JÖNKÖPING UNIVERSITY
School of Health and Welfare

Comparison of orthotic interventions for patients with congenital talipes equinovarus: a systematic review

PAPER WITHIN: *Orthotics*
AUTHOR: *Josefin Larsen & Matilda Strömbäck*
TUTOR: *Cleveland Barnett*
JÖNKÖPING *2021-04-30*

Sammanfattning

Syfte: Syftet med detta arbete är att undersöka olika ortos-lösningar med dess inkluderade resultat för patienter med kongenital idiopatisk klumpfot.
Metod: En systematisk sökning i databaserna MEDLINE, CINAHL, PubMed and Scopus genomfördes och väsentliga studier inkluderades utefter de förbestämda kriterierna. Inkluderade studiers validitet och möjliga partiskhet bedömdes samt att relevant data utifrån frågeställningen analyserades och besvarades.
Resultat: Resultatet presenterade 15 olika typer av ortoser, i 21 olika artiklar. De presenterade utfallsmåtten var inom komplians, återfall av deformationen, Pirani- och Dimeglio poäng, rörelseomfång samt inom funktionellt resultat.
Slutsats: Baserat på resultat går det inte att presentera en övervägande slutsats om vilken ortos-lösning som ger bäst resultat. Framtida forskning måste utveckla och framställa ortoser som tillfredsställer och möter patienternas behov.

Nyckelord: Klumpfot, PEVA, behandling, ortos, litteraturstudie

Summary

Aim: The aim with this study is to compare different orthotic solutions for pediatric patients with congenital talipes equnivarus.
Method: A search in the databases MEDLINE, CINAHL, PubMed and Scopus were made, and studies was included after the predefined criteria. The included studies were reviewed for risk of bias and validity, relevant data was extract and analyzed with regards to the research question.
Result: The result was presented on 15 different orthotic interventions in 21 different articles. The reported outcome measures were compliance, recurrence of deformity, Pirani- and Dimeglio score, range of motion and functional outcomes.
Conclusion: Based on the results a recommendation cannot be made on what orthosis that gives the best result. Future research should focus on developing and designing an orthosis that satisfies the patient's needs.

Keywords: Clubfoot, CTEV, treatment, orthosis, review

Table of content

Glossary

DBB= Dennis brown brace
DB= Dobb's brace
MS= Mitchell shoe
FAO= Foot abduction orthosis
FAB= Foot abduction brace
AFO= Ankle foot orhtosis
SFAB =Stenbeeck foot abduction brace
MPB= Mitchell-Ponseti brace
LLO= Lower leg orthosis

Introduction

During our clinical placement at two different prosthetic and orthotic clinics in Sweden, different orthotic treatment for children with congenital clubfoot was observed. This made us question why different orthotic design is used and if there is any proven evidence if it's leading to any different outcome. The common congenital deformity clubfoot occurs in somewhere between 1:250-1:1000 newborn children. Even if the misalignment of the foot or feet's is classified as sever, by giving treatment early the most cases end up with a good result (Edinger et al., 2019). Untreated cases may, however, contain lifelong disability, deformity, and pain as describe by Ansar et al. (2018). The primary standard method used for clubfoot treatment is the Ponseti method. Which includes several parts as manipulation with serial casting, sometimes Achilles tenotomy and bracing. Foot-abduction braces, as the Dennis brown brace is considered to currently be the most used brace (Edinger et al., 2019). In Sweden there is no define national clinical guidelines for treatment of clubfoot. But the Swedish pediatric orthopedic quality register collect data about the currently used treatment method and its result, which is presented in their annual reports. When more sufficient data is collected, presented and analyzed that is thought to be used as a base for national clubfoot-treatment guidelines in Sweden (SPOQ, n.d). Several systematic reviews are published within the area clubfoot treatment, for example Ganesa et al. (2017), Gelfer et al.(2019), Jowett et al.(2011) and Zionts & Dietz (2010). While all these articles either compare the Ponseti method or/and different parts of the Ponseti method, the general conclusion is that the recurrency of deformity is generated by non-adherence with bracing. The recurrence rate of clubfoot treated with the Ponseti method differs in research, but it has previous been stated between 11%-48% (Haft et al., 2007; Morcuende et al., 2004). Despite all previous research, there seems not to be an agreement of which orthosis that should be used in clinical practice.

Previous research

Within the area of different methods for congenital clubfoot treatment some reviews are published. Ganesa et al. (2017) compare the Ponseti method with the Kite technique. The Ponseti method indicated an effective treatment method for correction of clubfoot, but the rate of recurrence was notable high. Ganesan et al. correlated this with a lack of brace adherence and socioeconomic factors. Systematic reviews that presents and analysis within the Ponseti methods in itself are for example published by Jowett et al.(2011) and Gelfer et al.(2019). In these studies, all the different parts of the Ponseti-treatment are included, such as the serial casting, surgery, and bracing. Both studies demonstrate that the Ponseti method is the currently best presented method in terms of successfully results, but as well in these studies non-compliance with the brace is stated as the effect creating relapse of deformity. Zionts & Dietz (2010) has in their published systematic review focus on the orthotic interventions that are presented within the Ponseti method. The authors include different alternatives of foot abduction braces as the Markell brace, Dobbs dynamic and Steenbeek brace. The authors does present different aspects associated with recurrence of deformity in these foot abduction braces such as nonadherence with the brace (Zionts & Dietz, 2010). Generally seen, whether it is a systematic review focusing on a comparison of the Ponseti treatment or an analyzation of the including parts of the treatment, the evidence in the last decade presents high rates of relapses since lack of brace compliance.

Background

Talipes equinovarus

Talipes equinovarus also known as Clubfoot affect between 1:250- 1:1000 newborn children and is classified as one of the most common congenital limb deformities. Approximately half of all the clubfoot cases is bilateral, and approximately 20% of the cases are related with other congenital abnormalities. The male to female ratio in the deformity is 2:1. No exact cause has been found why clubfoot occurs, but different factors as environment and genetic inheritance are suspect as contributory. The misalignment is characterised by sever deformity in forefoot, midfoot and hindfoot. Adductus occurs in the forefoot, cavus occurs in the midfoot and the hindfoot is affected by a rigid equinovarus. This results in medial navicular and cuboid subluxation and internal rotation of the calcaneus which displace it from the heel pad. Plantar, medial, and posterior contractions is created by shortening of the skeletal structure. The misalignment is also related with hypoplasia in the calf as well as potential leg length discrepancy (Edinger et al., 2019). Clubfoot is divided into three different types: idiopathic which means that there is no known reason for it, neurogenetic which means that it is secondary due to a neurological condition like cerebral palsy and syndromic which means that it is due to another underlaying syndrome. Often children get diagnosed with ultrasound during the pregnancy but can also be diagnosed at or after birth. Children born with clubfoot needs to be treated and the treatment generally is divided into two parts, casting followed by bracing (John Hopkins University, 2021).

Methods for correction

Nowadays the most common way for treating clubfoot is by using the Ponseti method but historically other methods such as the French functional method and the Kite technique has been used. The French method included stretching, assisted movement, and taping for maintenance. The Kite technique included long leg casting for manipulation and maintenance. But as these methods has been insufficient to reduce the deformity in a large proportion the Ponseti method has been introduced, and now used as a standard method (Edinger et al., 2019). The Ponseti method treats the affect foot or feet by manipulating and holding through plaster casting. The effected foot is in generally casted five to seven times and each time the foot is corrected more until it is fully corrected or even overcorrected since the chance of relapse is high. The last cast is used for about 3 weeks. This procedure takes about two months if the case is not classified as severe. In most cases an Achilles tenotomy is done in conjunction with the last casting procedure to allow full ankle dorsiflexion. Then as ending the treatment an orthosis should be used. The first three to four months the brace is used around 23-hours per day. Then during sleep for the following years, number of years varies individually, but often somewhere between 2-5 years. The function of the orthosis it to maintain and hold the affected foot. By using an orthosis, the chances of getting a deformity relapse are decreased (Edinger et al., 2019). The splint that is often used after the casting procedure is called a Dennis Brown splint and is named after an Australian surgeon. The orthosis is also referred to as the foot abduction brace. The splint is made from two boots and a bar. The boots are placed on the bar and should be shoulder length from each other (The Royal Children's Hospital, 2007). A common setting could be 70° external rotation and 15° dorsiflexion (Edinger et al., 2019).

The Swedish pediatric orthopedic quality register describes 3 different orthotic interventions used for clubfeet in Sweden, the foot-abduction-orthosis, knee ankle foot orthosis and the ankle foot orthosis (SPOQ, 2018). The registry has in their annual reports not state if it is any correlation between used orthotic design and result (SPOQ, 2019). The biomechanical principal behind most of the designs and used orthoses includes external forces on the affected foot creating pronation, dorsiflexion, abduction, and an external rotation. In addition to the mentioned orthoses other interventions has been used, for example the shoes with straight or reverse lasts (Edinger et al., 2019).

Terminology within orthotic device

The international organization for standardization (ISO, 2020) has a standard of general terms that should be used for description within external orthoses. An orthosis should describe the body and joint that is encompassed. The orthosis should also be referred to as if it is prefabricated or custom fabricated and its biomechanical function. The biomechanical function can for example be to aim for prevention, maintenance or reduction of deformity (ISO, 2007).

Reported outcome following clubfoot correction

How researchers and clinicians use, measure, and evaluate the result of clubfoot treatment differ significantly. In Gelfer et al. (2020) systematic review the authors present and analyze the used outcome measures within idiopathic treatment according to the Ponseti method. In total the review present 36 different outcome measures (Gelfer et al., 2020). Below is a presentation of some of the commonly used outcome measures.

Recurrence of deformity

Recurrence of the clubfoot deformity that occurs at or after the initial treatment is commonly described as an outcome following the Ponseti method. The proportion of patients that suffer from recurrence differ, Morcuende et al.(2004) present a relapse rate of 11% and Haft et al. (2007) presented a relapse rate of 41%. Dobbs et al.(2004) also describe that the risk of recurrency is primarily during the orthotic procedure before the age of 4 years.

Compliance with the brace

There are different ways to investigate compliance, in the study done by Garg & Porter (2009) compliance is measured based on how adherent the patients and parents are to the treatment protocol based on reported hours of use. Other ways to measure compliance can be pressure sensors in the orthosis (Kuzma et al., 2020) and parents' ability to take on and off the orthosis (Manousaki et al., 2016).

Pirani score

A way to measure the outcome of the Ponseti treatment is to use the Pirani scoring method. Pirani uses a score from 0-1 for each category, where 0 is normal, 0.5 is abnormal and 1 is severe abnormality. The maximum score can range from 0 to 6 where 0 is no deformity and 6 is severe deformity. When conducting this test six different characteristics for clubfoot is evaluated at and they are divided between signs in the hindfoot and signs in the midfoot. The hindfoot is scored in posterior crease, amount of flesh in the heel, rigidity of plantarflexion. The midfoot is scored based on crease on the medial side of the foot, curvature of the foot and how prominent the talus bone is (Dyer & Davis, 2006)
Jain et al. (2017) has done a study regarding the validity of the Pirani score and concluded that it is a reliable assessment tool. Gelfer et al. (2019) have demonstrated that the Pirani score is a reliable way for surgeons to assess a clubfoot during treatment, thereby not yet validated for assessment done by other healthcare professions.

Dimeglio score

Dimeglio score is used for an assessment of clubfeet, grading on a scale of 0-20, where 0 is a corrected foot and 20 is the most severe clubfoot. The scoring is divided into 4 different parts: equines, varus, rotation around the talus and adduction of forefoot. Each of these can get a score of 0-5, 0 being no deformity and 5 being severe deformity. The scores can later be divided into severity where 0-5 is grade 1, 5-10 being moderate deformity and grade II, 10-15 being severe feet and grade III and 15-20 being very severe feet and grade IV. This method includes a checklist and there is training material for new users of the Dimeglio score (Diméglio et al., 1995).

PBS- score

The Swedish pediatric orthopedic quality register present PBS-score as a way of scoring a foot based on its appearance and function. This is a validated test for children with clubfoot. This scoring system is explained as a way of assessing ambulating children that have a history of clubfoot and should therefore not be used when assessing an infant (Böhm & Sinclair, 2019).

Functional outcome score

Another way to assess clubfoot in children is the Functional outcome score. This assessment method is like the PBS-score, looking at ambulating children and should therefore be used on children that has already gone through correction of clubfoot and is not relevant when looking at assessing a foot during the treatment (Dietz et al., 2009).

These presented outcome measures are only a selection of several different outcomes that are used within research and clinical practice of clubfeet-treatment (Gelfer et al., 2020). All these outcomes provide good conditions for evaluation of the treatment. However, this also creates difficulties when it comes to comparisons between different studies since all the presented outcomes are not translatable to each other.

Aim

The aim of this thesis is to investigate how different orthotic interventions for pediatric patients with idiopathic congenital talipes equinovarus differ in outcomes, including maintenance of foot alignment, foot mobility as well as compliance-rate.

Research question

In pediatric patients with congenital idiopathic talipes equinovarus does the traditional treatment of foot abduction orthosis compared to other presented orthotic interventions create a different result in reported outcomes?

Method

As described by Dickson et al. (2017) a systematic review has the design to detect and assess the primary available evidence in a research area. Thereby a systematic review was chosen as study design to address the aim and research question. The PRISMA Checklist for systematic reviews from 2009 was followed for structure during the process (Liberati et al., 2009).

Research question

The research question is developed from the PICO-model, which is commonly used when structuring a clinical research. PICO is an acronym where the P stands for population, I stand for intervention, C stands for comparison and O stands for outcome. By using this method, it could easily be broken down into the different parts that was going to be investigated (Del Mar et al., 2017). In this project the population were children with congenital idiopathic talipes equinovarus and the intervention was foot abduction orthosis compared to other orthotic interventions and the outcome was not specified, instead all reported outcome measures was of interest.

Eligibility criteria

After deciding the research question and area of interest the criteria for inclusion and exclusion were determined. The different criteria´s were based on the main concepts from the research question.
The included scientific articles all needed to be peer reviewed and written in English. All the articles also needed to be available in full text through Jönköping University library. The population of interested were participants with bilateral or unilateral congenital idiopathic talipes equinovarus, diagnosed before, at or after birth. All the patients needed to be treated according to the Ponseti method and thereby the treatment needed to include an orthotic intervention. The gender, age or ethnicity of the patients was not considered relevant as inclusion or exclusion. Reviews were excluded as well as articles published before the year 2000. The reason for only including evidence published after the year 2000 is based on the large research increase within this area that has taken place in the last two decades (Gelfer, et al., 2019).

Search strategy and data collection

An image of the final search can be seen in appendix 1, it includes all used terminology, MeSH-terms and number of hits for each database. Before the final search were compound and conducted, different scoping searches were made. Which are more simplified searches for prescribing the authors an overview over the published evidence, terminology and key issues withing the topic (Pilkington & Hounsome, 2017). Based on the scoping searches the final search begin to be constructed. The desired databases were MEDLINE, CINAHL, PubMed and Scopus as these are covering much in health science (Monash University Library, 2021). In discussion with Jönköping University Library these databases were confirmed as most relevant within the topic. The base in the final search was the research question that earlier were divided and structured according to the PICO-model. According to Booth (2017) a search can be conducted based on only the words in the PICO, but synonyms can also be included to make a broader search. It is common to only include the population (P) and the intervention (I) when conducting a search (Booth, 2017). As this project is based on a comparison between different orthotic interventions the authors desired to include the comparison (C) factors as well. In this case the search based on the desired PICO would be: P as children with congenital talipes equinovarus, I as foot abduction orthosis (often referred to Denis brown splint) and C as other orthotic solutions. In this project the O included all types of outcome measures but were desired to not include in the final search since the scoping searches identified a much smaller number of articles.

In all databases different searches was done for the synonymous of the population, intervention, and comparison. This was done separably to investigate the amount of result and its relevancy. Medical subject hedings (MeSH) were included in the allowing databases. As Dundar & Fleeman (2017b) describe, MeSH-terms is based on catalogues on different subject headings. By adding MeSH-terms different terminology within the same concept will be detected in the result, which gives the opportunity for a more comprehensive search (Dundar & Fleeman, 2017b). In MEDLINE and CINAHL MeSH-terms was marked by MH and brackets. The MeSH terms was put in quotation marks to indicate that the whole fraise should be coherent. Between synonyms OR was used and between different subjects AND was used. In PubMed the MeSH terms was identified by [MeSH] and the word in quotation marks to make sure that the words was coherent in the search. In Scopus the function of medical subheading

does not exist, the terms and synonymous were thereby separated and added for a broader search. Quotation mark was used for indication that the words needed to be coherent. When all the terms were combined the final search were made.

Articles that were written earlier than 2000 was directly excluded and results that was written in another language than English were excluded as well. By PubMed's exclude function, non-peer reviewed articles were directly excluded. The final search in all different databases were conducted 2021-03-22.

Screening process

Duplications were removed by the authors manually before the screening process started, see detailed number of duplications in figure 1. The screening process was done in three parts, that includes screening of title, abstract and full-text separately (Dundar & Fleeman, 2017a). The first part included identification and application of the different article-titles against the predetermined inclusion/exclusion criteria's (see criteria under *Eligibility criteria).* As a result of that, many articles that were not consider relevant in terms of the aim could be sorted out. In the next step, the abstract was read by the authors. Bas on the pre-determined criteria's the different articles were either included or excluded. The last step was to read the whole study and determine if it was relevant based on the pre-determined criteria. If a study passed all the three steps whey were included further in the review for data extraction, quality assessment and data analysis. See detail number of excluded articles present in figure 1. All these steps were made by both authors independently. The program Rayyan was used to present the result of articles from the final search. The authors were in that program blinded from each other's result when screening the title, abstract and full text. Any disagreement for either inclusion or exclusion that existed were primarily discussed between the two authors and secondarily with a third part, in this case the supervisor of the thesis.

Outcome

All the measured and presented outcomes related to the orthotic treatment were of interest in the analysis of the result. A specific outcome was not specified in this systematic review since this made the number of articles to limited, which was confirmed in the scoping searches.

Assessment risk of bias

SBUs critical appraisal templates are customized for systematic inspection of methods within different areas of healthcare (SBU, 2020). In support of identification with potential biases in the included articles, assessment templates published in (SBU, 2021) were used. Depending on the individually study design, suitable templates was applied.

In total 7 of the included articles did not state study design, see table 1. In these articles the SBU template for Assessment of non-randomized studies of interventions (effect of being assigned an intervention) were used. The authors of this thesis consider that template most appropriate accordingly to the individually information in the descriptive method sections. In total the three used templates were "Assessment of non-randomized studies of interventions (effect of being assigned an intervention (ITT))", "Assessment of non-randomized studies of interventions (effect of completing an intervention (per protocol))" and "Assessment of randomized trials (effect of being assigned an intervention (ITT))", see appendix 2 for templates. The assessment tables are written in Swedish, but are according to SBU either fully or partly translated of Cochrane risk-of-bias tools RoB:2 and ROBINS-1 (SBU, 2021). RoB:2 is design for randomized trials (Cochrane Methods Bias, n.d) and ROBINS-I are designed to use in non-randomized intervention studies (Sterne et al., 2016).

To ensure a common understanding of the included questions within the assessment tool the authors together conducted a test application of the different templates in three different non-included studies. When both authors agreed on the implication of all different questions, all 21 studies were appraised separately and independently. Different bias-related categories such as: bias from randomization or selection, confounding bias, classification of intervention bias and intended intervention bias, missing data bias, measuring outcome bias, reported result bias, reporting bias and conflict of interest were appraised and answered. The applicable answer was "yes", "probably yes", "probably no", "no" or "information is missing". An overall assessment as either high, moderate, or low risk of bias was decided in each category. The individual assessments were presented and discussed between the authors. Disagreements were investigated and a joint decision was made.

Internal validity, external validity and precision assessment

An assessment of the internal validity, external validity and precision in the included studies were made with support of the SBU "granskningsmallar" (translated = review templates), see appendix 3. The templates are based on question withing the categories internal validity, external validity, and precision. Each question was multi-answerable and had a defined score. The categories were defined by their total scores. To ensure a common understanding of the included questions within the assessment tool the authors together conducted a test application of the different templates in different non-included studies. When both authors agreed on the implication of all different questions, the assessment of the included articles was done individually. The individual assessments were presented and discussed between the authors. Disagreements were investigated and a joint decision was made.

Analysis of data

The primary outcome measures that were demonstrated in the included studies were recurrency of deformity, compliance with the brace, Diméglio score and Pirani score. Other outcomes that were included but only appeared in a small number of studies were functional outcomes and range of motion, see table 5 for a presentation of each studies and their outcome measures.

An analysis of each used and present orthotic intervention was made. Which included the orthosis design and function, based on the written description and if available also images and videos. The different orthosis that was equal and not equal in function and design could be identified and presented in the same category, see figure 2. Recurrence in relation to the used orthosis was in most of the included articles defined as recurrence of deformity needing additional treatment such as serial casting, Achilles tenotomy or/and re-bracing. The recurrence rate also described as relapse rate, were in the studies reported in either number of participants or number of feet. For clarification the results of recurrence for each article were decided to be presented in percentage of the total number of participants. To be able to analyze this data it was decided by the two authors to define different percentage limits for low, moderate, and high rate of relapse. 10% and below were decided as low rate of relapse, at 11% and up until 49% as moderate rate of relapse and at 50% and up till 100% as high rate of relapse. As previous studies investigating the recurrence rate of deformity within the Ponseti method has demonstrated big differences, the authors decided to define low rate of relapse as 10% and below, which approximately is the value that is presented in Haft et al.(2007) study. The other thresholds had no evidentiary support.

The definition and how to measure compliance and non-compliance differed between the included studies. For example, some studies defined compliance on which proportion of time the orthosis was used, while other conducted interviews with the parents/caregivers. Thereby, an analysis of each individual study and measuring tool was made. Some of the articles also presented skin problems in relation to the orthotic use. In several of those studies, various skin problems were presented in a subcategory of compliance. Any reported skin problems were thereby decided by the authors to be included and analysed in the category compliance. In the same manner as in the analysis of relapse, the compliance rate was presented in percentage of the total number of patients. To be able to analyze and draw any conclusions from the individual compliance rate an individual definition of poor-, moderate- and good compliance was determined by the two authors. The definitions were: poor rate of compliance as 50% and below, moderate rate of compliance at 51% and up till 84% and good rate of compliance as at 85% and up till 100%. No evidence was found for guiding when defining the thresholds.

Several studies used the Pirani score or/and Diméglio score for severity grading of the feet before initial treatment. Only the articles that presented both a score before and after orthotic treatment were included in the data analysis of these scores. The value before and after bracing was compared individual since it differed between the studies when the measuring was completed. Since the studies that included functional outcome used different outcome scoring systems, the result was analysed individually within the used scoring system. Range of motion were analysed in the study that presented a value before and after bracing, this value was analysed individual in relation to the orthotic intervention.

Ethical consideration

According to the World Medical Association (2018) all research including human subjects must follow the ethical principles stated in Declaration of Helsinki. Before conducting this study, a dialog with supervisor regarding ethical consideration with this thesis have been made. Appendix 5 has been filled out for an understanding of the ethical needs for conduction of this study. As this is a systematic review, no human objects are used to conduct a study, instead it is dependent on other studies and

their research with human subjects. This means that it must be taken into consideration what kind of ethical approval the included studies have. This systematic review also includes 4 retrospective studies. A retrospective study looks back in time to find their participants and is in some countries therefore said not to be need an ethical approval, since is does not count as an invasive research. 7 out of 24 articles does not state that they have ethical approval, instead an informed consent from parents is stated.

Result

Study selection

When the first search was conducted 140 studies was identified. These studies have later been through a screening process which can be followed in the figure below. In the end of this screening process there was 21 studies left. Reasons for exclusion during the screening process was wrong population, wrong intervention, or systematic reviews as study design.

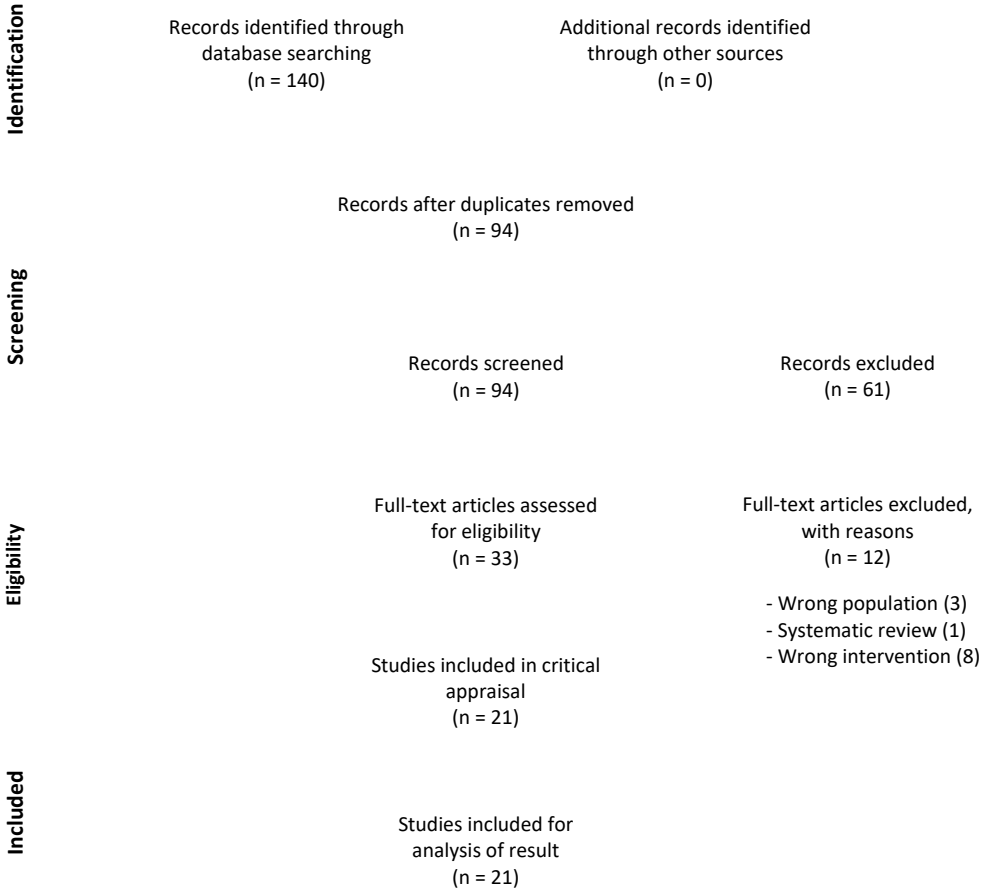


Figure 1. Flow chart of screened studies. Made from PRISMA template (Moher et al., 2009).

Study characteristics

Table 1. Table of included studies and their characteristics. "→" indicating change of orthosis. Participants include only the participants that were involved at the follow-up.

Study	Country of origin	Study design	Participants	Orthosis	Follow up time (avrage)
Abdi et al.(2017)	Tehran, Iran	Retrospective study	90	Accordion hinge DB	36 months
Berger et al.(2018)	Chiemgau, Germany	Retrospective study	45(LLO n=22, FAO n=3)	LLO, FAO	41,6 months
Bouchoucha et al. (2008)	Tunis, Tunisie	Prosepctive study	74	SFAB	
Changulani et al.(2006)	UK		66	DBB	18 months
Chong et al.(2014)	USA	Prospective randomized trial	30 (MS static n=15, MS dynamic n=15)	MS at static abduction brace, MS at dynamic abduction brace	18,7 months
Daun et al.(2018)	Malaysia	Cross-sectional study	37 (AFO n=18, DBB n=4 AFO+DBB n=3)	AFO, DBB, AFO +DBB	18,7 months
Dinesh et al.(2017)	Mangalore, India	Prospective study	25	SFAB	21 months
Emara & Diab(2015)	Cairo, Egypt		71	KAFO (custom-made)	52 months
Garg & Porter (2009)	St. Louis, USA	Case–control trial	114 (DFAB n=57, FAB n=57)	Dynamic FAB, FAB	24 months
George et al.(2011)	Liverpool, UK		27	Unilateral FAO	25 months
Janicki et al.(2011)	Toronto, Canada	Retrospective cohort study	45 (AFO n=17, DBB n=28)	Static AFO, DBB	60 months
Kuzma et al.(2020)	USA	Prognostic Retrospective Cohort Study	64	FAO	60 months
Lara et al.(2017)	Taubaté, Brazil	Retrospective Comparative Study	28 (DBB n=16, DB n=12)	DBB, DB	
Leeprakobboon et al.(2018)	Khon Kaen, Thailand	Prospective study	30	DUth brace	30 months
Manousaki et al.(2016)	Lund, Sweden	Prospectively followed cohort study	20	Dynamic KAFO (custom-made) → Dynamic AFO (custom-made)	
Ramírez et al.(2011)	San Paulo, Puerto Rico		53	DBB	48 months
Solanki et al.(2010)	India		28	modified AFO	
Sætersdal et al.(2012)	Norway	Multicenter clinical study	115 (FAB, n=64, Dynamic KAFO (Custom-made) (n=45), other brace (n=6))	FAB, Dynamic KAFO (Custom-made)	42 months
Sætersdal et al.(2017)	Norway	Multicenter clinical follow-up study	94 (FAB n=62, KAFO n=32)	FAB, Dynamic KAFO (Custom-made)	114 months
Thacker et al.(2005)	New York, USA		30	FAO	
Zionts, et al.(2012)	Los Angeles, USA	Prospective study	57	MPB	38,4 months

Result of orthotic intervention

In this study 15 different orthotic solutions could be identified which is specified in figure 2. Orthotic design and relapse rate for all studies are presented in table 6.

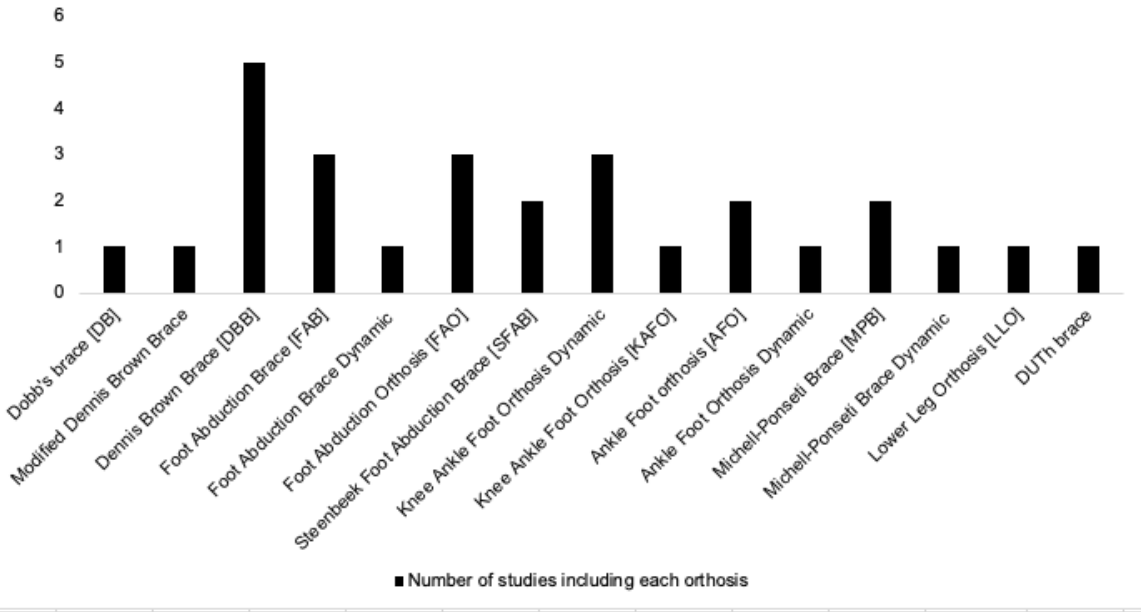


Figure 2. Diagram that demonstrates number of each orthotic design presented in the included studies. Vertical axes illustrate how many times each orthotic was used.

Dobb's brace

A traditional DB was used in one study (Lara et al., 2017) where they included 12 patients. The brace had the settings of 60°-70° external rotation on the effected side and 30°-40° external rotation on the non-effected side. They also had a dorsiflexion of 10°-15° on the effected side and the feet was placed as wide as the patients' shoulders were. The orthosis was prescribed to be used for full time during the three first months and then for approximately 12 hours each day and be worn for three to four years. The rate of recurrence was 8.33%. The authors recommend further use of the orthosis (Lara et al., 2017).

Dennis Brown Brace

DBB was used in seven studies (Reza Abdi et al., 2017; Changulani et al., 2006; Daun et al., 2018; Garg & Porter, 2009; Janicki et al., 2011; Lara et al., 2017; Ramírez et al., 2011). Garg & Porter (2009) had a historical control group using the DBB. All studies except Daun et al.(2018) and Garg & Porter (2009) stated that their settings were 70° external rotation on the effected side and 40°-45° external rotation on the non-effected side and that the feet was placed as wide as the patients shoulders. Lara et al.(2017) and Abdi et al.(2017) also included dorsiflexion of 10°-15° on both feet. The most common way was to prescribe the orthosis was for full time wear during the three first months and then during night and nap time until four years of age. Daun et al.(2018) did not report the prescribed wear time. Abdi et al.(2017) had a recurrence rate of 11% while Changulani et al.(2006) had a recurrence rate of 32%, Garg & Porter (2009) had a recurrence rate of 39%, Janicki et al.(2011) had a recurrence rate of 31%, Lara et al.(2017) had a recurrence rate of 8.33% and Ramírez et al.(2011)had a recurrence rate of 30%. Daun et al.(2018) did not report recurrence rate. All mentioned studies recommend to further use the DBB.

Foot abduction brace and Foot abduction orthosis

Three studies used a FAB (Garg & Porter, 2009; Sætersdal et al., 2017; Christian Sætersdal et al., 2012; C. Sætersdal et al., 2012). Garg & Porter (2009) had 57 patients using the FAB while Sætersdal et al.(2012) and Sætersdal et al.(2017) had 64 patients using a FAB. All studies had the orthosis prescribed to be used full time for three months and then for night and nap time until age four. Garg & Porter (2009) had a recurrence rate of 19% and Sætersdal et al.(2012) had a recurrence rate of 23.28%. Sætersdal et al.(2012) and Sætersdal et al.(2017) recommend the use of FAB.

In total four studies (Berger et al., 2018; George et al., 2011; Kuzma et al., 2020; Thacker, Scher, Sala, van Bosse, et al., 2005; Thacker, Scher, Sala, Van Bosse, et al., 2005) used FAO as an orthotic treatment.

George et al.(2011) had a unilateral FAO. Thacker et al.(2005) used a FAO with MJ Markell shoes. Berger et al.(2018), George et al.(2011), Kuzma et al.(2020) and Thacker et al.(2005) prescribed the brace to be used for full time for three months and then 10 hours each day for three to four years George et al.(2011) and Thacker et al.(2005) had the setting of 70° of abduction on the effected side and the non-effected side was put in 45° of abduction. Both feet were put into 15° of dorsiflexion. George et al.(2011) had recurrence of 31.4%, Kuzma et al.(2020) had recurrence of 40% and Thacker et al.(2005) had 57% recurrence.

SFAB was used in two studies (Bouchoucha et al., 2008; Dinesh et al., 2017). Bouchoucha et al. (2008) brace was set to 60°-70° of external rotation on the effected foot with 15° of dorsiflexion. Dinesh et al.(2017) had the settings 70° abduction on the effected side and 45° abduction on the non-effected side and both feet were also put into 15° of dorsiflexion and as wide as the patients' shoulders. Both studies prescribed the brace for full time for three months and then night and nap time, but it's not stated for how long. Bouchoucha et al. (2008)had a recurrence rate of 11% and Dinesh et al.(2017) had 5.3%. Both studies recommend the use of SFAB.

KAFO and AFO

Four studies (Emara & Diab, 2015; Manousaki et al., 2016; Sætersdal et al., 2017; C. Sætersdal et al., 2012) used a KAFO in their treatment of clubfoot. Emara & Diab(2015) had a static KAFO while (Manousaki et al., 2016) had joints in the ankle and Sætersdal et al.(2012) and Sætersdal et al.(2017) had joints in both the ankle and the knee. Emara & Diab (2015) and Sætersdal et al. (2012) prescribed the orthosis full time for three months and then until the age of four during night and nap time. Manousaki et al. (2016) prescribed the orthosis for 18 hours each day for the first two months and then 12 hours each day for eight months. Emara & Diab (2015) had a recurrence rate of 17.46% and Manousaki et al. (2016) had a recurrence rate of 0%. Sætersdal et al. (2012) and Sætersdal et al.(2017) had a recurrence rate of 23.28% and in their follow up study (Sætersdal et al., 2017) they concluded that patients using a FAB had better outcomes after 8-11 years of follow-up but better compliance with the custom-made dynamic KAFO. The authors in the mentioned studies still recommend the FAB.

AFO was used in four studies (Daun et al., 2018; Janicki et al., 2011; Manousaki et al., 2016; Solanki et al., 2010). Janicki et al. (2011) had AFO as treatment until the year 2002 when they got new guidelines to use the DBB. Manousaki et al. (2016) used a custom made AFO with ankle joints. Manousaki et al. (2016) used the AFO after the initial ten months of using KAFO and for 10 hours each day. They had no recurrence but only recommend using it if there are problems adhering to a DBB. Janicki et al. (2011) used the AFO full time until the age of five and had a recurrence of 83%. They do not recommend using an AFO. Solanki et al. (2010) used a custom-made low temperature AFO that could be remolded when the patients grow. It was used until the patients started to walk. Solanki et al. (2010) also stated that their AFO could be used when there are problems with adhering to a DBB.

Other orthotic solutions

MPB static was used for two studies (Chong et al., 2014; Zionts et al., 2012) and MPB dynamic was used in one study (Chong et al., 2014). Chong et al. (2014) had 15 patients using the MPB dynamic and 15 patients using the MPB static while Zionts et al. (2012) had 57 patients using the MPB static. Both braces had a quick release mechanism to easier be able to take on and off the bar. Zionts et al. (2012) had the settings of 60° abduction on the effected side and 30° of abduction on the non-effected side. The shoes were placed as wide as the patients shoulders. Both studies prescribed the braces to be used full time for three months and then during nighttime and naps. Chong et al. (2014) reported recurrence in 27% of the patients and Zionts et al. (2012) reported 28% recurrence.

Berger et al. (2018) used a custom-made LLO on 22 patients. The settings of the LLO were 40° external rotation to begin with but then ended up with an external rotation of 20° for best result. The braces were prescribed to be used for 24h/day the first three months and then 10h/day until the age of 3-4. Relapse rate was not reported. The authors states that LLO can be an alternative for DBB when there are problems with adherence.

DUTh brace was used by Leeprakobboon et al. (2018). DUTh brace is a custom-made orthosis made from two footplates attached with a bar. The distance between the plates should be as wide as the patients shoulders. The setting of the plates is 70° external rotation on the effected side and 40° of external rotation on the non-effected side. Both feet have 10° of dorsiflexion. The brace is prescribed to

be used 24h/day for three months and then during night and nap time. Relapse rate was not reported. The authors recommend the DUTh brace as an alternative in treatment of clubfoot.

Table 6. Table of relapse rate in each study. Green= low relapse rate, yellow=moderate relapse rate, red=high relapse rate, white=could not define.

Study	**Orthotic intervention**	**Relapse rate**	**Relapse rate graded**
Abdi et al.(2017)	Accordion hinge DBB	7,50%	
Berger et al.(2018)	LLO, FAO	did not report	
Bouchoucha et al. (2008)	SFAB	11%	
Changulani et al.(2006)	DBB	32%	
Chong et al.(2014)	MS at static abduction brace, MS at dynamic abduction brace	27%	
Daun et al.(2018)	AFO, DBB, AFO +DBB	did not report	
Dinesh et al.(2017)	SFAB	5,30%	
Emara & Diab(2015)	KAFO (custom-made)	17,46%	
Garg & Porter (2009)	Dynamic FAB	19%	
	FAB	39%	
George et al.(2011)	Unilateral FAO	31,40%	
Janicki et al.(2011)	DBB	31%	
Janicki et al.(2011)	Static AFO	83%	
Kuzma et al.(2020)	FAO	40%	
Lara et al.(2017)*	DBB	8,33%,	
	DB	5,26%	
Leeprakobboon et al.(2018)	DUth brace	did not report	
Manousaki et al.(2016)	Dynamic KAFO (custom-made) →Dynamic AFO (custom-made)	0%	
Ramírez et al.(2011)	DBB	30%	
Solanki et al.(2010)	modified AFO	did not report	
Sætersdal et al.(2012)	FAB, Dynamic KAFO (Custom-made)	23,28%	
Sætersdal et al.(2017)	FAB, Dynamic KAFO (Custom-made)	did not report	
Thacker et al.(2005)	FAO	57%	
Zionts, et al.(2012)	MPB	28%	

Result of presented outcomes

The present outcomes that could be identified in the included studies is Pirani score, Dimeglio score, compliance with brace, range of motion, recurrence of deformity, functional outcome, gait analysis and reported sin problems. Each individual study and its presented outcome are summarized in table 5.

Table 5. Table of the outcomes included in each study.

Study	Orthotic intervention	Outcome					
		Compliance with brace	Recurrence of deformity	Pirani score	Diméglio score	Range of motion	Functional outcome
Abdi et al.(2017)	Accordion hinge DBB	X	X				
Berger et al.(2018)	LLO, FAO	X					
Bouchoucha et al. (2008)	SFAB	X	X	X	X		X
Changulani et al.(2006)	DBB	X	X				
Chong et al.(2014)	MS at static abduction brace , MS at dynamic abduction brace	X	X				
Daun et al.(2018)	AFO, DBB, AFO +DBB	X		X			
Dinesh et al.(2017)	SFAB	X	X	X			
Emara & Diab(2015)	KAFO (custom-made)	X	X				
Garg & Porter (2009)	Dynamic FAB, FAB	X	X				
George et al.(2011)	Unilateral FAO	X	X				
Janicki et al.(2011)	Static AFO, DBB	X	X				
Kuzma et al.(2020)	FAO	X	X				
Lara et al.(2017)	DBB, DB		X	X			
Leeprakobboon et al.(2018)	DUth brace	X				X	
Manousaki et al.(2016)	Dynamic KAFO (custom-made) →Dynamic AFO (custom-made)	X	X				
Ramírez et al.(2011)	DBB	X	X				
Solanki et al.(2010)	modified AFO	X		X	X		
Sætersdal et al.(2012)	FAB, Dynamic KAFO (Custom-made)	X		X			
Sætersdal et al.(2017)	FAB, Dynamic KAFO (Custom-made)	X	X				X
Thacker et al.(2005)	FAO	X	X	X	X		
Zionts, et al.(2012)	MPB	X	X				

Result of compliance

Table 7 presents those studies that have reported compliance, that includes how it has been measured, the result for each intervention, the measured rate and its grade of compliance. Four studies gave examples on how to improve compliance. Changulani et al., (2006) and Chong et al. (2014) stated that education for the parents including how the brace works and the risks with not using it would improve compliance. Daun et al. (2018) stated that frequent follow-up until the compliance is high decreases the risk for low compliance later during the treatment. Berger et al. (2018) tried to implement a bed night routine in the patients with non-compliance and this showed positive results. Abdi et al. (2017) and Emara & Diab (2015) both stated that the ability to move the legs independent lowers the risk of non-compliance.

Six studies (Chong et al., 2014; Dinesh et al., 2017; Garg & Porter, 2009; George et al., 2011; Manousaki et al., 2016; Ramírez et al., 2011) could find a relationship between non-compliance and recurrence while two studies (Sætersdal et al., 2012; Kuzma et al., 2020) stated that there was no relationship between non-compliance and recurrence. 2 participants using the dynamic FAB and 11 participants using the FAB developed skin complication such as skin ulceration and blistering (Garg & Porter, 2009).

Ten studies (Bouchoucha et al., 2008; Dinesh et al., 2017; Chong et al., 2014; Daun et al., 2018; Garg & Porter, 2009; George et al., 2011; Janicki et al., 2011; Kuzma et al., 2020; Ramírez et al., 2011; Solanki et al., 2010) did not define compliance. But Bouchoucha et al. stated that 4 patients using the SFAB irregularly developed heel ulceration (Bouchoucha et al., 2008)

Nine studies (Abdi et al., 2017; Berger et al., 2018; Emara & Diab, 2015; Leeprakobboon et al., 2018; Manousaki et al., 2016; Sætersdal et al., 2012; Sætersdal et al., 2017; Thacker et al., 2005; Zionts et al., 2012) defined good compliance as when the patients adhered strictly to the included protocol. One study (Changulani et al., 2006) defined good compliance as wear time more than 10h/day. The participant using a FAO in Berger et al. (2018) that reported skin-problem were categorized as non-compliant with the intervention. The one patient that got affected by a serious cutaneous problem changed to a resin modified custom-made shoe on the FAO instead. None of the participants using an LLO reported any skin problems (Berger et al., 2018). One patient using the DUth brace got affected by skin breakdown (Leeprakobboon et al., 2018). 8 patients using the MPB developed skin problems in relation to the brace (Zionts et al., 2012).

Table 7. Table of how studies have reported compliance, which braces that were used and its rate of compliance. Green= good compliance, yellow=moderate compliance, red=poor compliance, white= could not be defined.

Study	Parent information	Wear and tear of orthosis	Skin breakdown	Medical record	Pressure sensor	Does not state	Orthotic intervention	Rate of compliance	Compliance grade
Abdi et al.(2017)	X		X				Accordion hinge DBB	89,7%	
Berger et al.(2018)	X	X					LLO, FAO	46%	
Bouchoucha et al. (2008)			X			X	SFAB	94,59%	
Changulani et al.(2006)	X						DBB	68%	
Chong et al.(2014)	X						MS at static abduction brace, MS at dynamic abduction brace	-	
Daun et al.(2018)						X	AFO, DBB, AFO +DBB	-	
Dinesh et al.(2017)	X						SFAB	94,7%	
Emara & Diab(2015)						X	KAFO (custom-made)	100%	
Garg & Porter (2009)	X		X				Dynamic FAB,	81%	
Garg & Porter (2009)	X		X				FAB	47%	
George et al.(2011)	X						Unilateral FAO	77,78%	
Janicki et al.(2011)						X	Static AFO, DBB	-	
Kuzma et al.(2020)	X						FAO	95%	
Kuzma et al.(2020)					X		FAO	77,1%	
Leeprakobboon et al.(2018)	X						DUth brace	-	
Manousaki et al.(2016)	X						Dynamic KAFO (custom-made) →Dynamic AFO (custom-made)	90%	
Ramírez et al.(2011)	X						DBB	53%	
Solanki et al.(2010)	X						modified AFO	-	
Sætersdal et al.(2012)				X			FAB, Dynamic KAFO (Custom-made)	61%	
Sætersdal et al.(2017)				X			FAB, Dynamic KAFO (Custom-made)	61%	
Thacker et al.(2005)	X						FAO	70%	
Zionts, et al.(2012)	X		X				MPB	-	

Result of Pirani score

The 7 articles seen in table 8, presented Pirani score as an outcome, measured before and after bracing. In general, all the present articles demonstrated a decrease between the first and last measured Pirani score. Solanki et al. (2010) did not present a measured mean score before casting. Daun et al. (2018) and Sætersdal et al. (2012) did not present a mean score measured between the end of the serial casting and initial bracing. Dinesh et al. (2017) presented the mean score after bracing in two categories, the patient that experienced compliance with the brace and the patients that not experienced compliance with the brace. Compared with the mean score after casting, the compliant group demonstrated decrease in mean score of total 0,625. However, the score in the non-compliant group demonstrated an increase of 0,75.

The patient that used an FAO (Thacker et al., 2005) and DUth brace (Leeprakobboon et al., 2018) demonstrated a slightly score increase after bracing compared with the measured score after serial casting.

Table 8. This table present each study including Pirani score, with mean score before and after bracing. "-" Indicates that data is not stated. n= number of participants

Study	**Orthotic intervention**	**Mean score before casting**	**Means score after casting**	**Mean score after bracing**
Bouchoucha et al. (2008)	SFAB	3,4	0,45	0,3
Daun et al.(2018)	AFO, DBB, AFO+DBB	2,18	-	0,19
Dinesh et al.(2017)	SFAB	3,79	0,65	Compliant group(n=36): 0,025 Non-compliant group (n=2): 1,4
Leeprakobboon et al.(2018)	DUth brace	5,6	0,07	0,1
Solanki et al.(2010)	Modified AFO	-	1,095	0,175
Sætersdal et al.(2012)	FAB	4,8	-	0,4
Sætersdal et al.(2012)	Dynamic KAFO (Custom-made)	4,8	-	0,6
Thacker et al.(2005)	FAO	5,25	0,65	0,75

Result of Diméglio score

The present articles in table 9 included Diméglio score as an outcome measure, measured before and after bracing. How the authors of the studies have presented the result based on the participants differs. Bouchoucha et al. (2008) demonstrated a total mean score for all 66 participants. Solanki et al. (2010) presented the result for each individual affected foot. Thacker et al. (2005) divided the result into two groups, non-compliant and compliant.

Table 9. This table presents Diméglio score before casting and after bracing. n=number of participants, "-" = data is not stated.

Study	**Group definition**	**Orthotic intervention**	**Mean score before casting**	**Mean score after casting**	**Mean score after brace**
Bouchoucha et al. (2008)	n=66	SFAB	12,9	1,3	0,7
Solanki et al.(2010)	Right side, number of feet = 22 Left side, number of feet=21	modified AFO	-	Right side=1,09 Left side =1,1	Right side =0,11 Left side =0,24
Thacker et al.(2005)	Compliant group, n =21 Non-compliant,n=9	FAO	Compliant group = 14,5 Non-compliant group=16,0	Compliant group=3.,5 Non-compliant group=4,0	Compliant group=1,0 Non-compliant group=6,0

Result of functional outcome

Bouchoucha et al. (2008) measured functional score with the functional classification of the Hospital for Joint Diseases. The mean score before casting were 13 and after full-time bracing 55,4. The follow-ups that were done in an average of 6,4 months after finished bracing, demonstrated a mean score of 54.

Sætersdal et al. (2017) used and evaluated two different clubfoot questionaries, Laaveg and Ponseti´s functional rating system and Roye´s Disease-specific Instrument. Both questionaries regarded the patient's satisfaction, pain and function and was answered from the parents. The score in total could be divided between 0-100 in both questionaries. The patients that were using the bilateral FAB demonstrated higher functional outcome based on the used outcome measurement-tools.

Table 11. Presentation of Bouchoucha et al. (2008) functional outcome.

Study	Orthotic intervention	Measurement tool	Mean score before casting	Mean score after full-time bracing	Mean score after follow-up (mean 6,4 months)
Bouchoucha et al. (2008)	SFAB	The functional classification of the Hospital for Joint Diseases	13	55,4	54

Table 12. Presentation of Sætersdal et al. (2017) functional outcome.

Study	Measurement tool	Orthotic intervention	Score
Sætersdal et al.(2017)	Laaveg and Ponseti´s functional rating system	FAB	87
Sætersdal et al.(2017)	Laaveg and Ponseti´s functional rating system	Dynamic KAFO	78
Sætersdal et al.(2017)	Roye´s Disease-specific Instrument	FAB	82
Sætersdal et al.(2017)	Roye´s Disease-specific Instrument	Dynamic KAFO	74

Result of range of motion

Leeprakobboon et al. (2018) included range of motion as an outcome, it included passive motions in the ankle dorsiflexion, forefoot abduction, derotation of calcaneopedal block and heel valgus. The abduction was measured in relation to the hindfoot and calcaneopedal block was measured in the horizontal plane.

Table 10. Presentation of Leeprakobboon et al.(2018) range of motion.

	Dorsiflexion	Forefoot abduction	Derotation of calcaneopadal block	Heel valgus
Initial application	13,08°	46,38°	28,93°	22,02°
12-months follow up	10,95°	47,23°	28,29°	19,25°

Result quality assessment

Bias is divided into two different sections, randomized studies, and not randomized studies. The randomized section only included one study. Conflict of interest was assessed as low if both questions was answered yes, moderate if either of the questions was answered no and high if both questions was answered no, see questions in appendix 2.

Chong et al. (2014) was the only randomized study included. The study was assessed as a randomized trial (effect of being assigned an intervention (ITT)) and the bias was assessed from the template that can be found in appendix 2. The study showed low risk of bias in both randomization and missing data but had moderate risk of bias based on intended intervention, in measuring outcome and in conflict of interest. Report bias is high because a study report or registration could not be found.

Table 2. Table of bias in randomized studies. Green= low risk of bias, yellow=moderate risk of bias and red=high risk of bias.

Study	Bias randomization	Bias intended intervention	Bias missing data	Bias measuring outcome	Bias reported result	Conflict of interest
Chong et al.(2014)						

There where 20 studies that was not randomized. These studies where assessed as either not randomized assigned interventions or as non-randomized studies of interventions (effect of being assigned an intervention (ITT)) or non-randomized studies of interventions (effect of completing an intervention (per protocol)). The templates used for assessing bias can be found in appendix 2. Confounders that could be identify in this review is different parental factors such as education, economics, and motivation to fulfill the treatment, social factors, esthetic factors with wearing a brace and the education and learning curve of the medical organization.

3 studies showed low bias of confounders. These studies presented a table of factors that could have affected the study. 11 studies showed moderate risk of bias. These studies presented possible confounders but did not discuss them or only had one. 6 studies showed high risk of confounding bias. These studies did not mention any confounders. 19 out of 20 studies shows low risk of bias based on selection of participants. Lara et al. (2017) is the only study that indicated high risk of bias regarding selection of participants. The reason for this is because they chose to exclude any participant that was not showing brace-compliance from the beginning and therefore might end up affecting the compliance-result. 11 out of 20 studies showed low risk of bias in classification of intervention. 8 studies had moderate risk of bias. Sætersdal et al. (2017) indicate high risk of bias in classification of intervention because it is a follow-up study based on their previous result and therefore part of the results already was known when conducting the study. 20 studies showed low risk of bias based on intended intervention. The bias based on intended intervention is hard to evaluate when it comes to retrospective studies since there is no way to know if the intervention was intended or changed. 18 studies showed low risk of bias based on missing data. Lara et al. (2017) and Solanki et al. (2010) showed high risk of bias based on missing data because they did not report the number of dropouts, or they had a high rate of dropouts. 18 out of 20 studies showed moderate risk of bias in measuring outcome. Two studies showed high risk of bias. The reason for the bias of measuring outcome is moderate or high in all studies might be because the intervention is something physical that depends on the design and can therefore not be changes in a way that blinds the intervention for the person assessing it. Reporting bias was observed for by inspection if the studies had a study report published before conducted the study or a study registry. This was done at ClinicalTrails.gov, International Clinical Trails Registry Platform and EudraCT. There could not be discovered a study report for any studies and therefore all studies are graded with high risk of report bias. Looking at conflict of interest two studies showed low risk of bias, 10 studies showed moderate risk of bias and 8 studies showed high risk of bias. This is summed up in table 3.

Table 3. Table presenting each studies and its assessed risk of bias. Green= low risk of bias, yellow=moderate risk of bias and red=high risk of bias.

Study	Bias confounding	Bias selection of participants	Bias classification of intervention	Bias intended intervention	Bias missing data	Bias measuring outcome	Bias reported result	Conflict of interest
Abdi et al.(2017)								
Berger et al.(2018)								
Bouchoucha et al. (2008)								
Changulani et al.(2006)								
Daun et al.(2018)								
Dinesh et al.(2017)								
Emara & Diab(2015)								
Garg & Porter (2009)								
George et al.(2011)								
Janicki et al.(2011)								
Kuzma et al.(2020)								
Lara et al.(2017)								
Leeprakobboon et al.(2018)								
Manousaki et al.(2016)								
Ramírez et al.(2011)								
Solanki et al.(2010)								
Sætersdal et al.(2012)								
Sætersdal et al.(2017)								
Thacker et al.(2005)								
Zionts, et al.(2012)								

3 studies showed clear external validity. 15 studies showed probable external validity and 3 studies showed uncertain external validity. 2 studies showed excellent internal validity, 8 studies showed great internal validity, 7 studies showed acceptable internal validity and 4 studies showed uncertain external validity. The result of the precision is assessed as probably underpowered study in 20 of 21 studies. This could be an effect of many studies being retrospective and therefore it is hard to calculate the study power. This is summed up in table 4.

Table 4. Table of internal validity, external validity and precision based on appendix 3. Dark Green= Clear external validity/Excellent internal validity. Green= Probable external validity/ Good internal validity/ Premeditated and sufficient study size. Yellow= Uncertain external validity/ Acceptable internal validity/ Sample size of uncertain adequacy. Red= External validity cannot be assessed/ Uncertain internal validity/ Probably underpowered study.

Study	**External validity**	**Internal validity**	**Precision**
Abdi et al.(2017)			
Berger et al.(2018)			
Bouchoucha et al. (2008)			
Changulani et al.(2006)			
Chong et al.(2014)			
Daun et al.(2018)			
Dinesh et al.(2017)			
Emara & Diab (2015)			
Garg & Porter (2009)			
George et al.(2011)			
Janicki et al.(2011)			
Kuzma et al.(2020)			
Lara et al.(2017)			
Leeprakobboon et al.(2018)			
Manousaki et al.(2016)			
Ramírez et al.(2011)			
Solanki et al.(2010)			
Sætersdal et al.(2012)			
Sætersdal et al.(2017)			
Thacker et al.(2005)			
Zionts, et al.(2012)			

Discussion

Effects of orthotic interventions

The aim of this thesis was to investigate how different orthotic interventions for pediatric patients with idiopathic congenital talipes equinovarus differ in outcomes, including maintenance of foot alignment, foot mobility as well as compliance-rate. The results demonstrated that non-compliance had an effect on recurrence of deformity. Generally, the evidence indicated a high use of several different orthotic interventions evaluated with different outcome measures. The use of different outcome measures makes it hard to draw a conclusion of which orthosis that is most suitable to prescribe in clubfoot treatment. Our initial idea was to investigate only Pirani score as an outcome of clubfoot treatment, but as the pilot search detected limited number of studies, this idea was remade for an inclusion of wider ranges of outcomes.

Based on the 21 included studies in this systematic review we can conclude that different orthotic interventions provided different in outcomes in treatment of idiopathic clubfoot. Compliance was reported in 20 of the included studies and thereby the outcome measure that were reported in the greatest number of times. 6 of the included studies did not demonstrate a concrete value of the compliance rate (Chong et al., 2014; Daun et al., 2018; Janicki et al., 2020; Leeprakobboon et al., 2018; Solanki et al., 2010; Zionts et al., 2012). The results of these studies were thereby difficult to include in the result analysis. The definition of poor-, moderate and good compliance as well as, low-, moderate and high rate of relapse has all weak scientific basis and has no clinical meaning. These definitions were created solely for support of the result analysis in this thesis. With that said do not use these definitions as guidelines to evaluate if an orthotic intervention is determinate as useful or not.

But analyzed according to our definition Berger et al. (2018) demonstrated poor compliance with the FAO and LLO brace. This was the only study that included wear and tear of the orthosis as measure of compliance in combination with parent information. Moderate compliance was presented in 9 studies (Changulani et al., 2006; Garg & Porter, 2009; George et al., 2011; Kuzma et al.; 2020, Ramirez et al., 2011; Sætersdal et al., 2012; Sætersdal et al., 2017; Thacker et al., 2005). Where George et al. (2011) differentiae in its orthotic design since this is the only unilateral brace of all the other braces that demonstrated moderate compliance. In terms of the method used for measure compliance, 8 of those 9 studies used caregiver interviews as at least one of their methods. Kuzma et al. (2020) used pressure sensors in combination with parent interviews, where the interviews demonstrated good compliance and the pressure sensor moderate compliance. Which might indicate that the information given by caregivers may be distorted from the truth. But as Kuzma et al. (2020) was the only article that measured compliance with help of pressure sensors it is difficult to draw a conclusion, thereby future research might investigate this additionally. Good compliance was demonstrated in 6 studies (Abdi et al., 2017; Bouchoucha et al., 2008; Dinesh et al., 2017; Emara & Diab, 2015; Kuzma et al., 2020; Manousaki et al., 2016). 2 of these studies presented SFAB as an orthotic intervention (Bouchoucha et al., 2008; Dinesh et al., 2017). Abdi et al. (2017), Emara & Diab (2015) and Manousaki et al. (2016) used a modified version of a KAFO and DDB. These modified orthoses all have been developed to increase the rate of compliance. In terms of function Emara & Diab (2015) prescribed a static orthosis compared with the rest that included dynamicity. Abdi et al. (2017) and Emara & Diab (2015) described that independent movement of each leg increases compliance. This can therefore indicate that it is not the orthosis dynamicity in itself that increases compliance it might instead be the fact that independent movement of the legs is available. None of the studies that reported compliance as an outcome presented it during different time frames. An interesting aspect to investigate further would be if brace compliance change during treatment time and in different patient ages. Ideas of how to increase brace compliance has been discussed in a few studies. Aspects that have been brought up is education for parents and sleep routines. For future research, these presented ideas should be investigated further for a potentially translation into clinical practice.

Dinesh et al. (2017) and Manousaki et al. (2016) stated that they confirmed an association between compliance and relapse. High rate of relapse was seen in 2 studies, Janicki et al. (2011) with the static AFO and Thacker et al. (2005) with FAO. The static AFO had by far the highest rate of recurrence of misalignment. This is as well the only included study that use a static AFO as orthotic intervention. Of all the studies that demonstrated low relapse rate three studies (Abdi et al., 2017; Dinesh et al., 2017; Manousaki et al., 2016) also presented good compliance. Abdi et al. (2017), Dinesh et al. (2017) and Lara et al. (2017) all used a bilateral orthotic intervention while Manokaski et al. (2016) used a unilateral. This indicated a tendency of the bilateral braces demonstrating a lower relapse rate. Eleven studies showed moderate relapse rate (Bouchoucha et al., 2008; Changulani et al., 2006; Chong et al.,

2014; Emara & Diab, 2015; Garg & Porter, 2009; George et al., 2011; Janicki et al., 2011: Kuzma et al., 2020; Ramírez et al., 2011: Sætersdal et al., 2012; Zionts et al., 2012). None of these studies demonstrated poor rate of compliance. On the other hand, a few of these studies did not present a value of compliance and could therefore not be included in the analysis for any potentially correlations. In terms of the presented Diméglio score, Pirani score, functional score and range of motion it is difficult to draw a conclusion since these outcomes are only presented in a limited number of studies.

Another outcome that could not be included in the analysis was the individual mean follow-up time. The mean follow-up time differed between 18 months to 114 months in the included studies. There was a number of studies that did not demonstrated the average follow-up time, as well as none of the studies presented a clear description of their definition of follow-up time. This data was thereby difficult to interpret and not possible to include in the analyze. An interest aspect to include as well in the analysis had been the mean follow-up time and possible correlations between it and potential recurrence of deformity. This is an important factor to include for development in future research within the area. As previously mentioned by Dobbs et al. (2004), the risk of recurrency with deformity is higher under the bracing time which shows that there might appeared a correlation in the included studies. By investigating the available evidence in clubfoot treatment, a strong indication for better outcomes is by using a brace protocol, including what orthotic intervention to apply in practice as well as directives on which time period to use it. This is hard to develop with the evidence that is available right now since the currently protocol differs a lot, but instead this opens for possibilities in future development. We have great understanding for the different socioeconomic factors in different countries around the world, this is a factor that needs to be remembered in a potential implementation of protocol since this can lead to unequal treatments in different social economic groups.

Translation of evidence into clinical practice

7 studies included Pirani score measured before and after bracing. We can clearly see that there is inconsistency in reporting Pirani score, it differs both when and how they have reported the score.
There was one study (Chong et al., 2014) that decided to not report Pirani score because the result could not be seen as reliable according to the authors. Even though Pirani score is a validated scoring system there is a lack on how to report it and how to use in in the medical field which creates results that are not trustworthy. Also, as previous mention Pirani score is stated to be a validated tool for surgeons. We have not found a source that demonstrate which profession that adapts and applies the orthoses around the world, nor what professions that usually are involved in the clubfoot treatment. But to the best of our knowledge, it is not only the surgeons that are involved. So, for a practical setting this outcome score probable is not the most appropriate right now. Dinesh et al. (2017) was the only study that reported Pirani score separated between a compliant group and non-compliant group. This showed a great difference between the two groups. It would have been interesting if more studies had reported this way because this would have given us a chance to include this further in the analysis.

The person that is responsible for fitting and evaluation of the orthosis probably gets direct feedback from parents and patients on the devices advantageous and disadvantages. So, standardize outcome measures would probably benefit for a systematically evaluation of the orthoses and its result in a clinical context. As mentioned, there are many indications that non-compliance with the different braces is appearing in the clinical practice, both demonstrated in this review and in previous research. But a common method for assessing this is still lacking. As mentioned under non-compliance, ulceration related to the orthose is appearing for some children in different orthotic designs. Prevention of ulceration is an important aspect to include in the future development of new interventions, since this might impact compliance with brace and recurrence of deformity. Even though the research strengthens the idea of developing one orthosis that is universal for treating clubfoot it is important that different options and solutions are available, and that the orthotist is inventive when complications appear. For example, it can be seen in the result that a KAFO can be done in different ways with different joints. A combination of different orthotic interventions can help when having struggles in brace adherence or ulcerations. The esthetic aspects of the orthotic designs are not mentioned in any of the included studies, this can be an important aspect to include in further studies. Even if the treatment of the children is starting in an early stage of life, these children are growing and developing an own opinion of the orthosis appearance. Also, what the caregivers thinks about the orthotic appearance can have an impact on compliance. In treating a diagnosis like clubfoot, it is important that all persons involved in the treatments process is being heard and especially the caregivers that spend the most time with the patients. Giving the parents a chance to be involved in the process of casting, choosing patterns for the orthosis can potentially give them an increased sense of participations.

There is a big research concern when it comes to evaluating this field of orthotic interventions. The population of patients with congenital idiopathic clubfoot is rather small, which probably is creating difficulties when requesting for special study design, such as a randomized controlled trail. In the included studies in this thesis, 4 articles were stated as retrospectives. Retrospective studies create research issues in terms of selection bias, classification bias, missing data bias and measuring outcome bias. Which is the reason for retrospective studies to generally be considered as a having a lower rate of evidence. The only included randomized controlled trail is the only study that presented study power. In table 3 it can clearly be seen that almost all studies indicated moderate risk of bias in measuring outcome. The studies that indicated high risk of bias in measuring outcome did not measure both groups in the same way. To be able to compare an intervention it is important to make sure that the groups are the same and handled the same way which was the lack in these studies. Overall few studies presented confounding factors. The studies that are green is the ones that presented a table with several confounding factors. This is a great way to present what could have affected the study and something we think more studies should take use of to eliminate confounding bias.

It is notable that research is taking place and the development in clubfoot treatment is increased the last decades. Hopefully in the future the deformity is corrected by using a compliant method and as well minimizing the risk of relapses.

Limitations

A general limitation in the presented outcome measure within clubfoot treatment is the lack of standardization. That it currently differs so much in what is evaluated and how the evaluation is done after clubfoot treatment is causing problems when comparisons of different result are to be made. In this thesis different outcomes such as Dimeglio score, range of motion and functional outcomes score are only presented in a limited number of studies and thereby not as comparable as for example compliance. Not included outcomes that had been interesting to investigate are for example gait analysis. Gait analysis was presented in Manousaki et al. (2016) but was excluded in this systematic review since that study had a comparison group which did not use an orthosis and was not diagnosed with congenital clubfoot. By using gait analysis as an outcome measure after bracing, valuable and interesting factors could be analyzed. Daun et al. (2018) used three different interventions but did not state an induvial result for the used orthoses. It had been interesting to analyze the induvial designs instead as a group.

When conducting the quality assessment, we made the decision to use the SBU tool instead of Cochrane because of the language barriers. We experience problems with interpretation of the Cochrane bias assessment tool and therefore struggled with the implementation of it. This decision was established since our first language is Swedish and therefore were thought to be providing a more trustworthy result using a Swedish template. The language could also have been a risk factor in the data extraction process, even if both authors understand English ell, potentially misinterpretations can still have appeared. Since the terminology within orthotic interventions differ hugely it is possible that other ways of naming an orthosis exist. The final search string was developed with the authors commonly known and seen terminology within this area.

To be able to conduct an analysis our own definitions of good, moderate and poor/low was done. This threshold we have done ourselves with only one limited source of information and could therefore be seen as a result bias. This might have influenced how a study result was analyzed. An example of this is Bouchoucha et al. (2008), that had a relapse rate of 11%, the limit for good relapse rate was set at 10%. This made the result form this study being categorized as moderate which could have changed the result.

Conclusions

In conclusion the studies presented in this systematic review do not provide enough information to make a judgement on which orthosis to recommend. The great variety of outcome measures create difficulties for a comparison between orthotic interventions. In a clinical context the previous and current evidence primarily seems to focus on the outcomes in terms of maintenance of the foot alignment. This thesis demonstrated non-compliance with the brace as the primary reported outcome measure. Future research must continue to develop orthotic interventions with focus on brace adherence for the patients.

Appendix list

1. Search process
2. Critical appraisal tool
    a. Bedömning av icke-randomiserade studier av interventioner (effekt av att tilldelas en intervention (ITT))
    b. Bedömning av icke-randomiserade studier av interventioner (effekt av att fullfölja en intervention (per protokoll))
    c. Bedömning av randomiserade studier (effekt av att tilldelas en intervention (ITT))
3. Validity control tool
4. Form for Self-Assessment of Ethical Issues in Degree Projects1 at the School of Health and Welfare

References

Abdi, R., Hajzargarbashi, R., Baghdadi, T., Shirazi, M. R., Omidnia, B., & Tajik, H. (2017). Effect of a modified dynamic accordion hinge Denis Browne brace on the success rate of the Ponseti method in idiopathic clubfoot: a preliminary study. *Journal of Pediatric Orthopaedics: Part B*, *26*(2), 112-115. https://doi.org/10.1097/BPB.0000000000000357

Berger, N., Lewens, D., Salzmann, M., Hapfelmeier, A., Döderlein, L., & Prodinger, P. M. (2018). Is unilateral lower leg orthosis with a circular foot unit in the treatment of idiopathic clubfeet a reasonable bracing alternative in the Ponseti method? Five-year results of a supraregional paediatric-orthopaedic centre. *BMC Musculoskelet Disord*, *19*(1), 229. https://doi.org/10.1186/s12891-018-2160-1

Bouchoucha, S., Smida, M., Saïed, W., Safi, H., Ammar, C., Nessib, M. N., & Ghachem, M. B. (2008). Early results of the Ponseti method using the Steenbek foot abduction brace: a prospective study of 95 feet. *J Pediatr Orthop B*, *17*(3), 134-138. https://doi.org/10.1097/BPB.0b013e3282fa5f0d

Changulani, M., Garg, N. K., Rajagopal, T. S., Bass, A., Nayagam, S. N., Sampath, J., & Bruce, C. E. (2006). Treatment of idiopathic club foot using the Ponseti method. Initial experience. *J Bone Joint Surg Br*, *88*(10), 1385-1387. https://doi.org/10.1302/0301-620x.88b10.17578

Chong, D. Y., Finberg, N. S., Conklin, M. J., Doyle, J. S., Khoury, J. G., & Gilbert, S. R. (2014). Prospective evaluation of the use of Mitchell shoes and dynamic abduction brace for idiopathic clubfeet. *J Pediatr Orthop B*, *23*(6), 501-504. https://doi.org/10.1097/bpb.0000000000000090

Dinesh, T. S., Kotian, P., Sujir, P., Joe, V., & Rajendra, A. (2017). Steenbeek foot abduction brace for clubfoot: Cost-effective but is it effective? A prospective study [Article]. *Asian Journal of Pharmaceutical and Clinical Research*, *10*(5), 99-102. https://doi.org/10.22159/ajpcr.2017.v10i5.16296

Dyer, P. J., & Davis, N. (2006). The role of the Pirani scoring system in the management of club foot by the Ponseti method. *J Bone Joint Surg Br*, *88*(8), 1082-1084. https://doi.org/10.1302/0301-620X.88B8.17482

Edinger, J., Sinha, A., & Fisher, M. (2019). Congenital and Acquired Disorders. In J. Webster & D. Murphy (Eds.), *Atlas of Orthoses and Assistive Devices* (Fifth Edition ed., pp. 303-312). Elsevier.

Emara, K. M., & Diab, R. A. (2015). A knee-ankle-foot orthosis in treatment of talipes equino varus (TEV) with better patient compliance [Article]. *European Orthopaedics and Traumatology*, *6*(4), 337-341. https://doi.org/10.1007/s12570-015-0323-1

Ganesan, B., Luximon, A., Al-Jumaily, A., Balasankar, S. K., & Naik, G. R. (2017). Ponseti method in the management of clubfoot under 2 years of age: A systematic review [Article]. *PLoS ONE*, *12*(6), Article e0178299. https://doi.org/10.1371/journal.pone.0178299

Garg, S., & Porter, K. (2009). Improved bracing compliance in children with clubfeet using a dynamic orthosis [Article]. *Journal of Children's Orthopaedics*, *3*(4), 271-276. https://doi.org/10.1007/s11832-009-0182-9

Gelfer, Y., Hughes, K. P., Fontalis, A., Wientroub, S., & Eastwood, D. M. (2020). A systematic review of reported outcomes following Ponseti correction of idiopathic club foot. *Bone & Joint Open*, *1*(8), 457-464. https://doi.org/10.1302/2633-1462.18.BJO-2020-0109.R1

Gelfer, Y., Wientroub, S., Hughes, K., Fontalis, A., & Eastwood, D. M. (2019). Congenital talipes equinovarus: a systematic review of relapse as a primary outcome of the Ponseti method. *Bone Joint J*, *101-B*(6), 639-645. https://doi.org/10.1302/0301-620X.101B6.BJJ-2018-1421.R1

George, H. L., Unnikrishnan, P. N., Garg, N. K., Sampath, J., & Bruce, C. E. (2011). Unilateral foot abduction orthosis: is it a substitute for Denis Browne boots following Ponseti technique? *J Pediatr Orthop B*, *20*(1), 22-25. https://doi.org/10.1097/BPB.0b013e32833fb8a5

Haft, G., Walker, C., & Crawford, H. (2007). Early clubfoot recurrence after use of the Ponseti method in a New Zealand population. In (Vol. 89(3), pp. 487-493). J Bone Joint Surg Am.

ISO. (2007). INTERNATIONAL STANDARD. In *Prosthetics and orthotics — Categorization and description of external orthoses and orthotic components* (Vol. ISO13404:2007).

ISO. (2020). INTERNATIONAL STANDARD. In *Prosthetics and orthotics —Vocabulary —Part 1:General terms for external limb prostheses and external orthoses* (Vol. ISO 8549-1:2020).

Jain, S., Ajmera, A., Solanki, M., & Verma, A. (2017). Interobserver variability in Pirani clubfoot severity scoring system between the orthopedic surgeons. *Indian J Orthop*, *51*(1), 81-85. https://doi.org/10.4103/0019-5413.197551

Janicki, J. A., Wright, J. G., Weir, S., & Narayanan, U. G. (2011). A comparison of ankle foot orthoses with foot abduction orthoses to prevent recurrence following correction of idiopathic clubfoot by the Ponseti method. *Journal of bone and joint surgery. British volume*, *93*(5), 700-704. https://doi.org/10.1302/0301-620X.93B5.24883

John Hopkins University. (2021). *Clubfoot*. John Hopkins Medicine. https://www.hopkinsmedicine.org/health/conditions-and-diseases/clubfoot

Jowett, C. R., Morcuende, J. A., & Ramachandran, M. (2011). Management of congenital talipes equinovarus using the Ponseti method: a systematic review. *J Bone Joint Surg Br*, *93*(9), 1160-1164. https://doi.org/10.1302/0301-620X.93B9.26947

Kuzma, A. L., Talwalkar, V. R., Muchow, R. D., Iwinski, H. J., Milbrandt, T. A., Jacobs, C. A., & Walker, J. L. (2020). Brace Yourselves: Outcomes of Ponseti Casting and Foot Abduction Orthosis Bracing in Idiopathic Congenital Talipes Equinovarus. *J Pediatr Orthop*, *40*(1), e25-e29. https://doi.org/10.1097/bpo.0000000000001380

Lara, L. C. R., Gil, B. L., Filho, L. C. A. T., & dos Santos, T. P. S. (2017). Comparison between two types of abduction orthotics in treating congenital clubfoot. *Acta Ortopedica Brasileira*, *25*(4), 125-128.

Leeprakobboon, D., Singjam, U., & Sangkomkamhang, T. (2018). The duth abduction brace: a new orthosis design for maintaining foot correction in idiopathic clubfoot [Article]. *Journal of the Medical Association of Thailand*, *101*(9), 1191-1196.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, *339*(2009), b2700-b2700. https://doi.org/10.1136/bmj.b2700

Manousaki, E., Czuba, T., Hägglund, G., Mattsson, L., & Andriesse, H. (2016). Evaluation of gait, relapse and compliance in clubfoot treatment with custom-made orthoses. *Gait & Posture*, *50*, 8-13. https://doi.org/10.1016/j.gaitpost.2016.08.005

Moher , D., Liberati , A., Tetzlaff , J., & Altman , D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. In: The PRISMA Group.

Monash University Library. (2021). *Databses by Subject: Health sciences*. Retrieved 14 April 2021 from https://guides.lib.monash.edu/subject-databases/health-sciences

Morcuende, J. A., Dolan, L. A., Dietz, F. R., & Ponseti, I. V. (2004). Radical reduction in the rate of extensive corrective surgery for clubfoot using the Ponseti method. *Pediatrics*, *113*(2), 376-380. https://doi.org/10.1542/peds.113.2.376

Pilkington, G., & Hounsome, J. (2017). Planning and Managing My Review. In A. Boland, G. Cherry, & R. Dickson (Eds.), *Doing a systematic review* (Vol. 2nd edition, pp. 22-41). SAGE.

Ramírez, N., Flynn, J. M., Fernández, S., Seda, W., & MacChiavelli, R. E. (2011). Orthosis noncompliance after the Ponseti method for the treatment of idiopathic clubfeet: A relevant problem that needs reevaluation [Article]. *Journal of Pediatric Orthopaedics*, *31*(6), 710-715. https://doi.org/10.1097/BPO.0b013e318221eaa1

SBU. (2020). *Utvärdering av metoder i hälso- och sjukvården och insatser i socialtjänsten: en metodbok.* Stockholm: Statens beredning för medicinsk och social utvärdering (SBU). Retrieved 2021-04-13 from https://www.sbu.se/metodbok.

SBU. (2021). Method. The process of producing a systematic review. In S. A. f. H. T. A. a. A. f. o. S. Service (Ed.).

Solanki, P. V., Sheth, B. A., Poduval, M., & Sams, S. B. A. (2010). Effectiveness of modified ankle foot orthosis of low-temperature thermoplastics in idiopathic congenital talipes equino varus [Article]. *Journal of Pediatric Orthopaedics Part B*, *19*(4), 353-360. https://doi.org/10.1097/BPB.0b013e3283387d16

SPOQ. (2018). Ortostyper vid behandling av klumpfot. In S. p. o. kvalitetsregistret (Ed.).

SPOQ. (2019). Årsrapport 2019. In S. Barnortopediskt & Q. SPOQ (Eds.).

SPOQ. (n.d). *Delregister för klumpfotsfelställning, PEVA*. https://spoq.registercentrum.se/om-spoq/presentationer-av-spoqs-delregister/delregistret-for-klumpfotsfelstallning-peva/p/S1ScUY7ae

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., Ramsay, C. R., Regidor, D., Rothstein, H. R., Sandhu, L., Santaguida, P. L., Schünemann, H. J., Shea, B., Shrier, I., Tugwell, P., Turner, L., Valentine, J. C., Waddington, H., Waters, E., Wells, G. A., Whiting, P. F., & Higgins, J. P. T. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *355*, i4919-i4919. https://doi.org/10.1136/bmj.i4919

Sætersdal, C., Fevang, J. M., & Engesæter, L. B. (2017). Inferior results with unilateral compared with bilateral brace in ponseti-treated clubfeet [Article]. *Journal of Children's Orthopaedics*, *11*(3), 216-222. https://doi.org/10.1302/1863-2548.11.160279

Sætersdal, C., Fevang, J. M., Fosse, L., & Engesæter, L. B. (2012). Good results with the Ponseti method...A multicenter study of 162 clubfeet followed for 2–5 years. *Acta Orthopaedica*, *83*(3), 288-293. https://doi.org/10.3109/17453674.2012.693015

Thacker, M. M., Scher, D. M., Sala, D. A., van Bosse, H. J., Feldman, D. S., & Lehman, W. B. (2005). Use of the foot abduction orthosis following Ponseti casts: is it essential? *J Pediatr Orthop*, *25*(2), 225-228. https://doi.org/10.1097/01.bpo.0000150814.56790.f9

https://doi.org/10.1097/01.bpo.0000150814.56790.f9

The Royal Children's Hospital. (2007). *Orthopaedic fact sheet 32, The Foot Abduction Brace for Clubfoot*. The Royal Children's Hospital (RCH), Victoria, Australia. https://www.rch.org.au/uploadedfiles/main/content/ortho/factshee

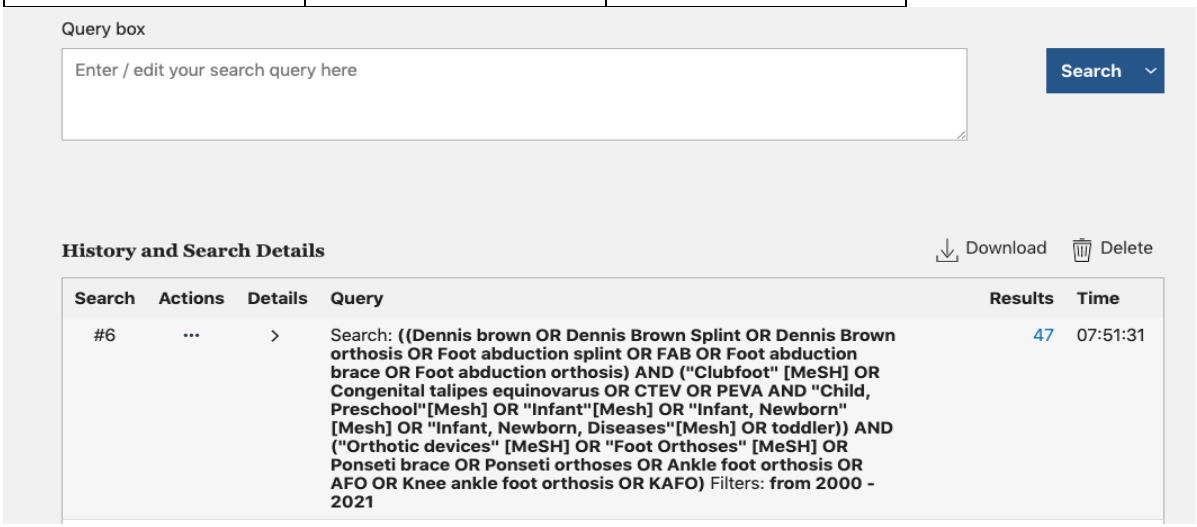
Zionts, L. E., & Dietz, F. R. (2010). Bracing following correction of Idiopathic clufoot using the Ponseti method. In (Vol. 18). American Academy of Orthopaedic Surgeons.

Zionts, L. E., Sangiorgio, S. N., Ebramzadeh, E., & Morcuende, J. A. (2012). The current management of idiopathic clubfoot revisited: Results of a survey of the POSNA membership [Article]. *Journal of Pediatric Orthopaedics*, *32*(5), 515-520. https://doi.org/10.1097/BPO.0b013e318259ff79

Appendix 1: Search strategy

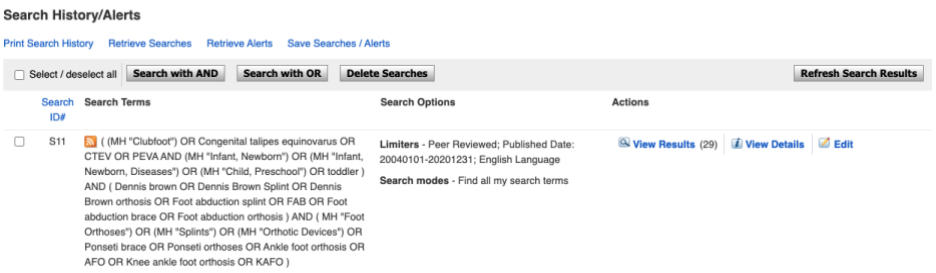
PubMed: 47 identified articles

P: Congenital talipes equinovarus	I: Foot abduction orthosis	C: Orthotic solution
"Clubfoot" [MeSH] OR Congenital talipes equinovarus OR CTEV OR PEVA AND "Child, Preschool"[Mesh] OR "Infant"[Mesh] OR "Infant, Newborn"[Mesh] OR "Infant, Newborn, Diseases"[Mesh] OR toddler	Dennis brown OR Dennis Brown Splint OR Dennis Brown orthosis OR Foot abduction splint OR FAB OR Foot abduction brace OR Foot abduction orthosis	"Orthotic devices" [MeSH] OR "Foot Orthoses" [MeSH] OR Ponseti brace OR Ponseti orthoses OR Ankle foot orthosis OR AFO OR Knee ankle foot orthosis OR KAFO



MEDLINE *and* CINAHL*: 29 identified articles*

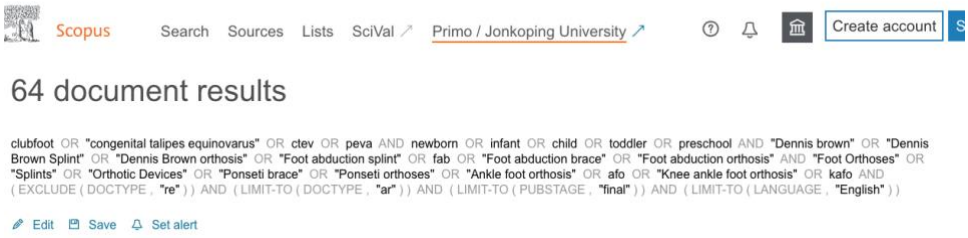
P: Congenital talipes equinovarus	I: Foot abduction orthosis	C: Orthotic solution
(MH "Clubfoot") OR Congenital talipes equinovarus OR CTEV OR PEVA AND (MH "Infant, Newborn") OR (MH "Infant, Newborn, Diseases") OR (MH "Child, Preschool") OR toddler	Dennis brown OR Dennis Brown Splint OR Dennis Brown orthosis OR Foot abduction splint OR FAB OR Foot abduction brace OR Foot abduction orthosis	(MH "Foot Orthoses") OR (MH "Splints") OR (MH "Orthotic Devices") OR Ponseti brace OR Ponseti orthoses OR Ankle foot orthosis OR AFO OR Knee ankle foot orthosis OR KAFO



Scopus: 64 identified articles

P: Congenital talipes equinovarus	I: Foot abduction orthosis	C: Orthotic solution
clubfoot OR "congenital talipes equinovarus" OR ctev OR peva AND newborn OR infant OR child OR toddler OR preschool	"Dennis brown" OR "Dennis Brown Splint" OR "Dennis Brown orthosis" OR "Foot abduction splint" OR fab OR "Foot abduction brace" OR "Foot abduction orthosis"	"Foot Orthoses" OR "Splints" OR "Orthotic Devices" OR "Ponseti brace" OR "Ponseti orthoses" OR "Ankle foot orthosis" OR afo OR "Knee ankle foot orthosis" OR kafo

JÖNKÖPING UNIVERSITY



Appendix 2: Critical appraisal tool, in Swedish

a) Bedömning av icke-randomiserade studier av interventioner (effekt av att tilldelas en intervention (ITT))

1A. Confounding

Identifiera viktiga confounders på det aktuella området för att besvara frågorna

Risk för bias från confounding bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
A1.1 effekten av interventionen har påverkats av viktiga confounders?	❑	❑	❑	❑	❑
Om svaret är "Nej" gå vidare till domän 1B.					
A1.2 deltagare bytte grupp eller avbröt behandlingen på grund av confounders som inte var synliga i baslinjen?	❑	❑	❑	❑	❑
A1.3 orsakerna till att deltagarna avbröt eller bytte grupp har påverkat utfallet?	❑	❑	❑	❑	❑
A1.4 man kontrollerade för alla viktiga confounders med acceptabla analysmetoder?	❑	❑	❑	❑	❑
A1.5 viktiga confounders var mätta med valida och reliabla metoder?	❑	❑	❑	❑	❑
A1.6 de data man använde för att kontrollera confounders var redovisade i studien?	❑	❑	❑	❑	❑
A1.7 man tog in och kontrollerade för nya variabler efter att interventionen inletts?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

1B. Selektion/gruppindelning

Risk för bias från selektion/ gruppindelning bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
B1.1 deltagaregenskaper (eller faktorer) som observerats efter att interventionen inletts påverkade valet av deltagare i studien/analysen?	❑	❑	❑	❑	❑
Om svaret är "nej" gå vidare till fråga B1.4.					
B1.2 dessa deltagaregenskaper (eller faktorer) hade samband med interventionen?	❑	❑	❑	❑	❑
B1.3 dessa deltagaregenskaper (eller faktorer) påverkades av utfallet eller av en orsak till utfallet?	❑	❑	❑	❑	❑
B1.4 intervention och uppföljning inföll vid samma fas i sjukdomsförloppet/utvecklingen för de flesta deltagarna?	❑	❑	❑	❑	❑
B1.5 lämpliga metoder som kan korrigera för selektionsbias användes?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

1C. Klassificering/avgränsning av interventionsgrupperna

Risk för bias från klassificering/ definition av interventionsgrupperna bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
C1.1 interventionsgrupperna var väl definierade?	❑	❑	❑	❑	❑
C1.2 informationen som användes för att definiera interventionsgrupperna samlades in innan resultatet av interventionen var känt (eller avblindat)?	❑	❑	❑	❑	❑
C1.3 definitionen av interventionsgrupperna kan ha påverkats av kännedom om utfallet?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

2. Avvikelser från planerade interventioner

Risk för bias från avvikelser från planerade interventioner bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
2.1 det fanns avvikelser från den planerade interventionen förutom vad som kan förväntas i klinisk rutin?	❑	❑	❑	❑	❑
Om bedömningen är "ja" eller "troligen ja" besvaras fråga 2.2					
2.2 avvikelserna var balanserade mellan grupperna?	❑	❑	❑	❑	❑
Om bedömningen är "nej" eller "troligen nej" besvaras fråga 2.3					
2.3. obalansen påverkade utfallet?	❑	❑	❑	❑	❑
Risk för bias	Låg ❑		Måttlig ❑		Hög ❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

3. Bortfall

Risk för bias från bortfall bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
3.1 resultat redovisades för alla eller nästan alla deltagare?	❑	❑	❑	❑	❑
Om svaret är "Ja" gå vidare till domän 4.					
3.2 man har visat att resultaten är robusta trots bortfallet (exempelvis med känslighetsanalyser)?	❑	❑	❑	❑	❑
3.3 bortfallet med stor sannolikhet är relaterat till utfallsmättet?	❑	❑	❑	❑	❑
3.4 såväl bortfallet som orsaker till bortfallet var likartat mellan grupperna?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

4. Mätning av utfall

Risk för bias från mätning av utfallet bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
4.1 datainsamlingen skilde sig åt mellan grupperna?	❑	❑	❑	❑	❑
4.2 de som mätte utfallet var medvetna om vilken intervention deltagarna fått?	❑	❑	❑	❑	❑
4.3 bedömningen med stor sannolikhet påverkades av detta?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

5. Rapportering

Risk för bias från rapportering bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
5.1 analyserna var genomförda enligt en plan som publicerats innan utfallsdata var tillgängliga?	❑	❑	❑	❑	❑
5.2 de rapporterade resultaten har valts ut från flera sätt att mäta utfallet (t.ex. olika skalor, tidpunkter)?	❑	❑	❑	❑	❑
5.3 de rapporterade resultaten har valts ut från olika analyser av samma utfall?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

Jäv/intressekonflikter (kan rapporteras narrativt)

	Ja	Nej	Kommentar
Deklarerar författarna att de saknar finansiella intressen som kan påverka utfallet?	❑	❑	
Deklarerar författarna att de saknar andra bindningar som kan påverka utfallet?	❑	❑	
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

b) Bedömning av icke-randomiserade studier av interventioner (effekt av att fullfölja en intervention (per protokoll))

1A. Confounding

Identifiera viktiga confounders på det aktuella området för att besvara frågorna

Risk för bias från confounding bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
A1.1 effekten av interventionen har påverkats av viktiga confounders?	❑	❑	❑	❑	❑
Om svaret är "Nej" gå vidare till domän 1B.					
A1.2 deltagare bytte grupp eller avbröt behandlingen på grund av confounders som inte var synliga i baslinjen?	❑	❑	❑	❑	❑
A1.3 orsakerna till att deltagarna avbröt eller bytte grupp har påverkat utfallet?	❑	❑	❑	❑	❑
A1.4 man kontrollerade för alla viktiga confounders med acceptabla analysmetoder?	❑	❑	❑	❑	❑
A1.5 viktiga confounders var mätta med valida och reliabla metoder?	❑	❑	❑	❑	❑
A1.6 de data man använde för att kontrollera confounders var redovisade i studien?	❑	❑	❑	❑	❑
A1.7 man tog in och kontrollerade för nya variabler efter att interventionen inletts?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

1B. Selektion/gruppindelning

Risk för bias från selektion/ gruppindelning bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
B1.1 deltagaregenskaper (eller faktorer) som observerats efter att interventionen inletts påverkade valet av deltagare i studien/analysen?	❑	❑	❑	❑	❑
Om svaret är "nej" gå vidare till fråga B1.4.					
B1.2 dessa deltagaregenskaper (eller faktorer) hade samband med interventionen?	❑	❑	❑	❑	❑
B1.3 dessa deltagaregenskaper (eller faktorer) påverkades av utfallet eller av en orsak till utfallet?	❑	❑	❑	❑	❑
B1.4 intervention och uppföljning inföll vid samma fas i sjukdoms-förloppet/utvecklingen för de flesta deltagarna?	❑	❑	❑	❑	❑
B1.5 lämpliga metoder som kan korrigera för selektionsbias användes?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

1C. Klassificering/avgränsning av interventionsgrupperna

Risk för bias från klassificering/ definition av interventions-grupperna bedöms som: Motivering: se stödfrågorna nedan	**Låg** ❑	**Måttlig** ❑	**Hög** ❑	**Oacceptabelt hög** ❑	
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
C1.1 interventionsgrupperna var väl definierade?	❑	❑	❑	❑	❑
C1.2 informationen som användes för att definiera interventionsgrupperna samlades in innan resultatet av inter-ventionen var känt (eller avblindat)?	❑	❑	❑	❑	❑
C1.3 definitionen av interventions-grupperna kan ha påverkats av kännedom om utfallet?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

2. Avvikelser från planerade interventioner

Risk för bias från avvikelser från planerade interventioner bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
2.1 det fanns avvikelser från den planerade interventionen förutom vad som kan förväntas i klinisk rutin?	❑	❑	❑	❑	❑
Om bedömningen är "ja" eller "troligen ja" besvaras fråga 2.2.					
2.2 avvikelserna var balanserade *mellan grupperna?*	❑	❑	❑	❑	❑
Om bedömningen är "nej" eller "troligen nej" besvaras fråga 2.3.					
2.3. obalansen påverkade utfallet?	❑	❑	❑	❑	❑
2.4. viktiga co-interventioner var balanserade mellan grupperna?	❑	❑	❑	❑	❑
2.5. interventionen var väl implementerad för de flesta deltagarna?	❑	❑	❑	❑	❑
2.6. deltagarna var följsamma till den intervention som de tilldelats?	❑	❑	❑	❑	❑
Om bedömningen på någon av frågorna 2.4–2.6 är "nej" eller "troligen nej" besvaras fråga 2.7					
2.7. studien använde en rimlig *metod för att analysera effekten av interventionen?*	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

3. Bortfall

Risk för bias från bortfall bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
3.1 resultat redovisades för alla eller nästan alla deltagare?	❑	❑	❑	❑	❑
Om svaret är "ja" gå vidare till domän 4.					
3.2 man har visat att resultaten är *robusta trots bortfallet (exempelvis med känslighetsanalyser)?*	❑	❑	❑	❑	❑
3.3 bortfallet *med stor sannolikhet är relaterat till utfallsmåttet?*	❑	❑	❑	❑	❑
3.4 såväl bortfallet som orsaker *till bortfallet var likartat mellan grupperna?*	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

4. Mätning av utfall

Risk för bias från mätning av utfallet bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
4.1 datainsamlingen skilde sig åt mellan grupperna?	❑	❑	❑	❑	❑
4.2 de som mätte utfallet var medvetna om vilken intervention deltagarna fått?	❑	❑	❑	❑	❑
4.3 bedömningen med stor sannolikhet påverkades av detta?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

5. Rapportering

Risk för bias från rapportering bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att...?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
5.1 analyserna var genomförda enligt en plan som publicerats innan utfallsdata var tillgängliga?	❑	❑	❑	❑	❑
5.2 de rapporterade resultaten har valts ut från flera sätt att mäta utfallet (t.ex. olika skalor, tidpunkter)?	❑	❑	❑	❑	❑
5.3 de rapporterade resultaten har valts ut från olika analyser av samma utfall?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

Jäv/intressekonflikter (kan rapporteras narrativt)

	Ja	Nej	Kommentar
Deklarerar författarna att de saknar finansiella intressen som kan påverka utfallet?	❑	❑	
Deklarerar författarna att de saknar andra bindningar som kan påverka utfallet?	❑	❑	
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

c) Bedömning av randomiserade studier (effekt av att tilldelas en intervention (ITT))

1. Randomisering

Risk för bias från randomiseringen bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
1.1 gruppindelningen var randomiserad?	❑	❑	❑	❑	❑
1.2 blivande grupptillhörighet inte kunde förutses, den var okänd tills deltagarna delats in (concealed allocation sequence)?	❑	❑	❑	❑	❑
1.3 baslinjen hade obalanser som tyder på brister i randomiseringsprocessen?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

2. Avvikelser från planerade interventioner

Risk för bias från avvikelser från planerade interventioner bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
2.1 deltagarna kände till vilken intervention de tilldelats under studiens gång?	❑	❑	❑	❑	❑
2.2 behandlarna kände till vilka interventioner deltagarna tilldelats under studiens gång?	❑	❑	❑	❑	❑
Om svaret är "nej" på både 2.1 och 2.2 gå vidare till fråga 2.5.					
2.3 kännedom om studien och gruppindelningen kunde leda till avvikelser som var obalanserade mellan grupperna (t.ex. förändringar i övrig vård eller avvikelser från klinisk praxis)?	❑	❑	❑	❑	❑
2.4 obalansen sannolikt påverkade utfallet?	❑	❑	❑	❑	❑
2.5 man använde en lämplig analysmetod för att uppskatta effekten?	❑	❑	❑	❑	❑
Om svaret på 2.5 är "nej" eller "troligen nej" besvara även 2.6					
2.6. resultatet påverkades allvarligt av att deltagarna inte analyserades i den grupp de randomiserats till?	❑	❑	❑	❑	❑
Risk för bias	Låg ❑		Måttlig ❑		Hög ❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

3. Bortfall

Risk för bias från bortfall bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
3.1 resultat redovisades för alla eller nästan alla deltagare?	❑	❑	❑	❑	❑
Om svaret är "ja" gå vidare till domän 4.					
3.2 man har visat att resultaten är robusta trots bortfallet (exempelvis med känslighetsanalyser)?	❑	❑	❑	❑	❑
3.3 bortfallet med stor sannolikhet är relaterat till utfallsmåttet?	❑	❑	❑	❑	❑
3.4 såväl bortfallet som orsaker till bortfallet var likartat mellan grupperna?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

4. Mätning av utfall

Risk för bias från mätning av utfallet bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
4.1 datainsamlingen skilde sig åt mellan grupperna?	❑	❑	❑	❑	❑
4.2 de som mätte utfallet var medvetna om vilken intervention deltagarna fått?	❑	❑	❑	❑	❑
4.3 bedömningen med stor sannolikhet påverkades av detta?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

5. Rapportering

Risk för bias från rapportering bedöms som:			Låg ❑	Måttlig ❑	Hög ❑
			Motivering: se stödfrågorna nedan		
Bedömer du att..?	**Ja**	**Troligen ja**	**Troligen nej**	**Nej**	**Information saknas**
5.1 analyserna var genomförda enligt en plan som publicerats innan utfallsdata var tillgängliga?	❑	❑	❑	❑	❑
5.2 de rapporterade resultaten har valts ut från flera sätt att mäta utfallet (t.ex. olika skalor, tidpunkter)?	❑	❑	❑	❑	❑
5.3 de rapporterade resultaten har valts ut från olika analyser av samma utfall?	❑	❑	❑	❑	❑
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

Jäv/intressekonflikter (kan rapporteras narrativt)

	Ja	Nej	Kommentar
Deklarerar författarna att de saknar finansiella intressen som kan påverka utfallet?	❑	❑	
Deklarerar författarna att de saknar andra bindningar som kan påverka utfallet?	❑	❑	
Om möjligt: Vilken är riktningen på bias för utfallet?	Gynnar intervention ❑	Gynnar kontroll ❑	Mot noll ❑	Från noll ❑	Går ej att bedöma ❑

Appendix 3: SBU Bilaga 2: Granskningsmallar och checklistor för bedömning av studier

Section A (randomized clinical trial)

External validity

Short form answer:

- ❑ Clear external validity (0)
- ❑ Probable external validity (1)
- ❑ Uncertain external validity (3)
- ❑ External validity cannot be assessed (5)

If uncertain, answer questions under Item 1.
Otherwise go to Internal validity (after Item 1)

1. Accrual of study subjects

a. Eligibility/inclusion criteria clearly stated (eg, if trial of treatment of a specified disease, is the definition acceptable)?
- ❑ Yes = 0
- ❑ No = 2

b. Consecutive eligible subjects?
- ❑ Yes = 0
- ❑ No = 1
- ❑ Not stated = 1

c. Numbers and reasons for non-participation given?
- ❑ Yes = 0
- ❑ No = 2

d. Exclusion criteria clearly stated and acceptable?
- ❑ Yes = 0
- ❑ No = 2

e. Are numbers of excluded persons given by reason (as prescribed in the CONSORT statement)?
- ❑ Yes = 0
- ❑ No = 2

Total sum of section 1
0 = Clear external validity
1 = Probable external validity
2–3 = Uncertain external validity
≥4 = External validity cannot be assessed

Internal validity

Short form answer:

- ❑ Excellent internal validity (0)
- ❑ Good internal validity (1)
- ❑ Acceptable internal validity (2)
- ❑ Uncertain internal validity (4)
- ❑ Uninformative due to flawed internal validity (10)

If uncertain, answer questions under Items 2–9.
Otherwise go to Precision (after Item 9)

2. Treatment/exposure assignment

a. Were details about randomization procedure given?
- ❑ Yes = 0
- ❑ No = 1

b. Could the randomization be manipulated?
- ❑ Yes (eg, tossing of coin or throwing of dice) = 1
- ❑ No (eg, opaque envelopes, computer-generated list kept by others than investigators) = 0

c. Did randomization lead to unpredictable treatment assignment?
- ❑ Yes = 0
- ❑ No, treatment could potentially be deduced in some or all = 2

d. Were there exclusions/withdrawals after randomization?
- ❑ Yes = 2
- ❑ No = 0

3. Comparability of groups

a. Was there an account of the comparability of groups with regard to all conceivable factors that might affect the outcome?
- ❑ Yes = 0
- ❑ No = 1

b. Were there any important differences?
- ❑ Yes = 2
- ❑ No = 0
- ❑ No data given = 0 (already scored under 3a)

c. Were any attempts in the analysis phase to adjust for imbalances between treatment arms with regard to important determinants for the outcome (eg, through multi-variate modelling)?
- ❑ Not needed (no important imbalances) = 0
- ❑ Yes = –1 (subtract 1 if you scored 2 under 3b)
- ❑ No, despite a need = 1

4. Blinding

a. Were there any attempts to blind the patients/investigators to treatment allocation?
- ❑ No (open study) = 2
- ❑ Only study subjects were blinded (single-blind) = 1
- ❑ Blinding only of investigators who evaluated the outcome ("blind observer") =0
- ❑ Double-blind = 0
- ❑ Triple-blind (breaking of the code first after completion of all analyses) = 0

b. Was there any reason to believe that the blinding had failed (eg, due to characteristic side-effects of active treatment or dissimilarities of active and reference tablets)?
- ❑ Yes = 1
- ❑ No = 0

c. Was the blinding tested (eg, through asking the subjects at the end of the study what they believed they had received)?
- ❑ Yes = 0
- ❑ No = 0

5. Compliance

a. Was there any account of the completeness of treatment/compliance?
- ❑ Yes = 0
- ❑ No = 2

b. Was the completeness acceptable (>80% of the subjects receiving >80% of the prescribed treatment)?
- ❑ Yes = 0
- ❑ No = 3
- ❑ Completeness/compliance data not given = 0 (scored under 5a)

6. Drop-outs/losses to follow-up

a. Was there an account of the numbers of subjects who dropped out (and the reasons for dropping out)?
- ❑ Yes = 0
- ❑ No = 3

b. What was the drop-out rate?
- ❑ <10% = 0
- ❑ 10–19% = 2
- ❑ 20–29% = 3
- ❑ ≥30% → study is deemed uninformative, excluded
- ❑ Drop-out rate not stated = 0 (scored under 6a)

7. Evaluation of outcome

a. Was there an acceptable definition of the outcome?
- ❑ Yes = 0
- ❑ No = 3

b. Was the outcome clinically relevant?
- ❑ Yes = 0
- ❑ Of questionable relevance = 2
- ❑ Irrelevant → study is deemed uninformative, excluded

c. Was the reporter of the outcome (eg, the investigator, the study subject) unaware of the treatment given?
- ❑ Yes = 0
- ❑ No = 2

d. Are there reasons to believe that there might have been misclassification of the outcome (eg, due to retrospective reporting over too long periods)?
- ❑ Yes = 1
- ❑ No = 0

8. Evaluation of side-effects

a. Was there acceptable reporting of side effects?
- ❑ Yes, with open-ended questions = 0
- ❑ Yes, with fixed response alternatives = 0
- ❑ Yes, response alternatives not stated = 0
- ❑ No = 3

9. Analysis

a. Was the main outcome variable defined in advance and was the conclusion of the study based on the analysis of this variable?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 2

b. Was there a prior hypothesis?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 1

c. Were the secondary variables defined in advance?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 1
- ❑ Not applicable, there was no secondary outcome variable = 0

d. Were all randomized subjects included in the analysis and retained in the treatment arm to which they were initially allocated ("intention-to-treat analysis")?
- ❑ Yes = 0
- ❑ No = 4

Total sum of Items 2–9 (internal validity)
0–1 = Excellent internal validity
2–4 = Good internal validity
5–7 = Acceptable internal validity
8–10 = Uncertain internal validity
≥10 = Uninformative due to flawed internal validity

Precision

Short form answer:

- ❑ Premeditated and sufficient study size (0)
- ❑ Sample size of uncertain adequacy (2)
- ❑ Probably underpowered study (4)

If uncertain, answer questions under Items 10–11

10. Smallest clinically relevant effect

a. Was the smallest clinically relevant effect defined?
- ❑ Yes = 0
- ❑ No = 1

b. Was the stated smallest clinically relevant effect reasonable?
- ❑ Yes = 0
- ❑ No = 1
- ❑ Not defined = 0 (scored under 10a)

11. Study power

a. Were the deliberations behind the sample size decision clearly described?
- ❑ Yes = 0
- ❑ No = 2

b. What was the power to detect a reasonably-sized smallest clinically relevant effect?
- ❑ Not stated because there was a strong and statistically significant effect = 0
- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ <70% = 3
- ❑ Not stated despite a non-significant finding = 4

Total sum of Items 10–11 (precision)
0–1 = Premeditated and sufficient study size
2–3 = Sample size of uncertain adequacy
≥4 = Probably underpowered study

Section B (observational cohort study or controlled clinical trial without randomisation)

External validity

Short form answer:

- ❑ Clear external validity (0)
- ❑ Probable external validity (1)
- ❑ Uncertain external validity (3)
- ❑ External validity cannot be assessed (5)

If uncertain, answer questions under Item 1.
Otherwise go to Internal validity (after Item 1)

1. Accrual/selection of study subjects

a. Was the studied exposure well defined (eg, if follow-up of a specified disease, is the definition of the disease acceptable)?
- ❑ Yes = 0
- ❑ No = 2

b. Eligibility/inclusion criteria clearly stated?
- ❑ Yes = 0
- ❑ No = 1

c. Consecutive eligible subjects included?
- ❑ Yes = 0
- ❑ No = 1
- ❑ Not stated = 1

d. Numbers and reasons for non-participation given?
- ❑ Yes = 0
- ❑ No = 1

e. Exclusion criteria clearly stated and acceptable?
- ❑ Yes = 0
- ❑ No = 1

f. Are numbers of excluded persons given by reason (as prescribed in the CONSORT statement)?
- ❑ Yes = 0
- ❑ No = 1

Total sum of section 1
0 = Clear external validity
1 = Probable external validity
2–3 = Uncertain external validity
≥4 = External validity cannot be assessed

Internal validity

Short form answer:

- ❑ Excellent internal validity (0)
- ❑ Good internal validity (1)
- ❑ Acceptable internal validity (2)
- ❑ Uncertain internal validity (4)
- ❑ Uninformative due to flawed internal validity (10)

If uncertain, answer questions under Items 2–6.
Otherwise go to Precision (after Item 6)

2. Exposure assessment

a. Was the studied exposure satisfactorily measured/recorded?
- ❑ Yes = 0
- ❑ Yes, with minor criticism = 1
- ❑ No = 3

b. Were all in the exposed group really exposed?
- ❑ Yes = 0
- ❑ Yes, probably = 1
- ❑ No, probably not = 2
- ❑ No = 2

c. Were all in the reference category really unexposed?
- ❑ Yes = 0
- ❑ Yes, probably = 1
- ❑ No, probably not = 2
- ❑ No = 2

3. Comparability of groups/selection bias/confounding

a. Was there an account of the comparability of groups with regard to factors that might conceivably affect the outcome (potential confounding factors)? (If only one cohort was studied and compared with the background population or historical controls – was there data to support the comparability with the reference category).
- ❑ Yes = 0
- ❑ No = 3

b. Did the investigators consider all important potential confounding factors (potential confounding factors = factors that are independent causes of/risk factors for/protective factors against the outcome, AND not a link in the causal chain between the studied exposure and the outcome)?
- ❑ Yes = 0
- ❑ Probably = 1
- ❑ No = 3
- ❑ No data given = 0 (already scored under 3a)

c. Were the relevant confounding factors satisfactorily measured/recorded?
- ❑ Yes = 0
- ❑ Yes, with minor criticism = 1
- ❑ No = 3

d. Were the potential confounding factors unevenly distributed among exposed and /non-exposed/ reference group (confounding arises if factors described under 3b are unevenly distributed among exposed and unexposed [ie, linked to the exposure])?
- ❑ Yes = 2
- ❑ No = 0
- ❑ No data given = 0 (already scored under 3a)

e. Were attempts in the analysis to adjust for imbalances between exposure groups with regard to potential confounding factors (eg, through restriction, stratified analyses, or multivariate modelling)?
- ❑ Not needed (no important imbalances) = 0
- ❑ Yes = –2 (subtract 2 if you scored 2 under 3d)
- ❑ No, despite a need = 2

4. Evaluation of outcome, ascertainment/detection bias

a. Was there an acceptable definition of the outcome?
- ❑ Yes = 0
- ❑ No = 3

b. Was the outcome clinically relevant?
- ❑ Yes = 0
- ❑ Of questionable relevance = 2
- ❑ Irrelevant → study is deemed uninformative, excluded

c. Were the evaluators of the outcome aware of exposure status of the cohort members?
- ❑ Yes = 1
- ❑ Probably = 1
- ❑ No = 0

d. Was there any reason to believe that there was important ascertainment/ detection bias (eg, exposure linked to smoking, and smoking, in turn, linked to higher frequency of health care visits, and thus a more intense surveillance)?
- ❑ Yes = 2
- ❑ No = 0

5. Losses to follow-up

a. Was there an account of the numbers of subjects who were lost to follow-up?
- ❑ Yes = 0
- ❑ No = 3

b. What proportion was lost to follow-up?
- ❑ <10% = 0
- ❑ 10–19% = 1
- ❑ 20–29% = 2
- ❑ 30–39 = 3
- ❑ ≥40% → study is deemed uninformative, excluded
- ❑ Proportion not stated = 0 (scored under 5a)

6. Analysis

a. Was the main outcome variable defined in advance and was the conclusion of the study based on the analysis of this variable?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 1

b. Was there a prior hypothesis?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 1

c. Was the statistical method adequate?
- ❑ Yes = 0
- ❑ No = 3

Total sum of Items 2–6 (internal validity)
0–1 = Excellent internal validity
2–3 = Good internal validity
4–6 = Acceptable internal validity
7–9 = Uncertain internal validity
≥10 = Uninformative due to flawed internal validity

Precision

Short form answer:

- ❑ Premeditated and sufficient study size (0)
- ❑ Sample size of uncertain adequacy (2)
- ❑ Probably underpowered study (4)

If uncertain, answer questions under Items 7–8

7. Smallest clinically relevant effect

a. Was the smallest clinically relevant effect defined?
- ❑ Yes = 0
- ❑ No = 1

b. Was the stated smallest clinically relevant effect reasonable?
- ❑ Yes = 0
- ❑ No = 1
- ❑ Not defined = 0 (scored under 10a)

8. Study power

a. Were the deliberations behind the sample size decision clearly described?
- ❑ Yes = 0
- ❑ No = 2

b. What was the power to detect a reasonably-sized smallest clinically relevant effect?
- ❑ Not stated because there was a strong and statistically significant effect = 0
- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ <70% = 3
- ❑ Not stated despite a non-significant finding = 4

Total sum of Items 7–8 (precision)
0–1 = Premeditated and sufficient study size
2–3 = Sample size of uncertain adequacy
≥4 = Probably underpowered study

Section C (case-control or cross-sectional studies)

External validity

Short form answer:

- ❑ Clear external validity (0)
- ❑ Probable external validity (1)
- ❑ Uncertain external validity (3)
- ❑ External validity cannot be assessed (5)

If uncertain, answer questions under Item 1.
Otherwise go to Internal validity (after Item 1)

1. Type of cases studied

a. Was there an acceptable definition of the outcome (that rendered subjects case/control status)?
   - ❑ Yes = 0
   - ❑ No = 2

b. Did the studied cases correspond to cases in the population to which the investigators wished to generalise their findings?
   - ❑ Yes = 0
   - ❑ Yes, probably = 1
   - ❑ No, probably not = 2
   - ❑ No, definitely not = 3

Total sum of section 1
0 = Clear external validity
1 = Probable external validity
2–3 = Uncertain external validity
≥4 = External validity cannot be assessed

Internal validity

Short form answer:

- ❑ Excellent internal validity (0)
- ❑ Good internal validity (1)
- ❑ Acceptable internal validity (2)
- ❑ Uncertain internal validity (4)
- ❑ Uninformative due to flawed internal validity (10)

If uncertain, answer questions under Items 2–6.
Otherwise go to Precision (after Item 6)

2. Study base (NOTE, not relevant to cross-sectional studies; if so, skip 2–3)

The study base is defined as the group of people (the "virtual cohort") who – if they developed the outcome condition – would necessarily have become cases in the study.

a. Was the study base (the "virtual cohort" [a defined source population followed for a defined time period] *that generated the cases*) well defined (geographically, age-wise, gender, other characteristics)?
   - ❑ Yes, quite clear (eg, an already established cohort, or definition through an existing, well-functioning population register) = 0
   - ❑ Yes, reasonably (eg, hospital-based study with strict catchment areas and no important selections of cases or controls) = 1
   - ❑ Yes, probably (eg, hospital-based study without clear catchment areas, and/or inability to rule out some less important selection among cases and/or controls; control selection via random digit dialing or through neighbourhood controls whereupon some minor mismatch [for instance socioeconomic] between cases and controls might have occurred) = 2
   - ❑ No, it is impossible to tell if the cases and controls come from the same study base and if there are important selection mechanisms for either of these categories = 4

b. Are the cases representative of all cases in the study base?
   - ❑ Yes, they represent all or virtually all new (incident) cases of the outcome that occurred in the study base = 0
   - ❑ Yes, although it is difficult to tell if they represent all cases, there is no reason to suspect that they are unrepresentative of all cases in the study base = 1
   - ❑ Yes, they represent prevalent cases in the study base, but there is no reason to suspect that they are unrepresentative = 1
   - ❑ No, there are reasons to suspect that they are unrepresentative of all cases in the study base = 3
   - ❑ No, definitely unrepresentative → study is deemed uninformative, excluded

c. Do the control subjects come from the very same study base as the cases?
- ❑ Yes, definitely = 0
- ❑ Yes, probably = 1
- ❑ Uncertain = 3
- ❑ Probably not = 4
- ❑ No, definitely not → study is deemed uninformative, excluded

d. Were the control subjects representative of the entire study base?
- ❑ Yes, they were selected randomly from a defined sampling frame (note that stratified random sampling in order to achieve frequency matching is acceptable) = 0
- ❑ Yes, probably, but they were selected in some other way = 1
- ❑ Uncertain = 3
- ❑ Probably not = 4
- ❑ No, the probability of being selected as control is linked to the subjects' exposure status → study is deemed uninformative, excluded

3. Non-participation

a. Were all eligible cases occurring in the study base identified and enumerated?
- ❑ Yes = 0
- ❑ Yes, probably = 1
- ❑ No = 3

b. What was the participation rate among all eligible cases?
- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ 60–69% = 3
- ❑ 50–59% = 4
- ❑ <50% → study is deemed uninformative, excluded
- ❑ Proportion not stated → study is deemed uninformative, excluded

c. Was anything done to insure that major selection bias was not introduced through non-participation among cases?
- ❑ Not needed because participation among cases was >80% = 0
- ❑ Participation ≤80%, but authors provide data about non-participants that seem to rule out important selection bias = –1 (subtract from sum)
- ❑ Participation ≤80%, and no data is given about non-participants = 0

d. What was the participation rate among all selected controls?
- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ 60–69% = 3
- ❑ 50–59% = 4
- ❑ <50% → study is deemed uninformative, excluded
- ❑ Proportion not stated → study is deemed uninformative, excluded

e. Was anything done to insure that major selection bias was not introduced through non-participation among controls?
- ❑ Not needed because participation among controls was >80% = 0
- ❑ Participation ≤80%, but authors provide data about non-participants that seem to rule out important selection bias = –1 (subtract from sum)
- ❑ Participation ≤80%, and no data is given about non-participants = 0

4. Participation in cross-sectional study (skip if regular case-control study)

- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ 60–69% = 3
- ❑ 50–59% = 4
- ❑ <50% → study is deemed uninformative, excluded
- ❑ Proportion not stated → study is deemed uninformative, excluded

5. Exposure assessment

a. How was exposure information collected?
- ❑ From existing databases with data obtained before cases developed outcome = 0
- ❑ Face-to-face or telephone interviews with interviewers blinded to case/control status = 0
- ❑ Face-to-face or telephone interviews where interviewers were aware of case/control status = 1
- ❑ Postal questionnaire = 2
- ❑ Other ways or not stated = 3

b. Use of substitute responders?
- ❑ No = 0
- ❑ ≤20% = 1
- ❑ >20% = 3

c. Are there good reasons to suspect biased recall (ie, cases remember/report exposures systematically different compared to controls)?
- ❑ No = 0
- ❑ No, probably not = 1
- ❑ Uncertain = 2
- ❑ Yes, recall bias likely = 4
- ❑ Yes, high probability of recall bias → study is deemed uninformative, excluded

6. Confounding

a. Did the investigators consider all important potential confounding factors (potential confounding factors = factors that are independent causes of/risk factors for/protective factors against the outcome, AND not a link in the causal chain between the studied exposure and the outcome)?
- ❑ Yes = 0
- ❑ Probably = 1
- ❑ No = 3
- ❑ No data given = 4

b. Were the relevant confounding factors satisfactorily measured/recorded?
- ❑ Yes = 0
- ❑ Yes, with minor criticism = 1
- ❑ No = 3

c. Were attempts in the study design or analysis to identify and handle confounding factors (eg, through matching, restriction, stratified analyses, or multivariate modelling)?
- ❑ Yes, adequately = 0
- ❑ Yes, but not sufficiently = 2
- ❑ No → study is deemed uninformative, excluded

7. Ascertainment/detection bias

a. Was there any reason to believe that there was important ascertainment/detection bias (eg, exposure linked to smoking, and smoking, in turn, linked to higher frequency of health care visits, and thus a more intense surveillance)?
- ❑ Yes = 2
- ❑ No = 0

8. Rare disease assumption

a. Was the rare disease assumption fulfilled (the outcome affected less than 10% of the population in the study base)?
- ❑ Yes = 0
- ❑ Unknown = 1
- ❑ No or probably not = 3 (effects are likely exaggerated!)

9. Analysis

a. Was there a prior hypothesis?
- ❑ Yes = 0
- ❑ No (or not mentioned in the report) = 1

b. Was the statistical method adequate?
- ❑ Yes = 0
- ❑ No = 3

Total sum of Items 2–9 (internal validity) – CASE-CONTROL STUDY
0–2 = Excellent internal validity
3–4 = Good internal validity
5–7 = Acceptable internal validity
8–10 = Uncertain internal validity
≥11 = Uninformative due to flawed internal validity

Total sum of Items 2–9 (internal validity) – CROSS-SECTIONAL STUDY
0–1 = Excellent internal validity
2–3 = Good internal validity
4–5 = Acceptable internal validity
6–8 = Uncertain internal validity
≥9 = Uninformative due to flawed internal validity

Precision

Short form answer:

- ❑ Premediated and sufficient study size (0)
- ❑ Sample size of uncertain adequacy (2)
- ❑ Probably underpowered study (4)

If uncertain, answer questions under Items 10–11

10. Smallest clinically relevant effect

a. Was the smallest clinically relevant effect defined?
- ❑ Yes = 0
- ❑ No = 1

b. Was the stated smallest clinically relevant effect reasonable?
- ❑ Yes = 0
- ❑ No = 1
- ❑ Not defined = 0 (scored under 10a)

11. Study power

a. Were the deliberations behind the sample size decision clearly described?
- ❑ Yes = 0
- ❑ No = 2

b. What was the power to detect a reasonably-sized smallest clinically relevant effect?
- ❑ Not stated because there was a strong and statistically significant effect = 0
- ❑ ≥90% = 0
- ❑ 80–89% = 1
- ❑ 70–79% = 2
- ❑ <70% = 3
- ❑ Not stated despite a non-significant finding = 4

Total sum of Items 10–11 (precision)
0–1 = Premeditated and sufficient study size
2–3 = Sample size of uncertain adequacy
≥4 = Probably underpowered study

Appendix 4: Form for Self-Assessment of Ethical Issues in Degree Projects1 at the School of Health and Welfare

JÖNKÖPING UNIVERSITY
School of Health and Welfare

Part A: Is this a research study?

The purpose of the questions in Part A is to determine if the study is to be regarded as research. A degree project is not ordinarily considered research and thus cannot be reviewed by the Swedish Ethical Review Authority. Under certain circumstances, however, a degree project may be research, namely if:

1. the intention is to publish it in a scientific journal **NO**
2. it addresses a scientific question and has a design that can answer that question **YES**
3. it is led by researchers within the discipline, either as part of a larger project or with a researcher as the supervisor. **YES**

All of these three must be fulfilled for the study to be considered research and be able to be reviewed by the Swedish Ethical Review Authority.

Is the study research in these three respects?

☐ YES (The study needs to be reviewed by the Swedish Ethical Review Authority.)

☐ **NO (Proceed to Parts B and C.)**

Part B: Does the degree project contain what is regarded as ethically sensitive according to the Ethical Review of Research Involving Humans?

The purpose of the questions in Part B is to explore whether the degree project raises any ethical issues that, *if it were research*, would require a review by the Swedish Ethical Review Authority, and whether the degree project complies with ethical principles.

		Yes	Not sure	No
1	Does the study intend to process what the General Data Protection Regulation (GDPR) considers to be sensitive personal data, i.e., data that at some stage can be linked to a person and that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, or information about an individual's health or sex life?	☐	☐	☒
2	Does the study intend to collect and process personal data relating to violations of the law that involve criminal offences, convictions in criminal proceedings, penal law sanctions, or administrative deprivation of liberty?	☐	☐	☒
3	Does the study entail a physical intervention on the participants (also that which is included in standard procedures, but also part of the research)?	☐	☐	☒
4	Is the purpose of the study to affect the participants physically or psychologically?	☐	☐	☒
5	Are there any risks for the participants (for example, the risk of physical harm or the risk of awakening traumatic memories)?	☐	☐	☒
6	Will the study use biological material that can be traced to an identifiable individual or deceased person (e.g., blood samples or tissue specimens)?	☐	☐	☒
7	Can voluntariness be questioned (e.g., vulnerable groups, such as children, people with cognitive impairment or mental disabilities, or individuals in a dependent relationship to the principal investigator, such as a patient or student)?	☐	☐	☒
8	Will the study involve individuals with limited autonomy (for instance individuals with cognitive difficulties, minors), whose understanding of the meaning of the consent is limited?	☐	☐	☒
	Selection of participants and social vulnerability			
9	Do the participants belong to a particularly vulnerable or disadvantaged group in society (a minority group)?	☐	☐	☒
10	Will the study involve the establishment of a personal register where data can be linked to a physical person?	☐	☐	☒
	Informed consent			
11	Will informed consent be obtained as a part of the study (in other words, will the participants receive full information about the study and/or the opportunity to opt out from participation?	☐	☐	☒
12	Is the study described in such a manner so that the participants understand its purpose and structure, and what participation in the project entails (e.g., number of visits, duration of the project, and written in easy-to-understand Swedish without technical terms or professional jargon)?	☐	☐	☒
13	Is it clearly stated in the written information to the participant that the participation in the study is entirely voluntary?	☐	☒	☐
14	Does the information letter contain persuasive formulations (implying that the person must or should participate, without showing full respect for the choice, for example mildly persuasive formulations, such as "thanks in advance")?	☐	☒	☐
15	Is it clearly stated that the participants may choose not to participate, without prejudice to the participants' being offered care or treatment or, if relating to students, it affecting their grades?	☐	☒	☐
16	Is it clearly stated that the participants may discontinue the participation at any time and without the need to state any reason, without prejudice to the participants' being offered care or treatment or, if relating to students, it affecting their grades?	☐	☒	☐
	Confidentiality and the security of the participants			
17	Are there procedures in place to ensure confidentiality in the collection of data?	☐	☒	☐
18	Are the results/findings described in such a manner so that the participants' identity remains confidential, meaning that they cannot be identified afterwards (including a minimal potential for reverse identification)?	☐	☒	☐
19	Are there procedures to ensure confidentiality?	☐	☒	☐
20	Are there clear routines for ensuring that the collected data material is handled according to GDPR?	☐	☒	☐
	Research results and/or findings			
21	Are there reasons to offer participants the opportunity to obtain a copy of or otherwise gain access to the research results/findings?	☐	☒	☐

The questions under section B, have been responded truthfully and discussed with supervisor.

Place and date: Sweden/UK, 4/3-2021

Signature Student **Signature Supervisor**

Part C: Application for an Advisory Opinion

Follow the instructions in the bulleted list. The application must be written under supervision and signed by the student(s) and supervisor. The text may not exceed 1,000 words, excluding the project plan (Times New Roman, 12 pt, 1.5 line spacing).

- Describe the ethical risks and problems identified in Part B.
- Describe and explain how these ethical risks and problems could be addressed.
- Attach the project plan and relevant attachments, such as questionnaires and information letters.

The project we are doing is a systematic review and will therefore not touch the questions asked in part B. We will ensure that the studies we include have had the questions about confidentiality in mind when doing their research and in that way our work will have not include anything that is not reviewed in an ethical way.

Sign the "Self-Assessment of Ethical Issues" form, scan it, and e-mail it along with what is requested in Part C to the Research Ethics Committee's secretary. All documents must be in PDF format. The dates for the meetings of the Research Ethics Committee can be found on the university's website.

The above questions have been carefully reviewed, truthfully answered, and discussed with the supervisor.

Place and date:

	Printed name	Signature
Student/Implementer:	Josefin Larsen	
Student/Implementer:	Matilda Stoomback	
Supervisor:	Cleveland Barnett	