Title Page

The effects of graded motor imagery and its components on chronic pain: a systematic review and meta-analysis**.**

K. Jane Bowering[1], Neil E. O'Connell[2], Abby Tabor[1,3], Mark. J Catley[1], Hayley B. Leake[1], G. Lorimer Moseley[1,4], Tasha R. Stanton[1,4].

[1] Sansom Institute for Health Research, University of South Australia

[2] Centre for Research in Rehabilitation, Brunel University

[3] King's College, London

[4] Neuroscience Research Australia

Corresponding Author:

Tasha R. Stanton

Email: Tasha.stanton@unisa.edu.au

Tel: +618 8302 2090

Running title: Effect of graded motor imagery on chronic pain

Indexing words:

Graded motor imagery, GMI, mirror therapy, motor imagery left/right judgements, chronic pain, systematic review

Abstract

Graded motor imagery (GMI) is becoming increasingly used in the treatment of chronic pain conditions. The objective of this systematic review was to synthesise all evidence concerning the effects of GMI and its constituent components on chronic pain. Systematic searches were conducted in 10 electronic databases. All randomised controlled trials (RCTs) of GMI, left/right judgement training, motor imagery, and mirror therapy used as a treatment for chronic pain were included. Methodological quality was assessed using the Cochrane risk of bias tool. Six RCTs met our inclusion criteria and the methodological quality was generally low. No effect was seen for left/right judgement training and conflicting results were found for motor imagery used as stand-alone techniques, but positive effects were observed for both mirror therapy and GMI. A meta-analysis of GMI versus usual physiotherapy care favoured GMI in reducing pain (2 studies, n = 63; effect size 1.06 (0.41, 1.71); heterogeneity, I^2 = 15%). Our results suggest that GMI and mirror therapy alone may be effective although this conclusion is based on limited evidence. Further rigorous studies are needed to investigate the effects of GMI and its components on a wider chronic pain population.

Perspective: This systematic review synthesises the evidence for GMI and its constituent components on chronic pain. This review may assist clinicians in making evidence-based decisions on managing patients with chronic pain conditions.

Key words

Graded motor imagery, GMI, mirror therapy, motor imagery, left/right judgements, chronic pain, systematic review

Introduction

Rapid advances in our understanding of the role of the brain in chronic pain has seen the development of treatments for chronic pain that directly target cortical reorganisation[32, 44]. The first of these treatments was developed in response to remarkable findings in amputees with phantom limb pain (PLP), which showed that pain was associated with reorganisation of the primary sensory cortex contralateral to the amputated limb. The normal representation of the amputated hand had been invaded by the representation of the lip[13]. This cortical reorganisation has also been demonstrated for chronic low back pain, in which representation of the painful side of the back was enlarged and shifted medially as compared with representation in healthy controls[12]. That primary sensory cortex receptive fields can be modified by tactile stimuli with a behavioural relevance (for example, eating or braille) is now well accepted[14]. Flor and colleagues aimed to exploit this plasticity in amputees with PLP by two weeks of sensory discrimination training, in which participants discriminated between stimuli of different frequency and at different locations on their stump[11, 15]. Their randomised controlled trial showed normalisation of cortical organisation and clinically important reduction of pain. This process, from discovery of altered sensory cortex organisation to targeted sensory discrimination training for clinical benefit, has been repeat in complex regional pain syndrome (CRPS)[17, 19, 34, 36].

As well as physiological evidence of disrupted somatotopic representation in chronic pain, there is behavioural evidence of disrupted spatial representation too – disrupted processing of stimuli delivered to healthy body parts held in the affected 'space'[33], the abnormality of the perceived size of the painful body part[21, 22, 29, 31], and poor voluntary movement and motor imagery performance[1, 6, 7, 27, 30, 38-40]. One treatment that was developed to directly target these cortical disruptions is graded motor imagery (GMI), a three-stage treatment that aims to gradually engage cortical motor networks without triggering the protective response, pain. This treatment gets its theoretical framework from the principle established in the physical therapies, of graded increase in activity. This principle is adapted in GMI to cater for both the overly sensitive nociception/pain systems and the disrupted cortical mechanisms mentioned above. Graded motor imagery was developed initially for an application to chronic limb, or phantom limb, pain, but has been extended clinically to chronic back pain, where a component of GMI has been used for some time[43].

The first stage of the GMI programme is left/right judgements of photographs that depict the affected area. For limb pain, this involves viewing an image of a limb and judging whether that image depicts a left or a right limb. Functional brain imaging studies in healthy subjects have shown that

this task selectively activates the pre-motor cortex without activating primary motor areas[37, 41, 45]. The second stage, motor imagery, requires imagined movement of the area. These imagined movements have demonstrated to activate similar motor cortical areas as the actual execution of that movement[10]. For the final stage, mirror therapy, patients place their affected limb inside a mirror box and watch movements of their non-affected limb in the mirror, giving the illusion of a moving, but pain-free, affected limb. This task activates the motor cortex and also provides a strong visual input to the cortex that the movements are occurring normally and without impediment[20]. While functional brain imaging studies have supported the proposed cortical activation for each stage of GMI in healthy subjects, no studies have investigated cortical activation of GMI stages in pain patients. These imaging studies nonetheless provide support for the possibility that similar sequential activation of cortical areas within each stage of the GMI programme could occur in pain patients.

Both GMI or its components have been used in the clinical setting to treat chronic pain conditions such as complex regional pain syndrome (CRPS), phantom limb pain (PLP), and back pain. However, an issue that remains to be addressed is whether the evidence supports or negates the use of GMI or its components in the treatment of a wider chronic pain population. A recent systematic review evaluating interventions for treating CRPS supported the use of GMI[8]. However, a recent clinical audit of CRPS multimodal management including, but not limited to, GMI clearly showed no benefit of treatment[16]. These conflicting findings, and that GMI has not, to our knowledge, been empirically evaluated in a wider chronic pain population, highlight the importance of systematic evaluation of the entire literature concerning GMI and its components. The aim of this review and meta-analysis was to synthesise all available literature regarding the efficacy of GMI programmes, or any of the three constituent components, on chronic pain. The results of this systematic review will enable clinicians to make evidence-based decisions on the use of GMI with chronic pain patients.

Materials and Methods

Data Sources

For this review, several health-based databases were searched from their relative inception through to January 2012. The electronic search was performed using the following databases: Medline (via OvidSP), Embase (via Ovid SP), CINAHL, Scopus, Academic Search Premier, Web of Science, Allied and Complementary Medicine, PubMed, the Cochrane Collaboration and the Physiotherapy Evidence Database (PEDro). A sensitive search was completed using a combination of keywords and

4

relevant subjects heading for GMI, its components, and chronic pain. The relevant subject headings were determined specific to each database. The complete Medline search strategy is provided in Appendix One. Searches were limited to English language and humans only. To attempt to identify grey literature (specifically non-indexed published trials, conference abstracts, and book chapters) experts were contacted and asked to contribute any materials not identified by database search. The references of all relevant articles were also hand-searched for further articles. We did not search the clinical trials registers for unpublished studies.

Study Selection

Four reviewers were paired and each pair independently screened the titles and abstracts of half of the potential studies – thus, all papers were screened by two people. Results of the screening process were compared within pairs. In this process, studies were retained if they evaluated GMI or at least one component of GMI. Following initial screening, the full text of potentially relevant studies were retrieved and reviewed independently by two reviewers. Studies were retained if they met the following criteria: human adult subjects (>18 years of age); clinically validated pain measure used; randomised controlled trial; subjects all had a chronic pain condition lasting longer than 3 months. No restrictions were placed on the comparison group used (i.e., placebo, wait-list control, or other active treatment). Any discrepancies were resolved through discussion, or if necessary, through consultation with a third independent reviewer.

Outcome measures

Pain intensity ratings were the primary outcome of interest for this review. This included self-reported measures such as the McGill pain questionnaire (MPQ), a visual analogue scale (VAS), a numerical rating scale (NRS), a neuropathic pain scale (NPS), or a categorical rating of pain (such as mild, moderate, severe). A rating of pain using one of these measures was required immediately pre-intervention and immediately post-intervention. Follow-up pain ratings were a secondary outcome of interest for this review.

Risk of Bias Assessment and Data Extraction

Two reviewers independently assessed the risk of bias of included studies using The Cochrane Collaboration's tool for assessing risk of bias. For the category for 'other' sources of bias, the reviewers were particularly concerned with similarity of pain scores at baseline as this is recommended by other quality assessment tools such as PEDro[9]. In the 'other' source of bias category we also included evaluation of sample sizes (i.e., less than 50 participants per treatment

5

arm considered a high risk of bias)[24]. These items were added as we anticipated that studies identified were likely to be small and, as such, these factors were more likely to represent a significant source of bias.

For all eligible studies, data extraction was completed independently by two reviewers using a customised data extraction form. This data extraction form was piloted before use. Data extracted included: participant characteristics such as age, gender, pain condition, and length of pain; the outcome measure used; the control and treatment intervention choices, and their length (minutes per each session), frequency (sessions per day/week), and total duration (weeks of intervention); baseline and immediate post-intervention pain scores, and follow-up pain scores if provided. Any disagreements regarding risk of bias or data extraction were resolved through discussion, or if necessary, through consultation with a third independent reviewer. If necessary, authors were contacted to provide further information.

Data Synthesis

We sought to pool data for pain relief from studies where adequate data were available. We planned *a priori* to pool data from studies comparing GMI programmes with usual care or no treatment, and to perform separate meta-analyses for studies that investigated similar individual components of GMI.

Data were pooled using Review Manager 5 software[5] using a random effects inverse-variance approach. A random effects model was chosen as it was anticipated and subsequently confirmed, that there would be differences in the populations and interventions studied which would suggest that the effects might differ somewhat across studies. Using the post-intervention means of each group and the pooled post-intervention standard deviations of pain scores, the standardised mean difference (Hedge's g) was calculated for each study to allow comparison between studies. Effect sizes were interpreted according to Cohen[4] (≤ 0.2 small, 0.5 moderate, ≥ 0.8 large). We used the Chi^2 test to detect statistically significant heterogeneity and the I^2 statistic to estimate the amount of heterogeneity. When heterogeneity was high, we did not pool the outcomes. Further, we considered it inappropriate to pool data from studies that used full GMI programmes with those that used individual components of GMI since it does not follow that the different types of interventions should be estimating the same effect size. We therefore planned separate meta-analyses for these types of studies considering both short-term (immediately post-intervention or the closest measure presented to that point) and follow-up (>4 weeks post-intervention) time points. We undertook a

6

sensitivity analysis to investigate the influence of using a random effects model by reanalysing the data using a fixed effects model.

In studies that evaluated a comprehensive GMI programme, the effect sizes for the first component (i.e., left/right judgements stage) were also calculated using post-intervention scores when individual participant data was present. It was decided *a priori*, that effect sizes would not be calculated for the second or third GMI treatment components (motor imagery and mirror therapy, respectively) because in these latter components, the methodological tenets of the RCT study design do not hold. Specifically, participants are not re-randomised following each component stage, so there are pre-intervention pain differences between groups in the latter stages. That the responses of the latter components were due to carry-over effects or continuing improvement from the previous treatment could therefore not be ruled out. We did not establish any *a priori* sensitivity or subgroup analyses as we anticipated identifying inadequate data to support this process.

Results

Study Description

The initial literature search yielded 6160 records following the removal of duplicates. Six thousand and fifteen studies were excluded in the initial screening of title and abstracts. One hundred and thirty-nine studies were then excluded following review of the full-text. The most prevalent reason for exclusion was that articles did not include primary research data; primarily, these were reviews, conference abstracts, and book chapters, all presented in a narrative form. Other reasons for exclusion were studies that included sample populations without chronic pain or did not include pain outcome measures, were not of RCT design, were non-English studies and those that included children. The screening and review process is shown in a PRISMA flow-diagram in Figure 1. Key data of the remaining 6 RCTs included are summarized in Table 1.

Characteristics of included studies

Three studies evaluated the effects of GMI on chronic pain[25, 26, 28]. Two of these studies compared a six week programme of GMI to usual physiotherapy care[25, 26]. The third study compared an ordered programme of GMI to an unordered programme of GMI[28]. Participants were instructed to spend ten minutes of each waking hour on the intervention. All studies collected follow-up data: one at six weeks post-intervention[26], one at twelve weeks post-intervention[28], and one at six months post-intervention[25]. These studies used varying methods of collecting participant pain scores. The author of each study was contacted, and numerical rating scale (NRS) data for each participant's pain level

7

was provided. These NRS data were used in the analyses. Only one study[25] provided data on adherence to the treatment programme. This study found both GMI and usual care groups had adherence rates of 75%.

Three other studies evaluated individual components of GMI[2, 3, 23]. No studies primarily evaluated left/right judgements; however, two studies[25, 26] evaluating GMI provided sufficient data to enable calculation of effect sizes for the two weeks of left/right judgement training. Three studies[2, 3, 23] evaluated the effects of mirror therapy on chronic pain. Two studies[2, 3] evaluated the effects of motor imagery. The time spent on the intervention differed between studies. In one study, participants completed five, 1-hour sessions of mirror therapy a week[23]. In the second study, participants spent thirty minutes per day doing either mirror therapy or motor imagery, depending on their group allocation[2]. In the third study, participants spent fifteen minutes per day doing either mirror therapy or motor imagery, depending on their group allocation[3]. Follow-up data from these studies were collected either at four weeks [2, 3] or six months[23]. All three studies used 100mm visual analogue scale (VAS) data to report participant's pain levels.

Characteristics of included populations

The participants in each study had experienced pain for greater than three months. The chronic pain conditions included CRPS[2, 25, 26, 28], PLP[3, 25], and pain following stroke[23]. Studies including children were excluded from this review. The mean age in each study ranged from 32 to 57 years. Overall, there were more females (n=90) than males (n=81) in the included studies.

Insert figure one

Table 1. Study Characteristics Data for Randomised Controlled Trials of Graded Motor Imagery or its Components for Chronic Pain

Study	Participants	Condition	Intervention	Outcome measures
Studies evaluating the components of GMI				
[23]Michielsen et al. (2011)	n = 40 Mean age = 57* Gender = 50% male	Chronic pain following stroke (mean time since stroke 3.9 years)	Exp: 6 week bilateral hand movement with mirror therapy programme. Practised 5xweek, 1 hour a session. Con: 6 week bilateral hand movements. Practised 5xweek, 1 hour a session.	100mm **VAS** Follow-up: 6 months
[2]Cacchio et al. (2009)	n = 24 Median age = 62 (53 to 71)** Gender = 46% male	CRPS	Exp: 4 week mirror therapy programme, 30 mins daily. Con: 4 week covered mirror programme, 30 mins daily. Exp2: 4 weeks of motor imagery, 30 mins daily	100mm VAS Follow-up: 4 weeks
[3]Chan et al. (2007)	n = 22 Mean age = 29 ±8.8*** Gender = 100% male	PLP	Exp: 4 week mirror therapy programme, 15 mins daily. Con: 4 weeks covered mirror programme, 15 mins daily. Exp2: 4 weeks of motor imagery, 15 mins daily	100mm **VAS** Follow-up: 4 weeks
Studies evaluating GMI				
[25]Moseley (2006) †	n= 50 Mean age = 41 ±16*** Gender = 36% male	CRPS, PLP following amputation or brachial plexus avulsion	Exp: laterality retraining, motor imagery, mirror therapy. 2 weeks each component, 10 mins for each waking hour. Con: usual physiotherapy/other treatment.	MPQ, **NRS** Follow-up: 6 months
[28]Moseley (2005)	n=20 Mean age = 32 ±11*** Gender = 30% male	CRPS1	Exp: sequential GMI. 2 weeks each component, 10 mins for each waking hour. Con: non-sequential GMI: MI, left/right, MI. 2 weeks each component, 10 mins for each waking hour. Con2: non-sequential GMI: left/right, mirror, left/right. 2 weeks each component, 10 mins for each waking hour.	NPS, **NRS** Follow-up: 12 weeks
[26]Moseley (2004) †	n=13 Mean age = 57 ±19*** Gender = 30% male	CRPS1	Exp: sequential GMI. 2 weeks each component, 10-15 mins for each waking hour. Con: usual physiotherapy/other treatment.	NPS, **NRS** Follow-up: 6 weeks

Exp = experimental group; Exp2 = secondary control groups; Con = control group; Con2 = secondary control group; n= number recruited (prior to drop-out or loss to follow-up); GMI = graded motor imagery; MI = motor imagery; left/right = left/right judgements; mirror = mirror therapy; CRPS = complex regional pain syndrome; PLP = phantom limb pain; VAS = visual analogue scale; MPQ = McGill Pain Questionnaire; NPS = neuropathic pain scale; NRS = numerical pain rating scale. * Range or Standard Deviation not provided; ** Range; ***Standard Deviation; †Due to the presence of individual participant post-intervention data, the left/right judgements component of treatment was also examined; Data used to calculate effect sizes are **bolded.**

Risk of Bias of Included Studies

The results of the risk of bias assessment are shown in Table 2. The study appraised to be at lowest risk of bias was that by Moseley[25], who met every criteria except the blinding of therapists and participants and the 'other' category, for its small sample size. None of the six included RCTs met the blinding of therapists and participants criterion. In therapy trials such as these, direct participant-therapist involvement means blinding is not feasible; hence, all six RCTs had non-blinded therapists and participants. While blinding in these trials is not feasible, it is still an inherent source of bias that must be highlighted for every study. No study was free of additional bias, as all studies had sample sizes less than 50. Michielsen et al.[23] presented additional bias in that they failed to report any baseline similarities or differences between groups on pain. Two other studies also failed to report whether groups had similar baseline pain levels[2, 3]. This lack of this information has implications for the validity of the observed effect sizes as it is uncertain whether differences found between groups may have been influenced by baseline group differences. These same studies also failed to provide information regarding whether the person who included/excluded participants was blinded to treatment allocation. Given the lack of participant/therapist blinding due to nature of the interventions within the studies, all studies were considered to have some inherent bias.

Table 2. Risk of Bias Assessment of Included Randomised Controlled Trials

	Random allocation	Concealed allocation	Blinding of: participants/ therapists	outcome assessors	Incomplete data	No selective outcome reporting	Free of additional bias
[23]Michielsen et al.(2011)	Y	Y	N	Y	Y	Y	N
[2]Cacchio et al. (2009)	U	U	N	N	U	Y	N
[3]Chan et al. (2007)	U	U	N	N	U	Y	N
[25]Moseley (2006)	Y	Y	N	Y	Y	Y	N
[28]Moseley (2005)	Y	U	N	Y	Y	Y	N
[26]Moseley (2004)	Y	U	N	Y	Y	N	N

Abbreviations: Yes (Y), low risk of bias; no (N), high risk of bias; Unclear (U), uncertain risk of bias.

Table 3. Effect Sizes (95% CI) for GMI and its Components on Chronic Pain when Compared to Control Groups

The effect sizes are standardised mean differences, calculated using Hedge's G (i.e., the difference in post-intervention pain scores between control and intervention groups divided by the pooled standard deviation of the two groups, each weighted for sample size). Effect sizes are grouped according to intervention type. Positive effect sizes indicate a lower pain score in the intervention group, favouring the intervention group. Negative effect sizes indicate a lower pain score in the control group, favouring the control group.

Study	Control	Number of participants		Post-intervention pain (mean±SD)		Effect size (95% CI)
		Control	Intervention	Control	Intervention	
Laterality judgement task						
Moseley (2006)[25]	Usual care	25	25	54±13	48±14	0.44 (-0.12, 1.00)
Moseley (2004) [26]	Usual care	6	7	61±10	57±15	0.29 (-0.81, 1.39)
Motor imagery						
Cacchio et al. (2009) [2]	N/A	-	8	-	-	-
Chan et al. (2007) [3]	Covered mirror therapy	6	6	34±22	58±20	-1.05 (-2.30, 0.19)
Mirror therapy						
Michielsen et al. (2011) [23]	Bilateral hand movements	19	17	9.2±14	8.8±10.8	0.03 (-0.62, 0.69)
Cacchio et al. (2009) [2]	N/A	-	8	-	-	-
Chan et al. (2007) [3]	Covered mirror therapy	6	6	34±22	17±21	0.73 (-0.46, 1.92)
	MI	6	6	58±20	17±21	1.85 (0.40, 3.29)^
Graded motor imagery						
Moseley (2006) [25]	Usual care	25	25	47±16	33±15	0.89 (0.31 to 1.47)^
Moseley (2005) [28]	MI, left/right, MI	6	7	40±10	33±8	0.73 (-0.41, 1.87)
	Left/right, mirror, left/right	6	7	42±9	33±8	0.99 (-0.19 to 2.17)
Moseley (2004) [26]	Usual care	6	7	58±12	38±10	1.70 (0.36, 3.04)^

N/A = not applicable, MI = motor imagery, left/right = left/right judgements, mirror = mirror therapy. ^$p<0.05$; For all Moseley studies, pain scores and effect estimates are for NRS results.

Insert figure two

Insert figure three

Outcomes

Four authors were contacted to gain additional information required to calculate the effect size of their intervention[2, 3, 23, 25, 26, 28]. One author could not be contacted, thus the effect size for this study could not be calculated[2]. The effect sizes for the remaining studies are presented in Table 3.

GMI Programme

Three studies evaluated the effects of a six-week GMI programme on chronic pain with all finding that GMI reduced pain when compared to usual physiotherapy care[25, 26] and unordered GMI[28]. The two studies comparing GMI to usual physiotherapy care both found large effect sizes (1.70 (0.36, 3.04)[26] and 0.89 (0.31, 1.47)[25]). In the study which compared a course of GMI to an unordered course of GMI[28], moderate to large effects in favour of the ordered GMI were found (0.73 (-0.41, 1.87) and 0.99 (-0.19, 2.17)).

The immediate post-intervention results of the two studies comparing GMI with usual care were pooled[25, 26]. The results of one study[28] evaluating GMI were not included in the meta-analysis because the control group intervention had pronounced differences; this heterogeneity meant pooling of these data were not appropriate. The heterogeneity of the pooled studies was low (I^2=15%), and produced a large pooled effect size (1.06 (0.41, 1.71)) (Figure 2). While the statistical heterogeneity of the studies was low, it must be noted that the chronic pain population in each study differed slightly; one included only CRPS participants[26] and the other a mix of CRPS, PLP and pain after brachial plexus avulsion[25]. Sensitivity analysis using fixed effects, rather than random effects, meta-analysis had no substantive impact on our findings (I^2=0%; effect size 0.97 (0.52, 1.42); test for overall effect, p<0.0001).

Follow-up data also suggests an effect of GMI further reducing pain with large effect sizes reported at 6 months for GMI when compared to usual physiotherapy care (1.59 (0.28, 2.90)[26] and 1.68 (1.02, 2.33)[25], and also at 12 weeks for GMI when compared to an unordered GMI programme (1.35 (0.09, 2.60) and 1.31 (0.06, 2.55))[28]. Pooling of these effect estimates was not appropriate as the follow-up in each study was conducted at a markedly different time point.

Left/right judgements

No studies were found that evaluated left/right judgements as the primary intervention, although two studies investigated the effects of left/right judgements as part of a GMI programme on chronic pain[25, 26]. Neither study found statistically significant effect estimates for left/right judgements

14

reducing pain when compared to usual care. However, the effect estimates produced were positive, albeit small (0.29 (-0.81, 1.39)[26] and 0.44 (-0.12, 1.00)[25]). The heterogeneity of the pooled studies was low (I^2=15%), and produced a similarly small effect estimate (0.41 (-0.09, 0.91))(Figure 3). Sensitivity analysis using fixed effects, rather than random effects, meta-analysis again had no substantive impact on our findings (I^2=0%; effect size 0.41 (-0.09, 0.91); test for overall effect, p=0.11).

Motor imagery

None of the included studies had a primary aim of evaluating the effects of motor imagery on chronic pain. However, in two studies, motor imagery was used as a secondary control group[2, 3] and was compared to covered mirror therapy (in which the participant is instructed to look at a mirror that is covered with a cloth so as to offer no reflection; controlling for attention). These studies found contrasting results. Chan et al.[3] found covered mirror therapy to be much more effective at reducing pain when compared to motor imagery, with a large effect found (-1.05 (-2.30, 0.19)). Interestingly, participants receiving motor imagery treatment had increased pain levels (compared to baseline pain). Similar findings were reported by Cacchio et al.[2], in which six out of eight participants experienced increased pain levels following four weeks of motor imagery. However, Cacchio et al.[2] found no difference between motor imagery and covered mirror therapy (five of eight participants had increased pain in covered mirror therapy group). All pain assessments were immediately post-intervention; no short- or long-term follow-up data was available. Both studies had small sample sizes and had a high risk of bias.

Mirror therapy

A total of three studies evaluated mirror therapy as a stand-alone treatment in chronic pain; in each study, mirror therapy was the primary treatment evaluated[2, 3, 23]. All three studies found positive effects of mirror therapy in reducing pain, despite using different control groups. The effect sizes ranged from trivial (0.03 (-0.62, 0.69)[23], bilateral hand movement control group), to moderate (0.73 (-0.46, 1.92)[3], covered mirror control group), to large (1.85 (0.40, 3.29)[3], motor imagery control group). Notably, this final effect size was the only statistically significant finding in the motor imagery analyses. This finding was further supported by Cacchio et al.[2], who reported seven of eight participants in the mirror therapy group experiencing decreased pain levels (compared to only one of eight participants in the covered mirror group and only two of eight participants in the mental imagery group having decreased pain levels).

15

The pooling of studies of mirror therapy demonstrated high levels of heterogeneity (I^2=63%) but no effect (p=0.07). Visual inspection of the forest plot showed that the one study that utilised a different comparison condition[3] (motor imagery as opposed to covered mirror therapy) was the most likely source of this variance. Post hoc sensitivity analysis removing this study from the analysis reduced this heterogeneity substantially (I^2=2%) and continued to demonstrate no effect (p=0.51). Sensitivity analysis using fixed effects, rather than random effects, meta-analysis again had no substantive impact on our findings (I^2=63%; effect size 0.42 (-0.011, 0.95); test for overall effect, p=0.12).

Only one study presented follow-up data[23], reporting a small, non-significant effect size (0.34 (-0.29, 0.96)) of mirror therapy compared to bilateral hand movements in patients with pain following stroke at 6 months follow-up. All three studies were considered to have a high risk of bias.

Discussion

This is the first review to systematically evaluate the effect of GMI or its components on pain outcomes in people with chronic pain. The limited number of small RCTs available have found mixed results for the effects of GMI or its components on chronic pain. Of the six RCTs identified, all contained some inherent bias. A key finding of this review was that the majority of studies evaluated the effect of GMI or its components in CRPS or PLP, thus it is unclear how GMI might relate to other chronic pain conditions. We will first consider our findings with respect to individual components of GMI and then consider our findings with respect to full GMI programmes.

Effect of individual GMI components on pain

Left/right judgements

Left/right judgements as a sole treatment appear to have no effect on chronic pain[25, 26]. That all effect sizes were positive raises the possibility that even the pooled data were underpowered to detect an effect, but one might conclude that such a small effect is of little clinical consequence.

Because left/right judgements have never been used as a stand-alone treatment for chronic pain, there have been no studies that evaluate only left/right judgements as a treatment for chronic pain. Because only data from the first stage of a GMI programme can currently be used to evaluate effect of left/right judgements, there is no data available on the long-term effect of this treatment. While left/right judgements alone may not produce statistically significant effects, they are an integral part

16

of the sequential GMI programme that our results suggest may be effective. Nonetheless, the clinical importance for left/right judgements *per* se, remains to be shown.

Motor imagery

Motor imagery appears less effective at treating chronic pain than covered mirror therapy[2, 3]. Covered mirror therapy was utilised in these studies as an inactive control condition. That two studies found an increase in pain relative to baseline following motor imagery and one observed greater improvements in an inactive control group suggests that motor imagery might have the potential to increase pain intensity. These findings are consistent with those of a separate pre-/post-treatment trial not included in this review, in which motor imagery increased pain and swelling in those with chronic arm pain[35] and speaks against the use of motor imagery alone as a treatment for chronic pain.

Mirror therapy

Mirror therapy is arguably the most studied component of GMI in terms of its effects on pain; however, much of the available literature concerns case studies, which were excluded from this review. The results of the included studies were consistently positive in favour of mirror therapy reducing pain[2, 3, 23] although there is wide variance in the reported effect sizes.

This variance may reflect differences between studies in the patient group and the choice of control treatment. For example, Michielsen et al.[23] recruited chronic pain patients with very low baseline pain scores, which are atypical of chronic pain populations and provide minimal room for improvement, creating the possibility of a floor effect. In contrast, the baseline pain scores for participants in the Chan et al.[3] study were high, providing the opportunity for greater pain reductions and therefore a larger effect size. Both the Chan et al.[3] and Cacchio et al.[2] studies suggest that mirror therapy is substantially more effective than motor imagery. However, motor imagery appeared to increase participants' pain levels, so the difference might reflect both the worsening in the control motor imagery group and the improvement in the mirror therapy group.

One important consideration when interpreting the effect of mirror therapy relative to a covered mirror control condition is the possible impact of variable placebo effects. That is, covering the mirror might imply to the patient that the mirror is the powerful component of treatment and, as such, the covered mirror condition might not be perceived as credible by the patient. As stated, blinding of therapists and participants in therapy interventions such as mirror therapy is near

impossible. Through matching the frequency and duration of therapy sessions for both the covered and active mirror groups, all studies achieved structural equivalence which is particularly important in situations where indistinguishable placebo controls are not possible[18]. While covered mirror therapy as a control may not be ideal, it is a pragmatic control.

Effect of full GMI programmes on pain

Our results suggest that a GMI programme likely has moderate effects when compared to unordered GMI[28], and large effects when compared to usual physiotherapy care[25, 26]. Both of the two identified studies evaluating GMI versus usual physiotherapy found a large effect size[25, 26] and clearly support the efficacy of GMI, at least as delivered within one clinical centre.

Recently published clinical audit data appear to contradict the GMI findings of this review. Prospective audit data from 32 patients treated at two interdisciplinary centres showed no reduction in pain after a multimodal approach that included GMI[16]; indeed some patients (30% in one centre and 50% in the other) actually reported an increase in their pain intensity following treatment. The authors proposed that variations in GMI protocol from other studies and logistic constraints may have led to the poor result. Nonetheless, this study, while less robust than an RCT, highlights that independent replication of the results of Moseley[26] and Moseley[25] in controlled trials remains a research priority.

That GMI produced moderate effects when compared to an unordered programme of GMI[28] is interesting. The order of GMI components seems to be important, which is consistent with its proposed mechanism[42]. Moreover, that there is such an effect relative to an unordered treatment control group[25, 26] suggests against the possibility that the effects of GMI are largely due to a placebo response. That is, unordered GMI might be a more appropriate placebo control treatment in future studies because it would capture much of the novelty of GMI, but it appears to have little effect. That this finding arises from a single small trial indicates that it also requires independent replication.

Given the limited data available, it is difficult to draw firm conclusions, but these data and those relating to the ordering of GMI components suggest that the gradual and progressive nature of GMI may be clinically important. Motor imagery particularly demands attention. Not only was no significant benefit observed with motor imagery, but unlike with left/right judgements there was no suggestion in the data of a trend towards pain relief with this intervention and some evidence to suggest a worsening of pain. This leads to the inevitable question of whether GMI might be more effective without a motor imagery stage. To our knowledge, no study has currently investigated this.

18

The majority of the evidence pertains to patients with CRPS, and we identified little evidence pertaining to the efficacy of GMI for other chronic pain conditions. Caution is advised when extrapolating these findings to the broader chronic pain population.

Limitations

Non-English studies were not included due to lack of translation resources and we did not search clinical trials registers for unpublished studies. However, experts in the area of GMI/chronic pain were consulted regarding any missing relevant publications or active research groups and did not identify any relevant contributions, so we would suggest that the chance of missing a study would seem low. The number of RCTs included was small, and the majority had a high risk of bias. The limited number of studies published in this area also raises the possibility of publication bias.

In terms of the evidence of the effectiveness of full GMI programmes for reducing chronic pain, perhaps the strongest limitation is that all of the included trials were completed by one research group with which we ourselves are affiliated[25, 26, 28]. To increase confidence in our findings, the need for further trials of GMI by independent research groups by independent research groups cannot be overstated. There was significant heterogeneity between the included study populations; the type and duration of chronic pain varied, and studies used a range of methods for sourcing and recruiting participants. Last, there were very few long-term follow-ups (i.e., all follow-ups were 6 months or earlier), which suggests that the effectiveness of these treatments in the longer term remains unknown.

In conclusion, while the results of this systematic review suggest that the effectiveness of GMI and its components is encouraging in CRPS and PLP, no evidence exists for these treatments in a wider chronic pain population. It is critical to acknowledge that more work is required – the theoretical framework underlying these treatments suggests the value of additional trials in a wider chronic pain population. It is difficult to be certain of the findings since there are very few studies of mixed risk of bias available. Differing methodologies and samples within each study significantly limit the generalisability of these findings to people with CRPS or PLP, although there seems to be good reason to extend this line of investigation into different chronic pain populations.

Acknowledgements

References

[1]Bray H,Moseley GL: Disrupted working body schema of the trunk in people with back pain. British Journal of Sports Medicine 45: 168-173, 2009.

[2]Cacchio A, De Blasis E, Necozione S, di Orio F,Santilli V: Mirror Therapy for Chronic Complex Regional Pain Syndrome Type 1 and Stroke. New England Journal of Medicine 361: 634-636, 2009.

[3]Chan BL, Witt R, Charrow AP, Magee A, Howard R, Pasquina PF, Heilman KM,Tsao JW: Mirror therapy for phantom limb pain. New England Journal of Medicine 357: 2206-2207, 2007.

[4]Statistical Power Analysis for the Behavioural Sciences, 2nd edition, Cohen J (Eds.), Lawrence Erlbaum Associates, Inc. , New Jersey, 1988.

[5]The Cochrane Collaboration. 2011. Review Manager (RevMan), Version 5.1.

[6]Coslett H, Medina J, Kliot D,Burkey A: Mental motor imagery and chronic pain: The foot laterality task. Journal of the International Neuropsychological Society : JINS 16: 603, 2010.

[7]Coslett HB, Medina J, Kliot D,Burkey AR: Mental motor imagery indexes pain: The hand laterality task. European Journal of Pain 14: 1007-1013, 2010.

[8]Daly AE, Bialocerkowski, AE: Does evidence support physiotherapy management of adult Complex Regional Pain Syndrome Type One? A systematic review. European Journal of Pain 13: 339-353, 2009.

[9]Physiotherapy Evidence Database. 1999. PEDro scale. http://www.pedro.org.au/english/downloads/pedro-scale/.

[10]Decety J: The neurophysiological basis of motor imagery. Behavioural Brain Research 77: 45-52, 1996.

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

[12]Flor H, Elbert T, Braun C,Birbaumer N: Extensive cortical reorganization in chronic back pain patients. NeuroImage 5: S216, 1997.

[13]Flor H, Elbert T, Knecht S, Wienbruch C, Pantev C, Birbaumer N, Larbig W,Taub E: Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. Nature 375: 482 - 484, 1995.

[14]Florence SL, Jain N,Kaas JH: Plasticity of Somatosensory Cortex in Primates. Seminars in Neuroscience 9: 3-12, 1997.

[15]Huse E, Preissl H, Larbig W,Birbaumer N: Phantom limb pain. The Lancet 358: 1015, 2001.

[16]Johnson S, Hall J, Barnett S, Draper M, Derbyshire G, Haynes L, Rooney C, Cameron H, Moseley GL, de C. Williams AC, McCabe C,Goebel A: Using graded motor imagery for complex regional pain syndrome in clinical practice: Failure to improve pain. European Journal of Pain 16: 550-561, 2012.

[17]Juottonen K, Gockel M, Silén T, Hurri H, Hari R,Forss N: Altered central sensorimotor processing in patients with complex regional pain syndrome. Pain 98: 315-323, 2002.

[18]Machado L, Kamper S, Herbert R, Maher C,McAuley J: Imperfect placebos are common in low back pain trials: a systematic review of the literature. European Spine Journal 17: 889-904, 2008.

[19]Maihofner C, Handwerker H, Neundorfer B,Birklein F: Patterns of cortical reorganization in complex regional pain syndrome. Neurology 61: 1707 - 1715, 2003.

[20]Matthys K, Smits M, Van der Geest JN, Van der Lugt A, Seurinck R, Stam HJ,Selles RW: Mirror-Induced Visual Illusion of Hand Movements: A Functional Magnetic Resonance Imaging Study. Archives of Physical Medicine and Rehabilitation 90: 675-681, 2009.

[21]Melzack R: Phantom limbs and the concept of a neuromatrix. Trends in Neurosciences 13: 88-92, 1990.

[22]Melzack R, Israel R, Lacroix R,Schultz G: Phantom limbs in people with congenital limb deficiency or amputation in early childhood. Brain 120: 1603-1620, 1997.

[23]Michielsen ME SR, van der Geest JN, Eckhardt M, Yavuzer G, Stam HJ, Smits M, Ribbers GM, Bussmann JBJ: Motor Recovery and Cortical Reorganization After Mirror Therapy in Chronic Stroke Patients. Neurorehabilitation and Neural Repair 25: 223-233, 2011.

[24]Moore RA, Eccleston C, Derry S, Wiffen P, Bell RF, Straube S,McQuay H: "Evidence" in chronic pain – establishing best practice in the reporting of systematic reviews. Pain 150: 386-389, 2010.

[25]Moseley G: Graded motor imagery for pathologic pain: a randomized controlled trial. Neurology 67:12: 2129 - 2134, 2006.

[26]Moseley G: Graded motor imagery is effective for long-standing complex regional pain syndrome: A randomised controlled trial. Pain 108: 192 - 198, 2004.

[27]Moseley G: I can't find it! Distorted body image and tactile dysfunction in patients with chronic back pain. Pain 140: 239-243, 2008.

[28]Moseley G: Is successful rehabilitation of complex regional pain syndrome due to sustained attention to the affected limb? A randomised clinical trial. Pain 114: 54 - 61, 2005.

[29]Moseley G, Gallagher, L, Gallace, A: Neglect-like tactile dysfunction in chronic back pain Neurology 79: 2012.

[30]Moseley GL: Why do people with complex regional pain syndrome take longer to recognize their affected hand? Neurology 62: 2182-2186, 2004.

[31]Moseley GL, Acerra N: Complex regional pain syndrome is associated with distorted body image of the affected part. Journal of the Neurological Sciences 238: S501-S501, 2005.

[32]Moseley GL,Flor H: Targeting Cortical Representations in the Treatment of Chronic Pain. Neurorehabilitation and Neural Repair 26: 646-652, 2012.

[33]Moseley GL, Gallace A,Spence C: Space-based, but not arm-based, shift in tactile processing in complex regional pain syndrome and its relationship to cooling of the affected limb. Brain 132: 3142-3151, 2009.

[34]Moseley GL,Wiech K: The effect of tactile discrimination training is enhanced when patients watch the reflected image of their unaffected limb during training. Pain 144: 314-319, 2009.

[35]Moseley GL, Zalucki N, Birklein F, Marinus J, van Hilten JJ,Luomajoki H: Thinking about movement hurts: The effect of motor imagery on pain and swelling in people with chronic arm pain. Arthritis Care & Research 59: 623-631, 2008.

[36]Moseley GL, Zalucki NM,Wiech K: Tactile discrimination, but not tactile stimulation alone, reduces chronic limb pain. Pain 137: 600-608, 2008.

[37]Parsons LM, Fox PT, Downs JH, Glass T, Hirsch TB, Martin CC, Jerabek PA,Lancaster JL: Use of implicit motor imagery for visual shape discrimination as revealed by PET. Nature 375: 54-58, 1995.

[38]Richter HO, R"ijezon U, Bj"rklund M,Djupsj"backa M: Long-term adaptation to neck/shoulder pain and perceptual performance in a hand laterality motor imagery test. Perception 39: 119-130, 2010.

[39]Schwoebel J, Coslett H, Bradt J, Friedman R,Dileo C: Pain and the body schema: effects of pain severity on mental representations of movement. Neurology 59: 775-778, 2002.

[40]Schwoebel J, Friedman R, Duda N,Coslett HB: Pain and the body schema: Evidence for peripheral effects on mental representations of movement. Brain 124: 2098-2104, 2001.

[41]Vingerhoets G, de Lange FP, Vandemaele P, Deblaere K, Achten E: Motor Imagery in Mental Rotation: An fMRI Study. NeuroImage 17: 1623-1633, 2002.

[42]Walz AD, Usichenko T, Moseley GL,Lotze M: Graded Motor Imagery and the Impact on Pain Processing in a Case of CRPS. Clin J Pain 2012.

[43]Wand BM, O'Connell NE, Di Pietro F,Bulsara M: Managing Chronic Nonspecific Low Back Pain With a Sensorimotor Retraining Approach: Exploratory Multiple-Baseline Study of 3 Participants. Physical Therapy 91: 535-546, 2011.

[44]Wand BM, Parkitny L, O'Connell NE, Luomajoki H, McAuley JH,Thacker M, Moseley, GL: Cortical changes in chronic low back pain: Current state of the art and implications for clinical practice. Manual Therapy 16: 15-20, 2011.

[45]Zacks J, Rypma, B, Gabrieli, JDE, Tversky, B, Glover, GH: Imagined transformations of bodies: an fMRI investigation. Neuropsychologia 37: 1029-1040, 1999.

Figure Legends

Figure 1.

The PRISMA flow-diagram describing the screening and review process.

Figure 2.

The pooled effect estimate for graded motor imagery versus usual care.

The effect sizes are calculated using the difference in post-intervention pain scores between control and intervention groups with the pooled post-intervention standard deviation. Positive effect sizes indicate a lower pain score in the intervention group, favouring the intervention group. Negative effect sizes indicate a lower pain score in the control group, favouring the control group

Footnotes: GMI = graded motor imagery; SD = standard deviation; CI = confidence interval; MI = motor imagery; LR = left/right judgements; MIRROR = mirror therapy.

Figure 3. The pooled effect estimate for left/right judgements versus usual care.

The effect sizes are calculated using the difference in post-intervention pain scores between control and intervention groups with the pooled post-intervention standard deviation. Positive effect sizes indicate a lower pain score in the intervention group, favouring the intervention group. Negative effect sizes indicate a lower pain score in the control group, favouring the control group

Footnotes: Left/right jud = left/right judgements; SD = standard deviation; CI = confidence interval.