

Title: Non-pharmacological conservative therapy for phantom limb pain: a systematic review of randomised controlled trials.

Author: Batsford S^{1,2}, Ryan CG³, Martin DJ³.

Affiliation

¹ School of Health and Social Care, Teesside University, Middlesbrough, UK.

² MSK CATS, South Tyneside NHS Foundation Trust, South Shields, Tyne and Wear, UK.

³ Health and Social Care Institute, Teesside University, Middlesbrough, UK.

Corresponding Author

Cormac Ryan

Reader in Physiotherapy

Institute of Health and Social Care

Teesside University

Middlesbrough

TS1 3BA

Tel: 07855424261

Email: c.ryan@tees.ac.uk

ABSTRACT

The aim of this paper was to investigate the effectiveness of conservative therapy for phantom limb pain (PLP). In this systematic review CINAHL, AMED, The Cochrane database of systematic reviews, PEDro, psychology and behavioural sciences collection, and MEDLINE were systematically searched for appropriate randomised controlled trials (RCTs). Selected papers were assessed for risk of bias and evidence was graded using the GRADE approach. Twelve RCTs met initial inclusion/exclusion criteria, of which five were of sufficient quality for final inclusion. There is conflicting evidence from two RCTs for the effectiveness of electromagnetic shielding limb liners on pain in the short-term. There is limited evidence supporting the effectiveness of both hypnosis in the short-term and Graded Motor Imagery (GMI) in the short-to-medium term. Additionally, there is limited evidence that a single session of mirror therapy has no immediate effect on PLP. Limb liner discomfort was the only adverse effect identified. This review identifies a range of conservative therapies, many of which demonstrate preliminary evidence of potential with respect to clinically worthwhile effects above control interventions and few, if any, adverse effects. However there is a paucity of high-quality evidence upon which to make any firm clinical conclusions.

Key terms: Phantom limb pain; amputee; conservative therapy; systematic review.

INTRODUCTION

Phantom limb pain (PLP) can be defined as when a person reports a pain sensation somewhere within a limb that has been amputated (Halbert, Crotty and Cameron, 2002). Approximately 60-80% of amputees develop PLP, with 5-10% of these individuals reporting severe pain (Nikolajsen, 2012). PLP is a long-term condition, which can significantly reduce quality-of-life and is associated with high healthcare and wider economic costs (Giummarra and Moseley, 2011; Moura et al, 2012; Nikolajsen, 2012).

PLP is a complex type of pain, which is poorly understood despite its existence in the literature since 1551 when French military doctor, Ambriose Pare, referred to it in his writings (Roulet, Noutte-Gualain, Brochet, Sztark, 2009). The exact underlying mechanisms of PLP remain largely unknown (Giummarra and Moseley, 2011; Nikolajsen, 2012). Three main mechanisms proposed in the literature are peripheral nerve injury, sensitisation of the central nervous system, and cortical reorganisation (Flor, 2002; Giummarra and Moseley, 2011; Nikolajsen, 2012). It is likely that all of these mechanisms contribute in part to PLP, but the proportional contribution has not yet been determined (Nikolajsen, 2012). The lack of clear understanding of the mechanisms of PLP is reflected in the lack of clear understanding of what works best to help with its management.

In practice treatments can be divided into pharmacological and non-pharmacological conservative therapy (Flor, 2002; Giummarra and Moseley, 2011): surgery is rare as it is costly and ineffective (Nikolajsen, 2012). While pharmacological treatments form the mainstay of practice (Giummarra and Moseley, 2011; Alviar, Hale and Dungca, 2011), a recent systematic review has questioned their clinical usefulness with limited

evidence of effectiveness accompanied by frequent reporting of moderate side-effects (Alvair, Hale, Dungca, 2011). Side-effects can include tiredness, dizziness, sweating, constipation, micturition difficulties, nausea, vertigo, itching, shortness of respiration, and sedation (Huse, Larbig, Flor, Birbaumer, 2001; Wu, 2002). Due to the lack of supportive research for the effectiveness of pharmacologic treatments, their high costs and relatively high occurrence of adverse effects, a greater focus on non-pharmacological conservative therapies is warranted.

An array of non-pharmacological conservative therapies for PLP has been reported in the literature including: electrotherapy, cognitive behavioral therapy, mirror therapy, sensory discrimination training, hypnotherapy, limb covers, bio-feedback, reflexology, guided imagery, acupuncture, virtual reality, and prosthesis use and training. However, there is a lack of clinical guidelines to guide practice. A comprehensive systematic review of the literature from 1966-1999 concluded that there was little evidence from appropriate randomised controlled trials to guide treatment choice, including the choice of non-pharmacological conservative therapies (Halbert, Crotty and Cameron, 2002). This was partially updated by a more recent review that limited its search to studies on “mind-body” techniques published between 1994 and 2010 (Moura et al. 2012). The review found studies on hypnosis, imagery, biofeedback and mirror therapy, but only two of these were RCTs and both were of low quality. As with the previous review (Halbert, Crotty and Cameron, 2002) there was little evidence from RCTs to guide treatment choice for this limited range of therapies: nor did the review provide an update of evidence for non-pharmacological therapies beyond the narrow range of “mind-body” techniques. Thus, we conducted a systematic review with the aim of

investigating the extent and nature of moderate-to-high quality evidence to guide the non-pharmacological conservative management of PLP.

METHODS

Methodology

The Cochrane handbook for systematic reviews of interventions (Higgins and Green, 2008) was used to guide the development and undertaking of this systematic review and the review is reported in line with PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, 2009). The PICOS acronym was used to structure the systematic search and develop inclusion/exclusion criteria. Studies containing adult males and females aged ≥ 18 years of age were included. The included studies had to directly measure PLP. Studies in which PLP was not clearly distinguished from residual limb pain or phantom sensations were excluded. Additionally, studies that included other conditions (for example studies investigating mirror therapy which may have included stroke participants) were excluded unless the data for amputees with PLP were presented separately. In those cases, only the data within that paper specific to adults with amputations who were experiencing PLP were included. The interventions included in the review were non-pharmacological conservative therapies and eligible comparison groups were usual care control, non-treatment control, placebo control and/or active control interventions. We accepted papers that included one or more of a broad range of outcomes to ensure the review was comprehensive and inclusive.

The primary outcome of interest was pain severity. A change of 30% or 2/10 (20/100mm) on the pain VAS/NRS was considered clinically relevant as recommended by the Cochrane Back Review Group (Furlan, Pennick, Bombardier, and van Tulder,

2009). Secondary outcome measures were not restricted in any way and could lie anywhere within the biopsychosocial spectrum including for example; quality of life, mood and depression, due to the complex bio-psychosocial effect that PLP has on those affected (Nikolajsen, 2012). These outcomes were split into physical function, psychological function and social function. Additionally, adverse effects were included as an outcome of interest as they are important for clinical decision-making (Loke, Price and Herxheimer, 2008) and a stated aim of our study.

This systematic review only included RCTs in order to reduce bias (Bradley and Nolan, 2008; O'Connor, Green, Higgins, and Deeks, 2008; Higgins et al 2011). Furthermore, RCTs rated as “low quality” (see below) were excluded (Higgins, Altman, and Sterne, 2008).

Search strategy

To ensure a thorough search, large databases and subject-specific databases were used (Dickersin, Scherer, Lefebvre, 1994; Bettany-Saltikov, 2010; Lefebvre, Manheimer, and Glanville, 2008) namely MEDLINE, EMBASE, CINAHL, AMED, PEDro, Psychology and Behavioral Sciences Collection. Hand searching the reference lists of key studies for additional trials was also carried out (Dickersin, Scherer, Lefebvre, 1994). The systematic search covered the period from the beginning of the database up to October the 4th 2014.

The search terms integrated into the systematic search were developed from the review question using the PICOS components (Bettany-Saltikov, 2010; Lefebvre, Manheimer, and Glanville, 2008). Words from the PICOS model were identified and then synonyms

for the components were identified from the population component (table 3.3). The specific search terms used were phantom limb pain (S1), phantom pain (S2), limb pain (S3), amput* (S4), pain (S5), and phantom limb (S7). Boolean operators of AND and OR were used to combine key words to ensure the search was both sensitive and specific (Bettany-Saltikov, 2010). The following limiters were applied: English language, human, all adult 19+ years, randomized controlled trial. The final search strategy was: S1 OR S2 OR S3 OR S4 OR S6 OR S7, where S6 was a combination of S4 AND S5.

Insert table 1 here

The relevant intervention, outcome and study components were then applied either as limiters or during the screening process of the studies identified. The inclusion criterion of RCT design was applied as a limitation after the keyword search. Other limiters applied were to include only English language articles, and adult human participants. Intervention and outcome components were not included in the keyword search. It was decided that including these could lead to missed relevant data. Intervention and outcome eligibility criteria were assessed while screening the identified studies by reading titles and abstracts during the selection of studies, after the keyword search and limitations were applied.

The Cochrane Collaboration's recommended process was used to select studies for inclusion (Higgins and Deeks 2008). Search results were merged and duplicates removed, titles and abstracts were reviewed and irrelevant studies removed. For the remaining potentially relevant articles the full text were screened and those fitting the

inclusion criteria were retained. These processes were carried out by one of the authors (SB). The remaining papers were then assessed for quality by two authors (SB and CR). Those considered to be of high risk of bias (low quality) were removed from the final review. If agreement between the two readers could not be met a third reviewer was used (DM) to reach a consensus.

Methodological quality assessment

Furlan, Pennick, Bombardier, and van Tulder's (2009) proforma for assessing the risk of bias was used to assess study quality. The level of agreement between SB and CR was assessed using the kappa co-efficient (Higgins, Altman, and Sterne 2008) interpreted as follows: 0.01-0.20 slight agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 substantial agreement, 0.81-0.99 almost perfect agreement (Viera and Garrett, 2005). Individual studies were then rated to be of high (>9/12), moderate (6-9/12) or low quality (<6/12). Those rated as low quality were excluded from the final review.

Data collection Process

Data extraction was carried out by SB using an adapted version of The Cochrane Collaboration's template for data collection. The data extraction form obtained data relating to study eligibility, methodological characteristics of included studies, participant characteristics, intervention group characteristics, outcome characteristics, risk of bias assessment, data analysis, and key study conclusions.

Data analysis and synthesis

A GRADE approach adapted from a previous Cochrane review (van Tulder, Furlan, Bombardier, and Bouter, 2003) was used to grade the quality of evidence (See table 2). A narrative synthesis approach was used throughout as no studies were sufficiently homogenous to justify useful meta-analysis (Bettany-Saltikov, 2010; Deeks, Higgins, and Altman, 2008).

Insert table 2 here

RESULTS

The electronic search yielded 1190 records and hand searching yielded an additional 13 records. The review of study titles and abstracts rejected 1166 studies because they did not meet the set eligibility criteria. Thirty-seven full texts were accessed. Of these, ten studies were excluded due to participants not having PLP; five studies were excluded as they were not RCTs; and ten studies were excluded as they were duplicates. The remaining 12 studies investigated the following interventions; TENS, mirror therapy, electromagnetically shielding limb liners, electromagnetic-own-signal treatment (EMOST), phantom exercises, transcranial magnetic stimulation, and sensory motor discrimination training. Of these 12 studies seven (Finsen et al 1988; Flor, Denke, Schaefer, and Grusser, 2001; Kern, Altkemper, and Kohl, 2006; Chan et al 2007; Ulger et al 2009; Ahmed, Mohamed, Sayed, 2011; Bokkon, Till, Grass, and Szabo, 2011) were judged to be of lower quality (see table 3) and were removed. Thus five studies were included in the final review (see table 4 and figure 1).

Insert figure 1 here

Insert tables 3 and 4 here

Characteristics of included studies

135 adult participants with PLP after amputation were included in the review. Moseley (2006) included 51 participants but only nine of these had PLP. Brodie, Whyte and Niven (2007) included 80 patients, however only 15 of these reported PLP. Both Moseley (2006) and Brodie, Whyte and Niven (2007) analysed the data of those with PLP separately from the rest of their samples. Therefore, these studies were included in this review but only the data relating to amputees with PLP was used.

Two studies examined limb liners that purport to shield against electromagnetic fields¹ (Hsiao, 2012; Conine, 1993), one study investigated Graded Motor Imagery (GMI) (Moseley, 2006), one study investigated the effects of hypnosis (Rickard, 2004) and one study examined a single session of mirror therapy (Brodie, Whyte and Niven, 2007).

All five studies assessed pain intensity using either a visual analogue scale (VAS), numerical rating scale (NRS) or the McGill Pain Questionnaire (MPQ). Secondary outcome measures were assessed in two of the studies both of which measured physical function (Moseley, 2006; Hsiao et al 2012) while one also measured psychological function (Hsiao et al. 2012).

The timing of outcome measures ranged from immediately (short-term) after a single intervention (Brodie, Whyte and Niven, 2007) to six months (moderate-term) after a six-week intervention (Moseley, 2006). There were no appropriate studies with long-term outcome measures (≥ 1 year) which were eligible for inclusion in this review.

It is not evident that adverse effects were formally planned to be assessed within the methods of any of the five included RCTs. One study on hypnosis reported “no unplanned problems or potentially negative experiences” (Rickard, 2004). Conine, Herschler, Alexander, and Crisp (1993) reported that six of the 52 participants involved in their trial withdrew due to discomfort from the limb liner they were given (this appeared to relate to both the placebo and active limb liner). Of the remaining three studies, none reported directly on the presence or absence of adverse events, though it was stated that there were no withdrawals in the mirror therapy study (Brodie, Whyte and Niven, 2007) or the high quality limb liner study (Hsiao et al. 2012). One participant in the control group withdrew in the GMI study (Moseley, 2006).

Data Analysis and Synthesis

Methodological quality summary

The methodological quality of the 12 trials which met the initial inclusion/exclusion criteria are shown in table 3. Level of agreement between the two assessors was Kappa = 0.73, classified as a substantial level of agreement (Viera and Garrett, 2005). Of the five trials that were of sufficient quality to be entered into the final review four were rated as moderate quality (Conine, Herschler, Alexander, and Crisp, 1993; Rickard, 2004; Moseley, 2006; Brodie, Whyte and Niven, 2007), and one as high quality (Hsiao et al 2012). The pain outcomes from these five studies are discussed in the follow paragraphs and the non-pain outcomes and adverse effects information from these studies are presented in table 5.

Insert table 5 here

Electromagnetically Shielding Limb Liner

Two studies, one high quality (Hsiao et al 2012) and one of moderate quality (Conine, Herschler, Alexander, and Crisp, 1993), assessed the effect of electromagnetically shielding limb liners.

Phantom limb pain intensity: The high quality trial (Hsiao et al 2012) reported no difference in pain between groups at 6 weeks (0-10 VAS; mean difference 0.8; 95%CI -1.4 to 3), and 12 weeks (0.2;-1.9 to 2.3) (Hsiao et al 2012). The moderate quality trial (Conine, Herschler, Alexander, and Crisp, 1993) showed a statistically significant reduction of pain with an electromagnetically shielding limb liner compared to a placebo (1.7; 1.0 to 2.4) though the time point at which these measures were taken is not clear but appears to be in the short-term following removal of the limb liner. Nor is it clear for how long the liner was worn.

In summary, there is conflicting evidence from two RCTs on the clinical effectiveness of electromagnetically shielding limb liners compared to a placebo limb liner for pain in the immediate to short-term. There is a moderate level of evidence from one high quality RCT of no effect for electromagnetically shielding limb liners compared to a placebo limb liner on physical and psychological function in the short-term.

Graded Motor Imagery (GMI)

One study investigated the effectiveness of GMI on PLP in comparison to usual care (Moseley, 2006). The methodological quality of the study was rated as moderate as an investigation of PLP. The data on PLP were analysed separately from data collected from people with other conditions. The participants with PLP constituted a very small

sample of nine, five receiving GMI and four in a usual care control group. Because of the extremely small sample size it is difficult to make any inferences about effectiveness, beyond noting that the data were in the direction of a comparatively greater relief of pain and improvement in function in the GMI group after treatment and at a six month follow up. However, for completeness we have quantified the results of Moseley (2006) below.

Phantom limb pain intensity: PLP was recorded on a 100mm VAS. The mean decrease in pain VAS between pre and post-treatment was greater in the GMI group compared to the control group (26mm vs. 12mm), with a mean group difference of 14mm. The mean change in pain VAS between pre-treatment and 6 month follow up was also greater in the GMI intervention group compared to the control group² (32mm vs. 5mm), with a mean group difference of 27mm.

Thus, there is limited evidence from one moderate quality that GMI is effective for pain and function compared to usual care control at six months but not immediately post treatment. It is important to note that this evidence comes from a very small sample size.

Mirror Therapy

One RCT rated as moderate quality compared the use of mirror therapy to a covered mirror control intervention in 15 people with PLP (Brodie, Whyte and Niven, 2007). As in the case above the data on PLP came from a subset of the overall sample in that study. Again, the sample size of these participants with PLP was very small with 7 receiving mirror therapy and 8 taking part as controls. Because of the extremely small

sample size we refrain from making any inference apart from simply noting that the data did not show any clear differences in pain intensity outcome between the groups. However, for completeness we have quantified the results of Brodie, Whyte and Niven (2007) below.

Phantom limb pain: the total pain-ranking index of the McGill Pain Questionnaire (MPQ) in the control group reduced by 14 points compared to 6 points in the mirror therapy group. When assessed using the 100mm pain VAS pain reduction in the control group was 4mm compared to 17mm in the mirror therapy group. These two measures are somewhat at odds with one another with the MPQ suggesting greater pain reduction for the control group and the VAS suggesting the reverse. However the magnitude of the difference between groups was small.

There is limited evidence from one moderate quality RCT that a single session of mirror therapy compared to a control was not effective for PLP immediately post intervention. It is important to note that this evidence comes from a very small sample size.

Hypnosis

One RCT rated as moderate quality compared the use of hypnosis to a waiting list control in 20 people with PLP (Rickard, 2004).

Phantom limb pain: In the short-term, the hypnosis group had a lower pain intensity than the waiting list control group as measured by the total pain rating index of the MPQ (10 [6] (mean [SD]) vs. 46 [15], $p < 0.001$), the present pain intensity score of the MPQ (1.3 [0.8] vs. 3.6 [1.0], $p < 0.001$) and the number of words chosen on the MPQ

(5.7 [3] vs. 19 [3, $p<0.001$]). Similarly, in the short-term, using the daily rating pain scales with a NRS from 0-100 the hypnosis group had a lower pain intensity than waiting list controls (12 [11] vs. 44 [14], $p<0.001$).

There is limited evidence from one moderate quality RCT that hypnosis, in comparison to a waiting list control, reduced PLP in the short-term.

DISCUSSION

This systematic review aimed to investigate the safety and effectiveness of non-pharmacological conservative therapies for pain and wellbeing in people with PLP. Twelve RCTs were initially identified five of which were included in the final review. These five RCTs involved 135 amputees treated for PLP and encompassed the following four forms of intervention: electromagnetically shielding limb liner, hypnosis, GMI and mirror therapy.

This is the first comprehensive review of non-pharmacological conservative therapies for this patient group in fifteen years. Thus, this study fills an important gap in the literature and provides much needed, up-to-date guidance for clinicians on the current state of the evidence with regards to both efficacy and safety. Though limited in quality, hypnosis and GMI are the interventions with the greatest level of scientific support. Both interventions have demonstrated clinically worthwhile improvements, hypnosis in the short term and GMI in the medium term.

Halbert, Crotty and Cameron's (2002) systematic review on PLP management included 12 studies, of which eight investigated pharmacological interventions and four

conservative interventions. Of the four conservative intervention studies in their review only one of those was included in our review (Conine, Herschler, Alexander, and Crisp, 1993). While one RCT was not included in our final review due to its low methodological quality (Finsen et al 1988), the two other papers were not included in our review as they were not RCTs (Katz and Melzack, 1991; Lundeburg, 1985).

Our review found conflicting evidence regarding the effect of electromagnetically shielding limb liner for PLP compared to a placebo in the short-term. In contrast Halbert, Crotty and Cameron (2002) reported that electromagnetically shielding limb liner produced greater reduction in PLP compared to a placebo. Halbert, Crotty and Cameron's (2002) findings were based solely on the moderate quality study by Conine, Herschler, Alexander, and Crisp (1993), which was included in our review but the raw data was re-analysed in our review to investigate between group rather than within group changes reported in the original paper (Doing this facilitated a more straightforward comparison with the Hsiao et al (2012) study but it had no material effect on the results of that RCT). In addition, our findings were also based on the negative findings in a high quality trial by Hsiao et al (2012) published after the Halbert, Crotty and Cameron's (2002) review. Given the negative findings of the methodologically stronger paper, it could be argued that on balance of the current evidence an effect seems unlikely. In addition, though not within the remit of this review, the proposed underlying mechanisms of action for electromagnetically shielding limb liners is unclear. However, within the GRADE framework used in our review classifying the evidence as conflicting is appropriate.

There is limited evidence from one moderate quality RCT (Moseley, 2006) of an effect of GMI on pain and function compared to control at six months post treatment but not immediately post treatment. A review by Morales-Osorio and Majia-Majia (2012) which aimed to determine the evidence of GMI treatment for PLP, provisionally identified 352 articles of which only one RCT, Moseley (2006), was included in their review (Moseley, 2006). Thus our findings reflect those of Morales-Osorio and Majia-Majia's (2012). This RCT on GMI included 51 participants but only nine of these had PLP due to amputation and our systematic review reports upon these nine only. With small samples differences are more likely to be spurious than with large samples, and less generalisable (Brutus, Aguinis and Wassmer, 2013). It is for this reason that we recommend caution in interpreting the results of this study.

Our review found Limited evidence from one moderate quality RCT (Brodie, Whyte and Niven, 2007) that a single session of mirror therapy compared to a control had no effect on PLP immediately post intervention. This is in contrast to Rothgangel et al's (2011) systematic review, which evaluated the effect of mirror therapy after stroke, PLP and CRPS. In relation to PLP Rothgangel et al (2011) found low quality evidence of mirror therapy for PLP. Of the 21 articles included in their study only two were related to PLP: Moseley (2006) and Chan et al (2007). Moseley's (2006) RCT focused on GMI which encompassed mirror therapy as a component part rather than mirror therapy alone. For that reason we dealt with it as a trial of GMI rather than mirror therapy. Chan et al (2007) reported small but statistically and clinically significant benefits after 4 weeks of mirror therapy but the study was excluded from our review due to its low quality (scored 3/12: the key limitation was that the randomisation and concealment methods were not described. See table 3). In comparison, the results of Brodie, Whyte

and Niven (2007), were taken immediately after a single intervention. Thus, it may be that to create a clinical effect of mirror therapy on PLP, treatment over a longer period of time with multiple inputs is required which may explain the negative results of Brodie, Whyte and Niven (2007).

Our review found limited evidence from one moderate quality RCT (Rickard, 2004) that hypnosis is effective for pain in the short-term compared to a waiting list control. The magnitude of improvement reported for hypnosis was far greater than seen for any of the interventions in the other studies reported in this review. This may have been related to the use of a waiting list control rather than a placebo/ or usual care control as used in those other trials. The hypnosis study also required participants to be rated as being hypnotizable based upon a standard screening tool. It was not reported how many individuals volunteered for the study but did not meet this inclusion criteria. This may reduce the generalizability of the findings of Rickard (2004). In addition Rickard (2004) was published as a PhD thesis rather than a peer-reviewed article.

With regard to safety, only one of the included studies reported any adverse effects (Conine, Herschler, Alexander, and Crisp, 1993). In this instance 6/52 participants withdrew due to the discomfort associated with wearing a limb liner (Placebo and active). It is worth noting that one of the RCTs excluded from this review due to low methodological quality, also reported side effects associated with limb liners where 2/30 participants withdrew from the study one due to pressure from the liner and one from skin irritation (Kern, Altkemper and Kohl, 2006). In contrast the study investigating hypnosis reported no adverse effects (Rickard, 2004) while the other studies did not directly report on adverse effects. If the lack of adverse effect reports

from the non-linear conservative therapy studies truly reflect a lack of adverse effects of these interventions then this provides some support for using these methods considering the more significant adverse effects associated with pharmacological interventions including: tiredness, dizziness, sweating, constipation, nausea, vertigo, itching and shortness of respiration (Huse, Larbig, Flor, Birbaumer, 2001). However, given the lack of detail in the literature it is possible that adverse effects were simply not reported rather than not present reinforcing the need to investigate this further.

Seven RCTs were identified within this study but not included in the final review as the RCTs were considered to be of lower quality scoring <6/12 on the quality assessment scale. These studies investigated the following interventions; TENS, mirror therapy, electromagnetically shielding limb liners, electromagnetic-own-signal treatment (EMOST), phantom exercises, transcranial magnetic stimulation, and sensory motor discrimination training. Considering the limited methodological quality of these trials their findings should be viewed with caution. However it is worth drawing attention to two interventions – transcranial magnetic stimulation (Ahmed et al. 2011) and sensory discrimination training (Flor et al. 2001). Both demonstrated clinically relevant improvements in pain with transcranial magnetic stimulation increasing beta-endorphin levels and sensory discrimination training improvements being directly correlated with improvements in cortical reorganization on FMRI. Given the magnitude of the effect and the accompanying physiological data demonstrating a plausible method of effect, these interventions warrant further investigation.

Strengths and Limitations

A strength of this review is its adherence to high quality methodological standards in accordance with the guidelines from the Cochrane Collaboration (Higgins and Green, 2008). This is the first study to do so for non-pharmacological interventions for patients with PLP and thus an important addition to the literature. To minimize bias, we limited our selection of studies to RCTs of at least moderate quality. As a result the overall number of trials included in the review was low (n=5), which limits the scope of recommendations that can be made. One could argue that focusing solely on RCTs was a limitation as it meant that information from other study designs was not taken into account. In doing so potentially important evidence may have been omitted. It has been argued that to focus solely on RCT's is blinkered and oversimplifies the complex individual specific nature of clinical care (Williams and Garner, 2002). However, RCTs are considered the optimal design to minimise research bias (Bradley and Nolan, 2008; O'Connor, Green, Higgins, and Deeks, 2008; Higgins et al 2011).

One major limitation of this review is the inability to make judgments on longer-term effects, due the complete lack of long-term outcome data within RCTs, which is important given the chronic nature of PLP. Similarly, no studies have reported on the effectiveness of interventions for social function. This is important given the chronicity of the condition and the isolating consequences of persistent pain. A key limitation of this review is that the studies included had very small sample sizes, which severely limited the quality of the work in itself despite other parts of the design being reasonably sound.

Clinical implications

This review shows that any decision to use conservative interventions for PLP can only be informed by direct evidence from RCTs to a very limited degree. There is some rudimentary support for the use of GMI and hypnosis. For both of these interventions the magnitude of the effect was above that which would be considered clinically relevant, which is encouraging. However, due to a lack of high quality consistent findings from the research, no firm clinical conclusions can be made at this time.

Future research

This study highlights that 15 years after the last comprehensive review of non-pharmacological conservative interventions for PLP, only one high quality RCT has been published. This raises serious questions about the approach to research in this clinical area. The limitations to the different studies do not appear to be methodological issues related to the specific clinical area but rather generic key quality issues such as appropriate randomisation and concealment.

There is a solid body of opinion based on low-level research, clinical observation and theoretical reasoning to make the case for carrying out research to investigate the effectiveness of non-pharmacological conservative therapies that show early promise. The limited evidence of potential benefit from GMI and hypnosis warrants further research on these interventions. Other interventions which show potential for reducing PLP, but were not tested within RCTs of sufficient methodological quality to warrant inclusion in the final review, such as sensory discrimination training and repetitive transcranial magnetic stimulation (Flor, Denke, Schaefer, and Grusser, 2001; Ahmed, Mohamed and Sayed, 2011). Thus, these interventions warrant future more rigorous investigation. These should take the form of RCTs informed by appropriate feasibility

studies to answer basic questions about comparisons with controls and placebo interventions. They should also include long-term follow-up periods.

CONCLUSION

In this systematic review we identified five RCTs (consisting of 135 participants) of moderate to high quality that have investigated the effectiveness of non-pharmacological conservative therapies for patients with PLP. There is limited evidence to support the use of GMI and hypnosis, while a single session of mirror therapy appears ineffective and the evidence supporting electromagnetically shielding limb liners is conflicting. Reporting of adverse effects is low for conservative therapy. These findings provide initial support for more rigorous testing of these types of interventions. However, overall, there is a paucity of high-quality evidence investigating the safety and effectiveness of conservative therapies to guide the clinical management of PLP.

Declaration of Interests

Two of the authors (CR and DM) are currently developing a medical device for the treatment of PLP. The remaining author (SB) has no conflict of interest to declare.

REFERENCES

Ahmed MA, Mohamed SA, Sayed D. 2011 Long-term antalgic effects of repetitive transcranial magnetic stimulation of motor cortex and serum beta-endorphin in patients with phantom pain. *Neurological Research* 33: 953-958.

Alviar MJM, Hale T, Dungca M 2011 Pharmacologic interventions for treating phantom limb pain. *Cochrane Database of Systematic Reviews*, Issue 12. Art. No.: CD006380. DOI: 10.1002/14651858. CD006380.pub2.

Bettany-Saltikov J 2010 Learning how to undertake a systematic review: part 2. *Nursing Standard* 24: 47-56.

Bokkon I, Till A, Grass F, Szabo AE 2011 Phantom pain reduction by low frequency and low intensity electromagnetic fields. *Electromagnetic Biology and Medicine* 30: 115-127.

Bradley E, Nolan P 2008 Evidence-based practice: implications and concerns. *Journal of Nursing Management* 16: 388-393.

Brodie EE, Whyte A, Niven C 2007 Analgesia through the looking glass? A randomised controlled trial investigating the effect of viewing a virtual limb upon phantom limb pain, sensation and movement. *European Journal of Pain* 11: 428-436.

Brutus S, Aguinis H, Wassmer U 2013 Self-reported limitations and future directions in scholarly reports: Analysis and recommendations. *Journal of Management* 39: 48-75.

Chan BL, Witt B, Charrow AP, Magee A, Howard R, Pasquina PF 2007 Mirror therapy for phantom limb pain. *New England Journal of Medicine* 357: 2206-2207.

Conine TA, Herschler C, Alexander SA, Crisp R 1993 The efficacy of Farabloc in the treatment of phantom limb pain. *Canadian Journal of Rehabilitation* 6: 155-161.

Deeks JJ, Higgins JPT, Altman DG 2008 Analysing data and undertaking meta-analysis. In Higgins JPT, Green S (ed). *Cochrane Handbook for Systematic Reviews of Interventions*, pp 243-296. Chichester, Wiley-Blackwell.

Dickersin K, Scherer R, Lefebvre C 1994 Identifying relevant studies for systematic reviews. *British Medical Journal* 309: 1286-1291.

Finsen V, Persen L, Lovlien M, Veslegaard EK, Simensen M, Gasvann AK, Benum PAL 1988 Transcutaneous electrical nerve stimulation after major amputation. *Journal of bone and joint surgery* 70-B: 109-112.

Flor H 2002 Phantom limb pain: characteristics, causes, and treatment. *The Lancet Neurology*, 1: 182-189.

Flor H, Denke C, Schaefer M, Grusser S 2001 Effect of sensory discrimination training on cortical reorganisation and phantom limb pain. *The Lancet* 357: 1763-1764.

Furlan AD, Pennick V, Bombardier C, van Tulder M 2009 Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *SPINE* 18: 1929-1941.

Giummarra MJ, Moseley LG 2011 Phantom limb pain and bodily awareness: current concepts and future directions. *Current Opinions in Anesthesiology* 24: 524-531.

Halbert J, Crotty M, Cameron ID 2002 Evidence for the optimal management of acute and chronic phantom pain: A systematic review. *Clinical Journal of Pain* 18: 84-92.

Higgins J, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne J 2011 The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal* 343: 1-9.

Higgins JPT, Altman DG, Sterne JAC 2008 Assessing risk of bias in included studies. In Higgins JPT, Green S (ed) *Cochrane Handbook for Systematic Reviews of Interventions*, pp187-241. Chichester, Wiley-Backwell.

Higgins JPT, Deeks JJ 2008 Selecting studies and collecting data. In Higgins JPT Green S (ed) *Cochrane Handbook for Systematic Reviews of Interventions*, pp 151-183. Chichester, Wiley-Backwell.

Higgins JPT, Green S (ed) 2008 Cochrane handbook for systematic reviews of interventions. Chichester: Wiley-Blackwell.

Hsiao AF, York R, Hsiao I, Hansen E, Hays RD, Ives J, Coulter ID 2012 A randomised controlled study to evaluate the efficacy of noninvasive limb cover for chronic phantom limb pain among veteran amputees. *Archives of physical medical rehabilitation* 93: 617-622.

Huse E, Larbig W, Flor H, Birbaumer N 2001 The effect of opioids on phantom limb pain and cortical reorganization. *Pain*, 90(1-2): 47-55.

Katz J, Melzack R 1991 Auricular transcutaneous electrical nerve stimulation reduces phantom limb pain. *Journal of Pain Symptom Management* 6: 73–83.

Kern U, Altkemper B, Kohl M 2006 Management of phantom pain with a textile, electromagnetically-acting stump liner: a randomised, double-blind, crossover study. *Journal of pain and symptom management* 32: 352-360.

Lefebvre C, Manheimer E, Glanville J 2008 Searching for studies. In Higgins, JPT, Green S (ed) *Cochrane Handbook for Systematic Reviews of Interventions*, pp 96-147. Chichester, Wiley-Backwell.

Loke YK, Price D, Herxheimer A 2008 Adverse effects. In Higgins JPT, Green S (ed) *Cochrane Handbook for Systematic Reviews of Interventions*, pp 433-448. Chichester, Wiley-Backwell.

Lundeberg T 1985 Relief of pain from a phantom limb by peripheral stimulation. *Journal of Neurology* 232: 79–82.

Moher D, Liberati A, Tetzlaff J, Altman DG 2009 Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 6(7), e1000097.

Morales-Osorio MA, Majia-Majia JM 2012 Treatment with graded motor imagery in phantom limb pain syndrome: a systematic review. *Rehabilitacion* 46: 310-316.

Moseley LG 2006 Graded motor imagery for pathologic pain. *Neurology* 67: 2129-2134.

Moseley LG 2006 Graded motor imagery for pathologic pain. *Neurology* 67: 2129-2134.

Moura VL, Faurot KR, Gaylord SA, Mann JD, Sill M, Lee MY 2012 Mind-Body Interventions for treatment of phantom limb pain in persons with amputation. *American Journal of Physical Medicine and Rehabilitation* 91: 701-714.

NHS (2014) NHS Choices website, Amputation. Available at: <http://www.nhs.uk/conditions/amputation/pages/introduction.aspx>. (Accessed 21/09/2015).

Nikolajsen L 2012 Post-amputation pain: Studies on mechanisms. *Danish Medical Journal* 59: 1-21.

O'Connor D, Green S, Higgins JPT, Deeks JJ 2008 Defining the review question and developing criteria for including studies. In Higgins JPT, Green S (ed) *Cochrane Handbook for Systematic Reviews of Interventions*, pp 83-94. Chichester, Wiley-Backwell.

Rickard JA 2004 *Effects of Hypnosis in the Treatment of Residual Stump and Phantom Limb Pain*. Pullman, WA: Washington State University.

Rothgangel AS, Braun SM, Beurskens AJ, Seitz RJ, Wade DT 2011 The clinical aspects of mirror therapy in rehabilitation: a systematic review of the literature. *International Journal of Rehabilitation Research* 34: 1-13.

Roulet S, Nouette-Gaulain K, Brochet B, Sztark F 2009 Phantom limb pain: from physiopathology to prevention. *Annales Francaises D'anesthesie et de Reanimation* 28: 460-72.

Ulger O, Topuz S, Bayramlar K, Sener G, Erbahceci F 2009 Effectiveness of phantom exercises for phantom limb pain. *Journal of Rehabilitation Medicine*, 41: 582-584.

Van Tulder M, Furlan A, Bombardier C, Bouter L, and the Editorial Board of the Cochrane Collaboration Back Review Group 2003 Updated Method guidelines for

systematic reviews in the Cochrane collaboration back review group. SPINE 28: 1290-1299.

Viera AJ, Garrett JM 2005 Understanding interobserver agreement: The kappa statistic. Family Medicine 37: 360-3.

Williams DDR, Garner J 2002 The case against 'the evidence': a different perspective on evidence-based medicine. The British Journal of Psychiatry 180: 8-12.

Wu C, Tella P, Staats P, Vaslav R, Kazim D, Wesselmann 2002 Analgesic effects of intravenous lidocaine and morphine on post-amputation pain. Anesthesiology 96: 841-848.

Footnotes

¹ It is proposed that such liners work by blocking severed nerve endings in the residual limb from irritation via environmental electrical and magnetic fields (Conine et al. 1993).

² These figures were calculated from a graphical figure presented in Moseley (2006) using the digitizing software DigitizIt (version 2.1 for windows XP – windows 8, <http://www.digitizeit.de/index.html>, Germany).