Surgical interventions for the early management of Bell's palsy
Menchetti, Isabella; McAllister, Kerrie; Walker, David; Donnan, Peter T.

Published in:
The Cochrane database of systematic reviews

DOI:
10.1002/14651858.CD007468.pub4

Publication date:
2021

Citation for published version (APA):
Surgical interventions for the early management of Bell's palsy (Review)

Menchetti I, McAllister K, Walker D, Donnan PT

Menchetti I, McAllister K, Walker D, Donnan PT.
Surgical interventions for the early management of Bell's palsy.
Cochrane Database of Systematic Reviews 2021, Issue 1. Art. No.: CD007468.
DOI: 10.1002/14651858.CD007468.pub4.

www.cochranelibrary.com
Surgical interventions for the early management of Bell's palsy

Isabella Menchetti1, Kerrie McAllister2, David Walker2, Peter T Donnan3

1College of Medicine, University of Saint Andrews, Fife, UK. 2Department of Otolaryngology, North Glasgow University NHS Trust, Glasgow, UK. 3Tayside Centre for General Practice, University of Dundee, Dundee, UK

Contact address: Isabella Menchetti, i.a.menchetti@sms.ed.ac.uk, Isabella.menchetti@gmail.com.

Editorial group: Cochrane Neuromuscular Group.

Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 1, 2021.

Citation: Menchetti I, McAllister K, Walker D, Donnan PT. Surgical interventions for the early management of Bell’s palsy. Cochrane Database of Systematic Reviews 2021, Issue 1. Art. No.: CD007468. DOI: 10.1002/14651858.CD007468.pub4.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background
Bell’s palsy is an acute unilateral facial paralysis of unknown aetiology and should only be used as a diagnosis in the absence of any other pathology. As the proposed pathophysiology is swelling and entrapment of the nerve, some surgeons suggest surgical decompression of the nerve as a possible management option; this is ideally performed as soon as possible after onset. This is an update of a review first published in 2011, and last updated in 2013. This update includes evidence from one newly identified study.

Objectives
To assess the effects of surgery in the early management of Bell’s palsy.

Search methods
On 20 March 2020, we searched the Cochrane Neuromuscular Specialised Register, CENTRAL, MEDLINE, Embase, ClinicalTrials.gov and WHO ICTRP. We handsearched selected conference abstracts for the original version of the review.

Selection criteria
We included all randomised controlled trials (RCTs) or quasi-RCTs involving any surgical intervention for Bell’s palsy. Trials compared surgical interventions to no treatment, later treatment (beyond three months), sham treatment, other surgical treatments or medical treatment.

Data collection and analysis
Three review authors independently assessed trials for inclusion, assessed risk of bias and extracted data. We used standard methodological procedures expected by Cochrane. The primary outcome was complete recovery of facial palsy at 12 months. Secondary outcomes were complete recovery at three and six months, synkinesis and contracture at 12 months, psychosocial outcomes at 12 months, and side effects and complications of treatment.

Main results
Two trials with 65 participants met the inclusion criteria; one was newly identified at this update. The first study randomised 25 participants into surgical or non-surgical (no treatment) groups using statistical charts. One participant declined surgery, leaving 24 evaluable participants. The second study quasi-randomised 53 participants; however, only 41 were evaluable as 12 declined the intervention they were allocated. These 41 participants were then divided into early surgery, late surgery or non-surgical (no treatment) groups using alternation. There was no mention on how alternation was decided. Neither study mentioned if there was any attempt to conceal allocation. Neither participants nor outcome assessors were blinded to the interventions in either study. There were no losses to follow-up in the first study. The second study lost three participants to follow-up, and 17 did not contribute to the assessment of secondary outcomes. Both studies were at high risk of bias.
Surgeons in both studies used a retro-auricular/transmastoid approach to decompress the facial nerve. For the outcome recovery of facial palsy at 12 months, the evidence was uncertain. The first study reported no differences between the surgical and no treatment groups. The second study fully reported numerical data, but included no statistical comparisons between groups for complete recovery. There was no evidence of a difference for the early surgery versus no treatment comparison (risk ratio (RR) 0.76, 95% confidence interval (CI) 0.05 to 11.11; P = 0.84; 33 participants; very low-certainty evidence) and for the early surgery versus late surgery comparison (RR 0.47, 95% CI 0.03 to 6.60; P = 0.58; 26 participants; very low-certainty evidence). We considered the effects of surgery on facial nerve function at 12 months very uncertain (2 RCTs, 65 participants; very low-certainty evidence).

Furthermore, the second study reported adverse effects with a statistically significant decrease in lacrimal control in the surgical group within two to three months of denervation. Four participants in the second study had 35 dB to 50 dB of sensorineural hearing loss at 4000 Hz, and three had tinnitus. Because of the small numbers and trial design we also considered the adverse effects evidence very uncertain (2 RCTs, 65 participants; very low-certainty evidence).

Authors’ conclusions
There is very low-certainty evidence from RCTs or quasi-RCTs on surgery for the early management of Bell’s palsy, and this is insufficient to decide whether surgical intervention is beneficial or harmful.

Further research into the role of surgical intervention is unlikely to be performed because spontaneous or medically supported recovery occurs in most cases.

**Plain Language Summary**

**Surgical operation for Bell’s palsy (idiopathic facial paralysis)**

**Review question**
What are the effects of early surgery in the management of Bell’s palsy, compared with no treatment, operations performed after three months, other types of surgery, sham (fake) treatment or treatment with medicines?

**Background**
Bell’s palsy is a paralysis of the facial muscles, usually one sided, that has no known underlying cause (idiopathic). The symptoms probably occur when a nerve in the face is trapped and swollen. People with Bell’s palsy generally recover, often with help from steroid medication, but there is a small group who do not. Some surgeons have thought that operating as soon as possible (within three months of the onset of paralysis) to free the nerve could improve recovery.

**Study characteristics**
Cochrane review authors collected and analysed studies to answer the review question. We found two studies to include in our review, which involved 65 people with Bell’s palsy. Our main measure of the effects of surgery was the complete recovery of paralysis at 12 months. The first study compared surgery with no treatment and the second study compared early surgery with late surgery and no treatment. Both studies reported no conflicts of interest. One study did not provide funding information, the other reported national and regional science programme and university funding.

**Key results and reliability of the evidence**
Evidence on the effects of early surgery for Bell’s palsy is of very-low reliability because the trials are small and have very serious limitations. We are unable to say whether recovery is better or worse after early surgery (before three months) than with no treatment, or after early surgery compared to later surgery.

One study reported side effects or complications of surgery. The reliability of this evidence was very low. Four participants who had surgery had some hearing loss, and three experienced tinnitus (ringing in the ears). The surgery group showed some loss of control of tear production in the eyes.

We have too little evidence to decide whether an operation would be helpful or harmful for people with early Bell’s palsy. There is unlikely to be further research into the role of an operation because Bell’s palsy usually recovers without treatment.

The evidence is up-to-date to March 2020.
### SUMMARY OF FINDINGS

#### Summary of findings 1. Surgery for Bell's palsy versus medical treatment (oral prednisolone) or no treatment

**Surgery compared with medical treatment for Bell's palsy**

**Patient or population:** Bell's palsy  
**Settings:** hospital attendance with idiopathic facial paralysis  
**Intervention:** early surgery (within 3 months from denervation onset)  
**Comparison:** no surgical treatment

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>Number of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complete recovery of facial nerve function at 12 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defined as: complete recovery HB grade I.</td>
<td></td>
<td></td>
<td>57 (2 RCTs)</td>
<td>⊗⊝⊝⊝ Very low o</td>
<td>We are uncertain whether surgery affects recovery of facial nerve function at 1 year as the certainty of evidence was very low.</td>
</tr>
<tr>
<td>Follow-up: 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st study: scale 0–5 (0 = no function, 5 = complete function for 3 facial muscles)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd study: HBGS (complete = HB I, fair = HB II–III and poor recovery = HB IV–VI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The 1st study (24 evaluable participants) compared surgery between the 2nd and 3rd week post onset to a no-treatment control group.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The 2nd study (33 evaluable participants) compared early surgery (within 2–3 months post paralysis onset) to no further treatment.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither study reported a statistically significant difference in recovery of facial nerve function between the surgical and non-surgical groups at 12 months (P &gt; 0.05). We calculated the RR for complete recovery at 12 months for the 2nd study; 1/18 participants in the early surgery group achieved complete recovery at 12 months compared to 1/15 in the no-treatment group (RR 0.76, 95% CI 0.05 to 11.11; P = 0.84).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Side effects and complications of treatment</strong></td>
<td>Measured by clinical assessment</td>
<td></td>
<td>57 (2 RCTs)</td>
<td>⊗⊝⊝⊝ Very low o</td>
<td>The numbers involved in the included studies were small. Statistical analysis was not possible.</td>
</tr>
<tr>
<td>Follow-up: 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The 1st study (24 evaluable participants) reported no complications of surgery.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The 2nd study (41 evaluable participants) reported postoperative complications such as sensorineural hearing loss, tinnitus, vertigo and dizziness. There were no reported surgical complications such as wound dehiscence, infection, bleeding and numbness. 4 participants had sensorineural hearing loss at high frequencies, with bone conduc-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

*Illustrative comparative risks are calculated using meta-analysis techniques when possible, otherwise by combining results presented in the studies. Relative effect is calculated as the ratio of the corresponding risks. Assumed risk is calculated as the ratio of the corresponding risks. Corresponding risk is calculated as the ratio of the corresponding risks.*
Surgical interventions for the early management of Bell’s palsy (Review)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: confidence interval; HB: House-Brackmann; HBGS: House-Brackmann grading system; RCT: randomised controlled trial; RR: risk ratio.

GRADE Working Group grades of evidence

High certainty: further research is very unlikely to change our confidence in the estimate of effect.

Moderate certainty: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low certainty: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low certainty: we are very uncertain about the estimate.

↓

Downgraded twice because of very serious limitations in study design: 1 study did not report the method of randomisation. Both studies had unclear allocation concealment. Blinding of participants was not possible and neither study blinded outcome assessors. 1 study did not follow up large numbers of participants. We further downgraded the evidence for imprecision as there were small numbers of participants in both studies. 1 study also reported the primary and secondary facial nerve recovery outcomes of this review indirectly, which warranted further downgrading for indirectness.

Summary of findings 2. Early surgery versus late surgery for Bell’s palsy

Early surgery compared with late surgery for Bell’s palsy

Patient or population: people with Bell’s palsy

Settings: hospital attendance with idiopathic facial paralysis

Intervention: early surgery (within 3 months from denervation onset)

Comparison: late surgery (later than 3 months from denervation onset)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumed risk</td>
<td>Corresponding risk</td>
<td>RR 0.47 (0.03 to 6.60)</td>
<td>26 (1 RCT)</td>
<td>very low</td>
<td></td>
</tr>
<tr>
<td>Late surgery</td>
<td>Early surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Complete recovery of facial nerve function at 12 months

Defined as: complete recovery HB grade I.

Follow-up: 12 months

Within the 1 study (33 evaluable participants), 18 undergoing surgery within 2–3 months of denervation (early surgery) were compared with 8 participants undergoing surgery > 3 months from denervation onset (late surgery).

There was no statistically significant difference in recovery of facial nerve function between the early surgical and late surgical groups at 12 months (P > 0.05). We calculated the RR for complete recovery at 12 months; 1/18 participants in the early

The evidence for recovery of facial nerve function with early vs late surgery was uncertain.
1st study: scale 0–5 (0 = no function, 5 = complete function for 3 facial muscles)

2nd study: HBGS (complete = HB I, fair = HB II–III, and poor recovery = HB IV–VI)

Surgery group achieved complete recovery at 12 months compared to 1/8 in the late surgery group.

**Side effects and complications of treatment**

Measured by clinical assessment

Follow-up: 12 months

The 1 study (41 evaluable participants) reported postoperative complications such as sensorineural hearing loss, tinnitus, vertigo and dizziness. There were no reported surgical complications such as wound dehiscence, infection, bleeding and numbness. 4 participants had sensorineural hearing loss at high frequencies, with bone conduction thresholds ranging from 35 dB to 50 dB at 4000 Hz. 3 participants reported tinnitus. There were no reported cases of postoperative vertigo or major labyrinthitis.

| Not reported | 26 (1 RCT) | Unable to rate certainty of evidence | The trial that compared early and late surgery did not report adverse events separately by timing of surgery. |

*The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

**CI:** confidence interval; **HB:** House-Brackmann; **HBGS:** House-Brackmann grading system; **RCT:** randomised controlled trial; **RR:** risk ratio

**GRADE Working Group grades of evidence**

**High certainty:** further research is very unlikely to change our confidence in the estimate of effect.

**Moderate certainty:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low certainty:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low certainty:** we are very uncertain about the estimate.

*Downgraded three times: twice because of very serious limitations in study design (high risk of bias in most domains) and once for imprecision from small-study size (26 participants).
BACKGROUND

Description of the condition

Bell’s palsy is an acute paralysis of one side of the face due to a lesion of the facial nerve. The condition is named after Sir Charles Bell, a Scottish surgeon (1774 to 1842). The cause is unknown and Bell’s palsy should only be used as a diagnosis in the absence of any other pathology. It was proposed in 1919 that the underlying pathology was that of a viral neuropathy (Antoni 1919). Herpes simplex virus has been suggested as the likely pathogen (McCormick 1972), and animal studies have suggested that reactivation of the virus may lead to demyelination of the nerve leading to reduced function (Adour 1975; Stjernquist-Desatnik 2006).

The condition affects 25 to 35 people per 100,000 population per year and is most common in the 30- to 45-year age group. It is also more common in pregnant women, people with diabetes or people with a respiratory tract infection (Theil 2001). Recovery in most people is good. One large review found that about 71% of people with Bell’s palsy will have normal function restored, 25% will have slight or mild sequelae and only 4% will have severe sequelae (Peitersen 2002). Contractures, facial disfigurement, with associated psychological difficulties, and facial pain remain the most common long-term problems (Morgenlander 1990).

Description of the intervention

Several studies have investigated which population might benefit most from surgery. In addition to simple clinical assessment of disease using the House-Brackmann grading system (HBGS) or similar, many studies have tried to assess the electrical function of the facial nerve. Electroneurography (ENG) is the most popular technique employed (Esslén 1977; Fisch 1984). In this, the degree of muscle response to an electrically evoked stimulus is assessed. Esslén 1977 and Fisch 1984 found that when 95% of the nerve had degenerated, the patient had a 50% chance of a poor outcome (less than 50% chance of recovery to House-Brackmann (HB) I or II) and would potentially benefit from surgical intervention (Stillman 1992).

How the intervention might work

Although Bell’s palsy is common, in the absence of an established aetiology, treatment continues to be based upon the presumed pathophysiology of swelling and entrapment of the nerve. Two double-blind randomised controlled studies (RCTs) have shown that early treatment with prednisolone but not aciclovir significantly improves the chances of complete recovery to 94% at nine months (Sullivan 2007; Engstrom 2008). Recent Cochrane Reviews on the use of corticosteroids (Madhok 2016) and antivirals (Gagyor 2019) in Bell’s palsy are consistent with these findings.

As the proposed pathophysiology involves entrapment of the nerve, this has led some surgeons to suggest that surgical decompression of the nerve is a suitable management option. The first recorded attempt at surgical decompression of the facial nerve for Bell’s palsy was in 1932 (Ballance 1932). Ballance 1932 recommended slitting the sheath in the distal descending segment of the nerve. This was consistent with theories of the site of the lesion at that time. Over the next few decades, the proposed site for operation migrated from the distal 1 cm at the stylomastoid foramen (in Ballance 1932) to the entrance of the Fallopian canal medially (Fisch 1972). In Fisch 1972, there were intraoperative electromyography (EMG) changes and oedematous swelling at this point proximal to the geniculate ganglion in up to 94% of participants. These findings have been further supported by magnetic resonance imaging (MRI) studies; there was gadolinium enhancement distal to the internal auditory canal across the labyrinthine and geniculate segments correlating to the greatest area of inflammation and oedema (Schwaber 1990). This study reported enhanced in these segments in both acute (less than four weeks of onset) and chronic (more than four weeks of onset) cases of Bell’s palsy, with enhancement seen in every case up to four months following onset (Schwaber 1990). Furthermore, Yanagihara 2001 stated that oedematous swelling of the facial nerve was commonly observed in patients operated on up to three months from the onset of paralysis.

Based on these findings, it would be logical to focus facial nerve decompression on the most affected segments. While multiple approaches exist, Fisch 1972 and May 1972 were among the first to outline the two most common surgical approaches: the transmastoid approach and the middle fossa craniotomy (MFC). Transmastoid decompression generally provides greater access to the tympanic or mastoid segments of the facial nerve and can be further subdivided into different approaches, such as the retro-auricular approach (da Franca Pereira 2016). As there is potential for the incus to be excised and replaced, hearing (conductive or sensorineural) loss is a common postoperative complication (Palombo 2012). May 1984 suggested that a transmastoid approach to decompression of the labyrinthine segment was of benefit. Conversely, MFCs allow exposure of the labyrinthine segment, geniculate ganglion and tympanic segment (da Franca Pereira 2016). However, it is more challenging to locate anatomical landmarks, and the procedure is quite invasive as it may require temporal lobe retraction (da Franca Pereira 2016). MFCs are less likely to damage hearing but have a much more extensive list of possible complications such as infection, seizures and cerebrospinal fluid (CSF) leakage (Hato 2012). Few prospective multi-centre trials evaluating the efficacy of MFC decompression exist (Fisch 1981; Gantz 1999); given the significant risks and complications associated with both procedures and the lack of evidence demonstrating efficacy, it is difficult for either approach to be routinely recommended in current practice. Two further studies published around the same time gave evidence both for (Giancarlo 1970) and against (McNeill 1974) operation. As current literature continues to remain divided, it is prudent to critically evaluate surgical approaches based on which segments are decompressed when considering their efficacy.

Likewise, there is no consensus on the ideal timing of facial nerve decompression surgery. Early reports from the 1930s advocated decompression within three months of paralysis onset, whereas some later reports suggest the procedure be carried out within 10 days (Adour 1982). Several non-randomised studies report conflicting results in regards to optimal timing. Giancarlo 1970 and Yanagihara 2001 both report improved recovery in a proportion of the participants operated on within three months. Other studies such as Fisch 1981 advocate surgical decompression within 24 hours of when nerve degeneration reaches 90% to 94% within one to 21 days after onset of paralysis. Due to the favourable outcome of the condition without surgical treatment and with medical management, the uncertainty surrounding both the affected facial nerve segment and timing of the operation, and also the potential for damage to the facial nerve and other ear structures during surgical decompression...
surgery, there has been a continued debate as to whether surgery has a role in the management of Bell’s palsy (Friedman 2000; Adour 2002).

**Why it is important to do this review**

Despite the debate on different surgical approaches and timing, there is a paucity of high-certainty evidence regarding facial nerve decompression surgery for acute Bell’s palsy. Few large studies have been carried out of these. Of those, one study convinced many surgeons that surgery did not have a place in the management of Bell’s palsy (May 1985). More recently, Gantz 1999 found that when selected using ENOG, people who would have had a bad outcome as predicted by ENOG had a better outcome if managed surgically compared with those who were not. Currently, in most people with Bell’s palsy the condition is managed medically with corticosteroids with or without aciclovir, as discussed above. Surgery, certainly in the UK, is rarely undertaken (Sullivan 2007).

This is an update of a review first published in 2011. We updated the review to incorporate new evidence.

**OBJECTIVES**

To assess the effects of surgery in the early management of Bell’s palsy.

**METHODS**

**Criteria for considering studies for this review**

**Types of studies**

We assessed RCTs and quasi-RCTs. Quasi-RCTs are trials in which allocation of participants is partly systematic (e.g. by medical record number or by alternation). We included quasi-RCTs because of the paucity of studies in this area. We reported other studies, including observational studies, in the Discussion. We have considered adverse events taking observational studies into account in the Discussion. We also planned to discuss economic issues in the ‘Discussion’, but there was insufficient information in the included studies.

We considered published or unpublished studies in any language.

**Types of participants**

We included adults or children who presented with an idiopathic facial palsy which was diagnosed as Bell’s palsy. We excluded those who were diagnosed as having herpes zoster, those who had a traumatic aetiology, or other identified aetiology from the review. This included any cases of recurrent and familial Bell’s palsy or Melkersson-Rosenthal syndrome.

**Types of interventions**

We included any surgical intervention carried out for Bell’s palsy. We compared these interventions to no treatment, later treatment (beyond three months), sham treatment, other surgical treatments or medical treatment. Where concomitant treatment was given, this was the same in both treatment and comparator groups. For this review, ‘early’ surgical intervention was defined as operations occurring within three months of facial paralysis onset, and ‘late’ surgical intervention was defined as operations occurring later than three months after facial paralysis onset. A time frame of three months or less was decided upon as throughout literature, there has been inconsistent reporting of the effects of decompression surgery in relation to specific time points, as explored in How the intervention might work. Given the fact that studies have reported cases with facial nerve inflammation as late as three to four months post onset (via gadolinium enhancement on MRI and intraoperative oedematous swelling) and that no high-quality evidence currently exists to discredit operating within a shorter timeframe, we decided studies investigating facial nerve decompression within three months should not be excluded from our analysis. Furthermore, the variability in facial nerve segments decompressed and surgical approaches used makes it more difficult to evaluate the efficacy of certain time points for establishing inclusion criteria. Therefore, we excluded studies where participants were operated on during the later stages of their recovery (after more than three months). In cases where studies reported outcomes from both early and late operations, we included the early treatment arm as the intervention group and the late treatment arm as a comparator.

Should future studies report on surgery within earlier time frames (immediately after onset, within two weeks of onset, etc.), these studies will be considered for inclusion of subsequent review updates.

**Types of outcome measures**

We did not use measurement of the review outcomes to determine study eligibility.

**Primary outcomes**

- Degree of recovery of facial nerve function and resolution of symptoms at 12 months as measured using the House-Brackmann scale, the Sunnybrook scale, the Yanagihara scale or other similar scale. Only complete recovery at 12 months was included, as surgical decompression is not typically first-line treatment in Bell’s palsy, and the risks involved with surgical decompression outweigh the benefits if recovery is not complete. Resolution of symptoms can be defined as ‘complete’ recovery. At present, the review includes only dichotomous data, but should any continuous data be reported in future studies, it will be appropriate to include it under this outcome as well, as a part of ‘degree of recovery of facial nerve function’.

**Secondary outcomes**

- Complete recovery at three and six months. Complete recovery was defined as HB I. As above, the review authors decided to only report on complete recovery as surgical decompression is not typically first-line treatment in Bell’s palsy, and the risks involved with surgical decompression outweigh the benefits if recovery is not complete.

- Synkinesis and contracture at 12 months.

- Psychosocial outcomes at 12 months. The preferred measures/scales for reporting on psychosocial outcomes vary depending on patient population. For adult populations, the preferred measures would be in relation to both mood and appearance: the HADS (Hospital Anxiety and Depression Scale) for adults for mood; and the DAS 24 (Derriford Appearance Scale) for adults for appearance, before and after. If these scales are not reported, any other scale in relation to those characteristics would be acceptable. For children we will report the Pediatric Quality of Life Inventory or other suitable validated scale.

- Side effects and complications of treatment.
Search methods for identification of studies

Electronic searches
The Cochrane Neuromuscular Information Specialist searched the following databases for this update:

- the Cochrane Neuromuscular Specialised Register via the Cochrane Register of Studies (CRS-Web; 20 March 2020; Appendix 1);
- the Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Register of Studies (CRS-Web; 20 March 2020; Appendix 2);
- MEDLINE OvidSP (1946 to 19 March 2020; Appendix 3);
- Embase OvidSP (1974 to 2020 week 11; Appendix 4);
- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (20 March 2020; www.clinicaltrials.gov; Appendix 1); and

Searching other resources
We reviewed the bibliographies of all trials identified.

We performed a handsearch of the following conference abstracts:

- American Academy of Otolaryngology – Head and Neck Surgery Annual Meeting 2006 and 2007; and
- British Academic Conference in Otolaryngology and ENT Expo Birmingham, UK, both 2006.

Data collection and analysis

Selection of studies
Two review authors (DW and KA) independently reviewed titles and abstracts identified by the search strategy. For a 2020 top-up search, an Information Specialist performed an initial screen of titles and abstracts for RCTs. The review authors obtained the full text of all relevant studies and assessed them independently. Three review authors (DW, KA and IM) assessed whether each trial met the inclusion criteria. They resolved any disagreements by discussion with the lead author (IM) where required. We included a PRISMA flow chart to illustrate the study selection process (Figure 1; Moher 2015).
Data extraction and management

For the previous version of the review, two review authors (DW and KA) independently extracted the data and entered them onto a specifically designed form. The extracted data included characteristics of study participants, methods, interventions used, outcomes and results. Two review authors (DW and KA) entered the data for Mechelse 1971 into Review Manager 5. For the update, IM and an independent systematic reviewer independently extracted data from the newly included study and documented differences between the extractions. IM entered data into Review Manager 5, checked outcome data entry and spot-checked characteristics of studies data. Differences were resolved via discussion between IM and a third review author (KM) who had not extracted data. For the update, we extracted additional characteristics of studies data (e.g. funding sources and conflicts of interest). For studies written in languages other than English, translators provided us with a translation for us to perform a data extraction. The lead author (IM) checked extracted numerical data against the trial report, as far as possible. For future updates, we will adapt the data extraction form as necessary and pilot it among review authors before use.

Assessment of risk of bias in included studies

Three review authors (IM, DW and KM) independently assessed the risk of bias across seven ‘Risk of bias’ domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and ‘other sources of bias’. We judged each study in relation to each domain, as at a high, low or unclear risk of bias. We resolved any disagreement by discussion with the lead author (IM). When considering treatment effects, three review authors (IM, DW and KM) considered the risk of bias for the studies that contributed to that outcome. The review authors made summary assessments for the ‘Risk of bias’ for each important outcome (across domains), according to the criteria outlined in Table 1 (Higgins 2011).
Measures of treatment effect
There were insufficient studies to enable any statistical analysis. If in future updates statistical analysis is possible, we will enter data into the Review Manager 5 when available and analyse the data using the standard statistical methods (Review Manager 2014). For continuously measured outcomes, we will use means to obtain mean differences (MDs) with 95% confidence intervals (CIs), or standardised mean differences for results across studies with outcomes that are conceptually the same but measured in different ways. For dichotomous outcome data, we will estimate the pooled risk ratio (RR) with 95% CI within Review Manager 5. We will calculate the number needed to treat for an additional beneficial effect (NNTB) or for an additional harmful effect (NNTH) if possible.

Unit of analysis issues
Where multiple trial arms are reported in a single trial, we included only the treatment arms relevant to the review topic, but list them all in the Characteristics of included studies table. The Li 2016a study had multiple eligible arms.

We describe the approach to unit-of-analysis issues in meta-analysis in Appendix 3.

Dealing with missing data
The review authors did not contact trial authors for additional data or to verify study characteristics.

Assessment of heterogeneity
We did not assess heterogeneity due to lack of data. See Appendix 3 for planned methods.

Assessment of reporting biases
There were insufficient studies to enable any statistical analyses or to allow use of funnel plots.

If searches had identified trial protocols, clinical trial registrations or abstracts indicating the existence of unpublished studies, we would have attempted to determine their status by contacting the investigators and would have reported on potential unpublished studies in the review.

We would also have considered the effects of reporting biases – whether of studies or selective outcome reporting – on our findings, during the GRADE assessment.

Data synthesis
We would have considered different types of surgery separately, other than minor variations in technique, such as within the retro-auricular/transmastoid. We considered different comparators separately, assessing the components of any comparator interventions (e.g. ‘usual care’ or ‘standard care’) to determine whether they were sufficiently similar for pooling to be feasible.

There were insufficient studies to enable any statistical analyses. We reported methods to be used if meta-analysis is possible in future in Appendix 3.

This review has a published protocol (Swan 2008). We have documented changes from the protocol in Differences between protocol and review.

Subgroup analysis and investigation of heterogeneity
We did not perform subgroup analyses or investigate heterogeneity for lack of data. See Appendix 3 for planned methods.

Sensitivity analysis
We did not perform sensitivity analyses for lack of data. See Appendix 3 for planned methods.

Summary of findings and assessment of the certainty of the evidence
We selected the following outcomes for presentation in ‘Summary of findings’ tables.

- Recovery of facial nerve function at 12 months.
- Side effects and complications of treatment.

We used Review Manager 5 to create a ‘Summary of findings’ table (Review Manager 2014).

We used the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness and publication bias) to assess the certainty of a body of evidence (studies that contributed data for the prespecified outcomes). We used methods and recommendations described in Chapters 11 and 12 of the Cochrane Handbook for Systematic Reviews of Interventions (Schünemann 2020a; Schünemann 2020b). We considered RCTs as high-certainty evidence if the five factors above were not present to any serious degree. We downgraded the certainty of evidence once if there were any individual GRADE considerations present to a serious degree and twice if very serious, to produce judgements of moderate, low or very low certainty. We justified all decisions to downgrade or upgrade the certainty of evidence using footnotes. Table 1 outlines our approach for summary assessments of the risk of bias for each important outcome (across domains) within and across studies (Higgins 2011).

RESULTS

Description of studies

Results of the search
The new searches retrieved 828 records, reduced to 703 after removal of duplicates. After three review authors (IM, KM, DW) screened the results of the updated search, we excluded 687 records and shortlisted 18 papers for full-text assessment, of which we subsequently excluded 17. We included one newly identified study and one study described in full-text in awaiting classification as we identified it from an updated search when the review was nearing completion (Hato 2015).

The previous version of the review included two studies (Adour 1971; Mechelse 1971). After reassessment of the allocation methods outlined in Adour 1971, we excluded this study based on non-random selection of the control groups, since the control group included participants who refused surgery. This is consistent with other excluded studies. Both Li 2016a and Mechelse 1971 excluded people who refused their allocated interventions; therefore, we deemed them eligible for inclusion.

Two trials met the inclusion criteria (Mechelse 1971; Li 2016a). See Figure 1 for a PRISMA flow chart summarising the study selection process.
Included studies

Populations

Mechelse 1971 randomly allocated 25 participants into surgical or non-surgical (control) groups, and after one participant refused surgery, had a total of 24 participants. Investigators selected participants based on a complete facial palsy and EMG showing no voluntary control of motor unit or a minimal applied current to evoke a motor response on the affected side 2.5 times that required on the unaffected side. These responses needed to be confirmed on two occasions a few days apart. The trial authors excluded participants with signs or symptoms suggestive of a cause for their facial paralysis other than Bell’s palsy; those with incomplete facial paralysis; and those with an abnormality on ear, nose and throat examination, skull X-ray or in blood and urine tests. This study reported the age and sex of the participants, and the side of the face affected, but did not report if there were statistically significant differences between the two groups at baseline. Two centres participated in the study: University Hospital, Leiden and Municipal Hospital, The Hague, Netherlands. There was no sponsorship noted. There were no declared conflicts of interest among the main researchers in the study (see Characteristics of included studies table).

Li 2016a included 41 participants. The trial quasi-randomised 53 participants via alternation into surgical or observational groups. Twelve people who initially refused the allocated intervention were excluded from the study. One of the treatment arms included participants operated on more than three months after the onset of denervation, and we included this group, which was not eligible under our selection criteria as an intervention, as a second comparator. The inclusion criteria of Li 2016a were based on a combination of patient demographics, previous treatment received, a clinical examination of facial nerve function and EMG examinations. More specifically, facial nerve function assessment involved a clinical examination using the HBGS to assess the severity of paralysis. The EMG assessment measured the percentage degeneration of compound muscle action potential (CMAP) amplitude and reduced voluntary EMG activity after two months. The investigators included participants if they met all of the outlined criteria: adults over 16 years of age, total paralysis at onset, previous course of oral prednisolone 1 mg/kg within 72 hours, disease duration longer than two months, the worst score of facial palsy was lower than HB V paralysis after two months, more than 95% degeneration of the CMAP amplitude with reduced voluntary EMG activity compared to the normal side after two months and no systemic disease such as severe diabetes. Additionally, the study excluded anyone thought to have facial paralysis of any other cause than Bell’s palsy. Via alternation, three groups were selected: participants who underwent decompression surgery two to three months after the onset of denervation (early surgical intervention group), participants who underwent decompression surgery more than three months after denervation onset (late surgical intervention group) and those managed with follow-up observation (see Types of interventions). This study reported the age, sex and side of palsy of the participants. Other than the numbers, the study stated no differences in age and sex distributions between the groups. The study was supported by Key Projects in the National Science & Technology Pillar Program of China during the Twelfth Five-Year Plan Period, The Natural Science Foundation of Shaanxi Province and Xi’an Jiaotong University basic scientific research operation expenses. There were no declared conflicts of interest among the main researchers in the study.

Intervention: surgical procedure

The participants in the surgical groups of both studies underwent facial nerve decompression using a retro-auricular/transmastoid approach.

Intervention: comparator

All the participants in Li 2016a received a course of oral prednisolone 1 mg/kg within 72 hours of palsy onset, before the allocated intervention. Thereafter, the control group received no further treatment (no treatment group). The surgical participants were subdivided based on the timing of their operations. We included the group of participants who were operated upon later than three months as a comparator as these participants were not eligible for the review. None of the participants in Mechelse 1971 received treatment prior to the allocated intervention, and therefore the control (non-surgical) group received no further treatment.

Outcomes

Mechelse 1971 used a scale of 0 to 5 to assess outcome (0 = no function, 5 = complete function) to assess frontalis, orbicularis oculi and orbicularis oris muscles. The authors did not stipulate what value on the clinical scale represented a satisfactory recovery, or what other values on the scale corresponded to in regards to facial muscle function. Mechelse 1971 predated the development of the HBGS.

Li 2016a used the HBGS to assess physician-based outcomes and the Facial Clinimetric Evaluation (FaCE) scale to assess patient-based outcomes. Using the HBGS, the recovery of facial function was categorised as complete, fair or poor. Grade I is equivalent to complete facial function recovery; Grades II and III indicate a fair recovery, classified as slight to moderate facial weakness and synkinesis; grades IV, V and VI are classified as poor recovery, indicating severe facial weakness, synkinesis, contracture, spasm, or a combination of these. The FaCE scale involves 15 statements, each using a 5-item Likert scale, where 1 corresponded to the lowest function and 5 corresponded to the highest. For each statement, each participant chose the most appropriate response regarding their impression of function. The statements were then grouped into six separate domains: social function, facial movement, facial comfort, oral function, eye comfort and lacrimal control. A total score was generated from all these domains, and a final value ranging from 0 (worst) to 100 (best) was then calculated.

Excluded studies

Upon review of the allocation methods outlined in Adour 1971, we excluded it along with several observational studies based on non-random selection of the control groups. Giancarlo 1970; Adour 1971; McNeill 1974; Fisch 1981; Brown 1982; May 1984; Aoyagi 1988; Gu 1994; Gantz 1999; Yanagihara 2001; Li 2015; and Kim 2016 all had control groups that were self-selected, in that people who refused surgery became the control group. We did not consider, therefore, that comparisons made between the surgical groups of participants and the control groups were valid, as the reasons people refused surgery may have been relevant to their outcomes. For example, in Yanagihara 2001, younger participants opted for surgery whereas older participants declined (see Characteristics of excluded studies table).
of excluded studies table). We excluded Zhu 2016 and Berania 2018 based on type of study, as they were both non-randomised retrospective reviews. We excluded Li 2016b and Ying 2011 as both studies included populations diagnosed with ‘hemi-facial spasm’ and did not report if participants were included with Bell’s palsy, and did not report outcomes separately by aetiology.

We excluded two non-English studies once we had reviewed translations. We assessed one German study, Heckmann 2012, using a proprietary translation application and excluded it, as the study itself was a literature review and primarily focused on the use of steroids rather than surgical interventions. We excluded one Chinese study based on its translation, due to a lack of control group (Liu 2013).

Studies awaiting classification

We listed one study as awaiting classification. We identified it from trials registries in a final search in 2020 and we are awaiting author correspondence to determine eligibility and status (whether completed or ongoing). See Characteristics of studies awaiting classification table.

Risk of bias in included studies

Both trials were at a high overall risk of bias. Li 2016a was at high risk in all domains except ‘other bias’. Mechelse 1971 lacked blinding and reporting was selective.

We summarised the review authors’ judgements about each ‘Risk of bias’ item for each included study in Figure 2.
Figure 2. Risk of bias summary: review authors’ judgements about each risk of bias item for each included study. Red = high risk of bias, yellow = unclear risk of bias and green = low risk of bias.

Allocation

Mechelse 1971 randomly allocated 25 participants into surgical or non-surgical groups. The report stated, “in both hospitals these participants were entered on a list, previously prepared by the statistical department, University of Leiden (head, Mr H. De Jonge), which randomly allocated them to surgical treatment or to a non-surgical control group.” One participant declined surgery and was removed from the study, totalling 24. Therefore, we judged the risk of selection bias from sequence generation as low.

Li 2016a allocated 53 participants through a quasi-randomised method via alternation. Participants allocated to the surgical group were then further subdivided into early and late surgical intervention groups. The report stated, “total random distribution could not be performed because some participants refused surgical
intervention.” Eleven participants allocated to the observational group refused, and one participant allocated to the surgical group refused. These 12 participants were removed from the study, resulting in 41 evaluable participants. We judged the risk of selection bias from sequence generation as high.

There was no allocation concealment in Li 2016a, as investigators used alternate assignment. Mechelse 1971 did not mention any method of allocation concealment. Therefore, we judged the risk of bias as unclear in both studies.

Blinding

It was not possible in either study to blind participants to the surgical or non-surgical intervention. Although blinding of outcome assessment was possible in surgical trials using an independent assessor, neither study report commented on whether blinding of the investigator or outcome assessor was performed. Therefore, we judged the risk of performance and detection bias for both studies as unclear.

Incomplete outcome data

Mechelse 1971 claimed to have followed all participants for one year with no losses (quote: “All patients were followed clinically and electromyographically for at least a year”). Therefore, attrition bias was judged as low.

Li 2016a removed 12 participants from the study upon allocation due to refusal of their allocated intervention (one surgical, 11 observational). Furthermore, one participant in the early surgical group (1/18) and two participants in the no treatment group (2/15) were lost to follow-up before the end of the 12-month study. Twenty-one of 41 participants completed the FaCE questionnaire at the end of follow-up. Therefore, we judged the risk of attrition bias as high.

Selective reporting

Mechelse 1971 report statistical comparison between the groups but did not report the statistical methods used. All prespecified outcomes were indirectly reported in this study as results ‘after the second week’, and results for specific time points were all inferred. Therefore, we judged reporting bias to be high.

Li 2016a reported all methods of statistical analysis, using SPSS 20. The report stated, “all tests were conducted at the 5% level of significance.” The “Fisher exact test was used for differences between proportions and non-parametric tests were used due to the non-normal distribution of scores.” All prespecified outcomes were acknowledged and fully reported on in this study, therefore, we judged reporting bias as low.

Other potential sources of bias

We identified no other sources of bias in either study.

Effects of interventions

See: Summary of findings 1 Surgery for Bell’s palsy versus medical treatment (oral prednisolone) or no treatment; Summary of findings 2 Early surgery versus late surgery for Bell’s palsy

Surgery versus control (no further treatment)

See Summary of findings 1.

Primary outcome

Degree of recovery of facial nerve function and resolution of symptoms at 12 months

Mechelse 1971 randomised 11 participants to the surgical group and 13 to the non-surgical control group (24 in total). They used a scale of 0 to 5 to assess outcome in three facial muscles (0 = no function, 5 = complete function) and reported data for each participant, but did not report a measure of overall recovery. All participants were followed up over a period of “at least one year.” They assessed participants daily during the first week and every other day during the following two weeks. For the next two months, they examined participants either weekly or fortnightly. After this period of three months, they adjusted each participant’s rate of follow-up ‘in relation to the rate of recovery,’ with total durations ranging from one to three years total. Participants ranged in recovery from 2 to 5 on their scale and, after the follow-up period of at least one year, there was no difference between the surgical and non-surgical group (P = 0.9). All results were reported ‘after the second week’; and were, therefore, indirect. No definitions were provided for facial muscle function scores of 1 to 4 on the scale used. There was no change in the degree of recovery of facial nerve function beyond that time point, and presumably at the 12-month time point as well.

Li 2016a used the HBGS to assess the degree of recovery in facial function at 12 months and classified recovery into complete (HB I), fair (HB II and III) and poor (HB IV to VI). They randomised 53 participants, 12 of whom refused their allocated treatment; 26 participants were allocated into one of two surgical intervention groups, and 15 to the control group, which received no further treatment. Eighteen participants in the ‘early’ surgical group underwent surgery between two and three months (early surgical intervention group), and eight underwent surgery more than three months following denervation (late surgical intervention group).

The study authors of Li 2016a reported that there were no significant differences in the “distribution of participants with complete, fair, or poor recovery at the end of the 12-month follow-up period” (P > 0.05). There were no other statistical comparisons between intervention groups. We were able to calculate an RR for complete recovery at 12 months, as there were numerical data for each participant group, at each stage of follow-up; one of 18 participants in the early surgery group achieved complete recovery at 12 months, compared to one of 15 participants in the no treatment group (RR 0.76, 95% CI 0.05 to 11.11; P = 0.84; very low-certainty evidence; Analysis 1.1). Only complete recovery was reported, as this is thought to have the greatest clinical applicability and significance and was considered full resolution of symptoms.

Secondary outcomes

For similar reasons as the primary outcome, all secondary outcome measure evidence was downgraded to ‘very low’ certainty of evidence. We downgraded evidence twice for study design limitations and imprecision of results mainly due to small participant populations resulting in insufficient power to detect statistically significant effects and high sources of bias in both studies. Indirectness of evidence reporting at each specified time frame resulted in the evidence being downgraded a third time (Overall completeness and applicability of evidence under ‘Certainty of the evidence’).
Complete recovery at three and six months

In Mechelse 1971, EMG follow-up showed that onset of recovery did not occur before the third month in either the surgical or the non-surgical group (24 participants). Facial nerve outcome scores were not directly documented for three and six months; however, the trial authors stated that from the second week onwards, there were no differences (see Primary outcomes). This outcome was not reported in full with very little supporting statistical analysis.

Li 2016a documented levels of recovery for all three groups at three and six months. There were no statistical comparisons in relation to complete recovery at both three and six months. We were able to calculate the RRs for complete recovery at three and six months for early surgery versus no treatment. One of 18 participants in the early surgery group compared to one of 15 participants in the no-treatment group reached complete recovery at three months (RR 0.83, 95% CI 0.06 to 12.22; P = 0.89; very low-certainty evidence; Analysis 1.2). There was no change in proportions of participants with complete recovery at the six-month interval in either group; one of 18 participants in the early surgery group compared to one of 15 participants in the no-treatment group reached complete recovery at six months (RR 0.83, 95% CI 0.06 to 12.22; P = 0.89; very low-certainty evidence; Analysis 1.3).

Synkinesis and contracture at 12 months

In Mechelse 1971, we calculated the RRs for synkinesis and contracture for comparisons between surgical and non-surgical groups as these statistical analyses were not provided in the text. Twelve of 13 participants in the non-surgical control group and nine of 11 participants in the surgical group reported synkinesis (RR 0.89, 95% CI 0.64 to 1.22; P = 0.46; very low-certainty evidence; Analysis 1.4). Six of 13 in the non-surgical group reported contracture, compared to eight of 11 in the surgical group (RR 1.58, 95% CI 0.79 to 3.14; P = 0.20; very low-certainty evidence; Analysis 1.5). These data were dichotomous, reported as a + (signifying yes) and a – (signifying no) for whether the participant experienced these complications during recovery.

Li 2016a broadly evaluated synkinesis and contracture using the FaCE score, which includes social function, facial movement, facial comfort, oral function, eye comfort, lacrimal control and an overall score. Twenty-one of 41 participants overall completed the FaCE questionnaire at the end of follow-up (11/18 in the early surgical group, 0/8 in the late surgical group and 10/13 in the no treatment group). All FaCE questionnaire results were reported as median (interquartile (IQR)) with P values. Within the early surgical treatment group, the study authors report a statistically significant improvement in the total FaCE score from the time of presentation (26.7, IQR 23.3 to 30.0) to the end of follow-up (61.7, IQR 60.0 to 65.0; P = 0.003). The study authors also reported a statistically significant improvement in the FaCE scores (P < 0.05) in all domains, with the exception of a statistically significant decrease in lacrimal control between first presentation (50.0, IQR 50.0 to 50.0) and the last follow-up visit (25.0, IQR 25.0 to 50.0) of the early surgical treatment group (P = 0.021). Within the no treatment group, the study authors reported a statistically significant improvement in the total FaCE score from first presentation (26.7, IQR 25.0 to 32.5) to the end of follow-up (60.0, IQR 58.3 to 66.7; P = 0.001). Similarly, all domains, excluding lacrimal control, showed a statistically significant improvement from presentation to end of follow-up (P < 0.05); however, the decrease in lacrimal control from presentation (25.0, IQR 25.0 to 50.0) to end of follow-up (50.0, IQR 25.0 to 62.5) was not statistically significant within the no treatment group (P = 0.102). When comparing FaCE scores of the surgical treatment (25.0, IQR 25.0 to 50.0) and no treatment groups (50.0, IQR 25.0 to 62.5) at the end of follow-up there was a statistically significant decrease in lacrimal control (P = 0.023). The trial authors reported no other statistically significant differences in the total FaCE score or the five remaining individual domains when comparing the two groups.

Psychosocial outcomes at 12 months

Neither study reported psychosocial outcomes.

Side effects and complications of treatment

Mechelse 1971 reported no complications of surgery, such as wound dehiscence, infection, bleeding and numbness.

Li 2016a reported the occurrence of postoperative complications such as sensorineural hearing loss, tinnitus, vertigo and dizziness. There were no reported surgical complications such as wound dehiscence, infection, bleeding and numbness. Four participants had sensorineural hearing loss at high frequencies, with bone conduction thresholds ranging from 35 dB to 50 dB at 4000 Hz. Three participants reported tinnitus. There were no reported cases of postoperative vertigo or major labyrinthitis. There was no comment on the timing of the operation in relation to these complications.

Early surgery versus late surgery

See Summary of findings 2.

In the multi-arm trial of Li 2016a, 18 participants in the 'early' surgical group underwent surgery between two and three months, and eight were allocated to a second group who underwent surgery more than three months following denervation ('late' surgical group).

See Risk of bias in included studies for summary of reasons for downgrading evidence to 'very low' certainty. More detailed explanation available in Overall completeness and applicability of evidence under 'Certainty of the evidence.'

Primary outcome

Degree of recovery of facial nerve function and resolution of symptoms at 12 months

Although Li 2016a reported outcomes for early surgery versus no treatment at 12 months, they did not report statistical comparisons between the early and late surgical groups. We were able to calculate the RR for complete recovery at 12 months for early surgery versus late surgery. One of 18 participants in the early surgery group compared to one of eight participants in the late surgery group reached complete recovery at 12 months (RR 0.47, 95% CI 0.03 to 6.60; P = 0.58; very low-certainty evidence; Analysis 2.1).

Secondary outcomes

Complete recovery at three and six months

Li 2016a did not report any further statistical comparisons between the early surgery and late surgery groups for complete recovery at both three and six months. We calculated RRs for complete recovery at three and six months for early surgery versus late surgery. In the early surgery group, one of 18 participants reached...
complete recovery at three months compared to one of eight participants in the late surgery group (RR 0.44, 95% CI 0.03 to 6.25; P = 0.55; very low-certainty evidence; Analysis 2.2). At the six-month interval, proportions were again unchanged; one of 18 participants reached complete recovery in the early surgical group compared to one of eight in the late surgical group (RR 0.44, 95% CI 0.03 to 6.25; P = 0.55; very low-certainty evidence; Analysis 2.3).

Sykinesis and contracture at 12 months
At the end of follow-up, none of the eight participants in the late surgical group provided a FaCE score, so comparison was not possible.

Psychosocial outcomes at 12 months
Li 2016a did not report psychosocial outcomes.

Side effects and complications of treatment
Li 2016a described complications of surgery as described under ‘Surgery versus control (no further treatment)’. There was no comment on the timing of the operation in relation to complications.

DISCUSSION

Summary of main results
The evidence from the one RCT and one quasi-RCT included in our review do not support or exclude benefit from early surgical intervention for the management of Bell’s palsy, as the certainty of the evidence was very low. The two included trials had relatively small numbers of participants: 24 in Mechelse 1971 and 41 in Li 2016a, and had serious study limitations with a high risk of bias in multiple domains. One trial compared early versus late surgical interventions, but we were unable to draw conclusions.

With regard to complications of surgery, Li 2016a reported on the incidence of 35 dB to 50 dB sensorineural hearing loss at 4000 Hz and tinnitus. Lacrimal control was decreased in the early surgical and no treatment groups at the end of the 12-month follow-up period; however, this only reached significance for the early surgery group. There were no complications reported in Mechelse 1971. The included studies did not report any other complications of surgery. We could not draw conclusions from the limited data.

Overall completeness and applicability of evidence
We found no evidence of any certainty on early surgical intervention for Bell’s palsy. Lack of data prevented meta-analysis. This review focused solely on early surgical interventions for Bell’s palsy occurring within three months, limiting the scope of evidence. Future review updates should consider expanding to include all time points to potentially facilitate a meta-analysis, with separate statistical tests evaluating efficacy at certain time points and if applicable, different surgical approaches as well.

Current evidence is incomplete and insufficient to make recommendations for clinical practice, therefore limiting our ability to address the effects of surgery in the early management of Bell’s palsy. Both studies used surgical approaches that did not decompress the segments of the facial nerve thought to be the site of injury in Bell’s palsy. It is difficult to draw conclusions based on this, as other types of surgical approaches such as MFC, which typically provides better access to the affected segments, carries a higher risk of complications due to its invasiveness, compared to the retro-auricular approach described in the two included studies. Furthermore, this side effect alone makes it difficult to justify undergoing surgical decompression if the recovery outcome is not complete. These factors greatly influence the ability of this review to address any of the specified objectives, as the data cannot be applied to current practice. This is further complicated by the insufficient data on the rates of complete recovery following surgical intervention. Ideally, future reviews should focus on evaluating the efficacy of facial nerve decompression at the affected segment at several time points, to create much clearer clinical guidance and prognostic information.

Certainty of the evidence
Considered individually, the trials had too few participants for sufficient statistical power to detect the magnitude of effect that might plausibly be expected.

Both included trials had very serious study limitations. Mechelse 1971 stated that the participants were randomised; however, the methods of randomisation were not clearly stated. Li 2016a stated that participants were quasi-randomised via alternation; however, there was no mention of how the order in which alternation occurred was determined. For both studies, allocation concealment was at unclear risk of bias and it was unclear if outcome assessors were blinded. Li 2016a was at risk of attrition bias. Mechelse 1971 appeared to have incomplete reporting of results; as previously outlined, Mechelse 1971 did not fully report direct results of the outcomes at the specified time points. These serious limitations throughout both studies prompted us to downgrade the evidence twice.

With regards to study size and imprecision, Mechelse 1971 had a small number of participants, although the numbers between the intervention groups were similar. Li 2016a had small, uneven numbers in each treatment arm of the study. The RRs demonstrated that the comparisons of interventions were not statistically significant and that further higher-powered research would be needed to detect a statistically significant effect (see Analysis 1.2; Analysis 1.3; Analysis 1.4; Analysis 1.5; Analysis 2.1; Analysis 2.2; Analysis 2.3). We downgraded the evidence for all outcomes for imprecision.

Mechelse 1971 provided no direct evidence for the specified outcomes of this review, assessing facial nerve recovery outcomes at three, six or 12 months. The abstract acknowledged the 12-month outcomes; however, all the results were reported vaguely as ‘after the second week’, and, therefore, served as an indirect measure of recovery at the three-, six- and 12-month time points. All the outcome measures and time points used in Li 2016a provided evidence for the specified outcomes in this review. The indirectness in results reporting in Mechelse 1971 resulted in a further downgrade, therefore, we deemed the evidence of very low certainty.

The inconsistency between Mechelse 1971 and Li 2016a regarding time-point reporting and assessment scales made combining the results for statistical analysis impractical (see Summary of findings 1). Mechelse 1971 used an assessment system with no stipulated score that constituted recovery. Even though Li 2016a used the standardised HBGS, each trial using this system creates their own definition of which grade constitutes a complete recovery, also
making it difficult to compare results across these studies (Gantz 1999; Yanagihara 2001; Liu 2013; Li 2015; Kim 2016; Zhu 2016; Berania 2018).

Overall, we downgraded the results from both studies to very low certainty of evidence for serious study limitations, imprecision, and, for the surgery versus no treatment comparison in Mechelse 1971, indirectness. We are uncertain whether surgery affects recovery of facial nerve function at three, six and 12 months, and are equally uncertain regarding its effects on synkinesis and contracture. Neither can we draw conclusions about the comparative frequency of adverse events between treatments.

Potential biases in the review process

This review presents several limitations. The insufficient amount of high-certainty evidence limited this review in a few ways; statistical analyses could not be conducted, and the lack of data for assessment prevented conclusions being drawn regarding any potential adverse effects in relation to the surgical procedures used. The incomplete reporting on all prespecified outcomes in Mechelse 1971 and the size of the included studies did not allow for full detection of serious or rare (or both) adverse events. While the searches conducted were comprehensive, we did not contact the authors of the included studies. Given the age of Mechelse 1971, a request for additional data seemed impractical. While Li 2016a did report all the participant recovery data in full, the decision to only comment on complete recovery at each time point further limited which data could be assessed by this review, and, therefore, which conclusions could be drawn. As explained in Primary outcomes, Secondary outcomes and Overall completeness and applicability of evidence, only complete recovery was commented on, as this information is most clinically useful to those considering surgery. Surgical decompression is not a first-line treatment in the management of Bell’s palsy, and the risks do not outweigh the benefits if the recovery is not complete. Last, in relation to limited data and study selection, Hato 2015 is currently awaiting assessment and, therefore, the eligibility of this trial remains unclear.

Moreover, in terms of study design, this review focused solely on early surgical interventions for Bell’s palsy occurring within three months of diagnosis, limiting the scope of evidence. The surgical approaches used in the included studies influence limitations at review level as well, since neither study used the surgical approach of decompressing the segments of the facial nerve hypothesised to be affected in Bell's palsy. This complicates the ability of this review to draw any relevant conclusions with regards to a superior surgical approach or time frame. Therefore, we could not draw any recommendations for clinical practice, as the evidence would not have been directly applicable. The heterogeneity that exists in current literature in relation to definitions of recovery, rating scales used and timelines of outcome measurement all present further difficulties for producing recommendations for clinicians in this review.

Future review updates should consider expanding to include all time points to potentially facilitate a meta-analysis, with separate statistical tests evaluating efficacy at certain time points and if applicable, the efficacy of different surgical approaches.

There is insufficient high-certainty evidence to determine on the effects of surgical interventions for the early management of Bell’s palsy.

Agreements and disagreements with other studies or reviews

Observational studies have shown contrasting results after surgical decompression of the facial nerve for Bell’s palsy. Some studies showed evidence of an improvement in facial nerve function in the surgical groups compared to control groups (Giancarlo 1970; Fisch 1981; Gantz 1999). These studies involved surgery on participants with complete facial paralysis and indicators of poor prognosis, but the numbers involved in each study were low. In contrast, other studies found no evidence of a difference between surgical and control groups (McNeill 1974; May 1984; Aoyagi 1988; Liu 2013; Zhu 2016). It is important to note that all these observational studies varied in both timing of surgery in relation to paralysis onset and which facial nerve segments were decompressed due to different surgical approaches.

Only a few observational studies have commented on postoperative complications after decompression of the facial nerve. In one study, one in 13 people who underwent surgery developed a 40 dB hearing loss at 8 kHz (Fisch 1981). In another, six in 19 participants who had surgical decompression had significant hearing loss postoperatively (McNeill 1974). Moreover, in another study of 36 surgical participants, four had a tympanic membrane perforation, two developed a postoperative haematoma and one developed a surgical site infection (Berania 2018). Conversely, in another study, the baseline hearing threshold of the 12 participants who underwent surgical decompression improved a mean of 9.7 dB postoperatively (Kim 2016); however, the increase was not statistically significant.

AUTHORS’ CONCLUSIONS

Implications for practice

There is only very low-certainty evidence from one randomised and one quasi-randomised controlled trial and this is insufficient to decide whether early surgical intervention is beneficial or harmful in the management of Bell’s palsy. The surgical approaches used in the included studies do not decompress the labyrinthine segment or the geniculate ganglion (or both), which further complicates the applicability of the surgical approach to current clinical practice.

Implications for research

Further research into the role of early surgical intervention is unlikely to be performed because spontaneous recovery occurs in most cases. However, novel surgical procedures and diagnostic tools may warrant further updates as these data become available.

ACKNOWLEDGEMENTS

Angela Gunn and Farhad Shokareh (FS), the Information Specialists of the Cochrane Neuromuscular Disease Group performed the searches. FS screened 2020 searches for RCTs for this update.

The Managing Editor of Cochrane Neuromuscular, Ruth Brassington, assisted with updating the methods section, based on a standard protocol developed by the group from an original created by Cochrane Airways.
Andrea Takeda acted as an independent systematic reviewer and extracted data for the newly included study and documented any differences between the extractions.

This project was supported by the National Institute for Health Research (NIHR) via Cochrane Infrastructure funding to Cochrane Neuromuscular. The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, National Health Service or the Department of Health. Cochrane Neuromuscular is also supported by the MRC Centre for Neuromuscular Disease.

With thanks to Jie Zhou, Department of Neurology, West China Hospital of Sichuan University, China and Ning Chen of West China Hospital of Sichuan University who kindly assisted with translations of Chinese studies.
Surgical interventions for the early management of Bell's palsy (Review)

REFERENCES

References to studies included in this review

Li 2016a (published data only)


Mechelse 1971 (published data only)


References to studies excluded from this review

Adour 1971 (published data only)


Aoyagi 1988 (published data only)


Berania 2018 (published data only)


Brown 1982 (published data only)


Fisch 1981 (published data only)


Gantz 1999 (published data only)


Giancarlo 1970 (published data only)


Gu 1994 (published data only)


Heckmann 2012 (published data only)


Kim 2016 (published data only)


Li 2015 (published data only)


Li 2016b (published data only)


Liu 2013 (published data only)


May 1984 (published data only)


McNeill 1974 (published data only)


Yanagihara 2001 (published data only)


Ying 2011 (published data only)


Zhu 2016 (published data only)

References to studies awaiting assessment

Hato 2015 (unpublished data only)
UMIN000016334. Facial nerve regeneration surgery using fibroblast growth factor in Bell's palsy and Hunt syndrome. upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000018954 (first received 26 January 2015).

Additional references

Adour 1975

Adour 1982

Adour 2002

Antoni 1919

Ballance 1932

Cohen 1988

da Franca Pereira 2016

Deeks 2020

DerSimonian 1986

Engstrom 2008

Esslen 1977

Fisch 1972

Fisch 1984

Friedman 2000

Gagyor 2019

Hato 2012

Higgins 2011

Higgins 2020

Madhok 2016
May 1972

May 1985

McCormick 1972

Moher 2015

Morgenlander 1990

Palombo 2012

Peitersen 2002

Review Manager 2014 [Computer program]

Schünemann 2020a

Schünemann 2020b

Schwaber 1990

Sillman 1992

Spiegelhalter 2000

Stjernquist-Desatnik 2006

Sullivan 2007

Theil 2001

References to other published versions of this review

McAllister 2011

Swan 2008

Characteristics of included studies [ordered by study ID]
### Methods

Quasi-randomised controlled study
- Follow-up 12 months, with additional assessments at 3- and 6-month intervals
- Location: study conducted through the Peking Union Medical College Hospital, China
- Dates: treated at hospital between 2003 and 2013

### Participants

53 participants initially selected for 3 groups
- Inclusion criteria: adults aged > 16 years, total paralysis at onset, previous treatment with oral prednisolone 1 mg/kg within first 72 hours, duration > 2 months, worst score of facial paralysis was lower than HB V after 2 months, > 95% degeneration with reduced voluntary EMG activity after 2 months, no systemic disease
- Once allocated, 1 participant declined surgery and 11 participants declined the ‘wait and see’ control; therefore, all 12 were removed from the study, resulting in 41 participants for evaluation.
- Early surgery group (within 2–3 months from onset): n = 18; aged 21–53 years; male:female ratio 10:8
- Late surgery group (> 3 months from onset): n = 8; aged 27–58 years; male:female ratio 4:4
- Control/no treatment group: n = 15; aged 25–62; male:female ratio 8:7
- No statistically significant baseline differences in the age and sex distributions between the 3 groups (P > 0.05)

### Interventions

Participants quasi-randomised into surgical and non-surgical groups via alternation
- Early surgery group: surgery performed within 2–3 months of denervation with decompression using a retro-auricular transmastoid approach
- Late surgery group: surgery performed > 3 months after denervation with decompression using a retro-auricular transmastoid approach
- Control group: no surgery
- All participants had received oral prednisolone 1 mg/kg within 72 hours of denervation onset to be included in the study. Any participants who declined their allocated interventions were removed from the study.

### Outcomes

HBGS used to measure physician-based outcomes
- FaCE scales used to measure patient-based outcomes
- Synkinesis and other clinical features assessed via the FaCE scale using patient-reported data
- Outcomes measured at 3, 6 and 12 months
- Statistical methods included Fisher exact test for differences between proportions. Non-parametric tests used due to non-normal distribution of scores. All tests conducted at 5% level of significance. All analyses performed using SPSS 20.

### Funding

Supported by Key Projects in the National Science & Technology Pillar Program of China during the Twelfth Five-Year Plan Period, the Natural Science Foundation of Shaanxi Province, and Xi’an Jiaotong University basic scientific research operation expenses.

### Conflicts of interest

No declared conflicts of interest among the main researchers in study.

### Risk of bias

Airway
- Li 2016a (Continued)
<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>High risk</td>
<td>Quote: &quot;all patients were allocated to the surgical treatment group or the follow-up group by a quasi-random method at the clinic (alternation) … total random distribution could not be performed because some patients refused surgical intervention.&quot; Comment: not true randomisation. Participants who declined their allocated interventions were removed from the study. There is also no mention as to how the order of alternation was determined, and we judged the risk of bias high.</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>No comment on method of allocation concealment. This in combination with a high risk of bias from random sequence generation is likely to have caused baseline imbalances between treatment arms in the study.</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>Unclear risk</td>
<td>Not specified whether blinding of participants to either a surgical or non-surgical intervention was possible.</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Unclear risk</td>
<td>Blinding of the investigator/outcome assessor was not reported and would have been difficult to perform in this study because of the surgical intervention involved.</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>High risk</td>
<td>Quote: &quot;one patient in the surgical group who declined surgery was removed from the study, and 11 patients in the follow-up group who refused the wait-and-see approach and went to other hospitals for treatment were also removed from the study … One patient in the early surgery group and two patients in the no treatment group were lost to follow-up … 11 patients in the early surgery group and 10 patients in the no treatment group completed the FaC E questionnaires at presentation and at the end of follow-up.&quot; Comment: 12 participants were lost to allocation and commented on. However, 1/18 participants in the early surgical group and 2/13 participants in the no treatment group were lost to follow-up after the 6-month assessment with no reasons given. Only 21/41 participants completed the FaC E questionnaire with no reasons given.</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Low risk</td>
<td>All prespecified outcomes were reported on; therefore, our judgement was low risk.</td>
</tr>
<tr>
<td>Other bias</td>
<td>Low risk</td>
<td>No other sources of bias identified.</td>
</tr>
</tbody>
</table>

**Mechelse 1971**

**Study characteristics**

Methods

Randomised controlled trial

After meeting inclusion criteria, participants were first allocated to a group with presumed ‘bad prognosis’ or to a group with presumed ‘good prognosis.’ For the ‘bad prognosis’ group, rate of recovery was followed for 1–3 years.

Location: 2 centres participated simultaneously: University Hospital, Leiden and Municipal Hospital, The Hague, Netherlands

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Participants

25 participants were selected based on a 'bad prognosis', defined as complete facial palsy and EMG showing no voluntary control of motor unit or a minimal applied current evoking a motor response on the affected side 2.5 times that of the unaffected side. These responses needed to be confirmed on 2 occasions a few days apart. 1 participant refused surgery and was excluded from the study, resulting in 24 evaluable participants total.

Surgery group: n = 11; mean age 45.3 years; male:female ratio 1:10; side of idiopathic facial palsy left:right = 7:4

Control group: n = 13; mean age 45 years; male:female ratio 8:5; side of idiopathic facial palsy left:right = 6:7

Study did not report if there were statistically significant differences between groups.

Interventions

Surgery group: decompressed using a retro-auricular approach 7–20 days after the onset of paralysis

Control group: no treatment

Outcomes

Recovery of facial nerve function at 12 months assessed using 0–5 scale (0 no function, 5 complete function). Study report did not state statistical methods used

Funding

No sponsorship noted

Conflicts of interest

No declared conflicts of interest among the main researchers in the study

Notes

Study also included a population of 242 non-randomised participants with 'presumed good prognosis' who did not meet the criteria for having a 'bad prognosis.' Quote: "They received no therapy... and were also followed clinically and electromyographically along with the patients of the controlled trial in order to validate our criteria."

Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “in both hospitals these patients were entered on a list, previously prepared by the statistical department, University of Leiden [head, Mr H. De Jonge], which randomly allocated them to surgical treatment or to a control group.” Comment: 25 participants were randomly allocated into surgical or control groups. 1 participant declined surgery and was removed from the study.</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>No comment on any attempted allocation concealment.</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias) All outcomes</td>
<td>Unclear risk</td>
<td>Not specified whether blinding of the participant to either a surgical or non-surgical intervention was possible.</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) All outcomes</td>
<td>Unclear risk</td>
<td>Report does not comment on blinding of the investigator/outcome assessor, which would have been difficult to perform in this study because of the surgical intervention involved.</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias) All outcomes</td>
<td>Low risk</td>
<td>Quote: “All patients were followed clinically and electromyographically for a least a year.” Comment: complete follow-up with no losses.</td>
</tr>
</tbody>
</table>
Selective reporting (reporting bias)  High risk  All prespecified outcomes were reported on indirectly via results “after the second week,” inferring that results at each of our prespecified time points would be the same.

Other bias  Low risk  No other sources of bias identified.

EMG: electromyography; FaCE: Facial Clinimetric Evaluation scale; HB: House-Brackmann; HBGS: House-Brackmann grading system; n: number of participants.

Characteristics of excluded studies [ordered by study ID]

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adour 1971</td>
<td>Participants who refused surgery were included and added to the control group. Therefore, allocation was not randomised.</td>
</tr>
<tr>
<td>Aoyagi 1988</td>
<td>Included participants were not randomised into surgical or non-surgical groups.</td>
</tr>
<tr>
<td>Berania 2018</td>
<td>Retrospective study. No control group included for comparison of outcomes.</td>
</tr>
<tr>
<td>Brown 1982</td>
<td>Retrospective audit. Study did not compare outcomes between groups or describe how surgery and no-surgery groups were selected.</td>
</tr>
<tr>
<td>Fisch 1981</td>
<td>Retrospective study. No description given as to how participants were allocated to the surgically treated or not surgically treated groups. No discussion of the non-surgical group. Only 14 participants underwent surgery.</td>
</tr>
<tr>
<td>Gantz 1999</td>
<td>No randomisation. Participants chose whether to have surgery. 11/30 participants offered surgery declined; therefore, only 19/169 evaluated had surgery.</td>
</tr>
<tr>
<td>Giancarlo 1970</td>
<td>No randomisation. Participants chose whether to have surgery. Also, report did not describe the degree of baseline clinical palsy, merely that the nerve was degenerating.</td>
</tr>
<tr>
<td>Gu 1994</td>
<td>Upon initial abstract screening, the abstract suggested potential randomisation. However, on full-text retrieval and translation, the study was non-randomised.</td>
</tr>
<tr>
<td>Heckmann 2012</td>
<td>Literature review. Study did not include comparison groups. Primarily focused on efficacy of steroid and antiviral use as opposed to surgical interventions.</td>
</tr>
<tr>
<td>Kim 2016</td>
<td>Retrospective study. No randomisation. Participants chose whether to have surgery. 22/34 participants offered surgery declined, therefore, only 12 had surgery.</td>
</tr>
<tr>
<td>Li 2015</td>
<td>Included participants were not randomised into surgical or non-surgical groups.</td>
</tr>
<tr>
<td>Li 2016b</td>
<td>Included participants were diagnosed with hemi-facial spasm.</td>
</tr>
<tr>
<td>Liu 2013</td>
<td>No control group included for comparison of outcomes.</td>
</tr>
<tr>
<td>May 1984</td>
<td>Prospective audit. No attempt at randomisation. Control group was self-selected, considered poor surgical risk or had already received corticosteroids. 50 participants underwent surgery with a control group of 35. However, statistical comparison was performed on only the poorest prognostic groups of 25 surgical and 13 non-surgical participants, which showed no statistically significant difference.</td>
</tr>
<tr>
<td>Study</td>
<td>Reason for exclusion</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>McNeill 1974</td>
<td>Retrospective audit. Participants chose whether to have surgery. Control group either declined surgery or were considered poor surgical risk.</td>
</tr>
<tr>
<td>Yanagihara 2001</td>
<td>Participants had trial of corticosteroids initially before being considered for surgery. No randomisation. Participants chose whether to have surgery. Report noted that younger participants opted for surgery and older participants refused surgery.</td>
</tr>
<tr>
<td>Ying 2011</td>
<td>Participants included those with ‘hemifacial spasm’ and outcomes were not reported separately by aetiology. In addition, timeframe from onset to procedure varied widely up to 180 days post onset, and was not reported separately for specific time periods.</td>
</tr>
<tr>
<td>Zhu 2016</td>
<td>Retrospective study. No randomisation. Participants chose whether to have surgery.</td>
</tr>
</tbody>
</table>

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

**Hato 2015**

**Methods**

Parallel, randomised, “placebo-controlled” (sic) (assessor blinded)

**Participants**

Inclusion criteria: adults with severe Bell’s palsy and Ramsay Hunt syndrome; aged > 15 years; with complete facial palsy (Yanagihara score ≤ 10); degree of denervation > 90% with ENOG and clinical follow-up unavailable for more than 12 months

Exclusion criteria: pregnancy; current neoplasms; chronic kidney disease; chronic liver disease; immune suppression; opt out decompression surgery

**Interventions**

Intervention group: facial nerve regenerating surgery with human recombinant bFGF

Control group: transmastoid conventional facial nerve decompression surgery

**Outcomes**

Primary: non-recovery rate 12 months after disease onset

Secondary: non-recovery rate 6 months after disease onset, incidence rate of sequelae 12 months after disease onset. Postoperative bleeding needing haemostasis in operating room and audiometric threshold elevation using air conduction pure tone audiometry at 250 Hz, 500 Hz, 1 kHz, 2 kHz and 4 kHz test frequencies

**Funding**

Quote: "Self-funding."

**Notes**

Author contacted 22 June 2020 regarding likely publication date, clarification of intervention, outcomes and population, and whether study is ongoing or if statistical analysis is completed.

bFGF: basic fibroblast growth factor; ENOG: electroneurography.

**DATA AND ANALYSES**
### Comparison 1. Early surgery versus no treatment

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Complete recovery at 12 months</td>
<td>1</td>
<td></td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>1.2 Complete recovery at 3 months</td>
<td>1</td>
<td></td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>1.3 Complete recovery at 6 months</td>
<td>1</td>
<td></td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>1.4 Synkinesis</td>
<td>1</td>
<td></td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>1.5 Contractures</td>
<td>1</td>
<td></td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>Totals not selected</td>
</tr>
</tbody>
</table>

#### Analysis 1.1. Comparison 1: Early surgery versus no treatment, Outcome 1: Complete recovery at 12 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery Events</th>
<th>Early surgery Total</th>
<th>No treatment Events</th>
<th>No treatment Total</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li 2016a</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td>13</td>
<td>0.76 [0.05, 11.11]</td>
</tr>
</tbody>
</table>

#### Analysis 1.2. Comparison 1: Early surgery versus no treatment, Outcome 2: Complete recovery at 3 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery Events</th>
<th>Early surgery Total</th>
<th>No treatment Events</th>
<th>No treatment Total</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li 2016a</td>
<td>1</td>
<td>18</td>
<td>1</td>
<td>15</td>
<td>0.83 [0.06, 12.22]</td>
</tr>
</tbody>
</table>

#### Analysis 1.3. Comparison 1: Early surgery versus no treatment, Outcome 3: Complete recovery at 6 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery Events</th>
<th>Early surgery Total</th>
<th>No treatment Events</th>
<th>No treatment Total</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li 2016a</td>
<td>1</td>
<td>18</td>
<td>1</td>
<td>15</td>
<td>0.83 [0.06, 12.22]</td>
</tr>
</tbody>
</table>
### Analysis 1.4. Comparison 1: Early surgery versus no treatment, Outcome 4: Synkinesis

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery</th>
<th>No treatment</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechelse 1971</td>
<td>9</td>
<td>11</td>
<td>0.89 [0.64, 1.22]</td>
</tr>
</tbody>
</table>

Risk Ratio: M-H, Random, 95% CI

Favours early surgery: 0.2
Favours no treatment: 5

### Analysis 1.5. Comparison 1: Early surgery versus no treatment, Outcome 5: Contractures

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery</th>
<th>No treatment</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechelse 1971</td>
<td>8</td>
<td>11</td>
<td>1.58 [0.79, 3.14]</td>
</tr>
</tbody>
</table>

Risk Ratio: M-H, Random, 95% CI

Favours early surgery: 0.2
Favours no treatment: 5

### Comparison 2. Early surgery versus late surgery

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Complete recovery at 12 months</td>
<td>1</td>
<td>17</td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>2.2 Complete recovery at 3 months</td>
<td>1</td>
<td>1</td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>2.3 Complete recovery at 6 months</td>
<td>1</td>
<td>1</td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
</tbody>
</table>

### Analysis 2.1. Comparison 2: Early surgery versus late surgery, Outcome 1: Complete recovery at 12 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery</th>
<th>Late surgery</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li 2016a</td>
<td>1</td>
<td>1</td>
<td>0.47 [0.03, 6.60]</td>
</tr>
</tbody>
</table>

Risk Ratio: M-H, Fixed, 95% CI

Favours early surgery: 0.01
Favours late surgery: 100
Analysis 2.2. Comparison 2: Early surgery versus late surgery, Outcome 2: Complete recovery at 3 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery</th>
<th>Late surgery</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events  Total</td>
<td>Events  Total</td>
<td></td>
</tr>
<tr>
<td>Li 2016a</td>
<td>1  18</td>
<td>1  8</td>
<td>0.44 [0.03 , 6.25]</td>
</tr>
</tbody>
</table>

Risk Ratio M-H, Fixed, 95% CI: 0.2

Risk Ratio M-H, Fixed, 95% CI: 0.5

Favours late surgery

Favours early surgery

Analysis 2.3. Comparison 2: Early surgery versus late surgery, Outcome 3: Complete recovery at 6 months

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Early surgery</th>
<th>Late surgery</th>
<th>Risk Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events  Total</td>
<td>Events  Total</td>
<td></td>
</tr>
<tr>
<td>Li 2016a</td>
<td>1  18</td>
<td>1  8</td>
<td>0.44 [0.03 , 6.25]</td>
</tr>
</tbody>
</table>

Risk Ratio M-H, Fixed, 95% CI: 0.2

Risk Ratio M-H, Fixed, 95% CI: 0.5

Favours late surgery

Favours early surgery

ADDITIONAL TABLES

Table 1. Risk of bias classification

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Interpretation</th>
<th>Within study</th>
<th>Across studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk of bias</td>
<td>Plausible bias unlikely to seriously alter the results.</td>
<td>Low risk of bias for all key domains.</td>
<td>Most information is from studies at low risk of bias.</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
<td>Plausible bias that raises some doubt about the results.</td>
<td>Unclear risk of bias for ≥ 1 key domains.</td>
<td>Most information is from studies at low or unclear risk of bias.</td>
</tr>
<tr>
<td>High risk of bias</td>
<td>Plausible bias that seriously weakens the confidence in the results.</td>
<td>High risk of bias for ≥ 1 key domains.</td>
<td>The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results.</td>
</tr>
</tbody>
</table>

Approach for summary assessments of the risk of bias for each important outcome (across domains) within and across studies (Higgins 2011).

APPENDICES

Appendix 1. ClinicalTrials.gov

Advanced search

Condition or disease: "Bell's palsy" OR "Bell palsy" OR "idiopathic facial paralysis" OR "facial paralysis" OR "facial palsy" OR "facial nerve"

Study type: Interventional Studies (Clinical Trials)

Intervention/treatment: "surgery" OR "surg" OR "operative" OR "operat" OR "decompression" OR "decompres"

17 studies
Appendix 2. World Health Organization International Clinical Trials Registry Platform (ICTRP)

Advanced search

(Bell's palsy OR Bell palsy OR idiopathic facial paralysis OR facial paralysis OR facial palsy OR facial nerve) in the Title

Recruitment status is ALL

107 records for 106 trials found

Appendix 3. Additional methods

Measures of treatment effect

If studies reported outcomes that were conceptually the same but were measured using different scales, we would have reported standardised mean differences with corresponding confidence intervals, ensuring that we entered data presented on a scale with a consistent direction of effect. We would have presented a rule-of-thumb interpretation alongside, whereby 0.2 represents a small effect, 0.5 a moderate effect and 0.8 a large effect (Cohen 1988), as described in Schünemann 2020b.

Unit of analysis issues

If in future updates two comparisons from the same trial (e.g. intervention A versus placebo and intervention B versus the same placebo group) are combined in the same meta-analysis, we will follow guidance in Chapter 23 of the Cochrane Handbook for Systematic Reviews of Interventions to avoid double-counting measurements. Our preferred approach will be to combine intervention groups if clinically appropriate, or if not, halve a control group (Higgins 2020).

Assessment of heterogeneity

There were insufficient studies to enable any statistical analysis. If we find sufficient studies at future updates, we will perform a Chi² test for homogeneity. If there is significant heterogeneity, we will attempt to identify the cause based on the characteristics of the studies included.

If we identify substantial unexplained heterogeneity, we will report it and explore possible causes by prespecified subgroup analysis. We will use the rough guide to interpretation as outlined in Chapter 10 of the Cochrane Handbook for Systematic Reviews of Interventions, as follows:

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

We will avoid the use of absolute cut-off values, but interpret the Chi² statistic in relation to the size and direction of effects and strength of evidence for heterogeneity (e.g. P value from the Chi² test, or CI for I² statistic) (Deeks 2020).

Assessment of reporting biases

If we can pool more than 10 trials, we will create and examine a funnel plot to explore possible small-study biases.

Data synthesis

If meta-analysis is possible in future updates, as a general rule, we will use a random-effects model in Review Manager 5 (Review Manager 2014). Random-effects models such as DerSimonian and Laird account for more uncertainty (DerSimonian 1986). We will perform a sensitivity analysis using the fixed-effect model in the presence of heterogeneity, to determine whether results are systematically different.

Subgroup analysis and investigation of heterogeneity

We did not plan subgroup analyses.

Sensitivity analysis

If statistical analysis becomes possible in the future, we will perform a sensitivity analysis omitting studies of high risk of bias. In addition, risk of bias could be incorporated into mixed models simultaneously allowing for differences in bias using Bayesian methods, utilised in WinBUGS (Spiegelhalter 2000).

We will assess the effects of fixed-effect versus random-effects meta-analysis if there is statistical heterogeneity.

Appendix 4. Embase (OvidSP) search strategy

Database: Embase <1974 to 2020 week 11>
Search strategy:

--------------------------------------------------------------------------------
1 crossover-procedure.sh. (62415)
2 double-blind procedure.sh. (170250)
3 single-blind procedure.sh. (38181)
4 randomized controlled trial.sh. (593541)
5 (random$ or crossover$ or cross over$ or placebo$ or (doubl$ adj blind$) or allocat$).tw,ot. (1749674)
6 trial.ti. (292985)
7 clinical trial/ (966802)
8 or/1-7 (2479571)
9 (animal/ or nonhuman/ or animal experiment/) and human/ (1928466)
10 animal/ or nonanimal/ or animal experiment/ (3929237)
11 10 not 9 (3270546)
12 8 not 11 (2315974)
13 limit 12 to (conference abstracts or em base) (1949659)
14 ((bell$ or facial or hemifacial$) adj3 (palsy$ or paralysis$ or paresis$ or spasm$)).mp. (29285)
15 (surg$ or operat$ or decompress$).mp. (4763691)
16 ((bell$ or facial or hemifacial$) adj3 (palsy$ or paralysis$ or paresis$ or spasm$)) adj5 (surg$ or operat$ or decompress$)).mp. (2225)
17 Bell Palsy/su or Facial Nerve Paralysis/su or Hemifacial Spasm/su (2309)
18 16 or 17 (3745)
19 13 and 18 (151)
20 remove duplicates from 19 (150)

Appendix 5. Cochrane Neuromuscular Specialised Register (CRS Web) search strategy

#1 MESH DESCRIPTOR Bell Palsy WITH QUALIFIER SU AND INREGISTER 2
#2 MESH DESCRIPTOR facial paralysis WITH QUALIFIER SU AND INREGISTER 4
#3 MESH DESCRIPTOR Hemifacial Spasm WITH QUALIFIER SU AND INREGISTER 1
#4 (bell NEAR2 palsy) or (facial NEAR2 palsy) or (facial NEAR2 paralysis) or (facial NEAR2 paresis) AND INREGISTER 171
#5 (hemifacial NEAR2 palsy) OR (hemifacial NEAR2 paralysis) OR (hemifacial NEAR2 paresis) AND INREGISTER 2
#6 (facial NEAR2 palsy) or (facial NEAR2 paralysis) or (facial NEAR2 paresis) AND INREGISTER 137
#7 #4 OR #5 OR #6 171
#8 (surg* or operati* or decompressi*):AB,E H,E M T,KW ,KY,M H,TI AND INREGISTER 956
#9 #8 AND #7 20
#10 #1 OR #2 OR #3 OR #9 150

Appendix 6. Cochrane Central Register of Controlled Trials (CENTRAL) (CRS Web) search strategy

#1 MESH DESCRIPTOR Bell Palsy WITH QUALIFIER SU AND CENTRAL:TARGET 2
#2 MESH DESCRIPTOR facial paralysis WITH QUALIFIER SU AND CENTRAL:TARGET 10
#3 MESH DESCRIPTOR Hemifacial Spasm WITH QUALIFIER SU AND CENTRAL:TARGET 4
#4 (bell NEAR2 palsy) or (facial NEAR2 palsy) or (facial NEAR2 paralysis) or (facial NEAR2 paresis) AND CENTRAL:TARGET 778
#5 (hemifacial NEAR2 palsy) OR (hemifacial NEAR2 paralysis) OR (hemifacial NEAR2 paresis) AND CENTRAL:TARGET 5
#6 (facial NEAR2 palsy) or (facial NEAR2 paralysis) or (facial NEAR2 paresis) AND CENTRAL:TARGET 692
#7 #4 OR #5 OR #6 AND CENTRAL:TARGET 779
#8 (surg* or operati* or decompressi*):AB,E H,E M T,KW ,KY,M H,TI AND CENTRAL:TARGET 266584
Appendix 7. MEDLINE (OvidSP) search strategy

Database: Ovid MEDLINE(R) ALL <1946 to 19 March 2020>

Search strategy:

1 randomized controlled trial.pt. (502288)
2 controlled clinical trial.pt. (93580)
3 randomized.ab. (473646)
4 placebo.ab. (206169)
5 drug therapy.fs. (2188387)
6 randomly.ab. (329654)
7 trial.ab. (498947)
8 groups.ab. (2024489)
9 or/1-8 (4661504)
10 exp animals/ not humans.sh. (4680615)
11 9 not 10 (4040593)
12 ((bell$ or facial or hem ifacial$) adj3 (pals$ or paralys$ or paresi$ or spasm$)).mp. (20548)
13 (surg$ or operat$ or decom pres$).mp. (3668969)
14 ((bell$ or facial or hem ifacial$) adj3 (pals$ or paralys$ or paresi$ or spasm$) adj5 (surg$ or operat$ or decom pres$)).mp. (1735)
15 bell palsy/su or facial paralysis/su or hem ifacial spasm/su (2855)
16 14 or 15 (3781)
17 11 and 16 (367)
18 remove duplicates from 17 (366)

WHAT’S NEW

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 March 2020</td>
<td>New citation required but conclusions have not changed</td>
<td>1 study included. Ian Swan withdrew from authorship. Isabella Menchetti joined the review as first author.</td>
</tr>
<tr>
<td>20 March 2020</td>
<td>New search has been performed</td>
<td>The search strategies were revised to be more specific and we updated the searches to March 2020. We included 1 new study (Li 2016a) and excluded 1 previously included study (Adour 1971), having reassessed it as a non-randomised controlled trial. We updated the methods to current standards.</td>
</tr>
</tbody>
</table>

HISTORY

Review first published: Issue 2, 2011

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 October 2014</td>
<td>Amended</td>
<td>Correction to Background: removed statement that the condition was first described by Sir Charles Bell.</td>
</tr>
<tr>
<td>17 October 2013</td>
<td>Amended</td>
<td>Corrected date of search in abstract</td>
</tr>
<tr>
<td>13 February 2013</td>
<td>New citation required but conclusions have not changed</td>
<td>New search to 29 October 2012. Published notes included</td>
</tr>
<tr>
<td>29 October 2012</td>
<td>New search has been performed</td>
<td>This review has been updated with a new search but no new relevant studies were identified. The text has been edited.</td>
</tr>
</tbody>
</table>

**CONTRIBUTIONS OF AUTHORS**

IM assessed risk of bias, certainty of evidence, undertook data collection and analysis, and wrote the review.

KMcA devised the search strategy, designed the protocol, assessed risk of bias, undertook data collection and analysis, and wrote the review.

DW devised the search strategy, designed the protocol, assessed risk of bias, undertook data collection and analysis, and wrote the review.

PTD provided statistical knowledge and expertise required for the protocol and review.

**DECLARATIONS OF INTEREST**

IM: none.

KMcA: none.

DW: none.

PTD has received grants from AbbVie, Gilead Sciences and SHIRE. Prof Donnan is a member of the New Drugs Committee of the Scottish Medicines Consortium.

**SOURCES OF SUPPORT**

**Internal sources**

- No sources of support supplied

**External sources**

- None, Other

**DIFFERENCES BETWEEN PROTOCOL AND REVIEW**

We have updated the ‘Risk of bias’ methodology since the protocol was published, in order to conform to the 2011 Cochrane methodology (Higgins 2011). We also added ‘Summary of findings’ tables and specified outcomes for inclusion in the table, in the Methods. We stated the comparisons that we would consider for inclusion and clarified this in the Methods.

We included searches of trials registries.

In this update, we also expanded most sections of the methods to fulfil MECIR reporting standards. We reported methods for use if meta-analysis becomes possible, in Appendix 3. Changes to the protocol included the following.

- **Types of outcomes**: clarified that the measurement of outcomes was not a selection criterion, but the outcomes were those of interest in eligible trials. We better defined ‘psychosocial outcomes’ by specifying the preferred measurement scales for both mood and appearance.
• **Primary outcome:** expanded on to include justification for only reporting on and calculating a risk ratio for complete recovery data in both groups, for both comparisons.

• **Selection of studies:** the Cochrane Neuromuscular Information Specialist performed an initial screen of titles and abstracts for randomised controlled trials.

• **Searching other resources:** we did not contact the authors of all included trials for further additional information or for information on unpublished trials. We did not perform handsearching.

• **Data extraction and management:** described an approach to papers requiring translation. An independent systematic reviewer performed the second data extraction for the newly included study and we resolved differences via discussion between lead author IM and review author KM.

• **Measures of treatment effect:** describes use of standardised mean difference if studies report different scales for conceptually the same outcome.

• **Unit of analysis issues:** described the approach to multi-arm trials.

• **Dealing with missing data:** clarified that no authors were contacted in relation to missing data.

• **Assessment of heterogeneity:** included a comment on interpretation of Chi² statistic.

• **Assessment of reporting biases:** described an approach to identification and reporting of unpublished studies, to analysis of reporting bias in the review. We also noted that funnel plots required 10 trials.

• **Data synthesis:** specified comparisons more clearly and described an approach to combining data; described GRADE methodology in more detail. We stated that we would report random-effects analyses, with sensitivity analysis using the fixed-effect model.

• **Sensitivity analysis:** stated that we would perform sensitivity analyses using risk of bias (previously ‘quality’) and specified a sensitivity analysis for fixed-effect versus random-effects models if heterogeneity was present.

Iain Swan withdrew from authorship. Isabella Menchetti joined the review as first author.

**INDEX TERMS**

**Medical Subject Headings (MeSH)**

Bell Palsy [*surgery]; Decompression, Surgical [*methods]; Facial Nerve [*surgery]; Randomized Controlled Trials as Topic

**MeSH check words**

Humans