Early Electrodiagnosis in the Management of Neonatal Brachial Plexus Palsy.
A Systematic Review.

Ruth VAN DER LOOVEN, MD; Laura LE ROY, Ms; Emma TANGHE, Ms; Christine VAN DEN BROECK, PhD; Martine DE MUYNCK, MD, PhD; Guy VINGERHOETS, PhD; Matthew PITT, MD; Guy VANDERSTRAETEN, MD, FRCP, PhD

1Department of Physical and Rehabilitation Medicine, Ghent University Hospital, C. Heymanslaan 10, 9000, Ghent, Belgium
2Department of Rehabilitation Sciences and Physiotherapy, Ghent University, C. Heymanslaan 10, 9000, Ghent, Belgium
3Department of Experimental Psychology, Faculty of Psychological and Educational Sciences, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium.
4Department of clinical Neurophysiology, Great Ormond Street Hospital for Children NHS Trust, London WC1N 3JH, United Kingdom.
5Faculty of Medicine and Health Sciences, Ghent University, C. Heymanslaan 10, 3K3, 9000 Ghent, Belgium.

Acknowledgments: We would like to thank Kristine Oostra, MD, PhD (Physical and Medicine, University Hospital Ghent, Belgium) and Wim Vanhove, MD (Orthopedics, University Hospital Ghent, Belgium) for their scientific support and sharing their clinical expertise.

Corresponding author: Ruth VAN DER LOOVEN, Child Rehabilitation Centre, Department of Physical and Rehabilitation Medicine, Ghent University Hospital, C. Heymanslaan 10, route 445, 9000 Ghent, Belgium. Email: ruth.vanderlooven@uzgent.be. Work phone: 003293326939. Mobile Phone: 0032477931644.

Running title: Early EDX in NBPP management

Ethical Publication Statement: We confirm that we have read the Journal’s position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Disclosure of conflicts of interest: None of the authors has any conflict of interest to disclose.
Abstract

Neonatal brachial plexus palsy (NBPP) is an important form of newborn morbidity with potentially disabling persistence. Neurosurgical intervention is indicated in selected patients. Early prognostic assessment would facilitate rational selection of those infants for surgery.

We conducted a systematic literature review to determine the prognostic value of early electrodiagnosis (EDX) in NBPP. We included 16 observational studies with a total sample size of 747 children. Risk of bias and quality of evidence were rated.

Wide variation was found in EDX techniques, outcome algorithms and decision making. Nevertheless, the most methodologically sound studies support the use of EDX, at standardized time frames, as key prognostic modality complementing clinical judgement and neuroimaging. An accurate knowledge of underlying anatomy of the nerve injury helps to counsel families and to guide reconstructive strategy.

Keywords  neonatal brachial plexus palsy, brachial plexus neuropathy, electrodiagnosis, prognosis, electrodiagnostic testing.
INTRODUCTION

Neonatal brachial plexus palsy (NBPP) is reported in 0.1 to 8.10 per 1000 live births worldwide.\textsuperscript{1-6} It is the result of a closed nerve stretch injury to the brachial plexus during the perinatal period. The most common lesions occur within the C5 and C6 spinal nerves (80% of patients), with a smaller group of patients having more extensive lesions ranging from C5–C7 and from C5–T1 (panplexopathy).\textsuperscript{7} Prevention is difficult due to the unpredictability and multifactorial nature of the risk factors. As severity varies from neurapraxia, axonotmesis, neurotmesis, to root avulsion, and the extent of injury varies between damage to one nerve or all roots, the impact of neonatal plexopathy ranges from temporary functional impairment to a lifelong total paralysis of one arm.\textsuperscript{3,5,8-13,14}

Early management includes parental counselling, family support, handling information, splinting, appropriate, early and supervised rehabilitation. Neurosurgical intervention is usually undertaken between 3 months and 12 months of age in children who have shown little or no significant improvement in the affected muscle groups.\textsuperscript{2,15-30} It is generally agreed that early surgical nerve repair can greatly improve the functional outcome in selected patients.\textsuperscript{15,20,22,24,31-33} Delay to time of surgery often results in progressive worsening of deformity in the shoulder joint as contractures progress quickly over time in a rapidly growing infant.\textsuperscript{34} Surgery in the first months of life is therefore indicated for serious neuroma formation, severe axonotmesis, neurotmesis and root avulsion. Early identification of those children is obligatory for surgical management and patient counselling. The major difficulty is the lack of reliable early indicators of prognosis. In young babies, obtaining a reliable physical examination can be difficult\textsuperscript{17,32} and falls short in defining the underlying anatomy of the injury. Imaging techniques, such as magnetic resonance imaging (MRI) and
computed tomography (CT) myelography, have approximately 60-80% accuracy.\textsuperscript{35,36} Disadvantageous is their need for general anesthesia, their availability and cost. The utility of ultrasound\textsuperscript{37} (US) has not been extensively examined but is promising.\textsuperscript{38-40} Although EDX is established in the assessment of babies with NBPP,\textsuperscript{41-45} its utility for early prognosis remains controversial.\textsuperscript{8,17,46-51} Opinions differ from no value at all\textsuperscript{26,52-54} to strong recommendation that all children undergo a combination of EDX techniques to determine which candidates would most benefit from an operation with nerve reconstruction.\textsuperscript{41,44,45,50,55-57} Even the value of intraoperative EDX remains to be determined.\textsuperscript{56} Because more stringent and exact EDX testing is available we wanted to revisit the place of early EDX in the prognostication of NBPP.

The aim of the present study is to systematically assess all available evidence on prognostic value of early EDX in detecting babies with root avulsion, neurotmesis, severe axonotmesis and neuroma formation.

**METHODS**

The Cochrane Database of Systematic Reviews was searched to ensure a similar review had not been undertaken. The PRISMA statement\textsuperscript{58} and MOOSE guidelines\textsuperscript{59} were strictly followed to report this review. The protocol was accepted for registration in PROSPERO on January 15\textsuperscript{th} 2018 (CRD42018076644).

**Information sources and search strategy**

The search strategy, using the PICO (Population, Intervention, Comparison and Outcome)\textsuperscript{60} format for clinical questions, was developed systematically in the MEDLINE database (PubMed interface), using medical subject headings as well as free text words (Supplementary Figure 1). A broad search term, including extensive number of synonyms for
NBPP, was preferred to decrease the likelihood of missing relevant articles. The search was translated in Web of Science and Embase (Embase.com Interface). The published guidelines to identify prognostic studies in MEDLINE and EMBASE were followed.\textsuperscript{61,62} If full text was not accessible, requests were send out to the home university and corresponding authors. We searched entries on ClinicalTrials.gov for ongoing studies. The final literature search was conducted on July 31, 2019. For each primary article retrieved, we examined the related articles and checked for additional pertinent references. In addition, we searched the reference list of encountered reviews for appropriate articles. Until the time of submission, automatic e-mail alerts were checked for possible additional relevant articles. A list of all identified studies is available on request. We imported citations into bibliographic software.\textsuperscript{63}

**Eligibility**

We included all published randomized controlled trials as well as observational studies that assessed the prognostic value of early EDX in detecting babies with neuroma formation, severe axonotmesis, neurotmesis and root avulsion. NBPP was defined as a closed nerve stretch injury to the cervical and/or first thoracic nerves (C5-T1). Language was restricted to English, Dutch, or French. No lower date limit or minimal length of follow-up period was applied. Case series with fewer than 10 patients and studies that were duplicates of other studies were excluded.

Two authors (R.V.D.L, G.VDS), trained medical specialists in electrodiagnosis, independently screened all articles by title, abstract and keywords using the Covidence web-based platform\textsuperscript{64} recommended by the Cochrane Organisation for systematic reviews. When all inclusion criteria (Supplementary Table 1) were met, the full text version of the article was assessed. Disagreements were resolved by consensus. When necessary, authors of the
original study were contacted if possible for further details. Inter-rater agreement was measured with Cohen’s kappa coefficient.

Data extraction

Data from each included trial were extracted independently by the 3 reviewers (R. V.D.L, E.T, and L.L.) on a data extraction form designed in accordance with the Cochrane Checklist of items. Any disagreements were solved after discussion. Details included: source, study design, sample size, participant demographics, EDX characteristics, outcome measures, length of follow-up, statistical analysis as well as any other miscellaneous data.

Risk of bias and Quality assessment

Three reviewers (R. V.D.L, E.T, L. L) independently assessed the risk of bias by using QUIPS (Quality In Prognosis Studies). The QUIPS tool was devised for prognostic factor review questions and has demonstrated acceptable inter-rater reliability. It consists of 6 domains of potential biases: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, statistical analysis and reporting.

The reviewers predefined some domains more specifically to judge the included articles. The study population should be constituted on a demographic basis, representing differing degrees of injury at different levels and with different recovery levels. The prognostic factor measurement should include the EDX technique, equipment, timing of intervention, skills of practitioner and prognostic outcome estimation. Assessment of the outcome should be accurate and reproducible, preferably using a specialized pre-defined assessment protocol. Ideally, outcome measures should take the variability of functional requirements of the different investigated muscles and nerves into account. As neurological recovery and motor improvement may continue beyond 2.5 years and sensory return beyond 3 years, the follow-up should be sufficiently long. Confounding factors were the concomitant
conservative management methods, development of secondary sequelae, the eligibility
criteria for surgery, the timing for surgery and the different types of surgery interventions.
The GRADE approach was used to rate the overall quality of evidence for the prognostic
value of the different electrodiagnosis techniques across studies.\textsuperscript{67} The GRADE method
criteria has been adapted for prognostic research.\textsuperscript{69}

RESULTS

Study identification

The PRISMA flow diagram (Figure 1) illustrates the search process. Main exclusion reasons
are provided in Figure 1. The Cohen’s Kappa to evaluate concordance of independent
reviewers was 0.6312, which indicates a substantial inter-rater agreement.

Study characteristics

The individual characteristics of the included studies are outlined in Supplementary Table 2.
Each study was published in English. All were observational with patients selected as an
‘inception’ cohort and followed forwards in time from diagnosis. Only 6 studies\textsuperscript{70-75} had a
prospective design. Most studies were completed in Europe,\textsuperscript{72-74,76-79} followed by three
studies performed in Asia\textsuperscript{75,80,81} and South-America.\textsuperscript{49,70,71} Three studies were conducted in
the USA.\textsuperscript{82-84}

Risk of bias assessment

Study participation : All of the studies are prone to selection bias by including a non-random
sample of the NBPP population. Six studies\textsuperscript{72,77,78,82-84} included the surgery group, five the
slow recovery or severe group,\textsuperscript{49,70,71,76,80} and five included NBPP children referred to a
tertiary centre.\textsuperscript{73-75,79,81} In two studies\textsuperscript{77,82} the population was restricted to upper trunk
lesions. Individual papers included from 13 to 100 children with NBPP, for a final sample size of 747.

Study attrition: Seven articles49,70,71,73,80,81,83 insufficiently described, if any, their participants lost to follow-up.

Prognostic factor measurement: Four authors49,73,74,79 used needle electromyography (EMG) and two70,71 the amplitude of the compound muscle action potential (CMAP) as a single technique. Five75,80,81,83,84 investigators combined EMG with complete nerve conductions studies (NCS), including motor and sensory action potentials and conduction velocities. Two authors72,76 used only motor NCS and EMG. None of the included studies adopted preoperative, and only one77 intraoperative, somatosensory evoked potentials (SEPs). Three authors77,78,82 performed intraoperative NCS, and Chin77 utilized a combination of intraoperative NCS, EMG and gross motor response to stimulation. The timing of the first (preoperative) EDX intervention varied from 5 days74 to 14 months76 after birth. Two studies73,75 provided only the mean value ranging from 2.6 weeks to 31 days and one author84 didn’t mention the exact timing. Intraoperative EDX ranged from 3 months77 to 19 months82 after birth. Strömbeck79 and Yilmaz81 performed serial EDX by conducting a second intervention later in time. Only 1 author74 performed the EDX three times over a period of 9 to 87 days after birth. In nine studies72-77,79,80,83 detailed information about the surface and/or needle electrodes, or set-up parameters is lacking. Eight studies49,75,77-80,83,84 mentioned the practitioners skills. Most studies71,72,76-78,81-83 reported prognostication by a clear outcome table, resulting from interpretation of the data acquired by the neurophysiologist. Three authors73,74,78 used dichotomisation of EMG results with a cutpoint chosen on analysis of the data, in particularly by splitting at the value which produced the largest difference in outcome between categories.
**Outcome Measurement**: The outcome measures varied widely. Two studies\(^{81,82}\) used the Active Movement Scale (AMS) scale.\(^{85}\) Birch\(^{72}\) preferred the Gilbert – Raimondi – Birch\(^{72}\) score whereas Bisinella\(^{76}\) and Chin\(^{77}\) opted for the Gilbert - Raimondi – Mallet score. Three studies\(^{49,70,71}\) adopted the Medical Research Council (MRC) scale. Toupchizadeh\(^{75}\) combined this with the Mallet score. Malessy\(^{73}\) and Strömbeck\(^{79}\) created their own protocol. Sensation, a prerequisite for full hand capacity, is only measured in 1 study.\(^{79}\) The gold standard evaluation for the presence of rupture or avulsion, operative exploration, was performed by 5 surgeons\(^{72,73,78,83,84}\). Four studies\(^{74,78,80,83}\) didn’t describe their outcome measurement. The mean follow-up time was 29 months. Five studies\(^{49,73,79,80,83}\) did not provide the duration of follow-up. Three authors\(^{81,83,84}\) compared the prognostic accuracy of radiologic and EDX findings.

**Study Confounding**: The confounding factors were addressed in only 2 studies\(^{75,78}\) and contributed largely to the overall risk of bias.

**Statistical Analysis and Reporting**: Three authors had a high risk of bias related to the statistical analysis and presentation of results.\(^{72,80,81}\) The QUIPS showed a low overall risk of bias for 4 studies\(^{75-78}\) and a moderate risk of bias for 10 studies\(^{49,70-74,79,82-84}\). Two articles\(^{80,81}\) presented a high overall risk of bias. (Table 1)

The overall quality of evidence of this review was approached by the GRADE in four categories (Table 2): EMG, NCS, EMG and NCS, EMG combined with CMAP and/or SNAP. All categories revealed very low quality due to study design (observational cohort studies), non-random selection of study population, failure of adequate control confounding, inadequately follow-up and publication bias. The intraoperative category is not considered due to the
diversity of applied techniques which makes pooling of data for rating a body of evidence impossible.

**DISCUSSION**

In this study we reviewed evidence on early prognostic value of EDX in the management of NBPP. The first relevant finding was the relative paucity of published literature on this topic. Furthermore, the methodological variety across the studies precludes the possibility of secondary data-analysis and represents a major limitation in drawing strong conclusions from the available evidence. Large differences were encountered in sample size, severity of primary involvement, applied treatment, surgery indications, EDX techniques and equipment, timing of EDX, definitions of recovery, outcome measurement tools, timing and length of follow-up. Evaluating its prognostic value, EDX outcome has been compared to many different outcome measures. Although direct surgical exploration may be considered as the gold standard for characterizing the extent of neurological injury type and levels, it carries significant morbidity and requires laminectomy to observe the intradural nerve roots. For this reason, we found this ‘gold standard’ evidence only for the subgroup of children with severe lesions in six articles.\(^{72,73,78,82-84}\) In the non-surgery population, outcome measures varied widely. As strength can recover up to 2 years after onset, the timing of outcome measures also impacts the studies.

Nevertheless, critical appraisal of the included studies supports an enhanced role of EDX for accurate estimation of prognosis. Thirteen\(^ {70-77,80-84}\) out of 16 articles demonstrated an additive predictive value of EDX. Prognostication of NBPP depends on the extent,\(^ {88}\) the severity of the lesion, recruitment in muscles supplied by the damaged nerve and the
individual speed of recovery.\textsuperscript{88,89} EDX should provide this information accurately. We will discuss the findings of the reviewed articles in the light of these prognostic criteria.

The extent of the lesion is a valuable guide to prognosis, which is for example much better for Narakas group 1 than for Narakas group 4.\textsuperscript{72} The extent is easily delineated by using a combination of NCS and EMG.\textsuperscript{18} Seven\textsuperscript{72,75,76,80,81,83,84} investigators combined EMG with NCS, CMAP and/or SNAP. Four studies\textsuperscript{49,73,74,79} performed solely EMG, which is useful for topographic diagnosis but a less quantitative source with regard to prognosis.\textsuperscript{22,25,27,30,90,91}

The severity of the lesion should be classified according to the degree of injury of axons and their supporting structures.\textsuperscript{92,93} Seddon\textsuperscript{92} and Sunderland\textsuperscript{93} defined, respectively, 3 and 5 grades of classification. EDX can help to distinguish between these lesions. Comparing\textsuperscript{89} distal and proximal CMAP gives an estimation of the conduction block. In the case of axonal loss, there are timing limitations for using CMAP: using this measure too early or too late can overestimate the number of viable axons. The CMAP’s can decrease within 24 hours and the SNAP’s a day later.\textsuperscript{45,91,94} None of the NCS in the included studies appears to have been performed so soon after birth and none of them related the possible impact of timing issue on CMAP amplitude. Normative data of upper-limb CMAP amplitudes in newborns and young infants are scarce. Therefore, Heise et al\textsuperscript{70,71} applied the axonal viability index, defined as the ratio of the amplitude of the CMAP of the involved side to that of the unaffected limb. This strategy has an added advantage in that it minimizes the effect of other confounding factors as age, gender, physical build and subcutaneous fat. The reproducibility of CMAP’s in newborns has been previously addressed, but in case of detecting NBPP children at risk for poor outcome, major differences are encountered between two arms. Sensory NCS have been stressed as an important prognostic tool for detecting root avulsions.\textsuperscript{95} Only 5
studies described the use of SNAP. Three investigators confirmed a very good sensitivity, ranging from 90% to 92.8%, for detecting nerve ruptures (complete transection or neurotmesis) and postganglionic lesions. The presence of normal SNAP or greater than 50% of normal compared to the uninvolved site or laboratory norms and EMG abnormalities in the correlated muscles is accepted as a main criteria for avulsion. The same authors found a sensitivity, ranging from 27.8% to 41.7%, together with a higher specificity, ranging from 41.9% to 85%. The difficulty arises when both preganglionic and postganglionic lesions are present, as this can lead to inaccurate labeling of the injury as an isolated postganglionic lesion. Serial EDX with persisting absence of MUAP’s and CMAP’s will further guide towards neurotmesis or root avulsions, both advocating for nerve surgery. Of course, the ideal workup requires the combination of clinical examination, EDX and imaging techniques as the strengths of each test often compensate for the other’s weaknesses. Smith et al evaluated the accuracy of EMG and imaging techniques together. They concluded that EDX outperformed imaging with regard to specificity and accuracy of identifying preganglionic injuries. The most marked difference was noted in the lower roots, with EDX being 87.5% specific at C8 and 78.6% at T1. Comparatively, imaging had a specificity of 29.4% at C8 and 57.1 at T1. Finally, the use of preoperative SEP has no value in determining severity of the plexus lesion as it might be confounded by the persistence of the fetal polyneuronal innervation of the somatosensory system in children with NBPP.

The third essential prerequisite for accurate prognosis estimation is information about recruitment in muscles supplied by the damaged nerve and the individual speed of recovery. Needle EMG adds information about muscle recruitment and monitors the nerve recovery process. However, in a developing child the use of EMG solely as prognostic factor might be confusing. Studies have shown underestimation and overestimation.
compared to clinical recovery outcome measures. Different explanations have been
previously published: pain inhibiting the muscle activation, the typical trick movements
observed in NBPP children, the baby's muscle physiology, presence of luxury innervation in
the newborn, developmental apraxia, aberrant reinnervation and equipment aspects
such as the needle recording area. Moreover EMG has important quantitative limitations.
As babies do not cooperate, recruitment analysis is seriously hampered. The recorded
potentials are often graded on an ordinal, though subjective, scale. They are difficult to
quantify, unless sophisticated qEMG models or macroEMG are performed. In the light of
these findings, we believe that the use of EMG solely, as performed in four included
studies, is not appropriate for prognostication. Needle EMG should be interpreted in
combination with NCS results. When performed at well-defined time frames, they inform the
clinician accurately about speed of recovery and the nerve recovery process.

Three authors performed intraoperative EDX to determine the anatomical extent and
severity of the NBPP. The combination of intraoperative root (SEP), mixed nerve action
potential (NAP) and evoked distal muscle response has an important prognostic value. Two studies demonstrate its use in improving the accuracy of intraoperative decision
making in case of neuroma in continuity. Such lesions represent an intermediate position
between axonotmesis and neurotmesis. Patients with less than 50% conduction across the
neuroma were treated with nerve resection and interposition grafting versus neurolysis
alone in patients with more than 50% conduction. This threshold is not yet validated
but future research could focus on analysing different cutoff levels. Furthermore, NAP
evoked by stimulating close to the foramen could differentiate between postganglionic and
preganglionic. The results of Pondaag et al are contrasting. In a consecutive series of 95
NBPP infants, the specificity for an absent NAP or CMAP ranged from 0.9 to 1.0 and its
sensitivity was <0.3. Chin et al\textsuperscript{77} augmented the sensitivity for C5 and C6 to 100%, respectively 87%, by combining NAPs, CMAP and gross motor response. Ideally, the reliability of the intraoperative EDX should be compared with the long term clinical outcome in a cohort of patients with no surgical intervention. Of course, this would be ethically unacceptable. Again, the use of different intraoperative EDX modalities combined with preoperative imaging and direct inspection further augments the prognostic accuracy.

The functional integrity of an anterior motor root cannot be monitored by SEP. Oberle\textsuperscript{107} described a technique with motor evoked potentials of the spinal muscles with a 100% sensitivity for anterior root lesions. This technique is not yet described in children with NBPP. Cortical motor evoked potentials could add information on the integrity of anterior roots, however special expertise in infants is required.

There are obviously challenges and limitations in using EDX as prognostic factor. EDX measures primarily reflect nerve and muscle function and not the internal architecture of the nerve, the quality of damaged nerve, or the anatomical variability.\textsuperscript{89} The reliability of the prognostic results of EDX should therefore be supplemented by clinical, neuroradiological and, in selected cases, surgical findings. Other challenges are the significant impact of the adequate timing and repetitions of EDX studies, the required skills and expertise of the neurophysiologist, the interpretation of data related to the pathophysiology and the developing child. The final EDX report should include all key prognostic information, preferably using standardized classification systems. The classification of Smith for preoperative EDX has been shown valid previously\textsuperscript{41,108} and could be combined with the cut-off values published by Heise et al.\textsuperscript{71} A similar classification system exists for intraoperative motor response to electrical stimulation.\textsuperscript{77} Some authors\textsuperscript{73,74,78} used dichotomisation of
EMG, which is in our opinion an incorrect simplification of a rather complex investigation. The reason for dichotomisation in the study of Van Dijk\textsuperscript{74} can be explained by their aim to develop EMG as a simple screening tool for referral of cases to specialized centres rather than a detailed prognostication tool.

Limitations
Evidence of publication and associated reporting bias is present, especially in the included studies with small sample size. As patients with inadequate recovery normally present to the recruiting hospitals, selection bias was present. Often a non-random sample of the NBPP population was included. Context-related quality aspects, such as technical equipment and skills of the investigator, might be an important cause of heterogeneity. EDX has, in common with all operator-dependent tests, intra-and interobserver variability. A clear and standardized definition of the prognostic factor was often lacking, which limits result comparability across studies. Some authors\textsuperscript{73,74,78} used dichotomisation of EMG, which introduces severe bias with creating overoptimistic results.\textsuperscript{109} Temporal influence (study duration and timing of intervention) might contribute to imprecision. The outcome of interest and the timing of its measurement caused selective outcome reporting. As NBPP involves a heterogeneous constellation of nerve injuries and is subject to an array of different treatment algorithms, the individual study results might be confounded. There is uncertainty regarding the degree to which the outcome is related to the natural history versus interventions. Moreover, the EDX might have changed the clinical decision management. The deleterious effects of the well-known secondary sequelae such as the medial rotation contracture or posterior (sub)luxation of the humerus, which may interfere substantially with neurological recovery, are important confounders. The deafferentation of
NBPP may also inhibit the development of normal motor control, which can remain abnormal even with natural or surgery repair.\textsuperscript{110}

The overall quality of evidence of this review, rated by the GRADE approach, is low. Therefore, conclusions should be drawn cautiously.

**Conclusion**

The most methodologically sound studies support the use of EDX, at standardized time frames, as a key prognostic modality complementing clinical judgement and neuroimaging. An accurate knowledge of underlying anatomy of the nerve injury helps to counsel families and the eventually planned reconstruction strategy.

We advocate for the following:

- the use of appropriately sized electrodes and stimulation probes.
- the use of a combination of NCS (including CMAP and SNAP) and EMG
- serial EDX at standardized timeframes, the first performed at the nadir of the loss of motor units and the second around the time of decision-making for surgery. The time of the nadir is currently unknown, and warrants further investigation.
- interpretation of results by an experienced and skilled neurophysiologist in the context of all the needle and NCS findings together. An extrapolation from adult neurophysiologic experience to the neonate is inappropriate.
- the following key elements in a final EDX report: diagnosis, timing of the lesion, location and extent of the injury, recruitment of affected muscles, evolution in time of the lesion and an accurate prognosis estimation.
Continued advances in imaging and EDX will further enhance accurate outcome prediction in NBPP.

We encourage further prospective, multicentre research with a large study population, clear in- and exclusion criteria, standardised EDX protocols, ICF based outcome measures at specific time points, a standardized conservative treatment scheme for all participants and precise measurement of the confounding factors. The length of follow-up should be at least 3 years. EDX should be serially performed by the same, neurophysiologist, with expertise in pediatric studies.

**Contribution to Authorship**  R.V.D.L conceived the study idea, contributed to the study conception, performed the searches, reviewed titles and abstracts, reviewed full text articles, extracted data, performed quality controls and drafted the manuscript. G.VDS reviewed titles, abstracts, full text articles, and reviewed the manuscript for important intellectual content.  E.T and L. L extracted data and performed quality controls. M.P contributed to the study conception and reviewed the manuscript for important intellectual content. C.VDB, M.D.M and G. V reviewed the manuscript for important intellectual content.

**Details of ethics approval** Not required.
Abbreviations

AMS: active movement scale

CMAP: compound muscle action potential

CT: computed tomography

EDX: electrodiagnosis

EMG: needle electromyography

MRC: medical research council

MRI: magnetic resonance imaging

MUAP: motor unit action potential

NAP: nerve action potential

NBPP: neonatal brachial plexus palsy

NCS: nerve conduction study

qEMG: quantitative electromyography

QUIPS: Quality In Prognosis Studies

SEP: somatosensory evoked potentials

SNAP: sensory nerve action potential

US: Ultrasound

References


64. Covidence Systematic Review Software, Veritas Health Innovation; Melbourne, Australia


**Figure Legend**

FIGURE 1: PRISMA flow diagram of study selection process in the systematic review

**Table Legend**

Table 1: QUIPS of the systematic review

L = low bias; M = moderate bias; H = high bias

Table 2: GRADE assessment of EDX techniques used in the systematic review

↓ -1 = downgraded with one level; ↑ +1 = upgraded with one level; 0 = item cannot be downgraded nor be upgraded; /= not applicable; LQ = low quality; VLQ = very low quality
<table>
<thead>
<tr>
<th>Publication</th>
<th>Study Participation</th>
<th>Study Attrition</th>
<th>Prognostic factor measurement</th>
<th>Outcome measurement</th>
<th>Study Confounding</th>
<th>Statistical Analysis and Reporting</th>
<th>Overall Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrisevic et al, 2014</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Birch et al, 2005</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Bisinella et al, 2003</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Chin et al, 2010</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Gopinath et al, 2002</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Heise et al, 2004</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Heise et al, 2007</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Heise et al, 2009</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Malessy et al, 2011</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Pondag et al, 2008</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Smith et al, 2018</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Strömbeck et al, 2007</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Toupchizadeh et al, 2010</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Van Dijk et al, 2012</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Vanderhave et al, 2012</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Yilmaz et al, 1999</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

L=low risk of bias; M= moderate risk of bias; H= high risk of bias.
<table>
<thead>
<tr>
<th>EDX Technique s (reference s)</th>
<th>Downgrade</th>
<th>Upgrade</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAR Limitation s (RoB)</td>
<td>Inconsistenc y of Results</td>
<td>Indirectnes s of Evidence</td>
</tr>
<tr>
<td>EMG (49,73,74,79)</td>
<td>LQ</td>
<td>High ↓-1</td>
</tr>
<tr>
<td>EMG and NCS (73,81,83,84)</td>
<td>LQ</td>
<td>High ↓-1</td>
</tr>
<tr>
<td>EMG and CMAP and/or SNAP (72,76,80)</td>
<td>LQ</td>
<td>High ↓-1</td>
</tr>
<tr>
<td>NCS (70,71)</td>
<td>LQ</td>
<td>High ↓-1</td>
</tr>
</tbody>
</table>

EDX= electrodiagnosis; EMG= needle electromyography; NCS= nerve conduction studies; CMAP= compound muscle action potential; SNAP= sensory nerve action potential; RoB= risk of bias; LQ= low quality; VLQ= very low quality; ↑= grade it up; ↓= lower it down; /= not applicable; 0= quality met criteria
Figure 1: PRISMA flow chart

Records identified through database searching (n = 2184):
- Pubmed: 493
- Embase: 1541
- Web of Science: 210

Additional records identified through other sources (n = 1)

Records after duplicates removed (n = 1892)

Records screened by abstract (n = 1892)

Records excluded (n = 1849)

Full-text articles assessed for eligibility (n = 43)

Full-text articles excluded, with reasons (n = 27)
- Intervention n = 4
- Population n = 1
- Outcome n = 4
- Design n = 12
- Language n = 3
- No full text n = 3

Studies included in qualitative synthesis (n = 16)