Supplementary data

Methods

Guideline committee

This guideline was developed and sponsored by the NOV, using governmental funding from the Quality Foundation of the Dutch Association of Medical Specialists in the Netherlands. The early preparative phase started in April 2011, and the guideline was officially authorized by the Dutch Orthopedic Association in February 2014.

The guideline committee consisted uniquely of members of the parents' association, pediatric orthopedic surgeons selected by the NOV, and a methodologist (KB) of the Knowledge Institute of Medical Specialists (KiMS). The methodologist was included to ensure a proper design and systematic evidence-based development of the guideline using the GRADE methodology, to meet all the criteria of the AGREE instrument.

Decisions within the guideline group were made by consensus. At the start of guideline development, all guideline committee members completed forms regarding conflicts of interests and this information was published together with the guideline.

Target group and aims

This guideline was developed for Dutch providers of healthcare for children with clubfoot, in particular (pediatric) orthopedic surgeons, but also pediatricians, gynecologists, obstetricians, and general practitioners. The main purpose of the guideline is to provide the best possible care for children with idiopathic clubfoot. This is achieved by informing healthcare providers about optimal treatment decisions, thereby reducing unwarranted variation in the delivery of care. The guideline is also meant to facilitate the development of uniform information for patients, parents, and caregivers.

Methodology and workflow

The guideline was developed in agreement with the criteria set by the advisory committee on guideline development of the Association of Medical Specialists in the Netherlands (Adviescommissie richtlijnen, 2011), which are based on the AGREE II instrument (Brouwers et al. 2010). The guideline was developed using an evidence-based approach endorsing the GRADE methodology, and meets all criteria of AGREE II. Grading of Recommendations Assessment, Development, and Evaluation (GRADE) is a systematic approach for synthesizing evidence and grading of recommendations offering transparency at each stage of the guideline development (Guyatt et al. 2011, Schunemann et al. 2014). The guideline development process involves a number of phases: a preparative phase, a development phase, a commentary phase, and an authorization phase. After authorization, the guideline has to be disseminated and implemented. Furthermore, uptake and use must be evaluated. Finally, the guideline must be kept up-to-date. It has to be revised every 5 years. Because the guideline consists of different modalities, stored in the Dutch Guideline Database, important changes can be incorporated earlier. Each phase involves a number of practical steps (Schunemann et al. 2014). As a first step in the early preparative phase, a broad forum discussion was held and all relevant stakeholders were consulted to define and prioritize key issues, which were extensively discussed in the guideline committee. The selected, high-priority, issues were translated into carefully formulated clinical questions. These questions defined patient problems, intervention, comparison, and outcomes. Furthermore, the patient outcomes relevant to decision-making were prioritized and minimal clinically important differences were defined.

In the development phase, the literature was systematically searched using the databases MEDLINE (Ovid), Embase, and the Cochrane Database of Systematic Reviews. Selection of the relevant literature was based on predefined inclusion and exclusion criteria and was carried out by one of the orthopedic surgeons (AB) in collaboration with the methodologist (KB). For each of the clinical questions, the evidence was summarized by the guideline methodologist using the GRADE approach. A systematic review was performed for each of the relevant outcomes and the quality of evidence was assessed in 1 of 4 grades (high, moderate, low, very low) by analyzing limitations in study design or execution (risk of bias), inconsistency of results, indirectness of evidence, imprecision, and publication bias. The evidence synthesis was complemented by a guideline committee member (AB) considering any additional arguments relevant to the clinical question, including patient values, preferences, and resource use (costs, organization of care issues). Evidence synthesis, complementary arguments, and concept recommendations were extensively discussed in the guideline committee. Then, final recommendations were formulated. The final recommendations are based on the balance between desirable and undesirable outcomes, the quality of the body of evidence across all relevant outcomes, values and preferences, and resource use. The strength of a recommendation reflects the extent to which the guideline panel was confident that desirable effects of the intervention would outweigh undesirable effects or vice versa, across the range of patients for whom the recommendation is intended. The strength of a recommendation is determined by weighing all relevant arguments together. This includes the weight of the body of evidence from the systematic literature analysis,

and also the weight of all complementary arguments formulated, the so called considerations. When using the GRADE approach, guideline panels must use judgement in integrating these arguments to make a strong or weak recommendation. Thus, although a low quality of the body of evidence from the systematic literature analysis will generally result in a weak recommendation, it does not a priori exclude a strong recommendation, and weak recommendations may also result from high-quality evidence (Schunemann et al. 2014).

After reaching consensus in the guideline committee, the concept guideline was subjected to peer review by all the relevant stakeholders: the commentary phase. Amendments were made and agreed upon by the guideline committee, and the final text was presented to the Dutch Parents' Clubfoot Association for approval and to the NOV for formal authorization. In this authorization phase, additional amendments were made to the guideline text based on the outcome of a general assembly of the NOV. The guideline was finally approved by the Dutch Parents' Clubfoot Association and officially authorized by the NOV.

Results of literature review and analysis

The following questions were formulated by the guideline committee:

1. What is the optimal treatment for clubfoot?
2. What is the importance of brace compliance and other patient-related factors in the successful treatment of clubfoot?
3. What is the optimal method to be used for the diagnosis and classification of a clubfoot?
4. Is screening for developmental dysplasia of the hip (DDH) in idiopathic clubfoot useful?
5. With respect to organization of care, what are the preconditions for optimal treatment of patients with clubfoot?

Below, the main (sub)questions are elaborated. For the complete recommendation on diagnosis and treatment of primary idiopathic clubfoot, see the article.

Clinical Question 1: What is the optimal treatment for clubfoot?

This clinical question consisted of several subquestions. The first subquestion should determine the preferable treatment, surgically (postero-medial release (PMR)) or non-surgically. The second subquestion judged the non-surgical treatments of clubfoot used in the Netherlands. Based on the available evidence, is the Kite or the Ponseti method preferred?

After first answering subquestions 1 and 2, some subquestions were added. In the Ponseti method, are accelerated treatment protocols useful and is the use of different plaster materials effective? Also, in the brace treatment there are differences within the Ponseti method, so the question of which brace needs to be used was addressed.

Our systematic literature analysis comparing surgical treatment (PMR) and non-surgical treatment (Ponseti) showed that the results of the Ponseti treatment were at least as effective as treatment with a PMR; level of evidence according to GRADE: LOW (Zwick et al. 2009, Adegbehingbe et al. 2010, Halanski et al. 2010, Clarke et al. 2011, Church et al. 2012). Mobility and also ankle and foot position in childhood are reported to be better after Ponseti treatment than after surgical treatment by PMR; level of evidence according to GRADE: LOW (Cooper and Dietz 1995, Dobbs et al. 2006, Edmondson et al. 2007, Zwick et al. 2009, van Gelder et al. 2010, Graf et al. 2010, Church et al. 2012). Serious clubfoot recurrence is reported less frequently after Ponseti treatment than after surgical treatment by PMR; level of evidence according to GRADE: LOW (Adegbehingbe et al. 2010, Halanski et al. 2010, Clarke et al. 2011, Church et al. 2012). A serious recurrence was defined as a recurrence with the need for intra-articular surgery (intracapsular or bony).

Quality of life in adulthood was reported to be higher after Ponseti treatment than after surgical treatment by PMR; level of evidence according to GRADE: VERY LOW (Cooper and Dietz 1995, Dobbs et al. 2006, Edmondson et al. 2007, van Gelder et al. 2010, Graf et al. 2010).

Cost-effectiveness studies considering advantages between 1 of the 2 treatments are scarcely available. One study reported 2 times lower costs for Ponseti treatment than for PMR in New Zealand. Also, in the USA, the costs related to the Ponseti method were lower than when an extensive surgical treatment was used (Halanski et al. 2010).

For any surgical treatment, one should consider local and systemic complications. Also, the severity and frequency at which these complications (risks) occur are factors that are important in deciding a favorable treatment. If a surgical treatment does not result in better treatment (effects vs. benefits) according to nonoperative treatment for the same deformity, the nonoperative modality is preferable. If one cannot proove significant advantages of a surgical procedure, the least invasive treatment should be recommended. If literature searches conclude that there is low evidence according to the GRADE classifications, these factors play an important role and can—although the literature-based evidence is low—lead to a strong recommendation.

Considering this aspect in relation to the above-mentioned literature, the guideline group came to a clear recommendation of the Ponseti method in preference to surgical strategies such as PMR. The systematic literature analysis also provided evidence in favor of using the Ponseti treatment in the primary correction of idiopathic clubfoot instead of using a treatment according to the Kite method; level of evidence according to GRADE: LOW (Segev et al. 2005, Sud et al. 2008, Sanghvi and Mittal 2009, Rijal et al. 2010).

In the Netherlands, besides the Ponseti method, Kite's method is also used as a non-surgical treatment of clubfoot. The literature analysis indicated that the Ponseti method was

more effective than Kite's method; level of evidence according to GRADE: LOW (Rijal et al. 2010).

Comparison of the standard Ponseti method with the accelerated Ponseti method showed similar effectiveness in primary correction of idiopathic clubfeet; level of evidence according to GRADE: MODERATE (Harnett et al. 2011). Because of greater and more widespread experience with the standard Ponseti method, using weekly serial manipulations, the guideline committee advised use of this method. The accelerated Ponseti method may be considered if poor compliance during the plaster phase is to be expected and if the hospital outpatient scheduling permits it.

In primary correction of idiopathic clubfoot, the available evidence indicates that the standard use of plaster of Paris in the Ponseti treatment is more effective than using synthetic plasters; level of evidence according to GRADE: LOW (Pittner et al. 2008). Moreover, long-term results are decisive and are not yet fully known for Ponseti treatment using materials other than plaster of Paris. Therefore, the guideline committee advised that the standard plaster of Paris should be used for the Ponseti treatment.

Ponseti treatment using a Dennis-Brown type of foot abduction brace resulted in greater effectiveness in preventing recurrences than Ponseti treatment followed by ankle/foot orthotics according to the literature analysis; level of evidence according to GRADE: VERY LOW (Janicki et al. 2011). The guideline committee recommended the use of a Dennis-Brown type of foot abduction orthotic with a bar, as used in the standard Ponseti method. Further scientific research will be needed to support future brace modifications.

Clinical Question 2: What is the importance of brace compliance and other patient-related factors in the successful treatment of clubfoot?

Firstly, the significance of the severity of clubfoot in relation to treatment success using the Ponseti method was addressed. Secondly, compliance regarding wearing of an abduction brace in relation to the success of Ponseti treatment was an important issue to investigate. Thirdly, the importance of other patient-related factors in relation to the treatment success in Ponseti treatment was investigated.

The literature search for these questions was similar to the search used in the treatment section. Due to heterogeneity between the studies and a lack of quantitative data required, a meta-analysis was not possible.

In the literature on wearing a foot abduction brace, different definitions of compliance are used. The duration of wearing the brace differs in different treatment protocols, and as a result non-compliance is defined differently in the studies that are included. Another important finding is that in all studies, patient compliance was reported by the parents and may have been overestimated. Even so, taking these limitations into account, a statistically significantly reduced risk of recurrence was found in compliant patients in all but one of the studies included. Only in the study by Halanski et al. was no statistically relevant relationship found, but the statistical power of this study was limited because of the small size of the patient sample (Halanski et al. 2010). Non-compliant patients regarding foot abduction brace wear are strongly positively associated with the risk of recurrence, although no strong scientific evidence is available; level of evidence according to GRADE: MODERATE (Dobbs et al. 2004, Morcuende et al. 2004, Ponseti et al. 2006, Changulani et al. 2006, Haft et al. 2007, Avilucea et al. 2009, Bor et al. 2009, Halanski et al. 2010, Ramirez et al. 2011).

A relatively large distance to the treatment center might possibly have a negative influence on treatment success. Treatment in a local urban patient group resulted in significantly greater improvements in Pirani score and a lower risk of recurrence; level of evidence according to GRADE: MODERATE (Avilucea et al. 2009).

The other statistically significantly positive associations are a married marital status of the parents, private insurance status, and high educational level and income level; level of evidence according to GRADE: MODERATE (Avilucea et al. 2009). It is important to mention the use of univariate data analysis; without correction for potential confounders, the relationship between the different factors remains unclear. Some of these factors could indirectly affect the treatment results, from their effects on compliance regarding wearing of the foot abduction brace. Considering the low statistical power (small study sample size), the univariate analyses without adjustment for confounding variables, the testing of a wide range of factors (multiple testing), and the large heterogeneity between the studies, the guideline committee concluded that these findings cannot be used to predict the success rate of Ponseti treatment.

The severity of clubfoot (determined at the start of the treatment) is a likely prognostic factor regarding the risk of recurrence (Dobbs et al. 2004, Morcuende et al. 2004, Changulani et al. 2006, Haft et al. 2007, Avilucea et al. 2009, Bor et al. 2009, Halanski et al. 2010, Ramirez et al. 2011, Zhang et al. 2012). The guideline committee therefore recommends that parents should be warned that in cases of severe clubfoot, the duration of the treatment and the number of plaster casts could be higher than usual. Caregivers are advised to register factors such as severity, mobility, and brace compliance. Each of these factors can influence the treatment outcome. By frequent contact between parents and caregivers this possible influence can be monitored and if necessary positively influenced.

Clinical Question 3: What is the optimal method to be used for the diagnosis and classification of a clubfoot?

Because there is international consensus on defining the deformities in clubfoot, a systematic literature analysis was not required. Instead, authoritative standard works in orthopedics were used (Hefti 2007, Herring 2007).

A clubfoot consists of 4 typical entities: equines (mid- and hindfoot), varus (hindfoot), cavus (midfoot), and adduction

(forefoot). Often there are typical folds in the sole of the foot, an altered heel formation, and an altered formation of calf and peroneal muscles. These typical, clinically addressable entities can reliably lead to the diagnosis of clubfoot.

The additional value of radiological examination as a diagnostic tool for clubfoot is limited (Ponseti et al. 2006) However, radiological examination can be of additional help when there is doubt about progress in treatment. Radiographs can be used to follow treatment progress (Simons 1993). The guideline committee supports the opinion that radiological examination is not primarily indicated in diagnosing clubfoot.

No uniform, generally accepted classification system for quantification of the severity of the clubfoot deformation is used—either in the Netherlands or worldwide. A classification system would ideally provide a solid prognosis for treatment outcomes and also a method to measure treatment outcomes. Ideally, the scoring system for clubfoot would be reliable and reproducible and would include separate information on the different parts of the foot and its position in 3 dimensions, and provide information on stiffness or flexibility. Last but not least, an ideal scoring system would have strong predictive value regarding treatment outcomes.

The severity of a clubfoot can be identified using various classification systems. The value of these classification systems was systematically analyzed according to patient-related prognostic factors in treatment outcomes. Whether or not the severity of the clubfoot is prognostic of relapse has not been reported. The classification systems according to Pirani (Pirani et al. 1995) and Diméglio (Dimeglio et al. 1995). are most frequently reported. No significant differences in reliability or reproducibility have been described (Wainwright et al. 2002), and there are currently no clear reasons to prefer one system to the other. The guideline committee recommends the simultaneous use of both of these classification systems, in order to facilitate comparison of data in the future. Scoring of the foot has to be performed at the start of the treatment as well as during the follow-up, to evaluate correction and notice negative alterations early on.

Clinical Question 4: Is screening for developmental dysplasia of the hip (DDH) in idiopathic clubfoot useful?

The systematic search revealed 5 studies in which the incidence of DDH was determined. All but 1 (Canavese et al. 2011) compared the incidence of DDH in clubfoot with already published data in a healthy population without clubfoot. Important again here is the definition of the disease: in one study, DDH was defined as an acetabular deficiency that must be treated; in other studies, every sonographic or radiological abnormality is mentioned as a DDH. The studies selected from our search mostly found an increased incidence of DDH in the clubfoot population, although the cases detected with DDH often do not need any treatment; level of evidence according to GRADE: LOW to MODERATE (Westberry et al. 2003, Carney and Vanek 2006, Paton and Choudry 2009, Perry et al. 2010, Canavese et al. 2011). New study results on this topic are expected in the near future. This new literature can be incorporated in a revision of the guideline. This is one of the great advantages of a dynamic, modular guideline format. New insights or new literature can easily be added.

C*linical Question 5: Regarding organization of care, what are the preconditions for optimal treatment of patients with clubfoot?*

Currently in the Netherlands, every orthopedic surgeon is allowed to treat children with clubfoot. Of the 800 orthopedic surgeons, around 75 are members of the Dutch Pediatric Orthopedic Association, but not necessarily every one of them is treating clubfoot on a regular basis. An unknown number of non-members also treat children with clubfoot. With an estimated incidence resulting in 200–300 clubfoot newborns every year, the annual number of clubfoot patients per orthopedic surgeon will be low.

For this clinical question the optimal process of patient referral, including dividing and allocating responsibilities according to the caregiver, has been addressed. The guideline committee formulated recommendations on how integrated care should be optimized and guaranteed. To answer this question no systematic research was conducted. An important issue was the question of how treatment should be organized nationally and the way in which centralization and specialization should be established. In addition, directly related issues concerning the qualifications and training of orthopedic surgeons are also mentioned. Finally, information and communication with patients, including the reimbursement of prenatal counseling, has been addressed.

The guideline group discussed the optimal process in referring patients with clubfoot. If, based on sonographic investigation, clubfoot pathology is suspected during pregnancy, the parents should be referred to a specialized member of a clubfoot treating team for prenatal counseling. Dutch insurance will carry the costs, as consultation is part of the mother's insurance in the period of the pregnancy. During this consultation, the protocol according to diagnostics and treatment can be discussed and the logistic process can be explained. Prognosis can also be discussed, although this is mainly covered by the severity of the pathology to be addressed after the child is born. If the clubfoot is diagnosed just after delivery, the gynecologist, midwife, or general practitioner must contact the local orthopedic surgeon who can transfer the child to a specialized center. According to the responsibilities within the care process, the guideline group stated that the orthopedic surgeon should be the leading member of the treating team. He or she should always diagnose the pathology and initiate the correct treatment.

Adequately trained members of the clubfoot team should perform serial manipulation and plastering themselves. Members of the team should also explain the features of the foot abduction brace and take care that there is a good fit. The

treating team should consist of at least 2 trained orthopedic surgeons, 2 trained plaster physicians, and—if necessary—a technician.

The optimization of the clubfoot care and aftercare should be coordinated by the orthopedic surgeon. Frequent contact with the patient and the parents appears to be essential. In the primary phase of the treatment until the start of brace wearing, the patient should be checked frequently, mostly on a weekly basis. After at least 2 weeks of starting with the brace, the child should be checked clinically, continuing with check-ups between 3 and 6 months. This is especially important due to the higher risk of residual deformities in non-compliers (See Clinical Question 2).

The guideline committee does not involve physiotherapists in the primary treatment initiated before the age of 6 months, because there is no evidence that it would have positive effects. A walking child with locomotive deficits can of course be treated by a physiotherapist, although thorough research on the effectiveness of this therapy would help formulation of recommendations on this subject.

Because of the population density in the Netherlands and the high number of hospitals in a relatively small area, patients never have to travel far to reach care—although, of course, the traffic increases the travel time. If care is centered because of a low number of patients, the implication is a longer travel time. Based on a questionnaire filled out by their members, the Dutch Clubfoot Parents' Association stated that if the care were to be of a higher quality, there would be no objection at all to such an investment of time (Questionnaire NVK 2012). Training of caregivers should be guaranteed by nationally or internationally certified courses, orthopedic surgeons should attend (certified, GAIA) refresher courses, and be members of the Dutch Pediatric Orthopedic Society. Because of the higher number of patients treated in a clubfoot center, the infrastructure can be kept at a high standard. The NOV should audit the centers on an annual basis and publish the results. A specialized center should have a website with up-to-date information, and should also refer to information that is available on international websites.