OPEN ACCESS

Proportion attributable to contextual effects in general medicine: a meta-epidemiological study based on Cochrane reviews

Yusuke Tsutsumi [ID] ,[1,2] Yasushi Tsujimoto,[3,4] Aran Tajika,[5] Kenji Omae,[3,6] Tomoko Fujii,[7,8] Akira Onishi,[9] Yuki Kataoka,[3,10] Morihiro Katsura,[11] Hisashi Noma,[12] Ethan Sahker,[7,13] Edoardo Giuseppe Ostinelli,[14] Toshi A Furukawa[7]

Check for updates

Abstract

Objectives Our objectives were to examine the magnitude of the proportion attributable to contextual effects (PCE), which shows what proportion of the treatment arm response can be achieved by the placebo arm across various interventions, and to examine PCE variability by outcome type and condition.

Design We conducted a meta-epidemiological study.

Setting We searched the Cochrane Database of Systematic Reviews with the keyword 'placebo' in titles, abstracts and keywords on 1 January 2020.

Participants We included reviews that showed statistically significant beneficial effects of the intervention over placebo for the first primary outcome.

Main outcome measures We performed a random-effects meta-analysis to calculate PCEs based on the pooled result of each included review, grouped by outcome type and condition. The PCE quantifies how much of the observed treatment response can be achieved by the contextual effects.

Public and patient involvement statement No patient or member of the public was involved in conducting this research.

Results We included 328 out of 3175 Cochrane systematic reviews. The results of meta-analyses showed that PCEs varied greatly depending on outcome type (I^2=98%) or condition (I^2=98%), but mostly lie between 0.40 and 0.95. Overall, the PCEs were 0.65 (95% CI 0.59 to 0.72) on average. Subjective outcomes were 0.50 (95% CI 0.41 to 0.59), which was significantly smaller than those of semiobjective (PCE 0.78; 95% CI 0.72 to 0.85) or objective outcomes (PCE 0.94; 95% CI 0.91 to 0.97).

Conclusions The results suggest that much of the observed benefit is not just due to the specific effect of the interventions. The specific effects of interventions may be larger for subjective outcomes than for objective or semiobjective outcomes. However, PCEs were exceptionally variable. When we evaluate the magnitude of PCEs, we should consider each PCE individually, for each condition, intervention and outcome in its context, to assess the importance of an intervention for each specific clinical setting.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Although the proportion attributable to contextual effects (PCE) is highly important to interpreting the results of clinical trials and selecting the appropriate treatment in the clinical setting, a comprehensive review of PCEs among several outcome types and conditions was not yet available.

WHAT THIS STUDY ADDS

⇒ This study showed that the overall PCE was 0.65 (95% CI 0.59 to 0.72).

⇒ The PCE of subjective outcomes was 0.50 (95% CI 0.41 to 0.59), while that of semiobjective and objective outcomes were 0.78 (95% CI 0.72 to 0.85) and 0.94 (95% CI, 0.91 to 0.97), respectively.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ The results suggest that much of the observed benefits in clinical trials are actually due to factors other than specific intervention effects. A smaller PCE may indicate that the effect of interventions on subjective outcomes is larger than on different types of outcomes.

⇒ We should consider each PCE individually, for each condition, intervention and outcome in its context, to assess the importance of an intervention for each specific clinical setting.

Introduction

Placebo has long been used as a dummy treatment in control groups of randomised controlled trials (RCTs) to control for non-specific factors.[1 2] Improvement seen in the intervention group (treatment response) can often be seen in the placebo group and is understood to be due to three contextual effects: the placebo effect, the natural course of the disease or regression to the mean.[3] While the

nature and degree of the placebo effect itself have been controversial,[4–6] assessing the amount of response due to contextual effects helps to explain the benefits specific to an intervention.

The proportion attributable to contextual effects (PCE) is a metric indicating what proportion of the treatment arm response can be achieved by the placebo arm. Response may be beneficial for a dichotomous outcome (eg, survival, remission) or a beneficial change for a continuous outcome (eg, reduction in pain, increase in quality of life (QoL)). Therefore, the PCE corresponds with the contextual effects (placebo effect +natural course +regression to the mean) divided by the intervention arm response (specific effect +placebo effect +natural course +regression to the mean). The PCE ranges from 0 to 1. A score of 0 means that none of the treatment response is due to the contextual effects. Conversely, a PCE score of 1 means that all of the treatment response is due to the contextual effects. A larger PCE reflects a larger contextual effect, or a smaller specific effect of the intervention.[7–9]

An RCT usually focuses on the specific treatment effect, which is the difference in the outcome between the treatment arm and the placebo arm. However, in clinical practice, the overall treatment effect includes not only the specific treatment effect, but also the contextual effects. As a result, a treatment that did not show a large specific effect in an RCT, can still show a larger response in clinical practice than a treatment with large beneficial contextual effects. This phenomenon is called the Efficacy Paradox.[10] Therefore, clinicians and patients should consider both the overall treatment effect and the PCE to interpret clinical trials properly and select appropriate treatments. A large PCE means that a large amount of response seen in patients receiving the treatment can also be seen in patients without receiving the active intervention. Some studies have shown that the PCE might reach 0.65–0.75.[11 12] However, there has been no systematic attempt to review PCEs across various current healthcare interventions.

In this study, we examine the magnitude of PCEs. This will inform how much of an intervention's beneficial effect can be achieved by the contextual effects. In addition, we evaluate PCE variability by outcome type and condition when contextual effects also show a beneficial effect. To achieve this, we systematically surveyed all relevant Cochrane reviews, calculated the PCEs in all fields of medicine, and compared them by the outcome, condition and degree of the certainty of evidence from existing reviews.

Methods

We followed the published reporting guideline for a meta-epidemiological study.[13]

Eligibility criteria

We included all systematic reviews (SRs) of randomised placebo-controlled trials published in the Cochrane Database of Systematic Reviews that showed statistically significant beneficial effects of the intervention over placebo in the first primary outcome. We regarded a two-tailed p<0.05 as statistically significant. We excluded interventions whose efficacy is not established because it would be meaningless to examine PCEs for non-beneficial interventions. When there were multiple comparisons for the first primary outcome due to multiple intervention arms in a review, we selected the first comparison. We included reviews that reported a risk ratio (RR) or an OR for dichotomous outcomes. We excluded reviews reporting other effect sizes, such as HRs in survival analyses. For continuous outcomes, we were able to calculate PCE only when: (1) the meta-analysis reported change scores, (2) the weighted mean of both intervention and control arm showed the same direction of change and (3) larger changes

in the outcome equated to more beneficial changes. We excluded initial reviews that had been updated (ie, we included only the most recent version), reviews of studies other than placebo-controlled trials (eg, sham-controlled trials, non-RCTs, diagnostic test accuracy studies or prognostic studies), overviews of reviews and methodological reviews. We extracted the numerical data from the forest plot of the first meta-analysis. Therefore, we were obliged to exclude reviews that did not show the forest plot of their primary outcome.

Search strategy and study selection

We searched the Cochrane Database of Systematic Reviews with the keyword 'placebo' in titles, abstracts and keywords on 1 January 2020 and selected all the available reviews, regardless of the publication date. Two authors independently performed the initial screening of the titles and abstracts of all studies identified by the search and examined the potential eligibility for inclusion. After initial screening, the same authors assessed the eligibility based on a full-text review. Disagreements were resolved by discussion between the authors, with another author acting as an arbiter when necessary.

Data extraction

Two authors independently used a structured data extraction form to collect data from the included studies. Differences were resolved by consensus. We extracted the pooled RR, SE, and 95% CI when the review reported RR. We extracted the pooled OR, SE, 95% CI and the average control event rate (CER) when the review reported OR. When the review reported a mean difference (MD) of change scores, we extracted the change score and the number of participants for the intervention and placebo arms separately. When the review reported the standardised MD, we extracted the change score of the outcome, SE and the number of participants for the intervention and placebo arms separately.

In addition, we extracted the following information: the number of participants and trials in the meta-analysis of the first primary outcome, the sample size of intervention and placebo arms, outcome data type (dichotomous or continuous), outcome type, condition, intervention type (pharmacological or non-pharmacological), Cochrane review group and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) category of the outcome.

We categorised outcome types and conditions following the categories used in previous studies.[13–15] We modified the category of outcome types, in which we translated harmful outcomes into equivalent beneficial outcomes (eg, from mortality to survival).

We categorised outcome types as follows, following typologies used in the literature.[14–16]

Objective outcome
Survival.

Semiobjective outcomes
No major morbidity events, improved obstetric outcomes, less resource use/shorter hospital stay, improved internal structure (structural outcome within the internal body such as radiograph outcomes), improved external structure (structural outcomes which can be externally observed such as eczema), improved biological markers, no unpleasant composite endpoint, no composite mortality/morbidity events, less drop-out from the treatment, no adverse events and others

Subjective outcomes

Pain relief, QoL improvement, mental health improvement, less consumption/satisfaction with care, cure of condition, no new signs of infection/disease and others

We categorised conditions as follows[14–16]: cardiovascular, central nervous system/musculoskeletal, digestive system, infectious disease, mental health and behavioural conditions, obstetrics and gynaecology, respiratory disease, urogenital and others.

Statistical analyses

The definition of beneficial and harmful outcome measures

Some reviews used an outcome measure in which the higher number of events is better (eg, survival), while others examined the same outcome but used the opposite measure in which the lower number of events is better (eg, death). We defined the outcome measure in which the higher number of events is better as a beneficial outcome and the outcome measure in which the lower number of events is better as a harmful outcome.

Calculation of the CE for each intervention over placebo

We defined and calculated the PCE to designate how much of the beneficial outcome observed in the intervention group is realised in the placebo group. We assumed the independence between the specific treatment effect and the contextual effects.

Beneficial dichotomous outcomes

When the meta-analysis reported an RR for a beneficial outcome (RR >1 expected), we defined the PCE as follows:

$$PCE = \frac{Control\ event\ rate\ (CER)}{Experimental\ event\ rate\ (EER)} \quad (1)$$

where the CER refers to the proportion of outcome events in the placebo group, and the experimental event rate (EER) refers to the proportion of outcome events in the intervention group. We calculated PCE as the inverse of the pooled RR.

When the meta-analysis reported an OR, we first converted the OR to an RR using the average CER obtained in the meta-analysis:

$$RR = \frac{OR}{1 - CER \times (1 - OR)} \quad (2)$$

PCE was then calculated using the RR according to the formula (1).

Harmful dichotomous outcomes

When the meta-analysis reported an RR for a harmful outcome, we converted the RR to an OR using the CER:

$$OR = \frac{RR \times (CER - 1)}{RR \times CER - 1} \quad (3)$$

Then we converted the OR of harmful outcomes to that of beneficial outcomes. For example, ORs for mortality were converted to ORs for survival, taking the inverse of the OR. Finally, we calculated the PCE using formulae (2) and (1).

When the meta-analysis reported an OR, we took the inverse of the OR to represent the OR for a beneficial outcome. Then we calculated the PCE using formulae (2) and (1).

For RRs and ORs, we calculated the 95% CI and SEs of the PCE using a formula for the variance estimators of log(RR) and log(OR), respectively. For dichotomous outcomes, the PCE means that the probability, given an individual had a positive outcome after treatment, would also have had a positive outcome after placebo.

Continuous outcomes

We define the PCE for the continuous outcome as follows in accordance with a previous study.[7]

$$PCE = \frac{the\ mean\ change\ score\ of\ the\ placebo\ arm}{the\ mean\ change\ score\ of\ the\ intervention\ arm} \quad (4)$$

In this study, we calculated the mean change score of each arm based on multiple studies rather than a single study. Therefore, we first calculated the standardised weighted change score means for both the intervention and placebo arms by meta-analysis, using the DerSimonian-Laird method for continuous outcomes.[17]

The PCE was then calculated as:

$$PCE = \frac{weighted\ standardized\ mean\ change\ score\ of\ the\ placebo\ arm}{weighted\ standardized\ mean\ change\ score\ of\ the\ intervention\ arm} \quad (5)$$

The conditions to calculate PCEs by formula (5) were that both groups showed the same direction of change (same positive or negative direction), and the change in the treatment group was greater than the change in the placebo group. Therefore, we excluded Cochrane reviews when the weighted mean of the intervention and control arms did not show the same direction of change or when the weighted standardised mean change score of the placebo arm was greater than that of the intervention arm.

To construct the CI of PCE, the ordinary Wald-type CI is unsuitable for the original scale because PCE is a ratio measure and usually has an asymmetric sample distribution. Thus, we adopted the approximate CI for log-transformed PCE. Using the Delta method, we obtained the SE estimator of the log(PCE) as follows[18]:

$$SE\left[\log(PCE)\right] = \sqrt{\frac{SE_{intervention}^2}{WSM_{intervention}^2} + \frac{SE_{placebo}^2}{WSM_{placebo}^2}} \quad (6)$$

where SE [log(PCE)] was the SE of log(PCE), $SE_{intervention}$ and $SE_{placebo}$ represent the SE of the weighted change score means of the intervention and placebo arms, respectively. $WSM_{intervention}$ and $WSM_{placebo}$ were the weighted standardised change score means of the intervention and placebo arms, respectively. The weighted standardised means were defined as pooled summaries of standardised mean change scores from the intervention and placebo arms of individual studies. The weights were defined by the corresponding meta-analysis methods. We could then obtain the Wald-type 95% CI for log(PCE) by the conventional normal approximation.[19] The 95% CI of PCE can be calculated by back-transformation (exponential transformation) of the confidence limits.

Meta-analyses of PCEs

We pooled the logarithm-transformed PCEs for each outcome type, condition and GRADE category, using the DerSimonian-Laird method.[17] We performed all analyses using meta (V.4.15–1) package of R V.4.0.0.[20 21]

Changes from the protocol

The study protocol was uploaded to the Department of Health Promotion and Human Behaviour website, Kyoto University Graduate School of Medicine/School of Public Health (http://ebmh.med.kyoto-u.ac.jp/r-meta.html, attached as online supplemental file 1) on 4 September 2018. Changes to the protocol are listed in online supplemental file 2, and all were minor.

Patient and public involvement

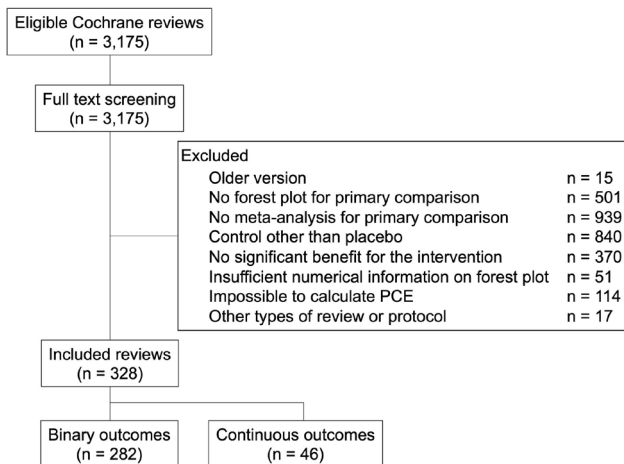No patient or member of the public was involved in conducting this research.

Figure 1 The flow chart for study selection. PCE, proportion attributable to contextual effect.

Results

Selected studies

Figure 1 shows the flowchart for study selection. The initial search identified 3175 Cochrane reviews, from which we excluded 2847 for reasons seen in figure 1. The final selection included 328 reviews. Of these, 282 reviews represent 1755 trials, with 2 625 184 participants reporting dichotomous outcomes. The remaining 46 reviews represent 259 trials, with 42 156 participants reporting continuous outcomes.

Over half of the reviews were published within the past 10 years. Approximately 70% of the included reviews reported subjective outcomes. Mental health and behaviour conditions was the most popular condition (table 1). Online supplemental file 3 shows all included SRs and their outcomes, conditions and PCEs calculated from the results of the meta-analysis from each review. For example, Bennett et al[22] examined the effect of tranexamic acid on mortality for upper gastrointestinal bleeding. They reported an RR of 0.6. The CER for mortality was 8.4% and converted to a positive outcome CER for survival at 91.6%. The EER of survival was 95.0%. Dividing the CER of survival by the EER of survival gives a PCE of 0.96. By contrast, Derry et al[23] examined the effect of ibuprofen plus codeine for acute postoperative pain and showed an RR of 4.1. Their CER for a 50% reduction of maximum pain was 18.0%, and the EER was 64.0%. Thus, the calculated PCE was 0.24.

PCEs by outcome types, conditions and GRADE categories

Outcome types

We first pooled PCEs across all outcome types, including both dichotomous and continuous outcomes (figure 2, table 2). The pooled PCE across all outcomes was 0.65 (95% CI 0.59 to 0.72). When the outcome types were divided into subgroups, objective outcomes showed a PCE of 0.94 (95% CI 0.91 to 0.97), which was higher than that of semiobjective outcomes (PCE 0.78; 95% CI 0.72 to 0.85, I^2=90%) and subjective outcomes (PCE 0.50; 95% CI 0.41 to 0.59, I^2=99%). The PCE of typical patient-reported outcomes such as pain, QoL, and mental health outcomes ranged between 0.44 and 0.74.

We next pooled PCEs across all outcome types when the outcomes were dichotomous (online supplemental figure S1) or continuous (online supplemental figure S2).

Conditions

Figure 3 shows the PCE by condition. PCEs ranged widely from 0.40 in 'anaesthesia' to 0.89 in 'cardiovascular disease' (I^2=98%). Online supplemental figure S3 and S4 present the results for dichotomous and continuous outcomes separately.

Grade categories

Among 156 Cochrane reviews that had reported a GRADE in the summary of findings table, the PCEs did not appear to be appreciably influenced by the certainty of evidence (online supplemental figure S5, $I^2 = 54\%$). Online supplemental figures S6 and S7 present these results for dichotomous and continuous outcomes separately.

Discussion

Key findings

Based on 328 Cochrane reviews representing 2014 trials and 2 667 340 participants, the overall PCE for various interventions in contemporary medicine was 0.65 (0.59 to 0.72). PCEs ranged from 0.28 to 0.94. The subjective outcomes showed lower PCEs than objective or semiobjective outcomes. They also varied depending on the condition. GRADE ratings did not appear to influence PCEs. These results suggest that PCEs should be considered according to the outcome type and condition when interpreting study results and determining the importance of interventions.

Comparisons with the previous literature

As previously stated, there is no study examining the PCE of several semiobjective outcomes. In our study, PCEs of the objective and semisubjective outcomes were mostly larger than 0.70. Several studies have reported PCEs for patient-reported outcomes in the literature. One study revealed a PCE of 0.72 for pain among burning mouth syndrome patients,[24] while another showed a PCE of 0.82 for antidepressant trials for depression.[25] When we examined PCEs for various patient-reported outcomes and conditions, the average PCE seemed to lie in the medium range between 0.40 and 0.75, which was lower than those reported in previous studies. The effects of the intervention on the subjective outcomes may be larger than those for the other types of outcomes. Previous studies suggest larger placebo effects for subjective outcomes than for objective outcomes.[6 26] These findings are in line with the current findings of contextual effects. This is because, regardless of the magnitude of the placebo effect, PCEs represent the proportion of the contextual effects of the improvements observed in the active intervention arm.

Clinical interpretations

The contextual effects have been known to contribute to the treatment response.[10 27] However, the magnitude of PCEs had not been examined in detail or quantified in the literature until this study. Our study presented the average PCEs across outcome types and conditions.

We found that subjective outcomes, including typical patient-reported outcomes such as pain, QoL, and mental health outcomes, showed PCEs of 0.44–0.74. This may reflect the fact that the specific intervention effect seen in subjective outcomes are likely to be of moderate magnitude and clinically important. On the other hand, 'hard' outcomes, including survival or morbidity events, showed PCEs above 0.80 or even 0.90. We assume PCEs seen in objective and semiobjective outcomes may reflect the fact that the natural disorder course plays a strong role in those outcomes. It would

Table 1 Characteristics of included studies

		Dichotomous outcome (n=282)	Continuous outcome (n=46)
No of included primary studies per meta-analysis, median (IQR)		4 (2–7)	4 (3–7)
No of included participants per meta-analysis, median (IQR)		557 (283–1490)	511 (229–1094.5)
Years of publication (%)	1998–2000	14 (5.0)	0 (0)
	2001–2005	27 (9.6)	10 (21.7)
	2006–2010	57 (20.2)	7 (15.2)
	2011–2015	101 (35.8)	12 (26.1)
	2016–2019	83 (29.4)	17 (37.0)
Outcome type (%)			
Objective outcome			
	Survival	7 (2.5)	0 (0)
Semiobjective outcomes			
	No major morbidity event	14 (5.0)	0 (0)
	Less drop-outs from the treatment	11 (3.9)	0 (0)
	Improved obstetric outcomes	9 (3.2)	0 (0)
	Less resource use/shorter hospital stay	9 (3.2)	0 (0)
	Improved internal structure	5 (1.8)	2 (4.3)
	Improved biological markers	4 (1.4)	6 (13.0)
	Improved external structure	4 (1.4)	1 (2.2)
	No unpleasant composite endpoint	4 (1.4)	0 (0)
	No adverse events	3 (1.1)	0 (0)
	No composite mortality/morbidity events	0 (0)	0 (0)
	Others (semiobjective)	1 (0.4)	6 (13.0)
Subjective outcomes			
	Pain relief	72 (25.5)	9 (19.6)
	Cure of condition	71 (25.2)	4 (8.7)
	Mental health outcomes improvement	35 (12.4)	12 (26.1)
	No new signs of infection/disease	17 (6.0)	0 (0.0)
	Quality of life	13 (4.6)	6 (13.0)
	Less consumption, satisfaction with care	1 (0.4)	0 (0.0)
	Others (Subjective)	2 (0.7)	0 (0.0)
Conditions (%)			
	Mental health and behavioural conditions	56 (19.9)	12 (26.1)
	Anaesthesia	48 (17.0)	3 (6.5)
	Central nervous system/ musculoskeletal	47 (16.7)	7 (15.2)
	Digestive system	29 (10.3)	0 (0)
	Infectious disease	26 (9.2)	1 (2.2)
	Respiratory disease	19 (6.7)	3 (6.5)
	Obstetrics and gynaecology	16 (5.7)	2 (4.3)
	Cardiovascular	12 (4.3)	7 (15.2)
	Urogenital	7 (2.5)	4 (8.7)
	Others	22 (7.8)	7 (15.2)
GRADE category (%)*	High	35 (25.9)	2 (9.5)
	Moderate	53 (39.3)	5 (23.8)
	Low	36 (26.7)	13 (61.9)
	Very low	11 (8.1)	1 (4.8)

*A total of 147 dichotomous outcomes and 25 continuous outcomes did not report the GRADE category in their summary of findings table.
GRADE, Grading of Recommendations Assessment, Development and Evaluation.

be relatively difficult for an intervention to show large effects in these 'hard' outcomes.

The pooled estimates are approximate indicators of how much PCE we can expect for each subgroup. However, the average PCEs are only a rough guide, and a starting point for clinical evaluation of an intervention effect. Generally, the higher the PCE, the greater the contextual effects. Therefore, watchful waiting or careful observation with follow-ups may be a reasonable choice,

rather than an active intervention, because much of the benefit can be expected from the placebo intervention. However, this general judgement should be tempered by the gravity of the outcome (eg, a small decrease in death may be as valuable as or even more valuable than a large decrease in some non-life-threatening consequences) and the burdens (side-effects and costs) of the alternative treatments. Moreover, our study revealed that PCEs were highly heterogeneous. We must consider the individual
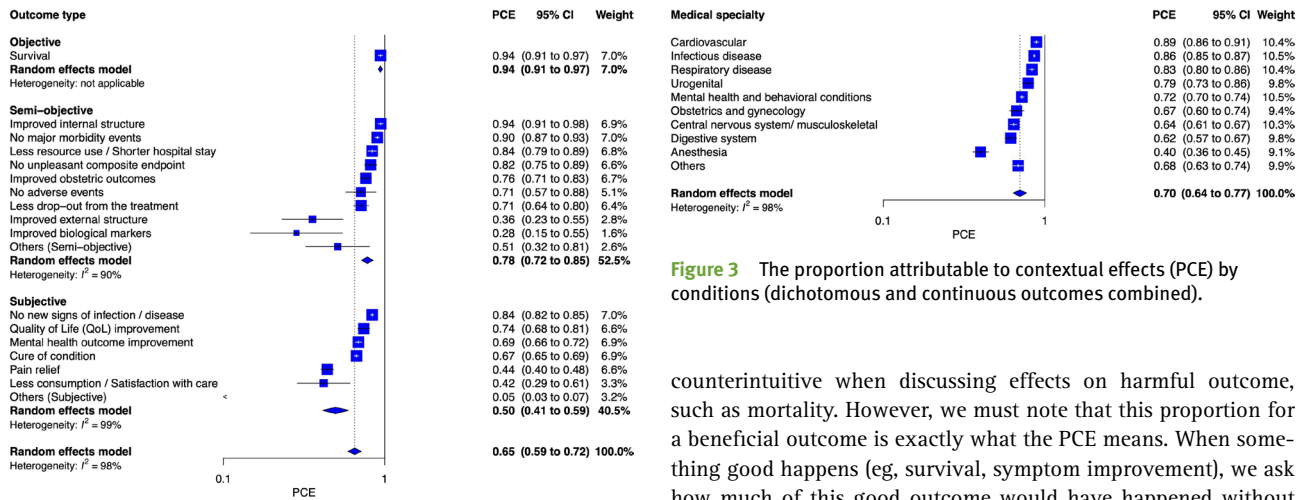
Figure 2 The proportion attributable to contextual effects (PCE) by outcome type (continuous and dichotomous outcomes combined)

Figure 3 The proportion attributable to contextual effects (PCE) by conditions (dichotomous and continuous outcomes combined).

condition, intervention and outcome when we evaluate the clinical significance of the PCE.

We also found that PCEs were diverse over several conditions, partly reflecting outcome types commonly used in each condition. For example, the PCEs of cardiovascular disease, infectious disease, and respiratory disease were over 0.80. Trials examining these conditions usually select objective or semiobjective outcomes such as survival, major morbidity, composite endpoints or severity of symptoms as the primary outcome. In fact, 30 of 68 included reviews among these conditions reported objective or semiobjective outcomes. On the other hand, conditions with subjective main outcomes, such as pain, mental health or behaviour conditions, showed moderate PCEs.

Additionally, we found no remarkable diversity of PCEs depending on the GRADE category. However, over half of the included reviews did not report the GRADE category with a summary of findings table of their first primary outcome, which led to the loss of power in this analysis. We also found that out of 328 Cochrane reviews, 282 reported a dichotomous outcome, such as mortality, as their first primary outcome, and we were only able to identify 46 Cochrane reviews with a primary continuous outcome reported first.

Strengths and limitations

There are some weaknesses to our study. First, we calculated a PCE as the proportion of the contextual effects over the intervention response for a beneficial outcome. This might appear
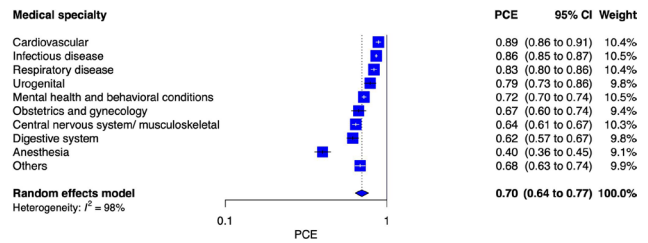
counterintuitive when discussing effects on harmful outcome, such as mortality. However, we must note that this proportion for a beneficial outcome is exactly what the PCE means. When something good happens (eg, survival, symptom improvement), we ask how much of this good outcome would have happened without the intervention (ie, it can be explained by the contextual effects). Second, we were unable to calculate PCE when the intervention showed improvement from baseline while the placebo group showed deterioration, or when the intervention showed deterioration while the placebo group showed greater deterioration. This may have led to overestimation or underestimation of PCE. This was inevitable, given the definition of PCE above. We wanted to quantify when something good happens following treatment and how much of that good could happen without the intervention. Third, there is intrinsic heterogeneity among outcomes belonging to the same type (eg, morbidity event) or used in the same condition (eg, respiratory disease). Each outcome may have different clinical importance depending on the context. Because PCE is a relative measure, the CER may also influence the absolute benefit. In other words, the absolute benefit is larger for events with large CERs, and smaller for events with smaller CERs for a given PCE. Therefore, each PCE of an intervention must be interpreted in its specific context and, ultimately, under each patient's preferences and values. However, it is scientifically meaningful to know the overall averages even when the apparent heterogeneity may be high. Fourth, measurement errors (random errors) are larger for subjective outcomes than semiobjective or objective outcomes. How such measurement errors may have contributed to the estimates of PCEs is not straightforward. However, by pooling across SRs, the overall averages would be less affected by such random errors. Fifth, our study calculated PCEs by outcome types and conditions independently. PCE is highly dependent on both the condition and outcome. For example, the PCE for pain in 'cardiovascular' may be larger than that in 'mental health and behavioural conditions' because the pain in cardiovascular is more likely

Table 2 The proportion attributable to contextual effects by outcome type

		Outcome category		All
		Dichotomous	Continuous	
Outcome type	Objective	0.94 (0.91 to 0.97) (NA)	NA	0.94 (0.91 to 0.97) (NA)
	Semiobjective	0.81 (0.75 to 0.87) (0.64 to 1.00)	0.54 (0.38 to 0.77) (0.15 to 1.00)	0.78 (0.72 to 0.85) (0.60 to 1.00)
	Subjective	0.49 (0.41 to 0.59) (0.26 to 0.94)	0.60 (0.54 to 0.66) (0.46 to 0.79)	0.50 (0.41 to 0.59) (0.26 to 0.94)
All		0.67 (0.61 to 0.74) (0.45 to 0.99)	0.59 (0.53 to 0.66) (0.47 to 0.75)	0.65 (0.59 to 0.72) (0.44 to 0.97)

(Parentheses) show 95% CIs, and (brackets) show 95% prediction intervals.
NA, not applicable.

due to an organic cause. Finer subgroup analyses examining their interactions were impossible in the current study, as we had too many subgroups, but this would be desirable for future research. Finally, the specific and contextual effects are not independent of each other but may interact with each other when, for example, the side effects of the intervention may enhance the non-specific effects.[28] Thus, the same outcome for the same disease may show different PCEs depending on the intervention. We did not have enough variability in the interventions in our dataset to appreciate such interactions. Further research would be warranted to clarify the nature and degree of the relationships between specific effects and contextual effects.

Our study also has several strengths. First, to the best of our knowledge, this is the first study to examine PCEs comprehensively across outcome types and conditions. Second, our study is based on Cochrane reviews which are considered the best resources of evidence. We searched all the Cochrane reviews that had reported the results of head-to-head meta-analyses of interventions against placebo. We pooled these meta-analyses to calculate the PCEs by outcome type, condition and GRADE category.

Conclusions

In conclusion, our study showed that the overall PCE was 0.65 (95% CI 0.59 to 0.72). PCEs were smaller for subjective outcomes than for objective or semiobjective outcomes. The results suggest that much of the observed benefit is actually due to factors including the placebo effect, the natural course of the disease, and regression to the mean, rather than just the specific effect of the interventions. Specific effects of interventions may be larger for subjective than for objective or semiobjective outcomes. However, PCEs were exceptionally variable. When we evaluate the magnitude of PCEs, we should consider each PCE individually, for each condition, intervention and outcome in its context, to assess the importance of an intervention for each specific clinical setting.

Author affiliations
[1]Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
[2]Department of Emergency Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan
[3]Department of Healthcare Epidemiology, Kyoto University Graduate School of Medicine/ School of Public Health, Kyoto, Japan
[4]Department of Nephrology and Dialysis, Kyoritsu Hospital, Kawanishi, Japan
[5]Department of Psychiatry, Kyoto University Hospital, Kyoto, Japan
[6]Department of Innovative Research and Education for Clinicians and Trainees (DiRECT), Fukushima Medical University Hospital, Fukushima, Japan
[7]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan
[8]Intensive Care Unit, Jikei University Hospital, Tokyo, Japan
[9]Department of Rheumatology and Clinical Immunology, Kobe University Graduate School of Medicine, Hyogo, Japan
[10]Hospital Care Research Unit/ Department of Respiratory Medicine, Hyogo Prefectural Amagasaki General Medical Center, Hyogo, Japan
[11]Department of Surgery, Okinawa Chubu Hospital, Okinawa, Japan
[12]Department of Data Science, The Institute of Statistical Mathematics, Tokyo, Japan
[13]Population Health and Policy Research Unit, Medical Education Center, Kyoto University Graduate School of Medicine, Kyoto, Japan
[14]Oxford Health NHS Foundation Trust, Warneford Hospital /Department of Psychiatry, University of Oxford, Oxford, UK

ORCID iD
Yusuke Tsutsumi http://orcid.org/0000-0002-9160-0241

References

1 **Beecher HK**. The powerful placebo. *J Am Med Assoc* 1955;**159**:1602–6.
2 **Blease CR**, Bishop FL, Kaptchuk TJ. Informed consent and clinical trials: where is the placebo effect? *BMJ* 2017;**356**:j463.
3 **Kirsch I**. The placebo effect revisited: lessons learned to date. *Complement Ther Med* 2013;**21**:102–4.
4 **Espay AJ**, Norris MM, Eliassen JC, *et al.* Placebo effect of medication cost in Parkinson disease: a randomized double-blind study. *Neurology* 2015;**84**:794–802.
5 **Kaptchuk TJ**, Kelley JM, Conboy LA, *et al.* Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ* 2008;**336**:999–1003.
6 **Hróbjartsson A**, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010;**1**:CD003974.
7 **Zou K**, Wong J, Abdullah N, *et al.* Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2016;**75**:1964–70.
8 **Zhang W**, Doherty M. Efficacy paradox and proportional contextual effect (PCE). *Clin Immunol* 2018;**186**:82–6.

9 Zhang W, Zou K, Doherty M. Placebos for Knee Osteoarthritis: Reaffirmation of "Needle Is Better Than Pill". *Ann Intern Med* 2015;**163**:392–3.

10 Walach H. Placebo Effects in Complementary and Alternative Medicine: The Self- Healing Response. In: Colloca L, Arve MA, Meissner K, eds. *Placebo and pain. academic press*. Elsevier, 2013: 189–202.

11 Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication. *Prev Treat* 1998;**1**:2a.

12 Jonas WB, Crawford C, Colloca L, *et al*. To what extent are surgery and invasive procedures effective beyond a placebo response? A systematic review with meta-analysis of randomised, sham controlled trials. *BMJ Open* 2015;**5**:e009655.

13 Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* 2017;**22**:139–42.

14 Davey J, Turner RM, Clarke MJ, *et al*. Characteristics of meta-analyses and their component studies in the Cochrane database of systematic reviews: a cross-sectional, descriptive analysis. *BMC Med Res Methodol* 2011;**11**:160.

15 Rhodes KM, Turner RM, Higgins JPT. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *J Clin Epidemiol* 2015;**68**:52–60.

16 Turner RM, Davey J, Clarke MJ, *et al*. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane database of systematic reviews. *Int J Epidemiol* 2012;**41**:818–27.

17 DerSimonian R, Laird N. Chapter 10: meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**:177–88.

18 Oehlert GW. A note on the delta method. *Am Stat* 1992;**46**:27–9.

19 Cox DR, Hinkley DV, Statistics T. 1St ED. *London. Chapman and Hall* 1974.

20 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;**22**:153–60.

21 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020.

22 Bennett C, Klingenberg SL, Langholz E, *et al*. Tranexamic acid for upper gastrointestinal bleeding. *Cochrane Database Syst Rev* 2014:CD006640.

23 Derry S, Karlin SM, Moore RA. Single dose oral ibuprofen plus codeine for acute postoperative pain in adults. *Cochrane Database Syst Rev* 2015;**2**:CD010107.

24 Kuten-Shorrer M, Kelley JM, Sonis ST, *et al*. Placebo effect in burning mouth syndrome: a systematic review. *Oral Dis* 2014;**20**:e1–6.

25 Kirsch I, Deacon BJ, Huedo-Medina TB, *et al*. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. *PLoS Med* 2008;**5**:e45.

26 Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? an analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 2001;**344**:1594–602.

27 Shaibani A, Frisaldi E, Benedetti F. Placebo response in pain, fatigue, and performance: possible implications for neuromuscular disorders. *Muscle Nerve* 2017;**56**:358–67.

28 Flaten MA, Simonsen T, Olsen H. Drug-Related information generates placebo and nocebo responses that modify the drug response. *Psychosom Med* 1999;**61**:250–5.