REVIEW

Component network meta-analysis identifies the most effective components of psychological preparation for adults undergoing surgery under general anesthesia

Suzanne C. Freeman[a,b,*], Neil W. Scott[b], Rachael Powell[c], Marie Johnston[b], Alex J. Sutton[a], Nicola J. Cooper[a]

[a]*NIHR Complex Reviews Support Unit, Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, UK*
[b]*Medical Statistics Team, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, UK*
[c]*Manchester Centre for Health Psychology and School of Health Sciences, University of Manchester, Manchester, UK*

Abstract

Objectives: To apply component network meta-analysis (CNMA) models to an existing Cochrane review of psychological preparation interventions for adults undergoing surgery and to extend the models to account for covariates to identify the most effective components for improving postoperative outcomes.

Study Design and Setting: Interventions consisted of between one and four components of psychological preparation: procedural information (P), sensory information (S), behavioral instruction (B), cognitive interventions (C), relaxation (R), and emotion-focused techniques (E). We used CNMA models to assess the effect of each component for three outcomes: length of stay, pain, and negative affect.

Results: We found evidence that the most effective component for reducing length of stay depends on the type of surgery and that R may improve pain. There was insufficient evidence that individual components contributed to the overall reduction in negative affect, but P and S emerged as the most likely beneficial components. Overall, we were unable to identify any one component as the most effective across all three outcomes.

Conclusion: The CNMA method allowed us to address questions about the effects of specific components that could not be answered using standard Cochrane methodology. © 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Network meta-analysis; Complex interventions; Psychological preparation; Surgery; Bayesian; Component

1. Introduction

Systematic reviews often study complex interventions comprising multiple components that may be delivered individually or in combination. One commonly accepted definition is that a complex intervention consists of "a number of components that may act both dependently and independently" [1,2].

Network meta-analysis (NMA) is an extension of pairwise meta-analysis (MA) methods to the setting where we have three or more interventions [3]. NMA incorporates both direct evidence from trials directly comparing two or more treatments (eg, a direct comparison between treatments A and B) and indirect evidence which arises when two treatments are both directly compared with the same third treatment (eg, indirect comparison of A and B, where each is directly compared with treatment C in separate trials). NMA uses a single statistical model to combine both the direct and indirect evidence within a network to estimate intervention effects for every treatment combination, regardless of whether two interventions have been directly compared or not [4].

In 2009, Welton et al. proposed four increasingly complex component network meta-analysis (CNMA) models (that increasingly make less strong assumptions) for assessing component effects of complex interventions [5]. The

What is new?

Key findings
- We used component network meta-analysis (CNMA) to complement standard pairwise meta-analysis techniques used in an existing Cochrane review of psychological preparation for surgery to overcome important methodological limitations of the original review and strengthen the value of the results for both application and contribution to the scientific evidence-base.

What this adds to what was known?
- CNMA allowed us to derive estimates of the effects of individual psychological components, even when delivered in combination

- Any of the investigated types of psychological preparation may be beneficial for reducing length of stay, pain, and negative affect, but certain components may be more important for individual outcomes (eg, relaxation for postoperative pain)

- Covariates such as type of surgery can be incorporated into the CNMA models and enhance the utility of the results

What is the implication and what should change now?
- CNMA is a little-used technique that could be more widely used in systematic reviews involving multi-component complex interventions

simplest model is equivalent to a pairwise MA and assesses the effect of any intervention. The models build in complexity introducing additive component effects, pairwise interactions between components, and unique effects for each observed combination of components. A citation review (July 2017) suggested that this article has been cited over 130 times in general methodology and NMA articles, but we have only identified one other research group that appears to have fully applied the CNMA method in practice [6—8].

In 2016, a meta-analysis published in the Cochrane Database of Systematic Reviews identified better postoperative outcomes for patients receiving psychological preparation before surgery [9]. The term "psychological preparation" refers to strategies designed to influence a person's thoughts, feelings, or actions. Johnston and Vögele [10] found that seven types of psychological preparation led to patient benefit for at least one postoperative outcome: providing procedural information (what, when, and how events will occur, aiming to reduce anxiety by reducing uncertainty); sensory information (sensations associated with a procedure—eg, what it will feel or smell like, expected to reduce anxiety by eliminating unexpected, and therefore worrying sensations); behavioral instruction (telling or teaching participants about actions they should perform to enhance the experience or outcomes); cognitive interventions (to change how an individual thinks, particularly aiming to reduce negative thoughts and thereby reduce negative emotions); emotion-focused interventions (to help the individual to manage their feelings); relaxation; hypnosis.

While surgery may be intended to improve (or diagnose) health conditions, it also has short-term impacts on outcomes such as pain and activity levels. How people think and feel before surgery impacts postoperative outcomes. For example, anxiety, depression, and catastrophizing have been found to predict postoperative pain [11—14]. There are a range of mechanisms by which psychological factors may influence postsurgical outcomes. Pain tends to be experienced as more severe when negative emotions are experienced [15,16]; behaviors are influenced by cognitions and emotions (eg, ambulatory behaviors postsurgery; using patient-controlled analgesia), and stress is associated with slower wound healing [17,18].

The 2016 review included 105 randomized controlled trials, where each active trial arm was classified as comprising between one and four psychological components. The outcomes were restricted to postoperative pain, negative affect (negative emotion, such as anxiety, depression), hospital length of stay, and behavioral recovery; findings for behavioral recovery were not included in a meta-analysis because of the small number of suitable studies and diversity of outcome measures. The analyses were limited to pairwise meta-analysis comparisons in which any form of intervention was grouped together in an "any intervention" arm and compared with "control". The review suggested a benefit of psychological intervention, but interpretation of the results for the effect of the individual components was more difficult. For each type of psychological component, subgroup analyses were conducted for two groups of trials: (1) those with an intervention group that contained only that component and (2) those with an intervention group including the component of interest, possibly in combination with other components. Most psychological interventions tended to be delivered in combination with other components, complicating the interpretation of these subgroup analyses. For all three outcomes meta-analyzed, better outcomes were identified with intervention compared with control. However, for all three outcomes, there was evidence of statistical heterogeneity. Further analysis was therefore required to identify individual component effects and to account for heterogeneity.

Further complexity in identifying component effects may arise from heterogeneity between trial results caused by factors other than intervention components including differences in baseline risk across trials. As defined by Achana et al. "baseline risk reflects the burden of disease in a study population and defines the average risk of a

patient to experience the outcome of interest if they have not been treated" that is, the expected outcome for a patient who does not receive psychological preparation [19]. Thus, if variables, in addition to those describing the intervention, modify treatment effects, then it is desirable to identify them and include them in the statistical model. In this article, we will assess the effect of baseline risk on treatment effectiveness using the mean effect for the control group. To illustrate, heterogeneity may arise as length of stay is likely to be longer in older trials or for some surgical procedures; control groups are subject to these influences, and their means reflect the baseline risk. In addition, there is ample evidence that effect sizes for an intervention are strongly influenced by the inputs to the control or "standard care" group and so controlling for baseline risk allows the effects of interventions to be investigated as if their control groups were equivalent [20].

In this article, we report extended analyses of the Cochrane review data on psychological preparation for adults undergoing surgery under general anesthesia. We increase our understanding of the data by applying CNMA models to identify the most effective components of psychological preparation for improving postoperative outcomes. Given the substantial between-study heterogeneity in this review, we also extend the models used by Welton et al. [5] to include baseline risk and type of surgery.

2. Methods

2.1. Study population

We used data from 71 trials included in the meta-analyses of length of stay, pain, and negative affect reported in the Cochrane review of psychological preparation and postoperative outcomes for adults undergoing surgery under general anesthesia (full reference list in Appendix A/ Appendix on the journal's web site at www.elsevier.com). The Cochrane review included an additional 34 trials, which were included in narrative syntheses only as there was insufficient information for them to be included in the meta-analyses. These trials are not considered in this article. Interventions consisted of between one and four components made up of the following psychological preparations: procedural information (P), sensory information (S), behavioral instruction (B), cognitive intervention (C), relaxation (R), and emotion-focused techniques (E). A seventh category used in the Cochrane review, hypnosis, was combined with the R category due to the low number of trials including this component, and because hypnosis procedures typically incorporate R elements.

2.2. Outcomes

Three outcomes were considered: length of stay, postoperative pain, and negative affect (negative emotion, eg, anxiety, depression). Length of stay was measured as the number of days a patient remained in hospital following surgery. Pain and negative affect were measured using a number of different standard scales across trials (details on which underlying scales were used in each study can be found elsewhere [9]). For these two outcomes results were synthesized as standardized mean differences (SMDs) using Hedges' g.

2.3. Included studies

We essentially used the same data as the Cochrane review but included some trials only in sensitivity analyses for the following reasons. One trial [21] was identified in the Cochrane review as an outlier and excluded from the primary analysis of pain. One trial [22] reporting change from baseline rather than final score was excluded (here) from the primary analysis of pain. Five trials [23–27] reporting change from baseline rather than final score were excluded from the primary analysis of negative affect. One trial [28] that provided binary outcome data only was excluded from all pain analyses. In addition, we were able to include one trial [29], which was excluded from the Cochrane analyses because it did not have a control arm.

2.4. Statistical analysis

Initially, four models were fitted using the contrast-based data approach of Welton et al., 2009 [5]. A linear regression model was used, where μ_j represents the mean response for the baseline intervention b_j (ie, the control arm) in trial j. Then θ_{jk}, the mean outcome for intervention k from trial j, was modeled as:

$$\theta_{jk} = \begin{cases} \mu_j, & \text{Intervention } b_j \\ \mu_j + \delta_{jk}, & \text{Intervention } k \end{cases}$$

where

$$\delta_{jk} \sim \text{Normal}\big((d_k - d_{b_j}), \tau^2\big) \tag{1}$$

δ_{jk} represents the mean difference (MD) in the change in outcome for intervention k relative to the baseline intervention b_j in trial j with between-trial variance τ^2. d_k is the MD of the outcome for intervention k relative to the reference treatment across the network (ie, the "treatment" effect for intervention k compared with control). The four models arise from making different assumptions about d_k.

Model 1 recreates the original Cochrane review by combining all interventions into one arm and comparing them to the control arm.

$$d_k = d$$

Model 2 allows additive main effects in which each component has a separate effect. In this model, the total effect of an intervention is equal to the sum of the relative component effects.

$$d_k = d_P + d_S + d_B + d_C + d_R + d_E$$

where

$$d_P = \begin{cases} 1, & \text{if intervention includes component } P \\ 0, & \text{otherwise} \end{cases}$$

Model 3 is an extension of model 2 with extra terms for combinations of pairs of components (a total of 13 pairwise interactions between components for length of stay, 12 for pain, and 11 for negative affect). This allows pairs of components to have a bigger or smaller effect than would be expected from the sum of their individual components.

$$d_k = d_P + d_S + d_B + d_C + d_R + d_E + d_{PS} + \ldots + d_{RE}$$

where

$$d_{PS} = \begin{cases} 1, & \text{if intervention includes components } P \text{ and } S \\ 0, & \text{otherwise} \end{cases}$$

Finally, model 4 is a standard NMA model in which each possible combination of components is considered to be a separate intervention so that each possible combination of components has its own effect, regardless of whether it is made up of one or four components. For example, in a network consisting of the six intervention combinations P, S, P + S, P + S + B, C + R and S + B + C + E:

$$d_k = d_P + d_S + d_{P+S} + d_{P+S+B} + d_{C+R} + d_{S+B+C+E}$$

where each d = 1 only for each specific disease combination and 0 otherwise (eg, for intervention P + S, the treatment effect is given by d_{P+S} only).

For outcomes combined on the SMD scale, we standardize each outcome by dividing the mean change in outcome for intervention k from trial j (α_{jk}) by the pooled standard deviation (SD_j) across all trial arms from trial j so that:

$$\theta_{jk} = \frac{\alpha_{jk}}{SD_j}$$

Models 1−4 are then fitted as previously described.

Throughout this article, component effects are presented compared with the reference treatment "control". In the Cochrane review trials, reporting final score and change from baseline were synthesized together. However, when considering standardized scales the "difference in standard deviation reflects differences in the reliability of the measurements" rather than differences in the measurement scale [30]. Therefore, in the present article, primary analyses for pain and negative affect included trials reporting final score only. All analyses took into account adjustment for multi-arm trials [31].

2.5. Inclusion of covariates

Control group mean (used as a measure of the expected outcome without intervention) and type of surgery were considered for inclusion into the CNMA models, adapting preexisting methods [19]. Control group mean was centered on the mean control group outcome across all trials. Therefore, the treatment effect model coefficients can be interpreted as the treatment effect at the average value of the covariate.

Type of surgery was split into three categories (cardiovascular [CV], orthopedic, and other) and fitted as a categorical covariate through a series of indicator variables with "other" treated as the reference group. Further details regarding the fitting and implementation of these models can be found in Appendix B/Appendix on the journal's web site at www.elsevier.com.

2.6. Simultaneous assessment across the multiple outcomes

To assess whether there is any agreement across the three outcomes about which component or components are most effective, we calculated the probability of being the best component for each outcome and present these results for all outcomes in a bar chart [32]. In addition, we calculated the probability of each component taking each rank from most effective to least effective and display these results as a line graph for each component.

3. Results

In this section, we present the results for length of stay and pain. Results for negative affect are presented in Appendix C/Appendix on the journal's web site at www.elsevier.com. Sensitivity analyses for pain are presented in Appendix D/Appendix on the journal's web site at www.elsevier.com. Additional tables and figures are included in Appendix E/Appendix on the journal's web site at www.elsevier.com.

3.1. Length of stay

We considered a network of 35 trials assessing 18 different combinations of components (Figure 1). Model 1 identified a reduction of −0.541 days (95% credible interval [CrI]: −0.913, −0.203) for any intervention compared with control (Table 1). However, this model was unable to identify the component effects, and based on the deviance information criteria (Table E1, Appendix E/Appendix on the journal's web site at www.elsevier.com) and clinical advice, model 2 was selected as the most appropriate model. From this model, P, S, B, C, and R were all identified as reducing length of stay although the CrIs for all components included zero suggesting that no single component was effective on its own (Table 1). Similarly, although E could increase length of stay, the CrI is wide and includes zero reflecting the uncertainty around the direction of effect. In addition, there was evidence of heterogeneity in this model ($\tau = 0.812$). The combinations P + S + B (MD: −0.956, 95% CrI: −1.619, −0.352) and P + S + R (MD: −1.016, 95% CrI: −1.998, −0.052) both had CrIs excluding zero and reduced length of stay by approximately 1 day each.
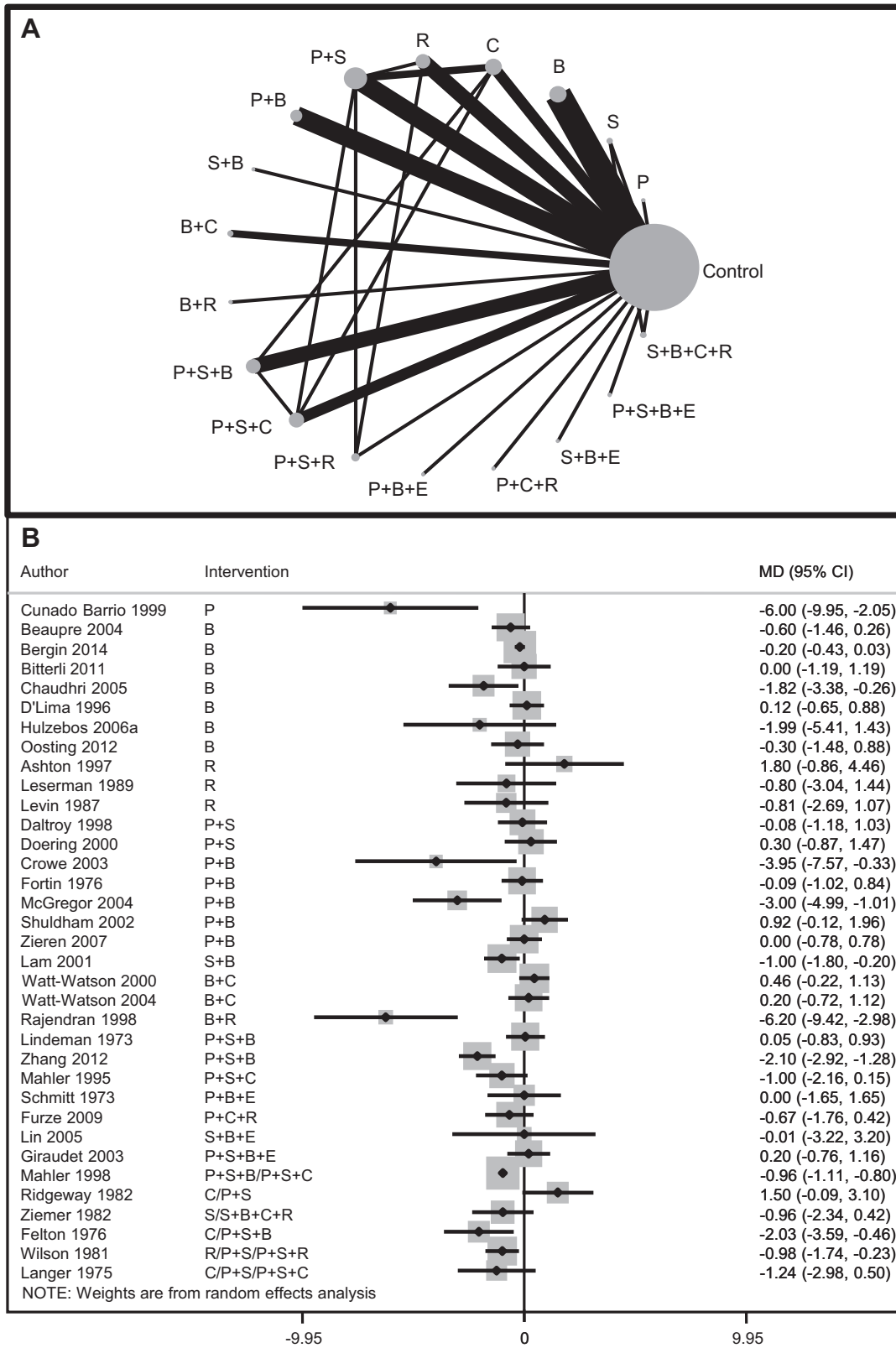
Fig. 1. (A) Length of stay network diagram. Node size is proportional to the number of patients randomized to each intervention, and line thickness is proportional to the number of trials directly comparing interventions. (B) Length of stay forest plot. P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques; MD, mean difference (days); CI, confidence interval.

Table 1. Posterior mean and 95% credible intervals for intervention effect for model 1 and component effects for model 2

Outcome	Summary	Model 1 (single effect) Intervention effect		Model 2 (additive main effects) Procedural information (P)		Sensory information (S)		Behavioral instruction (B)		Cognitive intervention (C)		Relaxation (R)		Emotion-focused techniques (E)	
		Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI
Length of stay (days)	MD	−0.541	*−0.913, −0.203*	−0.214	−0.988, 0.506	−0.322	−1.075, 0.507	−0.420	−0.973, 0.062	−0.020	−0.690, 0.680	−0.480	−1.349, 0.356	0.924	−0.521, 2.399
Pain	SMD	−0.145	*−0.274, −0.016*	0.103	−0.150, 0.352	−0.074	−0.338, 0.192	−0.054	−0.237, 0.125	0.039	−0.249, 0.329	−0.277	*−0.490, −0.058*	−0.388	−0.842, 0.074

Abbreviations: MD, mean difference; SMD, standardized mean difference; rI, credible interval.
Use of italics indicates a credible interval excluding zero.

3.2. Pain

For postoperative pain, we considered a network of 36 trials assessing nineteen different combinations of components (Figure 2). To fit model 1, which compared any intervention with control, we had to exclude the one trial, which compared two different forms of intervention without a control arm [29]. Model 1 identified evidence of a reduction in pain of −0.145 standardized units (95% CrI: −0.274, −0.016) for any intervention vs. control (Table 1). Based on the deviance information criteria (Table E1, Appendix E/Appendix on the journal's web site at www.elsevier.com) and clinical advice, model 2 was selected as the most appropriate model. From this model, we identified some evidence of heterogeneity ($\tau = 0.241$). R was identified as the most effective component reducing pain on average, by −0.277 and had a CrI excluding zero (95% CrI: −0.490, −0.058). The greatest reduction in pain was for the component E although the CrI included zero (SMD: −0.388; 95% CrI: −0.842, 0.074). The combinations B + R, S + R, S + B + E, and S + B + R all had CrIs excluding zero.

3.3. Inclusion of covariates: length of stay

The mean control group length of stay across all trials was 9.4 days. Model 2 was fitted including control group mean as a common effect across all components. In this model, parameters represent the mean difference for a trial with a control group length of stay of 9.4 days. For every 1 day increase in the control group length of stay, there is a mean reduction of −0.098 days (95% CrI: −0.157, −0.043) in the intervention group length of stay (Table 2). Components P, S, B, C, and R were all identified as reducing length of stay, whereas component E increased the length of stay. However, B was the only component with a CrI excluding zero. B reduced the length of stay in the intervention group by a mean difference of −0.561 days (95% CrI: −1.047, −0.111) when the control group length of stay was 9.4 days (Table 2, Figure 3). The inclusion of control group mean slightly reduced the amount of heterogeneity in the model ($\tau = 0.759$). Figure E2/Appendix on the journal's web site at www.elsevier.com presents estimates of the reduction in length of stay for combinations of control group length of stay and intervention. As the control group length of stay increased, the benefit of interventions in reducing length of stay increased (Figure E1, Appendix E/Appendix on the journal's web site at www.elsevier.com). For example, component R did not, on average, reduce length of stay when the control group length of stay was only 5 days but reduced length of stay by almost 1 day for a control group length of stay of 15 days (Table E2, Appendix E/Appendix on the journal's web site at www.elsevier.com).

Ten trials considered CV surgery, nine orthopedic surgery, and seventeen other types of surgery. Length of stay in the control groups of CV trials ranged from 5.1 to 9.9 days, in the orthopedic trials from 2.7 to 18 days, and in the other trials from 2.1 to 18.6 days. Model 2 was fitted including an independent effect for each component for type of surgery. Components P, S, and B were identified as having different effects for different types of surgery (Table E3 & Table E4, Appendix E/Appendix on the journal's web site at www.elsevier.com).

We included both control group mean and surgery within model 2. From this model control group mean was identified as being important, reducing length of stay by −0.099 days (95% CrI: −0.164, −0.038) for every 1 day increase in control group length of stay (Table 3). The model suggested that the effect of S differed for CV surgery compared with both other and orthopedic surgery (Table 3). For a control group length of stay of 5 days, S reduced length of stay in CV trials by 1.5 days. This reduction increased as the control group length of stay increased (Table 4). The effect of B differed for CV surgery compared with other surgery (Table 3). With a control group length of stay of 5 days, there was evidence that B was only effective in reducing length of stay for other types of surgery. However, when control group length of stay increased to 15 days, B reduced length of stay for all three types of surgery (Table 4). In contrast even with a control group length of stay of 15 days, S did not reduce length of stay for orthopedic surgery (Table 4). Including both control group mean and type of surgery reduced the amount of heterogeneity in the model further ($\tau = 0.535$).
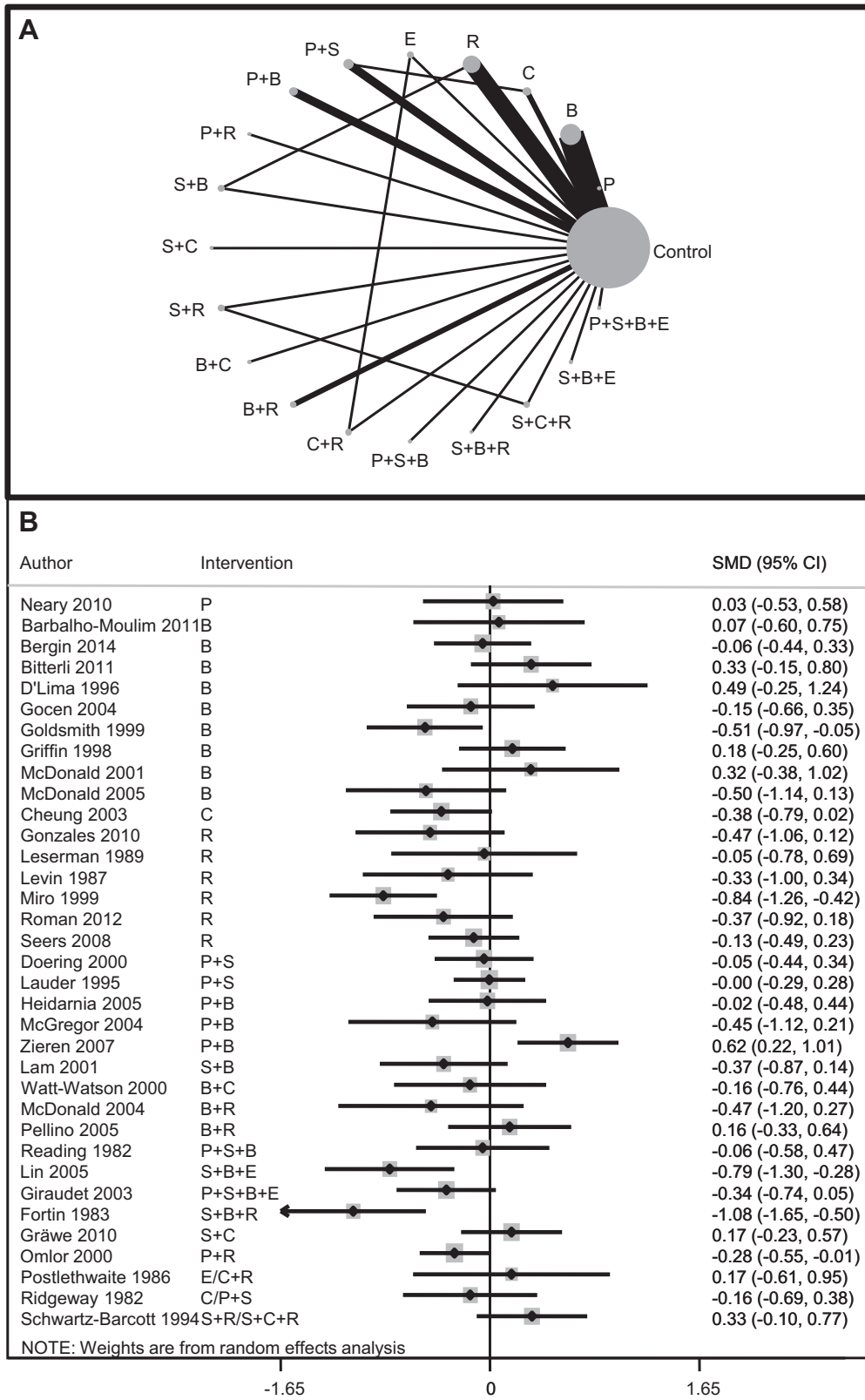
Fig. 2. (A) Pain network diagram. Node size is proportional to the number of patients randomized to each intervention, and line thickness is proportional to the number of trials directly comparing interventions. (B) Pain forest plot. P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques; SMD, standardized mean difference; CI, confidence interval.

Table 2. Posterior mean and 95% credible intervals for control group mean and component effects for model 2 including control group mean

Outcome	Summary	Control group mean		Procedural information (P)		Sensory information (S)		Behavioral instruction (B)		Cognitive intervention (C)		Relaxation (R)		Emotion-focused techniques (E)	
		Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI
Length of stay (days)	MD	−0.098	−0.157, −0.043	−0.308	−1.022, 0.360	−0.313	−1.004, 0.439	−0.561	−1.047, −0.111	−0.444	−1.106, 0.214	−0.368	−1.154, 0.412	1.296	−0.028, 2.700
Pain	SMD	−0.007	−0.057, 0.042	0.121	−0.162, 0.399	−0.078	−0.362, 0.207	−0.052	−0.257, 0.153	0.069	−0.270, 0.411	−0.274	−0.509, −0.035	−0.376	−0.866, 0.123

Abbreviations: MD, mean difference; SMD, standardized mean difference; CrI, credible interval.
Use of italics indicates a credible interval excluding zero.

For CV surgery, Figure E2 (Appendix E/Appendix on the journal's web site at www.elsevier.com) suggests that components S and E have the greatest probability of being the most effective components for reducing length of stay. However, the wide confidence intervals for E in both Table E3 (Appendix E/Appendix on the journal's web site at www.elsevier.com) and Table 4 highlight the fact that there are no CV trials which considered the component E. Likewise, there are no orthopedic trials which considered the components C and R. These results are driven by the prior distributions from the model and should be interpreted with caution (and hence their vast 95% CrIs). Following C and R, the component with the greatest probability of reducing length of stay in orthopedic trials was P. However, Table 4 shows that for P to be effective at reducing length of stay in orthopedic trials, the control group length of stay needs to be high. For other types of surgery, B had the greatest probability of being the most effective component.

3.4. Inclusion of covariates: pain

Model 2 was fitted including a common covariate effect of control group mean. The CrI for control group mean included zero (mean: −0.007, 95% CrI: −0.057, 0.042) (Table 2). However, when control group mean is included in the model, R is still the only component with a CrI excluding zero.
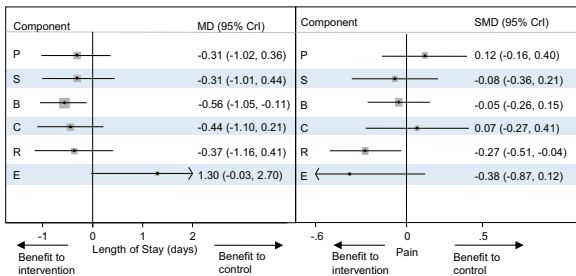

Fig. 3. Forest plot of component effects for length of stay and pain from model 2 including control group mean. P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques; MD, mean difference; SMD, standardized mean difference; CrI, credible interval.

Five trials considered CV surgery, twelve considered orthopedic surgery, and 20 considered other types of surgery. Model 2 was fitted including an independent effect for each component for type of surgery. However, there was no evidence to suggest that component effects differed by type of surgery (Table E5, Appendix E/Appendix on the journal's web site at www.elsevier.com).

3.5. Simultaneous assessment across the multiple outcomes

For each outcome, Figure 4 shows the probability that a particular component is the most effective. E had the greatest probability of being the most effective component for improving pain but had little effect on length of stay or negative affect (also shown in Figure E3, Appendix E/Appendix on the journal's web site at www.elsevier.com). However, in Table 1, R was the only component with a CrI excluding zero for pain and E had a wide CrI including zero. The probability of a component being the most effective fails to take into account the uncertainty surrounding the component effect. Therefore, in this case, the probability of E should be interpreted with caution and more emphasis placed on the component effect where the uncertainty surrounding the component effect can be quantified by the CrI [33].

In Table 3 and Figure E2 (Appendix E/Appendix on the journal's web site at www.elsevier.com), we identified that the most effective component for reducing length of stay depends on the type of surgery. Therefore, in Figure 4, the probability of being the most effective component is similar for P, S, B, C, and R. This is also shown in Figure E3 (Appendix E/Appendix on the journal's web site at www.elsevier.com), where these components have similar ranking profiles.

The component which appeared to be the most consistent across all three outcomes was R which had a probability of approximately 20% of being the best component for each outcome (Figure 4). However, both Figure 4 and Figure E3 (Appendix E/Appendix on the journal's web site at www.elsevier.com) show that no one component can be identified as the most effective component across all three outcomes.

Table 3. Length of stay: Posterior mean and 95% credible intervals for effect of control group mean, surgery, and component effects for model 2 including control group mean and surgery

Parameter	Model 2 including control group mean and type of surgery	
	Mean	95% CrI
Control group mean	−0.099	−0.164, −0.038
Effect of CV vs. other (P)	0.445	−1.140, 2.059
Effect of orthopedic vs. other (P)	−1.851	−4.329, 0.545
Effect of orthopedic vs. CV (P)	−2.295	−4.871, 0.193
Effect of CV vs. other (S)	−2.012	−3.729, −0.350
Effect of orthopedic vs. other (S)	2.497	−0.217, 5.307
Effect of orthopedic vs. CV (S)	4.509	1.654, 7.493
Effect of CV vs. other (B)	1.218	0.286, 2.183
Effect of orthopedic vs. other (B)	0.779	−0.099, 1.785
Effect of orthopedic vs. CV (B)	−0.439	−1.294, 0.539
Effect of CV vs. other (C)	−0.385	−1.543, 0.764
Effect of orthopedic vs. other (C)	0.219	−195, 195
Effect of orthopedic vs. CV (C)	0.604	−195, 196
Effect of CV vs. other (R)	0.066	−1.536, 1.778
Effect of orthopedic vs. other (R)	0.222	−196, 196
Effect of orthopedic vs. CV (R)	0.157	−196, 196
Effect of CV vs. other (E)	0.297	−196, 197
Effect of orthopedic vs. other (E)	−2.725	−5.905, 0.263
Effect of orthopedic vs. CV (E)	−3.023	−199, 193
P	−0.026	−0.864, 0.760
S	0.044	−0.702, 0.841
B	−1.199	−1.925, −0.521
C	−0.074	−1.021, 0.856
R	−0.612	−1.435, 0.208
E	2.106	0.239, 4.027

Abbreviations: P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques; CV, cardiovascularl; CrI, credible interval.

4. Discussion

The use of NMA has grown considerably in the last couple of decades, but the full potential of this flexible methodology has perhaps yet to be realized. One such extension, the CNMA approach of Welton et al. [5] has been little used to date but offers the potential to separate out component effects and thus has the potential to produce more clinically relevant results. This approach could be more widely used in systematic reviews involving complex interventions.

We have performed detailed further analyses of the data from a review on psychological preparation for adults undergoing surgery under general anesthesia [9], using CNMA methods to extend our understanding of the data and identify the most effective components for reducing length of stay, pain, and negative affect. The original Cochrane review identified that, on average, across all combinations of interventions evaluated, intervention was better than control for improving all three outcomes, but the effect of individual components was less clear given that the majority of the evidence-base concerned trials where these elements were delivered in combination.

The CNMA results support the findings of the original review in suggesting that any of the investigated types of psychological preparation may reduce length of stay in hospital by around half a day, an effect which may not only be clinically important but which may also represent substantial resource savings for health providers. Our results suggest that all psychological components, other than emotion-focused interventions may contribute to this benefit, with the strongest evidence for behavioral instruction.

Our results also confirmed benefits of any psychological intervention on both pain and negative affect (with effect sizes that may be considered low to moderate on a standardized scale [34]). There was evidence that relaxation may improve postoperative pain. However, emotion-focused techniques were ranked as the most likely to be of benefit despite the uncertainty surrounding the estimate of effectiveness. In addition, although E could be beneficial for pain, there was a possibility of a negative effect on length of stay. There was insufficient evidence that individual components contributed to the overall reduction in negative affect, but procedural and sensory information emerged as the most likely beneficial components. Overall, we were unable to identify any one component as the most effective across all three outcomes.

In clinical practice, the results support the use of psychological preparation for surgery to achieve emotional, pain, and resource use benefits. In particular, relaxation has been shown to be beneficial for reducing pain, and there is some indication that procedural information, sensory information, and behavioral instruction may be effective at reducing hospital stay for different types of surgery. In addition, combining components could maximize the benefit of receiving psychological intervention. However, some components and combinations have not been sufficiently evaluated to draw conclusions. Furthermore, greater gains are made when normal care results in longer length of stay or more negative affect, confirming the earlier findings of de Bruin et al. [20] in the context of HIV adherence. However, given the considerable uncertainty and residual heterogeneity in the CNMAs, further studies would be required before more detailed recommendations can be given. To that end, using a component-based model helps to inform specific gaps in the research and could potentially be used formally to inform which combinations of components should be evaluated in future trials [35].

Before this analysis, there was uncertainty surrounding which CNMA model would be the most clinically meaningful. Therefore we used the deviance information criteria as a guide for choosing the most appropriate model alongside clinical expertise. Clinical expertise was important because deviance information criteria is a statistical

Table 4. Predictions for reduction in length of stay for combinations of control group length of stay, type of surgery, and intervention

Control group length of stay (days)	Type of surgery	Intervention	Mean change in length of stay (days) with intervention (95% CrI)
5	CV	P	0.845 (−0.584, 2.251)
5	CV	S	−1.535 (−2.971, −0.059)
5	CV	B	0.445 (−0.236, 1.123)
5	CV	C	−0.028 (−0.725, 0.643)
5	CV	R	−0.114 (−1.532, 1.410)
5	CV	E	2.553 (−194, 199)
5	Orthopedic	P	−1.469 (−3.864, 0.831)
5	Orthopedic	S	2.987 (0.473, 5.588)
5	Orthopedic	B	0.012 (−0.580, 0.721)
5	Orthopedic	C	0.446 (−196, 196)
5	Orthopedic	R	0.0006 (−196, 196)
5	Orthopedic	E	−0.189 (−2.458, 1.994)
5	Other	P	0.405 (−0.372, 1.145)
5	Other	S	0.475 (−0.387, 1.405)
5	Other	B	−0.772 (−1.489, −0.089)
5	Other	C	0.348 (−0.588, 1.283)
5	Other	R	−0.187 (−1.035, 0.682)
5	Other	E	2.516 (0.497, 4.584)
15	CV	P	−0.137 (−1.559, 1.214)
15	CV	S	−2.517 (−4.096, −0.954)
15	CV	B	−0.538 (−1.407, 0.265)
15	CV	C	−1.010 (−2.010, −0.101)
15	CV	R	−1.096 (−2.554, 0.426)
15	CV	E	1.571 (−195, 198)
15	Orthopedic	P	−2.451 (−4.565, −0.421)
15	Orthopedic	S	2.005 (−0.719, 4.791)
15	Orthopedic	B	−0.970 (−1.749, −0.151)
15	Orthopedic	C	−0.536 (−197, 195)
15	Orthopedic	R	−0.981 (−197, 195)
15	Orthopedic	E	−1.171 (−3.604, 1.121)
15	Other	P	−0.577 (−1.615, 0.387)
15	Other	S	−0.507 (−1.242, 0.253)
15	Other	B	−1.754 (−2.632, −0.949)
15	Other	C	−0.634 (−1.707, 0.413)
15	Other	R	−1.169 (−2.097, −0.269)
15	Other	E	1.534 (−0.212, 3.325)

Abbreviations: P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques; CV, cardiovascular; CrI, credible interval.

measure of model fit, which identifies the model with the best fit to the data which may not always be the most clinically meaningful model. A 2003 study showed that only 6% of trials detected statistically significant treatment interactions between components [8,36]. We assessed pairwise interactions between components through model 3, which relaxes the additivity assumption for combinations of two components but not for combinations of three or more components. With few studies evaluating each component,

model 3 could be underpowered to identify any interactions. However, we felt it was important to consider the possibility of interactions although we were limited to considering pairwise interactions only, and this should be kept in mind when interpreting the results. For all three outcomes, the additive main effects model was selected as the most appropriate model. The additive effects model assumes that the effect of each component is independent of any other components it might be combined with. When
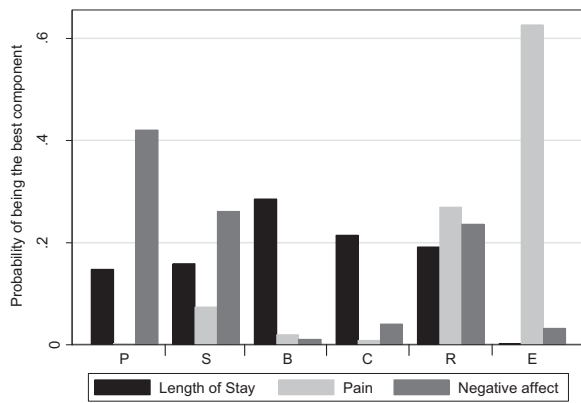
Fig. 4. Probability of being the most effective component for reducing length of stay, pain, and negative affect from model 2 including control group mean. P, procedural information; S, sensory information; B, behavioral instruction; C, cognitive intervention; R, relaxation; E, emotion-focused techniques.

the assumption of additivity is valid, then the additive effects model can increase the precision in the estimates of the component effects. However, when the assumption is violated, the component effects may be biased. Previous work by Thorlund et al. has shown that the additive effects model is valid as long as approximate additivity can be assumed [8].

A key assumption of NMA (and pairwise MA) is similarity in the definition of components across trials [37]. In this review, there was considerable variation in the definitions of interventions and outcomes, for example "behavioral instruction" could refer to strengthening exercises the patient is told to carry out before surgery, or instructions about using equipment postsurgery; exact measures for pain and negative affect, and the timing of assessment, varied across studies. For all three outcomes, there was some evidence of important statistical heterogeneity, and this was only partially explained by the inclusion of specific component effects and other covariates. For example, the heterogeneity in types of surgery was reflected in the variation in the mean length of stay. However, when type of surgery was taken into account heterogeneity remained in the model.

For both length of stay and negative affect, we identified control group risk as an important covariate for improving outcomes. However, a limitation of this analysis is that control group results will depend in part on baseline characteristics and in part on the components of standard care provided to the control participants. In the original Cochrane review, it was noted that control groups were often underspecified and referred to "treatment as usual" in the primary studies. In addition, "treatment as usual" was likely to have included some procedural information and behavioral instruction, but it would have been impossible to quantify. Nevertheless, where authors did specify intervention components as part of what the control group received, it had to be clear that the intervention group was receiving more of that component than the control group for that component to be coded as being present.

A main assumption of NMA is consistency of the direct and indirect evidence. One way to check consistency is through loops in the network where the treatment effect from the direct evidence can be compared with the treatment effect from the indirect evidence [38−40]. In this article, our networks were generally star shaped with the control group at the center of the "star" and head-to-head comparisons between interventions mainly from multi-arm trials. Therefore, we did not have treatment loops in which we could assess the assumption of network consistency. In addition, some comparisons were only informed by direct evidence from one trial. Indirect comparisons can require up to four times as much information as direct comparisons for the same precision. Therefore, it is possible that we may have been unable to identify some important effects due to lack of power [6].

In this article, we have fitted the CNMA models using a Bayesian framework. This has several advantages including the ability to estimate the probability that each component will be the best for each outcome. In addition, we were able to include the covariate baseline risk, correctly adjusting for potential regression to the mean due to measurement error, by using arm-based data, while synthesizing on the standardized mean difference scale; which we believe is a first for this sort of component modeling.

The analysis of complex reviews of interventions with multiple components remains challenging, particularly when time and resources are limited. Often only narrative synthesis may be possible, but when sufficient data for meta-analysis are available, there are important limitations when trying to interpret the results of multiple pairwise meta-analyses. In this article, we have illustrated the potential benefits of the more sophisticated but rarely-used CNMA methods, which we have successfully used to help elucidate the long-term effects of individual psychological components on surgery. We encourage wider adoption of these techniques and further methodological research in this area.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2018.02.012.

References

[1] Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655.

[2] Lewin S, Hendry M, Chandler J, Oxman AD, Michie S, Shepperd S, et al. Assessing the complexity of interventions within systematic reviews: development, content and use of a new tool (iCAT_SR). BMC Med Res Methodol 2017;17:76.

[3] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004;23:3105−24.

[4] Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331:1897−900.

[5] Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. Am J Epidemiol 2009; 169:1158−65.

[6] Mills EJ, Druyts E, Ghement I, Puhan MA. Pharmacotherapies for chronic obstructive pulmonary disease: a multiple treatment comparison meta-analysis. Clin Epidemiol 2011;3:107−29.

[7] Mills EJ, Thorlund K, Ioannidis JP. Calculating additive treatment effects from multiple randomized trials provides useful estimates of combination therapies. J Clin Epidemiol 2012;65:1282−8.

[8] Thorlund K, Mills E. Stability of additive treatment effects in multiple treatment comparison meta-analysis: a simulation study. Clin Epidemiol 2012;4:75−85.

[9] Powell R, Scott NW, Manyande A, Bruce J, Vogele C, Byrne-Davis LM, et al. Psychological preparation and postoperative outcomes for adults undergoing surgery under general anaesthesia. Cochrane Database Syst Rev 2016;CD008646.

[10] Johnston M, Vögele C. Benefits of psychological preparation for surgery: a meta-analysis. Ann Behav Med 1993;15(4):245−56.

[11] Arpino L, Iavarone A, Parlato C, Moraci A. Prognostic role of depression after lumbar disc surgery. Neurol Sci 2004;25(3):145−7.

[12] Bruce J, Thornton AJ, Scott NW, Marfizo S, powell R, Johnston M, et al. Chronic preopoerative pain and psychological robustness predict acute postoperative pain outcomes after surgery for breast cancer. Br J Cancer 2012;107:937−46.

[13] Granot M, Ferber SG. The roles of pain catastrophizing and anxiety in the prediction of postoperative pain intensity - a prospective study. Clin J Pain 2005;21(5):429−45.

[14] Munafò MR, Stevenson J. Anxiety and surgical recovery - Reinterpreting the literature. J Psychosomatic Res 2001;51(4):589−96.

[15] Rainville P, Bao QVH, Chrétien P. Pain-related emotions modulate experimental pain perception and autonomic responses. Pain 2005; 118:306−18.

[16] van Middendorp H, Lumley M, Jacobs JWG, Bijilsma JWJ, Geenen R. The effects of anger and sadness on clinical pain reports and experimentally-induced pain thresholds in women with and without fibromyalgia. Arthritis Care Res (Hoboken) 2010;62(10):1370−6.

[17] Maple H, Joseph C, Lee V, Simmonds S, Weinman J, Mamode N. Stress predicts the trajectory of wound healing in living kidney donors as measured by high-resolution ultrasound. Brain Behav Immun 2015;43:19−26.

[18] Walburn J, Vedhara K, Hankins M, Rixon L, Weinman J. Psychological stress and wound healing in humans: a systematic review and meta-analysis. J Psychosomatic Res 2009;67(3):253−71.

[19] Achana FA, Cooper NJ, Dias S, Lu G, Rice SJC, Kendrick D, et al. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. Stat Med 2013;32:752−71.

[20] de Bruin M, Viechtbauer W, Hospers HJ, Schaalma HP, Kok G. Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: implications for the interpretation and comparison of intervention effects. Health Psychol 2009; 28(6):668−74.

[21] Ma YI, Qin LJ, Han ZF. Relaxation training on stress response to abdominal surgery [[original title in Chinese script]]. Chin J Nurs (Zhonghua Hu Li Za Zhi) 1996;31(7):377−80.

[22] Guo P, East L, Arthur A. A preoperative education intervention to reduce anxiety and improve recovery among Chinese cardiac patients: a randomized controlled trial. Int J Nurs Stud 2012;49(2): 129−37.

[23] Ashton C Jr, Whitworth GC, Seldomridge JA, Shapiro PA, Weinberg AD, Michler RE, et al. Self-hypnosis reduces anxiety following coronary artery bypass surgery. A prospective, randomized trial. J Cardiovasc Surg 1997;38(1):69−75.

[24] Felton G, Huss K, Payne EA, Srsic K. Preoperative nursing intervention with the patient for surgery: outcomes of three alternative approaches. Int J Nurs Stud 1976;13(2):83−96.

[25] Giraudet-Le Quintrec J, Coste J, Vastel L, Pacault V, Jeanne L, Lamas J, et al. Positive effect of patient education for hip surgery: a randomized trial. Clin Orthop Relat Res 2003;(414):112−20.

[26] Lim L, Chow P, Wong CY, Chung A, Chan YH, Wong WK, et al. Doctor-patient communication, knowledge, and question prompt lists in reducing preoperative anxiety - a randomized control study. Asian J Surg 2011;34(4):175−80.

[27] Bitterli R, Sieben JM, Hartmann M, De Bruin ED. Pre-surgical sensorimotor training for patients undergoing total hip Replacement: a randomised controlled trial. Int J Sports Med 2011;32(9):725−32.

[28] Coslow BIF, Eddy ME. Effects of preoperative ambulatory gynecological education: clinical outcomes and patient satisfaction. J PeriAnesthesia Nurs 1998;13(1):4−10.

[29] Wells N. The effect of relaxation on postoperative muscle tension and pain. Nurs Res 1982;31(4):236−8.

[30] Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester (UK): John Wiley & Sons; 2008.

[31] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controled trials. Sheffield, UK: NICE Decision Support Unit, University of Sheffield; 2011: last updated September 2016.

[32] Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163−71.

[33] Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. BMJ 2013;346:f2914.

[34] Cohen J. Statistical power analysis for the behavioural sciences. Second edition. USA: Lawrence Erlbaum Associates; 1988.

[35] Nikolakopoulou A, Mavridis D, Salanti G. Planning future studies based on the precision of network meta-analysis results. Stat Med 2016;35:978−1000.

[36] McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials - a Systematic review. JAMA 2003;289: 2545−53.

[37] Del Giovane C, Vacchi L, Mavridis D, Filippini G, Salanti G. Network meta-analysis models to account for variability in treatment definitions: application to dose effects. Stat Med 2013;32:25−39.

[38] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997;50:683−91.

[39] Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. J Am Stat Assoc 2006;101:447−59.

[40] Salanti G, Higgins JPT, Ioannidis JPA. Evaluation of networks of randomized trials. Stat Methods Med Res 2007;17:279−301.