Cochrane Library

Cochrane Database of Systematic Reviews

Psychosocial interventions for recurrent abdominal pain in childhood (Review)

Abbott RA, Martin AE, Newlove-Delgado TV, Bethel A, Thompson-Coon J, Whear R, Logan S

Abbott RA, Martin AE, Newlove-Delgado TV, Bethel A, Thompson-Coon J, Whear R, Logan S.
Psychosocial interventions for recurrent abdominal pain in childhood.
Cochrane Database of Systematic Reviews 2017, Issue 1. Art. No.: CD010971.
DOI: 10.1002/14651858.CD010971.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER . . . 1
ABSTRACT . . . 1
PLAIN LANGUAGE SUMMARY . . . 2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON . . . 4
BACKGROUND . . . 7
OBJECTIVES . . . 8
METHODS . . . 8
RESULTS . . . 12
Figure 1. . . . 13
Figure 2. . . . 17
Figure 3. . . . 18
ADDITIONAL SUMMARY OF FINDINGS . . . 25
DISCUSSION . . . 29
AUTHORS' CONCLUSIONS . . . 31
ACKNOWLEDGEMENTS . . . 31
REFERENCES . . . 32
CHARACTERISTICS OF STUDIES . . . 39
DATA AND ANALYSES . . . 63
ADDITIONAL TABLES . . . 64
WHAT'S NEW . . . 66
HISTORY . . . 66
CONTRIBUTIONS OF AUTHORS . . . 67
DECLARATIONS OF INTEREST . . . 67
SOURCES OF SUPPORT . . . 68
DIFFERENCES BETWEEN PROTOCOL AND REVIEW . . . 68
NOTES . . . 68

[Intervention Review]

Psychosocial interventions for recurrent abdominal pain in childhood

Rebecca A Abbott[1], Alice E Martin[2], Tamsin V Newlove-Delgado[1], Alison Bethel[1], Joanna Thompson-Coon[1], Rebecca Whear[1], Stuart Logan[1]

[1]NIHR CLAHRC South West Peninsula (PenCLAHRC), University of Exeter Medical School, Exeter, UK. [2]Paediatrics, Royal Devon and Exeter Hospital, Exeter, UK

Contact address: Rebecca A Abbott, NIHR CLAHRC South West Peninsula (PenCLAHRC), University of Exeter Medical School, South Cloisters, St Luke's Campus, Exeter, England, EX1 2LU, UK. r.a.abbott@exeter.ac.uk, rebecca.abbott@pms.ac.uk.

Editorial group: Cochrane Developmental, Psychosocial and Learning Problems Group.
Publication status and date: New, published in Issue 1, 2017.

Citation: Abbott RA, Martin AE, Newlove-Delgado TV, Bethel A, Thompson-Coon J, Whear R, Logan S. Psychosocial interventions for recurrent abdominal pain in childhood. *Cochrane Database of Systematic Reviews* 2017, Issue 1. Art. No.: CD010971. DOI: 10.1002/14651858.CD010971.pub2.

ABSTRACT

Background

This review supersedes the original Cochrane review first published in 2008 (Huertas-Ceballos 2008).

Between 4% and 25% of school-aged children complain of recurrent abdominal pain (RAP) severe enough to interfere with their daily activities. No organic cause for this pain can be found on physical examination or investigation for the majority of such children. Although many children are managed by reassurance and simple measures, a large range of psychosocial interventions involving cognitive and behavioural components have been recommended.

Objectives

To determine the effectiveness of psychosocial interventions for reducing pain in school-aged children with RAP.

Search methods

In June 2016 we searched CENTRAL, MEDLINE, Embase, eight other databases, and two trials registers. We also searched the references of identified studies and relevant reviews.

Selection criteria

Randomised controlled trials comparing psychosocial therapies with usual care, active control, or wait-list control for children and adolescents (aged 5 to 18 years) with RAP or an abdominal pain-related functional gastrointestinal disorder defined by the Rome III criteria were eligible for inclusion.

Data collection and analysis

We used standard methodological procedures expected by Cochrane. Five review authors independently selected studies, assessed them for risk of bias, and extracted relevant data. We also assessed the quality of the evidence using the GRADE approach.

Main results

This review includes 18 randomised controlled trials (14 new to this version), reported in 26 papers, involving 928 children and adolescents with RAP between the ages of 6 and 18 years. The interventions were classified into four types of psychosocial therapy: cognitive behavioural therapy (CBT), hypnotherapy (including guided imagery), yoga, and written self-disclosure. The studies were carried out in the USA, Australia, Canada, the Netherlands, Germany, and Brazil. The majority of the studies were small and short term; only two studies included more than 100 participants, and only five studies had follow-up assessments beyond six months. Small sample sizes and the degree of assessed risk of performance and detection bias in many studies led to the overall quality of the evidence being rated as low to very low for all outcomes.

For CBT compared to control, we found evidence of treatment success postintervention (odds ratio (OR) 5.67, 95% confidence interval (CI) 1.18 to 27.32; Z = 2.16; P = 0.03; 4 studies; 175 children; very low-quality evidence), but no evidence of treatment success at medium-term follow-up (OR 3.08, 95% CI 0.93 to 10.16; Z = 1.85; P = 0.06; 3 studies; 139 children; low-quality evidence) or long-term follow-up (OR 1.29, 95% CI 0.50 to 3.33; Z = 0.53; P = 0.60; 2 studies; 120 children; low-quality evidence). We found no evidence of effects of intervention on pain intensity scores measured postintervention (standardised mean difference (SMD) -0.33, 95% CI -0.74 to 0.08; 7 studies; 405 children; low-quality evidence), or at medium-term follow-up (SMD -0.32, 95% CI -0.85 to 0.20; 4 studies; 301 children; low-quality evidence).

For hypnotherapy (including studies of guided imagery) compared to control, we found evidence of greater treatment success postintervention (OR 6.78, 95% CI 2.41 to 19.07; Z = 3.63; P = 0.0003; 4 studies; 146 children; low-quality evidence) as well as reductions in pain intensity (SMD -1.01, 95% CI -1.41 to -0.61; Z = 4.97; P < 0.00001; 4 studies; 146 children; low-quality evidence) and pain frequency (SMD -1.28, 95% CI -1.84 to -0.72; Z = 4.48; P < 0.00001; 4 studies; 146 children; low-quality evidence). The only study of long-term effect reported continued benefit of hypnotherapy compared to usual care after five years, with 68% reporting treatment success compared to 20% of controls (P = 0.005).

For yoga therapy compared to control, we found no evidence of effectiveness on pain intensity reduction postintervention (SMD -0.31, 95% CI -0.67 to 0.05; Z = 1.69; P = 0.09; 3 studies; 122 children; low-quality evidence).

The single study of written self-disclosure therapy reported no benefit for pain.

There was no evidence of effect from the pooled analyses for any type of intervention on the secondary outcomes of school performance, social or psychological functioning, and quality of daily life.

There were no adverse effects for any of the interventions reported.

Authors' conclusions

The data from trials to date provide some evidence for beneficial effects of CBT and hypnotherapy in reducing pain in the short term in children and adolescents presenting with RAP. There was no evidence for the effectiveness of yoga therapy or written self-disclosure therapy. There were insufficient data to explore effects of treatment by RAP subtype.

Higher-quality, longer-duration trials are needed to fully investigate the effectiveness of psychosocial interventions. Identifying the active components of the interventions and establishing whether benefits are sustained in the long term are areas of priority. Future research studies would benefit from employing active control groups to help minimise potential bias from wait-list control designs and to help account for therapist and intervention time.

PLAIN LANGUAGE SUMMARY

Psychosocial therapy for recurrent abdominal pain in childhood

Review question

Do psychosocial therapies reduce pain in children and adolescents with recurrent abdominal pain?

Background

Between 4% and 25% of school-aged children complain of recurrent abdominal pain severe enough to interfere with their daily activities. No organic cause for this pain can be found on physical examination or investigation for the majority of such children. Although many

children are managed by reassurance and simple measures, a large range of psychological and behavioural ('psychosocial') therapies have been recommended.

Methods and study characteristics

As of June 2016, we identified 18 randomised controlled trials (a type of scientific experiment in which people are randomly assigned to one of two or more treatments), which included 928 children and adolescents between the ages of 6 and 18 years. These studies compared a range of psychosocial therapy to usual care or some form of non-therapy control (such as education or breathing exercises). We identified four different kinds of psychosocial therapy: cognitive behavioural therapy, hypnotherapy, yoga, and written self-disclosure (a therapy that involves writing down thoughts and feelings about something distressing). The duration of the included studies ranged from five days to three months. The studies were conducted in the USA, Australia, Canada, the Netherlands, Germany, and Brazil.

Key results

We found that cognitive behavioural therapy and hypnotherapy may be effective in terms of reducing pain in the short term. There was little evidence of long-term benefit. There was no evidence that either therapy had a beneficial effect on quality of life, daily activities, or psychological outcomes such as anxiety and depression. Yoga therapy and written self-disclosure as a therapy had no effect on pain, quality of life, or daily activities. No adverse effects were reported from any of these therapies.

Quality of the evidence

We rated the overall quality of the evidence as low to very low for all outcomes. Many of the studies had small sample sizes or weaknesses in their study design. The authors reported no conflicts of interest in relation to funding.

Conclusion

Cognitive behavioural therapy and hypnotherapy warrant consideration by clinicians as part of the management strategy for children with recurrent abdominal pain. The overall quality of the evidence was low to very low. More high-quality research is needed to evaluate the particular aspects of the therapies that are effective and to establish whether benefits are maintained over time.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Cognitive behavioural therapy compared with control for children and adolescents with recurrent abdominal pain

Patient or population: children and adolescents with recurrent abdominal pain
Settings: mixed
Intervention: cognitive behavioural therapy
Comparison: usual care or wait-list control

Outcomes	Probable outcome with control or usual care	Probable outcome with CBT	OR (95% CI)	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
Treatment success: postintervention	**211 per 1000**	**494 per 1000**	Pooled OR 5.67 (1.18 to 27.32)	175 (4)	⊕◯◯◯ [1] **Very low**	Varied definitions of 'treatment success' used by authors (pain free; reduction of 10 points on API; Walker 1997).
Treatment success: medium-term follow-up (between 3 and 12 months)	**349 per 1000**	**551 per 1000**	Pooled OR 3.08 (0.93 to 10.16)	139 (3)	⊕⊕◯◯ [2] **Low**	Varied definitions of 'treatment success' used by authors (pain free; reduction of 10 points on API; Walker 1997).
Pain intensity: postintervention Lower score equals less pain.	The pain score in the CBT groups was, on average, **0.33 SDs lower** (95% CI -0.74 to 0.08) than in the usual care, wait-list, or education control groups	-		405 (7)	⊕⊕◯◯ [3] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference Varied measures used to assess pain intensity (FACES Pain Scale (Bieri 1990); visual analogue scale;

					Likert scale).
Pain intensity: medium-term follow-up (between 3 and 12 months) Lower score equals less pain.	The pain score in the CBT groups was, on average, **0.32 SDs lower** (95% CI -0.85 to 0.20) than in the usual care, wait-list, or education control groups	-	301 (4)	⊕⊕◯◯ [4] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference
QOL (physical subscale): postintervention Higher score equals better QOL.	The QOL score (physical subscale) in the CBT groups was, on average, **0.71 SDs higher** (95% CI -0.25 to 1.66) than in the usual care, wait-list, or education control groups	-	136 (3)	⊕◯◯◯ [5] **Very low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference 2 studies used PedsQL (Varni 2001), 1 study used KIDSCREEN (Ravens-Sieberer 2005).
QOL (psychosocial subscale): postintervention Higher score equals better QOL.	The QOL score (psychosocial subscale) in the CBT groups was, on average, **0.43 SDs higher** (95% CI -0.21 to 1.06) than in the usual care, wait-list, or education groups	-	136 (3)	⊕⊕◯◯ [6] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference 2 studies used PedsQL (Varni 2001), 1 study used KIDSCREEN (Ravens-Sieberer 2005).
Functional disability or activity limitations: postintervention Lower score equals less activity disability.	Functional disability in the CBT groups was, on average, **0.57 SDs lower** (95% CI -1.34 to 0.19) than in the usual care, wait-list, or education control groups	-	176 (4)	⊕◯◯◯ [7] **Very low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference,

			and 0.8 SD a large difference 3 different functional disability or activity limitation indices used (KINDL-R (Ravens-Sieberer 2005); CALI (Palermo 2004; Palermo 2016); FDI (Walker 1991)).

API: Abdominal Pain Index; **CALI:** Child Activity Limitations Interview; **CBT**: cognitive behavioural therapy; **CI:** confidence interval; **FDI:** Functional Disability Inventory; **KIDSCREEN**: Health Related Quality of Life Questionnaire for Children and Young People; **KINDL-R:** measure of health-related quality of life; **OR:** odds ratio; **PedsQL:** Pediatric Quality of Life Inventory; **QOL**: quality of life; **SD:** standard deviation

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]Downgraded three levels: high risk of bias across the studies in study design and outcome assessment; high level of unexplained heterogeneity (> 70%); and a low number of participants included in the analysis, wide CIs.
[2]Downgraded two levels: high risk of bias across the studies in study design and a low number of participants included in the analysis, wide CIs.
[3]Downgraded two levels: high level of unexplained heterogeneity (> 70%) and high risk of bias across the studies, with baseline differences in primary outcomes in the largest study.
[4]Downgraded two levels: high level of unexplained heterogeneity (> 70%) and high risk of bias across the studies, with baseline differences in primary outcomes in the largest study.
[5]Downgraded three levels: high risk of bias across the studies in study design; high level of unexplained heterogeneity (> 70%); and a low number of participants included in the analysis, wide CIs.
[6]Downgraded two levels: high risk of bias across the studies in study design and a low number of participants included in the analysis, wide CIs.
[7]Downgraded three levels: high risk of bias across the studies in study design; high level of unexplained heterogeneity (> 70%); and a low number of participants included in the analysis, wide CIs.

BACKGROUND

Description of the condition

Recurrent abdominal pain (RAP) is a common problem in paediatric practice. Between 4% to 25% of school-aged children suffer at some point from RAP that interferes with their activities of daily living (Konijnenberg 2005; Williams 1996; Youssef 2006). The condition is related to school absences, hospital admissions and, on occasion, unnecessary surgical intervention (Scharff 1997; Størdal 2005; Walker 1998). Symptoms sometimes continue into adulthood (Apley 1975; Walker 1995; Youssef 2008). The abdominal pain is commonly associated with other symptoms, including headaches, recurrent limb pains, pallor, and vomiting (Abu-Arafeh 1995; Devanarayana 2011; Hyams 1995). RAP can cause significant anxiety in parents and carers, who may become overwhelmed by fear of serious disease and feel helpless because they are unable to relieve their child's pain (Paul 2013).

It is generally accepted that RAP in children represents a group of functional gastrointestinal disorders that have an unclear aetiology. Children suffer either chronic or recurrent gastrointestinal symptoms not explained by a structural, biochemical, or inflammatory process. Apley first sought to define the condition in the 1950s, and suggested that the diagnostic label should be based on the presence of at least three episodes of severe abdominal pain (often, but not necessarily, associated with systemic symptoms), over three months (Apley 1958), with no established organic cause. More recently, an international consensus definition with a symptom-based classification system has been created: the Rome III criteria, which has specific categories for paediatric presentations (Rasquin 2006). We have used RAP throughout this review as an umbrella term to refer to the four categories included within this classification, which are: functional dyspepsia, irritable bowel syndrome, abdominal migraine, and functional abdominal pain syndrome. The pain classification for each of the Rome III diagnoses is defined by at least one episode per week for at least two months; this varies from Apley's original definition of RAP (Apley 1958). The Rome classification is not based on known pathophysiological differences between the conditions, but rather on the constellation of clinical features. The extent to which separating children into these categories defines groups that are distinct clinical entities who are likely to respond differently to treatment remains unclear. Nonetheless, this classification has been welcomed following the historical use of diverse terms, some of which imply causation, including: "abdominal migraine" (Bain 1974; Hockaday 1992; Symon 1986), "abdominal epilepsy" (Stowens 1970), "the irritable bowel syndrome in childhood" (Stone 1970), "allergic-tension-fatigue syndrome" (Sandberg 1973; Speer 1954), "neurovegetative dystonia" (Peltonen 1970; Rubin 1967), "functional gastrointestinal disorder" (Drossman 1995), and "the irritated colon syndrome" (Harvey 1973; Painter 1964).

Description of the intervention

The focus of this review is any intervention based on psychological or behavioural theory (a 'psychosocial' intervention). A variety of approaches have been used, including behavioural and cognitive behavioural techniques (Sanders 1994; Scharff 1997), psychotherapy (Vasquez 1992), family-centred approaches (Liebman 1976; Walker 1999; Wetchler 1992), multicomponent therapies (Edwards 1991; Finney 1989; Hicks 2003; Humphreys 1998), and more recently, a variety of what has been termed 'mind-body approaches', such as guided imagery, yoga, and hypnotherapy (Weydert 2006). Specific to the interventions found in this review, cognitive behavioural therapy (CBT) may involve the family or may focus only on the child. CBT can be carried out on an individual basis or in a group format, and can be performed face-to-face or remotely through the use of CDs and DVDs. CBT involves the teaching of coping and distraction strategies and relaxation techniques; identification and change of pain-related thoughts; and modification of family responses to pain (Groβ 2013). Through hypnosis, hypnotherapy in group or individual sessions helps the patient to relax and think about controlling their pain and strengthens the patient's self efficacy in managing their pain (Vlieger 2007). Guided imagery can also be carried out on a group or individual basis. Guided imagery is similar in concept to hypnotherapy, and is considered to be a form of self regulation therapy, which aims to induce deep relaxation and facilitate the creation of images to help bring resolution to pain and symptoms (Weydert 2006). In yoga therapy, the patient learns a series of physical postures along with a daily practice of breathing and meditation techniques (Kuttner 2006). Written self-disclosure therapy, sometimes called expressive writing, involves being given the opportunity in a quiet space, to write down thoughts and feelings about something deeply distressing, on three to four occasions over a couple of weeks, with no subsequent discussion or follow-up. (Wallander 2011).

How the intervention might work

The aetiology of pain-related functional gastrointestinal disorders is unclear. It has been suggested that visceral hypersensitivity (Di Lorenzo 2001; Van Ginkel 2001), autonomic dysfunction (Good 1995), and gut dysmotility may contribute, and this may be initiated by an inflammatory, infective, traumatic, or allergic trigger (Mayer 2002; Milla 1999). As with any chronic pain condition, it is likely that psychological factors are important in both presentation and treatment. Many clinicians believe that abdominal pain-related functional gastrointestinal disorders originate from, or are contributed to, by psychogenic factors (Friedman 1972; Raymer 1984). Historically, authors have suggested that children with RAP come from "psychosomatic families" (Osborne 1989). A population-based study by Ramchandani 2006 found that anxiety in parents, which added to a specific child temperament before

one year of age, was a strong predictor of RAP in childhood.

Children with RAP have been found to score higher than other children on questionnaires assessing psychopathological symptoms, especially internalising disturbances such as anxiety and other somatic complaints (Dufton 2009). Children with RAP have also been shown to have a high rate of psychiatric disorders such as anxiety disorders and depression (Campo 2004; Shelby 2013). Further evidence of psychological factors contributing to presentation of unexplained abdominal pain comes from Campo 2001, who suggested a strong association between RAP in childhood and anxiety in adult life. Children who suffer from RAP are more likely to have poor coping strategies for stressful situations (Walker 2007), and depressive symptoms have been linked with a poor ability to cope with RAP (Kaminsky 2006). The varied approaches to treating RAP therefore work on reducing the combination of anxiety and depression, improving coping strategies, and recognising and understanding RAP symptomology. A brief description of how each of the interventions addressed in this review might work follows below.

CBT aims to improve the child's mental health and coping strategies, specifically in helping them to understand the onset and progress of their RAP. It then offers the child a strategy to help manage it, along with anxiety management and specific behavioural techniques (Groβ 2013). CBT may take a family approach. Family therapy seeks to alter environmental factors that might reinforce the child's pain behaviour within the family and to identify and treat factors that may precipitate it (Van Slyke 2006; Walker 2006).

The mode of action for how hypnotherapy may help RAP is not completely understood and is likely to be from a combination of effects on gastrointestinal motility, visceral sensitivity, psychological factors, and direct effects within the central nervous system (Vlieger 2007). Hypnotherapy and guided imagery, a related therapy, may bring about cognitive changes through directly influencing cognitions, which helps to improve symptoms, or through influencing pain and gut functioning, leading to a change in cognition (Vlieger 2012). Alternatively, they both may help reduce stress and anxiety, which results in concomitant changes in the hypothalamic-pituitary-adrenal axis (Kennedy 2012). Guided imagery is a form of self regulation therapy, which along with deep relaxation, helps the patient to create images to help resolve their problems (Weydert 2006). It is has been further hypothesised that communication through images, along with deep relaxation, reduces anxiety, which impacts both the voluntary and autonomic nervous system hyper-reactivity that contributes to pain (Lee 1996).

Most forms of yoga involve a series of physical postures along with breathing and meditation techniques that are intended to reduce anxiety, improve body tone, and increase feelings of well-being (Kuttner 2006). In adults, yoga has been shown to help manage back pain and migraine (Williams 2005). In the limited research in paediatric populations, yoga has been shown to improve inattentive behaviour and self esteem, and decrease anxiety (Harrison 2004). As with hypnotherapy, reductions in stress and anxiety may affect perceived RAP through changes in the hypothalamic-pituitary-adrenal axis (Kennedy 2012).

Written self-disclosure, a therapy in which the patient writes down their thoughts and feelings about something deeply distressing, is hypothesised to help with pain through a number of mechanisms, including changes in insight, the creation of a story about emotional and painful experience, and adaptation of habituation to emotional stimuli (Pennebaker 2007).

There is no consensus about which of the numerous proposed causal pathways results in the heterogeneous presentations of chronic abdominal pain. Indeed, RAP is now considered within a biopsychosocial model, with physical, emotional, and environmental factors all likely to contribute to the manifestation of unexplained abdominal pain (McOmber 2007). When considering the diverse proposed mechanisms, it is unsurprising that a range of treatments have been suggested. In addition to the psychosocial interventions discussed above, a number of dietary and pharmacological approaches have been studied. Earlier reviews of the effectiveness of dietary and pharmacological interventions for RAP are currently being updated as companions to this updated review (Huertas-Ceballos 2009a; Huertas-Ceballos 2009b).

Why it is important to do this review

Recurrent abdominal pain in children is very common and is associated with a substantially reduced quality of life. In daily clinical practice there is no consensus on which treatments to offer patients, leading to an inconsistent approach. This review aimed to establish whether there is evidence for the effectiveness of psychosocial interventions in children with RAP, as new forms of psychosocial therapies become increasingly available. It updates an earlier version (Huertas-Ceballos 2008). Companion reviews addressing the effectiveness of dietary (Martin 2014a) and pharmacological (Martin 2014b) interventions for RAP are also being updated, so together they can guide clinicians, patients and their families in treatment decisions.

OBJECTIVES

To determine the effectiveness of psychosocial interventions for reducing pain in school-aged children with RAP.

METHODS

Criteria for considering studies for this review

Types of studies

Only fully randomised controlled trials (RCTs) were eligible. The control group in the RCT could be usual care, wait-list control, or an active form of control that is not considered to be a psychosocial intervention.

Types of participants

Children and adolescents aged 5 to 18 years with RAP or an abdominal pain-related functional gastrointestinal disorder as defined by the Rome III criteria (Rasquin 2006).
Recurrent abdominal pain is defined as at least three episodes of pain interfering with normal activities within a three-month period. The Rome III criteria recognises four abdominal pain-related categories: "abdominal migraine", "irritable bowel syndrome", "functional dyspepsia", and "functional abdominal pain syndrome" or "functional abdominal pain" (Rasquin 2006).

Types of interventions

Any psychosocial intervention (intervention based on psychological or behavioural theory) compared to usual care, wait-list control, or active control. Active control groups were deemed eligible if they were considered to be comparable to what a clinician may already provide or suggest, for example education, advice, or relaxation.

Types of outcome measures

Primary outcomes

1. Treatment success (as a dichotomous variable; yes or no).
2. Pain intensity (continuous or categorical variable).
3. Pain duration or pain frequency (continuous or categorical variable).

Treatment success would be defined by the trial author, which could be a complete absence of pain postintervention or a reduction in pain according to a specified, predefined threshold.
As there is no standard method for measuring pain in this condition, studies could have use any validated measurement of pain such as a Likert scale, a visual analogue scale, or a questionnaire such as the Abdominal Pain Index (Walker 1997), which exists in various versions and formats.
We expected studies to vary in their duration of postintervention follow-up. We therefore grouped studies according to duration of follow-up: immediate outcome measurement, short term (less than 3 months), medium term (between 3 and 12 months), and long term (12 months or more).

Secondary outcomes

1. School performance (to include measures such as school functioning, behaviour, or school attendance).
2. Social or psychological functioning (to include measures such as anxiety or depression).
3. Quality of daily life (to include measures such as quality of life or impairment in daily activities (functional disability or activity limitations)).

Studies could use any validated or appropriate measurement of these secondary outcomes. For example, for school functioning this could include the Connor's Teaching Rating Scale (Conners 1969), or could be the number of missed school days. For social or psychological functioning, this could include assessing psychological adjustment using scales such as the Child Behavior Checklist (Achenbach 1983), or depression and anxiety through scales such as the Child Depression Inventory or the Multidimensional Anxiety Scale for Children (Kovacs 1992; Reynolds 1985). Examples of scales considered valid for quality of life included such quality of life scales as the Pediatric Quality of Life Scale - Short Form 36 (Varni 2001), or the KINDer Lebensqualitätsfragebogen (KINDL-R in German; Ravens-Sieberer 1998), and for daily functional activity, scales such as the Functional Disability Inventory, in Walker 1991, or the Child Activity Limitations Index (Palermo 2004).

Search methods for identification of studies

Electronic searches

We ran the first literature searches in March 2013 and updated them in April 2014, March 2015, and again in June 2016. We searched the electronic databases and trial registers listed below.

- Cochrane Central Register of Controlled Studies (CENTRAL; 2015, Issue 2) in the Cochrane Library and which includes the Cochrane Developmental, Psychosocial and Learning Problems Specialised Register (searched 10 June 2016).
- Ovid MEDLINE In-Process & Other Non-Indexed Citations and Ovid MEDLINE (1946 to current; searched 9 June 2016).
- Embase Ovid (1974 to current; searched 9 June 2016).
- CINAHL Healthcare Databases Advanced Search (Cumulative Index to Nursing and Allied Health Literature; 1981 to current; searched 9 June 2016).
- PsycINFO Ovid (1806 to current; searched 9 June 2016).
- ERIC ProQuest (Educational Resources Information Center; 1966 to current; searched 9 June 2016).
- BEI ProQuest (British Education Index; 1975 to current; searched 9 June 2016).
- ASSIA ProQuest (Applied Social Sciences Index and Abstracts; 1987 to current; searched 9 June 2016).

- AMED Healthcare Databases Advanced Search (Allied and Complementary Medicine; 1985 to current; searched 9 June 2016).
- LILACS (Latin American and Caribbean Literature in Health Sciences; lilacs.bvsalud.org/en; searched 9 June 2016).
- OpenGrey (opengrey.eu; searched 9 June 2016).
- ClinicalTrials.gov (clinicaltrials.gov; searched 9 June 2016).
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP; apps.who.int/trialsearch; searched 9 June 2016).

The search terms were revised from the original Cochrane RAP reviews (Huertas-Ceballos 2008; Huertas-Ceballos 2009a; Huertas-Ceballos 2009b); consequently, searches were run for all available years. We used RCT filters where appropriate and imposed no language limits. We translated any non-English language studies identified so that they could be screened and considered for inclusion. The search strategies for each database are reported in Appendix 1.

Searching other resources

We used the Science Citation Index (Web of Science) for forward citation searching to identify papers in which the included articles had been cited, and we checked the reference lists of the included reports to identify any additional studies, including any ongoing or unpublished work.

Data collection and analysis

Selection of studies

Two review authors (RAA, AEM, TVND, AB, JTC or RW; see Differences between protocol and review), working in pairs, independently screened the titles and abstracts of all records retrieved by the search for relevance. We obtained full-text reports for all abstracts that appeared to be potentially eligible for inclusion, or for which more information was needed, and then selected these for inclusion against the agreed-upon eligibility criteria (see Criteria for considering studies for this review). Any disagreements were resolved through discussion with a third review author (JTC). We recorded our decisions in a study flow diagram (Moher 2009).

Data extraction and management

Two review authors (RAA, AEM, TVND, AB, JTC, or RW; see Differences between protocol and review), working in pairs, extracted the data (one review author extracted the data, and the second review author checked it for accuracy). RAA entered these data into Cochrane's statistical software, Review Manager 5 (Review Manager 2014). All review authors used the same data extraction form. We extracted the following data.

1. Study characteristics: number of participating children, inclusion and exclusion criteria, type of intervention and comparison, intervention characteristics (duration, frequency, setting), number of withdrawals, study design.
2. Participant characteristics: sex, age, diagnosis (e.g. RAP or syndrome defined by the Rome III criteria) (Rasquin 2006).
3. Outcome measures: measurement of pain and any secondary outcome measured (see Types of outcome measures).

We resolved any disagreements by discussion until a consensus was reached.

Assessment of risk of bias in included studies

We assessed the risk of bias of each included study using the Cochrane 'Risk of bias' tool (Higgins 2011a). We assessed each study for bias in each of the following domains: selection bias (random sequence generation and allocation concealment); performance bias (blinding of participants and personnel); detection bias (blinding of outcome assessment); attrition bias (incomplete outcome data); reporting bias (selective outcome reporting); and other sources of bias. Two review authors (RAA, AEM, TVND, AB, JTC, or RW; see Differences between protocol and review) independently assessed each study. Based on the methods detailed in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011a), we classified each category of bias as "low risk of bias", "high risk of bias", or "unclear risk of bias" (Table 1). We also assessed all included studies for other sources of bias that could have altered the estimate of treatment effect, for example, whether the data collection tools were valid, whether there was sufficient power in terms of appropriate sample size, whether baseline parameters were similar, and whether data analyses were appropriate. We considered a trial as having an overall low risk of bias if most of the above categories of bias were assessed as low risk of bias. We considered a trial as having an overall high risk of bias if several of the above categories were assessed as high risk of bias or unclear risk of bias. We resolved any disagreements by discussion until a consensus was reached.

Measures of treatment effect

We grouped psychosocial treatments for analysis according to the mode of therapy as described by the authors: CBT, hypnotherapy (including guided imagery), yoga, and written self-disclosure. We grouped all control conditions (usual care, wait-list, or active control) together, following the precedent set by Eccleston 2014.

Dichotomous data

We analysed dichotomous data (e.g. treatment success: yes or no) using odds ratios. The definition of treatment success varied across the studies and was sometimes referred to as pain improvement. We used the author definition of treatment success.

Continuous data

For continuous data (e.g. pain intensity or frequency), we analysed mean differences and standard deviations, if these were available or could be calculated and there was no clear evidence of skewness in the distribution. When different scales were used to measure the same clinical outcome, we combined standardised mean differences across the studies.

Unit of analysis issues

As we identified no cross-over trials, cluster RCTs, or multiple intervention groups, there were no unit of analysis issues. Our planned approach for dealing with these is provided in the Additional methods table in Appendix 2.

Dealing with missing data

In the few cases where there were missing data, such as standard deviations, or where children with RAP had been combined with children with general pain, we contacted the original investigators to inquire if the missing data were available. When we were unable to obtain the data from the original investigators, we did not impute values, as per our protocol (Martin 2014c). Studies in which authors provided additional data not originally reported are detailed in the Characteristics of included studies tables.

Assessment of heterogeneity

We anticipated finding considerable heterogeneity among included studies. We assessed clinical heterogeneity by examining the distribution of relevant participant characteristics (e.g. age, definition of RAP) and study differences (e.g. concealment of randomisation, blinding of outcome assessors, interventions or outcome measures). We described the statistical heterogeneity (observed variability in study results that is greater than that expected to occur by chance) by reporting the I^2 (Higgins 2003). The I^2 describes approximately the proportion of variation in point estimates due to heterogeneity rather than sampling error. An I^2 of more than 70% may indicate substantial heterogeneity. We used the Chi^2 test to further assess the strength of evidence of the heterogeneity. We regarded any result with a P value lower than 0.10 as indicating significant statistical heterogeneity. We interpreted this cautiously and used it to help quantify the impact of heterogeneity on the results of the meta-analysis (Higgins 2003), and ultimately on the GRADE quality rating. We also presented Tau^2 as an estimate of between-study variability (see Differences between protocol and review).

Assessment of reporting biases

We did not have more than 10 trials for each outcome and so did not perform these analyses (see Additional methods table in Appendix 2).

Data synthesis

We used Review Manager 5 for statistical analysis (Review Manager 2014). Two review authors (RAA, AEM, TVND, RW) independently entered data into Review Manager 5. For summary statistics for continuous data, we reported the mean differences or standardised mean differences using an inverse variance, random-effects model. For dichotomous data, we calculated the odds ratios using a random-effects model based on the Mantel-Haenszel method. We used a random-effects model because we anticipated significant statistical and clinical heterogeneity.

We conducted a meta-analysis for studies with equivalent psychosocial interventions, for example, studies assessing CBT where the same outcomes (albeit different assessment tools) were measured. We provided a narrative description of the results when, due to the heterogeneity of the psychosocial treatment used or the variety of methods used to measure pain, we did not consider a meta-analysis to be appropriate (DerSimonian 1986).

Assessment of the quality of evidence for outcomes across included studies

We used the GRADE approach to assess the overall quality of the body of evidence for a specific outcome (Schünemann 2011). We presented the findings in the 'Summary of findings' tables, which we completed for each main treatment comparison: Summary of findings for the main comparison; Summary of findings 2; and Summary of findings 3. The probable outcome of events was calculated per 1000 for both the control group and those receiving psychosocial therapies, similar to other reviews including participants with pain conditions (e.g. Eccleston 2014). We judged the studies included for each outcome using the following five criteria: risk of bias, indirectness, inconsistency, imprecision, and publication bias. We used limitations in the design and implementation to assess the overall risk of bias of included studies for each outcome. We downgraded an outcome if the majority of studies had unclear or high risk of bias. We assessed indirectness if a population, intervention, or outcome was not of direct interest to the review (e.g. using mostly wait-list controls). We determined inconsistency by the heterogeneity of results. If an outcome had a heterogeneity greater than 70%, we downgraded the outcome quality. We assessed imprecision by the number of children included in an outcome and confidence intervals. We downgraded outcomes when only a small number of children could be included in the analysis, or the analysis had wide confidence intervals. Finally, we downgraded for publication bias if studies failed to report outcomes in the published manuscript, or if there was a suspicion that null findings had not been published or reported (Schünemann 2011). We gave each outcome a quality marking ranging from 'very low' to 'high'. High-quality ratings are given when "further research is unlikely to change our estimate of effect". Moderate-quality ratings are given when "further research is likely to have an important impact on our confidence in the estimate of effect and may change

the estimate". Low-quality ratings are given when "further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate". Finally, very low-quality ratings are given when "we are very uncertain about the estimate" (Balshem 2011, p 404). We reported a maximum of seven 'most important outcomes' in each table (Guyatt 2013).
We presented all outcomes in Summary of findings for the main comparison; Summary of findings 2; and Summary of findings 3.

Subgroup analysis and investigation of heterogeneity

Of our planned subgroup analyses we were only able to perform analyses related to duration of follow-up.
After identifying a large number of included trials using wait-list controls, we decided post hoc that we would look, where sufficient data allowed (at least two studies per group), at the effects of comparator group (see Differences between protocol and review). Wait-list control studies have been criticised as being at increased risk of bias, as there may be an expectancy of benefit, which could overestimate the treatment effect (Cunningham 2013).

Sensitivity analysis

We were able to perform our planned sensitivity analysis (see Martin 2014c), to assess the effect of inadequate allocation concealment for one of the intervention types, CBT. The details of all other planned sensitivity analyses are archived for use in future updates of this review (see Additional methods table in Appendix 2).

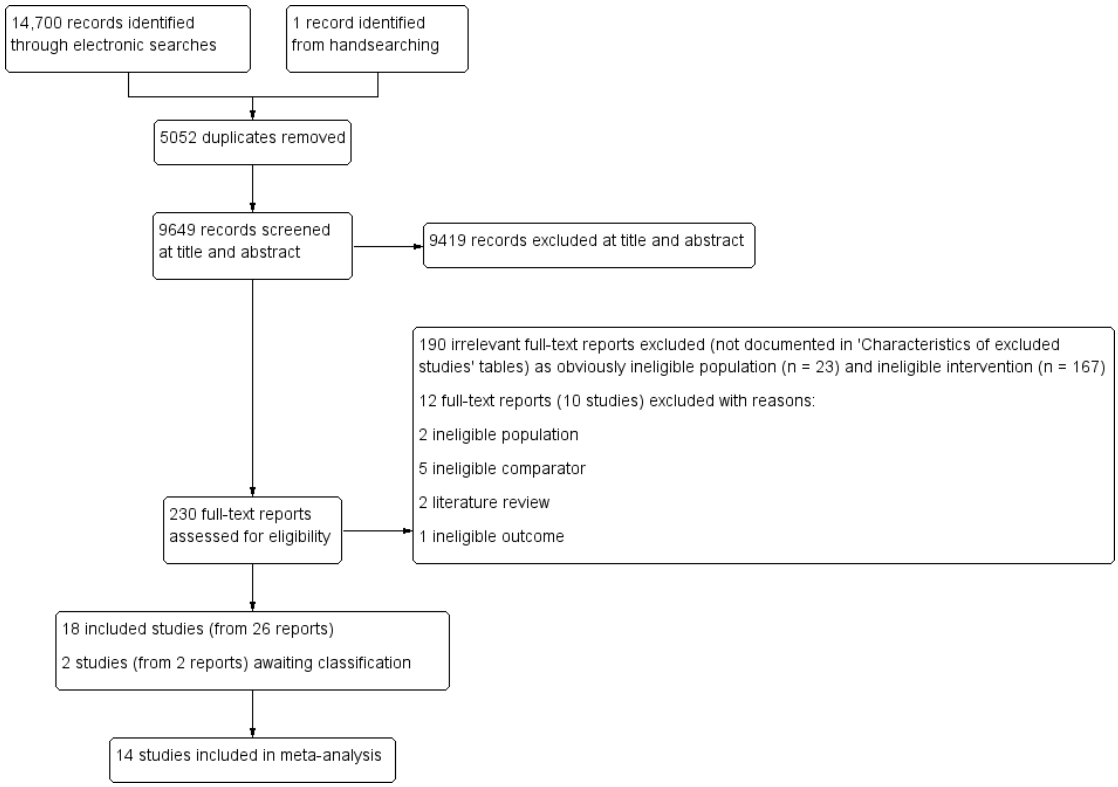
RESULTS

Description of studies

Results of the search

For this updated review, we chose to redesign the search strategy in order to include the recognised terms for different types of RAP, as defined by the Rome criteria (Rasquin 2006). We therefore ran new searches across the databases with no date restriction. The results of the searching and screening are shown in the PRISMA flow chart (Figure 1). We screened a total of 9649 titles and abstracts, and chose 230 full texts from these for further screening. We excluded 202 reports from these full texts. The majority of these (n = 190) clearly involved an ineligible population (adult) or ineligible intervention (dietary or pharmacological), and consequently are not described in the Excluded studies section. However, we have presented the details of the 12 full-text reports (describing 10 RCTs) that were excluded for less obvious reasons in the Excluded studies section.

Figure 1. PRISMA flow diagram.



The previous review, Huertas-Ceballos 2008, identified six RCTs (Duarte 2006; Hicks 2003; Humphreys 1998; Robins 2005; Sanders 1990; Sanders 1994), which were reported in 10 papers. For this new review, we included four of these original trials (reported across five papers) and excluded two trials: Hicks 2003 involved a population presenting with pain, but not specific to RAP. Correspondence with the author revealed that only six participants presented with RAP alone, but the outcome measure did not address RAP pain specifically (von Baeyer 2014 [pers comm]). We excluded Humphreys 1998 as the study involved four interventions with no control group.

Included studies

For a full description of the main study characteristics, including details on participants and setting, intervention aspects and outcome measures, see Characteristics of included studies.

We included 18 RCTs, reported in 26 papers. In addition to the four trials (reported in five papers) listed above, we included 14 new RCTs, reported in 21 papers (Evans 2014; Groß 2013; Gulewitsch 2013; Korterink 2016; Kuttner 2006; Levy 2010; Palermo 2009; Palermo 2016; van der Veek 2013; van Tilburg 2009; Vlieger 2007; Wallander 2011; Wassom 2009; Weydert 2006). One of these studies had been excluded from the earlier review on the grounds that the study assessed two psychosocial interventions with no comparator (Weydert 2006). However, we have included the study in this version of the review, as we did not consider the breathing technique comparator group to be a psychosocial intervention but rather an active control, and therefore found no grounds for exclusion.

Participants

The total number of children and adolescents with RAP randomised across the 18 included studies (26 reports) was 928; two studies randomised more than 100 participants (Levy 2010; van der Veek 2013), and three studies randomised fewer than 25 participants (Sanders 1990; Wassom 2009; Weydert 2006). The mean age at recruitment across the trials ranged from 9.4 to 14.9 years (standard deviation ranging from 1.1 to 2.9 years). Girls outnumbered boys in every trial. The majority of trials recruited children with a diagnosis under the broad umbrella of RAP; three trials recruited children specifically with functonal abdominal pain (Levy 2010; van der Veek 2013; van Tilburg 2009); two trials recruited children with irritable bowel syndrome (Evans 2014;

Kuttner 2006); one trial recruited children with functonal abdominal pain or irritable bowel syndrome (Vlieger 2007); and two trials recruited children with chronic pain more broadly, but only the results for those presenting with RAP are included here (Palermo 2009; Palermo 2016).

Settings

The majority of studies recruited children through paediatric gastroenterology or paediatric pain clinics. Three studies used a combination of clinics and community advertising (Evans 2014; Kuttner 2006; Sanders 1990); one study used community advertising (Gulewitsch 2013); and one study recruited children who were taking part in an existing epidemiological study (Groβ 2013).

Location

The studies were carried out across six countries: eight in the USA (Evans 2014; Levy 2010; Palermo 2009; Robins 2005; van Tilburg 2009; Wallander 2011; Wassom 2009; Weydert 2006), three in the Netherlands (Korterink 2016; van der Veek 2013; Vlieger 2007), two in Germany (Groβ 2013; Gulewitsch 2013), two in Australia (Sanders 1990; Sanders 1994), one in Canada (Kuttner 2006), one in Brazil (Duarte 2006), and one recruiting from both the USA and Canada (Palermo 2016).

Comparators

All 18 studies involved two treatment arms: an intervention assessed against a comparator. The comparator was usual medical care in six studies (Duarte 2006; Korterink 2016; Robins 2005; Sanders 1994; Vlieger 2007; Wallander 2011), a wait-list control in eight studies (Evans 2014; Groβ 2013; Gulewitsch 2013; Kuttner 2006; Palermo 2009; Sanders 1990; van Tilburg 2009; Wassom 2009), and an education or breathing control, or both, in four studies (Levy 2010; Palermo 2016; van der Veek 2013; Weydert 2006).

Outcomes

Outcomes were predominantly related to the primary outcome of pain, or the secondary outcome 'quality of daily life', which included functional disability or impairment due to pain and quality of life more generally.

Every trial reported on pain. Nine studies reported on treatment success (Groβ 2013; Gulewitsch 2013; Korterink 2016; Sanders 1990; Sanders 1994; van der Veek 2013; van Tilburg 2009; Vlieger 2007; Weydert 2006), defined either as a percentage reduction in pain from baseline or being pain free post-treatment. Every study reported on some measure of pain as a continuous outcome: pain intensity, pain duration or pain frequency, or a combination of these. The three most common pain scales used were versions of the FACES Pain Scale (Bieri 1990), the Pain Response Inventory (Walker 1997), and use of a standard visual analogue scale, typically with a range of 0 to 10.

Secondary outcomes varied considerably across the studies. The most common measure of functional disability related to pain was the Functional Disability Inventory (Walker 1991). One study used the KINDer Lebensqualitätsfragebogen (KINDL-R in German; Ravens-Sieberer 1998), one used the Paediatric Pain Disability Index (Hübner 2009), and one trial used the Child Activity Limitation Interview (Palermo 2004). Quality of life was measured using a variety of questionnaires: Pediatric Quality of Life Inventory (Varni 2001), KIDSCREEN (Ravens-Sieberer 2005), TNO AZL Child Quality Of Life questionnaire (Vogels 1998), KINDL-R (Ravens-Sieberer 2005), and the 36-Item Short Form Health Survey (Ware 1992). Six studies, reported in seven papers, evaluated aspects of school performance, either in the form of teacher ratings or missed school days (Korterink 2016; Sanders 1990; Sanders 1994; Vlieger 2007; Vlieger 2012; Weydert 2006).

Interventions

We categorised the interventions into four groups based on their content and the descriptions provided by the trial authors.

Cognitive behavioural therapy (CBT)

We classified 10 studies as CBT interventions, all of which we considered to be family interventions, involving both the child and a parent. The degree of parental involvement varied across the interventions from involvement in one session (Groβ 2013), approximately half of the provided sessions (Robins 2005; van der Veek 2013), or to attendance at every session (Duarte 2006; Levy 2010; Sanders 1990; Sanders 1994); or for the online interventions, an equal provision of parental and child modules (Palermo 2009; Palermo 2016; Wassom 2009). All of the CBT interventions aimed to help children cope autonomously with their pain experiences through a combination of CBT techniques, including the teaching of coping and distraction strategies; teaching of relaxation techniques; identification and change of negative pain-related thoughts; and modification of family responses to illness and wellness behaviours. Seven of the CBT interventions were carried out face-to-face (Duarte 2006; Groβ 2013; Levy 2010; Robins 2005; Sanders 1990; Sanders 1994; van der Veek 2013), whilst three were home based, with the CBT intervention facilitated via a website, in Palermo 2009 and Palermo 2016, or via a CD-ROM (Wassom 2009). Most CBT interventions involved a weekly or biweekly session that ranged from 30 to 90 minutes, and the intervention length ranged from three weeks, in Levy 2010, to eight weeks (Robins 2005; Sanders 1990). Two of the interventions were run as group-based sessions (Duarte 2006; Groβ 2013), whilst the remainder were conducted with the child or parent, or both, in one-to-one sessions. All studies, apart from one (Duarte 2006), reported having a homework component as part of the intervention. Four of the studies employed a wait-list control (Groβ 2013;

Palermo 2009; Sanders 1990; Wassom 2009); three employed a usual care control (Duarte 2006; Robins 2005; Sanders 1994); and two studies supplemented usual care with either extra education, in Levy 2010 and Palermo 2016, or education and medical support, in van der Veek 2013, to match the time and attention of the intervention group. All studies, apart from one (Duarte 2006), followed up with at least a three-month postintervention assessment, with four studies reporting a 12-month follow-up (Levy 2010; Robins 2005; Sanders 1994; van der Veek 2013). Two of the studies randomised more than 100 children (Levy 2010; van der Veek 2013); one study included 84 children (Robins 2005); and one randomised fewer than 20 children (Sanders 1990). The majority of studies randomised between 20 and 50 children.

Hypnotherapy (including guided imagery)

Four studies evaluated the effects of hypnotherapy, in Gulewitsch 2013 and Vlieger 2007, or guided imagery (van Tilburg 2009; Weydert 2006). Both hypnotherapy and guided imagery involve physical relaxation and behaviour modification through imagery. Vlieger 2007 randomised 52 children referred from paediatric gastroenterology clinics to either six-hourly sessions of individual, face-to-face hypnotherapy with a trained psychologist over three months, supported by daily practice at home (assisted by a CD-ROM of standardised hypnosis sessions) or to a usual care control group. To attempt to control for the therapist time, the usual care group also received six half-hour sessions of supportive therapy related to nutrition, pain, or stress issues. The children in this study were followed up at six and 12 months' postintervention, and in a subsequent article at five years (Vlieger 2012). Gulewitsch 2013 randomised 38 children recruited from public announcements in local newspapers and paediatric offices, to a brief group hypnotherapy intervention or wait-list control. The intervention was conducted by trained psychologists and consisted of two 90-minute group sessions for the child and two 90-minute group sessions for the parent, over four weeks. The children were educated on self instruction for relaxation; they practised standardised hypnotherapeutic trances and were advised to practise the trances with the help of a CD-ROM at home, at least five times a week during the four weeks. The parent sessions comprised information about pain and anxiety, triggers of pain, and positive educational strategies. Follow-up was undertaken two months after the end of intervention. The intervention by Weydert 2006 involved 22 children who were randomised to either four-weekly, face-to-face sessions with a therapist along with an audiotape from the first session to practise twice daily at home, or to a breathing control group. The breathing technique group in this study was designed to control for the time and attention of the therapist; the children were also provided with an audiotape to practise the techniques at home. This study also had a follow-up assessment at one-month postintervention. In the trial by van Tilburg 2009, 34 children were randomised to either wait-list control or home-based guided imagery therapy. In this eight-week trial, the initial guided imagery instruction was provided through a DVD, which the child and parent watched together, and subsequent daily practice was facilitated through a CD-ROM, which the child could listen to in his or her own space. Follow-up was immediately postintervention at eight weeks. For both guided imagery studies, children were recruited through paediatric gastroenterology clinics, with the Weydert 2006 study also recruiting through referral by general paediatricians.

Yoga

Three studies investigated the effects of yoga. Kuttner 2006 assessed the impact of daily yoga compared to wait-list control in 25 adolescents recruited from paediatric gastroenterology clinics and advertisements posted in the community. Those randomised to yoga received a one-hour instruction and demonstration session with a certified Hatha yoga instructor, and were given a series of 10 yoga positions and breathing techniques to perform, selected for their purported easing and self regulation on the abdomen and bowel. After the physical demonstration, the children were provided with a video demonstrating the same poses and breathing techniques, and were asked to practise them at home daily for four weeks. Follow-up was immediately postintervention. Evans 2014 assessed the impact of twice-weekly yoga compared to wait-list control in 29 adolescents recruited through community links and gastroenterology clinics. The intervention was Iyengar yoga, with classes held in a group format (maximum of six adolescents) for 90 minutes, twice a week, for six weeks. The adolescents received instruction in a series of postures taught with the use of props, and props were available to take home and practise with, although this was not mandatory. Follow-up was immediately postintervention and two months later. Korterink 2016 recruited 69 children through a gastroenterology outpatient clinic and assessed the impact of a 10-session, weekly yoga program (90 minutes per week) compared to usual medical care. The Hatha-based yoga involved a combination of classical yoga poses, meditation and breathing and relaxation exercises. Children were given a workbook with yoga exercises and were encouraged to practise at home on a daily basis.

Written self-disclosure

Wallander 2011 assessed the impact of written self-disclosure on abdominal pain frequency. In this trial, 63 children with RAP, recruited from paediatric pain clinics, were randomised to either three occasions of written self-disclosure or usual care. Children in the written self-disclosure group were asked (once in the clinic and on two further occasions at home, within a week of the first occasion) to write about their feelings of a distressing experience for 20 minutes. Both groups were followed up at three and six months' postintervention.

Excluded studies

We excluded 202 full texts (Figure 1). We excluded 167 because they described a dietary or pharmacological intervention and 23 because they involved adult populations. We excluded 10 studies reported in 12 full texts for the following reasons: one involved children with anxiety disorders (Warner 2011), and another one involved children with pain not specific to RAP (Hicks 2003); three were ineligible due to the comparator used (Alfvén 2007 compared psychological treatment with physiotherapy; Sieberg 2010 and Sieberg 2011 evaluated CBT against a CBT plus family therapy treatment with no control group used); two had no control groups (Humphreys 1998; van Barreveld 2015); two were literature reviews (Bursch 2008; Sato 2009); and one study had an ineligible outcome (Long 2009 reported on physical activity only). See Characteristics of excluded studies tables.

Risk of bias in included studies

We assessed each study for risk of bias in each of the following domains: random sequence generation (selection bias); allocation concealment (selection bias); blinding of participants and personnel (performance bias); blinding of outcome measures (detection bias); incomplete outcome data (attrition bias); selective outcome reporting (reporting bias); and other sources of bias (Figure 2; Figure 3).

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

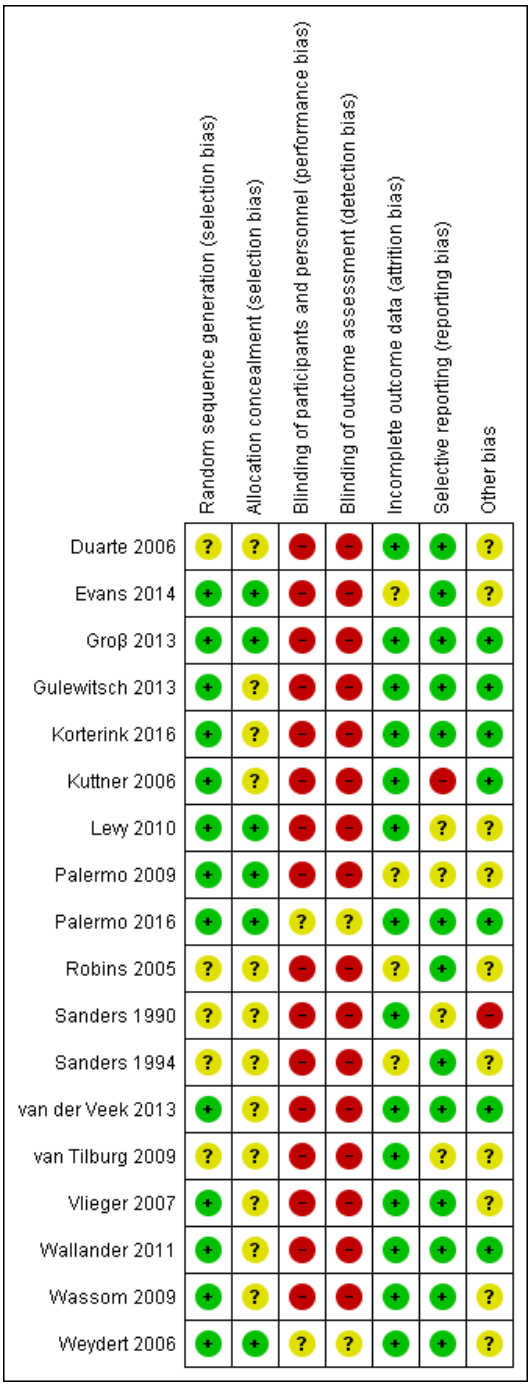
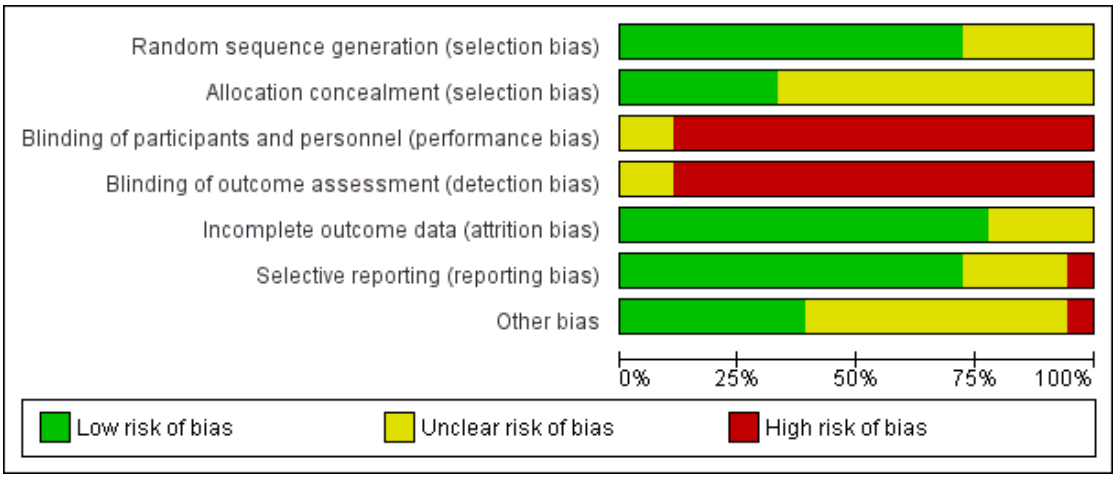


Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Allocation

We considered none of the 18 included studies to be at high risk of bias for either randomisation or allocation concealment.

Random sequence generation

We judged 13 studies to be at low risk of bias for randomisation, and the remaining five studies to be at unclear risk of bias (Duarte 2006; Robins 2005; Sanders 1990; Sanders 1994; van Tilburg 2009), because either it was not reported (n = 4), or a coin toss status method was used with no further information supplied.

Allocation concealment

Allocation concealment was not well reported. We judged six of the studies to be at low risk of bias, in which there was clear demonstration of an attempt to conceal allocation, either through the use of independent personnel not involved in the study (Evans 2014; Groβ 2013; Levy 2010; Weydert 2006), through the use of sealed, opaque envelopes (Palermo 2009), or through computer-generated randomisation and allocation programmed directly into the website for the internet-delivered study (Palermo 2016). We rated 12 studies where there was either insufficient or no detail about allocation to be at unclear risk of bias.

Blinding

Blinding of participants and personnel

As expected, due to the nature of psychosocial interventions, blinding of both participants and personnel was often not possible and we consequently judged 16 of the 18 included studies to be at high risk of performance bias. We judged two studies to be at unclear risk of bias (Palermo 2016; Weydert 2006). In the study by Palermo 2016, while it is claimed that the children were unaware whether they were receiving the active treatment or control, as their website automatically adapted to the arm to which they had been randomised, it is unknown what information was given at consent that could have made the children aware of their assignment. In the study by Weydert 2006, all of the treatments, regardless of group, were referred to as "relaxation techniques", which allowed blinding of the research associate collecting outcomes and some degree of masking of children (and parents) not previously aware of these therapies.

Blinding of outcome assessment

We considered most studies (16 out of 18) to be at high risk of bias for blinding of outcome assessment as the majority of outcomes were self reported, and children were aware of their treatment group. We judged two studies to be at unclear risk of detection bias

(Palermo 2016; Weydert 2006). In the study by Palermo 2016, as stated above, it is unknown what information was given at consent that could have made the children aware, and children were self reporting the primary outcomes. In the study by Weydert 2006, the researcher collecting the outcome data was unaware of treatment allocation, and attempts had been made to mask which treatment was being given by referring to them both as "relaxation techniques".

Incomplete outcome data

Fourteen of the studies reported attrition fully and were rated at low risk of bias. Four studies did not fully account for the drop in numbers through the study, or whether this differed between groups, and were rated at unclear risk of bias (Evans 2014; Palermo 2009; Robins 2005; Sanders 1994).

Selective reporting

Thirteen of the 18 studies were clear in their reporting of the primary outcomes and were therefore judged to be at low risk of reporting bias. We judged four studies to be at unclear risk, as they either presented their data as figures with little detail (Sanders 1990), or were missing some stated secondary outcomes (Levy 2010; Palermo 2009; van Tilburg 2009). We judged one study to be at high risk of bias, as the primary outcome data were missing (Kuttner 2006).

Other potential sources of bias

We rated the risk of other potential biases (such as validity of data collection tools, appropriate sample size, similarity of baseline details) as low in seven of the included studies (Groß 2013; Gulewitsch 2013; Korterink 2016; Kuttner 2006; Palermo 2016; van der Veek 2013; Wallander 2011). These studies used valid collection tools, reported calculation of sample sizes, and demonstrated no baseline differences of concern. We judged the risk of other sources of potential bias as unclear in 10 of the studies, as there was insufficient detail within the papers on which to judge the criteria. We rated one study at high risk of other potential bias due to baseline differences in the primary outcome of interest and uncertainty about whether these differences were accounted for, along with uncertainty about the adequacy of the sample size (Sanders 1990).

Effects of interventions

See: **Summary of findings for the main comparison** Cognitive behavioural therapy compared with control for children and adolescents with recurrent abdominal pain; **Summary of findings 2** Hypnotherapy compared with control for children and adolescents with recurrent abdominal pain; **Summary of findings 3** Yoga compared with control for children and adolescents with recurrent abdominal pain

We were able to perform 14 analyses across the included studies. Analyses were performed within intervention type. We were able to perform nine analyses for CBT intervention compared to control. With regard to the primary outcome of pain, we analysed effects on treatment success and pain intensity at postintervention, medium-term follow-up (between 3 and 12 months) and at long-term follow-up (12 months or more). For the secondary outcome 'quality of daily life', we analysed effects on quality of life (both physical and psychosocial domains) postintervention, and effects on functional disability due to pain at postintervention.

We were able to perform three analyses for hypnotherapy compared to control: effects on treatment success, pain intensity, and pain frequency postintervention.

We were able to perform two analyses for yoga therapy compared to control immediately postintervention: effects on pain intensity and effects on functional disability due to pain.

We only performed analyses on those studies that provided equivalent outcome data in comparable formats, therefore not all studies within intervention type were entered into the analyses. No analyses were possible for written self-disclosure, as there was only one study, so for this study we have presented a narrative description of the results. The heterogeneity across the interventions was mixed. Four analyses showed low heterogeneity (I^2 value less than 25%), and six analyses showed high heterogeneity (I^2 value 70% or more).

Post hoc subgroup analyses of the effect of comparator were possible for six analyses: three analyses of effects of CBT intervention (postintervention treatment success, pain intensity, and functional impairment) and three analyses of effects of hypnotherapy intervention (postintervention treatment success, pain intensity, and pain frequency).

We assessed the quality of evidence using the GRADE criteria (see Summary of findings for the main comparison; Summary of findings 2; Summary of findings 3). For CBT intervention (Summary of findings for the main comparison), three outcomes scored very low quality, meaning we are very uncertain of the estimates of effect on treatment success, physical quality of life, and functional disability postintervention. The remaining six outcomes (treatment success at medium- and long-term follow-up, pain intensity at postintervention and both medium- and long-term follow-up, and psychosocial quality of life postintervention) scored low quality, meaning future research is likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

For hypnotherapy and yoga, all outcomes (estimates of effect on treatment success, pain intensity, and pain frequency immediately postintervention, and functional disability at postintervention follow-up) scored low quality, therefore future research is likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate (Summary of findings

2; Summary of findings 3).

Comparison 1: cognitive behavioural therapy (CBT) versus control

Primary outcomes

Treatment success

Four studies presented dichotomous data relating to treatment success. The definition of treatment success varied across the studies, either being pain free or experiencing a reduction in pain over a certain threshold on the Abdominal Pain Index (Walker 1997). We combined data from a total of 175 children into an analysis of the effects of CBT intervention compared to control groups on treatment success immediately postintervention (Groβ 2013; Sanders 1990; Sanders 1994; van der Veek 2013). The pooled odds ratio (OR) for treatment success was 5.67 (95% confidence interval (CI) 1.18 to 27.32; P = 0.03; I^2 = 71%; Tau^2 = 1.69; P for heterogeneity = 0.01; Analysis 1.1), suggesting evidence of an effect for CBT on treatment success. However, due to the high risk of bias across the studies (unblinded allocation, unblinded outcome assessment), high level of unexplained heterogeneity (greater than (>) 70%), wide CIs, and the low number of participants included in the analysis, we rated the GRADE quality as very low, meaning we are very uncertain of this estimate of effect (see Summary of findings for the main comparison). We conducted subgroup analyses on treatment success postintervention according to study comparator. For two studies with active control or usual care (Sanders 1994; van der Veek 2013), the pooled OR for treatment success was 2.25 (95% CI 0.57 to 8.88; P = 0.25; 130 children). For two studies comparing intervention to wait-list control (Groβ 2013; Sanders 1990), the pooled OR for treatment success was 24.55 (95% CI 2.24 to 269.03; P = 0.009; 45 children). The difference between subgroups was not statistically significant (Chi^2 = 2.88; df = 1; P = 0.09; I^2 = 65.3%; Analysis 1.1).

Three of the four studies provided medium-term follow-up data on treatment success (Sanders 1990; Sanders 1994; van der Veek 2013). The pooled OR for medium-term treatment success was 3.08 (95% CI 0.93 to 10.16; P = 0.06; I^2 = 52%; Tau^2 = 0.57; P for heterogeneity = 0.12), based on data from 139 children (Analysis 1.2), suggesting there was insufficient evidence for the effect of CBT compared to control on medium-term treatment success. We rated the GRADE quality for this outcome as low, due to small sample sizes and variation in measurement of treatment success, meaning future research is likely to have an impact on our confidence in the estimate of effect. Two of the four studies provided long-term follow-up data on treatment success (Sanders 1994; van der Veek 2013). The pooled OR for long-term treatment success was 1.29 (95% CI 0.50 to 3.33; P = 0.60; I^2 = 35%; Tau^2 = 0.17; P for heterogeneity = 0.22), based on data from 120 children (Analysis 1.3), suggesting there was insufficient evidence for the effect of CBT on long-term treatment success. We rated the GRADE quality for this outcome as low, due to small sample sizes and variation in measurement of treatment success.

Pain intensity

We combined data from seven studies (405 children) to estimate the effects of CBT intervention compared to control groups on pain intensity postintervention (Groβ 2013; Levy 2010; Palermo 2009; Palermo 2016; Sanders 1994; van der Veek 2013; Wassom 2009). Pain intensity was measured in a number of ways: a visual analogue scale (ranging from 0 to 10), the FACES Pain Scale (Bieri 1990), and the Abdominal Pain Index (Walker 1997). The pooled standardised mean difference (SMD) of pain intensity across the studies was -0.33 (95% CI -0.74 to 0.08; P = 0.12; I^2 = 70%; Tau^2 = 0.19; P for heterogeneity = 0.003; Analysis 1.4), suggesting there was insufficient evidence for the effect of CBT on pain intensity immediately postintervention. The GRADE quality rating for this outcome was low, due to high unexplained heterogeneity and a high risk of bias across the studies. We conducted subgroup analyses on the effect of CBT on pain intensity postintervention according to study comparator. For four studies with active control or usual care (Levy 2010; Palermo 2016; Sanders 1994; van der Veek 2013), the pooled SMD for pain intensity was -0.04 (95% CI -0.39 to 0.31; P = 0.82; 337 children). For three studies comparing intervention to a wait-list control (Groβ 2013; Palermo 2009; Wassom 2009), the pooled SMD for pain intensity was -0.92 (95% CI -1.59 to -0.24; P = 0.008; 68 children). The difference between the subgroups was statistically significant (Chi^2 = 5.17; df = 1; P = 0.02; I^2 = 80.6%; Analysis 1.4). We conducted sensitivity analyses accounting for possible bias related to uncertainty about treatment allocation concealment. When the three studies with unclear allocation concealment were removed from the analysis, the pooled SMD of pain intensity was -0.28 (95% CI -1.00 to 0.45; Z = 0.75; P = 0.46; 4 studies; 247 children; Analysis 1.5), again suggesting insufficient evidence of effect of CBT on pain intensity immediately postintervention.

Three additional studies reported postintervention pain intensity outcome data (Duarte 2006; Robins 2005; Sanders 1990), which could not be pooled with the studies above due to insufficient data, such as missing standard deviations (SDs). Two studies reported significant benefits of decreased pain intensity with CBT compared to control (Robins 2005; Sanders 1990). Robins 2005 (86 children) found reduced scores (no SDs reported) on the Abdominal Pain Index for the CBT group (mean 16.2) compared to those given usual care (mean 19.5) postintervention (P < 0.05, exact P value not in report) and at 12 months (CBT: mean 15.7, usual care: mean 21.2, P < 0.05, exact P value not in report). Using a visual analogue scale, Sanders 1990 (16 children), reported reduced pain intensity for those receiving CBT compared to wait-

list at postintervention (P = 0.02, raw data not reported), but this was not sustained at three-month follow-up. Also using a visual analogue scale, Duarte 2006 (32 children) found no effect of treatment on pain intensity for CBT (mean 1.5) compared to control (mean 1.9) (P = 0.371 (no SDs reported)).

Four studies (301 children) provided medium-term follow-up data on the effectiveness of CBT intervention compared to control groups on pain intensity (Groβ 2013; Levy 2010; Palermo 2016; van der Veek 2013). The pooled SMD of pain intensity at medium-term follow-up was -0.32 (95% CI -0.85 to 0.20; P = 0.23; I^2 = 76%; Tau^2 = 0.20; P for heterogeneity = 0.007; Analysis 1.6), suggesting there was insufficient evidence for the effect of CBT on medium-term pain intensity. We rated the GRADE quality for this as low due to small sample sizes and substantial heterogeneity.

Three studies (308 children) provided long-term follow-up data on the effectiveness of CBT intervention compared to control groups on pain intensity (Levy 2010; Sanders 1994; van der Veek 2013). The pooled SMD of pain intensity was -0.04 (95% CI -0.39 to 0.31, P value = 0.82; I^2 = 52%, Tau^2 = 0.05; P value for heterogeneity = 0.13; Analysis 1.7), again suggesting insufficient evidence of effect. We rated the GRADE quality for this as low due to small sample sizes and substantial heterogeneity.

Pain duration

Only one of the above studies (104 children) reported on pain duration (van der Veek 2013), measured with a pain diary (score range 0 to 21). There was no evidence of effect of intervention compared with active control on pain duration at any time point (postintervention mean: 8.67 intervention, 6.84 control, P = 0.96; 6 months' mean: 5.34 intervention, 8.58 control, P = 0.25; 12 months' mean: 6.11 intervention, 6.89 control, P = 0.80 (no SDs reported)).

Secondary outcomes

School performance

Sanders 1990 (16 children) was the only study to report on children's school performance, as reported by teachers using the Conners' Teacher Rating Scale (Conners 1969). No difference was reported between children receiving CBT postintervention (mean 19.9 (SD 14.8)) or at three-month follow-up (mean 11.5 (SD 13.2)), compared to those in the wait-list control group (postintervention mean: 17.8 (SD 6.8); three-month follow-up mean: 15.8 (SD 13.5)).

Social or psychological functioning

Three studies reported outcomes related to social or psychological functioning, all of which found no effect of therapy when compared to control. Levy 2010 (200 children) found no effect of CBT compared to active control (education) on either child-reported depression or anxiety at postintervention (P > 0.05, data not shown, exact P value not in report). These outcomes were not reported in the follow-up paper (see Levy 2013 within Levy 2010). Sanders 1994 (44 children) found no effect of CBT compared to usual care on psychological adjustment (measured using the Child Behavior Checklist; Achenbach 1983). Neither internalising nor externalising behaviours were different at postintervention, or at 6- and 12-month follow-up (analyses not shown). Wassom 2009 (15 children) found no effect of CBT compared to wait-list control on either stress, as measured with a stress checklist inventory (Schanberg 2000), or on mood state, as measured with the Facial Affective Scale (McGrath 1991), at postintervention (stress: CBT mean 0.95 (SD 1.47), wait-list control mean 1.63 (SD 0.62), P > 0.05 (exact P value not in report); mood: CBT mean 0.33 (SD 0.13), wait-list control mean 0.44 (SD 0.15), P > 0.05 (exact P value not in report).

Quality of daily life

Three studies assessed the effectiveness of CBT family intervention on child quality of life postintervention (Groβ 2013; van der Veek 2013; Wassom 2009). Two studies, Groβ 2013 and Wassom 2009, used the Pedatric Quality of Life Inventory (Varni 2001), and one study, Wassom 2009, used KIDSCREEN (Ravens-Sieberer 2005). We ran separate analyses for the effects on physical and psychosocial domains of quality of life. Data were available from 136 children for both analyses. The pooled SMD for physical quality of life was 0.71 (95% CI -0.25 to 1.66; P = 0.15; I^2 = 79%; Tau^2 = 0.55; P for heterogeneity = 0.008; Analysis 1.8), and for psychosocial quality of life was 0.43 (95% CI -0.21 to 1.06; P = 0.19; I^2 = 58%; Tau^2 = 0.18; P for heterogeneity = 0.09; Analysis 1.9). Both analyses suggest insufficient evidence of effect for CBT on reported quality of life. GRADE quality ratings for the two quality of life outcomes were very low and low, respectively, mainly due to the high risk of bias across the studies, small sample size, and, in the case of physical quality of life, substantial heterogeneity. No data were available for this outcome from medium- or long-term follow-up.

Four studies measured functional impairment of daily activities due to pain. We pooled data from these studies (176 children) in the analysis of the effects of CBT intervention on pain-related functional impairment (Groβ 2013; Palermo 2009; Palermo 2016; van der Veek 2013). Although measured differently across the four studies (Palermo 2009 and Palermo 2016 used the Child Activity Limitation Interview (Palermo 2004); van der Veek 2013 used the Functional Disability Inventory (Walker 1991); and Groβ 2013 used the KINDL-R (Ravens-Sieberer 1998)), all four assessed whether treatment was effective in reducing pain-induced limitations on everyday activities. For CBT compared to control, the pooled SMD of functional impairment postintervention was

-0.57 (95% CI -1.34 to 0.19; P = 0.14; I^2 = 80%; Tau^2 = 0.47; P for heterogeneity = 0.002; Analysis 1.10), suggesting there was insufficient evidence for the effect of CBT on functional impairment. The GRADE quality rating for this effect estimate was very low due to substantial heterogeneity, small sample sizes, and concerns regarding a high risk of bias across the studies. We conducted subgroup analyses on the effect of CBT on functional impairment of daily activities postintervention according to study comparator. For two studies using active control or usual care (Palermo 2016; van der Veek 2013), the pooled SMD for functional impairment was -0.01 (95% CI -0.36 to 0.34; P = 0.96; 123 children). For two studies comparing intervention to wait-list control (Groß 2013; Palermo 2009), the pooled SMD for functional impairment was -1.31 (95% CI -2.10 to -0.52; P = 0.001; 53 children). The difference between the subgroups was statistically significant (Chi^2 = 8.59; df = 1; P = 0.003; I^2 = 88.4%; Analysis 1.10).

Comparison 2: hypnotherapy (including guided imagery) versus control

Four studies involving 146 children assessed the effectiveness of hypnotherapy, in Gulewitsch 2013 and Vlieger 2007, or guided imagery (van Tilburg 2009; Weydert 2006). All four studies measured treatment success as well as the absolute change in pain intensity and frequency.

Primary outcomes

Treatment success

All four studies presented dichotomous data relating to treatment success, the definition of which varied across the studies. Both Gulewitsch 2013 and Vlieger 2007 defined treatment success as an 80% decrease in pain intensity; van Tilburg 2009 defined it as a 50% reduction in baseline pain scores; and Weydert 2006 used a definition of fewer than four days of pain per month and no missed activities, as reported by the child. Two studies independently reported a higher likelihood of treatment success with hypnotherapy compared to control, whilst two did not. The pooled OR for treatment success was 6.78 (95% CI 2.41 to 19.07; P < 0.0003; I^2 = 23%; Tau^2 = 0.26; P for heterogeneity = 0.27; Analysis 2.1), suggesting evidence of an effect for hypnotherapy on treatment success. We rated the GRADE quality of this outcome as low due to the small number of participants across the studies and uncertain or high risk of bias within the studies (see Summary of findings 2). We conducted subgroup analyses on the effect of hypnotherapy on treatment success postintervention according to study comparator. For two studies with active control or usual care (Vlieger 2007; Weydert 2006), the pooled OR for treatment success postintervention was 10.51 (95% CI 2.88 to 38.33; P = 0.0004; 74 children). For two studies comparing hypnotherapy to wait-list control (Gulewitsch 2013; van Tilburg 2009), the pooled OR for treatment success was 5.77 (95% CI 0.64 to 52.05; P = 0.12; 72 children). The difference between subgroups was not statistically significant (Chi^2 = 0.21, df = 1, P = 0.65, I^2 = 0%; Analysis 2.1).

Weydert 2006 reported on follow-up at one month, when 70% of those who had received the guided imagery intervention reported treatment success compared to 15% in the comparator breathing group (risk ratio (RR) 7.3, 95% CI 1.1 to 48.6; P < 0.04, exact P value not in report). Long-term data from Vlieger 2012 in their five-year follow-up, which included 45 of the original 49 children, found 68% of the intervention group were symptom free compared to 20% in the control arm (P = 0.005).

Pain intensity

Pain intensity was measured using different scales: Gulewitsch 2013 used a numeric rating scale from 1 to 10; Vlieger 2007 used an affective facial pain scale ranging from 1 to 9; Weydert 2006 used the FACES Pain Scale (Bieri 1990); and van Tilburg 2009 used the Abdominal Pain Index (Walker 1997). Three of the four studies individually reported greater decreases in pain intensity in the intervention arm than in the control arm (Gulewitsch 2013; van Tilburg 2009; Vlieger 2007). The pooled SMD of pain intensity across the four studies (146 children) was -1.01 (95% CI -1.41 to -0.61; P < 0.00001; I^2 = 21%; Tau^2 = 0.03; P for heterogeneity = 0.28; Analysis 2.2), suggesting evidence of an effect for hypnotherapy on pain intensity scores immediately postintervention. We rated the GRADE quality for this as low, due to a high risk of bias across the studies in study design and outcome assessment (unblinded allocation and assessment, wait-list control) and a low number of participants included in the analysis and low number of events. We conducted subgroup analyses on the effect of hypnotherapy on pain intensity postintervention by study design comparator. For two studies with active control or usual care (Vlieger 2007; Weydert 2006), the pooled SMD for pain intensity was -1.00 (95% CI -1.90 to -0.11; P = 0.03; 74 children). For two studies using a wait-list control comparator (Gulewitsch 2013; van Tilburg 2009), the pooled SMD for pain intensity was -0.95 (95% CI -1.44 to -0.46; P = 0.0002; 72 children). The difference between subgroups was not statistically significant (Chi^2 = 0.01, df = 1, P = 0.92, I^2 = 0%; Analysis 2.2).

One study reported five-year follow-up data (Vlieger 2012). In this study, pain intensity remained significantly lower at five years (P < 0.001) in the group that had received three months of hypnotherapy (mean 2.9 (SD 4.4)) compared to the group that had received usual care (mean 7.7 (SD 5.3)).

Pain frequency

Pain frequency was measured using different scales: Gulewitsch 2013 and Weydert 2006 used a pain diary recording the number of days with pain over the past two weeks; Vlieger 2007 used a

score combining the number of days with length of pain episodes over seven days; and van Tilburg 2009 asked a question about the number of days with pain in the past week. Three of the four studies individually reported greater decreases in pain frequency in the intervention arm than in the control arm (Gulewitsch 2013; Vlieger 2007; Weydert 2006). The pooled SMD of pain frequency across the four studies (146 children) postintervention was -1.28 (95% CI -1.84 to -0.72; P < 0.00001; I^2 = 55%; Tau^2 = 0.18; P for heterogeneity = 0.08; Analysis 2.3), suggesting evidence of an effect of hypnotherapy on pain frequency. We rated the GRADE quality for the effects on pain frequency as low, due to a high risk of bias across the studies in study design and outcome assessment (limitations in study design) and a low number of participants included in the analysis and low number of events. We conducted subgroup analyses on the effect of hypnotherapy on pain frequency postintervention by study comparator. For two studies with active control or usual care (Vlieger 2007; Weydert 2006), the pooled SMD for pain frequency was -1.74 (95% CI -2.29 to -1.19; P < 0.00001; 74 children). For two studies using wait-list control (Gulewitsch 2013; van Tilburg 2009), the pooled SMD for pain frequency was -0.87 (95% CI -1.38 to -0.36; P = 0.0009; 72 children). The difference between the subgroups was statistically significant (Chi^2 = 5.22, df = 1, P = 0.02, I^2 = 80.8%; Analysis 2.3).

Vlieger 2012 also reported that, on five-year follow-up, pain frequency remained significantly lower (P = 0.001) in the group that had received hypnotherapy (mean 2.3 (SD 4.0)) compared to the group that had received usual care (mean 7.1 (SD 6.0)).

Pain duration

Only one study, Gulewitsch 2013, reported pain episode duration data. As with their data on pain frequency and pain intensity, they reported pain duration as significantly lower for the 20 children who had received hypnotherapy compared to those in the wait-list control group, with mean scores of 1.20 (SD 1.47) compared to 3.50 (SD 2.53), P = 0.014, respectively.

Secondary outcomes

All studies reported secondary outcomes measures that could not be pooled.

School performance

Two studies reported on missed school days. van Tilburg 2009 reported no differences (P = 0.2) in the number of missed school days immediately postintervention in the children who had received guided imagery (mean 0.7) compared to those in the wait-list control group (mean 1.7) (no SDs reported). Whilst not reported in the original study (Vlieger 2007), there were no differences in the number of children who had missed more than 6 days of school in the past 12 months between those who had received hypnotherapy compared to those who had received usual care at 5-year follow-up (Vlieger 2012); 3 out of 27 children compared to 7 out of 22 children respectively (P = 0.09, no SDs reported).

Social or psychological functioning

No studies reported on this outcome.

Quality of daily life

Two studies reported on quality of life. Gulewitsch 2013 found no difference in self reported health-related quality of life (as measured by the KINDL-R; Ravens-Sieberer 1998) for those receiving hypnotherapy compared to wait-list control (only summary data of analyses reported; F = 2.56, P = 0.120). van Tilburg 2009 found that children who had received guided imagery therapy reported an improved overall quality of life (mean 28.2), as measured by the Pediatric Quality of Life Inventory (Varni 2001), compared to those in the wait-list control group (mean 9.3) at postintervention (P = 0.49, no SDs reported). Long-term data from Vlieger 2012 were available on quality of life at 5-year follow-up, as measured using the TNO AZL Questionnaire for Children's Quality Of Life (Vogels 1998) for children under 16 years of age, and the TNO AZL Questionnaire for Adult's Quality of Life for children aged 16 years and older (Fekkes 2001). Whilst not reported in the original study (Vlieger 2007), there were no differences in quality of life at five-year follow-up between those who had received hypnotherapy compared to usual care control (raw data not reported).

Gulewitsch 2013 reported benefits of hypnotherapy on pain-related functional disability, as measured by the Paediatric Pain Disability Index (Hübner 2009), compared to wait-list control postintervention (hypnotherapy: mean 16.13 (SD 5.23), wait-list control: 22.44 (SD 6.33); P = 0.009). Weydert 2006 used a diary to collect data on days of missed activities due to pain, finding that children learning guided imagery had a greater reduction in days with missed activities compared to children in the active-control group at postintervention (guided imagery: 85%, active control: 15%, P = 0.02) and at one-month follow-up (guided imagery: 95%, active control: 77%, P = 0.05).

Comparison 3: yoga versus control

Three studies involving 122 children assessed the effectiveness of yoga compared to control on pain intensity, pain frequency, and functional disability (Evans 2014; Korterink 2016; Kuttner 2006).

Primary outcomes

Treatment success

Korterink 2016, which evaluated a 10-week yoga intervention and involved 69 children, was the only study to report on treatment

success. Treatment success was defined as a decrease of combined abdominal pain scores (frequency and intensity) of greater than 50%. No significant differences between those children who had undergone yoga compared to usual care were observed post-treatment (21.2% for yoga compared to 20% for control); however, by 12 months' follow-up significantly higher treatment success was reported by those in the intervention group compared to those in the usual care group (58.1% compared to 28.9% respectively, P = 0.01).

Pain intensity

Two studies measured pain intensity using a numeric rating scale (range 0 to 10) (Evans 2014; Kuttner 2006), and one study used the FACES scale (range 0 to 6; Bieri 1990) (Korterink 2016). None of the studies individually reported beneficial effects of yoga compared to control on pain intensity. The pooled SMD of pain intensity across the three studies (122 children) was -0.31 (95% CI -0.67 to 0.05; P = 0.09; I^2 = 0%; Tau^2 = 0.00; P for heterogeneity = 0.55; Analysis 3.1), suggesting no evidence of an effect of yoga therapy on pain intensity immediately postintervention. We rated the GRADE quality for this outcome as low due to the low number of participants in each study and concerns about risk of bias (non-blinding affecting risk of both performance and detection bias, and potential bias related to wait-list control design) within the studies (see Summary of findings 3).
Only Korterink 2016 provided long-term data, reporting pain intensity data from postintervention to 12 months' follow-up, and found no significant effect over time for the yoga intervention compared to usual care (P = 0.09)

Pain frequency

Korterink 2016 reported no significant effect of yoga compared to usual care on pain frequency across the three follow-ups (postintervention, 6 months, and 12 months; P = 0.20).

Pain duration

No studies reported on pain duration.

Secondary outcomes

School performance

Korterink 2016 reported a significant effect of yoga compared to usual care on school absenteeism across the three follow-ups (postintervention, 6 months, and 12 months; P = 0.03).

Social or psychological functioning

None of the three studies reported significant effects of yoga intervention on social or psychological functioning. Kuttner 2006 (25 children) reported no difference in change in anxiety score from baseline, as measured by the Revised Children's Manifest Anxiety Scale (Reynolds 1985), for those receiving yoga therapy (change score -0.28 (SDs not reported)) compared to wait-list controls (change score 0.13 (SDs not reported)) postintervention, and no difference in anxiety (P > 0.10), as measured by the Child Depression Inventory (Kovacs 1992). Evans 2014 reported no difference (P = 0.85) in psychological distress, as measured using the Brief Symptom Inventory-18 (Derogatis 2000), at postintervention for those receiving yoga therapy (change score -2.18 (95% CI -7.27 to 2.92)) compared to wait-list controls (change score -1.85 (95% CI -7.67 to 3.98)). Korterink 2016 reported a trend, though not significant, of psychological well-being, as assessed using KIDSCREEN (Ravens-Sieberer 2005), for those receiving yoga therapy compared to usual care (postintervention, 6 months, and 12 months; P = 0.06).

Quality of daily life

Two studies, Evans 2014 and Kuttner 2006, measured functional disability immediately postintervention, using the Functional Disability Inventory (Walker 1991). Evans 2014 reported no effect of yoga compared to wait-list control, whereas Kuttner 2006 reported a reduction for the yoga group compared to the active-control group (no raw data, P = 0.09 (according to their a priori statistical significance cutoff of P < 0.10)). The pooled SMD of functional impairment across both studies (53 children) was -0.32 (95% CI -1.07 to 0.43; P = 0.40; I^2 = 44%; Tau^2 = 0.13; P for heterogeneity = 0.18; Analysis 3.2) for yoga compared to control. As per above, we deemed the GRADE quality rating for this outcome to be low due to the low number of participants in each study and concerns about bias within the studies. There were no long-term follow-up data on functional disability from either study, as wait-list controls were entered into treatment either immediately postintervention or two months' postintervention completion. Korterink 2016 reported no significant effect of yoga compared to usual care on physical well-being, as assessed using KIDSCREEN (Ravens-Sieberer 2005), across the three follow-ups (postintervention, 6 months, and 12 months; P = 0.43).

Comparison 4: written self-disclosure versus usual care

One study (63 children) assessed the effectiveness of written self-disclosure therapy compared to usual care for RAP (Wallander 2011).

Primary outcomes

Treatment success

No data were reported for this outcome.

Pain intensity

No data were reported for this outcome.

Pain duration or pain frequency

No effect of treatment on the frequency of debilitating pain episodes (using a scale of 0 to 5, where 0 = none and 5 = every day) was found postintervention or at three months' follow-up. However, at six months' follow-up the frequency of such episodes was lower ($P < 0.05$, exact P value not in report) in those who had undergone written self-disclosure (mean 1.35 (SD 1.39)) compared to usual care (mean 2.32 (SD 1.72)).

Secondary outcomes

School performance and social or psychological functioning

No data were reported for these outcomes.

Quality of daily life

No differences were reported in quality of life measures or in somatisation severity.

Overall, given there being only a single study of short duration and our concerns over performance and detection bias due to lack of blinding, we have limited confidence in the observed results.

ADDITIONAL SUMMARY OF FINDINGS *[Explanation]*

Hypnotherapy compared with control for children and adolescents with recurrent abdominal pain

Patient or population: children and adolescents with recurrent abdominal pain
Settings: mixed
Intervention: hypnotherapy
Comparison: usual care or wait-list control

Outcomes	Probable outcome with control or usual care	Probable outcome with hypnotherapy	OR (95% CI)	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
Treatment success: postintervention	**136 per 1000**	**525 per 1000**	Pooled OR 6.78 (2.41 to 19.07)	146 (4)	⊕⊕◯◯ [1] **Low**	2 studies defined treatment success or remission as > 80% decrease in pain intensity. 1 study used the definition of "4 or less days of pain per month and no missed activities" and 1 study as "> 50% reduction in API" (Walker 1997).
Pain intensity: postintervention Lower score equals less pain.	The pain intensity score in the hypnotherapy groups was, on average, **1.01 SDs lower** (95% CI -1.41 to -0.61) than in the usual care or wait-list control groups		-	146 (4)	⊕⊕◯◯ [1] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference. 1.3 represents a large effect difference Pain intensity measured by 2 different scales (the FACES Pain Scale and the API (Bieri 1990; Walker 1997)).

Pain frequency: postintervention Lower score equals less pain.	The pain frequency score in the hypnotherapy groups was, on average, **1.28 SDs lower** (95% CI -1.84 to -0.72) than in the usual care or wait-list control groups	-	146 (4)	⊕⊕◯◯ [1] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference. 1.50 SD represents a large effect difference Pain frequency measured by different scales (bespoke pain diary recording the number of days; daily scale ranging from 0 to 3, summed over 7 days; and API, range 1 to 8 (Walker 1997)).

API: Abdominal Pain Index; **CI:** confidence interval; **OR:** odds ratio; **SD:** standard deviation

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Downgraded two levels: a high risk of bias across the studies in study design and outcome assessment (unblinded allocation and assessment, wait-list control) and a low number of participants included in the analysis or low number of events.

Yoga compared with control for children and adolescents with recurrent abdominal pain

Patient or population: children and adolescents with recurrent abdominal pain
Settings: mixed
Intervention: yoga
Comparison: wait-list control or usual care

Outcomes	Comparative effect of intervention versus comparator	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
Pain intensity: postintervention Lower score equals less pain.	The pain intensity score in the yoga groups was, on average, **0.31 SDs lower** (95% CI -0.67 to 0.05) than in the wait-list control groups	122 (3)	⊕⊕◯◯ [1] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference 2 studies measured pain intensity with a numeric rating scale, range 1 to 10, and 1 study used the FACES Pain Scale (0 to 5) (Bieri 1990).
Functional disability: postintervention Lower score equals less functional disability.	Functional disability in the yoga groups was, on average, **0.32 SDs lower** (95% CI -1.07 to 0.43) than in the wait-list control groups	53 (2)	⊕⊕◯◯ [1] **Low**	As a rule of thumb, 0.2 SD represents a small difference, 0.5 SD a moderate difference, and 0.8 SD a large difference Both studies used the Functional Disability Inventory (Walker 1991).

CI: confidence interval; **SD:** standard deviation

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Downgraded two levels: a high risk of bias across the studies in study design and a low number of participants included in the analysis or low number of events.

DISCUSSION

Summary of main results

We included 18 RCTs, reported in 26 papers, involving 928 children and adolescents with RAP in this updated review. All studies assessed one treatment arm of psychosocial intervention against either usual care, wait-list control, or some form of education or compensatory control. The duration of the interventions ranged from 1 to 12 weeks, with most reporting an intervention over 4 to 6 weeks. We combined all comparators as controls, and analysed the data within each intervention type: cognitive behavioural family therapy (CBT), hypnotherapy, yoga, and written self-disclosure. This update extends the evidence base in this area through the inclusion of 14 new studies, along with 4 from the original review (Huertas-Ceballos 2008). Fourteen new pooled analyses were possible.

This review provides some very low-quality evidence for the short-term effectiveness of CBT for the management of children and young people with RAP. When compared to children in control groups, CBT intervention resulted in greater short-term treatment success (pooled OR 5.67, 95% CI 1.18 to 27.32; 4 studies; 175 children). However, this effect was no longer evident at medium-term follow-up (pooled OR 3.08, 95% CI 0.93 to 10.16; Z = 1.85; P = 0.06; 3 studies; 139 children; low-quality evidence) or long-term follow-up (pooled OR 1.29, 95% CI 0.50 to 3.33; 2 studies; 120 children; low-quality evidence). Pooled analyses of other outcomes, pain intensity postintervention and at medium-term follow-up, quality of life measures postintervention, and pain-related functional impairment, did not provide robust evidence of effectiveness.

The review also provides some low-quality evidence for the short-term effectiveness of hypnotherapy. When compared to children in control groups, hypnotherapy resulted in greater treatment success (pooled OR 6.78, 95% CI 2.41 to 19.07; 4 studies; 146 children), along with reductions in both pain intensity (SMD -1.01, 95% CI -1.41 to -0.61; Z = 4.97; P < 0.00001; 4 studies; 146 children) and pain frequency (SMD -1.28, 95% CI -1.84 to -0.72; Z = 4.48; P < 0.00001; 4 studies; 146 children) postintervention. The only study of long-term effectiveness of hypnotherapy reported continued benefit of treatment compared to usual care after five years, with 68% reporting treatment success compared to 20% of controls (P = 0.005).

The review found no robust evidence of effectiveness for yoga therapy. Across three studies (122 children), when compared to children in control groups, yoga therapy resulted in no difference in pain intensity, pain-related functional impairment or measures of social or psychological functioning. A single small study (63 children) reported that written self-disclosure therapy was associated with beneficial effects at six months' follow-up, although not immediately postintervention.

There were no adverse effects for any of the interventions reported. Of the studies measuring psychological and behavioural outcomes, none found any deterioration of mood state or adjustment.

Overall completeness and applicability of evidence

This review highlights a few issues concerning the overall completeness and applicability of the evidence for the benefits and harms of psychosocial interventions for children and adolescents with RAP, that is, the lack of 1) trials conducted in specific subgroups of RAP, as defined by the Rome III criteria (Rasquin 2006); 2) trials assessing the same type of psychosocial intervention; and 3) sustained intervention and follow-up beyond the period of intervention.

It has been suggested that there are distinct subtypes of RAP (Rasquin 2006), and that these could guide treatment choice. We therefore thought it important in this review to estimate, if data allowed, whether RAP subtype predicted response to different treatment modalities. Unfortunately, the majority of studies included children within the broad diagnosis of RAP, which meant that children could be presenting with irritable bowel syndrome, functional abdominal pain, functional dyspepsia, or abdominal migraine. It was therefore not possible to investigate whether particular types of psychosocial interventions benefited particular subgroups of RAP.

Ten of the 18 studies assessed the effectiveness of CBT, whereas there was only a maximum of four studies for the other intervention types. Whilst we were able to pool the data across studies, we were not able to explore the effect of different delivery styles of intervention or dose of intervention, or whether specific components within intervention types were associated with more effectiveness. Reporting of fidelity to intervention was also lacking in the majority of studies.

Lastly, most of the interventions were relatively short in duration (four to six weeks), and very few had medium- or long-term follow-up, which limited our ability to assess whether any benefits are sustained in the long term.

Quality of the evidence

Eighteen studies involving 928 children assessing a variety of psychosocial interventions formed the basis of this evidence. We identified four types of psychosocial intervention, but only CBT was assessed across more than four studies (assessed in nine studies). As evaluated using the GRADE approach (Higgins 2011a), we found the overall quality of evidence within the review to range from very low to low due to the high risk of bias across the studies, such as unblinded participants and unblinded outcome assessment, along with some outcomes having a high level of unexplained heterogeneity (greater than 70%), wide confidence intervals, and a low

number of participants included in the analyses. Future research in this area is therefore likely to impact on our confidence in the estimate of effects observed in this review.

Potential biases in the review process

The present systematic review has many strengths. We developed a protocol for this review according to instructions provided in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). Our protocol was published before we embarked on the review itself (Martin 2014c). We conducted extensive searches of relevant databases and checked forward and backward citations of all included studies. We also contacted authors of included studies for additional data where the presented data was insufficient or missing in order to maximise our ability to pool data on comparable outcomes within comparable intervention types. Two review authors, working independently, selected trials for inclusion and extracted data. Disagreements were resolved by discussion with team members. We assessed the risk of bias in all trials according to the recommendations provided in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b).

We did not include studies that had a mix of children, adolescents, and young adults, where it was not possible to separate the data for children less than 18 years of age. Likewise, we did not include studies that did not specify recruiting children or adolescents, and which presented mean ages of the population sample exceeding 20 years of age. Both of these issues raise the possibility of bias in our review process, as we did not write to these authors asking whether they collected data for children less than 18 years of age. However, we believe this potential bias is not likely to change our conclusions.

Agreements and disagreements with other studies or reviews

The previous version of this review located only six RCTs (Huertas-Ceballos 2008), and due to issues of design in the included studies, was unable to pool data for analysis, thus confidence in the findings was limited. The review concluded that, despite few studies, there was some evidence to suggest that CBT may be a useful intervention for RAP. This update supports and extends these previous findings. We have been able to quantify the magnitude of effect of CBT on treatment success, reductions in pain intensity, and on improved quality of life, as well as present some limited evidence for longer-term effects. The update also presents a new evidence base for more contemporary mind-body therapy approaches for RAP such as hypnotherapy, including guided imagery, and yoga. The findings are in keeping with other systematic reviews of psychosocial interventions for children with RAP. Eccleston 2014 reviewed face-to-face psychological interventions for children with pain dichotomised as headache and non-headache pain. For 13 studies (852 children) of non-headache pain, which was predominantly RAP, they found a medium-sized beneficial effect of treatment on pain intensity (SMD -0.57, 95% CI -0.86 to -0.27; Z = 3.74; P = 0.0002), which compares well to the effects observed across the different interventions in this review. Of interest, they also graded this evidence as very low quality. Rutten 2015 reviewed non-pharmacological approaches for children with RAP. Whilst they did not pool any data, through reviewing individual studies they concluded that there is some evidence for hypnotherapy and CBT, but a lack of evidence for yoga and written self-disclosure. This updated review supports and extends these findings.

Issues for consideration

Overall, the evidence provided by the included studies is relatively weak. First, the majority of studies were short term, assessing outcome effects either immediately postintervention or within three months of the end of the intervention. Few studies investigated whether reported treatment effects were maintained beyond three months. Second, there was considerable variety in the definition and scales used to assess treatment success, as well as in the assessment of other outcomes such as pain intensity, frequency, and duration. For example, treatment success was sometimes assessed as being completely pain free (Sanders 1994), and sometimes as a percentage of reduction in pain from baseline (such as in Gulewitsch 2013 and Vlieger 2007). Likewise, the scales assessing quality of life, functional impairment, and other psychosocial outcomes varied across the studies, which in many cases limited our ability to pool the data. It would be helpful for there to be some consensus on the best instruments to be used consistently across the field of study of paediatric abdominal pain, especially for treatment trials. Third, even within each type of intervention, there was considerable variation in terms of length of weeks of therapy, sessions per week, and in the delivery of intervention. For example, for the CBT intervention, some studies had a one-to-one format between the therapist and child (such as in Levy 2010 and van der Veek 2013), and some used a group format (Duarte 2006; Groß 2013). Similarly, again for CBT, delivery of the intervention was either face-to-face (such as in Robins 2005 and van der Veek 2013), or remotely via a CD-ROM or website (Palermo 2009; Palermo 2016; Wassom 2009). This was also the case for the two guided imagery hypnotherapy studies, with the van Tilburg 2009 study delivering the intervention via CD-ROM and DVD, and Weydert 2006 using a face-to-face intervention. Fourth, there was some evidence of significant differences in outcome findings when studies were assessed according to comparator group used (wait-list control compared to usual care or active controls). Post hoc analyses on pain intensity and functional impairment within CBT intervention suggested that this could affect estimates of treatment effectiveness. Fifth, we did not undertake a global analysis by pooling all the data from all the psychosocial interventions. We consid-

ered that it was not appropriate to do so, as the intervention components and theories about how the specific interventions might work were not similar enough. Finally, it has been suggested that there are distinct subtypes of RAP (Rasquin 2006), and that these could guide treatment choice. We therefore thought it important in this review to estimate, if data allowed, whether RAP subtype predicted response to different treatment modalities. Although all participants met Rome III criteria for RAP (Rasquin 2006), with studies including children classified as having irritable bowel syndrome, functional abdominal pain, and functional dyspepsia, lack of sufficient data by subgroup meant that we were not able to investigate whether there were differences in responsiveness between these groups.

AUTHORS' CONCLUSIONS

Implications for practice

Overall, this review provides low-quality evidence that cognitive behavioural therapy (CBT) and hypnotherapy may be effective in reducing pain in the short term for children and adolescents presenting with recurrent abdominal pain (RAP). Sustained effects of both CBT and hypnotherapy on pain have also been reported, but the evidence to date is limited. There was little evidence that CBT or hypnotherapy affected school functioning, psychological well-being, or quality of life.

This review found no evidence to support the use of yoga or written self-disclosure for the treatment of RAP in children and adolescents.

The review evidence lends support to clinicians who want to consider CBT or hypnotherapy as part of a management strategy for children and adolescents with RAP. However, this evidence needs to be considered alongside the evidence for other approaches in the management of RAP, such as dietary and pharmacological treatments. Companion update reviews of the effectiveness of dietary and pharmacological treatment for RAP will be available soon (Martin 2014a; Martin 2014b).

Implications for research

The evidence for the effectiveness of all psychosocial interventions in children and adolescents with RAP remains weak; in particular, there is a dearth of long-term follow-up data. Further well-designed and reasonably powered trials, giving greater consideration to the nature of the control group, as well as attempts to reduce both performance and detection bias, would improve the rigour of the evidence base. While it may be difficult to remove potential bias in randomised controlled trials of psychotherapy, Button 2015 suggests that improvements can be made by integrating concepts of basic science within applied trials to adjust for these biases, such as elucidating the "active ingredients" of an intervention by using comparative treatments that have one or more components added or removed.

Future research could also consider whether specific aspects of interventions are associated with effectiveness, such as, but not exclusive to, dose, setting of therapy, and on-site versus remote interventions. Trials should also evaluate whether children who meet the criteria for the different forms of RAP according to the Rome III criteria respond differently to potential psychosocial interventions (Rasquin 2006), as well as children who present with and without psychiatric comorbidity. We found a higher proportion of girls to boys across all studies, but there was no exploration of whether the effects of interventions differed according to sex, which also warrants investigation.

Lastly, the precise mechanisms for how the various psychosocial interventions impact RAP are as yet unknown. Closer examination in research studies of changes to cognitive factors and mediating factors such as stress and anxiety throughout the intervention may help shed light on this.

ACKNOWLEDGEMENTS

We acknowledge the work by Angela Huertas-Ceballos, Colin MacArthur, and Stuart Logan on the original protocol (Huertas-Ceballos 2001), and the work by Angela Huertas-Ceballos, Stuart Logan, Cathy Bennett, Sarah See, Colin Macarthur and Morris Zwi on the original review (Huertas-Ceballos 2008).

We are also grateful to the authors who kindly responded to our requests for further information on the trials in which they were involved.

We also wish to warmly thank Geraldine Macdonald (Co-ordinating Editor), Joanne Wilson (Managing Editor), Gemma O'Loughlin (Assistant Managing Editor), and Margaret Anderson (Information Specialist) of Cochrane Developmental, Psychosocial and Learning Problems, for providing help and support.

We are also grateful for the comments and guidance of the reviewers and statisticians who reviewed this update.

REFERENCES

References to studies included in this review

Duarte 2006 *{published data only}*

Duarte MA, Penna FJ, Andrade EM, Cancela CS, Neto JC, Barbosa TF. Treatment of nonorganic recurrent abdominal pain: cognitive-behavioral family intervention. *Journal of Pediatric Gastroenterology & Nutrition* 2006;**43**(1): 59–64. ; DOI: 10.1097/01.mpg.0000226373.10871.76; PUBMED: 16819378]

Evans 2014 *{published and unpublished data}*

Evans S, Lung KC, Seidman LC, Sternlieb B, Zeltzer LK, Tsao JCI. Iyengar yoga for adolescents and young adults with irritable bowel syndrome. *Journal of Pediatric Gastroenterology and Nutrition* 2014;**59**(2):243–53. ; DOI: 10.1097/MPG.0000000000000366; NIHMS614814; PMC4146428]

Groβ 2013 *{published data only}*

Groβ M, Warschburger P. Evaluation of a cognitive-behavioral pain management program for children with chronic abdominal pain: a randomized controlled study. *International Journal of Behavioral Medicine* 2013;**20**(3): 434–43. ; DOI: 10.1007/s12529-012-9228-3; PUBMED: 22328460]

Gulewitsch 2013 *{published data only}*

Gulewitsch MD, Müller J, Hautzinger M, Schlarb AA. Brief hypnotherapeutic-behavioral intervention for functional abdominal pain and irritable bowel syndrome in childhood: a randomized controlled trial. *European Journal Pediatrics* 2013;**172**(8):1043–51. ; DOI: 10.1007/s00431-013-1990-y; PUBMED: 23568514]

Korterink 2016 *{published and unpublished data}*

Korterink JJ, Ockeloen LE, Hilbink MA, Benninga MA, Deckers-Kocken JM. Yoga therapy for abdominal pain-related functional gastrointestinal disorders in children: a randomized controlled trial. *Journal of Pediatric Gastroenterology and Nutrition* 2016;**63**(5):481–7. ; DOI: 10.1097/MPG.0000000000001230; PUBMED: 27050045]

Kuttner 2006 *{published and unpublished data}*

Kuttner L, Chambers CT, Hardial J, Israel DM, Jacobson K, Evans K. A randomized trial of yoga for adolescents with irritable bowel syndrome. *Pain Research & Management* 2006;**11**(4):217–23. ; PMC2673138; PUBMED: 17149454

Levy 2010 *{published and unpublished data}*

* Levy RL, Langer SL, Walker LS, Romano JM, Christie DL, Youssef N, et al. Cognitive-behavioral therapy for children with functional abdominal pain and their parents decreases pain and other symptoms. *American Journal of Gastroenterology* 2010;**105**(4):946–56. ; DOI: 10.1038/ajg.2010.106; PMC2887246; PUBMED: 20216531]

Levy RL, Langer SL, Walker LS, Romano JM, Christie DL, Youssef N, et al. Twelve-month follow-up of cognitive behavioral therapy for children with functional abdominal pain. *JAMA Pediatrics* 2013;**167**(2):178–84. ; DOI: 10.1001/2013.jamapediatrics.282; PMC4167735; PUBMED: 23277304]

Levy RL, Whitehead WE, Walker LS, Langer SL, Romano J, Christie DL, et al. RCT results of a family-based intervention for children with chronic abdominal pain. *Gastroenterology* 2009;**136**(5 Suppl 1):A74. ; DOI: 10.1016/S0016-5085(09)60332-2]

Palermo 2009 *{published and unpublished data}*

Palermo TM, Wilson AC, Peters M, Lewandowski A, Somhegyi H. Randomized controlled trial of an Internet delivered family cognitive behavioral therapy intervention for children and adolescents with chronic pain. *Pain* 2009; **146**(1-2):205–13. ; DOI: 10.1016/j.pain.2009.07.034; NIHMS136567; PMC2760656]

Palermo 2016 *{published and unpublished data}*

Palermo TM, Law EF, Fales J, Bromberg MH, Jessen-Fiddick T, Tai G. Internet-delivered cognitive-behavioral treatment for adolescents with chronic pain and their parents: a randomised controlled multicenter trial. *Pain* 2016;**157** (1):174–85. ; DOI: 10.1097/j.pain.0000000000000348; PMC4852469; PUBMED: 26335910]

Robins 2005 *{published data only}*

Robins PM, Smith SM, Glutting JJ, Bishop CT. A randomized controlled trial of a cognitive-behavioral family intervention for pediatric recurrent abdominal pain. *Journal of Pediatric Psychology* 2005;**30**(5):397–408. ; DOI: 10.1093/jpepsy/jsi063; PUBMED: 15944167]

Sanders 1990 *{published data only}*

Sanders MR, Morrison M, Rebgetz M, Bor W, Dadds M, Shepherd R. Behavioural treatment of childhood recurrent abdominal pain: relationships between pain, children's psychological characteristics and family functioning. *Behavior Change* 1990;**7**(1):16–24. ; DOI: 10.1017/S0813483900007373]

* Sanders MR, Rebgetz M, Morrison M, Bor W, Gordon A, Dadds M, et al. Cognitive-behavioral treatment of recurrent nonspecific abdominal pain in children: an analysis of generalization, maintenance, and side effects. *Journal of Consulting and Clinical Psychology* 1989;**57**(2):294–300. ; PUBMED: 2708618

Sanders 1994 *{published data only}*

Sanders MR, Shepherd RW, Cleghorn G, Woolford H. The treatment of recurrent abdominal pain in children: a controlled comparison of cognitive-behavioral family intervention and standard pediatric care. *Journal of Consulting and Clinical Psychology* 1994;**62**(2):306–14. ; PUBMED: 8201068

van der Veek 2013 *{published and unpublished data}*

* van der Veek SM, Derkx BH, Benninga MA, Boer F, de Haan E. Cognitive behavior therapy for pediatric functional abdominal pain: a randomized controlled trial. *Pediatrics*

2013;**132**(5):e1163–72. ; DOI: 10.1542/peds.2013-0242; PUBMED: 24127467]
van der Veek SM, Derkx BH, De Haan E, Benninga MA, Boer F. Cognitive behavioural therapy for pediatric functional abdominal pain: results of a randomized controlled trial. *Journal of Pediatric Gastroenterology and Nutrition* 2011;**53**(Suppl 2):S63–4.

van Tilburg 2009 *{published and unpublished data}*
van Tilburg MAL, Chitkara DK, Palsson OS, Turner M, Blois-Martin N, Ulshen M, et al. Audio-recorded guided imagery treatment reduces functional abdominal pain in children: a pilot study. *Pediatrics* 2009;**124**(5):e890–7. ; DOI: 10.1542/peds.2009-0028; PUBMED: 19822590]

Vlieger 2007 *{published and unpublished data}*
* Vlieger AM, Menko-Frankenhuis C, Wolfkamp SCS, Tromp E, Benninga MA. Hypnotherapy for children with functional abdominal pain or irritable bowel syndrome: a randomized controlled trial. *Gastroenterology* 2007;**133**(5): 1430–6. ; DOI: 10.1053/j.gastro.2007.08.072; PUBMED: 17919634]
Vlieger AM, Rutten JM, Govers AM, Frankenhuis C, Benninga MA. Long term follow-up of gut-directed hypnotherapy versus standard care in children with functional abdominal pain or irritable bowel syndrome. *Gastroenterology* 2012;**May**:S134.
Vlieger AM, Rutten JMTM, Govers AMAP, Frankenhuis C, Benninga MA. Long-term follow-up of gut-directed hypnotherapy vs. standard care in children with functional abdominal pain or irritable bowel syndrome. *American Journal of Gastroenterology* 2012;**107**(4):627–31. ; DOI: 10.1038/ajg.2011.487]
Vlieger AM, van den Berg MM, Menko-Frankenhuis C, Bongers MEJ, Tromp E, Benninga MA. No change in rectal sensitivity after gut-directed hypnotherapy in children with functional abdominal pain or irritable bowel syndrome. *American Journal of Gastroenterology* 2010;**105**:213–8.

Wallander 2011 *{published data only}*
Wallander JL, Madan-Swain A, Klapow J, Saeed S. A randomised controlled trial of written self-disclosure for functional recurrent abdominal pain in youth. *Psychology & Health* 2011;**26**(4):433–47. ; DOI: 10.1080/08870440903477212; PMC2911499; PUBMED: 20419562]

Wassom 2009 *{published data only}*
* Wassom MC. A minimal contact cognitive-behavioral intervention for abdominal pain-related functional gastrointestinal disorders: pilot-study of "Gutstrong". *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2009;**69**(8-B):5064.
Wassom MC, Schurman JV, Friesen CA, Rapoff MA. A pilot study of "Gutstrong" for adolescents with functional gastrointestinal disorders. *Clinical Practice in Pediatric Psychology* 2013;**1**(3):201–13. ; DOI: 10.1037/cpp0000025]

Weydert 2006 *{published data only}*
Weydert JA, Shapiro DE, Acra SA, Monheim CJ, Chambers AS, Ball TM. Evaluation of guided imagery as treatment for recurrent abdominal pain in children: a randomized controlled trial. *BMC Pediatrics* 2006;**6**:29. ; DOI: 10.1186/1471-2431-6-29; PMC1660537]

References to studies excluded from this review

Alfvén 2007 *{published data only}*
Alfvén G, Lindstrom A. A new method for the treatment of recurrent abdominal pain of prolonged negative stress origin. *Acta Paediatrica* 2007;**96**(1):76–81. ; DOI: 10.1111/j.1651-2227.2006.00028.x; PUBMED: 17187609]

Bursch 2008 *{published data only}*
Bursch B. Psychological/cognitive behavioral treatment of childhood functional abdominal pain and irritable bowel syndrome. *Journal of Pediatric Gastroenterology and Nutrition* 2008;**47**(5):706–7. ; DOI: 10.1097/01.mpg.0000338967.47679.e9; PUBMED: 18955885]

Hicks 2003 *{published data only}*
* Hicks CL. *Online Psychological Treatment for Pediatric Recurrent Pain: a Randomised Evaluation (PhD thesis)*. Saskatchewan, Canada: The University of Saskatchewan, 2003.
Hicks CL, von Baeyer CL, McGrath PJ. Online psychological treatment for pediatric recurrent pain: a randomized evaluation. *Journal of Pediatric Psychology* 2006; **31**(7):724–36. ; DOI: 10.1093/jpepsy/jsj065; PUBMED: 16093516]

Humphreys 1998 *{published data only}*
* Humphreys PA. *Components Analysis of Behavioural Treatment of Children Diagnosed With Recurrent Abdominal Pain (RAP) (PhD thesis)*. San Diego, CA: Californian School of Professional Psychology, 1998.
Humphreys PA, Gevirtz RN. Treatment of recurrent abdominal pain: components analysis of four treatment protocols. *Journal of Pediatric Gastroenterology and Nutrition* 2000;**31**(1):47–51. ; PUBMED: 10896070

Long 2009 *{published data only}*
Long A, Songhegyi K, Palermo T. Daytime activity levels in adolescents with chronic pain: treatment outcome in a randomized controlled trial of web-based cognitive-behavioral therapy. *Journal of Pain* 2009;**10**(4 Suppl):S63. ; DOI: 10.1016/j.jpain.2009.01.218]

Sato 2009 *{published data only}*
Sato AF, Clifford LM, Silverman AH, Davies WH. Cognitive-behavioral interventions via telehealth: applications to pediatric functional abdominal pain. *Children's Health Care* 2009;**38**(1):1–22. ; DOI: 10.1080/02739610802615724]

Sieberg 2010 *{published data only}*
Sieberg CB. Therapeutic treatments for children with recurrent abdominal pain: a multiple baseline pilot study. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2010;**71**:2061.

Sieberg 2011 *{published data only}*
Sieberg CB, Flannery-Schroeder E, Plante W. Children with co-morbid recurrent abdominal pain and anxiety disorders: results from a multiple-baseline intervention study. *Journal of Child Health Care* 2011;**15**(2):126–39. ; DOI: 10.1177/1367493511401640; PUBMED: 21685228]

van Barreveld 2015 *{published data only}*
van Barreveld M, Rutten J, Vlieger A, Frankenhuis C, George E, Groeneweg M, et al. Cost-effectiveness and cost-utility of home-based hypnotherapy using compact disc versus individual hypnotherapy by a therapist for pediatric irritable bowel syndrome and functional abdominal pain (syndrome). *Value Health* 2015;**18**(7):A628. ; DOI: 10.1016/j.jval.2015.09.2214; PUBMED: 26533527]

Warner 2011 *{published data only}*
Warner CM, Colognori D, Kim RE, Reigada LC, Klein RG, Browner-Elhanan KJ, et al. Cognitive-behavioral treatment of persistent functional somatic complaints and pediatric anxiety: an initial controlled trial. *Depression and Anxiety* 2011;**28**(7):551–9.

References to studies awaiting assessment

Tannen 2014 *{published data only}*
Tannen A. Cognitive training for children with functional abdominal pain and for their parents. *Pflege Zeitschrift* 2014;**67**(9):541. ; PUBMED: 25265696

Youssef 2009 *{published data only}*
Youssef NN, van Tilburg MA, Matta EN, Langseder A, Whitehead WE. Feasibility and efficacy of pilot study investigating a school nurse administered guided imagery program for childhood functional abdominal pain. *Gastroenterology* 2009;**136**(5 Suppl 1):A156–7. ; DOI: 10.1016/S0016-5085(09)60705-8]

Additional references

Abu-Arafeh 1995
Abu-Arafeh I, Russell G. Prevalence and clinical features of abdominal migraine compared with those of migraine headache. *Archives of Disease in Childhood* 1995;**72**(5):413–7. [PUBMED: 7618907]

Achenbach 1983
Achenbach TM, Edelbrook C. *Manual for the Child Behavior Checklist and Revised Behavior Profile*. Burlington: University of Vermont, Department of Psychiatry, 1983.

Apley 1958
Apley J, Naish N. Recurrent abdominal pains: a field survey of 1000 school children. *Archives of Disease in Childhood* 1958;**33**(168):165–70. [PMC2012205]

Apley 1975
Apley J. *The Child with Abdominal Pains*. 2nd Edition. Oxford: Blackwell Scientific Publications, 1975.

Bain 1974
Bain HW. Chronic vague abdominal pain in children. *Pediatric Clinics of North America* 1974;**21**(4):991–1000. [PUBMED: 4427792]

Balshem 2011
Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(4):401–6. [DOI: 10.1016/j.jclinepi.2010.07.015]

Bieri 1990
Bieri D, Reeve RA, Champion GD, Addicoat L, Ziegler JB. The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. *Pain* 1990;**41**(2):139–50. [DOI: 10.1016/0304-3959(90)90018-9]

Button 2015
Button KS, Munafò MR. Addressing risk of bias in trials of cognitive behavioral therapy. *Shanghai Archives of Psychiatry* 2015;**27**(3):144–8. [DOI: 10.11919/j.issn.1002-0829.215042; PMC4526826]

Campo 2001
Campo JV, Di Lorenzo C, Chappetta L, Bridge JD, Colborn K, Gartner JC, et al. Adult outcomes of pediatric recurrent abdominal pain: do they just grow out of it?. *Pediatrics* 2001;**108**(1):e1. [DOI: 10.1542/peds.108.1.e1]

Campo 2004
Campo JV, Bridge J, Ehmann M, Altman S, Lucas A, Birmaher B, et al. Recurrent abdominal pain, anxiety and depression in primary care. *Pediatrics* 2004;**113**(4):817–24. [PUBMED: 15060233]

Connelly 2006
Connelly M, Rapoff MA, Thompson N, Connelly W. Headstrong: a pilot study of a CD-ROM intervention for recurrent pediatric headache. *Journal of Pediatric Psychology* 2006;**31**(7):737–47. [DOI: 10.1093/jpepsy/jsj003; PUBMED: 16861397]

Conners 1969
Conners CK. A teacher rating scale for use in drug studies in children. *American Journal of Psychiatry* 1969;**126**(6):884–8. [DOI: 10.1176/ajp.126.6.884; PUBMED: 4900822]

Cunningham 2013
Cunningham JA, Kypri K, McCambridge J. Exploratory randomized controlled trial evaluating the impact of a waiting list control design. *BMC Medical Research Methodology* 2013;**13**:150. [DOI: 10.1186/1471-2288-13-150]

Derogatis 2000
Derogatis LR. *Brief Symptom Inventory (BSI) 18: Administration, Scoring and Procedures Manual*. Minneapolis: NCS Pearson, 2001.

DerSimonian 1986
DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177–88. [DOI: 10.1016/0197-2456(86)90046-2]

Devanarayana 2011
Devanarayana NM, Mettananda S, Liyanarachchi C, Nanayakkara N, Mendis N, Perera N, et al. Abdominal

pain - predominant functional gastrointestinal diseases in children and adolescents: prevalence, symptomatology, and association with emotional stress. *Journal of Pediatric Gastroenterology and Nutrition* 2011;**53**(6):659–65. [DOI: 10.1097/MPG.0b013e3182296033; PUBMED: 21697745]

Di Lorenzo 2001

Di Lorenzo C, Youssef NN, Sigurdsson L, Scharff L, Griffiths J, Wald A. Visceral hyperalgesia in children with functional abdominal pain. *Journal of Pediatrics* 2001; **139**(6):838–43. [DOI: 10.1067/mpd.2001.118883; PUBMED: 11743510]

Drossman 1995

Drossman DA, Creed FH, Fava GA, Olden KW, Patrick DL, Toner BB, et al. Psychosocial aspects of the functional gastrointestinal disorders. *Gastroenterology International* 1995;**8**(2):47–90.

Dufton 2009

Dufton LM, Dunn MJ, Compas BE. Anxiety and somatic complaints in children with recurrent abdominal pain and anxiety disorders. *Journal of Pediatric Psychology* 2009;**34**(2): 176–86. [DOI: 10.1093/jpepsy/jsn064; NIHMS58564; PMC2645470]

Eccleston 2014

Eccleston C, Palermo TM, Williams AC, Lewandowski Holley A, Morley S, Fisher E, et al. Psychological therapies for the management of chronic and recurrent pain in children and adolescents. *Cochrane Database of Systematic Reviews* 2014, Issue 5. [DOI: 10.1002/14651858.CD003968.pub4]

Edwards 1991

Edwards MC, Finney JW, Bonner M. Matching treatment with recurrent abdominal pain symptoms: an evaluation of dietary fiber and relaxation treatments. *Behavior Therapy* 1991;**22**(2):257–67.

Fekkes 2001

Fekkes M, Kamphuis RP, Ottenkamp J, Verrips E, Vogels T, Kamphuis M, et al. Health-related quality of life in young adults with minor congential heart disease. *Psychology & Health* 2001;**16**(2):239–51. [DOI: 10.1080/08870440108405502]

Finney 1989

Finney JW, Lemanek KL, Cataldo MF, Katz HP, Fuqua RW. Pediatric psychology in primary health care: brief targeted therapy for recurrent abdominal pain. *Behavior Therapy* 1989;**20**(2):283–91. [DOI: 10.1016/S0005-7894(89)80074-7]

Friedman 1972

Friedman R. Some characteristics of children with psychogenic pain. *Clinical Pediatrics* 1972;**11**(6):331–3. [DOI: 10.1177/000992287201100606]

Good 1995

Good PA. Neurologic investigations of childhood abdominal migraine: a combined electrophysiologic approach to diagnosis. *Journal of Pediatric Gastroenterology and Nutrition* 1995;**21 Suppl 1**:S44–8. [PUBMED: 8708868]

Guyatt 2013

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing summary of findings tables - binary outcomes. *Journal for Clinical Epidemiology* 2013;**66**(2):158–72. [DOI: 10.1016/j.jclinepi.2012.01.012; PUBMED: 22609141]

Harrison 2004

Harrison LJ, Manocha R, Rubia K. Sahaja yoga meditation as a family treatment program for children with Attention Deficit-Hyperactivity Disorder. *Clinical Child Psychology and Psychiatry* 2004;**9**(4):479–97. [DOI: 10.1177/1359104504046155]

Harvey 1973

Harvey RF, Pomare EW, Heaton KW. Effects of increased dietary fibre on intestinal transit. *Lancet* 1973;**301**(7815): 1278–80. [DOI: 10.1016/S0140-6736(73)91294-4]

Hicks 2001

Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale - Revised: toward a common metric in pediatric pain assessment. *Pain* 2001;**93**(2):173–83. [PUBMED: 11427329]

Higgins 2003

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**: 557–60. [DOI: 10.1136/bmj.327.7414.557]

Higgins 2011a

Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011b

Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011c

Higgins JPT, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. In: Higgins JP, Green S, editor (s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Hockaday 1992

Hockaday JM. Is there a place for "abdominal migraine" as a separate entity in the IHS classification? No!. *Cephalalgia* 1992;**12**(6):346–8. [PUBMED: 1473134]

Huertas-Ceballos 2009a

Huertas-Ceballos AA, Logan S, Bennett C, Macarthur C. Dietary interventions for recurrent abdominal pain (RAP) and irritable bowel syndrome (IBS) in childhood. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD003019.pub3]

Huertas-Ceballos 2009b
Huertas-Ceballos AA, Logan S, Bennett C, Macarthur C. Pharmacological interventions for recurrent abdominal pain (RAP) and irritable bowel syndrome (IBS) in childhood. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD003017.pub2]

Hyams 1995
Hyams JS, Treem WR, Justinich CJ, Davis P, Shoup M, Burke G. Characterization of symptoms in children with recurrent abdominal pain: resemblance to irritable bowel syndrome. *Journal of Pediatric Gastroenterology and Nutrition* 1995;**20**(2):209–14. [PUBMED: 7714688]

Hübner 2009
Hübner B, Hechler T, Dobe M, Damschen U, Kosfelder J, Denecke H, et al. Pain-related disability in adolescents suffering from chronic pain. Preliminary examination of the Pediatric Pain Disability Index (P-PDI). *Schmerz* 2009;**23**(1):20–32. [DOI: 10.1007/s00482-008-0730-0; PUBMED: 18941801]

Kaminsky 2006
Kaminsky L, Robertson M, Dewey D. Psychological correlates of depression in children with recurrent abdominal pain. *Journal of Pediatric Psychology* 2006;**31**(9): 956–66. [DOI: 10.1093/jpepsy/jsj103]

Kennedy 2012
Kennedy PJ, Clarke G, Quigley E, Groeger JA, Dinan TG, Cryan JF. Gut memories: towards a cognitive neurobiology of irritable bowel syndrome. *Neuroscience and Biobehavioural Reviews* 2012;**36**(1):310–40. [DOI: 10.1016/j.neubiorev.2011.07.001; PUBMED: 21777613]

Konijnenberg 2005
Konijnenberg AY, Uiterwaal CS, Kimpen JL, van der Hoeven J, Buitelaar JK, de Graeff-Meeder ER. Children with unexplained chronic pain: substantial impairment in everyday life. *Archives of Disease in Childhood* 2005;**90**(7): 680–6. [DOI: 10.1136/adc.2004.056820; PMC1720481; PUBMED: 15899922]

Kovacs 1992
Kovacs M. *Children's Depression Inventory: Manual.* New York: Multi-Health Systems, 1992.

Lee 1996
Lee LH, Olness KN. Effects of self-induced mental imagery on autonomic reactivity in children. *Journal of Developmental & Behavioral Pediatrics* 1996;**17**(5):323–7. [PUBMED: 8897220]

Liebman 1976
Liebman R, Honig P, Berger H. An integrated treatment program for psychogenic pain. *Family Process* 1976;**15**(4): 397–405. [PUBMED: 1026457]

Martin 2014a
Martin AE, Newlove-Delgado TV, Abbott RA, Bethel A, Thompson-Coon J, Nikolaou V, et al. Dietary interventions for recurrent abdominal pain in childhood. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010972]

Martin 2014b
Martin AE, Newlove-Delgado TV, Abbott RA, Bethel A, Thompson-Coon J, Nikolaou V, et al. Pharmacological interventions for recurrent abdominal pain in childhood. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010973]

Mayer 2002
Mayer EA, Collins SM. Evolving pathophysiologic models of functional gastrointestinal disorders. *Gastroenterology* 2002;**122**(7):2032–48. [PUBMED: 12055608]

McGrath 1991
McGrath PA. *Pain in Children: Nature, Assessment, and Treatment*. New York: Guilford Publications, 1991.

McOmber 2007
McOmber ME, Schulman MJ. Recurrent abdominal pain and irritable bowel syndrome in children. *Current Opinion in Pediatrics* 2007;**19**(5):581–5.

Milla 1999
Milla PJ. Acquired motility disorders in childhood. *Canadian Journal of Gastroenterology* 1999;**13 Suppl A**: 76–84A. [DOI: 10.1155/1999/610486]

Moher 2009
Moher D, Liberati A, Tetzlaff J, Altman DG for the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. [DOI: http://dx.doi.org/10.1136/bmj.b2535]

Muris 2002
Muris P, Meesters C, Schouten E. A brief questionnaire of DSM-IV-defined anxiety and depression symptoms among children. *Clinical Psychology and Psychotherapy* 2002;**9**(6): 430–42. [DOI: 10.1002/cpp.347]

Osborne 1989
Osborne RB, Hatcher JW, Richtsmeier AJ. The role of social modelling in unexplained pediatric pain. *Journal of Pediatric Psychology* 1989;**14**(1):43–61. [DOI: 10.1093/jpepsy/14.1.43]

Painter 1964
Painter NS, Truelove SC. Irritable bowel syndrome in childhood. *Gut* 1964;**5**:201.

Palermo 2004
Palermo TM, Witherspoon D, Valenzuela D, Drotar DD. Development and validation of the Child Activity Limitations Interview: a measure of pain-related functional impairment in school-age children and adolescents. *Pain* 2004;**109**(3):461–70. [DOI: 10.1016/j.pain.2004.02.023]

Paul 2013
Paul SP, Candy DC. Clinical update: recurrent abdominal pain in children. *Community Practice* 2013;**86**(11):48–51. [PUBMED: 24369571]

Peltonen 1970
Peltonen T. Recurrent abdominal pain. *Pediatrics* 1970;**46** (6):973.

Pennebaker 2007
Pennebaker JW, Chung CK. Expressive writing, emotional upheavals, and health. In: Friedman HS, Silver RC editor (s). *Foundations of Health Psychology*. New York: Oxford University Press, 2007:263–84.

Quay 1983
Quay HC, Peterson DR. Interim Manual for the Revised Behavior Checklist. *Interim Manual for the Revised Behavior Checklist*. Miami: University of Miami, 1983.

Ramchandani 2006
Ramchandani PG, Stein A, Hotopf M, Wiles NJ, The ALSPAC Study Team. Early parental and child predictors of recurrent abdominal pain at school age: results of a large population-based study. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;**45**(6):729–36. [DOI: 10.1097/01.chi.0000215329.35928.e0; PUBMED: 16721323]

Rasquin 2006
Rasquin A, Di Lorenzo C, Forbes D, Guiraldes E, Hyams JS, Staiano A, et al. Childhood functional gastrointestinal disorders: child/adolescent. *Gastroenterology* 2006;**130** (5):1527-37. [DOI: 10.1053/j.gastro.2005.08.063; PUBMED: 16678566]

Rasquin-Weber 1999
Rasquin-Weber A, Hymen PE, Cucchiara S, Fleisher DR, Hyams JS, Milla PJ, et al. Childhood functional gastrointestinal disorders. *Gut* 1999;**45**:II60–8. [DOI: 10.1136/gut.45.2008.ii60]

Ravens-Sieberer 1998
Ravens-Sieberer U, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. *Quality of Life Research* 1998;**7**(5):399–407. [PUBMED: 9691720]

Ravens-Sieberer 2005
Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Duer W, et al. KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research* 2005;**5**(3):353-64. [DOI: 10.1586/14737167.5.3.353; PUBMED: 19807604]

Raymer 1984
Raymer D, Weininger O, Hamilton JR. Psychological problems in children with abdominal pain. *Lancet* 1984;**1** (8374):439–40. [PUBMED: 6142160]

Review Manager 2014 [Computer program]
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Reynolds 1985
Reynolds CR, Richmond BO. *Revised Children's Manifest Anxiety Scale (RCMAS) Manual*. Los Angeles: Western Psychological Services, 1985.

Rubin 1967
Rubin LS, Barbero GL, Sibinga MS. Pupillary reactivity in children with recurrent abdominal pain. *Psychosomatic Medicine* 1967;**2**:111–20.

Rutten 2015
Rutten JM, Korterink JJ, Venmans LM, Benninga MA, Tabbers MM. Nonpharmacologic treatment of functional abdominal pain disorders: a systematic review. *Pediatrics* 2015;**135**(3):2014–123. [DOI: 10.1542/peds.2014-2123]

Sandberg 1973
Sandberg DH. Additional references on the tension-fatigue syndrome. *Pediatrics* 1973;**51**(2):308–9. [PUBMED: 4572072]

Sanders 1981
Sanders MR, Glynn T. Training parents in behavioral self-management: an analysis of generalization and management. *Journal of Applied Behavior Analysis* 1981;**14** (3):223–37. [PMC1308209]

Schanberg 2000
Schanberg LE, Sandstrom MJ, Starr K, Gil KM, Lefebvre JC, Keefe FJ, et al. The relationship of daily mood and stressful events to symptoms in juvenile rheumatic disease. *Arthritis Care Research* 2000;**13**(1):33–41. [PUBMED: 11094924]

Scharff 1997
Scharff L. Recurrent abdominal pain in children: a review of psychological factors and treatment. *Clinical Psychology Review* 1997;**17**(2):145–66. [PUBMED: 9140713]

Schünemann 2011
Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Shelby 2013
Shelby GD, Shirkey KC, Sherman AL, Beck JE, Haman K, Shears AR, et al. Functional abdominal pain in childhood and long-term vulnerability to anxiety disorders. *Pediatrics* 2013;**132**(3):475–82. [DOI: 10.1542/peds.2012-2191]

Speer 1954
Speer F. The allergic tension-fatigue syndrome. *Pediatric Clinics of North America* 1954;**11**:1029–38. [PUBMED: 13204094]

Stone 1970
Stone RT, Barbero GJ. Recurrent abdominal pain in childhood. *Pediatrics* 1970;**45**(5):732–8.

Stowens 1970
Stowens D. Recurrent abdominal pain. *Pediatrics* 1970;**46** (6):968–9.

Størdal 2005
Størdal K, Nygaard EA, Bentsen BS. Recurrent abdominal pain: a five year follow-up study. *Acta Paediatrica* 2005;**94** (2):234–6. [PUBMED: 15981760]

Sutton 2000
Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta analysis. *BMJ* 2000;**320**(7249):1574–7. [DOI: 10.1136/bmj.320.7249.1574]

Symon 1986
Symon DN, Russell G. Abdominal migraine: a childhood syndrome defined. *Cephalalgia* 1986;**6**(4):223–8. [PUBMED: 3802189]

Thompson 1989
Thompson WG, Dotevall G, Drossman DA, Heaton KW, Druis W. Irritable bowel syndrome: guidelines for diagnosis. *Gastroenterology International* 1989;**2**:92–5.

Van Ginkel 2001
Van Ginkel R, Voskuijl WP, Benninga MA, Taminiau JA, Boeckxstaens GE. Alterations in rectal sensitivity and motility in childhood irritable bowel syndrome. *Gastroenterology* 2001;**120**(1):31–8. [PUBMED: 11208711]

Van Slyke 2006
Van Slyke DA, Walker LS. Mothers' response to children's pain. *Clinical Journal of Pain* 2006;**22**(4):387–91. [DOI: 10.1097/01.ajp.0000205257.80044.01; PUBMED: 16691093]

Varni 2001
Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Medical Care* 2001;**39**(8):800–12. [PUBMED: 11468499]

Vasquez 1992
Vasquez R, Calvo M, Martinez A, Chavez P. Recurrent abdominal pain in children [Dolor abdominal recurrente en ninos]. *Actualizaciones Pediatricas* 1992;**2**(1):22–7.

Vlieger 2012
Vlieger AM, Rutten JM, Govers AM, Frankenhuis C, Benninga MA. Long-term follow-up of gut-directed hypnotherapy vs. standard care in children with functional abdominal pain or irritable bowel syndrome. *Gastroenterology* 2012;**107**(4):627–31. [DOI: 10.1038/ajg.2011.487; PUBMED: 22310221]

Vogels 1998
Vogels AGC, Verrips GH, Verloove-Vanhorick SP, Fekkes M, Kamphuis RP, Koopman H, et al. Measuring health related quality of life in children: the development of the TACQOL parent form. *Quality of Life Research* 1998;**7**: 457–65.

von Baeyer 2014 [pers comm]
Von Baeyer C. [personal communication] [Re: Paper 2006 'Online psychological treatment for pediatric recurrent pain: a randomized evaluation']. Email correspondence with: Ms R Whear (RAP review team, University of Exeter, UK) 2 April 2014.

Walker 1991
Walker LS, Greene JW. The Functional Disability Inventory: measuring a neglected dimension of child health status. *Journal of Pediatric Psychology* 1991;**16**(1):39–58. [PUBMED: 1826329]

Walker 1993
Walker LS, Garber J, Greene JW. Psychosocial correlates of recurrent childhood pain: a comparison of pediatric patients with recurrent abdominal pain, organic illness, and psychiatric disorders. *Journal of Abnormal Psychology* 1993; **102**(2):248–58. [PUBMED: 8315137]

Walker 1995
Walker LS, Garber J, Van Slyke DA, Greene JW. Long-term health outcomes in patients with recurrent abdominal pain. *Journal of Pediatric Psychology* 1995;**20**(2):233–45. [PUBMED: 7760222]

Walker 1997
Walker LS, Smith CA, Garber J, Van Slyke DA. Development and validation of the Pain Response Inventory for children. *Psychological Assessment* 1997;**9**(4):392–405. [DOI: 10.1037/1040-3590.9.4.392]

Walker 1998
Walker LS, Guite JW, Duke M, Barnard JA, Greene JW. Recurrent abdominal pain: a potential precursor of irritable bowel syndrome in adolescents and young adults. *Journal of Pediatrics* 1998;**132**(6):1010–5. [PUBMED: 9627595]

Walker 1999
Walker LS. The evolution of research on recurrent abdominal pain: history, assumptions and conceptual model. In: McGrath PJ, Finley GA editor(s). *Chronic and Recurrent Pain in Children and Adolescents. Progress in Pain Research and Management*. Vol. **13**, Seattle: IASP Press, 1999:141–71.

Walker 2006
Walker LS, Williams SE, Smith CA, Garber J, Van Slyke DA, Lipani TA. Parent attention versus distraction: impact on symptom complaints by children with and without chronic functional abdominal pain. *Journal of Abnormal Child Psychology* 2006;**122**(1-2):43–52. [DOI: 10.1016/j.pain.2005.12.020; PMC3232036; PUBMED: 16495006]

Walker 2007
Walker LS, Smith CA, Garber J, Claar RL. Appraisal and coping with daily stressors by pediatric patients with chronic abdominal pain. *Journal of Pediatric Psychology* 2007;**32**(2): 206–16. [DOI: 10.1093/jpepsy/jsj124; NIHMS311267; PMC3150731]

Ware 1992
Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36) I. Conceptual framework and item selection. *Medical Care* 1992;**30**(6):473–83. [PUBMED: 1593914]

Wetchler 1992
Wetchler JL. Family treatment of a thirteen-year-old with chronic stomach pain. *Journal of Family Psychotherapy* 1992; **3**(4):1–14. [DOI: 10.1300/j085V03N04˙01]

Williams 1996
Williams K, Chambers M, Logan S, Robinson D. Association of common health symptoms with bullying

in primary school children. *BMJ* 1996;**313**(7048):17–9. [PMC2351438]

Williams 2005

Williams KA, Petronis J, Smith D, Goodrich D, Wu J, Ravi N, et al. Effects of Iyengar yoga therapy for chronic low back pain. *Pain* 2005;**115**(1-2):107–17. [DOI: 10.1016/j.pain.2005.02.016; PUBMED: 15836974]

Youssef 2006

Youssef NN, Murphy TG, Langseder AL, Rosh JR. Quality of life for children with functional abdominal pain: a comparison study of patients' and parents' perceptions. *Pediatrics* 2006;**117**(1):54–9. [DOI: 10.1542/peds.2005-0114; PUBMED: 16396860]

Youssef 2008

Youssef NN, Atienza K, Langseder AL, Strauss RS. Chronic abdominal pain and depressive symptoms: analysis of the national longitudinal study of adolescent health. *Clinical Gastroenterology and Hepatology* 2008;**6**(3):329–32. [DOI: 10.1016/j.cgh.2007.12.019; PUBMED: 18258491]

References to other published versions of this review

Huertas-Ceballos 2001

Huertas-Ceballos A, MacArthur C, Logan S. Psychosocial interventions for recurrent abdominal pain (RAP) in childhood. *Cochrane Database of Systematic Reviews* 2001, Issue 2. [DOI: 10.1002/14651858.CD003014]

Huertas-Ceballos 2008

Huertas-Ceballos A, Macarthur C, Logan S. Psychosocial interventions for recurrent abdominal pain (RAP) and irritable bowel syndrome (IBS) in childhood. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD003014.pub2]

Martin 2014c

Martin AE, Newlove-Delgado TV, Abbott RA, Bethel A, Thompson-Coon J, Nikolaou V, et al. Psychosocial interventions for recurrent abdominal pain in childhood. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010971]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Duarte 2006

Methods	RCT with usual care control **Follow-up:** postintervention (4 months) follow-up
Participants	**Location:** Brazil **Setting:** paediatric gastroenterology clinic **Sample size:** 32 children (15 intervention, 17 control) **Sex:** 10 boys, 22 girls **Dropouts/withdrawals:** 0 **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** intervention: 9.9 (SD 2.2; range not reported) years; control: 8.8 (SD 2.0; range not reported) years. Data not reported for groups overall
Interventions	**Intervention:** cognitive behavioural family intervention (group format), 50 minutes/month x 4 (2 x child, 2 x parent) **Control:** usual care (usual paediatric care and advice on nutrition, intestinal parasite prophylaxis, and accident prevention), 50 minutes/month x 4
Outcomes	1. Pain frequency and intensity (50-centimetre visual analogue scale)
Notes	**Study dates:** January 2003 to December 2004 **Funding:** not stated **Declarations of interest:** not reported

Risk of bias *Risk of bias*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	**Comment:** method not reported
Allocation concealment (selection bias)	Unclear risk	**Comment:** method not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of group allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** unclear whether parents could influence visual analogue scale, and who was assessing pain thresholds
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** follow-up complete

Duarte 2006 (*Continued*)

Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Unclear risk	**Comment:** no power calculations, compliance not reported

Evans 2014

Methods	RCT with wait-list control **Follow-up:** postintervention (6 weeks) and 2 months' follow-up
Participants	**Location:** USA **Setting:** recruited through physician referrals, advertisements in gastroenterology clinics, and community bulletin boards **Sample size:** 30 children (18 intervention, 12 control) and 21 young adults (data for these not extracted) **Sex:** 5 boys, 25 girls **Dropouts/withdrawals:** 1 control **Diagnosis:** RAP diagnosed using Rome III (Rasquin 2006) **Mean age:** intervention: 16.4 (SD not reported; range 14 to 17) years; control: 15.9 (SD not reported; range 14 to 17) years. Data not reported for groups overall
Interventions	**Intervention:** Iyengar yoga, 2 x 90-minute classes per week for 6 weeks, with a maximum of 6 participants per class, held at a paediatric pain yoga studio on a university campus **Control:** wait-list control (no other details reported)
Outcomes	1. Pain intensity (visual analogue scale, 0 to 10) 2. Functional disability (Functional Disability Inventory; Walker 1991) 3. Quality of life (36-Item Short Form Health Survey (SF-36); Ware 1992) 4. Psychological distress (Brief Symptom Inventory-18; Derogatis 2000)
Notes	**Study dates:** recruitment from October 2009 and May 2013 **Funding:** National Center for Complementary and Alternative Medicine grant K01AT005093, an Oppenheimer Seed Grant for Complementary, Alternative and Integrative Medicine, and by the University of California, Los Angeles Clinical and Translational Research Center, Clinical and Translational Science Institute Grant UL1TR000124 **Declarations of interest:** Authors report no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** randomiser program used by researcher not involved in project
Allocation concealment (selection bias)	Low risk	**Comment:** randomiser program used by researcher not involved in project

Evans 2014 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants aware of treatment
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** participants self reporting outcomes
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	**Comment:** not all participants fully accounted for, high attrition rates
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Unclear risk	**Comment:** queries regarding data analysis procedures

Groβ 2013

Methods	RCT with wait-list control **Follow-up:** postintervention (6 weeks) and 3 months' follow-up
Participants	**Location:** Germany **Setting:** recruited from an ongoing epidemiological study of schoolchildren **Sample size:** 29 children (15 intervention, 14 control) **Sex:** 4 boys, 25 girls **Dropouts/withdrawals:** 0 **Diagnosis:** RAP diagnosed using Rome III (Rasquin 2006) **Mean age:** intervention: 9.2 (SD 1.5; range 6.6 to 11.2) years; control 10.1 (SD 1.4; range 8.0 to 11.9) years. Data not reported for groups overall
Interventions	**Intervention:** cognitive behavioural pain management program for child and parent: 'Stop the pain with Happy Pingu' (group format), 90 minutes/week x 6. Children had a CD-ROM with relaxation exercises to do as homework **Control:** wait-list control (no other details reported)
Outcomes	1. Pain frequency, intensity and duration (pain diary) 2. Health-related quality of life (PedsQL; Varni 2001) 3. Pain impairment (KINDL-R; Ravens-Sieberer 2005)
Notes	**Study dates:** not reported **Funding:** grant from Potsdam Graduate School **Declarations of interest:** Authors report no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**

Groβ 2013 (*Continued*)

Random sequence generation (selection bias)	Low risk	**Comment:** computer-aided randomisation
Allocation concealment (selection bias)	Low risk	**Comment:** randomisation carried out by person not involved in the study
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of treatment and could bias reporting
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self report measures
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Low risk	**Comment:** all outcome data reported
Other bias	Low risk	**Comment:** well reported

Gulewitsch 2013

Methods	RCT with wait-list control **Follow-up:** postintervention (2 weeks) follow-up
Participants	**Location:** Germany **Setting:** recruited from community and outpatient clinics **Sample size:** 38 children (20 intervention, 18 control) **Sex:** 14 boys, 24 girls **Dropouts/withdrawals:** 0 **Diagnosis:** FAP and IBS according to Rome III (Rasquin 2006) **Mean age:** intervention: 9.1 (SD 1.7; range not reported) years; control: 9.7 (SD 1.8; range not reported) years. Data not reported for groups overall
Interventions	**Intervention:** brief hypnotherapeutic behavioural intervention (group format), 90 minutes/week x 4 (2 child focused, 2 parent focused) with homework for children to practice hypnotherapeutic trances at home, 5 times a week **Control:** wait-list control (no further details reported)
Outcomes	1. Number of days with pain (pain diary) 2. Pain index (Abdominal Pain Index; Walker 1997) 3. Pain intensity and duration (1 to 10 scale) 4. Health-related quality of life (KINDL-R in German; Ravens-Sieberer 1998) 5. Treatment success (> 80% reduction in Abdominal Pain Index; Walker 1997) 6. Pain disability (P-PDI; Hübner 2009)

Gulewitsch 2013 (*Continued*)

Notes	**Study dates:** not reported **Funding:** none stated **Declarations of interest:** authors report no conflicts of interest

Risk of bias		***Risk of bias***
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** computerised random number generator
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of treatment and issues re: wait-list expectancy
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self report
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** all participants accounted for
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Low risk	**Comment:** well reported

Korterink 2016

Methods	RCT with usual care control **Follow-up:** postintervention, 6 months, and 12 months
Participants	**Location:** the Netherlands **Setting:** recruited from outpatient clinics **Sample size:** 69 children (35 intervention, 34 control) **Sex:** 15 boys, 54 girls **Dropouts/withdrawals:** 4 intervention, 16 controls **Diagnosis:** abdominal pain - FGID Rome III (Rasquin 2006) **Mean age:** intervention: 12.2 (SD 2.9; range not reported) years; control: 12.1 (SD 2.7; range not reported) years. Overall range for both groups: 8 to 18 years
Interventions	**Intervention:** weekly Hatha yoga sessions of 1.5 hr (series of poses and breathing techniques) for 10 weeks (group format), daily practice at home encouraged **Control:** usual medical care

Korterink 2016 (*Continued*)

Outcomes	1. Pain intensity - diary (0 to 5 FACES Pain Scale; Hicks 2001) 2. Pain frequency - diary (0 to 4 scale) 3. Functional disability (Functional Disabilty Inventory; Walker 1991) 4. Depression (Children's Depression Inventory - short form; Kovacs 1992) 5. Anxiety (Revised Children's Manifest Anxiety Scale; Reynolds 1985)
Notes	**Study dates:** February 2012 to August 2013 **Funding:** partially funded by an unrestricted grant from VGZ Health Care Insurance, the Netherlands **Declarations of interest:** no conflicts of interest

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** computerised random number generator
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of treatment
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self report
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** there was a high proportion of dropouts from the control group, however the authors used several methods to attempt to account for this
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Low risk	**Comment:** well reported

Kuttner 2006

Methods	RCT with wait-list control **Follow-up:** postintervention (4 weeks) follow-up
Participants	**Location:** Canada **Setting:** recruited from gastroenterology clinics and the community **Sample size:** 28 children (14 intervention, 14 control) **Sex:** 8 boys, 20 girls **Dropouts/withdrawals:** 3 controls **Diagnosis:** IBS according to Rome I (Thompson 1989)

Kuttner 2006 (*Continued*)

	Mean age: intervention: 14.4 (SD 2.1; range not reported) years; control: 13.8 (SD 1. 9; range not reported) years. Overall range for both groups: 11 to 18 years
Interventions	**Intervention:** 1 hour of yoga instruction (series of poses and breathing techniques), followed by 4 weeks of daily practice, with video (individual format) **Control:** wait-list control (no other details reported)
Outcomes	1. Pain intensity (0 to 10 scale) 2. Functional disability (Functional Disabilty Inventory; Walker 1991) 3. Depression (Children's Depression Inventory - short form; Kovacs 1992) 4. Anxiety (Revised Children's Manifest Anxiety Scale; Reynolds 1985)
Notes	**Study dates:** not reported **Funding:** personal grants from British Columbia Research Institute, Canadian Institutes of Health Research, and the Michael Smith Foundation for Health Research **Declarations of interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** random sequence used
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants aware of allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** aware of treatment and issues regarding wait-list expectancy
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	High risk	**Comment:** data not reported on main outcome, pain intensity
Other bias	Low risk	**Comment:** well-reported paper

Levy 2010

Methods	RCT with active control **Follow-up:** postintervention (1 month), 3 months and 6 months follow-ups
Participants	**Location:** USA **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 200 children (100 intervention, 100 control) **Sex:** 12 boys, 188 girls **Dropouts/withdrawals:** 16 intervention, 16 control (postintervention); further 6 intervention, 8 control (at 6 months) **Diagnosis:** FAP according to Rome III (Rasquin 2006) **Mean age:** intervention: 11.1 (SD 2.6; range not reported) years; control: 11.3 (SD 2.5; range not reported) years. Overall range for both groups: 7 to 17 years
Interventions	**Intervention:** cognitive behavioural intervention for parent and child targeting response to pain, 3 x 75 minutes/week (non-group format) **Control:** active control of 3 x 75 minutes education sessions (education on gastrointestinal anatomy and function, nutrition guidelines, and reading food labels)
Outcomes	1. Pain intensity (FACES Pain Scale - Revised; Hicks 2001): child and parent 2. General disability due to pain (Functional Disability Inventory; Walker 1991): child and parent 3. Depression (Child Depression Inventory; Kovacs 1992) 4. Anxiety (Multidimensional Anxiety Scale for Children; Reynolds 1985) 5. Pain coping (Pain Response Inventory; Walker 1997)
Notes	**Study dates:** recruited 2005 to 2009 **Funding:** grant 5R01HD036069 from the National Institutes of Health - National Institute of Child Health and Human Development **Declarations of interest:** One author was a member of the Board of Directors at the Rome Foundation at the time of the study (listed as a potential competing interest)

Risk of bias ***Risk of bias***

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	**Comment:** computerised random number generator
Allocation concealment (selection bias)	Low risk	**Comment:** separate researcher performed randomisation
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** groups aware of treatment allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** although groups similar in intervention format, aware of therapy and self reporting outcome

Levy 2010 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Unclear risk	**Comment:** no data on child-reported pain
Other bias	Unclear risk	**Comment:** parent reports of pain not similar at baseline, but authors report that analyses have taken this into account

Palermo 2009

Methods	RCT with wait-list control **Follow-up:** postintervention (8 to 10 weeks) and 3 months' follow-up
Participants	**Location:** USA **Setting:** recruited from paediatric pain clinics **Sample size:** 24 children (14 intervention, 10 control); children with headache not included in analysis **Sex:** 7 boys, 17 girls **Dropouts/withdrawals:** none reported **Diagnosis:** chronic idiopathic abdominal pain **Mean age:** whole group 15.0 (SD 2.2; range 11 to 17) years. Data not reported by intervention group
Interventions	**Intervention:** cognitive behavioural family intervention (non-group format), delivered over the Internet, 30 minutes/week x 8 for both child and parent (4 hours total for each) **Control:** wait-list control; usual care (visits to physicians as needed)
Outcomes	1. Pain intensity (0 to 10 scale) 2. Activity limitations (Child Activity Limitations Interview; Palermo 2004)
Notes	**Study dates:** not reported **Funding:** grant HD050674 from the National Institutes of Health - National Institute of Child Health and Human Development **Declarations of interest:** Authors report no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** online random number generator
Allocation concealment (selection bias)	Low risk	**Comment:** sealed envelopes, postbaseline assessment

Palermo 2009 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of group allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** aware of treatment and issues regarding wait-list expectancy
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	**Comment:** some data on satisfaction missing
Selective reporting (reporting bias)	Unclear risk	**Comment:** some outcomes missing
Other bias	Unclear risk	**Comment:** queried validity of data collection tool

Palermo 2016

Methods	RCT with education control **Follow-up:** postintervention (8 to 10 weeks) and 6 months' follow-up
Participants	**Location:** USA and Canada **Setting:** recruited from paediatric pain clinics **Sample size:** 31 children (17 intervention, 14 control); children with headache not included in analysis **Sex:** 11 boys, 20 girls **Dropouts/withdrawals:** none reported **Diagnosis:** chronic idiopathic abdominal pain **Mean age:** intervention: 13.7 (SD 1.3; range 11 to 17) years; control 14.5 (SD 2.0; range 11 to 17) years. Data not reported for groups overall
Interventions	**Intervention:** Internet-delivered cognitive behavioural child and parent intervention (non-group format), 30 minutes/week x 8 for both child and parent (4 hours total for each) **Control:** Internet-delivered education child and parent intervention (non-group format) , 30 minutes/week x 8 for both child and parent (4 hours total for each)
Outcomes	1. Pain intensity (0 to 10 scale) 2. Activity limitations (Child Activity Limitations Interview; Palermo 2004)
Notes	**Study dates:** September 2011 to April 2014 **Funding:** grant HD062538 from the National Institutes of Health - Eunice Kennedy Schriver National Institute of Child Health and Human Development **Declarations of interest:** Authors report no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**

Palermo 2016 (*Continued*)

Random sequence generation (selection bias)	Low risk	**Comment:** computerised random number generator
Allocation concealment (selection bias)	Low risk	**Comment:** computerised random number generator linked automatically to web program
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	**Comment:** unsure what information participants were given with consent; it is therefore difficult to know whether they were truly 'unaware' of allocation as is suggested by author
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	**Comment:** unaware of intervention but self reporting outcome
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Low risk	**Comment:** all variables presented
Other bias	Low risk	**Comment:** baseline variables similar, well-reported paper

Robins 2005

Methods	RCT with usual care control **Follow-up:** postintervention (3 months), 6 months' and 12 months' follow-up
Participants	**Location:** USA **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 86 children (46 intervention, 40 control) **Sex:** 30 boys, 39 girls (data reported for completers only) **Dropouts/withdrawals:** 11 control **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** intervention: 10.8 (SD 2.5; range not reported) years; control: 11.9 (SD 2.3; range not reported) years. Overall range for both groups: 6 to 16 years
Interventions	**Intervention:** cognitive behavioural therapy, for child and parent (non-group format), 40 minutes x 5, every 2 weeks, as well as usual medical care **Control:** usual medical care (usual individualised care including visits with physicians and advice on diet)
Outcomes	1. Pain (Abdominal Pain Index; Walker 1997): child and parent 2. Functional disability due to pain (Functional Disability Inventory; Walker 1991): child

Robins 2005 (*Continued*)

Notes	**Study dates:** August 1998 to April 2000 **Funding:** grant from the Nemours Research Programs **Declarations of Interest:** not reported

Risk of bias ***Risk of bias***

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	**Comment:** coin flip status, with a witness, but procedure unclear
Allocation concealment (selection bias)	Unclear risk	**Comment:** no details provided of how coin flip was managed, or whether done in advance
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** only clinicians delivering usual medical care were blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self report outcomes and participants aware of intervention
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	**Comment:** clear numbers at follow-up not reported
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Unclear risk	**Comment:** issues relate to differences at baseline, sufficient numbers not recruited, larger loss to follow-up than expected, and merged results

Sanders 1990

Methods	RCT with wait-list control **Follow-up:** postintervention (2 months) and 3 months' follow-up
Participants	**Location:** Australia **Setting:** recruited from GP referrals and community advertisements **Sample size:** 16 children (8 intervention, 8 control) **Sex:** 4 boys, 12 girls **Dropouts/withdrawals:** none reported **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** intervention: 9.1 years (SD not reported; range not reported); control: 9.9 years (SD not reported; range not reported). Overall range for both groups: 6 to 12 years

Sanders 1990 (*Continued*)

Interventions	**Intervention:** cognitive behavioural therapy for parent and child (non-group format), 60 minutes/week x 8 **Control:** wait-list control (no further details reported)
Outcomes	1. Pain intensity (10-centimetre visual analogue scale): child and parent 2. Parent/child interaction (Family Observational Schedule; Sanders 1981) 3. Behaviour (Revised Behavior Problem Checklist; Quay 1983) 4. Adjustment (Children's Depression Inventory; Kovacs 1992) 5. Teacher rating (Conners' Teacher Rating Scale; Conners 1969) 6. Treatment success (pain free)
Notes	**Study dates:** not reported **Funding:** not reported **Declarations of Interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	**Comment:** method not reported
Allocation concealment (selection bias)	Unclear risk	**Comment:** method not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants and therapists not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** authors state no missing data
Selective reporting (reporting bias)	Unclear risk	**Comment:** no primary data reported (only shown in figures)
Other bias	High risk	**Comment:** no power calculations reported, small sample size, unclear whether accounted for baseline difference

Sanders 1994

Methods	RCT with usual care control **Follow-up:** postintervention (8 weeks), 6 months' and 12 months' follow-up
Participants	**Location:** Australia **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 44 children (22 intervention, 22 control) **Sex:** 16 boys, 28 girls **Dropouts/withdrawals:** none reported **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** intervention: 9.0 (SD 1.6; range not reported) years; control: 9.9 (SD 2.4; range not reported) years. Overall range for both groups: 7 to 14 years
Interventions	**Intervention:** cognitive behavioural therapy for parent and child (non-group format), 50 minutes/week x 6 **Control:** usual medical care (typically 4 to 6 visits with the gastroenterologist providing caring, supportive advice)
Outcomes	1. Pain intensity (10-centimetre visual analogue scale): child 2. Pain intensity (parent observation record): parent 3. Treatment success (reporting being pain free) 4. Child adjustment (Child Behavior Checklist; Achenbach 1983) 5. Pain interference with activities (scale 0 to 7): parent and child
Notes	**Study dates:** not reported **Funding:** funded by the National Health and Medical Research Council of Australia (grant 53091) **Declarations of interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	**Comment:** not reported
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants aware of allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** self report
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	**Comment:** missing numbers for primary outcomes
Selective reporting (reporting bias)	Low risk	**Comment:** all listed outcomes reported

Sanders 1994 (*Continued*)

Other bias	Unclear risk	**Comment:** study did not report calculating sample size sufficiency, and limited detail on method

van der Veek 2013

Methods	RCT with active control **Follow-up:** postintervention (8 weeks), 6 months' and 12 months' follow-up
Participants	**Location:** the Netherlands **Setting:** recruited from paediatric gastroenterology outpatient clinics **Sample size:** 104 children (52 intervention, 52 control) **Sex:** 29 boys, 75 girls **Dropouts/withdrawals:** 6 intervention, 4 control **Diagnosis:** FAP according to Rome III (Rasquin 2006) **Mean age:** intervention: 11.9 (SD 2.6; range not reported) years; control: 11.9 (SD 2.9; range not reported) years. Overall range for both groups: 8 to 17 years
Interventions	**Intervention:** cognitive behavioural therapy for parent and child (non-group format), 45 minutes/week x 6 **Control:** active control consisting of medical and dietary advice, 20 to 30 minutes/week x 6
Outcomes	1. Pain index (Abdominal Pain Index; Walker 1997) 2. Functional disability (Functional Disability Inventory; Walker 1991) 3. Anxiety and depression (Child Anxiety and Depression Scale - short version; Muris 2002) 4. Quality of life (KIDSCREEN-27; Ravens-Sieberer 2005) 5. Treatment success (pain intensity and pain duration reductions > 80%)
Notes	**Study dates:** not reported **Funding:** Dutch Digestive Foundation grant SW0 0509 **Declarations of interest:** Authors report no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** computerised randomisation
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** aware of group allocation

van der Veek 2013 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** no blinding and could influence self report
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Low risk	**Comment:** well reported

van Tilburg 2009

Methods	RCT with wait-list control **Follow-up:** postintervention (8 weeks) and 12 months' follow-up
Participants	**Location:** USA **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 34 children (19 intervention, 15 control) **Sex:** 9 boys, 23 girls (data for completers only) **Dropouts/withdrawals:** 1 intervention, 1 control **Diagnosis:** FAP using Rome II criteria (Rasquin-Weber 1999) **Mean age:** intervention: 10.6 (SD 3.0; range not reported) years; control: 9.9 (SD 2.2; range not reported) years. Overall range for both groups: 6 to 15 years
Interventions	**Intervention:** self directed "guided imagery" at home, child only (individual), 3 biweekly sessions (25 to 30 minutes) plus booster session, plus 3 x 10 minutes daily homework, x 8 weeks **Control:** wait-list control (usual medical care prescribed by physicians)
Outcomes	1. Pain index (Abdominal Pain Index, scale 0 to 40; Walker 1997) 2. Gastrointestinal symptoms (change score scale 0 to 7) 3. Functional disability (Functional Disability Inventory; Walker 1991) 4. School attendance 5. Quality of life (PedsQL; Varni 2001) 6. Treatment success (50% reduction on the Abdominal Pain Index; Walker 1997)
Notes	**Study dates:** March 2006 to March 2007 **Funding:** National Institutes of Health grants R24 DK067674 and RR00046 **Declarations of interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	**Comment:** no information on randomisation. Children picked an envelope, not sure

van Tilburg 2009 (*Continued*)

		whether these were shuffled
Allocation concealment (selection bias)	Unclear risk	**Comment:** not enough information provided
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** aware of treatment and issues regarding wait-list expectancy
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Unclear risk	**Comment:** secondary outcome data missing, uncertain when outcomes were measured
Other bias	Unclear risk	**Comment:** nothing reported on sample size adequacy

Vlieger 2007

Methods	RCT with usual care control **Follow-up:** postintervention (3 months), 6 months' and 12 months' follow-up
Participants	**Location:** the Netherlands **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 53 children (28 intervention, 25 control) **Sex:** 13 boys, 39 girls (data for completers only) **Dropouts/withdrawals:** 1 intervention, 0 control **Diagnosis:** FAP or IBS according to Rome II criteria (Rasquin-Weber 1999) **Mean age:** intervention: 13.2 (SD 2.5; range not reported) years; control: 13.4 (SD 2.9; range not reported) years. Data not reported for groups overall
Interventions	**Intervention:** hypnotherapy (gut directed) for child only (non-group format), 30 minutes x 6, over 3 months, with daily homework **Control:** usual care (education, dietary and pain medication advice, including 6 x 30 minute sessions of supportive therapy relating to possible triggers)
Outcomes	1. Pain intensity (scale 0 to 21) 2. Pain frequency (scale 0 to 21) 3. Treatment success (pain intensity and pain frequency reduction > 80%)
Notes	**Study dates:** October 2002 to June 2005 **Funding:** no external funding source **Declarations of interest:** not reported

Vlieger 2007 (*Continued*)

Risk of bias		*Risk of bias*
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** computerised random number generator
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants aware of treatment allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** outcomes analysed by someone blinded to treatment allocation, but outcome self reported by unblinded participants
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Low risk	**Comment:** all stated outcomes reported
Other bias	Unclear risk	**Comment:** no sample size calculations reported, and no information on compliance or validity of data tool

Wallander 2011

Methods	RCT with usual care control Follow-up: no immediate postintervention follow-up, 3 months' and 6 months' follow-up
Participants	**Location:** USA **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 63 children (36 intervention, 27 control) **Sex:** 19 boys, 44 girls **Dropouts/withdrawals:** 4 intervention, 3 control **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** whole group: 13.6 (SD 1.9; range 11 to 18) years. Data not reported by intervention group
Interventions	**Intervention:** written disclosure for child only (individual), 30 minutes x 3 sessions (1 at clinic, 2 at home), over 5 days **Control:** usual medical care (individualised as usual, involving dietary advice, education, and support)

Wallander 2011 (*Continued*)

Outcomes	1. Pain frequency (Abdominal Pain Frequency Rating; Walker 1993) 2. Quality of life - physical (PedsQL; Varni 2001) 3. Quality of life - psychosocial (PedsQL; Varni 2001)
Notes	**Study dates:** not reported **Funding:** funded in part by National Institute of Diabetes and Digestive and Kidney Diseases/National Institutes of Health grant RO3 DK61481-01A1 **Declarations of interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** computer generated
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** participants not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** participants and personnel not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounted for all participants
Selective reporting (reporting bias)	Low risk	**Comment:** stated outcomes reported
Other bias	Low risk	**Comment:** well reported

Wassom 2009

Methods	RCT with wait-list control **Follow-up:** postintervention (1 month) and 3 months' follow-up
Participants	**Location:** USA **Setting:** recruited from paediatric gastroenterology clinics **Sample size:** 20 children (9 intervention, 11 control) **Sex:** 4 boys, 11 girls (data for completers only) **Dropouts/withdrawals:** 2 intervention, 3 control **Diagnosis:** RAP according to Rome III (Rasquin 2006) **Mean age:** intervention: 11.9 (SD 2.6; range not reported) years; control: 11.9 (SD 2.9; range not reported) years (data for completers only). Overall range for both groups: 6 to 15 years

Wassom 2009 (*Continued*)

Interventions	**Intervention:** cognitive behavioural program Gutstrong, directed at child, but parental involvement (non-group format). At-home program via CD-ROM, over 4 weeks **Control:** wait-list control
Outcomes	1. Pain intensity (scale 1 to 10; Connelly 2006) 2. Pain-free days 3. Quality of life (PedsQL; Varni 2001) 4. Mood (Facial Affective Scale; McGrath 1991) 5. Stress (Stress Inventory; Schanberg 2000)
Notes	**Study dates:** not reported **Funding:** grant through the Children's Miracle Network (Kansas University Medical Center) **Declarations of interest:** not reported

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** uniform random numbers table
Allocation concealment (selection bias)	Unclear risk	**Comment:** not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	**Comment:** no blinding of participants or personnel reported
Blinding of outcome assessment (detection bias) All outcomes	High risk	**Comment:** no blinding of participants or personnel reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** not all participants had finished before results were reported
Selective reporting (reporting bias)	Low risk	**Comment:** as point above
Other bias	Unclear risk	**Comment:** no sample size calculations made, and some discrepancy over whether the results of the reported analyses were planned analyses

Weydert 2006

Methods	RCT with active control **Follow-up:** postintervention (4 weeks) and 2 months' follow-up
Participants	**Location:** USA **Setting:** recruited from GP referrals and paediatric gastroenterology clinics **Sample size:** 27 children (16 intervention, 11 control) **Sex:** 7 boys, 15 girls **Dropouts/withdrawals:** 2 intervention, 3 control (all prior to starting allocation) **Diagnosis:** RAP diagnosed using Apley criteria (Apley 1958) **Mean age:** intervention: 11.1 years (SD not reported; range not reported); control: 11.0 (SD not reported; range not reported) years. Data not reported for groups overall
Interventions	**Intervention:** guided imagery and progressive muscle relaxation, for child only (individual), 60 minutes/week x 4, with daily homework **Control:** breathing technique training, 60 minutes/week x 4, to control for therapist time and attention, with daily homework
Outcomes	1. Number of days with pain (diary) 2. Pain intensity (FACES Pain Scale; Bieri 1990) 3. Missed days of school (diary) 4. Depression (Child Depression Inventory; Kovacs 1992) 5. Anxiety (Multidimensional Anxiety Scale for Children; Reynolds 1985) 6. Treatment success (4 or fewer days of pain with no missed activities during each month)
Notes	**Study dates:** 2000 to June 2002 **Funding:** National Center for Complementary and Alternative Medicine grant **Declarations of interest:** Authors reported no conflicts of interest.

Risk of bias ***Risk of bias***

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Comment:** blocks of 4 (tokens in a hat)
Allocation concealment (selection bias)	Low risk	**Comment:** separate group responsible for randomisation
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	**Comment:** personnel blinded to group (both called "relaxation techniques"), but unsure of the degree of participant blinding (depends on how treatments were explained)
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	**Comment:** researcher recording outcomes was blind to treatment allocation. Although participants self reported outcome, attempts were made to blind participants

Weydert 2006 (*Continued*)

		to their treatment allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	**Comment:** accounts for participants
Selective reporting (reporting bias)	Low risk	**Comment:** all outcomes reported
Other bias	Unclear risk	**Comment:** sample size calculations not reported, but otherwise well reported

FAP: functional abdominal pain
FGID: functional gastrointestinal disorders
GP: general practitioner
IBS: irritable bowel syndrome
KINDL-R: KINDer Lebensqualitätsfragebogen
P-PDI: Paediatric Pain Disability Index
PedsQL: Pediatric Quality of Life Inventory
RAP: recurrent abdominal pain
RCT: randomised controlled trial
SD: standard deviation

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Alfvén 2007	Ineligible comparator: compared psychological treatment with physiotherapy
Bursch 2008	No primary data: literature review
Hicks 2003	Ineligible population: not RAP-specific pain
Humphreys 1998	No control group
Long 2009	Ineligible outcome: outcome was physical activity, no measure of pain
Sato 2009	No primary data: literature review
Sieberg 2010	Ineligible comparator: compared CBT with CBT plus family therapy (no control)
Sieberg 2011	Ineligible comparator: compared CBT with CBT plus family therapy (no control)
van Barreveld 2015	Ineligible comparator: no control group
Warner 2011	Ineligible population: children with RAP and psychological disorders

CBT: cognitive behavioural therapy
RAP: recurrent abdominal pain

Characteristics of studies awaiting assessment *[ordered by study ID]*

Tannen 2014

Methods	Unknown
Participants	Children with functional abdominal pain; further details unknown
Interventions	Cognitive training; further details unknown
Outcomes	Unknown
Notes	Unknown

Youssef 2009

Methods	RCT with active control (pilot study) **Follow-up:** postintervention (1 week) and 3 months' follow-up
Participants	**Location:** not reported **Setting:** schools **Sample size:** not reported **Sex:** not reported **Dropouts/withdrawals:** not stated **Diagnosis:** FAP, defined as 3 episodes of pain interfering with activity for 3 months in the past year **Mean age:** not reported
Interventions	**Intervention:** guided imagery, 6 sessions in a week (first session 15 minutes, then 5 x 7 minutes) **Control:** rest and relaxation, 6 sessions a week
Outcomes	1. Pain (measure not reported) 2. Disability (measure not reported)
Notes	**Study dates:** not reported **Funding:** not reported **Declarations of interest:** not reported

FAP: functional abdominal pain
RCT: randomised controlled trial

DATA AND ANALYSES

Comparison 1. Cognitive behavioural therapy (CBT) versus control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Treatment success: postintervention	4	175	Odds Ratio (M-H, Random, 95% CI)	5.67 [1.18, 27.32]
1.1 Active control or usual care	2	130	Odds Ratio (M-H, Random, 95% CI)	2.25 [0.57, 8.88]
1.2 Wait-list control	2	45	Odds Ratio (M-H, Random, 95% CI)	24.55 [2.24, 269.03]
2 Treatment success: medium-term follow-up (3 to 12 months)	3	139	Odds Ratio (M-H, Random, 95% CI)	3.08 [0.93, 10.16]
3 Treatment success: long-term follow-up (12 months or more)	2	120	Odds Ratio (M-H, Random, 95% CI)	1.29 [0.50, 3.33]
4 Pain intensity: postintervention	7	405	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.74, 0.08]
4.1 Active control or usual care	4	337	Std. Mean Difference (IV, Random, 95% CI)	-0.04 [-0.39, 0.31]
4.2 Wait-list control	3	68	Std. Mean Difference (IV, Random, 95% CI)	-0.92 [-1.59, -0.24]
5 Pain intensity: postintervention sensitivity analysis for allocation concealment (low risk of bias)	4	247	Std. Mean Difference (IV, Random, 95% CI)	-0.28 [1.00, 0.45]
6 Pain intensity: medium-term follow-up (3 to 12 months)	4	301	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.85, 0.20]
7 Pain intensity: long-term follow-up (12 months or more)	3	308	Std. Mean Difference (IV, Random, 95% CI)	-0.04 [-0.39, 0.31]
8 Quality of life (physical subscale) : postintervention	3	136	Std. Mean Difference (IV, Random, 95% CI)	0.71 [-0.25, 1.66]
9 Quality of life (psychosocial subscale): postintervention	3	136	Std. Mean Difference (IV, Random, 95% CI)	0.43 [-0.21, 1.06]
10 Functional disability or activity limitations: postintervention	4	176	Std. Mean Difference (IV, Random, 95% CI)	-0.57 [-1.34, 0.19]
10.1 Active control or usual care	2	123	Std. Mean Difference (IV, Random, 95% CI)	-0.01 [-0.36, 0.34]
10.2 Wait-list control	2	53	Std. Mean Difference (IV, Random, 95% CI)	-1.31 [-2.10, -0.52]

Comparison 2. Hypnotherapy (including guided imagery) versus control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Treatment success: postintervention	4	146	Odds Ratio (M-H, Random, 95% CI)	6.78 [2.41, 19.07]
1.1 Active control or usual care	2	74	Odds Ratio (M-H, Random, 95% CI)	10.51 [2.88, 38.33]
1.2 Wait-list control	2	72	Odds Ratio (M-H, Random, 95% CI)	5.77 [0.64, 52.05]
2 Pain intensity: postintervention	4	146	Std. Mean Difference (IV, Random, 95% CI)	-1.01 [-1.41, -0.61]
2.1 Active control or usual care	2	74	Std. Mean Difference (IV, Random, 95% CI)	-1.00 [-1.90, -0.11]
2.2 Wait-list control	2	72	Std. Mean Difference (IV, Random, 95% CI)	-0.95 [-1.44, -0.46]
3 Pain frequency: postintervention	4	146	Std. Mean Difference (IV, Random, 95% CI)	-1.28 [-1.84, -0.72]
3.1 Active control or usual care	2	74	Std. Mean Difference (IV, Random, 95% CI)	-1.74 [-2.29, -1.19]
3.2 Wait-list control	2	72	Std. Mean Difference (IV, Random, 95% CI)	-0.87 [-1.38, -0.36]

Comparison 3. Yoga versus control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity: postintervention	3	122	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.67, 0.05]
2 Functional impairment: postintervention	2	53	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-1.07, 0.43]

ADDITIONAL TABLES

Table 1. Assessment of risk of bias in included studies

Domain	'Risk of bias' judgement		
Selection bias	**Low**	**High**	**Unclear**
Random sequence generation	If the study details any of the following methods. 1. Simple randomisation (such as coin-tossing, throwing dice or dealing previously shuffled cards, a list of random numbers, or computer generated random numbers). 2. Restricted randomisation: blocked, ideally with varying	If the study details randomisation by an inadequate method such as alternation, assignment based on date of birth, case record number, and date of presentation. These may be referred to as 'quasi-random'	If there is insufficient detail to judge the risk of bias as low or high

Table 1. Assessment of risk of bias in included studies (*Continued*)

	block sizes or stratified groups, provided that within groups randomisation is not affected.		
Allocation concealment	If the study details concealed allocation sequence in sufficient detail to determine that allocations could not have been foreseen in advance of or during enrolment	If the study details a method where the allocation may have been known prior to assignment	If there is insufficient detail to judge the risk of bias as low or high
Performance bias	**Low**	**High**	**Unclear**
Blinding of participants and personnel	If the study details a method of blinding the participants and personnel. This requires sufficient detail to show that neither participants nor personnel were able to distinguish the therapeutic intervention from the control intervention	Considering the nature of the interventions, we do not expect it to be possible for participants and therapists to be blinded. The effect of this is addressed in the discussion	If there is insufficient detail to judge the risk of bias as low or high
Detection bias	**Low**	**High**	**Unclear**
Blinding of outcome assessment	If the study details a blinded outcome assessment. This may only be possible for outcomes that are externally assessed	If the outcome assessment is not blinded. We expect this may be unavoidable for self rated outcomes of unblinded interventions	If there is insufficient detail to judge the risk of bias as low or high
Attrition bias	**Low**	**High**	**Unclear**
Incomplete outcome data	If the study reports attrition and exclusions, including the numbers in each intervention group (compared with total randomised participants), reasons for attrition or exclusions and any re-inclusions, and if the impact of missing data is not believed to alter the conclusions, and there are acceptable reasons for the missing data	We may judge the risk of attrition bias to be high due to the amount, nature or handling (such as per-protocol analysis) of incomplete outcome data	If there is insufficient detail to judge the risk of bias as low or high (e.g. the number of children randomised to each treatment is not reported)
Reporting bias	**Low**	**High**	**Unclear**
Selective outcome reporting	If we judge there to be complete reporting, as found on comparison of the protocol and pub-	If the reporting is selective, so that some outcome data are not reported	If there is insufficient detail to judge the risk of bias as low or high (e.g. protocols are unavail-

Table 1. Assessment of risk of bias in included studies (*Continued*)

	lished study, if available		able)
Other sources of bias	**Low**	**High**	**Unclear**
Four additional possible sources of bias: 1. Were the data collection tools valid? 2. Was there sufficient power in terms of appropriate sample size? 3. Were groups equal at baseline (primary outcome)? 4. Were data analyses appropriate?	Three or more of these judged to be at low risk of bias.	One or more of these judged to be at high risk of bias.	If there is insufficient detail to judge the risk of bias as low or high

WHAT'S NEW

Last assessed as up-to-date: 9 June 2016.

Date	Event	Description
16 August 2016	New search has been performed	Following an updated search in June 2016, we added 2 new studies
4 February 2016	New citation required and conclusions have changed	We found an additional 12 studies.
4 February 2016	New search has been performed	This review supersedes the previous review (see published notes), following a new protocol, and new search in March 2013 and updated searches in April 2014 and March 2015

HISTORY

Protocol first published: Issue 2, 2014

Review first published: Issue 1, 2017

Date	Event	Description
6 November 2015	Amended	First full draft of review

CONTRIBUTIONS OF AUTHORS

Review design: RAA, AEM, SL.

Review co-ordination: RAA, AEM.

Data collection:

1. Search strategy design: AEM, AB.
2. Searches: AEM, AB.
3. Search results screening: RAA, AEM, TNVD, AB, JTC, RW.
4. Retrieval of papers: AEM, AB.
5. Paper screening and appraisal, and extraction of data: RAA, AEM, TVND, AB, JTC, RW.
6. Writing to authors for additional information: RAA, AEM, AB, RW.
7. Entering the data into Review Manager 5: RAA, AEM, TVND, RW.

Analysis of the data: RAA, AEM, TVND, SL.

Interpretation of the data:

1. Methodological perspective: RAA, AEM, TVND, AB, JTC, RW.
2. Clinical perspective: AEM, TVND, SL.

DECLARATIONS OF INTEREST

The work of the evidence synthesis team is funded by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care South West Peninsula (PenCLAHRC). However, the funder had no role in the review itself.

Rebecca A Abbott: none known.

Alice E Martin: none known.

Tamsin V Newlove-Delgado: none known.

Alison Bethel: none known.

Joanna Thompson-Coon: none known.

Rebecca Whear: none known

Stuart Logan: none known.

The authors who practice clinical paediatrics are Alice E Martin and Stuart Logan. Alice is a Paediatric Trainee and works under the guidance of various Consultant Paediatricians. Stuart is a Consultant Paediatrician and treats children according to current best evidence, in light of their preference. There are therefore no conflicts of interest with this review.

SOURCES OF SUPPORT

Internal sources

- None, Other.

External sources

- None, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Rebecca Whear (RW) was added to the review team after registration of the protocol. RW was involved in screening abstracts and full texts, data extraction, writing to authors, entering data into Review Manager and contributed to discussions pertaining to methods.

We presented Tau², an estimate of between-study variability, as requested by the Cochrane editorial team.

In this review we have referred to the "proportion of participants that improved with treatment" as "treatment success".

The table below provides details of analyses that were employed post hoc and not specified in the protocol (Martin 2014c). These analyses were deemed appropriate due to the nature of bias that wait-list control groups can incur, and due to the fact that many psychosocial interventions chose to use wait-list controls.

Post hoc method employed	Reason for use
Sensitivity analyses Where data allowed, we performed sensitivity analyses to assess the robustness of conclusions in relation to the possible bias introduced by the use of wait-list controls	Wait-list control studies have been criticised as being at increased risk of bias, as there may be an expectancy of benefit, which could overestimate the treatment effect

NOTES

This is a new review, which supersedes a previously published review (Huertas-Ceballos 2008).