLP nasal changed from being marginally significant to insignificant. In the nonclinical sample, cOLP pills moved from being significant to non significant, as only one study with a cOLP pills group remained in the network. Otherwise results and heterogeneity measures remained comparable. To investigate the impact of including studies with subclinical populations within the clinical sample, we conducted a sensitivity analysis by excluding studies with subclinical samples. In principle, the results remained unchanged with a trend for slightly bigger effect sizes when subclinical studies were excluded (see eFigure 3-6 in the supplement for the results of sensitivity analyses). Surprisingly, heterogeneity increased from I^2 = 26.5% (clinical all) to I^2 = 32.6% (clinical without subclinical). Furthermore, owing to the great variance of included conditions within each of the two networks, we performed subgroup analysis for two broad areas: pain (i.e., chronic back pain, experimental pain, irritable bowel syndrome, knee ostheoarthritis) and psychological (i.e., depression, fatigue, conditions, well-being, insomnia, test anxiety, sadness, relaxation, stress). The results for the clinical pain network (11 studies) showed comparable results to the ones of the whole network, except the treatment programme changed to being significantly better than NT, whereas OLP- moved to being significantly worse than NT. Interestingly, heterogeneity was reduced from I^2 = 26.5% (clinical all) to I^2 = 0% (clinical pain). Within the nonclinical pain sample (N = 4), results did also change only marginally, with OLP nasal not being significantly better than NT anymore. Heterogeneity as well decreased from I^2 = 66% (nonclinical all) to I^2 = 51.7% (nonclinical pain). Within the psychological subsamples results could in general also be replicated (clinical psychological = 10 and nonclinical psychological = 3 studies), with the exception of DP being bigger in the nonclinical sample and the effect size of OLP- changing from -0.03 to 0.30 in the clinical network. Heterogeneity decreased within the clinical sample from I^2 = 26.5% (clinical all) to I^2 = 0% (clinical psychological) and in the nonclinical network from I^2 = 66% (nonclinical all) to I^2 = 0% (nonclinical psychological). Overall, very few studies were included in the networks of these subgroup-analyses.

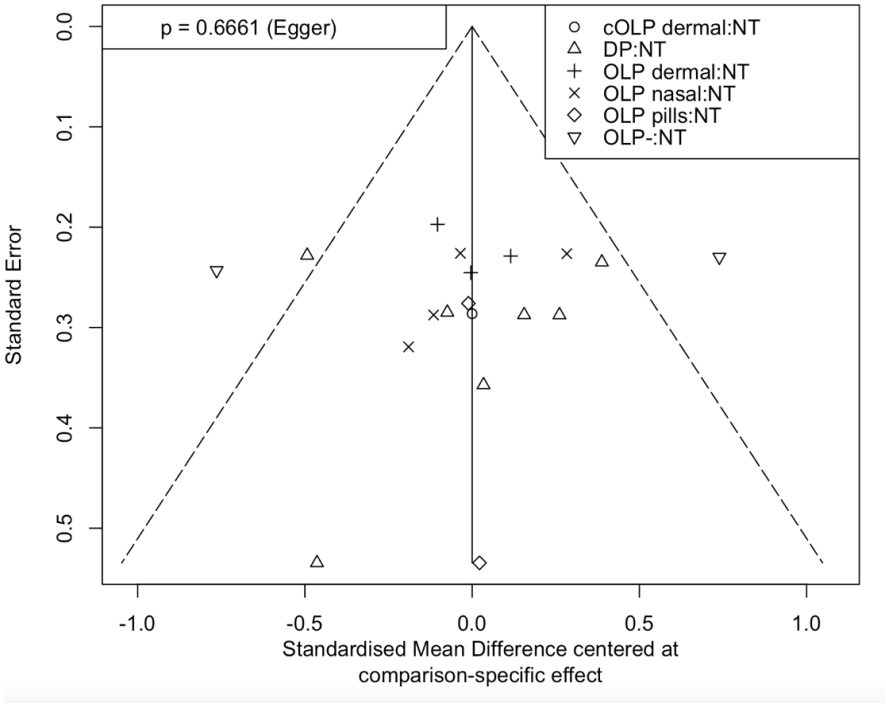
eFigure 1. Flowchart



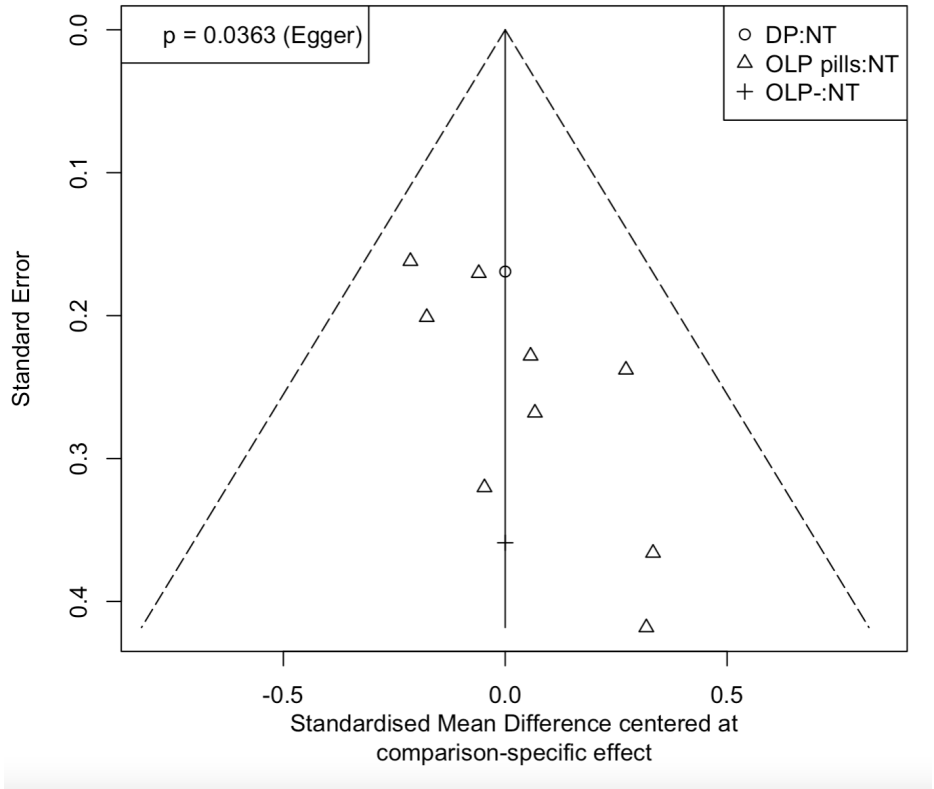
From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

eFigure 2. Funnel plots with accompanying Egger test

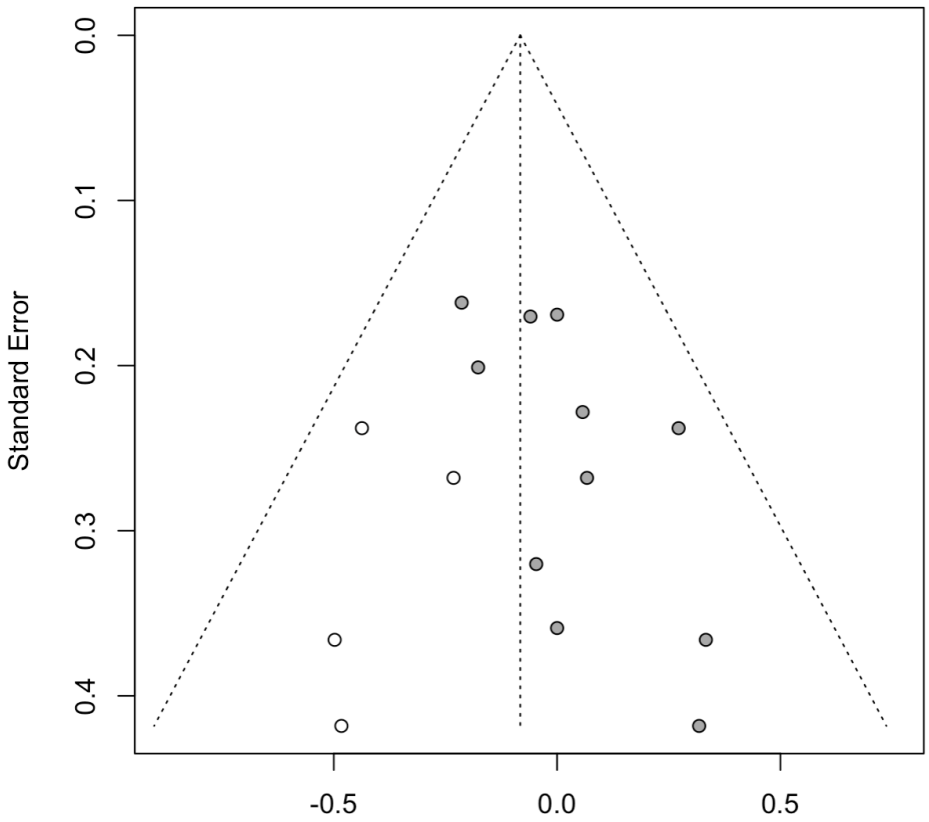
Nonclinical network



Clinical network



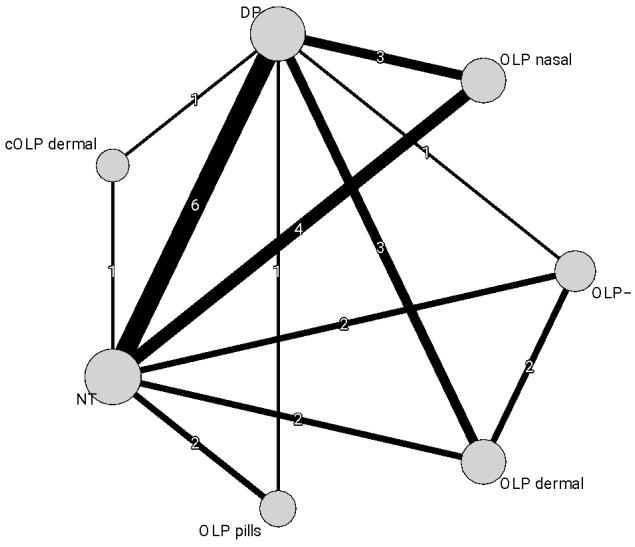
Note: Funnel plot with reference NT, i.e. this plot only includes studies with NT as a control group depicting available comparisons with DP and OLP.



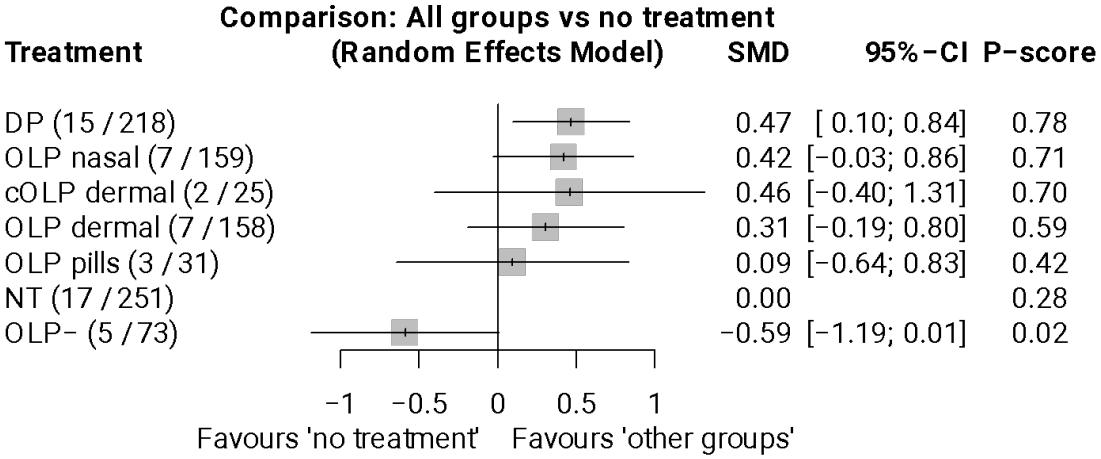
Note: Funnel plot with reference NT after using the trim and fill method. The plot depicts the four comparisons (white dots) NT vs. OLP pills that are missing in order for the funnel to be symmetric.

eFigure 3. Plots of low and moderate risk of bias only (sensitivity analysis)

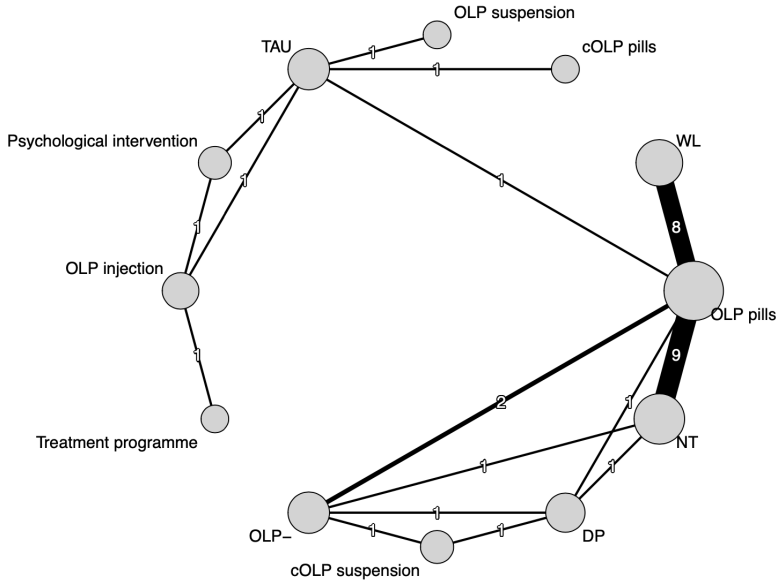
Netgraph of nonclinical network meta-analysis on low and moderate risk of bias only



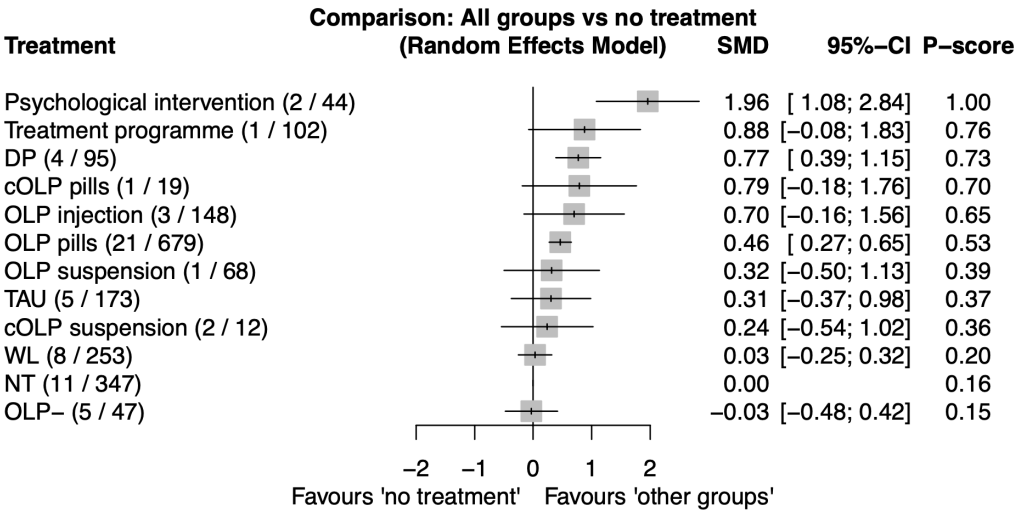
Forest plot of nonclinical network meta-analysis on low and moderate risk of bias only



Netgraph of clinical network meta-analysis on low and moderate risk of bias only

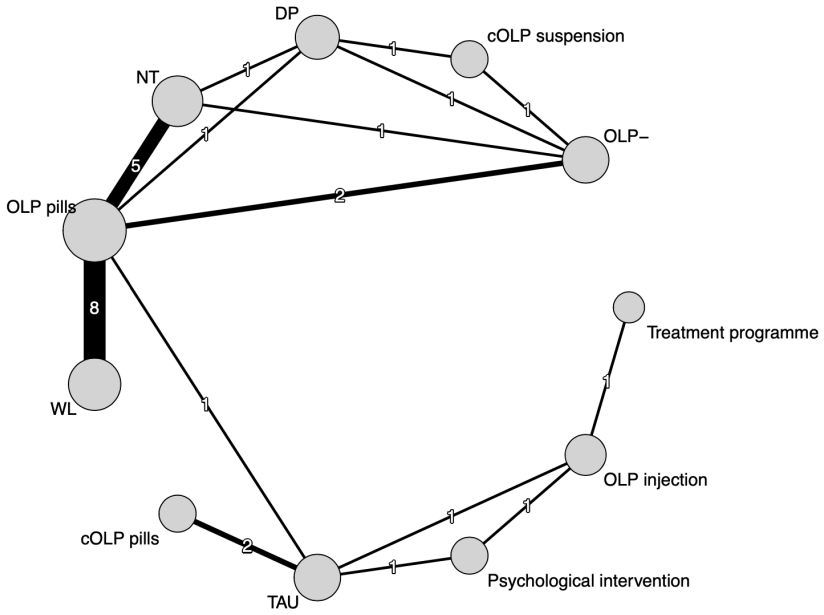


Forest plot of clinical network meta-analysis on low and moderate risk of bias only

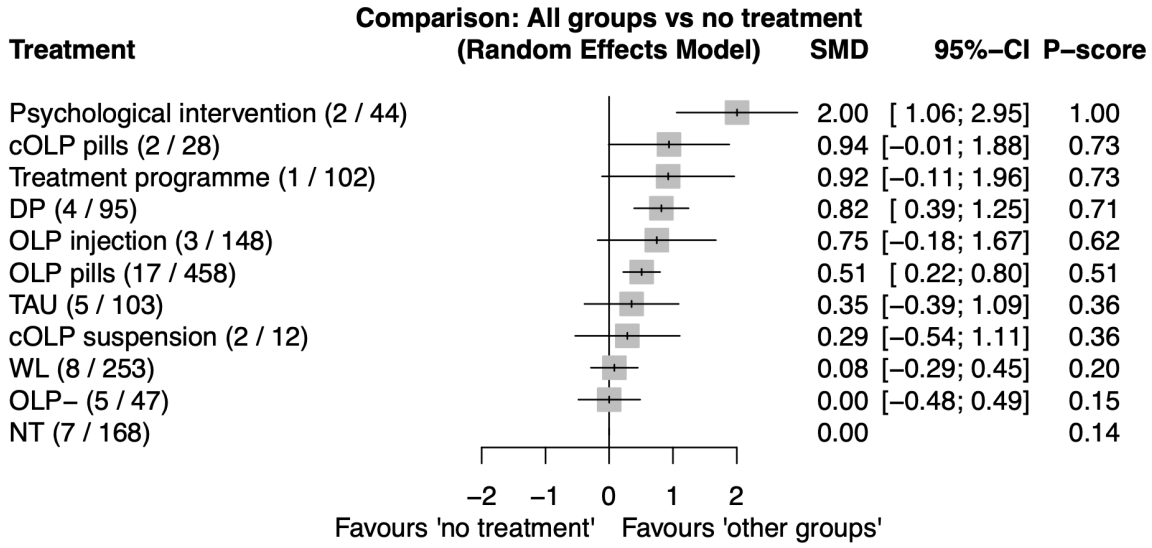


eFigure 4. Plots of clinical network without subclinical trials (sensitivity analysis)

Netgraph of network meta-analysis on clinical studies only

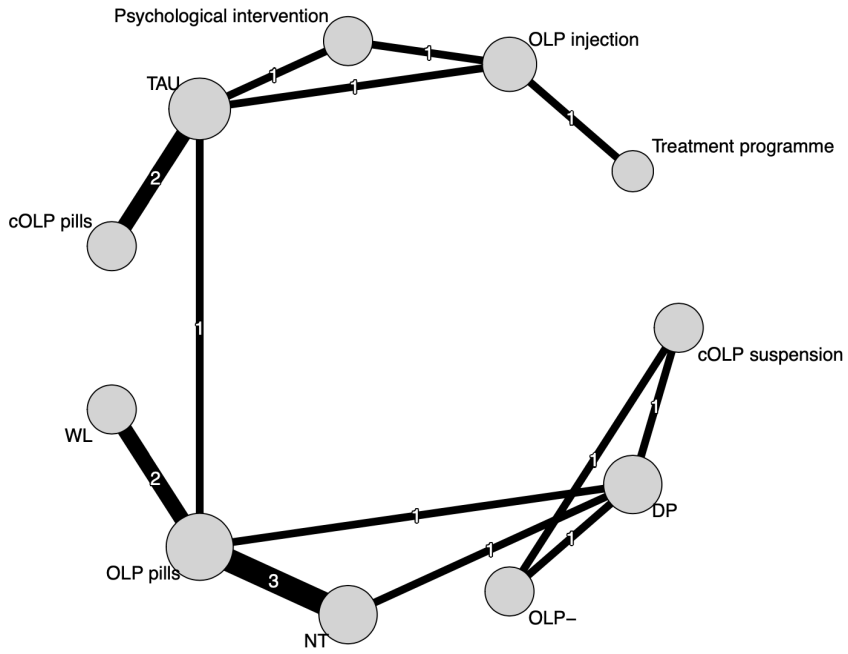


Forest plot of network-meta-analysis on clinical studies only

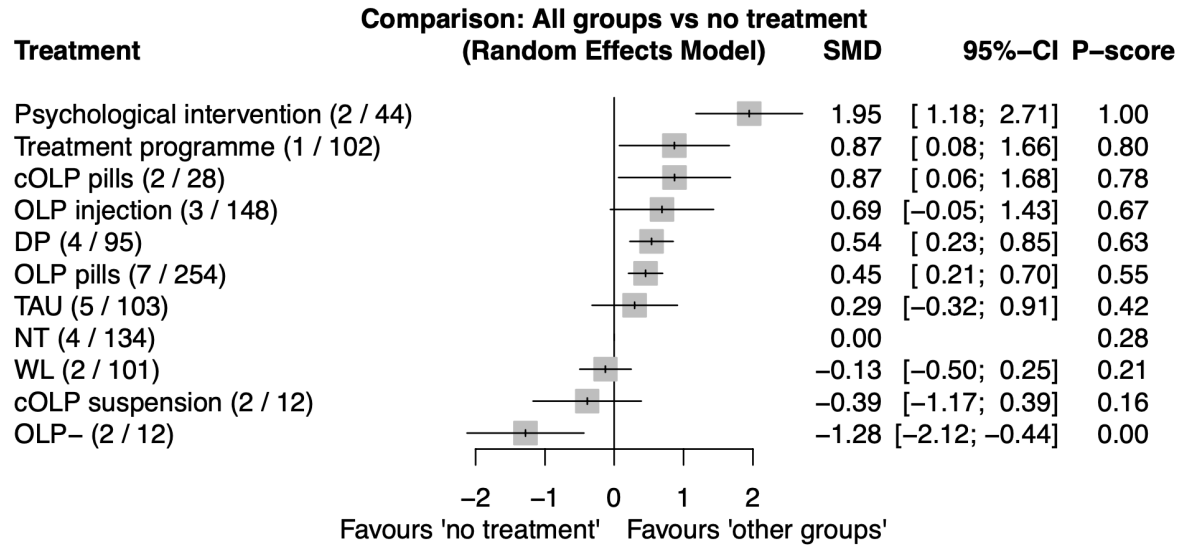


eFigure 5. Plots of pain trials (sensitivity analysis)

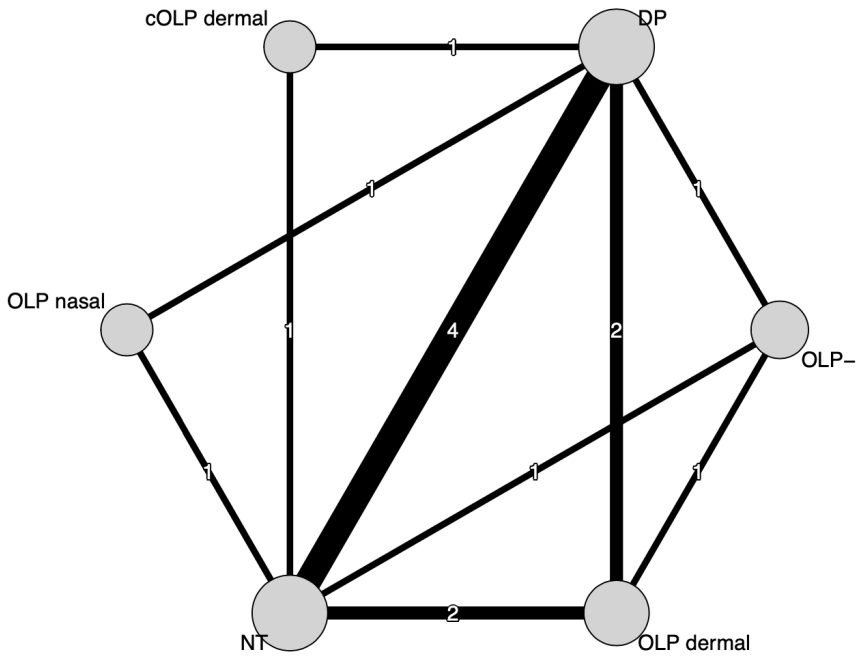
Netgraph of network meta-analysis on clinical pain studies only



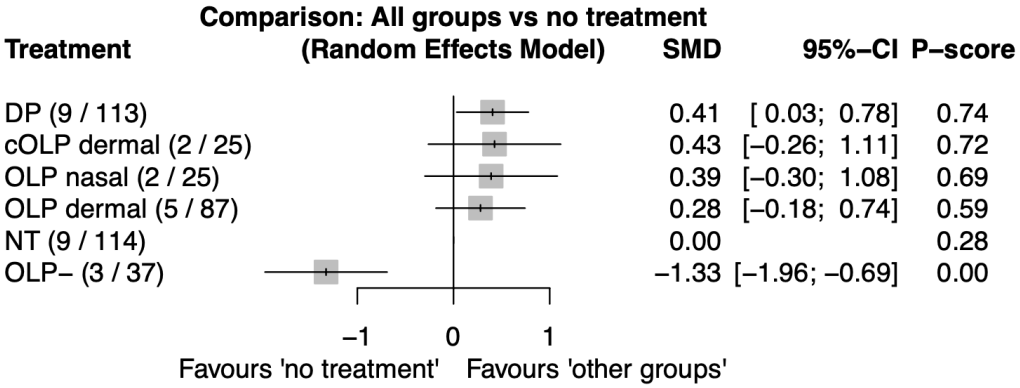
Forest plot of network meta-analysis on clinical pain studies only



Netgraph of network meta-analysis on nonclinical pain studies only

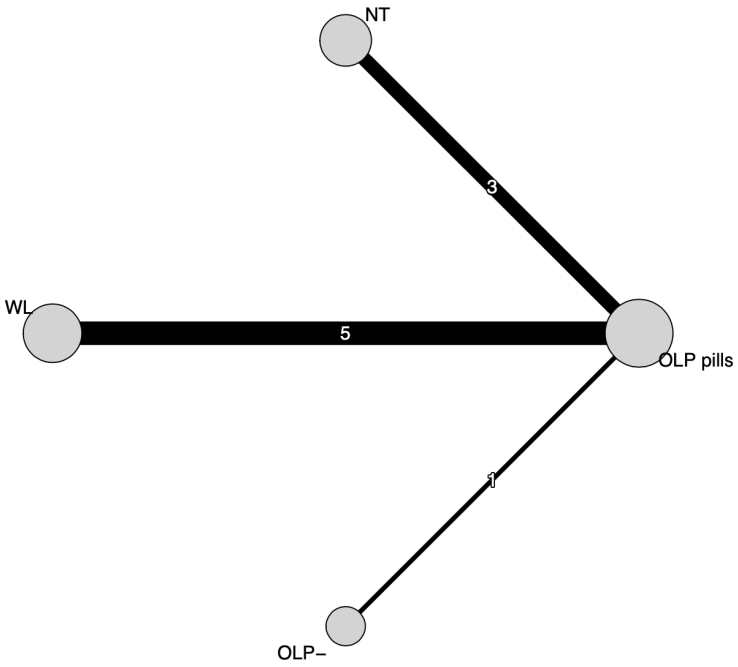


Forest plot of network meta-analysis on nonclinical pain studies only

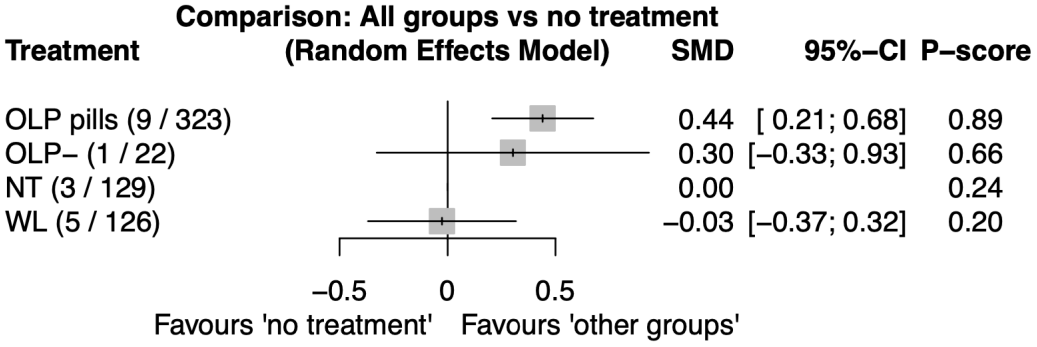


eFigure 6. Plots of psychological trials (sensitivity analysis)

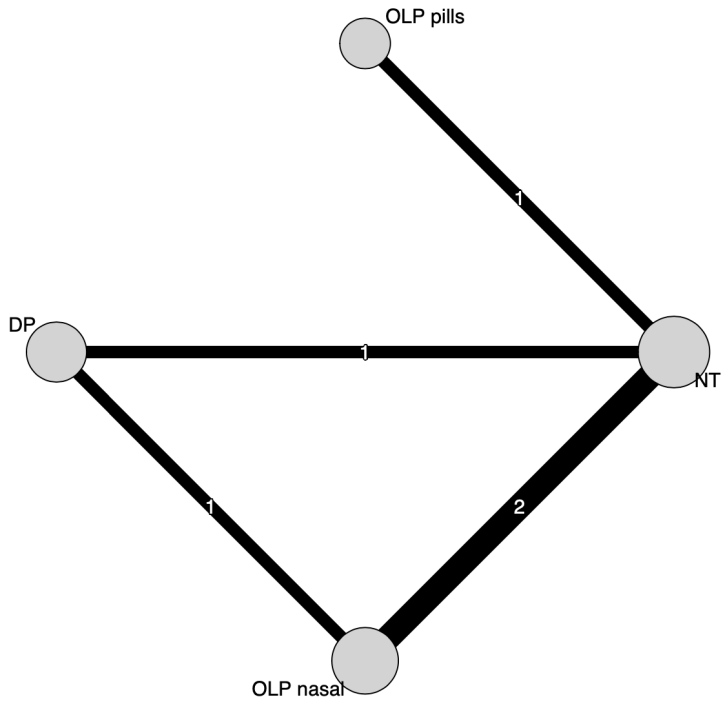
Netgraph of network meta-analysis on clinical psychological studies only



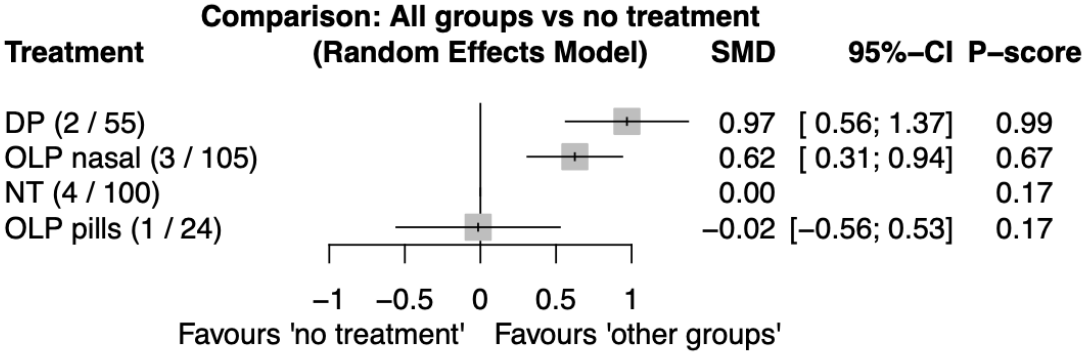
Forest plot of network meta-analysis on clinical psychological studies only



Netgraph of network meta-analysis on nonclinical psychological studies only



Forest plot of network meta-analysis on nonclinical psychological studies only



eTable 1. Demographics and study characteristic

Author, Year	Country	Condition/ Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Ashar, 2021	USA	chronic low back pain	clinical	135 (53.67)	41.10 (15.67)	28	OLP injection (injection)	psychological intervention		TAU	pain intensity (NRS 0-10)	only PO	low
Bandak, 2022	Denmark	knee osteoarthritis	clinical	206 (45.65)	68.40 (8.25)	56	OLP injection (injection)			Treatment programme (exercise and education)	pain subscale (KOOS 0-100)	only PO	some concerns
Barnes, 2019	Australia	experimental nausea	nonclinical	61 (52.74)	21.50 (4.65)	2	OLP nasal (vapor) (semi + fully open)	DP (vapor)		NT	nausea (VAS 0-10)	only PO	some concerns
Carvalho, 2016	Portugal	chronic low back pain	clinical	83 (71.05)	44.25 (13.45)	21	OLP pills (pill)			WL	pain intensity (NRS 0-10)	most frequent	some concerns

Author, Year	Country	Condition/ Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Disley, 2021	UK	experimental pain	nonclinical	75 (86.67)	21.05 (5.04)	1	OLP nasal (spray)	DP (spray)		NT	pain intensity (VAS 0-100)	most frequent	some concerns
El Brihi, 2019	Australia	well-being	subclinical	88 (80.00)	19.00 (3.90)	7	OLP pills (capsule) (different doses merged)			NT	emotional distress (DASS)	most frequent	some concerns
Flowers, 2021	USA	acute pain (following spine surgery)	clinical	41 (NA)	60.15 (13.05)	17	cOLP pills (pill)			TAU	worst daily pain (mini-BP; NRS 0-10)	most frequent	some concerns
Friehs, 2022	Germany	experimental sadness	nonclinical	147 (70.26)	23.56 (4.25)	7	OLP nasal (spray sesame oil) (personal + scientific)	DP (spray)		NT (personal + scientific)	sadness subscale (PANAS-X)	only PO	some concerns
Haas, 2022	Germany	primary insomnia	clinical	45 (84.39)	30.07 (NA)	2	OLP pills (pill)			OLP- (pill)	subjective total sleep time in minutes	only PO	some concerns

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Hahn, 2022	Germany	experimental sadness	nonclinical	84 (100.00)	24.74 (5.15)	1	OLP nasal (spray)			NT	sadness subscale (PANAS-X)	only PO	some concerns
Hoeneme yer, 2018	USA	cancer-related fatigue	clinical	73 (69.00)	57.20 (11.80)	21	OLP pills (pill)			WL	cancer related fatigue (FSI-14)	most frequent	some concerns
Ikemoto, 2020	Japan	chronic low back pain	clinical	48 (61.55)	66.75 (66.75)	84	OLP pills (pill)			TAU	pain intensity (NRS 0-10)	most frequent	some concerns
Kaptchuk, 2010	Israel	irritable bowel syndrome	clinical	80 (69.50)	46.50 (18.00)	21	OLP pills (pill)			NT	IBS symptom severity scale (IBS-SSS)	baseline available	some concerns
Kelley, 2012	USA	major depressive disorder	clinical	20 (70.00)	38.80 (12.60)	14	OLP pills (capsule)			WL	depression severity (HAM-D-17)	only PO	some concerns

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Kleine-Borgmann, 2021	Germany	chronic low back pain	clinical	122 (NA)	59.33 (14.56)	21	OLP pills (capsule)			WL	pain intensity (NRS 0-10)	only PO	some concerns
Kleine-Borgmann, 2019	Germany	well-being & cognitive enhancement	subclinical	154 (67.50)	24.03 (2.79)	21	OLP pills (pill)			NT	stress (PSQ-20)	baseline available; most informative	some concerns
Klinger, 2017	Germany	chronic low back pain + experimental pain	clinical	48 (75.00)	50.89 (15.07)	1	cOLP suspension (saline cotton swab)	OLP- (saline cotton swab)		DP (conditioned + unconditioned), (saline cotton swab)	pain intensity (NRS 0-10)	only PO	some concerns
Kube, 2020	Germany	experimental pain	nonclinical	100 (49.50)	24.56 (5.66)	1	OLP dermal (cream) (expectancy + hope)	DP (cream)		NT	pain intensity (VAS 0-100)	only PO	high

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Kube, 2021	Germany	allergic rhinitis	clinical	54 (68.68)	31.48 (12.67)	14	OLP pills (tablet) (augmented + limited)			WL (augmented + limited)	self-reported allergic symptoms (CSMS)	only PO	some concerns
Leibowitz, 2019	USA	experimental itch	nonclinical	NA (63.50)	24.55 (NA)	NA	OLP dermal (cream) (expectation + rationale)	OLP-(cream)		NT	physiological allergic reaction (size of the wheal)	only PO	some concerns
Lembo, 2021	USA	irritable bowel syndrome	clinical	211 (72.93)	42.00 (18.00)	42	OLP pills (pill)	DP (pill)		NT	IBS symptom severity scale (IBS-SSS)	only PO	some concerns
Locher, 2017	Switzerland	experimental pain	nonclinical	151 (68.00)	27.15 (9.51)	1	OLP dermal (cream)	OLP-(cream)	DP (cream)	NT	pain intensity (VAS 0-100)	most frequent	low

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Meeuwis, 2021	Netherlands	experimental itch	nonclinical	55 (85.45)	21.89 (2.50)	1	OLP dermal (patch)			DP (patch)	mean itch (NRS 0-10)	only PO	some concerns
Meeuwis, 2019	Netherlands	experimental itch	nonclinical	45 (82.60)	21.80 (2.70)	7	OLP dermal (tonic)			DP (tonic)	AUC itch (NRS 0-10)	only PO	some concerns
Morales-Quezada, 2020	USA	acute pain (spinal cord injury and polytrauma)	clinical	19 (30.00)	47.30 (16.78)	6	cOLP pills (capsule)			TAU	opioid consumption (MEDC)	only PO	high
Mundt, 2017	USA	experimental pain	nonclinical	75 (57.33)	22.75 (5.89)	1	cOLP dermal (cream)	DP (cream)		NT	pain intensity (VAS 0-100)	only PO	some concerns
Nitzan, 2020	Israel	major depressive disorder	clinical	38 (NA)	49.91 (17.27)	56	OLP pills (capsule)			WL	depression severity (QIDS)	only PO	some concerns

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Olliges, 2022	Germany	knee osteoarthritis	clinical	40 (60.15)	67.02 (9.47)	21	OLP pills (capsule)			NT	pain intensity (NRS 0-10)	most frequent	some concerns
Pan, 2020	Germany	menopausal hot flushes	subclinical	100 (100.00)	54.55 (NA)	28	OLP pills (pill)			NT	hot flushes composite score	most informative	some concerns
Schaefer, 2018	Germany	allergic rhinitis	clinical	46 (77.80)	24.67 (6.37)	14	OLP pills (pill)	OLP- (pill)		NT (with rationale + without rationale)	allergic symptoms composite score	only PO	some concerns
Schaefer, 2016	Germany	allergic rhinitis	clinical	25 (84.00)	26 (9.90)	14	OLP pills (pill)			NT	allergic symptoms composite score	only PO	some concerns
Schaefer, 2019	Germany	test anxiety	subclinical	58 (86.60)	22.90 (2.85)	14	OLP pills (pill)			NT	test anxiety (PAF)	most informative	some concerns
Schaefer, 2021	Germany	experimental acute stress	nonclinical	53 (53.31)	26.33 (8.77)	21	OLP pills (pill)			NT	acute stress (0-100)	most frequent	some concerns

Author, Year	Country	Condition/Diagnosis	Sample used in analysis	N (% female) per group	Mean age in years (SD) per group	Treatment duration in days	Intervention 1	Intervention 2	Intervention 3	Control	Outcome used for analysis	Rationale for choice of outcome	Risk of Bias
Schienle, 2021	Austria	relaxation	subclinical	148 (71.00)	24.40 (2.70)	14	OLP suspension (sunflower oil)			TAU	PMR exercise quality: relaxation	baseline available; most informative	some concerns
Swafford, 2019*	USA	muscle strength	nonclinical	21 (47.60)	22.52 (3.00)	7	OLP pills (capsule)	DP (capsule)		NT	isometric peak torque	authors judgment	some concerns
Yennurajalingam, 2022	USA	cancer-related fatigue	clinical	84 (67.00)	56.00 (13.00)	7	OLP pills (tablet)			WL	cancer related fatigue (FACIT-F)	only PO	some concerns
Zhou, 2019	USA	cancer-related fatigue	clinical	40 (92.50)	47.30 (12.40)	22	OLP pills (tablet)			WL	cancer related fatigue (FACIT-F)	only PO	some concerns

Note. cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; PO, Primary Outcome; TAU, Treatment as Usual; WL, Wait List; *, crossover study

eTable 2. Individual study data

Nonclinical network

author	year	merged groups	data from author	group	population	age mean	age sd	% female	country	continuous outcome	n	mean change	sd change
Barnes	2019	yes (fully & semi open)	yes	OLP nasal	experimental nausea	20.3	3.26	58.62	Australia	Self-report nausea, 6-item composite scale	29	6.14	9.78
Barnes	2019	no	yes	DP	experimental nausea	21.3	5.2	NA	Australia	Self-report nausea, 6-item composite scale	17	8.18	11.36
Barnes	2019	no	yes	NT	experimental nausea	22.9	5.5	NA	Australia	Self-report nausea, 6-item composite scale	15	2.86	11.07
Disley	2021	no	no	OLP nasal	experimental pain	21.05	5.04	86.666	UK	Pain Intensitiy, VAS	25	-0.12	20.35
Disley	2021	no	no	DP	experimental pain	21.05	5.04	86.666	UK	Pain Intensitiy, VAS	26	0.08	21.60
Disley	2021	no	no	NT	experimental pain	21.05	5.04	86.666	UK	Pain Intensitiy, VAS	24	-7.79	17.31
Friehs	2022	yes (personal & scientific)	no	OLP nasal	experimental sadness	24.56	6.55	69.79	Germany	Sadness subscale PANAS-X score total score 0-50	63	-2.20	9.60

Friehs	2022	yes (personal & scientific)	no	DP	experimental sadness	23.02	3.31	58.18	Germany	Sadness subscale PANAS-X score total score 0-50	55	1.00	6.06
Friehs	2022	no	no	NT	experimental sadness	23.1	2.9	82.8	Germany	Sadness subscale PANAS-X score total score 0-50	29	-6.00	8.22
Hahn	2022	no	no	OLP nasal	experimental sadness	23.67	3.31	100	Germany	Sadness subscale PANAS-X score total score 0-50	42	-4.27	8.76
Hahn	2022	no	no	NT	experimental sadness	25.81	6.98	100	Germany	Sadness subscale PANAS-X score total score 0-50	42	-12.01	10.87
Kube	2020	yes (Expectancy & Hope)	no	OLP dermal	experimental pain	25.16	6.41	62	Germany	Pain Intensitiy, VAS	50	-0.02	13.63
Kube	2020	no	no	DP	experimental pain	23.6	4.81	48	Germany	Pain Intensitiy, VAS	25	7.29	13.48
Kube	2020	no	no	NT	experimental pain	24.92	5.76	38.5	Germany	Pain Intensitiy, VAS	25	-2.70	13.92
Leibowitz	2019	no	yes	NT	experimental itch	24.55	NA	63.5	USA	Physiological allergic reaction (size of the wheal)	40	-1.65	1.09

Leibowitz	2019	no	yes	OLP-	experimental itch	24.55	NA	63.5	USA	Physiological allergic reaction (size of the wheal)	36	-1.61	0.83
Leibowitz	2019	yes (expectation & rationale)	yes	OLP dermal	experimental itch	24.55	NA	63.5	USA	Physiological allergic reaction (size of the wheal)	72	-1.56	0.88
Locher	2017	no	no	NT	experimental pain	27.9	8.52	73	Switzerland	Subjective heat pain intensity	40	1.89	3.33
Locher	2017	no	no	OLP-	experimental pain	28.27	11.34	65	Switzerland	Subjective heat pain intensity	37	-3.11	3.46
Locher	2017	no	no	OLP dermal	experimental pain	25.7	7.76	73	Switzerland	Subjective heat pain intensity	37	2.97	3.46
Locher	2017	no	no	DP	experimental pain	26.65	10.25	62	Switzerland	Subjective heat pain intensity	37	1.81	3.46
Meeuwis	2021	no	no	OLP dermal	experimental itch	21.67	2.6	85.19	Netherlands	Self reported mean itch, NRS	27	0.55	1.53
Meeuwis	2021	no	no	DP	experimental itch	22.11	2.39	85.71	Netherlands	Self reported mean itch, NRS	28	0.81	1.48
Meeuwis	2019	no	no	OLP dermal	experimental itch	21.8	2.7	82.6	Netherlands	AUC itch	22	49.71	223.04
Meeuwis	2019	no	no	DP	experimental itch	21.8	2.7	82.6	Netherlands	AUC itch	23	58.27	259.11

Mundt	2017	no	no	NT	experimental pain	22.75	5.89	57.33	USA	Mean pain intensity ratings, VAS	25	-6.14	11.33
Mundt	2017	no	no	DP	experimental pain	22.75	5.89	57.33	USA	Mean pain intensity ratings, VAS	25	1.25	12.46
Mundt	2017	no	no	cOLP dermal	experimental pain	22.75	5.89	57.33	USA	Mean pain intensity ratings, VAS	25	0.21	11.64
Schaefer	2021	no	no	OLP pills	experimental acute stress	25.25	7.28	58.33	Germany	Perceived stress, VAS	24	-31.29	35.07
Schaefer	2021	no	no	NT	experimental acute stress	27.41	10.25	48.28	Germany	Perceived stress, VAS	29	-30.76	34.60
Swafford	2019	no	yes	DP	muscle strength	22.52	3	47.6	USA	Peak torque of experiment 1	7	5.20	55.11
Swafford	2019	no	yes	OLP pills	muscle strength	22.52	3	47.6	USA	Peak torque of experiment 1	7	5.80	58.98
Swafford	2019	no	yes	NT	muscle strength	22.52	3	47.6	USA	Peak torque of experiment 1	7	4.90	32.06

Clinical network

author	year	merged groups	data from author	group	population	age mean	age sd	% female	country	continuous outcome	n	mean change	sd change
Ashar	2021	no	no	Psycho-logical intervention	chronic low back pain	42.6	16.2	58	USA	Pain intensity, VAS	44	3.04	1.23
Ashar	2021	no	no	OLP injection	chronic low back pain	39.4	14.9	49	USA	Pain intensity, VAS	44	1.32	1.51
Ashar	2021	no	no	TAU	chronic low back pain	41.3	15.9	54	USA	Pain intensity, VAS	47	0.78	1.36
Bandak	2022	no	no	OLP injection	knee osteoarthritis	66.7	8.2	47.2	Denmark	Pain score, KOOS (baseline-week 9)	104	7.30	15.22
Bandak	2022	no	no	Treatment program	knee osteoarthritis	70.1	8.3	44.1	Denmark	Pain score, KOOS (baseline-week 9)	102	10.00	15.07
Carvalho	2016	no	no	OLP pills	chronic low back pain	44.4	13.2	70.7	Portugal	Pain intensity, NRS	41	1.49	1.68
Carvalho	2016	no	no	WL	chronic low back pain	44.1	13.7	71.4	Portugal	Pain intensity, NRS	42	0.24	1.61
El Brihi	2019	yes (OLP 1/d & 4/d)	yes	OLP pills	well-being	19	3.9	80	Australia	Emotional distress (DASS)	61	7.30	9.16

El Brihi	2019	no	yes	NT	well-being	19	3.9	80	Australia	Emotional distress (DASS)	27	0.20	10.26
Flowers	2021	no	no	cOLP pills	acute pain (after spine surgery)	59.1	13.1	NA	USA	Worst daily pain (mini-BP; 0-10)	19	-0.60	2.26
Flowers	2021	no	no	TAU	acute pain (after spine surgery)	61.2	13	NA	USA	Worst daily pain (mini-BP; 0-10)	22	-1.50	1.44
Haas	2022	no	no	OLP pills	primary insomnia	31.04	NA	86.96	Germany	Subjective total sleep time in minutes	23	24.83	91.13
Haas	2022	no	no	OLP-	primary insomnia	29.09	NA	81.82	Germany	Subjective total sleep time in minutes	22	11.31	104.21
Hoenemeyer	2018	no	no	OLP pills	cancer-related fatigue	58.4	11.2	72	USA	FSI, Fatique Symptom Severity)	38	18.60	23.01
Hoenemeyer	2018	no	no	WL	cancer-related fatigue	56	12.4	66	USA	FSI, Fatique Symptom Severity)	35	6.10	22.75
Ikemoto	2020	no	no	OLP pills	chronic low back pain	68.2	68.2	65.4	Japan	Pain intensity, NRS	24	1.10	1.90
Ikemoto	2020	no	no	TAU	chronic low back pain	65.3	65.3	57.7	Japan	Pain intensity, NRS	24	0.80	1.90
Kaptchuk	2010	no	no	OLP pills	irritable bowel syndrome	47	18	65	Israel	IBS-SSS 0-500	37	92.00	99.00

Kaptchuk	2010	no	no	NT	irritable bowel syndrome	46	18	74	Israel	IBS-SSS 0-500	43	46.00	74.00
Kelley	2012	no	no	OLP pills	MDD	38.8	12.6	70	USA	Depression severity, HAM-D	11	1.64	4.52
Kelley	2012	no	no	WL	MDD	38.8	12.6	70	USA	Depression severity, HAM-D	9	-0.67	4.00
Kleine-Borgmann	2021	no	no	OLP pills	chronic low back pain	60.28	15.15	NA	Germany	Composite pain intensity score	63	0.62	1.81
Kleine-Borgmann	2021	no	no	WL	chronic low back pain	58.37	13.97	NA	Germany	Composite pain intensity score	59	-0.11	1.29
Kleine-Borgmann	2019	no	no	OLP pills	well-being & cognitive enhancement	23.97	2.83	68	Germany	Perceived Stress Questionnaire, PSQ20	79	-11.90	19.67
Kleine-Borgmann	2019	no	no	NT	well-being & cognitive enhancement	24.08	2.74	67	Germany	Perceived Stress Questionnaire, PSQ20	75	-16.74	17.22
Klinger	2017	no	no	OLP-	chronic low back pain + experimental pain	50.83	17.01	75	Germany	Back pain rating, NRS	12	-1.16	1.83
Klinger	2017	no	no	cOLP suspension	chronic low back pain + experimental pain	50.33	15.17	75	Germany	Back pain rating, NRS	12	0.67	2.12

Klinger	2017	yes (cond. & uncond. DP)	no	DP	chronic low back pain + experimental pain	51.52	13.05	75	Germany	Back pain rating, NRS	24	2.58	2.12
Kube	2021	yes (augmented & limited)	no	OLP pills	allergic rhinitis	26.95	10.56	64.3	Germany	Self-reported allergic symptoms, CSMS	28	2.20	3.81
Kube	2021	yes (augmented & limited)	no	WL	allergic rhinitis	36	14.77	73.05	Germany	Self-reported allergic symptoms, CSMS	26	2.90	3.76
Lembo	2021	no	no	OLP pills	irritable bowel syndrome	42.2	17.8	71.9	USA	IBS-SSS 0-500	68	90.60	89.50
Lembo	2021	no	no	NT	irritable bowel syndrome	40	17	73.3	USA	IBS-SSS 0-500	72	52.30	87.00
Lembo	2021	no	no	DP	irritable bowel syndrome	43.8	19.2	73.6	USA	IBS-SSS 0-500	71	100.30	99.60
Morales-Quezada	2020	no	yes	cOLP pills	acute pain (spinal cord injury and polytrauma)	44.9	16.93	30	USA	Opiod consumption, MDEC	9	66.00	99.55
Morales-Quezada	2020	no	yes	TAU	acute pain (spinal cord injury and polytrauma)	49.7	16.62	30	USA	Opiod consumption, MDEC	10	3.76	56.51

Nitzan	2020	no	no	OLP pills	major depressive disorder	48.17	16.86	NA	Israel	Depression severity,QIDS total score	18	1.95	5.06
Nitzan	2020	no	no	WL	major depressive disorder	51.65	17.68	NA	Israel	Depression severity, QIDS total score	20	0.45	4.12
Olliges	2022	no	yes	OLP pills	knee osteoarthritis	64.19	9.3	57.1	Germany	Pain intensity, NRS	21	0.44	1.35
Olliges	2022	no	yes	NT	knee osteoarthritis	69.84	9.63	63.2	Germany	Pain intensity, NRS	19	-0.28	1.99
Pan	2020	no	no	OLP pills	menopausal hot flushes	54.2	NA	100	Germany	Hot flush score, composite score	50	6.02	9.71
Pan	2020	no	no	NT	menopausal hot flushes	54.9	NA	100	Germany	Hot flush score, composite score	50	3.26	8.79
Schaefer	2018	no	no	OLP pills	allergic rhinitis	25	9	69.2	Germany	Allergic symptoms composite score	13	0.78	0.67
Schaefer	2018	no	no	OLP-	allergic rhinitis	23	3	69.2	Germany	Allergic symptoms composite score	13	0.43	0.90
Schaefer	2018	yes (with & without rationale)	no	NT	allergic rhinitis	26	7.11	95	Germany	Allergic symptoms composite score	20	0.05	1.03

Schaefer	2016	no	no	OLP pills	allergic rhinitis	26	9.9	84	Germany	Allergic symptoms composite score	11	0.88	0.93
Schaefer	2016	no	no	NT	allergic rhinitis	26	9.9	84	Germany	Allergic symptoms composite score	14	0.23	0.72
Schaefer	2019	no	no	OLP pills	test anxiety	22.3	2.3	80.6	Germany	Test anxiety, PAF	31	4.39	9.35
Schaefer	2019	no	no	NT	test anxiety	23.5	3.4	92.6	Germany	Test anxiety, PAF	27	0.07	6.00
Schienle	2021	no	no	OLP suspension	relaxation	24.4	2.7	71	Austria	Exercise quality relaxation	68	1.29	0.97
Schienle	2021	no	no	TAU	relaxation	24.4	2.7	71	Austria	Exercise quality relaxation	80	1.28	0.93
Yennurajalingam	2022	no	no	OLP pills	cancer-related fatigue	57	12	74	USA	Fatigue, FACIT-F	42	6.60	7.60
Yennurajalingam	2022	no	no	WL	cancer-related fatigue	55	14	60	USA	Fatigue, FACIT-F	42	2.10	9.40
Zhou	2019	no	no	OLP pills	cancer-related fatigue	47.3	12.4	92.5	USA	Fatigue, FACIT-F	20	4.30	10.43
Zhou	2019	no	no	WL	cancer-related fatigue	47.3	12.4	92.5	USA	Fatigue, FACIT-F	20	1.20	10.15

Note. cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; TAU, Treatment as Usual; WL, Wait List.

eTable 3. Head to head comparisons

Nonclinical network

	cOLP dermal	DP	NT	OLP dermal	OLP nasal	OLP pills	OLP-
cOLP dermal		-0.09 [-1.00; 0.82]	0.54 [-0.37; 1.45]	.	.	.	.
DP	-0.03 [-0.83; 0.78]		**0.47 [0.12; 0.82]**	0.10 [-0.34; 0.54]	0.21 [-0.29; 0.72]	-0.01 [-1.28; 1.26]	**1.44 [0.57; 2.30]**
NT	0.47 [-0.33; 1.28]	**0.50 [0.17; 0.83]**		-0.20 [-0.69; 0.29]	**-0.50 [-0.94; -0.06]**	0.00 [-0.73; 0.74]	**0.70 [0.09; 1.30]**
OLP dermal	0.21 [-0.66; 1.08]	0.24 [-0.15; 0.62]	-0.26 [-0.67; 0.14]		.	.	**0.86 [0.26; 1.46]**
OLP nasal	0.04 [-0.83; 0.92]	0.07 [-0.36; 0.50]	**-0.43 [-0.84; -0.02]**	-0.17 [-0.70; 0.37]		.	.
OLP pills	0.38 [-0.68; 1.43]	0.40 [-0.34; 1.15]	-0.10 [-0.80; 0.60]	0.17 [-0.63; 0.96]	0.33 [-0.47; 1.13]		.
OLP-	**1.07 [0.12; 2.02]**	**1.10 [0.53; 1.66]**	**0.60 [0.05; 1.15]**	**0.86 [0.31; 1.41]**	**1.03 [0.37; 1.69]**	0.69 [-0.19; 1.57]	

Note. Column headers are identical to row headers. Cells contain the network estimates (SMDs) from network meta-analysis (direct and indirect evidence) in the lower triangle and the direct treatment estimates (SMDs) from pairwise comparisons in the upper triangle. Comparisons considered for RQ2 (modalities) are highlighted in yellow, for RQ3 (expectation) in green and for RQ4 (comparator) in blue. Legend: cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; TAU, Treatment as Usual; WL, Wait List.

Clinical network

	cOLP pills	cOLP suspension	DP	NT	OLP injection	OLP pills	OLP suspension	OLP-	Psych. intervent.	TAU	Treatment programme	WL
cOLP pills		.	.	.	.	.	.	.	.	**0.58 [0.02; 1.15]**	.	.
cOLP suspension	0.65 [-0.50; 1.81]		**-0.93 [-1.71; -0.15]**	.	.	.	.	**0.89 [0.02; 1.76]**	.	.	.	.
DP	0.12 [-0.81; 1.06]	-0.53 [-1.26; 0.20]		**0.52 [0.07; 0.97]**	.	0.11 [-0.34; 0.55]	.	**1.82 [0.98; 2.66]**	.	.	.	.
NT	**0.89 [0.01; 1.76]**	0.23 [-0.54; 1.01]	**0.76 [0.39; 1.14]**		.	**-0.47 [-0.66; -0.29]**	.	-0.42 [-1.19; 0.34]	.	.	.	.
OLP injection	0.19 [-0.57; 0.95]	-0.47 [-1.59; 0.66]	0.06 [-0.84; 0.96]	-0.70 [-1.54; 0.14]		.	.	.	**-1.26 [-1.79; -0.72]**	0.39 [-0.12; 0.91]	-0.18 [-0.58; 0.23]	.
OLP pills	0.42 [-0.43; 1.28]	-0.23 [-1.00; 0.54]	0.30 [-0.07; 0.67]	**-0.46 [-0.65; -0.28**	0.24 [-0.58; 1.06]		.	0.23 [-0.28; 0.75]	.	0.16 [-0.48; 0.80]	.	**0.43 [0.22; 0.64]**
OLP suspension	0.57 [-0.15; 1.29]	-0.08 [-1.18; 1.01]	0.45 [-0.41; 1.31]	-0.32 [-1.12; 0.49]	0.38 [-0.29; 1.06]	0.15 [-0.63; 0.93]		.	.	0.01 [-0.43; 0.45]	.	.
OLP-	0.92 [-0.04; 1.87]	0.26 [-0.51; 1.03]	**0.79 [0.30; 1.29]**	0.03 [-0.41; 0.47]	0.73 [-0.20; 1.66]	**0.49 [0.07; 0.92]**	0.35 [-0.54; 1.23]		.	.	.	.
Psych. intervent.	**-1.07 [-1.85; -0.28]**	**-1.72 [-2.87; -0.58]**	**-1.19 [-2.11; -0.27]**	**-1.96 [-2.82; -1.09]**	**-1.26 [-1.79; -0.72]**	**-1.49 [-2.34; -0.65]**	**-1.64 [-2.34; -0.94]**	**-1.99 [-2.93; -1.04]**		**1.65 [1.11; 2.20]**	.	.
TAU	**0.58 [0.02; 1.15]**	-0.07 [-1.08; 0.93]	0.46 [-0.28; 1.20]	-0.30 [-0.97; 0.36]	0.39 [-0.12; 0.91]	0.16 [-0.48; 0.80]	0.01 [-0.43; 0.45]	-0.34 [-1.11; 0.44]	**1.65 [1.11; 2.20]**		.	.
Treatment programme	0.01 [-0.85; 0.87]	-0.65 [-1.84; 0.55]	-0.11 [-1.10; 0.87]	-0.88 [-1.81; 0.06]	-0.18 [-0.58; 0.23]	-0.41 [-1.33; 0.50]	-0.56 [-1.35; 0.23]	-0.91 [-1.92; 0.10]	**1.08 [0.41; 1.75]**	-0.57 [-1.23; 0.08]		.
WL	0.86 [-0.02; 1.74]	0.20 [-0.60; 1.00]	**0.73 [0.31; 1.16]**	-0.03 [-0.31; 0.25]	0.67 [-0.18; 1.51]	**0.43 [0.22; 0.64]**	0.28 [-0.52; 1.09]	-0.06 [-0.54; 0.41]	**1.92 [1.06; 2.79]**	0.27 [-0.40; 0.95]	0.85 [-0.09; 1.78]	

Note. Column headers are identical to row headers. Cells contain the network estimates (SMDs) from network meta-analysis (direct and indirect evidence) in the lower triangle and the direct treatment estimates (SMDs) from pairwise comparisons in the upper triangle. Comparisons considered for RQ2 (modalities) are highlighted in yellow, for RQ3 (expectation) in green and for RQ4 (comparator) in blue. Legend: cOLP, conditioned Open-Label Placebo; DP, Deceptive Placebo; NT, No Treatment; OLP, Open-Label Placebo with rationale; OLP-, Open-Label Placebo without expectation induction; TAU, Treatment as Usual; WL, Wait List.

eReferences

1. Salanti, G., Del Giovane, C., Chaimani, A., Caldwell, D. M. & Higgins, J. P. Evaluating the quality of evidence from a network meta-analysis. *PloS One* **9**, e99682 (2014).
2. Papakonstantinou, T., Nikolakopoulou, A., Higgins, J. P., Egger, M. & Salanti, G. Cinema: software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell Syst. Rev.* **16**, e1080 (2020).
3. Nikolakopoulou, A. *et al.* CINeMA: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Med.* **17**, e1003082 (2020).
4. Sterne, J. A. C. *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj* **366**, l4898 (2019).
5. Chiocchia, V. *et al.* ROB-MEN: a tool to assess risk of bias due to missing evidence in network meta-analysis. *BMC Med.* **19**, 1–13 (2021).
6. Cuijpers, P., Noma, H., Karyotaki, E., Cipriani, A. & Furukawa, T. A. Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: a network meta-analysis. *JAMA Psychiatry* **76**, 700124–707 (2019).

B. Study 2

Buergler, S., Sezer, D., Bagge, N., Kirsch, I. Locher, C., Carvalho, C., & Gaab, J. (2023). Imaginary pills and open-label placebos can reduce test anxiety by means of placebo mechanisms. *Scientific Reports, 13*(1), 2624. https://doi.org/10.1038/s41598-023-29624-7

scientific reports

Check for updates

OPEN

Imaginary pills and open-label placebos can reduce test anxiety by means of placebo mechanisms

Sarah Buergler[1]✉, Dilan Sezer[1], Niels Bagge[2], Irving Kirsch[3], Cosima Locher[4,5], Claudia Carvalho[6] & Jens Gaab[1]

Placebos have been shown to be beneficial for various conditions even if administered with full transparency. Hence, so-called open-label placebos (OLPs) offer a new way to harness placebo effects ethically. To take this concept one step further, this study aimed at evaluating placebo effects without the use of a physical placebo, i.e., by imagining taking a pill. Healthy students (*N* = 173) with self-reported test anxiety were either randomized to an imaginary pill (IP; n = 55), an OLP (n = 59) or a control group (CG; n = 59). Both intervention groups were instructed to take two pills daily for three weeks. Primary outcome was test anxiety, secondary outcomes were sleep quality, general well-being and test performance. Groups test anxiety differed at study-endpoint, $F(2,169) = 11.50$, $p < .001$. Test anxiety was lower in the intervention groups compared to the CG, $t(169) = -4.44$, $p < .001$, $d = -0.71$. The interventions did not differ significantly, i.e., both were similarly efficacious, $t(169) = 0.61$, $p = .540$, $d = 0.11$. The interaction between group and time in explaining test anxiety was significant, $F(5,407.93) = 6.13$, $p < .001$. OLPs and IPs reduced test anxiety in healthy participants compared to the CG. This finding opens the door for a novel and ethical method to harness placebo effects.

Placebo effects are clinically highly relevant and the need to harness these effects has been voiced[1]. In this regard, open-label placebos (OLPs) administered with full disclosure and transparency can be deemed both ethical and feasible as they avoid the use of deception[2]. Interestingly, meta-analyses show medium sized to large clinically relevant effects of OLPs in patients with various clinical conditions compared to control groups[3,4]. Thus, if placebos also work without deception, it implies that it is not necessarily the pill serving as a symbol for a real medication that triggers these effects. The investigation of underlying mechanisms by eliminating the physical treatment constituent (i.e., the pill itself) can reveal the power of the purely psychological component of a placebo. For this reason, we aimed to evaluate placebo effects without the use of a placebo by having participants imagine taking a pill rather than actually taking one.

The concept of an imaginary pill (IP) was first introduced by De Shazer in 1984 in the context of clinical hypnosis[5]. More recently, Niels Bagge, a Danish clinician, independently introduced the same idea without hypnosis[6]. Although seemingly farfetched, recent data supports its plausibility: For instance, pharmacological placebos can be effective even when only possessed, but not applied[7]. Also, psychotherapeutic, non-pharmacological placebos have been shown to be effective[8] and the idea of triggering placebo effects without a placebo pill is discussed in sports performance[9], healthcare[10] and in research on the moderating role of mind-sets[11]. Additionally, a study by Peerdeman et al.[12] indicated that mental imagery of reduced pain can induce placebo-like expectancy effects on pain. Thus, placebos can also be purely psychological in nature and still produce beneficial effects. With regard to the underlying mechanisms of such psychological placebos, it yet needs to be investigated, whether their efficacy is purely mediated by the meaning that is attributed to these rituals or the expectations of improvement that are being formed as a consequence[13,14]. Despite the elimination of the physical stimulus, it is plausible that an IP relies in principle on the same underlying mechanisms as an OLP. Besides expectation, conditioning could for instance play a role, as even imagining something can activate corresponding brain areas and associated learning mechanisms (e.g.[15]). In addition, placebo mechanisms have also been discussed in relation to the theory of embodied cognition, which states that our experiences are not only consciously stored

[1]Division of Clinical Psychology and Psychotherapy, Faculty of Psychology, University of Basel, Basel, Switzerland. [2]Institute for Emotion-Focused Therapy, Roskilde, Denmark. [3]Program in Placebo Studies, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA. [4]Department of Consultation-Liaison Psychiatry and Psychosomatic Medicine, University Hospital Zurich, University of Zurich, Zurich, Switzerland. [5]Faculty of Health, University of Plymouth, Plymouth, UK. [6]Department of Clinical and Health Psychology, Instituto Superior de Psicologia Aplicada (ISPA), Lisbon, Portugal. ✉email: sarah.buergler@unibas.ch

as memories, but also directly imprinted in our bodies without any cognitive process being involved[16]. Thus, a placebo effect could result from the unconscious internal act of imagining a specific change in body state[16], which underlines the potential of IPs to harness placebo effects. In conclusion, this study provides a next step towards the quantification of the combined effect of various plausible psychological mechanisms within placebo research, omitting the physical treatment component.

In light of the high prevalence of test anxiety, affecting for example 53% of German freshman medical students[17], and its negative impact on educational performance[18,19], this condition is suitable to test the effects of an IP and OLP intervention. Evidence suggests that OLPs can effectively reduce test anxiety in healthy college students[20] and can have a positive impact on subjective well-being, whereas no improvement of exam performance by the intervention was found[21]. Furthermore, placebo effects in psychopharmacological treatments of anxiety disorders in general[22,23], social anxiety[24,25], generalized anxiety disorders[26] and panic disorders[27,28] are moderate to large.

In the present study, we set out to test the efficacy of an IP and OLP intervention in reducing test anxiety in a randomized controlled trial with healthy participants. To pursue this research question, we applied a previously used OLP intervention (e.g. in[29,30]) and further developed an IP intervention that was based on knowledge derived from placebo and imagination research to compare them to a control group (CG). We hypothesized that students receiving the OLP and IP intervention would show greater decreases in test anxiety from baseline to study endpoint (shortly before the exam) compared to students in the CG. We further expected students in the intervention groups to show higher general well-being, higher sleep quality and higher test performance than students in the CG.

Results

Sample characteristics and study flow. As shown in Fig. 1, of the 283 interested participants, 33 did not provide an e-mail contact and six did not give informed consent. The remaining 244 participants completed the online screening, of which 15 did not fulfill at least one inclusion criteria and 18 were excluded due to other reasons. Hence, 211 participants were randomized, of whom 178 received the intervention and completed the baseline assessment (T1; see Fig. 1 for reasons of exclusion). Five participants were excluded from the analyses as there was missing data (mostly due to nonattendance at exams because of COVID-19). Hence, an *N* of 173 was used for the final analyses (IP = 55, OLP = 59, CG = 59).

Table 1 depicts participants demographic and baseline characteristics. Participants' age ranged from 18 to 47 years, with a mean of 22.70 (± 4.18) years. The majority were female (85.55%) and undergraduate psychology students (87.86%). The three groups did not significantly differ in any of the demographic characteristics or primary and secondary outcomes at baseline. All outcomes were within the normal range of scores in the anxiety, well-being, and sleep questionnaires, indicating that our sample was healthy displaying an average test anxiety score.

Primary and secondary outcomes at study endpoint (T4). Figure 2 shows mean improvement from baseline (T1) to endpoint (T4) per group on the primary outcome. Table 2 depicts all primary and secondary outcomes for all groups and assessments.

The overall analyses of covariance (ANCOVA) showed that groups significantly differed in test anxiety at study endpoint T4, $F(2, 169) = 11.50$, $p < .001$. Planned contrasts indicated that the mean changes in test anxiety were significantly greater in the intervention groups (OLP/IP) compared to the CG at study endpoint T4, $t(169) = -4.44$, $p < .001$, $d = -0.71$. However, changes in test anxiety did not differ between the two intervention groups, $t(169) = 0.61$, $p = .540$, $d = 0.11$. These results held true for all subscales of the test anxiety questionnaire (see supplementary Table S1).

Regarding secondary outcomes, the groups differed significantly in terms of general well-being at study endpoint T4, $F(2, 169) = 9.37$, $p < .001$ with the same pattern across all subscales. Changes in general well-being were significantly greater in the intervention groups (OLP/IP) compared to the CG at study endpoint T4, $t(169) = -3.98$, $p < .001$, $d = -0.64$, but did not differ between the two intervention groups, $t(169) = 0.38$, $p = .707$, $d = 0.07$.

No significant between-group effect was found for total sleep quality, $F(2, 169) = 0.902$, $p = .408$, or in any of its component subscales. Nevertheless, although the overall between-group effect on the subjective sleep quality at study endpoint failed to reach statistical significance, $F(2, 169) = 2.73$, $p = .068$, contrasts indicated that both intervention groups showed better subjective sleep quality compared to the CG, $t(169) = -2.40$, $p = .017$, $d = -0.39$, with no significant difference between the intervention groups, $t(169) = 0.06$, $p = .952$, $d = 0.01$.

With respect to the test performance, 120 participants had a continuous test score (IP = 41, OLP = 35, CG = 44). Mean grade was 4.82 (± 0.83) ranging from 2.5 to 6.0 (IP = 4.94 ± 0.83, OLP = 4.92 ± 0.76, CG = 4.62 ± 0.87). Figure 3 depicts the participants grades per group. The overall ANOVA showed no significant group effect on test score, $F(2, 117) = 1.98$, $p = .143$. The contrasts, however, indicated that the intervention groups (OLP/IP) had higher test scores compared to the CG, $t(117) = 1.98$, $p = .050$, $d = 0.38$, whereas the intervention groups did not differ, $t(117) = -0.12$, $p = .908$, $d = -0.03$. Binary test scores (pass/fail) revealed that 155 (89.60%) of all participants passed the exam (IP = 87.27%, OLP = 96.61%, CG = 84.75%).

Primary and secondary outcomes over time (T1–T4). Figure 4 shows the course of test anxiety outcomes over time. There was a statistically significant interaction between group and time (T1–T4) for test anxiety, $F(5, 407.93) = 6.13$, $p < .001$. Bonferroni adjusted post-hoc p-values showed that the simple main effect of group was significant after one week (T2; $p_{adj} < .001$), two weeks (T3; $p_{adj} < .001$) and three weeks (T4; $p_{adj} < .001$) after

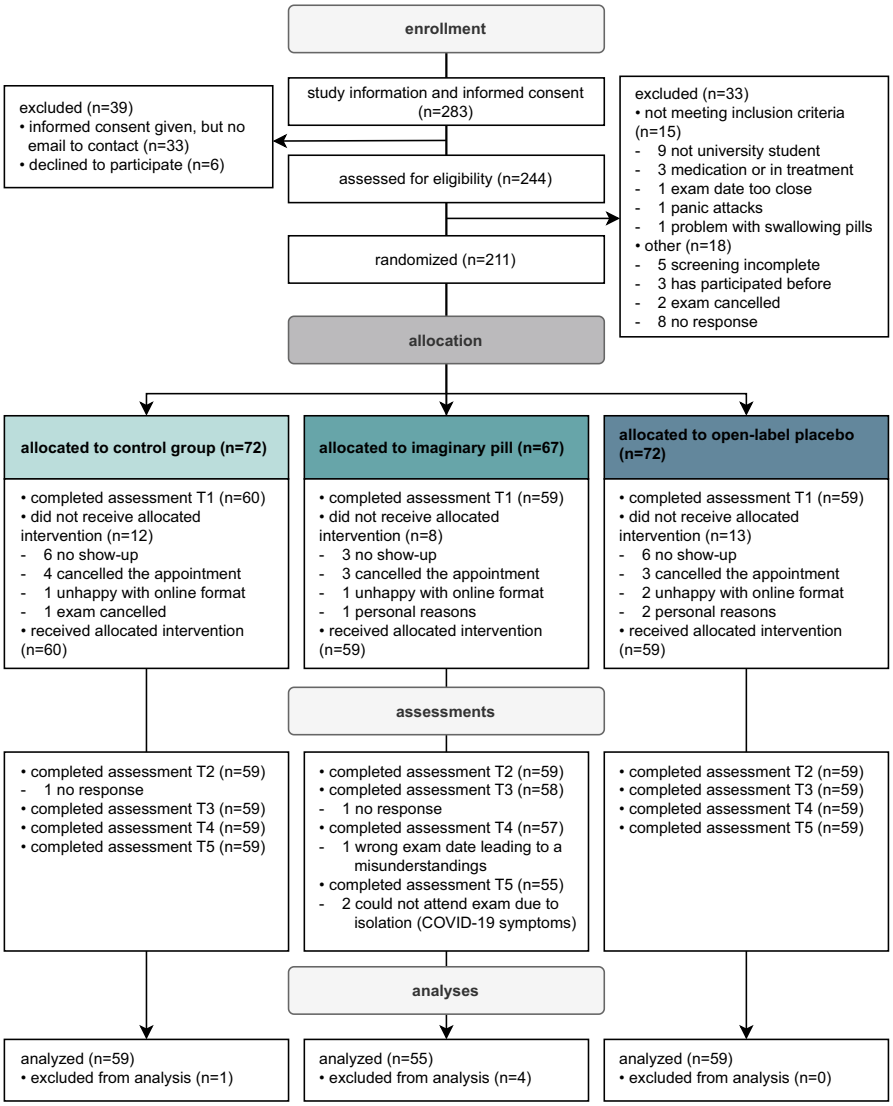


Figure 1. CONSORT diagram. Flow of the study, including reasons for exclusions.

	IP	OLP	CG
N (% female)	55 (82%)	59 (90%)	59 (85%)
Age in years, M (SD)	23.20 (4.30)	22.00 (3.48)	22.95 (4.67)
Psychology students, N (%)	49 (89%)	52 (88%)	51 (86%)
Test anxiety (PAF), M (SD)	45.53 (6.81)	48.34 (6.94)	47.36 (6.37)
Sleep quality (PSQI), M (SD)	5.69 (2.83)	6.02 (2.92)	5.85 (3.05)
General well-being (ASS-SYM), M (SD)	45.40 (20.90)	49.03 (22.52)	48.86 (20.76)

Table 1. Demographics and baseline scores of primary and secondary outcomes per group. *ASS-SYM* Änderungssensitive Symptomliste (general well-being), *CG* control group, *IP* imaginary pill, *M* mean, *OLP* open-label placebo, *PAF* Prüfungsangstfragebogen (test anxiety questionnaire), *PSQI* pittsburgh sleep quality index, *SD* standard deviation.

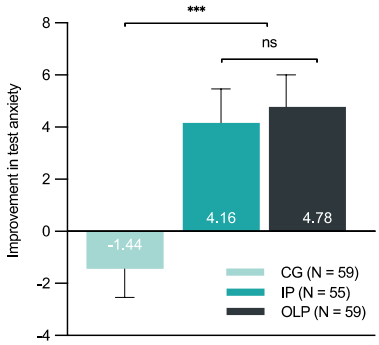


Figure 2. Mean improvement in test anxiety (PAF: test anxiety questionnaire) from baseline (T1) to endpoint (T4) per group. Results indicate a significant improvement for the OLP and IP group compared to the CG. *Note.* *CG* control group, *IP* imaginary pill, *ns* = not significant, *OLP* open-label placebo, ***$p < .001$. Error bars represent standard error of the mean.

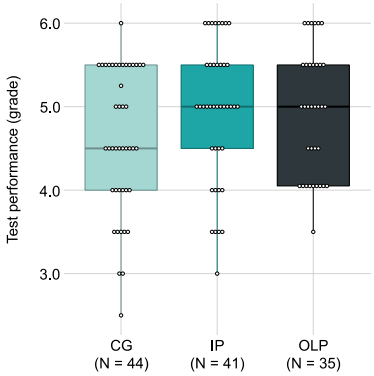


Figure 3. Boxplot showing continuous grades of the participants per group. Every dot represents a participants' grade with higher grades being better (ranging from 1.0 to 6.0). *Note.* Median is represented by the bold line within the box and upper/lower quartiles mark the end of the box. *CG* control group, *IP* imaginary pill, *OLP* open-label placebo.

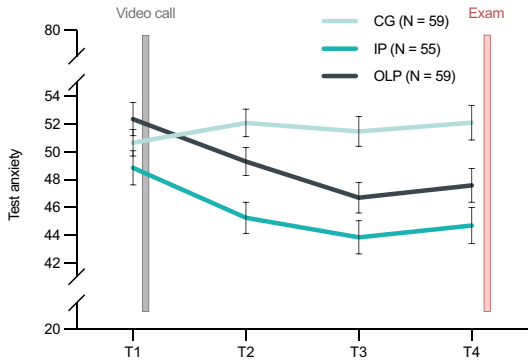


Figure 4. Course of test anxiety over time. Mean test anxiety per group from baseline (T1) through midpoints (T2, T3) to study endpoint (T4). *Note.* Error bars represent standard error of the mean. *CG* control group, *IP* imaginary pill, *OLP* open-label placebo.

		T1	T2	T3	T4	T5
	Group (*N*)	M (SD)				
PAF	IP (55)	48.85 (9.20)	45.25 (8.26)	43.84 (8.86)	44.69 (9.72)	39.85 (10.18)
	OLP (59)	52.36 (9.18)	49.31 (7.81)	46.70 (8.49)	47.58 (9.39)	41.64 (10.60)
	CG (59)	50.66 (7.37)	52.08 (7.70)	51.47 (8.23)	52.10 (9.56)	45.78 (9.97)
ASS-SYM	IP (55)	45.40 (20.90)	42.55 (22.57)	37.29 (21.9)3	39.25 (23.04)	
	OLP (59)	49.03 (22.52)	46.56 (19.48)	44.58 (19.96)	42.85 (21.99)	
	CG (59)	48.86 (20.76)	52.85 (24.36)	51.81 (22.02)	55.80 (28.58)	
PSQI	IP (55)	5.69 (2.83)	5.69 (2.48)	5.49 (2.46)	5.54 (2.71)	
	OLP (59)	6.02 (2.92)	5.49 (2.52)	5.88 (3.08)	5.86 (2.82)	
	CG (59)	5.85 (3.05)	6.36 (2.94)	6.36 (2.90)	6.22 (3.00)	

Table 2. Mean values for primary and secondary outcomes per group at all assessed timepoints. *ASS-SYM* Änderungssensitive Symptomliste (general well-being), *CG* control group, *IP* imaginary pill, *M* mean, *OLP* open-label placebo, *PAF* Prüfungsangstfragebogen (test anxiety questionnaire), *PSQI* pittsburgh sleep quality index, *SD* standard deviation.

randomization, but not at baseline (T1; $p_{adj} = .098$). Furthermore, there was also a statistically significant effect of time on test anxiety scores for the IP ($p_{adj} < .001$) and OLP ($p_{adj} < .001$) group, but not for the CG ($p_{adj} = .318$).

Regarding the secondary outcomes, a statistically significant interaction was found between group and time (T1–T4) in general well-being, $F(5, 422.71) = 3.58$, $p = .004$. Considering the Bonferroni adjusted p-values, the simple main effect of group was significant at T3 ($p_{adj} = .004$) and T4 ($p_{adj} = .004$), but not at T1 ($p_{adj} = .598$) and T2 ($p_{adj} = .061$). Also, the effect of time was significant with an increase of general well-being in the IP ($p_{adj} < .01$), but not in the OLP ($p_{adj} = .191$) group, whereas general well-being in the CG showed a trend to decrease ($p_{adj} = .071$). There was no significant interaction between group and time on sleep quality scores, $F(5, 443.85) = 0.90$, $p = .485$.

Rating of test anxiety at follow up, opinion on treatment idea, side-effects and adherence. Regarding the retrospective evaluation of the test situation (T5), the overall ANCOVA showed a significant overall effect of group, $F(2, 169) = 5.89$, $p = .003$. Planned contrasts indicated that mean retrospective test anxiety scores were rated significantly lower in the intervention groups (OLP/IP) compared to the CG at T5, $t(169), = -3.29$, $p = .001$, $d = -0.53$. However, retrospective test anxiety scores did not differ between the two intervention groups, $t(169) = 0.10$, $p = .918$, $d = 0.02$.

Table 3 provides an overview of the evaluation of the idea (positive, negative, neutral) towards the two interventions in the context of the open questions. The two independent raters had concordant judgments for 91.2% of the answers. A third rater was included for the remaining 8.8%.

No negative side-effects were reported, other than in the IP group, in which three subjects mentioned additional effects immediately after pill intake (i.e., dry mouth, goose bumps, warmth radiating from the abdomen). These effects were suggested during the pill intake in the study contact and were part of the IP response to demonstrate the effect of the pill (see supplementary material).

Regarding adherence, one participant (1.7%) in the OLP group and five participants (9.1%) in the IP group reported less than 80% adherence (i.e., forgot 9 or more pills).

Influence of study contact duration, treatment provider and moderation of treatment expectancy on primary outcome. Study contact duration was significantly associated with changes in test anxiety from T1 to T4, $F(1, 168) = 5.84$, $p = .017$. However, when including treatment group as an additional factor in the model, contact duration was no longer significant, $F(1, 166) = 0.01$, $p = .942$ and group remained significant, $F(2, 166) = 8.00$, $p < .001$. Treatment provider was not associated with changes in test anxiety, $F(1, 171) = 0.80$, $p = .373$.

Mean expectancy of relief across the 20 items of the test anxiety questionnaire was significantly different across the three groups, $F(2, 170) = 14.86$, $p < .001$. Participants receiving an intervention (IP/OLP) expected less symptoms compared to the CG, $t(169) = -5.76$, $p < .001$, $d = -0.92$, whereas scores of the two intervention groups were comparable, $t(169) = 1.47$, $p = .144$, $d = 0.28$. Mean expectancy of relief measures significantly correlated with endpoint test anxiety (T4; $r = 0.56$, $p < .001$). When including expectancy of relief as an additional covariate into

	OLP (*N* = 59) *N* (%)	IP (*N* = 55) *N* (%)
Positive	39 (66.1%)	38 (69.1%)
Negative	8 (13.6%)	10 (18.2%)
Neutral	12 (20.3%)	7 (12.7%)

Table 3. Ratings of the open-ended questions. Judgement of the idea regarding the respective interventions. *IP* imaginary pill, *OLP* open-label placebo.

the overall model, expectancy of relief was significantly associated with test anxiety, $F(1, 168) = 21.14$, $p < .001$, but treatment group remained significant, $F(2, 168) = 12.87$, $p < .001$.

Discussion

The present randomized controlled trial tested the effects of an IP against an OLP intervention and a CG on test anxiety in healthy students. We found that both IP and OLP significantly reduced test anxiety compared to the CG with a moderate-to-large effect size (d = − 0.71). These findings were comparable across all subscales of the test anxiety questionnaire (i.e., worry, emotionality, interference and lack of confidence). Interestingly, the beneficial effect was apparent over the course of the three weeks, starting after only one week of intervention. While study contact duration and treatment provider did not appear to be critical for changes in test anxiety, the observed effects were associated with treatment expectancy as this measure positively correlated with changes in test anxiety (r = 0.56). The retrospective assessment of the exam situation (follow-up T5) supports the superiority of the two interventions over the CG, as it indicated less retrospectively perceived anxiety during the exam situation. Consistent with the effects on our primary outcome, general well-being was significantly augmented in both intervention groups compared to the CG with a moderate to large effect (d = − 0.64). Overall sleep quality, however, was not affected by the intervention, i.e., all three groups showed comparable sleep quality during the three weeks. Test performances (i.e., continuous grades) were significantly better in the intervention groups compared to the CG with a small effect (d = 0.38). Overall, OLP and IP showed comparable results on all assessed outcomes. These findings question the necessity of the pill to produce positive treatment effects.

The effect sizes of the two interventions in the present study are slightly higher compared to a previous OLP trial testing openly prescribed placebos in test anxiety against no treatment with a between group effect size of d = 0.54[20], whereby test anxiety scores in both studied populations indicate average, non-clinical test anxiety[31]. The remarkable and rapid decreases in test anxiety in the intervention groups of the present study are noteworthy. The observed effect is comparable to the moderate-to-large effect (g = − 0.76) of a meta-analysis on various psychological interventions for test-anxious university students (i.e., psychological, study skill training, and/or combined intervention packages) against control conditions[32].

Extended or different placebo paradigms such as IPs aid to understand the mechanisms of OLP by systematically manipulating the treatment setting and application. As OLP and IP groups showed comparable results in all outcomes, the necessity of a physical placebo to produce positive treatment effects is called into question. Psychological components, for their part, may be sufficient on their own to exploit placebo effects which is supported by studies showing that triggering placebo effects without a physical treatment component is possible[8,11,33]. Research on placebo-like expectancy effects in pain analgesia is consistent with this: Peerdeman et al. (2017) showed less experienced pain in participants receiving instructions to vividly imagine a warm and impermeable glove preventing pain from cold before a cold pressor test, compared to a control imagery group instructed to imagine their hand without any reference to pain or cold water. This effect was mediated by expected pain[12]. Along these lines, expectancy of relief was also significantly associated with test anxiety in our study. However, the treatment group remained significant even after expectancy was included in our linear model, implying that not only expectancy but also other factors must account for the group-specific improvement in test anxiety. The effects can, for example, be discussed in the context of the embodied cognition theory, which states that a placebo effect can result unconsciously from embodied experiences by an internal act of imagining a particular state change in the body[16]. Similarly, conditioning effects may have played a role in our study, as even imagining something can activate corresponding brain areas (e.g.[15]). The Western cultural understanding of a pill underpins this line of reasoning as a pill in itself has a therapeutic meaning—learning from an early age to associate the pill and its effects, whereas no physical pill is required to trigger positive processes. Notably, mental imagery relies on similar neural processes to those of actual perception[34,35]. The ability to generate internal representations that retain the essential features of a perceptual experience suggests that mental imagery may have similar effects to actual experiences[12]. Consistent with the response expectancy theory[14] the findings of this study extend previous research on the mechanisms of placebo effects by showing that placebo effects on test anxiety can be induced not only by a physical cue, but also by imagining a pill and its effects. Overall, it can be suggested that OLP and IP may rely on the same underlying mechanisms (e.g., expectations, conditioning, embodied cognition), whereas these mechanisms can be triggered even in the absence of a physical pill.

Due to the COVID-19 pandemic, the study contact took place by means of a virtual clinical encounter. The present study is not the first to provide the OLP treatment remotely: Kube and colleagues[36], however, failed to replicate previous findings of OLP effects on allergic rhinitis[37,38]. They concluded that remote OLP provision is feasible, yet their effectiveness might be lower, as a physical encounter between patient and provider might be a prerequisite for OLPs to be effective[36]. However, our findings demonstrate that providing OLP and IP remotely is not only feasible but can also yield significant effects. A potential reason for the better effects in this online intervention compared to Kube et al. might be the younger sample (22.7 vs. 31.1 years) consisting only of students who may be more accustomed to online interactions. Whether the effects would be different with physical contact remains unclear and should be tested in a follow-up study.

This is the first study to conceptually extend ethically feasible placebo treatments by testing an IP intervention for test anxiety, taking OLP research a step further. It moreover corroborates important findings on OLP efficacy in a remote setting on a large sample. A manual including a five-step procedure was developed by our team to implement the IP intervention (see supplementary material). Manualized instructions used in the study further allowed for the control of many incidental factors to make accurate inferences about the interventions tested. Weekly assessments of primary and secondary outcomes moreover enabled observation of placebo effects over time. Also, there were less than 3% participants with missing data and reported nonadherence was low, especially in the OLP group.

Nevertheless, several aspects of this study need to be considered: Due to sample restriction and recruitment locations and routes, a largely female, young, academic sample resulted, limiting generalizability of the findings. Also, most outcome measures were self-reported and rather subjective than objective, raising questions of report and social desirability bias. Disappointment effects may have further played a role in the CG as they were not offered future treatment. In fact, 52.5% reported to be disappointed due to being allocated to the CG. However, given that test anxiety can be assumed to increase as the exam approaches[39], but the CG showed stable scores over time, it seems that despite disappointment this group also benefited from taking part in the study. In addition, adherence was self-reported, so we had no option to verify the reported values, eliminating the influence of social desirability bias. Further, because of planning reasons, a short time gap between study contact and start of the intervention occurred in some participants. However, this gap was kept to a minimum. Also, the conduct of the present study coincided with the start of the COVID-19 pandemic, which necessitated meeting with participants virtually. Due to the remote setting, the participants in the OLP group received the envelope with the placebo pills in advance: although not knowing about the contents of the envelope and being instructed not to open the envelope until the study contact occurred, we had no way of controlling this behavior. Nevertheless, positive effects of interventions could be observed and implementation remotely was feasible. Considering this, we assessed changes in test anxieties due to the pandemic-related circumstances which were almost evenly distributed across participants—with some reporting unchanged (33.5%), higher (34.1%) or less (32.4%) anxiety. Comparisons of within changes of participants should, however, control for these complicated circumstances. Future investigations should test OLP and IP with physical contact and no pandemic-related restrictions.

The present study is the first to conceptually expand on previous OLP studies by eliminating the physical pill as a treatment component and testing an IP intervention. Results indicate a moderate-to-large effect of both interventions on test anxiety and general well-being in a large cohort of 173 healthy students. These findings demonstrate that placebo effects can be harnessed without the use of a physical pill. The IP intervention could thus serve as a stand-alone or adjunct treatment to maximize and boost placebo effects in clinical practice, as indicated by the ethical principle of "beneficence"[1,40,41]. As an ethical, cost-effective, easily applicable and fully patient-centered method, the IP intervention has potential and should be tested in other settings, conditions and populations.

Methods

Experimental design. Between March 2020 and July 2021, we conducted an online randomized controlled, parallel group trial at the Division of Clinical Psychology and Psychotherapy at the Faculty of Psychology, University of Basel, Switzerland, in order to test the effects of an IP and OLP intervention compared to a CG in healthy students with test anxiety. Written informed consent was obtained from each participant before participation in the study. The Ethics Committee of the Faculty of Psychology, University of Basel approved the design and informed consent of the study. This study was carried out in accordance with the protocol and principles enunciated in the current version of the Declaration of Helsinki. The study was registered at ClinicalTrials.gov: NCT04250571 (31/01/2020).

Study population. Participants were recruited via web- and print-based advertisement (title: "Efficacy study of two treatment methods for test anxiety") and registered online for the study. Potential enrollees had to be students of the University of Basel aged between 18 and 65 years. To meet inclusion criteria, participants had to have an exam at the end of the semester, have self-reported test anxiety, being healthy by self-report (i.e., no known current or chronic primary pain disorders or psychiatric disorders) and be sufficiently proficient in German. Exclusion criteria were use of medications (psychoactive or narcotic), being in psychological or psychiatric treatment, taking psychotropic drugs, being a master student in Psychology (due to prior knowledge about placebo mechanisms), allergy to one of the ingredients of the placebo pills (see supplementary material), and problems swallowing pills. All participants were reimbursed either financially or with credit points.

Study procedure. The study procedure is depicted in Fig. 5. Interested participants were directed to an online survey page providing information about the nature and purpose of the study. Upon providing online informed consent, participants were checked online for inclusion and exclusion criteria. Eligible participants were randomly assigned to one of three study group. Baseline assessments of primary and secondary outcomes were completed online two or less days before the study contact (T1). The study contact, in which participants received one of three interventions (four to three weeks before the exam), was held online via the standard video call software of the University of Basel, zoom (https://zoom.us/), the use of which was approved by the ethics committee. Expectancy of relief was assessed immediately after the study contact. Study contacts were distributed over the time period of four to three weeks before the exam for resource management reasons (number of treatment provider). However, treatment started exactly three weeks before the exam as indicated by the receipt of a reminder e-mail in both intervention groups, i.e., the treatment duration was the same. Again, two weeks (T2), one week (T3) and two or less days before the exam (T4) all three groups were asked to complete online assessments of primary and secondary outcomes, as well as to answer one question regarding their intervention adherence. After the exam, there was a final online assessment (T5) to evaluate retrospective experiences of the exam situation, to assess side-effects during the treatment period and to answer open questions respective to the group (e.g., possible feelings of disappointment to be assigned to the CG, see supplementary material). Finally, all participants were asked about their examination grade (approximately two months after the exam).

Study arms. In total, there were three study arms, i.e., CG, OLP, and IP.

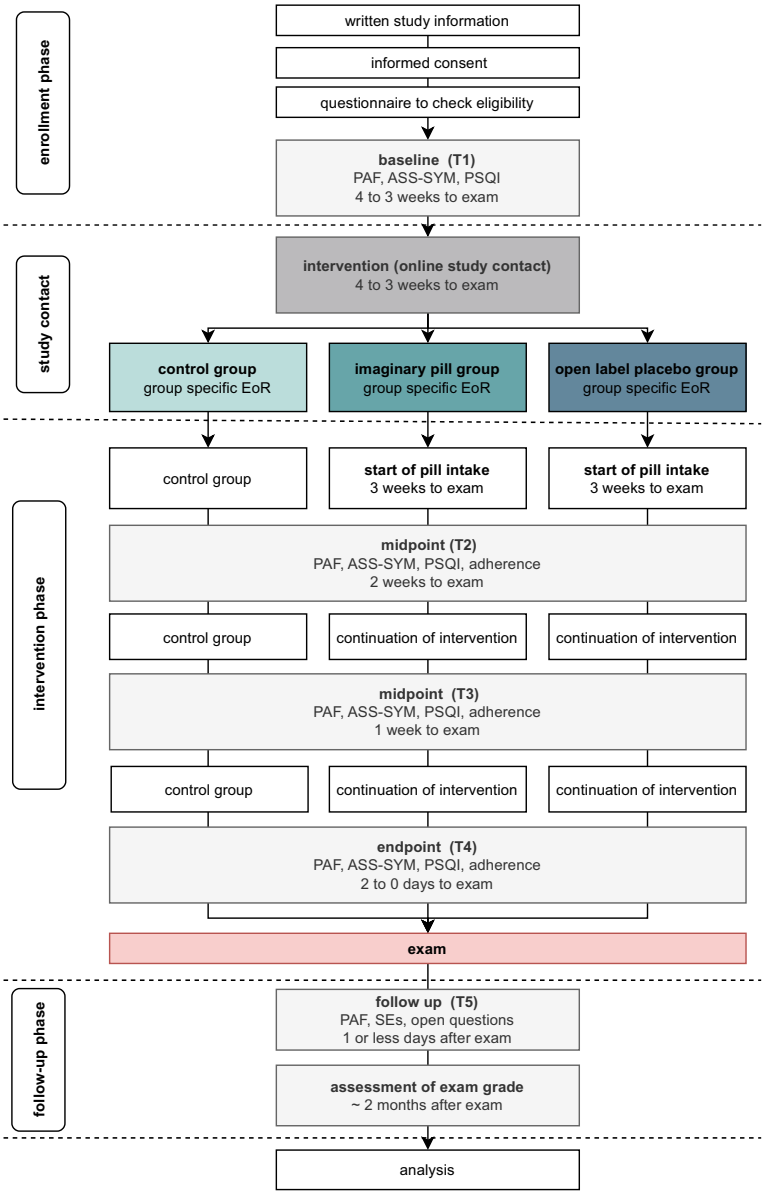


Figure 5. Procedure of the study. *Note: ASS-SYM* Änderungssensitive Symptomliste (general well-being), *EoR* expectancy of test anxiety relief, *PAF* Prüfungsangstfragebogen (test anxiety questionnaire), *PSQI* pittsburgh sleep quality index, *SEs* side-effects, *SDD* sociodemographic data.

Participants allocated to the IP group did not take a physical pill, but imagined taking a pill along with verbal suggestions from the treatment provider during the study contact. Hence, the idea of IPs has resemblance to the clinical application of hypnosis[42]. Participants in this group received a procedure in accordance with the technique by de Shazer[5] and a structure proposed by Niels Bagge[6]. Detailed formulation and translation to German was performed by the local study team (SB, DS, CL, JG). The instruction consisted of a procedure including five steps: (1) identifying the persons' problem and the desired state, (2) building trust in the treatment, (3) constructing a personally meaningful pill, (4) taking the IP, (5) suggestions for self-administration in real life and building adherence (see supplementary material). Importantly, step 2 consisted of teaching participants about findings of (open-label) placebo and imagination research. At the end of the intervention, participants in the IP

group had to describe their individual elaborated pill (size, shape, pill kind, color, packaging) and its effects in an interactive document. They sent the completed document back to the treatment provider and were able to print it out for their own use. Participants were asked not to take any physical aids, such as a candy, to facilitate their imagination, ensuring that the groups remained distinguishable in their specific ingredients. Participants were instructed to take two IPs a day for three weeks until the exam takes place and received daily e-mail reminders during that period to remember their IP intake.

In the OLP group participants obtained the information that they were receiving inert blue pills (i.e. "P-Dragees" blau Lichtenstein manufactured by Zentiva Pharma GmbH) and were given a treatment rationale in accordance with previous OLP studies (e.g.,[29,30]; see supplementary material), that encompassed four discussion points. In order to keep the OLP rationale similar to the one of the IP, a brief introduction was added at the beginning of the intervention, elaborating on what comprises the persons' problem (how do symptoms express themselves; how does the person wish to feel). Hence, the rationale was structured as follows: (1) identifying the OLP-sensitive problem, (2) deceptive as well as OLPs are efficacious, (3) one mechanism of placebo is conditioning, (4) an open attitude towards the treatment can be helpful but is not necessary for its effect, (5) taking the pill faithfully is important. Participants were instructed to take two placebo pills a day for three weeks until the exam takes place. The placebo pills were sent in an envelope to participants via postal mail prior to the online study contact or if participants did not wish to disclose their postal addresses, they were given the option of a personal handover by a member of the study team. Participants did not know about the content of the envelope and were instructed to not open the envelope until the study contact takes place. Daily e-mails were also sent to this group as a reminder to take the placebo pills.

In order to control for factors not considered characteristic for the intervention, the CG was fashioned according to the intervention groups (i.e., characteristic components were the pill intake and intervention-specific rationales[43]). Participants were (1) reminded of the importance of this group, (2) asked about the nature of their exam, (3) about their problem (i.e., test anxiety) and the wished-for state, (4) and about learning strategies. The design of this group attempted to keep interaction time comparable and to account for the structural equivalence between the CG and intervention groups, e.g., by allowing the CG to talk about the problem (i.e., test anxiety) to enable a "fair" comparison of groups[44]. Despite the interventional nature of this study arm, no advice or problem-solving task was given (see supplementary material).

Study contacts on zoom were carried out by five female treatment providers. Although not all treatment providers had the same number of study contact appointments, the proportion of participants per group were evenly distributed among them. Average duration of interventions was 31 minutes (IP = 44 min, OLP = 29 min, CG = 20 min).

Randomization and blinding. A random allocation sequence was created by SB using the built-in random number generator in Microsoft Excel for Mac, version 16.53. Participants were enrolled in the pre-generated list in order of their study registration and assigned by master students to interventions accordingly. All participants were informed about their assigned group at the study contact via zoom. Due to the study design, the providers were unblinded to the treatment they were administering. However, the encounter was kept constant in all groups through a standardized protocol. Also, except for the study contact on zoom, all communication was via e-mail contact (e.g., sending links for online assessments), using the same e-mail templates for all three groups to ensure the same type of interaction.

Outcome measures. The primary outcome was test anxiety measured by means of the "Prüfungs-Angst Fragebogen" (PAF; English: "test anxiety questionnaire"[31]). The questionnaire consists of 20 items with four subscales (worry, emotionality, interference, lack of confidence) with scores ranging from 20 to 80 points. Each item is rated on a 4-point Likert scale (*1—almost never* to *4—almost always*). Secondary outcomes were sleep quality and general well-being. Sleep quality was assessed by means of the "Pittsburgh Sleep Quality Index" (PSQI[45,46]). The PSQI is an 18 item self-rating questionnaire forming 7 subscales. To fit our time frame, we adjusted the time interval to the last week (7 days). To assess general well-being the ASS-SYM symptom list was used (Änderungssensitive Symptomliste[47]). This list is composed of 48 items and 6 subscales. Lower scores indicate less symptoms (i.e., higher general well-being). All measures were assessed four to three weeks prior to the exam (T1; baseline assessment), two and one week prior the exam (T2–T3; midpoint assessments) and two or less days prior the exam (T4; endpoint assessment).

Test performance of each participant was collected as another secondary outcome (approximately two months after the exam). Students received as a test performance either a continuous grade, ranging from a minimum of 1 (very poor) to a maximum of 6 (very good) in the Swiss grading system, or a binary test score (pass or fail), where a grade greater than or equal to 4 (sufficient) is considered a pass. Other outcomes of interests included sociodemographic data (SDD) assessed at T1. Immediately after the intervention, expectancy of test anxiety relief[48] according to the received intervention was assessed using an ad-hoc constructed questionnaire with each item of the primary outcome on a numeric rating scale from 1 to 4 (e.g., based on the intervention you have received, how strong would you expect the following symptoms to be present before your next exam on a scale from 1—*almost never* to 4—*almost always*) as e.g. used in[49]. Furthermore, within the intervention groups, adherence was assessed weekly with a single item asking for how often someone forgot the actual or imagined pill intake in the last week. In total, each participant assigned to one of the two intervention groups had to take 42 pills (i.e., 2 pills × 21 days). A sum score was computed to determine overall adherence. Adherence was defined as > 80% (i.e., 9 or more missed pills). Additional variables were collected on the same day or at most one day after the exam (T5; follow up assessment) including retrospective experience of the exam using the test anxiety questionnaire (i.e., the wording of the introduction was changed as follows: please read through each statement

and choose from the four answers 1—*almost never* to 4—*almost always* the one that indicates best how you were feeling during the exam), side-effects (i.e., (1) did you experience side effects, (2) if yes, give a description, (3) how severe were they from 0—*none* to 100—*very severe*, (4) when was the onset, (5) how long did the they last, (6) was there a connection with participation in our study?) and open questions respective to group allocation for example about the idea of intervention (i.e., what do you think about the idea of taking placebo/imaginary pills?; see supplementary material for all open-ended questions). All outcome variables were assessed by means of online surveys using Limesurvey (limesurvey.org).

Statistical analyses. Statistical analyses were carried out using the open-source software environment RStudio. For all analyses, significance level was set at $\alpha = 5\%$. Using a conservative power analysis on the basis of an *F*-Test and an ANCOVA for three groups, we calculated that a total sample size of $N = 206$ for a power of 0.9 and a total sample size of $N = 158$ for a power of 0.8 would be necessary to detect a medium effect size of $f = 0.25$ (i.e., $d = 0.5$) with an alpha-level of 0.05, using the statistical software G*Power. On this basis we decided on a total sample size of a minimum of 165 participants. Considering dropouts (e.g., due to increased nonattendance because of the COVID-19 pandemic), we planned to include and randomize slightly more than 55 per group ($N \sim 60$). Cohen's *d* was used to assess the size of effects.

Initially planned multiple imputation was not conducted, as there were less than 3% participants with missing data and the missingness appeared to be completely random (e.g., due to nonattendance at exams because of COVID-19). The five participants with missing data were thus not considered for analyses (see Fig. 1 for reasons for exclusion).

To assess differences in changes from baseline (T1) to endpoint (T4; primary analyses) and follow-up (T5) in test anxiety across the three groups, two separate omnibus tests (ANCOVA) using treatment group as the independent factor and baseline (T1) as covariate to control for baseline differences[50] were computed to test for overall effects. Orthogonal contrasts were computed to evaluate intergroup differences in the change from baseline (T1) to study endpoint (T4). The contrasts were: CG < IP + OLP and IP < OLP. To evaluate changes over time, we conducted a two-way mixed analysis of variance (ANOVA) using group as between-subject factor and time (T1-T4) as within-subject factor. Bonferroni adjustments accounted for multiple testing within post-hoc tests.

To analyze test performance across groups, we followed a two-step approach as there were continuous (grades) as well as binary (pass/fail) test scores. First, we performed an analysis only with participants having a continuous test score (1–6) using an ANOVA to test for overall effects and above-mentioned contrasts. Second, all continuous variables were transformed into a binary test score (pass ≥ 4; fail < 4) and reported as percentages of passing.

In order to investigate differences in treatment expectancy of relief across groups, an overall ANOVA was performed using the expectancy scores as outcome and group as between-subject factor. A priori contrasts were then used to explore differences across groups. Furthermore, we calculated correlations in order to investigate possible relationships between treatment expectancy of relief and test anxiety[51] and computed a linear model with test anxiety from T1 to T4 as dependent factor and expectancy of relief as independent factor to investigate their impact on the effects.

To investigate the influence of study contact duration and treatment provider on test anxiety we used a linear model with the corresponding variable as independent factor and changes in test anxiety from T1 to T4 as dependent variable. To analyze the open-ended questions about attitudes toward the idea about the two interventions, two independent raters rated each statement as "positive," "negative," or "neutral". When ratings differed, a consensus was reached by a third rater.

Data availability

Access to data from this study may be obtained by contacting the corresponding author.

Received: 2 November 2022; Accepted: 7 February 2023
Published online: 14 February 2023

References

1. Enck, P., Bingel, U., Schedlowski, M. & Rief, W. The placebo response in medicine: Minimize, maximize or personalize?. *Nat. Rev. Drug Discov.* **12**, 191–204 (2013).
2. Blease, C., Colloca, L. & Kaptchuk, T. J. Are open-label placebos ethical? Informed consent and ethical equivocations. *Bioethics* **30**, 407–414 (2016).
3. Charlesworth, J. E. G. *et al.* Effects of placebos without deception compared with no treatment: A systematic review and meta-analysis. *J. Evid. Based Med.* **10**, 97–107 (2017).
4. von Wernsdorff, M., Loef, M., Tuschen-Caffier, B. & Schmidt, S. Effects of open-label placebos in clinical trials: A systematic review and meta-analysis. *Sci. Rep.* **11**, 3855 (2021).
5. de Shazer, S. The imaginary pill technique. *J. Strateg. Syst. Ther.* **3**, 30–34 (1984).
6. Bagge, N. *The Imaginary Placebo Pill: Isolating Psychological Components of Placebo Pills for Treatment in Clinical Practice* [Conference presentation abstract]. Second official conference of the Society for Interdisciplinary Placebo Studies (SIPS). (Leiden, NL, 2019).
7. Wai-Lan Yeung, V., Geers, A. L. & Colloca, L. Merely possessing a placebo analgesic improves analgesia similar to using the placebo analgesic. *Ann. Behav. Med.* **54**, 637–652 (2020).
8. Gaab, J., Kossowsky, J., Ehlert, U. & Locher, C. Effects and components of placebos with a psychological treatment rationale—Three randomized-controlled studies. *Sci. Rep.* **9**, 1421 (2019).
9. Davis, A. J., Hettinga, F. & Beedie, C. You don't need to administer a placebo to elicit a placebo effect: Social factors trigger neurobiological pathways to enhance sports performance. *Eur. J. Sport Sci.* **20**, 302–312 (2020).
10. Chen, P.-H.A. *et al.* Socially transmitted placebo effects. *Nat. Hum. Behav.* **3**, 1295–1305 (2019).
11. Crum, A. J. & Langer, E. J. Mindset matters: Exercise and the placebo effect. *Psychol. Sci.* **18**, 165–171 (2007).

12. Peerdeman, K. J., van Laarhoven, A. I. M., Bartels, D. J. P., Peters, M. L. & Evers, A. W. M. Placebo-like analgesia via response imagery. *Eur. J. Pain* **21**, 1366–1377 (2017).
13. Moerman, D. E. & Jonas, W. B. Deconstructing the placebo effect and finding the meaning response. *Ann. Intern. Med.* **136**, 471–476 (2002).
14. Kirsch, I. Response expectancy theory and application: A decennial review. *Appl. Prev. Psychol.* **6**, 69–79 (1997).
15. Ogino, Y. *et al.* Inner experience of pain: Imagination of pain while viewing images showing painful events forms subjective pain representation in human brain. *Cereb. Cortex* **17**, 1139–1146 (2007).
16. Thompson, J. J., Ritenbaugh, C. & Nichter, M. Reconsidering the placebo response from a broad anthropological perspective. *Cult. Med. Psychiatry* **33**, 112–152 (2009).
17. Bischofsberger, L. *Prävalenz und Ausprägungen von Prüfungsangst bei Studienanfänger/-Innen der Medizin*. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) (2022).
18. Ergene, T. Effective interventions on test anxiety reduction: A meta-analysis. *Sch. Psychol. Int.* **24**, 313–328 (2003).
19. von der Embse, N., Jester, D., Roy, D. & Post, J. Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *J. Affect. Disord.* **227**, 483–493 (2018).
20. Schaefer, M. *et al.* Open-label placebos reduce test anxiety and improve self-management skills: A randomized-controlled trial. *Sci. Rep.* **9**, 13317 (2019).
21. Kleine-Borgmann, J. *et al.* Effects of open-label placebos on test performance and psychological well-being in healthy medical students: A randomized controlled trial. *Sci. Rep.* **11**, 2130 (2021).
22. Sugarman, M. A., Loree, A. M., Baltes, B. B., Grekin, E. R. & Kirsch, I. The efficacy of paroxetine and placebo in treating anxiety and depression: A meta-analysis of change on the Hamilton rating scales. *PLoS ONE* **9**, e106337 (2014).
23. Sugarman, M. A., Kirsch, I. & Huppert, J. D. Obsessive-compulsive disorder has a reduced placebo (and antidepressant) response compared to other anxiety disorders: A meta-analysis. *J. Affect. Disord.* **218**, 217–226 (2017).
24. Faria, V. *et al.* Do you believe it? verbal suggestions influence the clinical and neural effects of escitalopram in social anxiety disorder: A randomized trial. *EBioMedicine* **24**, 179–188 (2017).
25. Oosterbaan, D. B., van Balkom, A. J. L. M., Spinhoven, P. & van Dyck, R. The placebo response in social phobia. *J. Psychopharmacol.* **15**, 199–203 (2001).
26. Cuijpers, P. *et al.* Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clin. Psychol. Rev.* **34**, 130–140 (2014).
27. Ahmadzad-Asl, M. *et al.* Systematic review and meta-analysis of the placebo effect in panic disorder: Implications for research and clinical practice. *Aust. N. Z. J. Psychiatry.* **56**, 1130–1141 (2022).
28. Kim, S., Wollburg, E. & Roth, W. T. Opposing breathing therapies for panic disorder: A randomized controlled trial of lowering vs raising end-tidal PCO2. *J. Clin. Psychiat.* **73**, 13479 (2012).
29. Kaptchuk, T. J. *et al.* Placebos without deception: A randomized controlled trial in irritable bowel syndrome. *PLoS ONE* **5**, e15591 (2010).
30. Carvalho, C. *et al.* Open-label placebo treatment in chronic low back pain: A randomized controlled trial. *Pain* **157**, 2766–2772 (2016).
31. Hodapp, V., Rohrmann, S. & Ringeisen, T. *PAF–Prüfungsangstfragebogen [the brief German test anxiety inventory]* (Hogrefe, 2011).
32. Huntley, C. D. *et al.* The efficacy of interventions for test-anxious university students: A meta-analysis of randomized controlled trials. *J. Anxiety Disord.* **63**, 36–50 (2019).
33. Kong, J. *et al.* Enhancing treatment of osteoarthritis knee pain by boosting expectancy: A functional neuroimaging study. *NeuroImage Clin.* **18**, 325–334 (2018).
34. Kosslyn, S. M., Ganis, G. & Thompson, W. L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2**, 635–642 (2001).
35. Mcnorgan, C. A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. *Front. Hum. Neurosci.* **6**, 285 (2012).
36. Kube, T., Hofmann, V. E., Glombiewski, J. A. & Kirsch, I. Providing open-label placebos remotely—A randomized controlled trial in allergic rhinitis. *PLoS ONE* **16**, e0248367 (2021).
37. Schaefer, M., Harke, R. & Denke, C. Open-label placebos improve symptoms in allergic rhinitis: A randomized controlled trial. *Psychother. Psychosom.* **85**, 373–374 (2016).
38. Schaefer, M., Sahin, T. & Berstecher, B. 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 (2018).
39. Lotz, C. & Sparfeldt, J. R. Does test anxiety increase as the exam draws near?—Students' state test anxiety recorded over the course of one semester. *Pers. Individ. Differ.* **104**, 397–400 (2017).
40. Evers, A. W. *et al.* Implications of placebo and nocebo effects for clinical practice: Expert consensus. *Psychother. Psychosom.* **87**, 204–210 (2018).
41. Evers, A. W. *et al.* What should clinicians tell patients about placebo and nocebo effects? Practical considerations based on expert consensus. *Psychother. Psychosom.* **90**, 49–56 (2021).
42. Lynn, S. J., Laurence, J.-R. & Kirsch, I. Hypnosis, suggestion, and suggestibility: An integrative model. *Am. J. Clin. Hypn.* **57**, 314–329 (2015).
43. Grünbaum, A. The placebo concept. *Behav. Res. Ther.* **19**, 157–167 (1981).
44. Baskin, T. W., Tierney, S. C., Minami, T. & Wampold, B. E. Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *J. Consult. Clin. Psychol.* **71**, 973 (2003).
45. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R. & Kupfer, D. J. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res.* **28**, 193–213 (1989).
46. Riemann, D. & Backhaus, J. *Behandlung von Schlafstörungen. Ein psychologisches Gruppenprogramm (Materialien für die psychosoziale Praxis). Beltz Psychologie Verlags Union*. (Weinheim, 1996).
47. Krampen, G. *Änderungssensitive Symptomliste zu Entspannungserleben, Wohlbefinden, Beschwerden-und Problembelastungen: ASS-SYM*. (Hogrefe, 2006).
48. Papakostas, G. I. & Fava, M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur. Neuropsychopharmacol.* **19**, 34–40 (2009).
49. Locher, C. *et al.* Is the rationale more important than deception? A randomized controlled trial of open-label placebo analgesia. *Pain* **158**, 2320–2328 (2017).
50. Vickers, A. J. & Altman, D. G. Analysing controlled trials with baseline and follow up measurements. *BMJ* **323**, 1123–1124 (2001).
51. Pearson, J., Naselaris, T., Holmes, E. A. & Kosslyn, S. M. Mental imagery: Functional mechanisms and clinical applications. *Trends Cogn. Sci.* **19**, 590–602 (2015).

Acknowledgements

The authors would like to thank Berfin Bakis for her assistance with the data monitoring, assistance in conducting the study and support in writing the manuscript. We also thank all Masterstudents: Annina Dettwiler, Alina Burkart, Yannick Schönenberger, Marlon Enzmann, Alex Busch, and interns Clara Bruhin, Jana Hochreuter, for aiding with conducting the study and analyzing the data.

Author contributions

Conceptualization: S.B., D.S., N.B., I.K., C.L., C.C., J.G.; Methodology: S.B., D.S., N.B., I.K., C.L., C.C., J.G.; Formal Analyses: S.B., D.S., I.K., C.L.; Investigation: S.B., D.S.; Writing—original draft: S.B., D.S.; Writing—review & editing: S.B., D.S., N.B., I.K., C.L., C.C., J.G.; Visualization: S.B.; Supervision: N.B., I.K., C.L., C.C., J.G.; Project Administration: S.B.

Funding

This work was not supported by any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. CL received funding from Swiss National Science Foundation (SNSF): PZ00P1_201972.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-29624-7.

Correspondence and requests for materials should be addressed to S.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary materials for

Imaginary pills and open-label placebos can reduce test anxiety by means of placebo mechanisms

Sarah Buergler, Dilan Sezer, Niels Bagge, Irving Kirsch, Cosima Locher, Claudia Carvalho, & Jens Gaab

This file includes:

Supplementary Text

Supplementary Table

Supplementary Text

Ingredients of placebo pills

Lactose-Monohydrat, Magnesiumstearat, mikrokristalline Cellulose, hochdisperses Siliciumdioxid, weisser Ton, Macrogolglycerolhydroxystearat, Arabisches Gummi, Montanglycolwachs, Povidon, Talkum, Titandioxid, Patentblau-V-Aluminiumsalz, Calciumcarbonat, Sucrose, Glukosesirup, Maisstärke, Macrogol 6000

Rationales

Imaginary Pill (IP) rationale (translated from German in English)	Open-label Placebo (OLP) rationale (translated from German in English)	Control Group (CG) rationale (translated from German in English)
1. Identifying the IP-sensitive problem and the desired state: Before I explain the concept of the open administration of placebos and how we use these placebo effects with the imaginary pill intervention, I would like to know more about your test anxiety and your preparation stress. Could you use a previous exam situation to describe what the symptoms feel like? I ask you now to think back to that bad exam/situation. Can you now describe to me, based on this previous exam situation, what the symptoms of your exam anxiety felt like? What are the sensations in your body? What are your thoughts and emotions when you experience this anxiety? Now, thinking of the upcoming exam, if you had to specify how strong these symptoms are from 0-10 (not at all – very strong) in this moment, what would you say? And now can you describe how you would like to feel in your exam phase/during the exam? Can you describe to me a specific situation where you feel this? What are your feelings in this situation? What are your thoughts? What physical sensations do you have in this situation? Try to put yourself in this positive state. [Check to see if the person really knows how they want to feel. Get as precise as possible].	**1. Identifying OLP-sensitive problem** Before I explain the concept of the open administration of placebos and how we use these placebo effects with the imaginary pill intervention, I would like to know more about your test anxiety and your preparation stress. Could you use a previous exam situation to describe what the symptoms feel like? I ask you now to think back to that bad exam/situation. Can you now describe to me, based on this previous exam situation, what the symptoms of your exam anxiety felt like? What are the sensations in your body? What are your thoughts and emotions when you experience this anxiety? Now, thinking of the upcoming exam, if you had to specify how strong these symptoms are from 0-10 (not at all – very strong) in this moment, what would you say? And now can you describe how you would like to feel in your exam phase/during the exam? Can you describe to me a specific situation where you feel this? What are your feelings in this situation? What are your thoughts? What physical sensations do you have in this situation? Ok, what you say is all very understandable and I hope we can help you with our intervention to reach the positive state you just described. Is it okay if I now explain to you the concept of open placebos?	**1. Explaining importance of group** As you already know from the study information, we randomly assign all study participants to one of the three study groups. You have been assigned to the control group, which means you will not receive a treatment. This group and your participation is very important for our study. Only through the control group can we see how symptoms naturally behave when you do not receive a treatment. So, we are also asking you to fill out all the online surveys accordingly. You will still receive weekly surveys.
2. Building trust in the treatment Is it ok if I now explain the concept of open-label placebos and how you can use this placebo effect with the imaginary pill for your goal – the positive experience, we've just talked about? OK, we know from clinical research that placebos have significant	**2. Deceptive and OLPs are effective** We know from clinical research that placebos have significant effects on pain, depression and anxiety and that these effects can even be demonstrated in changes in brain activity and the release of neurotransmitters. As mentioned earlier, scientists	**2. Nature of exam** In your case, we would be interested in how the exam stress and anxiety manifests itself and what your general learning strategies are. I will possibly make some notes on this. Before we get to your exam

effects on pain, depression and anxiety and that these effects can even be demonstrated in changes in brain activity and the release of neurotransmitters. As mentioned earlier, scientists previously assumed that placebo pills can only help if they are given covertly, i.e. with deception. Now, however, more recent studies suggest that this is not the case. This means that placebos can work even if the patient knows that it is a placebo. We are incorporating this approach in our study. Many double-blind randomized studies show that the placebo effect is very effective for many complaints. This means that placebos can relieve pain, cramps and gastrointestinal complaints, among other problems, and also have a very positive effect on mood. Especially for chronic back pain and irritable bowel syndrome, the open-label placebo treatment has been shown to be very effective, even in patients where nothing else has worked. Here at the division, a study has already been carried out in which a placebo cream was used for the treatment of heat-induced pain. And there too we found large placebo effects. A positive placebo effect has also been shown for test anxiety. This has been shown recently by a study from Germany, where they tested open placebos also in students. We are now considering the possibility that if placebos work, even though we know that they are placebos, then we could simply omit the sugar pill and imagine the pill and still have all the placebo effects. A reaction to placebos is not only triggered by the placebo pill itself, but also by the imaginative meaning that is both consciously and automatically attributed to the placebo pill. Imagination research shows, for example, that the idea of something activates the same areas of the brain as when one actually sees or experiences something. A study has also shown that the idea of exercising in a gym can already lead to muscle growth. Accordingly, it is possible to imagine taking this pill and achieve a similar effect as if you were taking a real pill. And this is exactly what I would like to discuss and practice with you. Okay for you?	previously assumed that placebo pills can only help if they are given covertly, i.e. with deception. Now, however, more recent studies suggest that this is not the case. This means that placebos can work even if the patient knows that it is a placebo. We are incorporating this approach in our study. Many double-blind randomized studies show that the placebo effect is very effective for many complaints. This means that placebos can relieve pain, cramps and gastrointestinal complaints, among other problems, and also have a very positive effect on mood. Especially for chronic back pain and irritable bowel syndrome, the open-label placebo treatment has been shown to be very effective, even in patients where nothing else has worked. Here at the division, a study has already been carried out in which a placebo cream was used for the treatment of heat-induced pain. And there too we found large placebo effects. A positive placebo effect has also been shown for test anxiety. This has been shown recently by a study from Germany, where they tested open placebos also in students.	anxiety itself, I'd like to ask you questions about the nature of the exam: - What format does the exam take? Is it written or oral? - Are you generally more afraid of written/oral (repeat what was said) exams, compared to exams that have a different format? - Is it a repetition exam? - *If not yet clear:* Does the test anxiety also have to do with the subject in which the exam takes place? In which subject is the exam? What does this subject involve? - What makes the exam so difficult or scary for you? - How often are you afraid of an exam to this extent or are you stressed because of the exam (in every learning phase or especially now)? (Possibly why especially now?)

3. Constructing a personally meaningful pill The first step is to find an imaginary pill for you. Recall the positive state, that you described earlier and the experience of relief you would like to feel. Suppose there was a pill that could bring you to that state, what effects would that pill have, how would it help you reach that state? Imagine there was a pill that could have all these positive effects. What would this pill look like (regarding color, shape and size)? And is the pill packaged also? [Wait and trust, that the person will come up with a pill. If a picture of such a pill is cannot be formed, then offer pill characteristics to choose from, for example: "The pill could be round, oval, (...)" etc.]	**3. One mechanism of placebos: Conditioning** Next, I would like to explain in more detail why placebos can alleviate symptoms. A very important explanation is that the body automatically reacts to the intake of medication. From an early age we learn that pills and effects are related, it results in a learning effect, so to speak. Accordingly, swallowing the pill alone can lead to symptom relief. The physiological reaction of our body to placebos is comparable to this. We know that when placebos work, they release neurotransmitters such as endorphins and dopamine, automatically activating specific areas of the brain. These neurotransmitters, in turn, can relieve symptoms or have a positive effect on mood.	**3. Talking about the problem (test anxiety) and the wished-for state** - Can you tell me specifically about a bad exam (it can also be a lecture or something similar) that you have had in the past and where you were very afraid? [Ask person to actually name an exam, the more specific the better]. I ask you now to think back to that bad exam/situation [wait until person remembers]. Now, using that previous exam situation, can you describe to me what the symptoms of your exam anxiety felt like? What were the sensations in your body based on your experience? What are thoughts and emotions that went through your mind? - Could you use a previous exam situation to describe what the symptoms feel like? I ask you now to think back to that bad exam/situation. Can you now describe to me, based on this previous exam situation, what the symptoms of your exam anxiety felt like? What are the sensations in your body? What are your thoughts and emotions when you experience this anxiety? Now, thinking of the upcoming exam, if you had to specify how strong these symptoms are from 0-10 (not at all – very strong) in this moment, what would you say? - And now can you describe how you would like to feel in your exam phase/during the exam?
4. Taking the IP Now imagine the pill described in detail as if it were a real pill. You can ascribe so much reality to the pill that taking it is experienced as if you were swallowing a real pill. It may take some practice. The effect may be stronger and the procedure easier for you if you have done it several times. Now I would suggest that you take your imaginary pill, to try this. You can close your eyes, if you want to. Just think of it as a regular pill. Imagine the pill and how it is packaged. Imagine how you take the pill out of the packaging and how you hold it in your hand. Bring it to your mouth. Swallow the pill slowly. Now it is in your body and starts to work. Maybe you can already feel the effects of the pill. Try to feel what the pill does to you. Maybe the pill has also other effects,	**4. An open attitude towards the treatment can be helpful but is not necessary** It's also absolutely okay if you have doubts that placebos work. As mentioned before, placebos can work automatically, which means they can work even if you have doubts.	**4. Learning strategies** - Now I'm still wondering what your general learning strategies are: Do you work in study groups or more alone or both? - Do you study with summaries, mind maps, study plans or flashcards? Thank you very much for your answers to the many questions, it is very informative.

such as making your mouth dry. You might get warm or a little dizzy. You have now had your first experience of such an imaginary pill taking. Try to remember this state so that you can recall it on your own. Now, if you had to indicate again after taking your imaginary pill how strong at the moment your symptoms are from 0-10 when you think about the upcoming exam, what would you say?		
5. Suggestions for self-administering in real life and building adherence For the effect of this intervention it is now important that you take such an imaginary pill twice a day from the time when you receive a reminder per e-mail: once in the morning and once in the evening, in order to reach the desired state (up to the exam). Before taking the pill, take a little time to recall the image of the pill you have just described. I also ask you to fill out the announced surveys once a week until the exam. You will receive an e-mail with the link at the right time, so that you remember to do it. Then one more thing: In order for us to really be able to identify what the effects of an imaginary pill are, we ask you not to take sweets like Sugus or Tiktak to make it easier for you to imagine. As said before, it is best to simply take your time and take the imaginary pill twice a day for three weeks. We will also send you daily reminders that you are reminded to take the imaginary pill. I am aware that the concept of the imaginary pill may sound strange to you at first. But we would like to find out whether you can reach the desired state if you imagine taking a pill every day to relieve your test anxiety. We have developed this procedure here at the university in collaboration with experts from all over the world and we really believe in the effectiveness of this treatment. Also, because there are already several cases from the clinic where the imaginary pill treatment has shown very good effectiveness. Therefore, I would like to encourage you to give the imaginary pill a chance and see what happens.	**5. Taking the pill faithfully is important** Therefore, it is important that you take the placebos regularly and according to the prescription. This means for you that you have to take the placebo pills faithfully in order to feel an effect. It is important for you to know that for some people the effects occur earlier and for others later. When you take the pills, we recommend that you also be aware of what the pills are supposed to help you against, i.e. to achieve the positive state you described earlier. I am aware that this may sound unfamiliar to you at first. However, we want to find out what happens to your symptoms when you take placebo pills every day. Therefore, I would like to encourage you to give the open placebo treatment a chance and see what happens. Now you may open the parcel and take out the box with the placebo pills in it. Please take two pills every day for the next three weeks (until the exam) from the time you receive a reminder by e-mail. It is best to take the pills at the same time in the morning and in the evening. There is also a small envelope in the package, which you can open right away. On the envelope you will find information on how to take the pills. As we said, we will send you daily reminders to remember to take your pills. If you take two pills a day for three weeks, that's a total of 42 pills. There are 50 pills in the package, which means there are 8 pills too many. You don't have to send them back to us (you can take them at a later date, for example). During the three weeks please take always two pills per day and not more or less. I also ask you to fill out the announced surveys once a week until the exam. You will receive an e-mail with the link at the right time to make sure you remember to do this.	

Open-ended questions

Open-ended questions in open-label placebo group

1. What do you think about the idea of taking placebo pills? (open-ended question)
2. Do you find the placebo pill has generally helped you to be less anxious/stressed before the exam? Yes/No
3. For which symptoms did the placebo pill help to which extent 0% (the pill did not help at all) - 100% (the pill helped 100%)
   - Concerning excitement (emotional and physical tension)
   - Concern (thoughts about failure, self-doubt)
   - Regarding distraction (distraction from the task by irrelevant thoughts)
   - Regarding confidence (self-worth)
4. Did you assume that the placebo pills would work or were you skeptical? (open-ended question)
5. What do you think was in the placebo pills ? (open-ended question)
6. What did you learn by participating in this treatment study ? (open-ended question)
7. Do you have any other comments ? (open-ended question)

Open-ended questions in imaginary pill group

1. In general, how open are you to taking a pharmacological pill for your test anxiety ? 0% (not at all open) to 100% (very open) (slider).
2. Do you find the imaginary pill helped you to be less anxious/stressed before the exam ? Yes/No
3. For which symptoms did the imaginary pill help to what extent 0% (the pill did not help at all) - 100% (the pill helped 100%)
   - Regarding excitement (emotional and physical tension)
   - Regarding concerns (thoughts about failure, self-doubt)
   - Regarding distraction (distraction from the task by irrelevant thoughts)
   - Regarding confidence (self-worth)
4. How difficult was it for you to imagine the imaginary pill? 1 (very easy) - 7 (very difficult)
5. How well could you imagine the following aspects of the imaginary pill (0 - not at all well to 100 almost identical to a real pill):
   - Seeing the pill (visualization)
   - Tasting the pill
   - Feeling the pill
   - Effects of the pill
6. Did you find it easier to visualize and take the imaginary pill during the study?
   - Yes it was easier
   - It did not change
   - No it became more difficult
8. What do you think about the idea of taking an imaginary pill ? (open-ended question)
9. Did you assume that the imaginary pill would work or were you skeptical ? (open-ended question)
10. Did you learn anything from participating in this treatment study? If yes, what? (open-ended question)
11. Do you have any other comment? (open-ended question)

Open-ended questions in control group

1. Were you disappointed that you were in the control group? Yes/No
2. Is there anything else you would like to comment on? (open-ended question)

Supplementary Table

Table S1

Mean values for subscales of the test anxiety questionnaire for all assessed timepoints.

		T1	T2	T3	T4	T5
	Group (n)			*M (SD)*		
worry	IP (55)	13.27 (3.36)	12.36 (3.65)	12.18 (3.58)	12.36 (3.99)	10.42 (10.18)
	OLP (59)	14.42 (2.96)	13.85 (3.00)	13.00 (3.44)	13.66 (3.47)	10.86 (10.60)
	CG (59)	13.95 (3.50)	14.05 (2.75)	13.89 (3.48)	14.14 (3.65)	11.49 (9.97)
emotionality	IP (55)	11.00 (3.25)	9.69 (2.81)	9.24 (2.84)	9.80 (3.13)	9.58 (3.47)
	OLP (59)	12.08 (3.52)	10.97 (2.78)	10.07 (2.83)	10.24 (3.15)	10.14 (3.33)
	CG (59)	11.25 (2.89)	11.83 (3.39)	11.75 (2.88)	12.10 (3.84)	11.93 (3.99)
interference	IP (55)	10.42 (3.26)	10.29 (3.11)	9.36 (3.25)	9.38 (3.31)	6.85 (2.38)
	OLP (59)	11.34 (3.14)	10.66 (2.82)	9.95 (3.13)	10.20 (3.14)	7.20 (2.72)
	CG (59)	11.31 (3.14)	11.98 (3.17)	11.59 (3.40)	11.69 (3.27)	8.34 (2.91)
lack of confidence	IP (55)	14.16 (2.94)	12.91 (2.71)	13.05 (2.65)	13.15 (3.05)	13.00 (3.49)
	OLP (59)	14.51 (2.47)	13.83 (2.64)	13.68 (2.68)	13.47 (2.93)	13.44 (3.53)
	CG (59)	14.15 (2.32)	14.22 (2.09)	14.24 (2.74)	14.17 (2.83)	14.02 (2.96)

Note. ASS-SYM, Änderungssensitive Symptomliste (general well-being); CG, control group; IP, imaginary pill; M, mean; OLP, open-label placebo; PAF, Prüfungsangstfragebogen (test anxiety questionnaire); PSQI, pittsburgh sleep quality index; SD, standard deviation.

C. Study 3

Buergler, S., Sezer, D., Busch, A., Enzmann, M., Bakis, B., Locher, C., Bagge, N., Kirsch, I., Carvalho, C., & Gaab, J. (2023). A qualitative study on imaginary pills and open-label placebos in test anxiety. *PLOS ONE (in Review)*

A qualitative study of imaginary pills and open-label placebos in test anxiety

Sarah Buergler[1*], Dilan Sezer[1], Alexander Busch[1], Marlon Enzmann[1], Berfin Bakis[1], Cosima Locher[2], Niels Bagge[3], Irving Kirsch[4], Claudia Carvalho[5], & Jens Gaab[1]

[1] Division of Clinical Psychology and Psychotherapy, Faculty of Psychology, University of Basel, Basel, Switzerland

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

[3] Institute for Emotion-Focused Therapy, Roskilde, Denmark

[4] Program in Placebo Studies, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, United States

[5] Department of Clinical and Health Psychology, Instituto Superior de Psicologia Aplicada (ISPA), Lisbon, Portugal

*Corresponding author: Sarah Buergler

Email: sarah.buergler@unibas.ch

Abstract

Background: The efficacy of open-label placebos (OLPs) has been increasingly demonstrated and their use holds promise for applications compatible with basic ethical principles. Taking this concept one step further an imaginary pill (IP) intervention without the use of a physical pill was developed and tested in a randomized controlled trial (RCT). To explore participants' experiences and views, we conducted the first qualitative study in the field of IP.

Methods: A reflexive thematic analysis (RTA) of semi-structured interviews with test anxious students (*N*=20) was nested in an RCT investigating an IP and OLP intervention. In addition, open-ended questions from the RCT were evaluated (*N*=114) to corroborate the RTA, and pill characteristics were included to more accurately capture the IP experience.

Results: Four key themes were identified: (1) attitude towards the intervention, (2) applicability of the intervention, (3) experience of effects, (4) characteristics of the imagination. The IP intervention was well-accepted, easily applicable, and various effects, pill characteristics and appearances were described. While many participants did not desire a physical pill, either due to the absence of the imagination component or aversion to pills, the approach was considered to be cognitively and time demanding, which in turn, however, encouraged the establishment of a therapeutic ritual that protected against the increase in test anxiety during the preparation phase. OLP findings were comparable, and especially the importance of a treatment rationale was stressed in both groups, counteracting an initial ambivalent attitude. The RTA findings were supported by the open-ended questions of the RCT.

Conclusion: IPs appear to be a well-accepted and easily applicable intervention producing a variety of beneficial effects. Thus, the IP approach might serve as an imaginary based alternative to OLPs warranting further investigations on its application to harness placebo effects without a physical pill.

Introduction

Traditionally, placebos are tied to randomized controlled trials (RCTs) as a concealed and inert treatment (sugar pill, saline injection) to isolate the specific efficacy of a verum by controlling for therapeutic noise, e.g., expectancy, spontaneous remission, or regression to the mean [1]. However, placebos have moved beyond being a mere methodological tool. The placebo effect has been established as a means to gain positive therapeutic outcomes in clinical trials [1], and its utilization is widespread among doctors and medical specialists [2-4]. Nonetheless, the deceptive nature of the placebo treatment is controversial and faces multiple ethical hurdles [5]: It threatens patients' autonomy, practitioners' obligatory veracity, and finally patient's trust in the therapeutic relationship [6]. In this context, the prescription of open-label placebos (OLPs; [7]) holds the promise of a treatment that harnesses placebo effects in a transparent and ethical way [8]. The clinical potential of placebos and specifically OLPs is evident and further research on their beneficial effect is being called for [9, 10].

Recent meta-analyses found medium to large effect sizes for OLP in clinical conditions [9] and a medium sized effect in self-reported outcomes in experimental nonclinical conditions [11]. Thus, placebos also seem to work when given without deception. Yet, what contributes to the positive effects of an OLP – consisting of a pill, a rationale, and a therapeutic interaction [8] – remains unclear. Through the elimination of the physical pill, it is possible to examine whether the pill is a necessary factor in producing placebo effects.

The concept of an imaginary pill (IP) was first introduced by De Shazer in 1984 in the context of clinical hypnosis [12] and more recently, Niels Bagge, a Danish clinician, independently introduced the same idea based on the concepts of placebo research [13]. To test the efficacy, we examined the effects of an IP intervention and compared it to OLP in an RCT in healthy students with test anxiety. We found comparable positive effects for IP and OLP with a medium to large effect size (d = 0.71) in comparison to a control group [14].

However, apart from intervention efficacy, the promising findings on IP warrant qualitative research to include participants' perspectives regarding the interventions' acceptability (i.e., credibility, trust, and belief), applicability and the experience of effects of the IP. In the field of OLP research, several qualitative studies have already been conducted wherein it was found that participants have a mixed reception to the OLP treatment idea, with reactions ranging from skeptical to curious and hopeful [15-21]. For instance, irritable bowel patients in the OLP group were more ambivalent and self-reflective as compared to those in the deceptive placebo group [15], whereas women with menopausal hot flushes had an overall positive experience with the OLP treatment [21]. In contrast, participants in an experimental pain study were rather skeptical about the efficacy of the intervention, despite beneficial treatment effects [20]. This might be explained by a lack of trust in the competence of the providing health care professionals, as well as perceived self-efficacy in solving a problem, as Druart and colleagues' qualitative study on experimental pain showed [22]. In postoperative pain, patients perceived an OLP both as trustworthy and ethical [23], but experiences and perceptions of the treatments' efficacy appear to vary widely [16]. A large range of reactions is also displayed by physicians reviewing the idea of OLPs [4]. Hence, an additional qualitative study in this research field may expand the current database and include participants' experiences of a placebo intervention without a physical pill.

The aim of this embedded qualitative study was to generate initial knowledge of views and experiences toward the novel IP intervention and to compare these views with the experience of individuals who received OLPs (N = 20). Further, we sought to corroborate these findings with data from open-ended questions of the RCT (N = 114) and to provide insights into individual characteristics and appearances of imagined pills from RCT participants.

Methods

Study design

This qualitative study was embedded within an RCT testing the efficacy of IP and OLP interventions compared to a control group in reducing participants' test anxiety over three weeks [14]. In short, participants were randomly assigned to one of the three groups and received their intervention in an online treatment session. During this session, both intervention groups received a treatment explanation (i.e., rationale) and participants in the IP group practiced taking their IP together for the first time with the assistance of the treatment provider. Both intervention groups were instructed to apply their respective placebo procedure twice daily for three weeks. After study completion, randomly selected participants were contacted via email and provided with relevant information regarding the planned qualitative study. Upon receiving informed consent, semi-structured interviews were scheduled and held online (due to pandemic restrictions). The study design and informed consent was approved by The Ethics Committee of the Faculty of Psychology, University of Basel, Switzerland, and was carried out in accordance to the principles expressed in the Declaration of Helsinki. The study was registered at ClinicalTrials. gov: NCT04250571 (31/01/2020). Access to study data was limited to the study personnel and all identifying data was anonymized. This qualitative study was reported in accordance with the standards for reporting qualitative research (SRQR) checklist [24].

Study participants

Of the 114 participants in the intervention groups of the RCT, twenty participants (i.e., OLP n = 10 and IP n = 10) were randomly selected to take part in the qualitative study. In case participants either did not want to take part in the study or did not respond to the recruitment request, the next person on the randomized list was contacted until ten participants per group were reached. The sample consisted of healthy students at the University of Basel, who had an exam at least four weeks ahead. Being a master student, insufficient German skills and problems swallowing pills were reasons for exclusion. Incentive to take part in this nested qualitative study were credit points or a fixed monetary compensation (20 Swiss Francs). Recruitment took place between June 2020 and June 2021. Table 1 depicts sociodemographic characteristics of the twenty study participants.

Table 1. Sociodemographic characteristics of the study participants.

Group	N	Age (±SD)	N (%) female	Mean interview duration in minutes (±SD)	N (%) Psychology student
IP	10	24,5 (±5,5)	6 (60%)	42,2 (±9,3)	10 (100%)
OLP	10	22,9 (±5,2)	8 (80%)	37,1 (±7,5)	9 (90%)

Note. IP imaginary pill, *OLP* open-label placebo, *SD* standard deviation.

Interview procedure

The semi-structured interviews were conducted by SB. The platform used was zoom (https://zoom.us), which allowed the audio recording of the session. The interviews aimed at receiving a comprehensive insight into the subjective views and experiences about the interventions. It was stressed that there are no right or wrong answers and that participants can talk freely and criticism or suggestions for improvement of the intervention are welcome. In some cases, non-predefined questions were asked for comprehension issues and for an overall agreeable and open conversation atmosphere. The interviews included questions about the following study time points: (a) the treatment session, (b) the three-week intervention phase, (c) the exam situation, and (d) possible future use. Interview questions only varied slightly between IP and OLP participants and can be found as supporting information (S1 appendix: interview questions). The interviews were conducted in Swiss German and transcribed verbatim in the German language with preservation of typical Swiss German expressions, following an integrative approach [25]. Quotations in the results were translated from German to English.

Qualitative analysis

The interviews were exported to and analyzed with the software MAXQDA (https://www.maxqda.com). A reflexive thematic analysis (RTA) approach [26, 27] was employed. The number of interviews (*N*=20) was considered as sufficient according to practical guidelines [28] and allowed us to reach a desired 'saturation', where no new information was generated by more interviews. An inductive-deductive hybrid approach was used to analyze the qualitative data [29]: First, an inductive and data-driven coding process was used to map the content of the interviews for a phenomenon with limited research literature as accurately as possible [30-32], and in a next step, a more deductive approach was used to generate themes that were relevant and meaningful to the research question [33]. With the aim of developing a coding system and key themes across the interviews, the following steps were performed: (1) Prior to coding, two study team members (AB, ME) familiarized themselves with the dataset by reading through the interviews thoroughly. (2) Initial codes were generated by AB and ME. The codes (i.e., short segments of the interview conveying significant information about the topic) were chosen based on their perceived meaningfulness and

relatedness to the research question. To maintain characteristic features of IP and OLP interviews, two unique code systems were developed. (3) To minimize bias, 50% of the interviews were coded by two independent coders (AB, ME; [34]). Codes of the independent coders were regularly compared, and the code system was updated based on agreement. (4) A common consensus about the code system was found in regular group meetings, including SB and DS. (5) Codes were grouped by AB and ME into categories and these further into main categories. These steps were done based on inductive similarity and belonging of codes and deductive relatedness to the research question. (6) Based on the main categories, initial themes for both code systems were formed by ME. As a result of their congruence, both code systems (i.e., IP and OLP) were merged. (7) Upon further revision of the themes, the final system was created by AB and SB.

Qualitative analyses of open-ended questions from RCT and imaginary pill characteristics

In order to back the qualitative analyses, open-ended questions completed at the last assessment time point of the RCT were further included in this analysis (S2 appendix: open-ended questions of RCT). This sample included all OLP and IP participants of the trial (*N*=114). Results of two open-ended questions (i.e., "Why did you find the explanation that the IP/OLP can work helpful?"; "Did you assume that the IP/OLP would work or were you skeptical?") were included in the current analysis. Further analyses on open-ended questions, such as treatment credibility and intervention idea or learnings due to study participation, can be found in the supporting information (S1-S3 Tables: open-ended questions). All of the responses were imported to the online software MAXQDA2022 and a descriptive method of data analysis was used [35]. Data analysis was conducted by thematically coding participant's comments, without a thorough theme analysis [36]. Two coders (BB, CB) repeatedly read the comments to achieve familiarization and then used an inductive, open coding process in which descriptive labels (i.e., codes) were assigned to each comment. Both coders worked independently, and for comments that communicated numerous meanings, several codes were used. Next, codes were examined and compared to find commonalities and discrepancies. To offer a comprehensive overview of the comments, similar codes were sorted into higher level categories.

Since the IPs were in participants' minds and not visible to observers, we wanted to capture the inner experience and characteristics of the participants' pills including their appearances with graphical illustrations. To do so, all IP participants in the RCT had to complete an interactive document to describe the characteristics of their pill (type, shape, color, packaging, size) and its effects immediately after the treatment session. The frequencies of each feature and effect were summarized in a table. Nine IPs were selected and illustrated by BB using the software blender 3.2.2 (https://www.blender.org).

Results

Qualitative analysis

Of a total of 37 contacted, 17 (OLP n = 8 and IP n = 9) did not respond to the recruitment request or declined to participate. The twenty study participants who took part in the interview had a mean age of 23.7 (± 5.4) years, 70% of them were female and 95% psychology students (see Table 1). The interviews had an average length of 39.4 minutes (± 9.1).

The interviews consisted of 106'946 words in total. Based on this material, 816 codes were generated, of which 651 were used for the analysis. The codes were summarized into 36 categories, which in turn were subsumed to 14 main categories. These categories were integrated into 4 key themes: (1) attitude towards the intervention, (2) applicability of the intervention, (3) experience of effects and (4) characteristics of the imagination. Due to congruent content between the IP and OLP coding systems, theme 1 to 3 capture both interventions, while theme 4 relates only to the IP group. Table 2 presents a detailed presentation of the key themes and main categories, including all listed categories. In the following, key themes and corresponding main categories are presented with a selection of relevant quotations.

Table 2 Key themes with underlying categories.

Key themes	**Main categories**	**Categories**
Attitude towards the intervention	Acceptance	Approval of the intervention
		Open attitude towards the interventions
	Expectation	Expectations towards the interventions / How expectations were met?
	Skepticism	Doubts towards the interventions
		Attitude towards IP from the OLP group
		Importance of the rationale for trust in OLP
		Importance of the treatment session (e.g., pill intake) for trust in IP
Applicability of the intervention	Preferences of pill intake	Schedule of the pill intake
		Ritualization of the pill intake
		Desire to individualize the pill intake
		Desire for physical placebo
	Integration of the pill in daily life	Difficulty to remembering pill intake in daily life
		Easy integration of pill intake into daily life
		Difficult integration of pill intake into daily life

		Reminder emails help with integration into daily life
		Intervention requires cognitive effort
	Application over time	Pill intake was more difficult over time
		Pill intake was easier over time
		Motivation-changes over time
		Refresher-appointment would be helpful
Experience of effects	Increased self-confidence	Increased self-awareness/-efficacy
		Increased self-confidence
	Stress relief	Stress reduction and increased relaxation
	Increased concentration	Increased mental focus/concentration
	Reaction to IP effect	Astonished that IP works *[only coded in IP group]*
	Modulating factors on the effect	Positive influence of a routine
		Effect dependent on the daily context
		Effect dependent on exam proximity
		Effect-changes over time
	Effects on the exam	Positive effect on learning
		Positive effect on examination situation
		Uncertainty about the effect on examination situation
		No effect on examination situation
Characteristics of the imagination	Individual aspects of imagination	Individual aspects that supported the imagination
	Realness of imagination	Body sensations during imagination
		Vivid imagination

Theme 1 Attitude towards the intervention

Acceptance

The IP intervention found wide acceptance among participants and various reasons were expressed why this treatment felt right for them. The imagined pill intake was described as easy and effortless, taking only a few minutes. It was viewed as a new and interesting idea, practical and flexible in its use. Also, participants expressed that the IP intervention was helpful and efficacious, and one participant preferred the IP over OLP due to reluctance to take real pills. The level of approval with the assigned intervention was also high among OLP participants. Example responses include:

> ” […] actually, I was amazed that you can have such an impact in this short amount of time.” (IP; Subject 14)

> “No, so for me [taking a physical pill] would have been bad. […] So, I liked it much more [than OLP] since it was imaginary.” (IP; Subject 17)

> “I kind of always looked forward to it, it was always like: ‘Uh, you still have to take your placebos.’” (OLP; Subject 1)

The pills (imaginary and physical) were associated multiple times with a drug or a symbol for healing or health improvement. It was emphasized that people are socialized with pills in their culture as a mean to get better and one participant reflected on the common familiarity with them:

> “[…] yes, the [imaginary] pill strongly resembles a drug, and I would say that most of the people have already taken a drug in their lives. So, it’s a good choice, a good method.” (IP; Subject 15)

Many approached the IP intervention with an open attitude. Participants showed a mixture between curiosity and excitement about the previously unknown approach. Some employed a mindset of ‘let’s give it a try’, as there were no costs or harmful side effects involved.

Contrary to IP, half of OLP interviewees had pre-existing familiarity with (open-label) placebos. Regardless of prior unfamiliarity, however, there was an openness and curiosity about the intervention:

> “I have heard about placebo before – the placebo effect – and that studies are done about it. But I didn't know that you can also do that with open placebos, that you can disclose that. I had never heard of that before. I actually found that exciting, that it is being tried out like that.” (OLP; Subject 9)

Approval towards the IP intervention was also apparent as the majority of participants showed openness to reapplying the intervention for the next exam phase in the same or a slightly altered way. Moreover, nearly all interviewees reported that they could imagine an extension of the IP intervention beyond test anxiety. Especially regulation of other emotions (insecurity, sorrow) as well as pain were regularly mentioned to be possible areas of application. Most participants would furthermore recommend the intervention to others, however, only to people that show open-mindedness:

> “To certain people in my environment, I would definitely recommend it, yes, others, definitively not, because, well, I think that it is something that you really have to make up your mind to do it and believe in it.” (IP; Subject 16)

These results are consistent with those of the OLP group, in which open-mindedness was also considered a prerequisite for the application of an OLP intervention.

Expectation

Most participants' expectations of IP were met. Seven participants reported that their expectations were fulfilled, with three of them claiming it was exceeded. Some, however, reported initial skepticism and therefore lower expectations, which were then easier to be met:

> “I would say, exceeded, yes. Well, simply because of the skepticism that I had in the beginning. Or I underestimated that it can really work, maybe it is better to say it like that.” (IP; Subject 16)

Similarly, almost all OLP participants reported fulfilled expectations. Again, initially low expectations were existent in this group:

> “I must say almost more than I would have thought. [...] Somehow, I had the feeling my anxiety is too big that it could have a big effect. But in the end, I think it helped me and gave me a certain security, also in myself. And from that point of view, I would say it was more than I expected initially.” (OLP; Subject 8)

Skepticism

Many participants in the IP group reported that they felt skeptical or doubtful about the intervention, especially in the beginning. While some participants expressed insecurities towards an unknown method, others mentioned doubts regarding the required resources (e.g., time, cognitive capacities). Comparably many OLP participants expressed initial skepticism.

When asked about the other group's intervention, a majority of IP participants preferred taking a pill the imaginative way. A physical component was not wished for, either due to the absence of the imagination component (which they liked), or because of a reluctance to taking real pills. Further, two IP participants expressed that the imagination was satisfactory and helpful, thus not wishing anything to be different:

> "So, it did help me, the imaginary pill, so I don't know if a real placebo would have helped more than this." (IP; Subject 18)

> [about whether OLP would have been better] ”I don't think so, because then the imagination would have disappeared. I don't know what would have happened if I had taken a Tic Tac [...] but ehm, I think the pure imagination that I needed for it, would then have vanished. And that's why I found it really cool that it is all about imagination, yeah.” (IP; Subject 16)

However, feedback from six IP participants indicated that the imaginative method is cognitively taxing and requires concentration and calmness. Three of those interviewees thus expressed that, even though the IP worked for them, taking a physical pill might have been easier and less tiring. On the other hand, eight participants in the OLP group expressed openness and interest towards an IP intervention, yet, assumed

that an imagination might be more challenging. In this context, three participants also stated that they preferred taking a physical pill.

> [about taking an IP] “Yes, I mean the critical part in me says: [...] ‘I need a physical correlation here, something that someone sends me, what someone has packed, what is produced [...] It would probably be more difficult for me to believe in it.’ [...] And then there's another part in me saying: ‘Mind over matter.’ [...] Just because you can touch it [...] doesn't really make any difference." (OLP; Subject 1)

The treatment session where participants received the intervention reduced skepticism and strengthened the trust in the intervention. Almost all IP participants reported that – besides the explanation – the joint IP intake was helpful, fostered trust, and countered their initial skepticism. The comments of the IP participants showed that the single practice with the provider within the RCT was enough to gain security to continue the practice independently.

> “This exercise that was done together with the pill taking was kind of really cool, because I had this feeling of ‘huh that really works?’ [...] I wouldn't have expected that before.” (IP; Subject 12)

Along these lines, almost all OLP interviewees were convinced by the rationale, which helped to reduce skepticism at the beginning and during the intervention. The most memorable and convincing discussion points in the treatment rationale were that the effects can occur even with a skeptical attitude (n=5) and that a possible mechanism is assumed to be conditioning (n=2), which can be explained by an automatic biological mechanism (n=1).

> “So, the biological [discussion point] made sense for me, with the neurotransmitters, I think that's the one that kind of sold me the most and that's also the one that I thought about the most during the study as well.” (OLP; Subject 5)

> “I believe that an additional thing was that it was said that placebos also work, even if you don't believe in it. Simply taking the pills can help. And I think that was something that was most notably decisive.” (OLP; Subject 6)

Theme 2 Applicability of the intervention

Preferences of pill intake

An external structure was provided to the participants for the pill intake, that comprised daily reminders in the morning and evening, encouraging a fixed intake schedule. The majority of IP participants preferred this method over a variable pill intake schedule (i.e., on demand intake). Mainly because a fixed schedule supported ritualization, reduced forgetfulness, and prevented negative feelings (stress, nervousness). Only

a few participants voiced a preference for a pill on demand. The voiced preferences were similar in the OLP group.

> “I also don't think that you can build up a routine just like that for a few days. Instead, I think it takes a little longer [...]. I am rather of the opinion that you don't use it in an acute way, but rather for a longer period of time. Maybe that it just becomes like a part of life itself.” (IP; Subject 15)

Integration of the pill in daily life

Around half of the participants in both groups combined the pill intake with a form of ritualization. This was either an integration into an existing ritual (morning routine, breakfast, etc.), a link with a suitable situation (start of learning session, break during learning) or simply with drinking a glass of water. The latter was even present in the IP group, where three participants either drank a glass of water or performed swallowing motions when taking their IP, supporting the imagination:

> “I actually had to make a swallowing motion to make it really go down in my head. Well, you were allowed to do that, and I didn't take anything else. I had to do something so that I felt that it's really something real.” (IP; Subject 20)

Three out of four IP participants who did not integrate the pill intake into a ritual reported that it might have been helpful to do so. One of them explained that the intervention would have been easier if repeatedly done in the same context, one that is preferably quiet. Another IP participant, who embedded the IP intake within a ritual, reinforced the need for such a ritual:

> "It's supposed to be a kind of ritual. If you just swallow a pill, it works for maybe three seconds, or not even, I don't know. And from my point of view, imagination is something that, if you don't do it often and practice it, is just [...] it needs a calm mood. You can't just say: ‘Now I'm at the train station and pop this imaginary pill’, it has to be an environment. It takes a lot more time, I think, for you to really feel this effect. […] It needs such a ritual." (IP; Subject 19)

Various reasons were conveyed about challenges to integrate the IP into daily life. Also, a few participants sometimes forgot to take their IP. These difficulties were also mentioned in the OLP group. However, the need for cognitive effort while taking the pill was solely emphasized in the IP group.

Application over time

The time course affected not only the experienced effect of the IP but also its application. For instance, a participant shared that motivation for the pill intake dropped after some time due to a lower level of suffering. However, motivation for applying the intervention also increased when it was experienced as helpful. In addition, five participants in the IP group expressed a wish for a so-called ‘refresher-appointment’ (i.e., a second meeting in the middle of the intervention phase that would renew the placebo information and the

IP exercise). They explained that this would be helpful for remembering the rationale, to consolidate and remember the desired feeling connected to the IP, and to generally increase motivation. Besides the interactive document in which participants recorded their IP, a few participants suggested ideas for alternative or idiosyncratic reminders to recall specific placebo information. For instance, one IP interviewee proposed having an auditory aid:

> “But maybe also an idea would be, if you had that [the suggestions about the IP intake] as an audio recording or something, so instead of a refresher, on the phone or in person, you could just listen to that again.” (IP; Subject 14)

While the timing also affected the experienced OLP effect and its application, only two participants in the OLP group mentioned the idea of such an additional appointment.

Theme 3 Experience of effects

Increased self-confidence and concentration, stress relief and reaction to IP effect

The IP participants reported increased self-confidence and concentration as well as decreased stress as beneficial effects during the 3-week intervention phase. Participants who reported an increase in self-confidence experienced more optimism or could draw strength and courage and thus felt supported through the pill. Others reported a feeling of security and saw the intervention similarly to an anchor that prevented the emergence of more negative feelings:

> "So, it felt less that it had an effect, but more that it actually prevented other things from being triggered. [...], that it just didn't develop as strongly. So, because the stress with me is rather high, the closer the exam comes. And the feeling didn't really come at all." (IP; Subject 13)

Stress relief was among the most prominent positive IP effects leading to increased relaxation and calmness, as well as reduction of anxiety and stress. This was felt on a psychological (e.g., less nervous thoughts) and physical level (e.g., general relaxation and loosening, reduction of shaking). At the same time, an increase of concentration during the learning phase was reported. Here, clearer thoughts, wakefulness, better focus and increased inner calmness were described by roughly half the participants. The same kind of effects were mentioned in the OLP group.

Besides the mentioning of the beneficial effects, eight participants of the IP group were positively surprised, astonished, or impressed by those effects. These reactions were usually connected to the initial skepticism and novelty of the intervention, and the unfamiliar realization that the mind can have such an effect on the body:

> “So that placebos work, I already knew that before. But I never tried it myself, because I never took placebos myself and then I was positively surprised that it can work so strongly, that what the mind – stupidly said – can simply do.” (IP; Subject 16)

Modulating factors on the effect

The strength of the IP’s effect varied based on different factors: Most importantly, daily fluctuations in sleep, stress levels, and fatigue, or even simply the time of day sometimes made pill intakes more difficult and, as a consequence, the pill effect less strong.

> “This [strength of the effect] actually depended on the day. I think when [...] I had a stressful day or so, then it was difficult for me to sit down and concentrate [...]. Yes, so I just think that it depends very much on the day.” (IP; Subject 12)

> “It depended a little bit on how I had the space, how I took my time and how much I could focus. And if, for example, I left the house and was doing something and then: ‘Oh well, I still have to take the pill!’ then I took it and then it [the effect] came much less strongly, but if you are almost meditative, or you really take time to experience the ritual completely, then it has a much stronger effect.” (IP; Subject 19)

OLP participants reported that the exam proximity changed the experience of effects (see also *application over time* in theme 2). With the increase in pressure and nervousness due to the approaching exam, four respondents showed enhanced motivation for the pill intake and increased effects of the OLP:

> "As the exam was approaching, you became more nervous and you needed something that would help you [..] so it was really like that, the need increased and the effect also went up." (OLP; Subject 2).

On the other hand (and similar to IP participants who reported the need for a refresher appointment as the effect diminished after the treatment session) some OLP participants reported a decrease in effect due to an increase in skepticism, which in turn negatively impacted motivation:

> "Well, at first I believed in it, at the beginning, let's say the first week, I was kind of motivated to do it, ... it's just that afterwards I believed in it less, I became more and more skeptical, I thought about it more, yes." (OLP; Subject 5)

Effects on the exam

Participants had mixed perspectives on the pills’ effect during the exam situation. In the IP group, nine participants were either unsure if the pill had a positive effect on the exam itself or thought it might have been subliminal. Some felt less nervous and more calm compared to previous exam situations, but could not certainly attribute it to the pill. Half of the IP participants expressed that they rather felt a helpful and buffering effect during the learning phase but not during the exam.

Similarly, five OLP interviewees thought that the intervention helped indirectly for the exam by facilitating the learning phase. For the other five participants, the OLP induced noticeable effects during or right before the examination that ranged from better mood, increased calmness to reduced anxiety.

Theme 4 Characteristics of the imagination

Individual aspects of imagination, realness of imagination

All IP respondents described a vivid and detailed imagination, which was maintained during the intervention phase. Concerning the pill, the majority of participants imagined characteristics of the pill that resembled real pills, including form, size, color, and taste. Participants were also able to regularly reproduce the desired state connected to the pill:

> “I found it impressive how I was able to revive this feeling from this situation that I imagined [during the treatment session]. That I truly felt it for real and that it continued to work later when I took the imaginary pills [by myself].” (IP; Subject 14)

> "It was like a mental picture, but I could really picture it quite well, so like the color, shape, size and then also the taste in the mouth and also the physical effect, so really quite quite well." (IP; Subject 19)

Five respondents found it helpful to link the desired effect of the IP with the positive situation identified during the treatment session. For example, some participants focused on specific situations, such as playing an instrument (e.g., ‘when I take the IP, I want to feel the way I feel when I play the cello') or more complex imaginary scenarios, by being mindful and focused at present:

> "So, I always imagined it like this, as if I would now go from the desk in my room to the kitchen and open the drawer, take out the pill, have it in my hand, put it in my mouth and then take a sip of water. I really imagined myself doing the movement but just stayed at my desk. [...] Each time I could say to myself ‘I'm going to take a little round pill that's bright yellow. It has no smell, no taste’. I think that helped me to make this mental journey, to imagine the whole way to the kitchen every time, every step." (IP; Subject 18)

Remarkably, some participants had physical sensations in their bodies during their IP intake. Based on suggestions during the exercise in the treatment session, one participant felt the imagined pill sliding down her throat. This sensation reinforced the trust in the pill for the interviewee, who was surprised that an imagination can induce such a bodily feeling. Other participants reported experiencing side effects such as a dry mouth, goosebumps or warmth radiating from the abdomen. These side effects were suggested during the pill intake in the treatment session to demonstrate the IP response and were in fact found to be

helpful in convincing people of the pill's efficacy. Overall, realistic and detailed imagination were regularly reproduced and maintained during the three weeks of intervention:

> "[...] when I thought of it, it didn't take a long time for the image to appear, but it appeared like when I would think of a piece of paper or of a pen, as if I had already had it for real or seen it before [....]. So, if I would draw it now, I'm not good at drawing, but if I had to draw it, I would be able to." (IP; Subject 11)

Qualitative analyses of open-ended questions from RCT and imaginary pill characteristics

To corroborate the RTA, two open-ended questions from the RCT were analyzed qualitatively. A total of 221 responses to the two questions were received from 114 participants, with the responses typically being brief comments or phrases. Participants' characteristics can be found in the main study [14]. In both groups, the most commonly cited reasons why the treatment explanations were helpful were that 'the explanations made sense', 'led to a better understanding', 'created belief', and 'strengthened previous knowledge/belief' (see table 3). Table 4 depicts the responses to the question concerning skepticism towards the efficacy of the received intervention: While 'belief in the effect' was frequently expressed, a comparable number of statements included skepticism.

Table 5 shows the characteristics and effects of the IPs recorded in the interactive document after the treatment session: A wide range of different pill shapes, colors and packaging and particularly various effects were imagined. The most common type was a small, white pill with a round shape packed in a blister. However, more distinctive shapes (i.e., star-shaped), colors (i.e., colorful), and packaging (i.e., tins) were also reported. The most prevalent effect of the IP was 'relaxation' followed by 'focus', 'confidence' and 'better mood'. Visualizations of the pills can be seen in Fig 1.

Table 3 Helpfulness of explanation: Why did you find the explanation that the imaginary pill / open-label placebo helpful?

	OLP (N=82)	IP (N=77)	Total N=159*
The explanation…	N (%)		
… created faith	**10 (12.2)**	**8 (10.39)**	**18 (11.32)**
… gave new knowledge	7 (8.45)	3 (3.9)	10 (6.29)
… led to better understanding	**13 (15.85)**	**9 (11.69)**	**22 (13.84)**
… was helpful	5 (6.1)	1 (1.3)	6 (3.77)
… was believable	3 (3.66)	2 (2.6)	5 (3.14)
… made imagination easier	5 (6.1)	4 (5.19)	9 (5.66)
… strengthened previous knowledge/beliefs	**7 (8.45)**	**9 (11.69)**	**16 (10.06)**
… mentioned previous studies	5 (6.1)	7 (9.09)	12 (7.55)

… was conforming to personal interest	5 (6.1)	0 (0.0)	5 (3.14)
… made sense	**8 (9.76)**	**15 (19.48)**	**23 (14.47)**
… created an expectation	4 (4.88)	2 (2.6)	6 (3.77)
… focused on positive aspects	2 (2.44)	1 (1.3)	3 (1.89)
… gave security	1 (1.22)	0 (0.0)	1 (0.63)
… was credible	2 (2.44)	0 (0.0)	2 (1.26)
… left an open outcome	1 (1.22)	1 (1.3)	2 (1.26)
… showed researcher allegiance	0 (0.0)	1 (1.3)	1 (0.63)
… improved mindfulness	3 (3.66)	6 (7.79)	9 (5.66)
N/A	1 (1.22)	8 (0.39)	9 (5.66)

Note. Statements in bold are those that were mentioned by more than 10% of the respondents in total. *N (%) indicates the number (percent) of participants who mentioned the respective topic in their answer. Since several answers were coded per question, one person can be included several times in the data. *IP* imaginary pill, *OLP* open-label placebo.

Table 4 Skepticism towards treatment: Did you assume that the imaginary pill / open-label placebo would work or were you skeptical?

	OLP (N=71)	IP (N=71)	Total N=142*
	N (%)		
not skeptical	0 (0)	3 (4.2)	3 (2.1)
believed in effect	**20 (28.2)**	**28 (39.4)**	**48 (33.8)**
open minded	8 (11.3)	7 (9.9)	15 (10.6)
hopeful	4 (5.6)	2 (2.8)	6 (4.3)
curious	1 (1.4)	0 (0)	1 (0.7)
neutral	2 (2.8)	0 (0)	2 (1.4)
unsure	1 (1.4)	0 (0)	1 (0.7)
slightly skeptical	1 (1.4)	7 (9.86)	8 (5.6)
skeptical at first	2 (2.8)	2 (2.8)	4 (2.8)
grew more skeptical over time	1 (1.4)	1 (1.4)	2 (1.4)
skeptical	**30 (42.3)**	**21 (29.6)**	**51 (35.9)**
very skeptical	1 (1.4)	0 (0)	1 (0.7)

Note. Statements in bold are those that were mentioned by more than 30% of the respondents in total; *N (%) indicates the number (percent) of participants who mentioned the respective topic in their answer. Since several answers were coded per question, one person can be included several times in the data. *IP* imaginary pill, *OLP* open-label placebo.

Table 5 Overview of the characteristics of the pills in the imaginary pill group and their frequencies.

Kind	**N**	**Effects**	**N**
Pill	39	Relaxation	49
Capsule	10	Focus	34
Lozenge	5	Confidence	14
Shape		Better mood	11
Round	32	Alertness	5
Oval	19	Motivation	5
Globule	1	Optimism	4
Star shaped	1	No tremor	4
Octagon	1	Control	4
Color		Stops rumination	4
White	20	Lowering pulse	3
Blue	10	Security	2
Colorful	10	Soothes intestinal problems	2
Green	4	Good sleep	2
Pink	3	Feeling of freedom	2
Purple	2	Positive stress	2
Red	2	Energy	2
Yellow	2	Competence	1
Grey	1	Light chest	1
Packaging		Connected to nature	1
Blister	24	Narcotizing	1
Glass	11	Heightened blood flow	1
Tin	10	Less stress	1
Box	7	No sweating	1
No package	1		
Sachet	1		
Size			
Small	29		
Medium	18		
Very small	4		
Big	3		

Note. Features might be mentioned several times.

Fig 1. Visualization of a selection of imaginary pills. Imaginary pills of nine study participants and their packaging (if existing).

Discussion

This qualitative study – nested within a 3-week RCT in test anxious students – sought out to assess participants' views and experiences towards a novel IP intervention and to compare them to those of participants receiving OLPs. Further, we aimed to corroborate these findings with data from open-ended questions of the RCT and to give insights into individual characteristics and appearances of imaginations from study participants.

Four key themes were identified: (1) attitudes towards the intervention, (2) applicability of the intervention, (3) experience of effects, (4) characteristics of the imagination (see table 2). Regarding the attitude (theme 1) and applicability (theme 2) of the IP intervention, most interviewees showed a high level of acceptance and mentioned its easy application with a minimal effort required to achieve a positive effect. While a few, however, expressed a preference for a physical pill because imagination requires more time and cognitive effort, the majority of participants did not desire a physical component but were intrigued by the 'power' of imagination. For a few participants, taking a physical pill might have caused the imagination to disappear. Participants approached the IP intervention with an open attitude and with a mind-set of 'let's give it a try' and 'it won't harm me if it doesn't work', the latter underlining the advantage of such placebo interventions having no known side effects. Most interviewees showed openness to reapplying the IP intervention, to extending its use to other symptoms (i.e., pain, sorrow, migraine), and recommending it to relatives and

friends with open mindsets – a finding that shows that the IP offers flexibility and empowerment of users. Initial skepticism was commonly present in the IP group but seemed to decrease for some participants over the course of time. The results from the open-ended questions were consistent with this ambivalent finding: Many respondents in both groups reported being skeptical, while just as many also believed in the effect, with many expressing both (see table 4). The treatment session seemed to be crucial to counteract skepticism: Besides the joint intake of the IP, the treatment rationale was perceived as convincing. This finding is supported by the open-ended questions, which revealed that the explanations led to better understanding, strengthened previous knowledge, made sense, and created faith in the intervention (see table 3). In terms of IP application, the intake was often ritualized and/or embedded in a pre-existing ritual and a preference for a fixed intake-schedule was generally voiced. This preference supports the choice of a fixed schedule over a variable one and reflects the use of many (pain) medications which can be taken on a regular basis and/or on demand. Thus, the understanding of IP also seems to align with normal pill prescribing. The symbol of an (imaginary) pill itself elicited familiarity among participants, as it was described that pill-taking is embedded in their culture. Further, the IP intervention was preferred by a participant with aversion to pills.

Regarding the experience of effects (theme 3) and characteristics (theme 4) of the IPs, a wide range of beneficial effects as well as various individual pill properties were reported. Mostly, the IP intervention helped with stress relief and increased participants' concentration and self-confidence. The experienced effects matched those, that participants recorded in the interactive document after the treatment session (see table 5). The IP approach further seemed to serve a preventive function during the learning phase, partly by providing a soothing and sustaining daily ritual, some of which allegedly prevented the onset of severe symptoms altogether. The quality and intensity of the IP effects varied, as factors such as daily context, exam proximity or treatment session proximity influenced the pill-intake. However, imagining the pill was perceived as easy and participants could continuously maintain the visual image of the pill which they had imagined at the treatment session. With respect to the appearance of IPs, a wide range of different pills were described (e.g., shapes, colors and packaging; see table 5), with the most common pill being small, round, and white, packaged in a blister. This type of pill and packaging (i.e., blister) represents a very common form in Switzerland. While more distinctive shapes, colors, and packaging were also described (see fig 1 for a selection), it is noteworthy that the majority of IPs correspond to a conventional appearance of pills. This realistic appearance could, in turn, underpin the credibility and familiarity of the intervention.

Overall, the qualitative findings derived from the RTA could be confirmed with the open-ended questions from the RCT and the results from the IP group can be applied to those participants receiving OLPs: The OLP intervention was equally well received, described as easy to apply, and initially met with ambivalence (openness and skepticism), although initial skepticism was countered by the treatment rationale provided. Similarly, the reported effects of both groups were comparable. However, some intervention-specific

themes emerged in the qualitative interviews, particularly related to only one intervention: Imagining a pill was, for instance, considered more demanding and challenging by some participants compared to taking a physical pill, whereas imagination was also regarded as a central part of the intervention. Additionally, the IP intervention was described as a possible alternative for individuals who have an aversion to pills. In terms of effects, whereas both interventions were perceived as beneficial during the learning phase, solely OLP participants additionally reported to have felt a supportive effect during the exam. This finding suggests that OLPs may be more helpful in acute situations, whereas IPs may be more protective, counteracting the onset of symptoms in a preventive manner.

This is the first qualitative study on the use of IPs. However, qualitative research exists in the field of OLP: There, participants of OLP intervention arms rated OLPs to be acceptable and reported an overall positive experience [21] expressing curiosity towards this intervention [15]. Yet, participants also expressed ambivalent feelings: on the one side, they were open about taking an OLP, on the other side, skepticism coexisted with open attitudes [15,16,23]. Taken together, these results are congruent with experiences reported in the OLP group of the current study which further supports the finding that ambivalent attitudes and the cooccurrence of belief and skepticism are a common theme in OLP administration. It is noteworthy that the treatment rationales in this study addressed and appeared to mitigate skepticism successfully. Especially the explanation of underlying mechanisms (e.g., conditioning) as well as the information that an open attitude helps but is not necessary were most remembered and perceived as convincing. In line, a recent survey identified that mentioning of classical conditioning and brain mechanisms within the OLP rationale to be perceived as plausible [37]. However, these findings are contrasting the ones of Locher and colleagues [20], where participants seldom emphasized classical conditioning as a mechanism of placebo effects. Instead, other factors such as general attitude and beliefs were voiced. Nevertheless, mindsets such as ‘I have nothing to lose’, ‘let’s give it a try’ or ‘it won’t harm me if it doesn’t work’ (a Swiss German saying) were in this study equally used to approach an OLP as was the case within the study of Locher and colleagues [20]. These mindsets are related to open-mindedness and the idea of hope. Hope is a factor that has already been suggested to be important in OLP effects [19], as patients, who had no previous success with medication for their symptoms, could adopt a try-out-attitude of ‘what if it helps?’ [15]. This heuristic of ‘losing little if it doesn’t work, gaining a lot if it does’, that seems to be used frequently, could be an additional rationale-perspective next to ‘an open attitude helps, but is not necessary’. Overall, the current findings highlight the importance of the treatment rationale, which is not only important for the building of expectations (thus, leading to higher effects) – studied in several quantitative OLP trials [15,38-41] – but is also important to counteract skepticism. Our findings further underline the need to adapt the proposed explanation for future studies, possibly with new and updated OLP mechanisms that resonate with participants beliefs [20].

Whereas the treatment rationale seems to be decisive in order to produce placebo effects, the component of a physical pill might apparently not be. IP and OLP are, however, not only equally efficacious (difference n.s. d=0.11; [14]) but, as this qualitative analysis shows, both interventions are similarly well-accepted and easy to apply. Yet, IP offers an advantage over OLP in that it does not require a physical pill to be taken, reducing cost, facilitating accessibility and increasing customization, thus providing flexibility and empowerment to users. This well-accepted approach could therefore serve, for example, as a viable solution for health care providers faced with unclear regulatory requirements or lack of guidelines regarding the administration of OLPs, an issue raised by US physicians [4]. Besides that, IP can be an appropriate intervention for individuals who have difficulties taking pills, a particularly underestimated issue that affects many individuals [42], some of whom not only resort to non-adherence but also alter their medication regimen, potentially compromising safety and efficacy [43]. Nevertheless, the intake of IPs can be demanding, requiring time, rest, and resources. Consequently, this approach may be more appropriate for a clientele that has mental and physical capacities and that is not too impaired by symptoms. The potentially challenging and time-consuming part of imagination, however, was also viewed as a central part of the IP approach and thus crucial for producing beneficial effects. Therefore, the higher demands in IP have both a negative and a positive aspect: on the one hand, IP has its costs (i.e., time, resources), on the other hand, the imaginative ritual and the small break it creates in everyday life can have a positive effect in and of itself. Thus, both the moralizing treatment explanation and the therapeutic ritual may be central to the efficacy of IP [44]. Whether these results would be reproducible at all and in a sample with fewer available resources (e.g., in an inpatient setting) – not only in terms of efficacy but also acceptability and applicability – should be investigated in future research.

Limitations

The current study is subject to certain limitations. First, a sample consisting of healthy, young, female, and academic participants limits the generalizability of the findings. Higher education is related to more placebo knowledge [37] and may lead to expectations influencing the placebo effect. Likewise, it is possible that only those participants who also had a positive experience with the intervention responded to the interview request, ruling out potential negative views (i.e., 17 out of 37 contacted did not respond/declined). In addition, psychology students may be more responsive to such procedures than other student groups. Second, the nature of the reflexive thematic analysis involves an inductive approach and thus the researchers' subjective interpretation of the data, which might lead to biases and inconsistencies. However, the double coding and the continuous exchange in the research group partly addressed those concerns. Third, the interviewer (SB), who was the principal investigator of the RCT and is conducting research on placebo effect, might have influenced participants' way of answering questions. However, it was stressed that skepticism was welcome and that there are no right or wrong answers and an overall agreeable and open conversation atmosphere was created. Nevertheless, a certain social desirability effect on the sides

of the interviewees might have changed some answers in favor of the treatments. Fourth, the qualitative results in this study were not compared with the quantitative findings of the RCT, i.e., it is unclear whether the subjective statements are consistent with data obtained from the RCT. Lastly, a sample size of twenty is considered small in quantitative research. However, in the case of a reflexive thematic analysis it can be considered as sufficient and it most likely reached the state of 'saturation' [45], enabling a detailed and meaningful analysis on participants' perspectives on IP and OLP treatments.

Conclusion

The present qualitative study is the first to provide insights into participants' views and experiences on a novel IP intervention against test anxiety. Overall, the IP intervention was well-accepted (theme 1), easy applicable (theme 2) and manifold effects experienced on a physical, mental and emotional level (theme 3) as well as characteristics of the imagination (theme 4) were described. Whereas some IP participants did not wish for a physical pill, either due to the absence of the central component of imagination within this intervention or because of a reluctance to real pills, others viewed the IP intake as cognitively and timely demanding, requiring concentration and calmness. This, on the other hand, promoted the therapeutic ritual, which was perceived as supportive and protective against the increase in test anxiety during the preparation phase. The OLP findings were in principle comparable to those of the IP group regarding overall acceptance, application and effects. While IP effects, however, were mainly perceived during the learning phase, OLP effects increasingly appeared during the exam itself. In both groups, initial openness to the intervention was often accompanied by a certain amount of skepticism, while the treatment rationale was seen as important to counteract this ambivalence and build trust in these (new) interventions. Hence, the IP intervention seems to be an accepted, easy to apply and cost-effective intervention and might serve as an imaginary based alternative to OLPs. This warrants further investigations on IP application in both anxiety related domains, as well as for other clinical and nonclinical conditions to harness placebo effects without a physical pill.

Additional statements

Acknowledgments: The authors would like to thank Clara Bruhin (CB) for her assistance with analyzing the open-ended questions and assistance in conducting the qualitative study. Furthermore, the authors thank the interns Sabrina Michel and Jana Borer, for aiding with transcribing and coding the material.

Author Contributions: Conceptualization: SB, DS, NB, IK, CL, CC, JG; Data analysis: AB, ME, SB, DS, BB; Investigation: SB; Project Administration: SB; Supervision: CL, JG, NB, IC, CC; Visualization: SB, BB; Writing–original draft preparation: AB, ME, SB, DS, BB; Writing–review & editing: CL, JG, CC, IK, NB.

References

1. Blease C, Annoni M. Overcoming disagreement: a roadmap for placebo studies. Biol Philos. 2019;34(2). doi: 10.1007/s10539-019-9671-5
2. Howick J, Bishop FL, Heneghan C, Wolstenholme J, Stevens S, Hobbs FDR, et al. Placebo Use in the United Kingdom: Results from a National Survey of Primary Care Practitioners. PLoS ONE. 2013;8(3):e58247. doi: 10.1371/journal.pone.0058247
3. Tilburt JC, Emanuel EJ, Kaptchuk TJ, Curlin FA, Miller FG. Prescribing "placebo treatments": results of national survey of US internists and rheumatologists. BMJ. 2008;337:a1938–8. doi: 10.1136/bmj.a1938
4. Bernstein MH, Locher C, Stewart-Ferrer S, Buergler S, DesRoches CM, Dossett ML, et al. Primary care providers' use of and attitudes towards placebos: An exploratory focus group study with US physicians. Br J Health Psychol. 2020;25(3):596–614. doi: 10.1111/bjhp.12429
5. Annoni M. The Ethics of Placebo Effects in Clinical Practice and Research. Int Rev Neurobiol. 2018;463–84. doi: 10.1016/bs.irn.2018.07.031
6. Gaab J, Blease C, Locher C, Gerger H. Go open: a plea for transparency in psychotherapy. Psychol Conscious Theory Res Pract. 2016;3:175–98. doi: 10.1037/cns0000063
7. Park, L. C., Covi, L. Nonblind placebo trial. Arch Gen Psychiatry. 1965;12(4), 336–345. doi: 10.1001/archpsyc.1965.01720340008002
8. Blease, C. R., Bernstein, M. H., Locher, C. Open-label placebo clinical trials: Is it the rationale, the interaction or the pill? BMJ Evid Based Med. 2020;25(5), 159–165. doi: 10.1136/bmjebm-2019-111209
9. von Wernsdorff M, Loef M, Tuschen-Caffier B, Schmidt S. Effects of open-label placebos in clinical trials: a systematic review and meta-analysis. Sci Rep. 2021;11(1). doi: 10.1038/s41598-021-83148-6
10. Enck P, Bingel U, Schedlowski M, Rief W. The placebo response in medicine: minimize, maximize or personalize? Nat Rev Drug Discov. 2013;12(3):191–204. doi: 10.1038/nrd3923
11. Spille, L., Fendel, J. C., Seuling, P., Göritz, A. S., Schmidt, S. Open-label placebos: A systematic review and meta-analysis of experimental studies with non-clinical samples. 2022 [Forthcoming]
12. de Shazer, S. The imaginary pill technique. Journal of Strategic and Systemic Therapies. 1984;3,30–34.
13. Bagge, N. The Imaginary Placebo Pill: Isolating Psychological Components of Placebo Pills for Treatment in Clinical Practice [abstract]. Second official conference of the Society for Interdisciplinary Placebo Studies (SIPS): 2019; Leiden, NL.
14. Buergler S, Sezer D, Bagge N, Kirsch I, Locher C, Carvalho C, et al. Imaginary pills and open-label placebos can reduce test anxiety by means of placebo mechanisms. Sci Rep. 2023;13(1):2624. doi: 10.1038/s41598-023-29624-7
15. Haas JW, Ongaro G, Jacobson E, Conboy LA, Nee J., Iturrino J et al. Patients' experiences treated with open-label placebo versus double-blind placebo: A mixed methods qualitative study. BMC Psychol. 2022;10(1),20. doi: 10.1186/s40359-022-00731-w
16. Hruschak V, Flowers KM, Patton M, Merchantz V, Schwartz E, Edwards R, et al. Experiences of Patients Taking Conditioned Open-Label Placebos for Reduction of Postoperative Pain and Opioid Exposure After Spine Surgery. Int J Behav Med. 2022. Available from: https://pubmed.ncbi.nlm.nih.gov/35915346/
17. Kaptchuk et al. 2009 Kaptchuk TJ, Shaw J, Kerr CE, Conboy LA, Kelley JM, Csordas TJ, et al. "Maybe I Made Up the Whole Thing": Placebos and Patients' Experiences in a Randomized Controlled Trial. Cult Med Psychiatry. 2009;33(3):382–411. doi: 10.1007/s11013-009-9141-7

18. Kaptchuk TJ, Miller FG. Open label placebo: can honestly prescribed placebos evoke meaningful therapeutic benefits? BMJ. 2018;k3889. doi: 10.1136/bmj.k3889
19. Kaptchuk TJ. Open-Label Placebo: Reflections on a Research Agenda. Perspect Biol Med. 2018;61(3), 311–34. doi: 10.1353/pbm.2018.0045
20. Locher C, Buergler S, Nascimento AF, Kost L, Blease C, Gaab J. Lay perspectives of the open-label placebo rationale: A qualitative study of participants in an experimental trial. BMJ Open. 2021;11(8), e053346. doi: 10.1136/bmjopen-2021-053346
21. Pan Y, Frank ML, Kaptchuk TJ, Nestoriuc Y. "Let's see what happens:"—Women's experiences of open-label placebo treatment for menopausal hot flushes in a randomized controlled trial. PLOS ONE. 2022;17(11):e0276499. doi: 10.1371/journal.pone.0276499
22. Druart L, Vauthrin O, Pinsault N, Locher C, Blease C. "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. Br J Health Psychol. 2022. Available from: https://pubmed.ncbi.nlm.nih.gov/36086862/
23. Bernstein MH, Fuchs N., Rosenfield M., Weiss AP, Blease C, Locher C et al. Treating Pain With Open-Label Placebos: A Qualitative Study With Post-Surgical Pain Patients. J Pain. 2021;22(11), 1518–1529. doi: 10.1016/j.jpain.2021.05.001
24. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51. doi: 10.1097/ACM.0000000000000388.
25. Kruse J, Schmieder C. Qualitative Interviewforschung: Ein integrativer Ansatz. 2. Auflage. Beltz Juventa; 2014
26. Braun V, Clarke V. Using Thematic Analysis in Psychology. Qual Res Psychol. 2006;3(2):77–101. doi: 10.1191/1478088706qp063oa
27. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health. 2019;11(4):589–97. doi: 10.1080/2159676X.2019.1628806
28. Braun V, Clarke V, Hayfield N, Terry G. Thematic Analysis. In P. Liamputtong (editor). Handbook of Research Methods in Health Social Sciences. Springer; 2019. pp. 843–860. doi: 10.1007/978-981-10-5251-4_103
29. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods. 2006;5:80–92.
30. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-88. doi: 10.1177/1049732305276687
31. Oakley A, Strange V, Bonell C, et al. Process evaluation in randomised controlled trials of complex interventions. BMJ. 2006;332:413–6. doi: 10.1136/bmj.332.7538.413
32. O'Cathain A. A practical guide to using qualitative research with randomized controlled trials. 1 ed. Oxford University Press, 2018.
33. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs. 2008;62:107–15. doi: 10.1111/j.1365-2648.2007.04569.x
34. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. SAGE Publications, Inc.; 2011
35. Sandelowski, M. Whatever happened to qualitative description? Res Nurs Health. 2000;23(4), 334–40. doi: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G
36. Marks DF, Yardley L. Research Methods for Clinical and Health Psychology. SAGE Publications, Inc; 2004
37. Smits RM, Veldhuijzen DS, Olde Hartman T, Peerdeman KJ, Van Vliet LM, Van Middendorp H, et al. Explaining placebo effects in an online survey study: Does "Pavlov" ring a bell?. PLOS ONE. 2021;16(3):e0247103. doi: 10.1371/journal.pone.0247103

38. Klinger R, Kothe R, Schmitz J, Kamping S, Flor H. Placebo effects of a sham opioid solution. PAIN. 2017 Oct;158(10):1893–902. doi: 10.1097/j.pain.0000000000000977
39. Leibowitz KA, Hardebeck EJ, Goyer JP, Crum AJ. The role of patient beliefs in open-label placebo effects. Health Psychol. 2019;38(7):613–22. doi: 10.1037/hea0000751
40. Locher C, Frey Nascimento A, Kirsch I, Kossowsky J, Meyer A, Gaab J. Is the rationale more important than deception? A randomized controlled trial of open-label placebo analgesia. Pain. 2017;158(12):2320–8. doi: 10.1097/j.pain.0000000000001012
41. Schaefer M, Denke C, Harke R, Olk N, Erkovan M, Enge S. Open-label placebos reduce test anxiety and improve self-management skills: A randomized-controlled trial. Sci Rep. 2019;9(1). doi: 10.1038/s41598-019-49466-6
42. Messerli M, Aschwanden R, Buslau M, Hersberger KE, Arnet I. Swallowing difficulties with medication intake assessed with a novel self-report questionnaire in patients with systemic sclerosis – a cross-sectional population study. Patient Preference Adherence. 2017;11:1687–99. doi: 10.2147/PPA.S142653
43. Schiele JT, Quinzler R, Klimm H-D, Pruszydlo MG, Haefeli WE. Difficulties swallowing solid oral dosage forms in a general practice population: prevalence, causes, and relationship to dosage forms. Eur J Clin Pharmacol. 2013;69(4):937–48. doi: 10.1007/s00228-012-1417-0
44. Frank JD. Psychotherapy — The Transformation of Meanings: Discussion Paper. J R Soc Med. 1986;79(6):341–6. doi: 10.1177/014107688607900611
45. Hennink MM, Kaiser BN, Marconi VC. Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? Qualitative Health Research. 2016;27(4):591–608. doi: 10.1177/1049732316665344

Supporting information

S1 Appendix. Interview questions.

S2 Appendix. Open-ended questions of RCT.

S1-S3 Tables. Open-ended questions.

Supporting information for

A qualitative study of imaginary pills and open-label placebos in test anxiety

Sarah Buergler, Dilan Sezer, Alexander Busch, Marlon Enzmann, Berfin Bakis, Cosima Locher, Niels Bagge, Irving Kirsch, Claudia Carvalho, & Jens Gaab

This file includes:

S1 Appendix: Interview questions

S2 Appendix: Open-ended questions of RCT

S1-S3 Tables: Open-ended questions

S1 Appendix: Interview questions

OLP group

---

1. Intervention [Treatment 3 weeks prior to the exam]

Teaser

Okay, then I'd like to start by asking you a few questions about the intervention three weeks before the exam with *** [person who made the video call]:

- What do you remember most about it?

Rationale

We gave you an explanation why placebos may also have an effect if administered openly.

- Can you remember the explanation? Can you tell me what you remember? [Complete if necessary: Discussion points to OLP etc.]
- Does this explanation make sense to you?
- What about the explanation made sense to you, what didn't?

2. Intervention phase [Treatment during 3 weeks]

The next questions refer to the three-week period before the exam during which you took the placebos.

Pill intake

- Did you find it easy to take the placebos? Tell me about it.
- Did the daily reminder emails help you to remember to take it as prescribed or would you have thought of taking it the anyway?
- Have you linked the intake with a ritual (e.g. breakfast in the morning)? If so with what was it?
- Was there a general feeling or a recurring thought when taking the pills? If so, what was it?
- Would you have wished that the placebo pills were not to be taken at fixed times, but rather in situations where you would have needed them more (before learning or in case of increased anxiety)?

Effect of the pill

- Did you believe in the effects of the pill or were you skeptical about them?
- When you took the placebo, did you sometimes remember the explanation given during the intervention appointment? If so, what about it? If not, why not?
- During intake, did you think about the effect?
- Do you feel that the pills have helped you during the examination phase?
    - With regards to physical symptoms, thoughts or emotions?

Pill as such

- Well, you got blue mid-sized placebo pills for your test anxiety. Is this kind of pill credible to you, or/and do you think there is another kind of pill that would have helped you better?

3. **Exam Situation**

Effect of pill

- How did you feel during the exam?
- Do you think the pill helped in the exam situation?
  - With regards to physical symptoms, thoughts or emotions?

4. **Future use**

Assessment of effectiveness and future behaviour

With regards to a future use of placebo:

- Would you use this method in a future situation where you need help (be it again before an exam or any other difficult situation)? If yes: Why? If not: Why not?
- Would you recommend this method to others? If so, to whom and why to this person?

Changes in the understanding of placebo through the study

You gave us a definition of placebos at our treatment appointment a few weeks ago. I looked it up and I'm going to read it to you right now:

"Definition of placebo given during treatment appointment"

- If you hear this definition now, would you define placebos as such again, or would you change something about it? If yes: What?
- Would you say that this study has changed your understanding of placebo? If yes: What?

IP group [red indicates differences to OLP]

---

1. **Intervention** [Treatment 3 weeks prior to the exam]

Teaser

Okay, then I'd like to start by asking you a few questions about the intervention three weeks before the exam with *** [person who made the video call]:

- What do you remember most about it?

Rationale

We gave you an explanation why imaginary pill may have an effect.

- Can you remember the explanation? Can you tell me what you remember?
  [Complete if necessary: Discussion points to OLP and imagination research etc.]
- Does this explanation make sense to you?
- What about the explanation made sense to you, what didn't?

Exercise of intake

We practiced the imaginary pill intake together during the intervention appointment.

- How was this exercise for you?
- Was this exercise helpful for the intake at home?
    - If so, why and how could it have been improved?
    - If not, why not and what would have helped you?

2. Intervention phase [Treatment during 3 weeks]

The next questions refer to the three week period before the exam during which you took the imaginary pills.

Pill intake

- Did you find it easy to take the imaginary pills? Were you able to access the image of the pill clearly? Tell me about it.
- Did the daily reminder emails help you to remember to take it as prescribed or would you have thought of taking it the anyway?
- Have you linked the intake with a ritual (e.g. breakfast in the morning)? If so with what was it?
- Was there a general feeling or a recurring thought when taking the pills? If so, what was it?
- Would you have wished that the imaginary pills were not to be taken at fixed times, but rather in situations where you would have needed them more (before learning or in case of increased anxiety)?

Effect of the pill

- Did you believe in the effects of the pill or were you skeptical about them?
- When you took the placebo, did you sometimes remember the explanation given during the intervention appointment? If so, what about it? If not, why not?
- During intake, did you think of the positive state you described back in the video call?
- Do you feel that the pills have helped you during the examination phase?
    - With regards to physical symptoms, thoughts or emotions?

Pill as such

- At the intervention appointment, you described your idea of the pill and the state it should bring you to [i.e.: repeat pill characteristics and the feeling associated with them] Did anything change in the effect or appearance of the pill when you took it at home?
- Is the pill the right shape or would you rather have imagined something else to get it in the desired state? (For example, a ritual like physical exercise?)

3. Exam situation

Effect of pill

- How did you feel during the exam?
- Do you think the pill helped in the exam situation?
    - With regards to physical symptoms, thoughts or emotions?

4. Future use

Assessment of effectiveness and future behaviour

With regards to a future use of imaginary pill:

- Would you use this method in a future situation where you need help (be it again before an exam or any other difficult situation)? If yes: Why? If not: Why not?
- Would you recommend this method to others? If so, to whom and why to this person?

S2 Appendix: Open-ended questions of RCT

Open-ended questions in open-label placebo group

1. Intervention Credibility: Did you find the explanation of why the placebo intervention can work helpful?
    - **a. Yes, why? (open-ended question)**
    - b. No, why? (open-ended question)
2. What do you think about the idea of taking placebo pills? (open-ended question)
3. Do you find the placebo pill has generally helped you to be less anxious/stressed before the exam? Yes/No
4. For which symptoms did the placebo pill help to which extent 0% (the pill did not help at all) - 100% (the pill helped 100%)
    - Concerning excitement (emotional and physical tension)
    - Concern (thoughts about failure, self-doubt)
    - Regarding distraction (distraction from the task by irrelevant thoughts)
    - Regarding confidence (self-worth)
5. **Did you assume that the placebo pills would work or were you skeptical? (open-ended question)**
6. What do you think was in the placebo pills ? (open-ended question)
7. What did you learn by participating in this treatment study ? (open-ended question)
8. Do you have any other comments ? (open-ended question)

Open-ended questions in imaginary pill group

1. Intervention Credibility: Did you find the explanation of why the placebo intervention can work helpful?
    - **a. Yes, why? (open-ended question)**
    - b. No, why? (open-ended question)
2. In general, how open are you to taking a pharmacological pill for your test anxiety ? 0% (not at all open) to 100% (very open) (slider).
3. Do you find the imaginary pill helped you to be less anxious/stressed before the exam ? Yes/No

4. For which symptoms did the imaginary pill help to what extent 0% (the pill did not help at all) - 100% (the pill helped 100%)
   - Regarding excitement (emotional and physical tension)
   - Regarding concerns (thoughts about failure, self-doubt)
   - Regarding distraction (distraction from the task by irrelevant thoughts)
   - Regarding confidence (self-worth)
5. How difficult was it for you to imagine the imaginary pill? 1 (very easy) - 7 (very difficult)
6. How well could you imagine the following aspects of the imaginary pill (0 - not at all well to 100 almost identical to a real pill):
   - Seeing the pill (visualization)
   - Tasting the pill
   - Feeling the pill
   - Effects of the pill
7. Did you find it easier to visualize and take the imaginary pill during the study?
   - Yes, it was easier
   - It did not change
   - No, it became more difficult
9. What do you think about the idea of taking an imaginary pill ? (open-ended question)
10. **Did you assume that the imaginary pill would work or were you skeptical? (open-ended question)**
11. Did you learn anything from participating in this treatment study? If yes, what? (open-ended question)
12. Do you have any other comment? (open-ended question)

S1-S3 Tables: Open-ended questions

S1 Table. Idea of the intervention. What do you think about the idea of taking an imaginary pill / placebo pill?

	OLP N=59	IP N=55	Total N=114
The idea is…		N (%)	
… excellent	10 (16.9)	7 (12.7)	17 (14.9)
… good	29 (49.2)	27 (49.1)	56 (49.1)
… fair	11 (18.6)	11 (20)	22 (19.3)
… poor	9 (15.3)	10 (18.2)	19 (16.7)

Note. IP imaginary pill, *OLP* open-label placebo

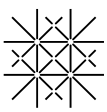
S2 Table. Credibility of explanation. How credible did you find the explanation of why an imaginary pill / open label placebo can work?

	Total N=114
The explanation was…	N (%)
… extremely credible	10 (8.77)
… very credible	63 (55.26)
… moderately credible	34 (29.82)
… hardly credible	2 (1.75)
… minimally credible	4 (3.51)
… not credible at all	1 (0.88)

Helpfulness of explanation How helpful did you find the explanation of why the imaginary pill / open label placebo can work?	Total N=114
The explanation was…	N (%)
… extremely helpful	11 (9.65)
… very helpful	62 (54.39)
… moderately helpful	36 (31.58)
… hardly helpful	2 (1.75)
… minimally helpful	2 (1.75)
… not helpful at all	1 (0.88)

S3 Table. Learning during treatment. Did you learn anything from participating in this treatment study? If yes, what?

	OLP (N=51)	IP (N=50)	Total N=101
I learned…	N (%)		
… how powerful our psyche/ imagination can be	8 (15.7)	12 (24)	20 (19.8)
… that mindfulness can be very helpful	10 (19.6)	7 (14)	17 (16.8)
… how to deal with anxiety	4 (7.8)	6 (12)	10 (9.9)
… to have more self confidence	5 (9.8)	4 (8)	9 (8.9)
… that IPs can actually work and be used daily	0 (0)	9 (18)	9 (8.9)
… that OLPs can actually work	8 (15.7)	0 (0)	8 (7.9)
… more about the placebo mechanism	7 (13.7)	0 (0)	7 (6.9)
… that daily routines help to clear thoughts	2 (3.9)	4 (8)	6 (5.9)
… that adherence is important but difficult	3 (5.9)	2 (4)	5 (5)
… how important expectations are	3 (5.9)	1 (2)	4 (4)
… something about myself	1 (2)	2 (4)	3 (3)
… a potentially new therapy method	0 (0)	2 (4)	2 (2)
… to always stay open minded	0 (0)	1 (2)	1 (1)



Universität
Basel

Fakultät für
Psychologie