OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

PEDIATRICS

Multilevel surgery for children with cerebral palsy: A meta-analysis

Journal:	*Pediatrics*
Manuscript ID	2018-3390.R1
Article Type:	Review Article
Date Submitted by the Author:	17-Dec-2018
Complete List of Authors:	Amirmudin, Noor Amirah; Royal College of Surgeons in Ireland , Department of Epidemiology and Public Health Medicine Lavelle, Grace; Brunel University London, Institute of Environment, Health and Societies Theologis, Tim; Oxford University Hospitals NHS Foundation Trust, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science Thompson, Nicky; Oxford University Hospitals NHS Foundation Trust, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science Ryan, Jennifer; Royal College of Surgeons in Ireland , Department of Epidemiology and Public Health Medicine; Brunel University London, Institute of Environment, Health and Societies
Keyword/Topic:	Neurologic Disorders < Neurology, Orthopedic Medicine

Multilevel surgery for children with cerebral palsy: A meta-analysis

Noor Amirah Amirmudin,[a] Grace Lavelle PhD,[b] Tim Theologis PhD,[c] Nicky Thompson PhD,[c] Jennifer M. Ryan PhD[a,b]

[a]Department of Epidemiology and Public Health Medicine, RCSI Dublin, Ireland
[b]Department of Clinical Sciences, Brunel University London, UK
[c]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science, Oxford University Hospitals NHS Foundation Trust, UK

Corresponding author: Jennifer M. Ryan, Population Health Sciences, Beaux Lane House, Lower Mercer Street, Dublin 2, Ireland; telephone: +35314022413; email: jenniferryan@rcsi.com

Short title: Multilevel Surgery for Children with CP.

Financial disclosure statement: The authors have no financial relationship relevant to this article to disclose.

Funding source: Funding to conduct this review was received from RCSI's Research Summer School programme.

Conflict of interest statement: The authors have no conflicts of interest relevant to this article to disclose.

Protocol Registration: This protocol for this systematic review is registered with PROSPERO: CRD42018103131

Abbreviations: CP: Cerebral palsy; ICF: International Classification of Functioning, Disability and Health; ICIDH: International Classification of Impairments, Disabilities, and Handicaps, MLS: multi-level surgery; RCT: randomised controlled trial

Table of Contents Summary: This review summarises the literature on multilevel surgery for children with cerebral palsy, including the effects of and satisfaction with outcomes following multilevel surgery.

Contributors' Statement page: Noor Amirah Amirmudin designed the search strategy, selected the articles, extracted the data, assessed the quality of the evidence, performed the data analysis, and drafted the initial manuscript. Jennifer M. Ryan conceptualised and designed the review, co-ordinated and supervised the review, designed the search strategy, selected the articles, extracted the data, assessed the quality of the evidence, performed the data analysis, interpreted the data analysis, drafted the initial manuscript, reviewed and revised the manuscript. Grace Lavelle conceptualised and designed the review, co-ordinated and supervised the review, designed the search strategy, interpreted the data, reviewed and revised the manuscript. Tim Theologis and Nicky Thompson conceptualised and designed the review, interpreted the data, reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Abstract

Context: Multilevel surgery (MLS) is standard care for reducing musculoskeletal disorders among children with spastic cerebral palsy (CP).

Objective: To summarise the literature on MLS for children with CP, including the effects of MLS and satisfaction with outcomes following MLS.

Data Sources: MEDLINE, Embase, CINAHL and Cochrane Central Register of Controlled Trials were searched.

Study Selection: Studies reporting effects of or satisfaction with MLS in children with CP.

Data Extraction: Two authors screened and extracted data on gross motor function, gait speed, summary statistics of gait (e.g. Gait Profile Score [GPS]), range of motion, strength, spasticity, participation, quality of life, satisfaction, and adverse events.

Results: 74 studies (3551 participants) were identified; one was a randomised controlled trial (RCT; n=19) and the remainder were cohort studies. Pooled analysis of cohort studies showed that MLS did not have a long-term effect on gross motor function (SMD: 0.38, 95% CI -0.25 to 1.01, p=0.24), or gait speed (SMD: 0.12, 95% CI -0.01 to 0.25, p=0.08), but did improve gait (SMD: -0.80, 95% CI -0.95 to -0.65, p<0.001). The RCT also found no effect of MLS on gross motor function but improvements in the Gait Profile Score at 1 year. Only five studies reported participation and quality of life and seventeen adequately reported adverse events.

Limitations: Data were largely from cohort studies. There was moderate heterogeneity in meta-analyses.

Conclusions: Findings suggest that gait but not gross motor function improves following MLS. RCTs and improved reporting of studies of MLS are required.

Cerebral palsy (CP) is characterised by abnormal fine and gross motor functioning.[1] The incidence of CP is approximately 2-3 per 1000 live births worldwide.[2,3] Musculoskeletal disorders, including contractures of muscle tendon units and bony deformities, are a secondary impairment of CP, which contribute to restricted mobility.[1] Multilevel surgery (MLS) followed by intensive rehabilitation is considered standard care for reducing musculoskeletal disorders among children with spastic CP.[4]

To date, two systematic reviews have examined the literature on MLS for children with CP. The first, conducted in 2010, included 31 studies and concluded that there was a trend towards improvements in passive range of motion, gait kinematics and kinetics, and gait efficiency, but little evidence for improvements in gross motor function or quality of life (QoL).[5] However, study quality was variable and a meta-analysis of data was not performed. The authors also identified variability in the quality of reporting of surgical procedures, rehabilitation, and adverse events. The second review examined the effect of MLS specifically on gait parameters.[6] Authors concluded from a narrative synthesis that there was a trend towards improvements in gait parameters, but reported variability in study quality.

The World Health Organisation's International Classification of Functioning, Disability and Health (ICF) framework is useful for considering the impact of CP on the individual.[7] Developed in 2001, the ICF classifies health-related domains using the terms "body functions and structure", "activity" and "participation". The ICF replaced the International Classification of Impairments, Disabilities, and Handicaps (ICIDH), which conceptualised a health condition leading to impairment, disability and handicap in a linear manner. The ICF places greater emphasis on the role of the social and physical environment on functioning, and conceptualises that functioning results from a complex interaction between the person with the health condition and their context (consisting of personal and environmental factors). When developed, it was envisaged that one use of the ICF would be the evaluation of interventions. However, the review in 2010 found that few studies evaluate the effect of MLS across multiple domains of the ICF.

Given the impact of MLS on children with CP and their families,[8] a summary of the current literature on MLS and systematic evaluation of the effects of MLS is required. Thus, the aim of this review was to summarise the current literature on MLS for children with CP, including the effects of MLS and satisfaction with outcomes following MLS. We also aimed to examine the use of the ICF and ICIDH by authors when reporting outcomes of MLS.

Method

Study selection criteria

We included studies meeting the following criteria: 1. included children with CP. We chose to define a child as a person aged 0-20 yr, in order to be inclusive of all studies evaluating MLS in children as we expected variation in the age range across studies; 2. reported outcomes before and after MLS or reported satisfaction with MLS; 3. additional interventions, such as Botulinum toxin injections, were not performed simultaneously to surgery. We defined MLS as "two or more soft-tissue or bony surgical procedures at two or more anatomical levels".[9] Original peer-reviewed articles published in English were included. To provide a comprehensive overview of the current literature we did not limit studies to randomised controlled trials (RCTs). However, qualitative studies, reviews, commentaries, conference abstracts, and case reports were excluded.

Search strategy

We searched Ovid MEDLINE, Embase Ovid, CINAHL and Cochrane Central Register of Controlled Trials (CENTRAL) from January 2000 to June 2018. Search terms relating to children, CP and surgery were combined. Each search was adapted for the respective database (see Appendix). Additionally, reference lists of previous reviews were searched.

Two investigators screened titles and abstracts independently against eligibility criteria. Where articles met eligibility criteria or there was doubt over inclusion, full texts were obtained. Any disagreement regarding the inclusion of an article was resolved through discussion.

Data extraction

Two authors (NAA, JMR) extracted data independently using a standardised piloted data extraction form. Data on study design, study population (e.g. Gross Motor Function Classification System [GMFCS] level), intervention (e.g. procedures performed), rehabilitation received, duration of follow-up, adverse events, and outcomes were extracted. We noted if studies reported outcomes according to ICIDH or ICF frameworks.

We extracted data on gross motor function as measured by the Gross Motor Function Measure (GMFM-66 or GMFM-88), the Functional Mobility Scale (FMS), and the Gillette Functional Assessment Questionnaire (FAQ). Although some studies assessed gross motor function using the GMFCS we did not consider the GMFCS a measure of gross motor function as it is primarily a classification system. We extracted data on gait speed and summary statistics of gait i.e., the Gait

Deviation Index (GDI), Gait Profile Score (GPS), or Gillette Gait Index (GGI).[10-12] Where studies reported more than one summary score we extracted data on the GPS. We chose to extract data on gait summary scores rather than individual kinematic or kinetic variables as many studies reported a large number of variables obtained from three-dimensional gait analysis. Further, we extracted data on participation, QoL and satisfaction with surgery. Since the publication of the ICF there has been lack of consensus regarding measurement tools to assess participation. For the purpose of this review, we extracted data on measures that mapped to the recently developed family of participation-related constructs.[13] Finally, we extracted data on passive range of motion, muscle strength, and spasticity, as these are commonly evaluated before and after MLS as part of a clinical examination and inform clinical decision making on surgery. We extracted data on passive knee extension and passive dorsiflexion only as these were the most commonly reported joints. As authors typically reported muscle strength and spasticity for more than one muscle group, we ranked muscle groups in order of frequency of reporting across all included studies. For each study, we then extracted data for strength and spasticity, respectively, for the most frequently reported muscle group. If data were not reported for the most frequently reported muscle group we extracted data for the next most frequently reported muscle group, and so on. If data were reported on both limbs, data on the most affected limb or the right limb were extracted.

Data on short-term (≤6 months post-surgery), intermediate-term (6 months to 1 year post-surgery), and long-term (>1 year post-surgery) effects were extracted. If more than one assessment was performed during any of these time-periods, the last measurement taken during that period was extracted. Where studies compared two groups that received MLS, we extracted pre- and post-surgery data for the whole sample, if available, or pre- and post-surgery data for the two groups separately. In these cases, we report results according to number of groups included in meta-analysis.

Quality assessment

The Methodological Index for Non-Randomised Studies (MINORS) checklist was used to assess the quality of studies based on 8 or 12 items.[14] Items related to the description of the aim, selection, prospective data collection and sample size calculation, appropriateness of follow-up period, and loss to follow-up. In addition, the adequacy of reporting of surgical procedures, previous surgery, adverse events, and rehabilitation was rated as 0 (not reported), 1 (reported but inadequate), or 2 (reported adequately).

Data analysis

Where sufficient data were obtained, we conducted separate meta-analyses for each outcome at each time-point.

Weighted effect sizes were calculated using the inverse-variance method. A random effects model was used as there was large variation between studies in terms of procedures, populations, and context.[15] Effect sizes were calculated as Hedges' g^* to correct for bias associated with small sample size.[16] The standard deviation for the change score (i.e. mean difference [MD] between baseline and follow-up values) (SD_{change}) was used to calculate g^*. Where this was not provided SD_{change} was estimated using statistics provided (e.g. t values, confidence intervals [CI], standard errors, p values) or standard deviations for the baseline and follow-up score ($SD_{baseline}$ and SD_{fu}, respectively), and the correlation between baseline and follow-up scores (r). Where r was not provided, we estimated r using either individual data or using $SD_{baseline}$, SD_{fu} and SD_{change}. If this was not possible we imputed r as 0.7 for the analysis of short-term effect on gross motor function and imputed r as 0.3 for all other analyses, based on the strength of correlations observed in studies with similar samples, outcome measures, and follow-up periods.[17] However, sensitivity analyses were performed by imputing r as 0.1, 0.3 and 0.7 to examine the robustness of the findings to this assumption.[17] Finally, the standard error of g^* was calculated using the provided, estimated or imputed r. We conducted a subgroup analysis of studies that included children with bilateral CP only. We assessed statistical heterogeneity and its impact using the Chi^2 test and the I^2 statistic.[18] To aid interpretation we proposed an effect size of 0.2, 0.5 and 0.8 represents a small, moderate, and large effect, respectively,[19] although this was used cautiously.

Results

Database searches identified 2364 articles; 2 articles were obtained from reviewing reference lists (Figure 1). After removal of duplicates, 1,823 titles and abstracts were screened. 164 full text articles were obtained. Of these, 75 articles reporting findings from 74 studies were included. A description of each study is presented in Table 1. MINORS score ranged from 4[20] to 21[9] (supplemental table).

<u>Study Design</u>

We identified one RCT comparing MLS to progressive resistance training at 1-year.[9] This study also reported outcomes for the MLS group at 2-years. Fifty-two studies were retrospective cohort studies. Where these studies included a comparison group they compared results between males and females,[21] participants with diplegia and hemiplegia,[22] participants with more or less than a

5° increase in anterior pelvic tilt,[23] or between patients who received different surgical procedures as part of MLS.[20,24-31] Seventeen studies were prospective cohort studies, of which one used a single-subject AB design[32] and two compared two surgical procedures as part of MLS.[33,34] Four studies were RCTs comparing two surgical procedures as part of MLS[35,36] or comparing two types of rehabilitation following MLS.[37-39] We treated these as cohort studies in pooled analyses as they were not comparisons of MLS to control.

Participants

In total, 3551 participants were included in 74 studies. Sample sizes ranged from 7[40] to 314.[41] We considered 58 studies (78%) were small (n<50), 8 medium sized (n=50-100), and 8 large (n>100). Participants were 3 to 20 years at time of surgery (mean age 6.0yr[42] to 14.0yr[26]). Where sex was reported, the percentage of males ranged from 20%[43] to 100%;[44] overall 62% were male.

Where type was reported (76% of included studies), all participants had spastic CP. Eighty-six percent of studies reported anatomical distribution. Two studies included children with unilateral CP only.[45,46] Twelve studies included children with unilateral and bilateral CP. The remaining studies included children with bilateral CP only.

Forty-seven studies included children in GMFCS levels I-III. One included children in level IV only[46] and the remainder included participants in levels I-IV, I-V or II-IV.[28,30,42-44,47-50] Seventeen studies did not report GMFCS level; all children were described as ambulatory with or without a walking aid except for Khan,[51] who reported all participants were non-ambulatory.

Description of surgery and rehabilitation

Description of surgery was rated adequate in 48 studies (65%). Soft tissue procedures only were performed in seven studies.[42,50,52-56] Bony and soft tissue procedures were performed in 64 studies. Three studies did not state the type of procedures performed.[57-59] Where reported, the number of procedures conducted per participant ranged from 2[54] to 18;[60] the mean per patient was 2.2[25] to 12.9.[34] Provision of previous surgery was adequately reported in thirty-three studies.

Nineteen studies (26%) did not report any information regarding rehabilitation provided following surgery. Twenty-five studies (34%) provided inadequate description of rehabilitation. Thirty studies (40%) reported the frequency and/or duration of rehabilitation provided. Where reported, the content of rehabilitation included active and passive movements, strengthening exercises, stretching, balance, and gait training.

Outcome assessment

The duration of follow-up varied widely between studies. Where described, the mean duration of follow-up was 1.0 years[30] to 21.3 years.[61] We extracted data on short-term outcomes from three studies,[39,40,50] on intermediate-term outcomes from 26 studies, and on long-term outcomes from 45 studies. No study described outcomes in terms of the ICIDH framework. Seven studies described outcomes in terms of the ICF framework but there was inconsistency in the measures used to assess each domain.[9,32,40,59,62-64] Authors reported assessing the collective domain of "activity and participation" using the FMS,[59,63] GMFM-66,[9,64] Pediatric Evaluation of Disability Inventory (PEDI),[40] time spent in upright positions measured with an activity logger,[9] and self-reported walking ability.[32]

Fifteen studies assessed gross motor function; ten studies used the GMFM-66 or GMFM-88,[9,25,28,32,33,38-40,50,64] three studies used the FAQ,[30,52,65] and two studies used the FMS.[59,63] Twenty-nine studies assessed gait speed and thirty-five assessed gait using a summary score. Four studies assessed participation,[40,52,58,65] three assessed QoL,[9,32,52] and four reported satisfaction with surgery.[32,65-67] Twenty-two studies assessed passive range of motion (knee extension [n=15] and ankle dorsiflexion [n=17]), fourteen assessed muscle strength, and six assessed spasticity.[23,35,38,50,68,69]

Seventeen studies (23%) adequately reported adverse events. These included intra-operative haemorrhage, poor post-operative pain management, hardware removal for pain, superficial wound infection, nerve palsy, avascular necrosis, permanent reflex sympathetic dystrophy syndrome, and occurence of genu recurvatum. Three studies reported adverse events according to the Clavien-Dindo classification.[35,45,70] In these studies, 68 Grade I, 52 Grade II, nine Grade III, and one Grade IV complication were reported among 265 children.

Effects of multilevel surgery

Data from 54 studies were included in meta-analysis; data from one RCT are reported separately.[64] Results from a further 8 studies are reported descriptively as insufficient information was provided to include them in meta-analysis.[52,59,63,65-67,71,72] Data on 12 studies are not reported as eight studies assessed individual gait variables or GMFCS level,[20,29,42,47,49,55,73,74] and four studies did not report results from statistical analyses to allow interpretation of findings.[32,51,58,60]

Gross motor function

One RCT of 19 children, found that gross motor function as measured by the GMFM-66 did not differ between children receiving MLS and children receiving resistance training at 1-year post-surgery (MD between groups: 0.3, 95% CI -4.5 to 5.0). However, at 2-years post-surgery the children receiving MLS demonstrated an improvement in GMFM-66 from baseline (MD: 4.9, 95% CI 0.98 to 8.7). Pooled analysis indicated that gross motor function did not improve in the short-term (standardised mean difference [SMD]: 0.01, 95% CI -0.48 to 0.50, p=0.97, I^2=92%, p<0.001, n=52), intermediate-term (SMD: 0.51, 95% CI -0.56 to 1.58, p=0.35, I^2=97%, p<0.001, n=104), or long-term (SMD: 0.38, 95% CI -0.25 to 1.01, p=0.24, n=103; Figure 2). Two studies (n=118) reported that FAQ improved by a mean of 1.06 (<0.001) and 0.5 (p=0.002), respectively,[52,65] and one study (n=75) reported no change in FAQ (p>0.05),[30] in the intermediate-term.

One study (n=156) reported that 5 years after MLS, 23%, 18% and 19% of participants required less assistance at 5m, 50m and 500m, respectively.[63] A second study (n=66) found evidence of a poorer rating on the FMS at 6 months compared to pre-surgery at all distances.[59] This study found no evidence of change in FMS at 12 months, but evidence of an improvement in rating at 24 months at all distances.

Gait speed

One study assessed the short-term change in normalised walking speed following surgery and found it declined (MD: -0.04, p=0.002, n=20).[39] There was no evidence that walking speed changed in the intermediate-term (SMD: -0.08, 95% CI -0.22 to 0.07, p=0.30, I^2=43%, p=0.02, n=481) or long-term (SMD: 0.12, 95% CI -0.01 to 0.25, p=0.08; Figure 3).

Gait

GPS improved in children who received MLS compared to those who received resistance training at 1-year (MD between groups: -5.5, 95% CI -7.8 to -3.4, n=19). Pooled analysis indicated that gait improved at 1-year following MLS (SMD: -0.78, 95% CI -0.98 to -0.57, p<0.001, I^2=70%, p<0.001, n=497). There was also improvement in gait in the long-term (SMD: -0.80, 95% CI -0.95 to -0.65, p<0.001; Figure 4). One study, not included in pooled analysis, reported improvements in GDI in the intermediate-term (MD: 12.1, p<0.001, n=39) and long-term (MD: 10.3, p<0.001, n=39).[72]

Passive range of motion

Passive ankle dorsiflexion was greater among children who received MLS compared to those who received resistance training at 1-year (MD between groups=7, 95% CI 3 to 12, p<0.001, n=19). Knee extension declined in both groups (MD: -1 degrees for MLS group and MD: -3 degrees for control group) but a comparison between groups was not conducted.

There was evidence of intermediate-term improvements in knee extension (SMD: 0.60, 95% CI 0.43 to 0.77, p<0.001, I^2=0%, p=0.58, n=228) and dorsiflexion (SMD: 0.48, 95% CI 0.10 to 0.86, p=0.01, I^2=49%, p=0.12, n=96) following MLS. Long-term improvements in knee extension (SMD: 0.43, 95% CI 0.24 to 0.61, p<0.001; Figure 5) and dorsiflexion (SMD: 0.47, 95% CI 0.17 to 0.76, p=0.002, I^2=76%, p<0.001, n=324) were also observed. One study, not included in pooled analysis, reported an improvement in dorsiflexion in the intermediate-term (p<0.001, n=19).[71]

Strength

Plantar flexor strength improved by a larger amount in children following MLS compared to resistance training at 1-year (MD between groups: 1.9 kg, 95% CI 0.01 kg to 3.9 kg; n=19).[9] Although quadriceps strength, hip extensor strength, and hip abductor strength were also measured in the RCT, comparisons between groups were not conducted.

One study (n=20) found evidence of a decline in strength of the hip flexors (MD: -0.2 Nm/kg, $p<0.001$), hip extensors (MD: -0.5 Nm/kg, $p<0.001$), hip abductors (MD: -0.1 Nm/kg, p=0.002), knee flexors (MD: -0.6 Nm/kg, $p<0.001$), and knee extensors (MD: -0.5 Nm/kg, $p<0.001$) at 6 months post-surgery.[39] The second study reported an increase in general muscle strength at 6 months post-surgery (median 2 vs 3 on a scale of 0-4, p=0.020; n=25).[50] Pooled analysis found no evidence for change in muscle strength in the intermediate-term (SMD:-0.22, 95% CI -0.50 to 0.06, p=0.12, I^2=51%, p=0.08, n=117). We were unable to pool data from two studies; one reported no change[75] and the second reported a decline in muscle strength.[69]

Pooled analysis found no evidence of change in muscle strength in the long-term (SMD: -0.39, 95% CI -0.79 to 0, p=0.05, I^2=74%, $p<0.001$, n=165). We were unable to include three studies in pooled analysis; two found no change[75,76] and one reported an improvement ($p<0.05$).[69]

Spasticity

All studies that assessed short-term (n=25)[50] and intermediate-term (n=145)[35,38,68,69] changes in spasticity reported reduced spasticity ($p<0.05$). Two studies reported reductions and one reported no change in spasticity in the long-term (n=118).[23,68,69]

Participation and quality of life

Evidence for improvement in participation was observed at 6-months,[40] 1-year,[40,52,65] and 2-years[40] following MLS (Table 2). Intermediate and long-term improvements in QoL were also observed (Table 2).

Satisfaction

Satisfaction with the outcome of surgery was rated as mean (SD) 7.9 (2.0) on a scale of 0-10 (with 10 being completely satisfied) by 279 parents.[66] Sixty-one parents rated satisfaction with functional and cosmetic outcomes as mean (SD) 6.8 (2.0) and 6.6 (2.0) out of 10, respectively.[65] Of 11 parents, 91% reported that they were satisfied with surgical outcomes.[67] When 11 children were asked if the "results from SEMLS and rehabilitation were worth the effect" 9 and 10, respectively, said yes at 1- and 2-years post-MLS.[32]

<u>Sensitivity analyses</u>

Imputing the correlation co-efficient as 0.7 instead of 0.3 resulted in a change in conclusions for the change in gait speed in the long-term (SMD: 0.17, 95% CI 0.04 to 0.30, $p=0.01$, $I^2=78\%$, $p<0.001$), and strength in the long-term (SMD: -0.58, 95% CI -1. 0 o -0.15, $p=0.008$, $I^2=91\%$, $p<0.001$).

<u>Subgroup analysis</u>

When children with bilateral CP only were included in pooled analyses, there was evidence that muscle strength declined in the intermediate-term (SMD: -0.35, 95% CI -0.64 to -0.07, $p=0.02$,

I^2=9%, p=0.33, n=40) and long-term (SMD: -0.49, 95% CI -0.96 to -0.02, p=0.04, I^2=75%, p<0.001, n=134). All other analyses of children with bilateral CP only produced similar effect sizes and identical inference to that obtained from the primary analysis.

Discussion

In summary, we found no evidence from a meta-analysis that gross motor function improved following MLS, but some evidence of improvements in gait and passive range of motion. It should be noted however, that there was significant heterogeneity associated with pooled analysis of gross motor function and inconsistent findings across studies; while many studies found no evidence of change in gross motor function following MLS some found evidence of improvement. Conversely, there was some evidence that muscle strength declined in the 6 months following MLS and was not different to pre-surgery levels at 1-year following MLS. Although only three and two studies assessed participation and QoL, respectively, there was some evidence that participation and QoL improved following MLS. Overall, parents reported being satisfied with outcomes of MLS, although there was considerable variation in satisfaction ratings. Adverse events following MLS appear relatively rare; only one adverse event that resulted in a "life-threatening complication (including CNS complications) requiring IC/ICU management" (Grade IV complication according to the Clavien-Dindo classification[77]) was reported. However, adverse event reporting was incomplete, with only 23% of studies adequately reporting adverse events.

Our findings support two previous reviews of MLS for CP.[5,6] Both concluded that gait kinematics improved following MLS, although neither conducted a quantitative synthesis of data. With a similar search strategy to that used in the current review, McGinley identified 31 studies that examined the effect of MLS for children with CP.[5] Despite including twice as many studies in the current review, we identified similar issues with the evidence base to those identified in 2010. There was consistently inadequate reporting of participant characteristics, surgical interventions, previous surgery, and rehabilitation. MLS is a complex intervention, consisting of interacting components, which are particularly difficult to evaluate. For example there are difficulties standardising the design and delivery of MLS and rehabilitation, and the causal association linking the intervention and outcome is lengthy, complex and likely context dependent.[78] The results of this review suggest that a pattern of initial deterioration in gross motor function occurs in the year after MLS, followed by a return to baseline at 1-2 years post-MLS, and potentially further improvements in the longer-term. However, it is difficult to comment on the long-term impact of MLS on gross motor function given the relatively small number of patients who were assessed beyond 1 year after surgery. This complex and lengthy interaction between the intervention and outcome indicate that the content of MLS and rehabilitation are likely crucial to explaining effectiveness or otherwise. Improved reporting of these factors is therefore essential to improving the evidence base for MLS. Future studies should use reporting guidelines such as the CONSORT statement[79] and the TIDiER checklist[80] to ensure that the study and the intervention are described in sufficient detail to appraise and replicate.

A lack of RCTs is a significant limitation to the evidence base. We identified just one RCT of 19 children. All other evidence was from cohort studies, with many being retrospective reviews of clinical records. As a result, studies often excluded patients without routine follow-up assessments suggesting that the findings are subject to significant selection bias. While effect estimates may be biased as a result of excluding participants from the analysis,[81] the extent of attrition bias is unknown as the majority of studies did not report the proportion of eligible participants included in analyses. Further, few prospective cohort studies reported loss to follow-up. In addition, establishing causality is difficult without a comparison group to control for the potential impact of known and unknown confounding variables on the outcome. It should be noted however, that findings from the RCT largely align with pooled analyses of data from cohort studies. In particular, the mean difference of -5.5 in GPS score, observed between the intervention and control group at 1-year,[9] was very similar to the improvement in GPS observed in cohort studies (-3.7 to -7.1).[35,45,57,64]

The pilot RCT demonstrates that it is feasible to conduct RCTs of MLS. However, researchers still face several barriers to conducting well-designed RCTs. These include ethical concerns regarding allowing people to experience progression of musculoskeletal deformities without intervening, difficulty recruiting patients given the relatively small number of children that undergo MLS annually, and difficulties obtaining funding to conduct a multi-centre trial with sufficient duration of follow-up to evaluate the impact of MLS. Further, when designing future studies, careful consideration should be given to the comparator in order to overcome ethical concerns and provide clinically meaningful results. Although the current review suggests that MLS improves gait, it does not establish if MLS is the most appropriate approach to improving gait.

We found very few studies that assessed outcomes according to ICF domains, despite all but three studies being published after the introduction of the ICF in 2001. The lack of prospective studies may explain why so few studies assessed outcomes across ICF domains. Lack of consensus regarding how to assess each ICF domain is also a barrier to using this framework. Regardless of whether or not a framework was used to evaluate MLS, it is concerning that only 10 studies assessed activity using the GMFM, which is considered the criterion measure of gross motor function. Additionally only four studies assessed participation and three studies assessed QoL. More studies assessed passive range of motion, strength and spasticity, even though these are arguably less meaningful outcomes to participants. Indeed, passive range of motion and spasticity are associated with no change or very small changes in gross motor function.[82] In order to provide comprehensive information on the effects of MLS to families and professionals, future studies need to prospectively plan to collect data on a range of outcomes. While the ICF provides a standardised framework for the evaluation of interventions, it may be unrealistic to expect MLS to improve activity and participation, given that activity and participation result from a complex interaction between the individual and the environment. However, even if MLS does not improve activity or participation, it may prevent further activity limitations and participation restrictions by preventing deterioration in musculoskeletal deformities.

The lack of prospective data collection may explain why approximately 70% of studies did not report information on adverse events. This review can therefore not make conclusions regarding the safety of MLS. Standardised recording and reporting of adverse events should be implemented in future studies to ensure consistent and deliberative reporting of safety. Additional limitations of the evidence base are that all included studies are at high risk of bias for blinding of outcome assessment and the majority had small sample sizes. Lack of blinding may exaggerate the effect of MLS, particularly when outcomes are assessed using self-report measures, such as participation and QoL.[83] Sample sizes may have impacted results by either inflating effect sizes or resulting in statistically non-significant findings.[84,85]

There are also limitations to this review. We did not include grey literature or studies published in any language other than English. We included studies that reported outcomes before and after MLS even if the primary aim of the study was not to assess the effect of MLS. In order to include some data in meta-analyses we imputed correlation coefficients. However, sensitivity analyses, demonstrated that our findings were generally robust for different imputed values. There was evidence of at least moderate heterogeneity in all models, which may be explained by the large amount of clinical and methodological diversity in included studies.

Conclusions

This review represents the most comprehensive summary of the evidence on MLS for children with CP to date. Findings suggest that MLS is associated with changes in gait but not gross motor function in children with CP. However, the review identified considerable limitations to the evidence base that need to be addressed in future trials. Specifically, there is a need for RCTs and improvements in reporting of trials.

References

1. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl. 2007;109:8-14.

2. Kirby RS, Wingate MS, Van Naarden Braun K, Doernberg NS, Arneson CL, Benedict RE, et al. Prevalence and functioning of children with cerebral palsy in four areas of the United States in 2006: a report from the Autism and Developmental Disabilities Monitoring Network. Res Dev Disabil. 2011;32(2):462-9.

3. Surveillance of Cerebral Palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). Dev Med Child Neurol. 2000;42(12):816-24.

4. Graham HK, Rosenbaum P, Paneth N, Dan B, Lin JP, Damiano DL, et al. Cerebral palsy. Nat Rev Dis Primers. 2016;2:15082.

5. McGinley JL, Dobson F, Ganeshalingam R, Shore BJ, Rutz E, Graham HK. Single-event multilevel surgery for children with cerebral palsy: a systematic review. Dev Med Child Neurol. 2012;54(2):117-28.

6. Lamberts RP, Burger M, du Toit J, Langerak NG. A Systematic Review of the Effects of Single-Event Multilevel Surgery on Gait Parameters in Children with Spastic Cerebral Palsy. PloS One. 2016;11(10):e0164686.

7. World Health Organization. Towards a common language for functioning, disability and health: ICF. Geneva: 2002.

8. Iversen AS, Graue M, Clare J. Parents' perspectives of surgery for a child who has cerebral palsy. J Pediatr health Care 2009;23(3):165-72.

9. Thomason P, Baker R, Dodd K, Taylor N, Selber P, Wolfe R, et al. Single-event multilevel surgery in children with spastic diplegia: a pilot randomized controlled trial. J Bone Joint Surg Am. 2011;93(5):451-60.

10. Schutte LM, Narayanan U, Stout JL, Selber P, Gage JR, Schwartz MH. An index for quantifying deviations from normal gait. Gait Posture. 2000;11(1):25-31.

11. Baker R, McGinley JL, Schwartz MH, Beynon S, Rozumalski A, Graham HK, et al. The gait profile score and movement analysis profile. Gait Posture. 2009;30(3):265-9.

12. Schwartz MH, Rozumalski A. The Gait Deviation Index: a new comprehensive index of gait pathology. Gait Posture. 2008;28(3):351-7.

13. Adair B, Ullenhag A, Rosenbaum P, Granlund M, Keen D, Imms C. Measures used to quantify participation in childhood disability and their alignment with the family of participation-related constructs: a systematic review. Dev Med Child Neurol. 2018;60(11):1101-16.

14. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. ANZ J Surg. 2003;73(9):712-6.

15. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88.

16. Cumming G. Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis. New York: Routledge; 2012.

17. Borenstein M, Hedges, LV, Higgins JPT, Rothstein HR. Effect sizes based on means. In Introduction to Meta-Analysis. Chichester: John Wiley & Sons; 2009.

18. Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0: The Cochrane Collaboration; 2011.
19. Cohen J. Statistical Power Analysis in the Behavioural Sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associations, Inc; 1988.
20. Blumetti FC, Morais Filho MC, Kawamura CM, Cardoso MO, Neves DL, Fujino MH, et al. Does the GMFCS level influence the improvement in knee range of motion after rectus femoris transfer in cerebral palsy? J Pediatr Orthop B. 2015;24(5):433-9.
21. Zwick EB, Svehlik M, Kraus T, Steinwender G, Linhart WE. Does gender influence the long-term outcome of single-event multilevel surgery in spastic cerebral palsy? J Pediatr Orthop B. 2012;21(5):448-51.
22. Saraph V, Zwick EB, Zwick G, Dreier M, Steinwender G, Linhart W. Effect of derotation osteotomy of the femur on hip and pelvis rotations in hemiplegic and diplegic children. J Pediatr Orthop B. 2002;11(2):159-66.
23. Bohm H, Hosl M, Doderlein L. Predictors for anterior pelvic tilt following surgical correction of flexed knee gait including patellar tendon shortening in children with cerebral palsy. Gait Posture. 2017;54:8-14.
24. Desailly E, Thevenin-Lemoine C, Khouri N. Does patella lowering improve crouch gait in cerebral palsy? Comparative retrospective study. Orthop Traumatol Surg Res. 2017;103(5):741-6.
25. Feger MA, Lunsford CD, Sauer LD, Novicoff W, Abel MF. Comparative effects of multilevel muscle tendon surgery, osteotomies, and dorsal rhizotomy on functional and gait outcome measures for children with cerebral palsy. PM R. 2015;7(5):485-93.
26. Laracca E, Stewart C, Postans N, Roberts A. The effects of surgical lengthening of hamstring muscles in children with cerebral palsy - the consequences of pre-operative muscle length measurement. Gait Posture. 2014;39(3):847-51.
27. Mallet C, Simon AL, Ilharreborde B, Presedo A, Mazda K, Pennecot GF. Intramuscular psoas lengthening during single-event multi-level surgery fails to improve hip dynamics in children with spastic diplegia. Orthop Traumatol Surg Res. 2016;102(4):501-6.
28. Taylor D, Connor J, Church C, Lennon N, Henley J, Niiler T, et al. The effectiveness of posterior knee capsulotomies and knee extension osteotomies in crouched gait in children with cerebral palsy. J Pediatr Orthop B. 2016;25(6):543-50.
29. Kwon DG, Lee SY, Kim TW, Chung CY, Lee KM, Sung KH, et al. Short-term effects of proximal femoral derotation osteotomy on kinematics in ambulatory patients with spastic diplegia. J Pediatr Orthop B. 2013;22(3):189-94.
30. Truong WH, Rozumalski A, Novacheck TF, Beattie C, Schwartz MH. Evaluation of conventional selection criteria for psoas lengthening for individuals with cerebral palsy: a retrospective, case-controlled study. J Pediatr Orthpo. 2011;31(5):534-40.
31. Dreher T, Wolf SI, Maier M, Hagmann S, Vegvari D, Gantz S, et al. Long-term results after distal rectus femoris transfer as a part of multilevel surgery for the correction of stiff-knee gait in spastic diplegic cerebral palsy. J Bone Joint Surg Am. 2012;94(19):e142(1-10).
32. Akerstedt A, Risto O, Odman P, Oberg B. Evaluation of single event multilevel surgery and rehabilitation in children and youth with cerebral palsy - A 2-year follow-up study. Disabil Rehab. 2010;32(7):530-9.

33. Thompson N, Stebbins J, Seniorou M, Wainwright AM, Newham DJ, Theologis TN. The use of minimally invasive techniques in multi-level surgery for children with cerebral palsy: preliminary results. J Bone Joint Surg Br. 2010;92(10):1442-8.
34. Dreher T, Vegvari D, Wolf SL, Klotz M, Muller S, Metaxiotis D, et al. Long-term effects after conversion of biarticular to monoarticular muscles compared with musculotendinous lengthening in children with spastic diplegia. Gait Posture. 2013;37(3):430-5.
35. Klotz MCM, Krautwurst BK, Hirsch K, Niklasch M, Maier MW, Wolf SI, et al. Does additional patella tendon shortening influence the effects of multilevel surgery to correct flexed knee gait in cerebral palsy: A randomized controlled trial. Gait Posture. 2018;60:217-24.
36. Dreher T, Gotze M, Wolf SI, Hagmann S, Heitzmann D, Gantz S, et al. Distal rectus femoris transfer as part of multilevel surgery in children with spastic diplegia--a randomized clinical trial. Gait Posture 2012;36(2):212-8.
37. Patikas D, Wolf SI, Armbrust P, Mund K, Schuster W, Dreher T, et al. Effects of a postoperative resistive exercise program on the knee extension and flexion torque in children with cerebral palsy: a randomized clinical trial. Arch Phys Med Rehabil. 2006;87(9):1161-9.
38. Patikas D, Wolf SI, Mund K, Armbrust P, Schuster W, Doderlein L. Effects of a postoperative strength-training program on the walking ability of children with cerebral palsy: a randomized controlled trial. Arch Phys Med Rehabil. 2006;87(5):619-26.
39. Seniorou M, Thompson N, Harrington M, Theologis T. Recovery of muscle strength following multi-level orthopaedic surgery in diplegic cerebral palsy. Gait Posture. 2007;26(4):475-81.
40. Buckon CE, Thomas SS, Piatt JH, Jr., Aiona MD, Sussman MD. Selective dorsal rhizotomy versus orthopedic surgery: a multidimensional assessment of outcome efficacy. Arch Phys Med Rehabil. 2004;85(3):457-65.
41. Sung KH, Lee J, Chung CY, Lee KM, Cho BC, Moon SJ, et al. Factors influencing outcomes after medial hamstring lengthening with semitendinosus transfer in patients with cerebral palsy. J Neuroeng Rehabil. 2017;14(1):83.
42. Godwin EM, Spero CR, Nof L, Rosenthal RR, Echternach JL. The gross motor function classification system for cerebral palsy and single-event multilevel surgery: is there a relationship between level of function and intervention over time? J Pediatr Orthop. 2009;29(8):910-5.
43. Lehtonen K, Maenpaa H, Piirainen A. Does single-event multilevel surgery enhance physical functioning in the real-life environment in children and adolescents with cerebral palsy (CP)?: patient perceptions five years after surgery. Gait Posture. 2015;41(2):448-53.
44. Gough M, Schneider P, Shortland AP. The outcome of surgical intervention for early deformity in young ambulant children with bilateral spastic cerebral palsy. J Bone Joint Surg Br. 2008;90(7):946-51.
45. Schranz C, Kruse A, Kraus T, Steinwender G, Svehlik M. Does unilateral single-event multilevel surgery improve gait in children with spastic hemiplegia? A retrospective analysis of a long-term follow-up. Gait Posture. 2017;52:135-9.
46. Graham HK, Baker R, Dobson F, Morris ME. Multilevel orthopaedic surgery in group IV spastic hemiplegia. J Bone Joint Surg Br. 2005;87(4):548-55.
47. Dreher T, Wolf S, Braatz F, Patikas D, Doderlein L. Internal rotation gait in spastic diplegia--critical considerations for the femoral derotation osteotomy. Gait Posture 2007;26(1):25-31.
48. Khouri N, Desailly E. Rectus femoris transfer in multilevel surgery: technical details and gait outcome assessment in cerebral palsy patients. Orthop Traumatol Surg Res. 2013;99(3):333-40.

49. Ng BKW, Chau WW, Hung ALH, Lam TP, Cheng, JCY. Soft tissue release and osteotomies in the treatment of patients with spastic diplegic cerebral palsy. J Orthop Trauma Rehabil. 2018;24:80-3.
50. Chang CH, Chen YY, Yeh KK, Chen CL. Gross motor function change after multilevel soft tissue release in children with cerebral palsy. Biomed J. 2017;40(3):163-8.
51. Khan MA. Outcome of single-event multilevel surgery in untreated cerebral palsy in a developing country. J Bone Joint Surg Br. 2007;89(8):1088-91.
52. Cuomo AV, Gamradt SC, Kim CO, Pirpiris M, Gates PE, McCarthy JJ, et al. Health-related quality of life outcomes improve after multilevel surgery in ambulatory children with cerebral palsy. J Pediatr Orthop. 2007;27(6):653-7.
53. Adolfsen SE, Ounpuu S, Bell KJ, DeLuca PA. Kinematic and kinetic outcomes after identical multilevel soft tissue surgery in children with cerebral palsy. J Pediatr Orthop. 2007;27(6):658-67.
54. Bernthal NM, Gamradt SC, Kay RM, Wren TA, Cuomo AV, Reid J, et al. Static and dynamic gait parameters before and after multilevel soft tissue surgery in ambulating children with cerebral palsy. J Pediatr Orthop. 2010;30(2):174-9.
55. Lofterod B, Terjesen T. Changes in lower limb rotation after soft tissue surgery in spastic diplegia. Acta Orthop. 2010;81(2):245-9.
56. Steinwender G, Saraph V, Zwick EB, Uitz C, Linhart W. Assessment of hip rotation after gait improvement surgery in cerebral palsy. Acta Orthop Belg. 2000;66(3):259-64.
57. Rutz E, Baker R, Tirosh O, Brunner R. Are results after single-event multilevel surgery in cerebral palsy durable? Clin Orthop Relat Res. 2013;471(3):1028-38.
58. Schwartz MH, Rozumalski A, Steele KM. Dynamic motor control is associated with treatment outcomes for children with cerebral palsy. Dev Med Child Neurol. 2016;58(11):1139-45.
59. Harvey A, Graham HK, Morris ME, Baker R, Wolfe R. The Functional Mobility Scale: ability to detect change following single event multilevel surgery. Dev Med Child Neurol. 2007;49(8):603-7.
60. Firth GB, Passmore E, Sangeux M, Thomason P, Rodda J, Donath S, et al. Multilevel surgery for equinus gait in children with spastic diplegic cerebral palsy: medium-term follow-up with gait analysis. J Bone Joint Surg Am. 2013;95(10):931-8.
61. Svehlik M, Steinwender G, Kraus T, Saraph V, Lehmann T, Linhart WE, et al. The influence of age at single-event multilevel surgery on outcome in children with cerebral palsy who walk with flexed knee gait. Dev Med Child Neurol. 2011;53(8):730-5.
62. Gannotti ME, Gorton GE, Nahorniak MT, Masso PD, Landry B, Lyman J, et al. Postoperative gait velocity and mean knee flexion in stance of ambulatory children with spastic diplegia four years or more after multilevel surgery. J Pediatr Orthop. 2007;27(4):451-6.
63. Harvey A, Rosenbaum P, Hanna S, Yousefi-Nooraie R, Graham KH. Longitudinal changes in mobility following single-event multilevel surgery in ambulatory children with cerebral palsy. J Rehabil Med. 2012;44(2):137-43.
64. Thomason P, Selber P, Graham HK. Single Event Multilevel Surgery in children with bilateral spastic cerebral palsy: a 5 year prospective cohort study. Gait Posture. 2013;37(1):23-8.
65. Lee KM, Chung CY, Park MS, Lee SH, Choi IH, Cho TJ, et al. Level of improvement determined by PODCI is related to parental satisfaction after single-event multilevel surgery in children with cerebral palsy. J Pediatr Orthop. 2010;30(4):396-402.

66. Lee SH, Chung CY, Park MS, Choi IH, Cho TJ, Yoo WJ, et al. Parental satisfaction after single-event multilevel surgery in ambulatory children with cerebral palsy. J Pediatr Orthop. 2009;29(4):398-401.
67. Stephan-Carlier A, Facione J, Speyer E, Rumilly E, Paysant J. Quality of life and satisfaction after multilevel surgery in cerebral palsy: confronting the experience of children and their parents. Ann Phys Rehabil Med. 2014;57(9-10):640-52.
68. Dreher T, Brunner R, Vegvari D, Heitzmann D, Gantz S, Maier MW, et al. The effects of muscle-tendon surgery on dynamic electromyographic patterns and muscle tone in children with cerebral palsy. Gait Posture. 2013;38(2):215-20.
69. Dreher T, Buccoliero T, Wolf SI, Heitzmann D, Gantz S, Braatz F, et al. Long-term results after gastrocnemius-soleus intramuscular aponeurotic recession as a part of multilevel surgery in spastic diplegic cerebral palsy. J Bone Joint Surg Am. 2012;94(7):627-37.
70. Dreher T, Thomason P, Svehlik M, Doderlein L, Wolf SI, Putz C, et al. Long-term development of gait after multilevel surgery in children with cerebral palsy: a multicentre cohort study. Dev Med Child Neurol. 2018;60(1):88-93.
71. Klotz MC, Wolf SI, Heitzmann D, Krautwurst B, Braatz F, Dreher T. Reduction in primary genu recurvatum gait after aponeurotic calf muscle lengthening during multilevel surgery. Res Dev Disabil. 2013;34(11):3773-80.
72. Svehlik M, Steinwender G, Lehmann T, Kraus T. Predictors of outcome after single-event multilevel surgery in children with cerebral palsy: a retrospective ten-year follow-up study. Bone Joint J. 2016;98-B(2):278-81.
73. Gough M, Eve LC, Robinson RO, Shortland AP. Short-term outcome of multilevel surgical intervention in spastic diplegic cerebral palsy compared with the natural history. Dev Med Child Neurol. 2004;46(2):91-7.
74. Klotz MCM, Hirsch K, Heitzmann D, Maier MW, Hagmann S, Dreher T. Distal femoral extension and shortening osteotomy as a part of multilevel surgery in children with cerebral palsy. World J Pediatr. 2017;13(4):353-9.
75. Ounpuu S, DeLuca P, Davis R, Romness M. Long-term effects of femoral derotation osteotomies: an evaluation using three-dimensional gait analysis. J Pediatr Orthop. 2002;22(2):139-45.
76. Ounpuu S, Solomito M, Bell K, DeLuca P, Pierz K. Long-term outcomes after multilevel surgery including rectus femoris, hamstring and gastrocnemius procedures in children with cerebral palsy. Gait Posture. 2015;42(3):365-72.
77. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240(2):205-13.
78. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;337:a1655.
79. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
80. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687.

81. Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Burgi E, Scherer M, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. BMJ. 2009;339:b3244.
82. Vos RC, Becher JG, Voorman JM, Gorter JW, van Eck M, van Meeteren J, et al. Longitudinal Association Between Gross Motor Capacity and Neuromusculoskeletal Function in Children and Youth With Cerebral Palsy. Arch Phys Med Rehabil. 2016;97(8):1329-37.
83. Savovic J, Jones H, Altman D, Harris R, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess. 2012;16(35):1-82.
84. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76.
85. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ. 2013;346:f2304.

Table 1. Description of included studies

Author	Design[a]	Participants				Surgery Type[b]	Procedure per patient, mean±SD (range)	Duration of follow-up, yr
		n (M,F)	Age, yr mean±SD (range)	Distribution (%)	GMFCS level (%)			
Adolfsen 2007	Retro	31 (NS)	8.5±2 (5-15)	D (65), H (32), Q (3)	NS	Soft tissue	4.4	*Mean (range)* 1.9 (0.7- 6.4)
Akerstedt 2010	Prosp single-subject	11 (10,1)	13.8 (9-18)	Uni (9), Bi (91)	I (27), II (55), III (8)	Mixed	NS	2
Ancillao 2017	Unclear	9 (7,2)	11.1±2.4	Bi (100)	II-III	Mixed	NS	*Mean±SD* 1.2±0.4
Bernthal 2010	Prosp	23 (NS)	9.2 (6-14)	NS	NS	Soft tissue	4.2 (2-8)	1
Blumetti 2015	Retro (RFT, non-RFT)	216 (NS)	DRFT: 12.6±5.8 Non DRFT: 10.3±5.6	D (100)	I–III	Mixed	NS	*Mean±SD* DRFT: 3.38±3.19; Non- DRFT: 3.4±2.74
Bohm 2017	Retro (PT, PN)	32 (NS)	PT: 13.5±1.9 PN: 13.7±2.3 (8-18)	Bi (100)	I (16), II (84)	Mixed	NS	*Mean±SD* 2.0±0.2
Braatz 2013	Prosp	30 (18,12)	11.6±2.9	Bi (100)	I-III	Mixed	6.4	*Mean±SD* 1.1±0.2
Buckon 2004	Prosp	7 (4,3)	6.6 (4.3-11)	D (100)	I (29), II(14), III (57)	Mixed	NS	2
Chang 2017	Prosp	25 (20,5)	8.6 (4-12)	Bi (52), Q (48)	II (36), III (48), IV (16)	Soft tissue	NS	0.5
Cuomo 2007	Prosp	57 (NS)	9.5 (5-15)	NS	NS	Soft tissue	NS	*Mean (range)* 1.3 (0.8 -2.5)
Desailly 2017	Retro (PTS, no-PTS)	41 (NS)	PT: 10.8±1.6 No PTS: 10.8±1.9	NS	I-III	Mixed	NS	*Mean±SD* 2.3±0.8
Dreher 2007	Prosp	30 (22,8)	10.0± 2.9 (6-16)	D (100)	I (10), II (67), III (3),	Mixed	10.5	*Mean±SD* 11.3±0.7

					IV (20)			
Dreher 2012a	Prosp (DRFT, non-DRFT)	32 (19,13)	10.9±2.8 (6–16)	D (100)	I (9), II (53), III (38)	Mixed	7.1	*Mean±SD* DRFT: 1.0±0.6; Non-DRFT: 1.2±0.2
Dreher 2012b	Retro	39 (26,13)	10.2±3.5 (6-16)	D (100)	I (15), II (54), III (31)	Mixed	10.1	*Mean±SD* 8.1±1.8
Dreher 2012c	Retro (C-DRFT, P-DRFT)[c]	53 (36,17)	11.0±3.4	D (100)	I-III	Mixed	NS	*Mean±SD* 8.8±2.3
Dreher 2012d	Retro	44 (26,18)	9.8±3.4	D (100)	I (16), II (64), III (20)	Mixed	NS	*Mean±SD* 8.6 ±2
Dreher 2013a	Prosp (CBM, MTL)	42 (28,14)	CBM: 11.3±3.1 MTL: 11.1±5.4 (6–16)	D (100)	I (14), II (57), III (29)	Mixed	12.9	*Mean±SD* CBM: 9.2±2.5 MTL: 9.1±2.6
Dreher 2013b	Retro	42 (27,15)	9.8±2.8 (6-16)	D (100)	I (14), II (67), III (19)	Mixed	NS	*Mean±SD* 3.2±1.1
Dreher 2018	Retro	231 (142,89)	10.6±2.9 (5-16)	Bi (100)	I (8), II (62), III (29)	Mixed	8±3	*Mean±SD* 9.1±3.0
Feger 2015	Retro (O, MTL)[d]	48 (27,21)	O: 9.9±3.6 MTL: 9.0±3.6	D (65), H (29), Other (6)	I (29), II (35), III (35)	Mixed	O: 3.0±1.5 (1-5) MTS: 2.2±1.3 (1-4)	*Mean±SD (range)* O: 1.7±0.6 (1.0-2.9) MT: 1.9±0.6 (1.0-3.0)
Firth 2013	Retro	40 (25,15)	9.3 (4.8-15.1)	D (100)	II (68), III (32)	Mixed	9.1 (5-18)	*Mean (range)* 7.5 (4.4-14.6)
Gannotti 2007	Retro	20 (19,1)	9.5±3.0 (4.9-16.0)	D (100)	II (80), III (20)	Mixed	3.6	*Minimum* 4
Gannotti 2010	Retro	11 (5,6)	9.9 (7-13)	D (73), Q (27)	II (55), III (45)	Mixed	NS	*Mean (range)* 13 (11-15)
Godwin 2009	Retro	84 (35,49)	5.99±2.59 (3.25-16.33)	D (51), H (21), Q (27)	I (15), II (18), III (20), IV (33),	Soft tissue	5.45	5

					V (13)			
Gough 2004	Retro	12 (8,4)	9.8±3.3 (5.5-15.3)	D (100)	NS	Mixed	NS	*Mean (range)* 1.5 (0.8-2.4)
Gough 2008	Retro	13 (13,0)	6.4 (5.5-8.9)	Bi (100)	II (54), III (38), IV (8)	Mixed	5.1	*Mean* 4.2
Dobson 2005	Prosp	17 (14,3)	12.1 (7.1-17.1)	H (100)	IV (100)	Mixed	4.2 (2-7)	*Mean±SD* 2.9±0.9
Harvey 2007	Retro	66 (34,32)	10±2.5 (6–16)	D (100)	I (27), II (36), III (36)	NS	8 (4-12)	2
Harvey 2012	Retro	156 (99,57)	11.1±2.47 (6–19)	NS	I/II (62), III (38)	Mixed	7.6±2.1	5
Khan 2007	Prosp	85 (53,32)	8.5 (5-12)	D (100)	-	Mixed	NS	*Mean (range)* 3.5 (2-5)
Khouri 2013	Retro	25 (NS)	12±3	NS	I (27), II (42), III (23), IV (4)	Mixed	NS	*Mean±SD* 2.1±0.8
Klotz 2013	Retro	19 (13,6)	9.4±4.9 (4-23)	Bi (100)	II (100)	Mixed	NS	*Mean±SD (range)* 1.1±0.5 (0.8-2.8)
Klotz 2017	Retro	22 (14,8)	12.1±3.1 (6-16)	Bi (100)	I (9), II (41), III (50)	Mixed	NS	*Mean±SD (range)* 1.3±0.3 (1.0-1.9)
Klotz 2018	Prosp (PTS, no-PTS)	20 (NS)	10.4±2.6 (6–18)	Bi (100)	II (55), III (45)	Mixed	NS	*Mean±SD* 1.1±0.1
Kwon 2013	Retro (FDO, no-FDO)	53 (38,15)	FDO: 6.8±1.5 no FDO: 7.4±2.4	D (100)	I (68), II (32)	Mixed	NS	*Mean±SD* 1.1±0.4
Laracca 2014	Retro[e]	30 (17,13)	Group 1: 14.0 (11–17); Group 2: 12.6 (10–14)	D (100)	NS	Mixed	G1: 7.7 G2: 6.8	*Mean (range)* G1: 2.3 (1.1-4.3) G2: 1.8 (0.7-3.2)
Lee 2010	Prosp	61 (40,21)	10.2±3.8 (5.1–18)	D (13), H (87)	I (46), II (30), III (24)	Mixed	NS	*Mean±SD (range)* 2.2±0.5 (1.1-2.8)
Lee 2009	Retro	279 (191,88)	NS	D (58),	I (69),	Mixed	5.4±3.0	*Mean±SD*

				H (42)	II (26), III (5)			6.0±3.7
Lehtonen 2015	Retro	10 (2,8)	NS	D (100)	II (30), III (60), IV (10)	Mixed	4.2 (2–7)	5
Lofterød 2010	Retro	28 (16,12)	12 (7–19)	D (100)	I (21), II (50), III (29)	Soft tissue	6	*Mean (range)* 1.2 (1.0-2.1)
Mallet 2016	Retro (IMPL, no-IMPL)	34 (26,8)	13±2.8 (6.8–18.5)	D (100)	II (88), III (12)	Mixed	NS	*Mean±SD (range)* 2.4±2.0 (1.0–8.7)
Metaxiotis 2004	Prosp	20 (13,7)	11.5 (5.6-17.0)	D (100)	NS	Mixed	12.05	*Mean (range)* 3.1 (2.0-4.5)
Ng 2018	Retro	20 (14, 6)	12.7±3.1 (5.6-18.2)	D (100)	I (23), II (9), III (36), IV (27), V (5)	Mixed	NS	*Mean±SD (range)* 5.45±3.06 (1.48-13.11)
Õunpuu 2002	Prosp	20 (NS)	8.1±2.9 (5–15)	NS	NS	Mixed	NS	5
Õunpuu 2015	Retro	22 (11,11)	8.0±2.7	D (64), T (5), H (18), Q(14)	I (22.5), II (55), III (22.5)	Mixed	NS	*Mean±SD* 11±2
Patikas 2006	Prosp (EG, CG)	39 (27,12)	ST: 10.6±3.2 C: 8.9±1.9 (6-16)	D (100)	I (31), II (46), III (23)	Mixed	NS	2
Patikas 2007	Prosp	34 (22,12)	10.1±3.0 (6-16)	D (53), H (47)	NS	Mixed	NS	*Mean±SD (range)* 2.53±1.23 (1–5)
Rodda 2006	Retro	10 (7,3)	12.0 (7.9-16.2)	D (100)	II (30), III (70)	Mixed	7 (5-8)	*Minimum 5*
Rutz 2012	Retro	107 (61,46)	10.6±2.7 (6–17)	Bi (100)	II (69), III (31)	Mixed	8 (4–14)	*Mean±SD (range)* 5.0±2.7 (2–12)
Rutz 2013a	Retro	121 (73,48)	10.7±2.7	Bi (100)	II (66), III (34)	Mixed	7.6±2.1	*Mean±SD* 1.3±1.0
Rutz 2013b	Retro	14 (10,4)	13 (7–18)	D (100)	I (7), II (71), III (21)	NS	7.4±2.8 (4-15)	*Mean (range)* 1.8 (1-3)

Saraph 2000	Retro	22 (NS)	12.6 (7.4-16.6)	D (100)	NS	Mixed	8.2 (3-8)	*Mean (range)* 2 (2.1-4.0)
Saraph 2001	Retro	12 (NS)	12.7±3.3	D (100)	NS	Mixed	6.2	*Mean±SD* 3.2±0.6
Saraph 2002a	Retro (D, H)	22 (NS)	11.9 (9.2–15.5)	D (36), H (64)	NS	Mixed	7.4	*Mean (range)* 3.1 (3.0–3.9)
Saraph 2002b	Retro	25 (NS)	13.6 (6.0–15.5)	D (100)	NS	Mixed	8.2	*Mean (range)* 3.3 (3.0-3.9)
Saraph 2005	Retro	32 (NS)	11.1 (8.7–13.5)	D (100)	NS	Mixed	8.1	*Mean±SD* 4.4±1.1
Saraph 2006	Retro	11 (NS)	12.4 (9.5–17.2)	D (27), H (73)	NS	Mixed	7.5	*Mean (range)* 3.2 (2.3–4.0)
Schranz 2017	Retro	14 (9,5)	12.1±3.3 (6-17)	Uni (100)	I (43), II (57)	Mixed	5.1	*Mean±SD (range)* 9.4±2.2 (6-13)
Schwartz 2016	Retro	176 (99,77)	10.0±3.4	NS	I (14), II (48), III (38)	NS	NS	*Mean±SD* 1.4±0.4
Seniorou 2007	Prosp (EG, CG)	20 (10,10)	12.5±2.5 (7–16)	D (100)	I (15), II (65), III (20)	Mixed	9.1	1
Steinwender 2000	Retro	16 (NS)	10.2 (6-14)	D (100)	NS	Soft tissue	5.69	*Mean (range)* 3.4 (2.0-5.7)
Stephan-Carlier 2014	Retro	12 (10,2)	14.0±2.5 (11-18)	D (83), H (17)	I (25), II (25), III (50)	Mixed	8 (3-16)	*Mean (range)* 2.5 (1.0-4.8)
Sung 2013	Retro	29 (18,11)	8.3±2.6 (5.4–16.3)	D (100)	I (24), II (66), III(10)	Mixed	7.3	*Mean (range)* 11.8 (10.0–13.3)
Sung 2017	Retro	314 (198,116)	7.9 ± 3.7 (3.4–20.0)	D (100)	I-III	Mixed	6.2	*Mean±SD (range)* 2.7±2.9 (1.0-14.7)
Svehlík 2011	Retro	32 (17,15)	10.3±3.1 (5.7–15.5)	Bi (100)	II (38), III (62)	Mixed	NS	*Mean±SD* 21.3±3.3
Svehlik 2016	Retro	39 (22,17)	10.3±3.1 (5.7-15.5)	Bi (100)	II (51), III (49)	Mixed	7.5	*Minimum* 10
Taylor 2016	Retro (DFEO, PKC)	31 (NS)	13 (7–18)	D (87), H (3), Q (10)	I (6), II (32), III (58), IV (3)	Mixed	NS	*Mean (range)* PKC: 2 (0.8-4); DFEO: 1 (1–2)

Thomason 2011	RCT	19 (12,7)	9.7 (6-12)	D (100)	II-III	Mixed	8±4	2
Thomason 2013	Prosp	18 (11,7)	10.2±1.7 (7–17)	Bi (100)	II (72), III (28)	Mixed	8±3	5
Thompson 2010	Prosp (MI-SEMLS, SEMLS)	20 (12,8)	11.0±1.9 (7.8-14.3)	D (100)	I (15), II (65), III (20)	Mixed	NS	1
Truong 2011	Retro (Control, Psoas)[f]	87 (NS)	NS (3-8)	NS	I/II (66), III/IV (34)	Mixed	NS	*Mean±SD* 1.0±0.3
Van Drongelen 2013	Retro	14 (8,6)	9.3±2.0 (6.6-12.0)	D (100)	I (14), II (64), III (21)	Mixed	NS	*Mean (range)* 2.0 (1.3-3.2)
Zwick 2001	Retro	17 (NS)	11.2 (5.7-16.4)	D (100)	NS	Mixed	NS	*Mean (range)* 3.8 (2.6-5.7)
Zwick 2012	Retro (M, F)	34 (19,15)	M: 9.9±3.2, F: 11.3±2.8	D (100)	II-III	Mixed	M: 4.09±2.85 F: 3.92±2.36	*Minimum* 10

[a]Comparison groups in parentheses if applicable; [b]mixed: soft tissue and bony surgery; [c]compared group that had distal rectus femoris transfer to correct decreased peak knee flexion in swing phase (C-DRFT) and group that had prophylactic distal rectus femoris transfer (P-DRFT); [d]compared group that had osteotomy with or without muscle-tendon procedure (O) to group that had muscle-tendon procdure only (MTL); [e]compared patients treated prior to year 2000 (group 1) and patients treated after 2000 (group 2); [f]compared patients who had psoas lengthening as part of multilevel surgery (psoas) and patients who did not have psoas lengthening as part of multilevel surgery (control)

Bi: Bilateral; CBM: conversion of biarticular muscles; CG: control group; D: Diplegia; Dist: Distribution; DFEO: distal femoral extension osteotomy; DRFT: distal rectus femoris transfer; EG: exercise group; F: female; FDO: femoral derotation osteotomy; GMFCS: Gross motor function classification system; H: Hemiplegia; IMPL: intramuscular psoas lengthening; M: male; MI-SEMLS: minimally; MTL: muscle tendon lengthening; invasive single event multilevel surgery; NS: not stated; PKC: posterior knee capsulotomy; PN: no anterior pelvic tilt group; Prosp: prospective; PT: anterior pelvic tilt group; PTS: patellar tendon shortening; Q: Quadriplegia; Retro: retrospective; RCT: randomised controlled trial; RFT: rectus femoris transfer; SD: standard deviation; SEMLS: single event multilevel surgery; Uni: Unilateral.

Table 2. Changes in participation and quality of life following multilevel surgery.

Study	Sample size	Outcome	Outcome measure	Domain	Result
Intermediate-term results					
Buckon 2004	7	Participation	PEDI	Social skills	MD: 7.41, 95% CI 2.56 to 12.26
Buckon 2004	7	Participation	PEDI	Self-care	MD: 5.5, 95% CI 0.63 to 10.37
Cuomo 2007	57	Participation	PODCI	Global function	MD: 8.25, p<0.001
Lee 2010	61	Participation	PODCI	Global function	MD: 2.6, p=0.02
Cuomo 2007	57	Quality of life	PedsQL		MD: 10.03, p<0.001
Long-term results					
Buckon 2004	7	Participation	PEDI	Self-care	MD: 8.17, 95% CI 2.35 to 13.99
Buckon 2004	7	Participation	PEDI	Mobility	MD: 7.34, 95% CI 0.39 to 14.2
Buckon 2004	7	Participation	PEDI	Social skills	MD: 7.67, 95% CI 3.10 to 12.24
Thomason 2011	11	Quality of life	CHQ		MD: 22, 95% CI 4 to 39

PEDI: Pediatric Evaluation of Disability Inventory; MD: mean difference; CI: confidence interval; PODCI: Pediatric Outcomes Data Collection Instrument; PedsQL: Pediatric Quality of Life instrument; Child Health Questionnaire.

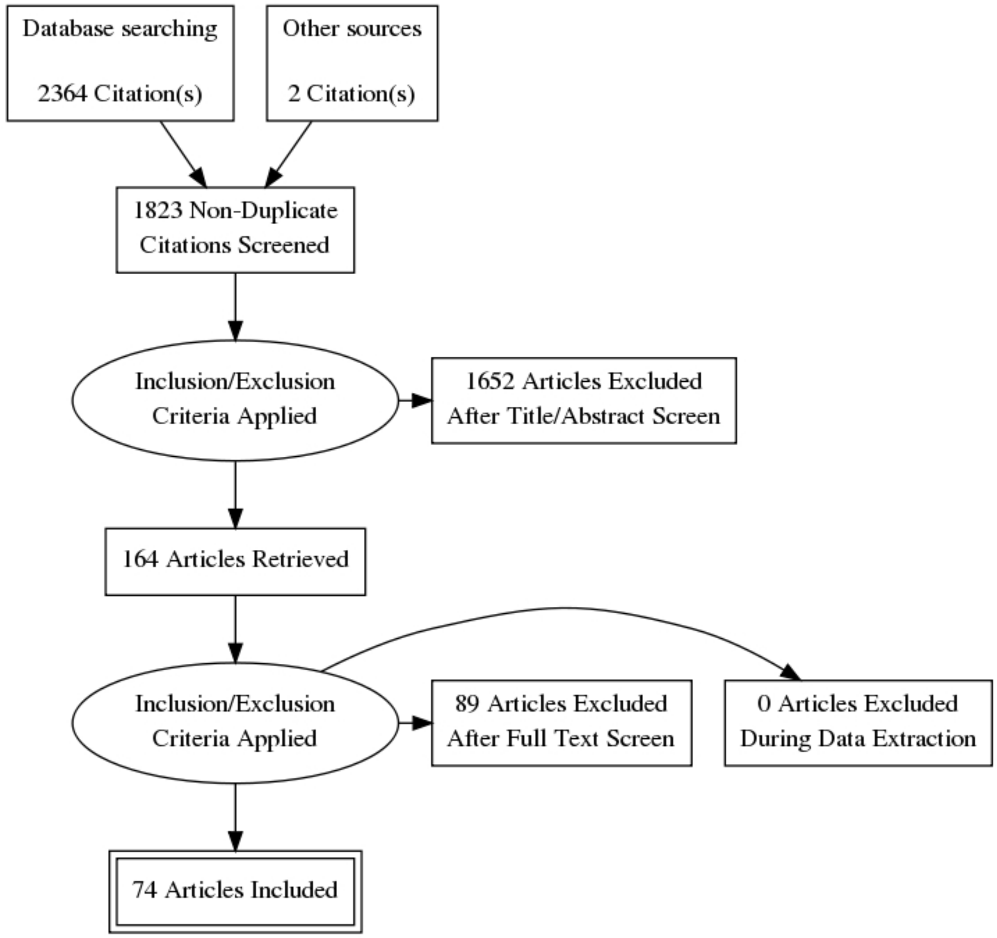


Figure 1. Study flow diagram

253x238mm (72 x 72 DPI)

Study or Subgroup	Std. Mean Difference	SE	Weight	Std. Mean Difference IV, Random, 95% CI
Buckon 2004	2.34	0.41	15.0%	2.34 [1.54, 3.14]
Feger 2015 MTL	0.51	0.21	18.3%	0.51 [0.10, 0.92]
Feger 2015 O	0.2	0.36	15.9%	0.20 [−0.51, 0.91]
Taylor 2018 DFEO	−0.17	0.34	16.3%	−0.17 [−0.84, 0.50]
Taylor 2018 PKC	−0.13	0.27	17.4%	−0.13 [−0.66, 0.40]
Thomason 2013	−0.29	0.29	17.1%	−0.29 [−0.86, 0.28]
Total (95% CI)			**100.0%**	**0.38 [−0.25, 1.01]**

Heterogeneity: $Tau^2 = 0.52$; $Chi^2 = 34.41$, df = 5 ($P < 0.00001$); $I^2 = 85\%$

Test for overall effect: Z = 1.17 (P = 0.24)

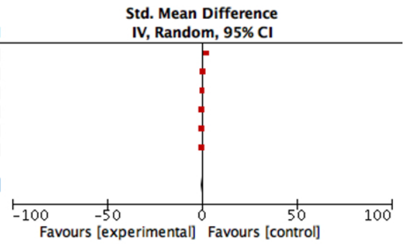
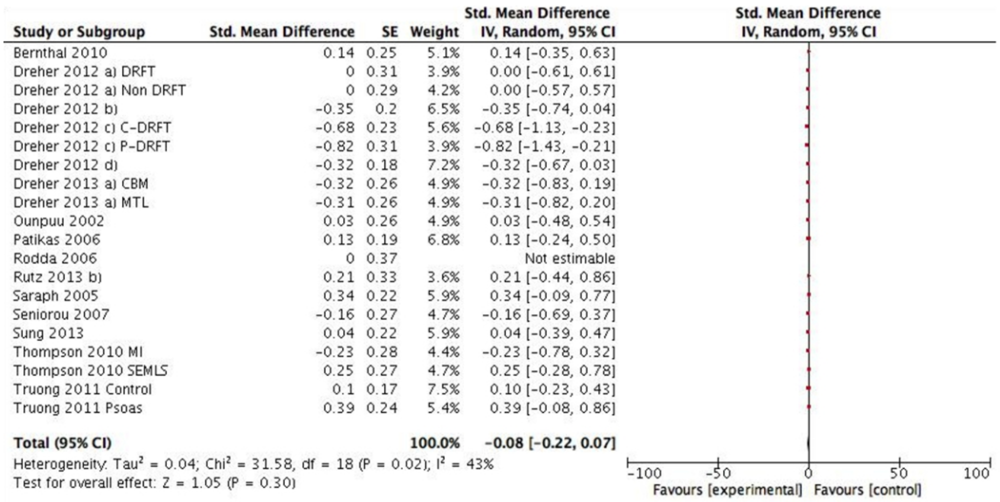


Figure 2. Long-term effect of multilevel surgery on gross motor function

298x72mm (96 x 96 DPI)



Study or Subgroup	Std. Mean Difference	SE	Weight	Std. Mean Difference IV, Random, 95% CI
Bernthal 2010	0.14	0.25	5.1%	0.14 [-0.35, 0.63]
Dreher 2012 a) DRFT	0	0.31	3.9%	0.00 [-0.61, 0.61]
Dreher 2012 a) Non DRFT	0	0.29	4.2%	0.00 [-0.57, 0.57]
Dreher 2012 b)	-0.35	0.2	6.5%	-0.35 [-0.74, 0.04]
Dreher 2012 c) C-DRFT	-0.68	0.23	5.6%	-0.68 [-1.13, -0.23]
Dreher 2012 c) P-DRFT	-0.82	0.31	3.9%	-0.82 [-1.43, -0.21]
Dreher 2012 d)	-0.32	0.18	7.2%	-0.32 [-0.67, 0.03]
Dreher 2013 a) CBM	-0.32	0.26	4.9%	-0.32 [-0.83, 0.19]
Dreher 2013 a) MTL	-0.31	0.26	4.9%	-0.31 [-0.82, 0.20]
Ounpuu 2002	0.03	0.26	4.9%	0.03 [-0.48, 0.54]
Patikas 2006	0.13	0.19	6.8%	0.13 [-0.24, 0.50]
Rodda 2006	0	0.37		Not estimable
Rutz 2013 b)	0.21	0.33	3.6%	0.21 [-0.44, 0.86]
Saraph 2005	0.34	0.22	5.9%	0.34 [-0.09, 0.77]
Seniorou 2007	-0.16	0.27	4.7%	-0.16 [-0.69, 0.37]
Sung 2013	0.04	0.22	5.9%	0.04 [-0.39, 0.47]
Thompson 2010 MI	-0.23	0.28	4.4%	-0.23 [-0.78, 0.32]
Thompson 2010 SEMLS	0.25	0.27	4.7%	0.25 [-0.28, 0.78]
Truong 2011 Control	0.1	0.17	7.5%	0.10 [-0.23, 0.43]
Truong 2011 Psoas	0.39	0.24	5.4%	0.39 [-0.08, 0.86]
Total (95% CI)			**100.0%**	**-0.08 [-0.22, 0.07]**

Heterogeneity: $Tau^2 = 0.04$; $Chi^2 = 31.58$, df = 18 (P = 0.02); $I^2 = 43\%$
Test for overall effect: Z = 1.05 (P = 0.30)

Figure 3. Long-term effect of multilevel surgery on gait speed

316x158mm (96 x 96 DPI)

Study or Subgroup	Std. Mean Difference	SE	Weight	Std. Mean Difference IV, Random, 95% CI
Ancillao 2017	0.99	0.48	1.6%	0.99 [0.05, 1.93]
Bohm 2017 PN	−0.27	0.3	2.7%	−0.27 [−0.86, 0.32]
Bohm 2017 PT	−0.68	0.31	2.6%	−0.68 [−1.29, −0.07]
Desailly 2017 Non PTS	−1.44	0.49	1.6%	−1.44 [−2.40, −0.48]
Desailly 2017 PTS	−1.65	0.25	3.1%	−1.65 [−2.14, −1.16]
Dreher 2018	−1	0.1	4.5%	−1.00 [−1.20, −0.80]
Dreher 2012 b)	−0.66	0.21	3.5%	−0.66 [−1.07, −0.25]
Dreher 2012 c) C–DRFT	−0.59	0.22	3.4%	−0.59 [−1.02, −0.16]
Dreher 2012 c) P–DRFT	−0.66	0.29	2.8%	−0.66 [−1.23, −0.09]
Dreher 2012 d)	−0.63	0.2	3.6%	−0.63 [−1.02, −0.24]
Dreher 2013 a) CBM	−0.45	0.27	3.0%	−0.45 [−0.98, 0.08]
Dreher 2013 a) MTL	−0.7	0.29	2.8%	−0.70 [−1.27, −0.13]
Firth 2013	−1.72	0.29	2.8%	−1.72 [−2.29, −1.15]
Gannotti 2010	−0.14	0.36	2.3%	−0.14 [−0.85, 0.57]
Gough 2008	−1.48	0.47	1.6%	−1.48 [−2.40, −0.56]
Khouri 2013	−1.25	0.32	2.6%	−1.25 [−1.88, −0.62]
Laracca 2014 G1	−0.74	0.34	2.4%	−0.74 [−1.41, −0.07]
Laracca 2014 G2	−0.21	0.29	2.8%	−0.21 [−0.78, 0.36]
Lehtonen 2015	−1.14	0.32	2.6%	−1.14 [−1.77, −0.51]
Mallet 2016 IMPL	−0.79	0.33	2.5%	−0.79 [−1.44, −0.14]
Mallet 2016 No IMPL	−1.36	0.25	3.1%	−1.36 [−1.85, −0.87]
Patikas 2007 H	−0.76	0.34	2.4%	−0.76 [−1.43, −0.09]
Patikas 2007 D	−0.84	0.32	2.6%	−0.84 [−1.47, −0.21]
Rutz 2012	−1.08	0.14	4.1%	−1.08 [−1.35, −0.81]
Rutz 2013 a)	−1.15	0.13	4.2%	−1.15 [−1.40, −0.90]
Schranz 2016	−0.48	0.33	2.5%	−0.48 [−1.13, 0.17]
Sung 2013	−1.03	0.27	3.0%	−1.03 [−1.56, −0.50]
Sung 2017	−0.84	0.08	4.6%	−0.84 [−1.00, −0.68]
Svehlík 2011	−0.81	0.17	3.9%	−0.81 [−1.14, −0.48]
Taylor 2018 DFEO	−0.47	0.28	2.9%	−0.47 [−1.02, 0.08]
Taylor 2018 PKC	−0.8	0.72	0.9%	−0.80 [−2.21, 0.61]
Thomason 2011	−1.66	0.55	1.3%	−1.66 [−2.74, −0.58]
Thomason 2013	1.48	0.41	2.0%	1.48 [0.68, 2.28]
Van Drongelen 2013	−1.15	0.41	2.0%	−1.15 [−1.95, −0.35]
Zwick 2012 Female	−0.97	0.29	2.8%	−0.97 [−1.54, −0.40]
Zwick 2012 Male	−0.74	0.23	3.3%	−0.74 [−1.19, −0.29]
Total (95% CI)			**100.0%**	**−0.80 [−0.95, −0.65]**

Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 112.13$, $df = 35$ ($P < 0.00001$); $I^2 = 69\%$

Test for overall effect: $Z = 10.79$ ($P < 0.00001$)

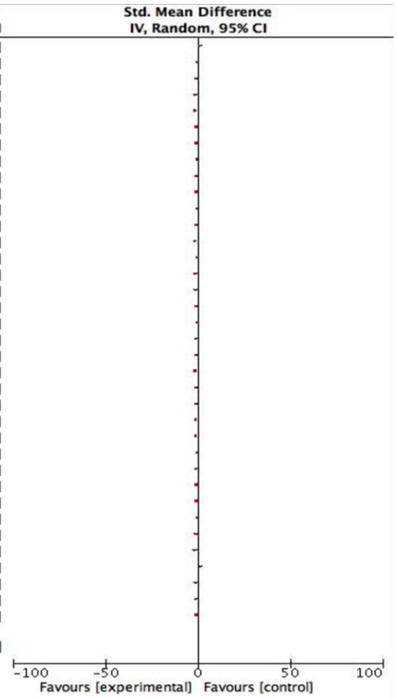


Figure 4. Long-term effect of multilevel surgery on gait

317x221mm (96 x 96 DPI)

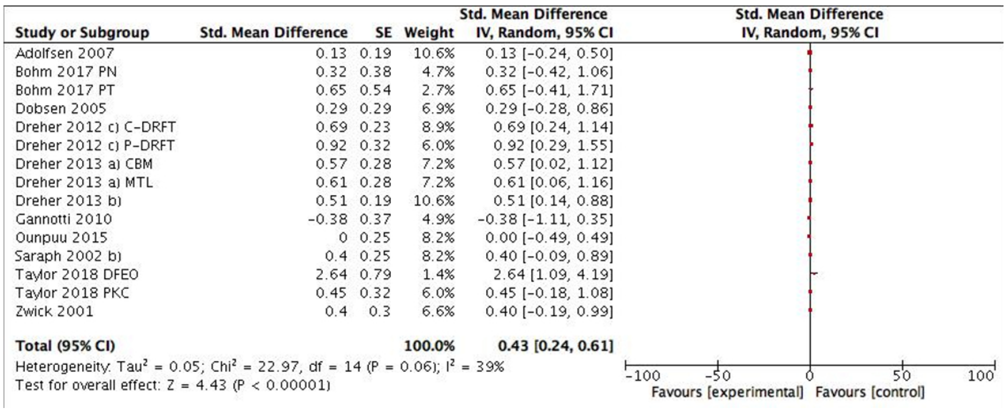


Figure 5. Long-term effect of multilevel surgery on knee extension range of motion.

299x120mm (96 x 96 DPI)

Supplemental table. Quality assessment of included studies

Author	Quality of reporting[a]				MINORS[b]
	Surgery	Previous surgery	Adverse events	Rehabilitation	
Adolfsen 2007	2	1	0	0	7
Akerstedt 2010	1	0	2	2	12
Ancillao 2017	1	1	0	0	6
Bernthal 2010	2	0	0	2	12
Blumetti 2015	1	0	0	0	4
Bohm 2017	1	0	0	0	5
Braatz 2013	1	2	0	1	10
Buckon 2004	2	0	0	2	10
Chang 2017	2	2	0	2	10
Cuomo 2007	1	0	0	1	11
Desailly 2017	1	0	0	1	5
Dreher 2007	2	2	0	1	8
Dreher 2012a	2	2	0	1	13
Dreher 2012b	2	0	2	1	5
Dreher 2012c	2	0	0	1	7
Dreher 2012d	2	0	1	1	7
Dreher 2013a	2	2	2	1	7
Dreher 2013b	1	2	0	0	7
Dreher 2018	1	2	2	0	6
Feger 2015	2	2	0	2	8
Firth 2013	2	1	0	1	7
Gannotti 2007	2	1	0	0	7
Gannotti 2010	1	0	0	0	9
Godwin 2009	2	0	0	0	9

Gough 2004	1	0	0	1	6
Gough 2008	2	0	1	1	6
Dobson 2005	2	1	2	2	9
Harvey 2007	0	0	0	0	10
Harvey 2012	1	0	0	1	10
Khan 2007	2	0	2	2	6
Khouri 2013	2	0	0	0	4
Klotz 2013	2	2	0	1	9
Klotz 2017	2	2	0	1	8
Klotz 2018	2	2	2	1	12
Kwon 2013	2	2	0	1	8
Laracca 2014	2	1	0	1	5
Lee 2010	1	2	0	1	12
Lee 2009	2	2	0	0	9
Lehtonen 2015	1	0	0	2	9
Lofterød 2010	2	2	0	0	7
Mallet 2016	1	1	0	0	7
Metaxiotis 2004	2	2	2	2	8
Ng 2018	1	0	0	0	7
Õunpuu 2002	2	2	2	1	7
Õunpuu 2015	2	0	0	1	8
Patikas 2006	2	2	1	2	9
Patikas 2007	2	2	0	0	6
Rodda 2006	2	2	0	2	10
Rutz 2012	1	0	0	1	7
Rutz 2013a	1	1	1	1	7
Rutz 2013b	2	1	2	2	9
Saraph 2000	2	0	2	2	4
Saraph 2001	2	0	0	2	6
Saraph 2002a	2	2	0	2	6
Saraph 2002b	2	2	0	2	5

Saraph 2005	2	2	0	2	6
Saraph 2006	2	2	0	2	4
Schranz 2017	2	2	2	2	8
Schwartz 2016	0	0	0	0	7
Seniorou 2007	2	1	0	2	11
Steinwender 2000	2	2	0	2	7
Stephan-Carlier 2014	1	0	0	1	7
Sung 2013	2	0	0	2	8
Sung 2017	2	0	0	1	8
Svehlík 2011	1	2	0	2	5
Svehlik 2016	1	2	0	2	5
Taylor 2016	2	0	2	2	8
Thomason 2011	2	2	2	2	21
Thomason 2013	2	2	2	2	13
Thompson 2010	2	1	2	2	8
Truong 2011	1	2	2	0	5
Van Drongelen 2013	1	0	0	0	8
Zwick 2001	2	2	0	2	9
Zwick 2012	1	2	0	2	5

[a]assessed as 0 (not reported), 1 (reported but inadequate), 2 (reported and adequate); [b]total possible score is 16 for all studies except for Thomason 2011;

Appendix 1 Search strategy

CINAHL

1. AB Cerebral palsy
2. Orthopaedic surgery OR
3. Orthopedic surgery OR
4. Orthopaedics OR
5. Orthopedics OR
6. Surgery OR
7. SEMLS OR
8. Single event multilevel surgery OR
9. Single event multi level surgery OR
10. Multilevel OR
11. Multi level OR
12. AB (2 to11)
13. 1 AND 12

Ovid MEDLINE

1. Keyword: Cerebral palsy
2. Orthopaedic surgery OR
3. Orthopedic surgery OR
4. Orthopaedics OR
5. Orthopedics OR
6. Surgery OR
7. SEMLS OR
8. Single event multilevel surgery OR
9. Single event multi level surgery OR
10. Multilevel OR
11. Multi level OR
12. Keyword (2 to 11)
13. Map term to subheading 12
14. 1 AND 13

EMBASE

1. Title, abstract, keywords: Cerebral palsy
2. Orthopaedic surgery OR
3. Orthopedic surgery OR
4. Orthopaedics OR
5. Orthopedics OR
6. Surgery OR
7. SEMLS OR
8. Single event multilevel surgery OR
9. Single event multi level surgery OR
10. Multilevel OR
11. Multi level OR
12. Title, abstract, keywords (2 to 11)
13. 1 AND 12

Cochrane Central Register of Controlled Trials (CENTRAL)

1. Title, abstract, keywords:Cerebral palsy
2. Orthopaedic surgery OR
3. Orthopedic surgery OR
4. Orthopaedics OR
5. Orthopedics OR
6. Surgery OR
7. SEMLS OR
8. Single event multilevel surgery OR
9. Single event multi level surgery OR
10. Multilevel OR
11. Multi level OR
12. Title, abstract, keywords (2 to 11)
13. 1 AND 12